



INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Otimização do Planeamento Curricular: Análise do Desempenho Escolar de Estudantes de Informática

Mariana Matos Fonseca

Mestrado em Ciência de Dados

Orientador:

Doutor Pedro de Paula Nogueira Ramos, Professor Catedrático,
ISCTE - Instituto Universitário de Lisboa

Orientador:

Doutor Sérgio Miguel Carneiro Moro, Professor Catedrático,
ISCTE - Instituto Universitário de Lisboa

Setembro, 2024

iscte

BUSINESS
SCHOOL

iscte

TECNOLOGIAS
E ARQUITETURA

Departamento de Ciências e Tecnologias da Informação
**Otimização do Planeamento Curricular: Análise do Desempenho
Escolar de Estudantes de Informática**

Mariana Matos Fonseca

Mestrado em Ciência de Dados

Orientador:

Doutor Pedro de Paula Nogueira Ramos, Professor Catedrático,
ISCTE - Instituto Universitário de Lisboa

Orientador:

Doutor Sérgio Miguel Carneiro Moro, Professor Catedrático,
ISCTE - Instituto Universitário de Lisboa

Setembro, 2024

Agradecimentos

“Se podes olhar, vê. Se podes ver, repara.”

- Livro dos Conselhos

Quero começar por agradecer aos Professores Pedro Ramos e Sérgio Moro, orientadores desta dissertação, que me acompanharam ao longo de todo o processo e que me ajudaram a ultrapassar os diversos obstáculos que foram surgindo. Não teria conseguido realizar este estudo sem o seu apoio e sem os ensinamentos que precedem a realização da dissertação.

Neste sentido, quero também agradecer aos Professores Carlos Coutinho, Rui Neto Marinheiro e Sancho Moura Oliveira por se mostrarem disponíveis para dar sugestões acerca da interpretação de resultados e de possíveis análises, e por serem a primeira “prova” de que este estudo pode ter informações úteis para os professores do ISCTE-IUL.

À empresa que me acompanhou ao longo destes dois anos de Mestrado, o meu grande agradecimento. Ao Paulo Ribeiro, ao Francisco Sanches, à Catarina Ferreira, ao João Bartolomeu, e ao Emanuel Couto. Sem os conselhos, apoio, disponibilidade e amizade, não seria possível concluir esta etapa, muito menos em paralelo com uma experiência profissional tão enriquecedora.

Aos amigos que o curso me deu, Gil, André e Mariana, que espero levar comigo para além do mesmo. Obrigada pela companhia, pelos convites, pela ajuda, e por partilharem os azares e os sucessos comigo. E às amigas que me ouviram falar de dados durante 2 anos, e que continuaram a motivar-me e a apoiar-me, mesmo quando eu tinha menos tempo para elas. Obrigada, Ana, Daniela, Inês, Alexandra, Beatriz e Manuela.

À minha querida família, que mesmo sem perceber muitas vezes o que eu estava a fazer ou a estudar, ajudavam da melhor maneira que podiam. À minha avó, à minha madrinha, aos meus pais, e à minha tia Roberta.

Ao meu namorado Paulo, o meu pilar. Obrigada por estares comigo em todos os momentos. Partilhamos desafios e conquistas desde muito antes do Mestrado, e espero que continuemos muito para além dele. Continuo a inspirar-me na tua determinação, inteligência e curiosidade, e só posso agradecer o teu apoio incondicional.

Não posso deixar de reconhecer a sorte que tenho em ter uma rede de apoio com a qual posso contar em momentos mais desafiantes. Não teria chegado aqui sozinha e, por isso, só vos posso agradecer.

E, por fim, quero agradecer-me a mim. Desde que comecei a estudar que me esforço para dar o meu melhor em tudo o que faço, quer seja ao nível curricular, quer fora deste. Todos nós temos momentos nos quais é necessário sermos resilientes, e a elaboração desta dissertação foi o culminar dessa resiliência. Estou muito orgulhosa por tê-la concluído.

Espero que gostem!

Resumo

Esta dissertação investiga as inscrições de alunos em licenciaturas numa universidade pública, de forma a detetar padrões no desempenho escolar de estudantes, relacionando-os com áreas científicas. Utilizando ferramentas estatísticas e de análise de dados, e após entrevistas com os coordenadores das licenciaturas, concluiu-se que áreas como Informática Aplicada, Multimédia, Gestão Estratégica e Business Intelligence apresentam os melhores desempenhos, com notas mais homogêneas e medianas que entre 13 e 15,5 valores. Contrariamente, áreas como Redes, Matemática, Física e Eletromagnetismo, Contabilidade e Finanças e Sistemas Operativos tendem a apresentar desempenhos mais baixos e maior variabilidade, com medianas entre 10,5 e 12 valores, sendo Redes uma área que implica maior dificuldade para os alunos dos vários cursos. A análise revela relações significativas entre disciplinas, e indica que os alunos têm mais dificuldades na área de Informática comparativamente a Gestão. Introdução à Programação é uma unidade curricular essencial no bom desempenho dos alunos das diferentes licenciaturas. É possível concluir que os “bons alunos” são consistentes no seu bom desempenho nas diferentes unidades curriculares, sendo que os horários diurnos apresentam melhores resultados e maior homogeneidade.

Palavras-chave: sucesso académico, ensino superior, *data mining*, desempenho dos estudantes

Classificação JEL: I21 (Análise de Educação); C55 (Grandes Conjuntos de Dados: Modelação e Análise)

Abstract

This dissertation investigates the enrollment of undergraduate students at a public university to detect patterns of student performance and relate them to scientific areas. Using statistical and data analysis tools, and after interviews with the degree coordinators, it was concluded that areas such as Applied Computing, Multimedia, Strategic Management and Business Intelligence show the best performance, with more homogeneous grades and medians between 13 and 15.5. On the contrary, areas such as Networking, Mathematics, Physics and Electromagnetism, Accounting and Finance and Operating Systems tend to show lower performance and greater variability, with medians between 10.5 and 12 points, with Networking being an area of greater difficulty for students on the various courses. The analysis reveals significant relationships between subjects and indicates that students have more difficulties in Computing compared to Management. Introduction to Programming is an essential curricular unit in the good performance of students in the different degrees. It can be concluded that the “good students” are consistent in their good performance in the different curricular units, with the daytime timetables showing better results and greater homogeneity.

Keywords: academic success, higher education, data mining, student performance

JEL Classification: I21 (Analysis of Education); C55 (Large Data Sets: Modeling and Analysis)

Índice

Agradecimentos.....	i
Resumo.....	iii
Abstract.....	v
Índice de Tabelas.....	ix
Índice de Figuras.....	xi
Lista de Abreviaturas.....	xiii
1. Introdução.....	1
2. Revisão da Literatura.....	3
2.1. Critérios de Pesquisa.....	3
2.2. Artigos Analisados.....	4
2.3. Identificação de Temáticas.....	4
2.3.1. Temas Transversais.....	4
2.3.1.1. Distribuição Geográfica dos Estudos.....	4
2.3.1.2. Análise aos Alunos.....	5
2.3.1.3. Análise às Disciplinas.....	6
2.3.2. Questões Levantadas.....	6
2.3.3. <i>Research Gap</i>	6
3. Metodologia e Tratamento de Dados.....	7
3.1. Caracterização do Problema.....	7
3.2. Preparação dos Dados.....	8
3.3. Análise dos Dados.....	9
3.3.1. Questão 1.....	10
3.3.1.1 Engenharia Informática.....	10
3.3.1.2. Engenharia Informática Pós-Laboral.....	11
3.3.1.3. Engenharia de Telecomunicações e Informática.....	12
3.3.1.4. Engenharia de Telecomunicações e Informática Pós-Laboral.....	13
3.3.1.5. Informática e Gestão de Empresas.....	14
3.3.1.6. Informática e Gestão de Empresas Pós-Laboral.....	15
3.3.1.7 Análise Geral.....	16
3.3.2. Questão 2.....	17
3.3.2.1. Engenharia Informática.....	17
3.3.2.2. Engenharia Informática Pós-Laboral.....	18
3.3.2.3. Engenharia de Telecomunicações e Informática.....	19
3.3.2.4. Engenharia de Telecomunicações e Informática Pós-Laboral.....	20

3.3.2.5. Informática e Gestão de Empresas.....	20
3.3.2.6. Informática e Gestão de Empresas Pós-Laboral	21
3.3.3. Questão 3	21
3.3.3.1. Engenharia Informática.....	22
3.3.3.2. Engenharia Informática Pós-Laboral	22
3.3.3.3. Engenharia de Telecomunicações e Informática	22
3.3.3.4. Engenharia de Telecomunicações e Informática Pós-Laboral.....	23
3.3.3.5. Informática e Gestão de Empresas.....	23
3.3.3.6. Informática e Gestão de Empresas Pós-Laboral	23
3.3.4. Questão 4	24
3.3.4.1. Engenharia Informática.....	24
3.3.4.2. Engenharia Informática Pós-Laboral	24
3.3.4.3. Engenharia de Telecomunicações e Informática	25
3.3.4.4. Engenharia de Telecomunicações e Informática Pós-Laboral.....	25
3.3.4.5. Informática e Gestão de Empresas.....	25
3.3.4.6. Informática e Gestão de Empresas Pós-Laboral	26
4. Resultados e Discussão.....	27
4.1. Questão 1	27
4.2. Questão 2	28
4.3. Questão 3	31
4.4. Questão 4	32
5. Conclusões.....	33
5.1. Limitações e Trabalhos Futuros.....	35
Referências Bibliográficas.....	39
Anexos	43
Anexo A: Análise de Literatura	43
Anexo B: Tratamento dos Dados	45
Anexo C: Análise de Dados da Questão 1	50
Anexo D: Análise de Dados da Questão 2.....	73
Anexo E: Análise de Dados da Questão 3	79
Anexo F: Análise de Dados da Questão 4	85

Índice de Tabelas

Tabela 1. Caracterização dos Artigos Analisados.....	43
Tabela 2. Descrição dos Objetivos, Amostra e Resultados dos Artigos Analisados.....	44
Tabela 3. Identificação dos Locais de Estudo dos Artigos Analisados.....	44
Tabela 4. Descrição das Tabelas da Base de Dados.....	45
Tabela 5. Descrição da Base de Dados.....	46
Tabela 6. Tipo de Dados.....	49
Tabela 7. Descrição dos Cursos.....	49
Tabela 8. Unidades Curriculares por Curso.....	49
Tabela 9. Pares Relevantes por Curso.....	50
Tabela 10. Distribuição de Unidades Curriculares e Áreas por <i>Cluster</i> EI.....	51
Tabela 11. Distribuição de Unidades Curriculares e Áreas por <i>Cluster</i> EI PL.....	54
Tabela 12. Distribuição de Unidades Curriculares e Áreas por <i>Cluster</i> ETI.....	58
Tabela 13. Distribuição de Unidades Curriculares e Áreas por <i>Cluster</i> ETI PL.....	62
Tabela 14. Distribuição de Unidades Curriculares e Áreas por <i>Cluster</i> IGE.....	65
Tabela 15. Distribuição de Unidades Curriculares e Áreas por <i>Cluster</i> IGE PL.....	69
Tabela 16. Pares Relevantes, Áreas e Valores de Diferença EI.....	73
Tabela 17. Pares Relevantes, Áreas e Valores de Diferença EI PL.....	74
Tabela 18. Pares Relevantes, Áreas e Valores de Diferença ETI.....	75
Tabela 19. Pares Relevantes, Áreas e Valores de Diferença ETI PL.....	76
Tabela 20. Pares Relevantes, Áreas e Valores de Diferença IGE.....	77
Tabela 21. Pares Relevantes, Áreas e Valores de Diferença IGE PL.....	78
Tabela 22. Análise Estatística dos 10 Alunos com Maior Média EI.....	85
Tabela 23. Média do Último Colocado em EI ao Longo dos Anos Letivos.....	85
Tabela 24. Análise Estatística dos 10 Alunos com Maior Média EI PL.....	85
Tabela 25. Média do Último Colocado em EI PL ao Longo dos Anos Letivos.....	85
Tabela 26. Análise Estatística dos 10 Alunos com Maior Média ETI.....	86
Tabela 27. Média do Último Colocado em ETI ao Longo dos Anos Letivos.....	86
Tabela 28. Análise Estatística dos 10 Alunos com Maior Média ETI PL.....	86
Tabela 29. Análise Estatística dos 10 Alunos com Maior Média IGE.....	86
Tabela 30. Média do Último Colocado em IGE ao Longo dos Anos Letivos.....	87
Tabela 31. Análise Estatística dos 10 Alunos com Maior Média IGE PL.....	87
Tabela 32. Média do Último Colocado em IGE PL ao Longo dos Anos Letivos.....	87

Índice de Figuras

Figura 1. Fases da Revisão Sistemática da Literatura, com Base no Método PRISMA.	43
Figura 2. Organização da Metodologia.	45
Figura 3. Modelo Relacional	46
Figura 4. Scree Plot EI.	50
Figura 5. Gráfico de Dispersão EI.	50
Figura 6. Médias das Áreas EI	52
Figura 7. Áreas com Menor Média EI.	52
Figura 8. Áreas com Maior Percentagem de Alunos Reprovados EI.	53
Figura 9. Áreas com Maior Percentagem de Alunos Não Avaliados EI.	53
Figura 10. Scree Plot EI PL.	53
Figura 11. Gráfico de Dispersão EI PL.	54
Figura 12. Médias das Áreas EI PL.	55
Figura 13. Áreas com Menor Média EI PL.	55
Figura 14. Áreas com Maior Percentagem de Alunos Reprovados EI PL.	56
Figura 15. Áreas com Maior Percentagem de Alunos Não Avaliados EI PL.	56
Figura 16. Scree Plot ETI.	56
Figura 17. Gráfico de Dispersão ETI.	57
Figura 18. Médias das Áreas ETI.	59
Figura 19. Áreas com Menor Média ETI.	59
Figura 20. Áreas com Maior Percentagem de Alunos Reprovados ETI.	60
Figura 21. Áreas com Maior Percentagem de Alunos Não Avaliados ETI.	60
Figura 22. Scree Plot ETI PL.	61
Figura 23. Gráfico de Dispersão ETI PL.	61
Figura 24. Médias das Áreas ETI PL.	63
Figura 25. Áreas com Menor Média ETI PL.	63
Figura 26. Áreas com Maior Percentagem de Alunos Reprovados ETI PL.	64
Figura 27. Áreas com Maior Percentagem de Alunos Não Avaliados ETI PL.	64
Figura 28. Scree Plot IGE.	64
Figura 29. Gráfico de Dispersão IGE.	64
Figura 30. Médias das Áreas IGE.	66
Figura 31. Comparação das Médias de Gestão e Informática IGE.	66
Figura 32. Áreas com Menor Média IGE.	67
Figura 33. Áreas com Maior Percentagem de Alunos Reprovados IGE.	67
Figura 34. Áreas com Maior Percentagem de Alunos Não Avaliados IGE.	67
Figura 35. Scree Plot IGE PL.	68
Figura 36. Gráfico de Dispersão IGE PL.	68
Figura 37. Médias das Áreas IGE PL.	70
Figura 38. Comparação das Médias de Gestão e Informática IGE PL.	70
Figura 39. Áreas com Menor Média IGE PL.	71
Figura 40. Áreas com Maior Percentagem de Alunos Reprovados IGE PL.	71
Figura 41. Áreas com Maior Percentagem de Alunos Não Avaliados IGE PL.	71
Figura 42. Comparação das Médias de Programação nos Diferentes Cursos	72
Figura 43. Boxplots por Área EI.	79
Figura 44. Boxplots por Área EI PL.	80

Figura 45. Boxplots por Área ETI.....	81
Figura 46. Boxplots por Área ETI PL.....	82
Figura 47. Boxplots por Área IGE.....	83
Figura 48. Boxplots por Área IGE PL.....	84

Lista de Abreviaturas

EDM	Data Mining Educacional
QI	Questões de Investigação
UC	Unidade Curricular
KI	Integração de Conhecimento (“Knowledge Integration” (Robbiano et al., 2020))
GPA	Média Final (<i>Grade Point Average</i>)
LMS	Sistema de Gestão de Aprendizagem Online da Universidade (“University Online Learning Management System” (Wells et al., 2021))
SSE	Soma dos Erros Quadrados

CAPÍTULO 1

Introdução

Existem diversas áreas nas quais as técnicas de *data mining* têm um papel relevante. A análise de dados é crucial para a tomada de decisões, independentemente do tipo de indústria – estando o sistema educativo incluído (Goyal & Vohra, 2012).

Apesar de *data mining* ter sido implementado com sucesso no mundo empresarial há algum tempo, o seu uso na educação superior é relativamente novo, e o interesse em utilizar esta técnica para investigação na área da educação tem vindo a crescer nos últimos anos (Osmanbegovic & Suljic, 2012). Neste sentido, Data Mining Educacional (EDM) é uma área de investigação independente que tem emergido, o que culminou no estabelecimento de uma conferência anual e de um jornal sobre o tema, em 2008 (Baker, 2010).

Sendo que *data mining* tem como objetivo extrair amostras implícitas e interessantes (Azuaje et al., 2006), tendências e informação dos dados, todos os participantes do processo educativo são beneficiados (Osmanbegovic & Suljic, 2012).

Um dos maiores desafios que as instituições de ensino enfrentam atualmente é a melhoria da qualidade de decisões por parte da gestão, um processo cada vez mais complexo devido ao crescimento exponencial de dados relativos à educação (Delavari et al., 2008; Goyal & Vohra, 2012), mas a importância dos dados não está limitada à gestão. Ao converter os dados em conhecimento, todas as partes são beneficiadas, ou seja, estudantes, professores, administração, e comunidade social (Osmanbegovic & Suljic, 2012).

A previsão do sucesso escolar é um tópico que tem sido largamente estudado em diferentes áreas científicas (Trstenjak & Donko, 2014). O termo “sucesso académico” é um dos mais utilizados na investigação e avaliação educacional no ensino superior (York et al., 2015). No entanto, a definição de “sucesso académico” é bastante abrangente e pode incluir diferentes aspetos. A maioria dos estudos sugere que o maior indicativo do sucesso dos alunos é terminar o seu curso (Shahiri et al., 2015), mas a obtenção de competências específicas de preparação de carreira podem ser também um indicativo de um aluno bem sucedido (Levine et al., 2023).

Apesar de ser uma temática bastante abrangente, o ponto de partida é analisar os dados disponíveis de forma a conhecer melhor o desempenho dos estudantes no processo de avaliação formal. Só com essa informação é que se criam as condições indicadas para que o corpo docente da instituição seja capaz de rever práticas de ensino e avaliação que poderão garantir o melhor aproveitamento dos alunos.

Deste modo, o presente estudo procura responder a quatro questões de investigação (QI) que foram definidas com o intuito de orientar a mesma:

QII – O desempenho dos alunos varia consoante o curso e as áreas científicas das unidades curriculares (UC)?

Q12 – Existem UC cujo desempenho dos alunos seja dependente do desempenho que os mesmos alunos tiveram em UC precedentes?

Q13 – Como é que se comportam as notas dos alunos dentro de cada área científica de cada curso, em termos de mediana, das notas da maioria dos alunos, de notas mínimas e máximas, e de variabilidade?

Q14 – Terão os “bons alunos” um bom desempenho em todas as UC?

É nos estudantes do ISCTE-IUL que este estudo se irá focar, mais propriamente em três licenciaturas da área de Informática, sendo que são analisados os alunos com horário diurno e pós-laboral. Será analisado o desempenho dos alunos, de forma a comparar entre disciplinas, cursos, horários e anos letivos. Para a execução do estudo, são analisadas as inscrições dos alunos da universidade pública, recorrendo ao sistema de informação académico, denominado de Fénix. Adicionalmente, são realizadas entrevistas aos coordenadores dos cursos analisados, para melhor entender os resultados obtidos. O objetivo final passa por transmitir informação de suporte para que seja possível uma melhor adequação do plano curricular e dos mecanismos de avaliação das UC ao contexto dos alunos do ISCTE-IUL.

O presente estudo é composto por 4 capítulos. Primeiramente, na Revisão de Literatura, são identificados estudos semelhantes e *research gaps* na investigação científica. No capítulo referente à Metodologia e Tratamento de Dados, explora-se a metodologia e os dados utilizados, explicam-se os tratamentos aplicados e as análises realizadas de forma a responder às questões de investigação definidas. No capítulo de Resultados e Discussão, revisitam-se os resultados obtidos e faz-se uma apreciação crítica dos mesmos, tendo em conta as entrevistas. De forma resumida, foi possível concluir que a área científica Redes apresenta dificuldades para os alunos analisados, que a UC Introdução à Programação é bastante influente no desempenho dos alunos dos vários cursos, que as áreas com menores resultados apresentam maior variabilidade, e que os “bons alunos” têm tendência a ter bons resultados em todas as disciplinas, sendo que os resultados são mais homogêneos se encontram nos horários diurnos. Por fim, no capítulo das Conclusões é feita uma análise geral das ilações obtidas, e consideram-se as limitações do presente estudo e os trabalhos futuros a realizar.

CAPÍTULO 2

Revisão da Literatura

O objetivo do presente capítulo prende-se na identificação e análise da literatura com maior relevância para o estudo. Apenas através desta análise é possível compreender o conhecimento científico que existe na área da educação, nomeadamente das características dos estudantes da área de Informática do ensino superior.

Este capítulo está dividido em três partes. A primeira tem como propósito a definição dos métodos de pesquisa utilizados, sendo possível identificar qual a literatura com maior relevância para sustentar o estudo. A segunda parte trata de classificar os artigos selecionados para que, na terceira parte, seja possível verificar as temáticas transversais aos diferentes artigos, bem como detetar o *research gap* na literatura científica.

2.1. Critérios de Pesquisa

Com o intuito de conhecer os estudos que têm sido realizados na área da educação, mais especificamente no contexto do sucesso escolar no ensino superior, e utilizando ferramentas de *big data* e *data mining*, foi elaborada uma pesquisa. Por se tratar de dados relativos à plataforma Fénix, utilizada no ISCTE-IUL, a pesquisa começou pela análise de duas dissertações que utilizaram dados semelhantes aos utilizados no presente estudo, (Gil, 2019; Miguel et al., 2020), e das referências das mesmas.

Adicionalmente, foi realizada a revisão sistemática da literatura, baseada na abordagem PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis)¹, durante o mês de Junho, nas bases de dados Scopus² e Web Of Science³. De forma a ajustar a investigação à temática pretendida, foram pesquisados os conceitos “education”, “student”, “computing”, “success”, “big data” ou “data analytics”, “analyzing” ou “analysing”, “courses”, “grades”, “curriculum” ou “curricular”. A pesquisa foi realizada através das consultas ("education" AND "big data" AND "student" AND "computing" AND "success"), ("education" AND "data anaylites" AND "student" AND "computing" AND "success") e (“analyzing” AND “courses” AND “student” AND “grades” AND “education” AND “computing” AND (“curriculum” OR “curricular”)), filtrando apenas os artigos dos últimos 10 anos. A utilização de palavras-chave em inglês permitiu alargar o espectro da pesquisa.

Tal como é possível observar no Anexo A - Figura 1, através das mesmas, foi possível obter 103 documentos do repositório Scopus e 34 documentos do repositório Web Of Science, resultando num total de 137 artigos. Após a eliminação dos 19 documentos duplicados, restaram 118 estudos.

¹<http://www.prisma-statement.org/>

²<http://scopus.com/>

³<http://www.webofscience.com/>

De seguida, realizou-se a triagem dos artigos encontrados através da leitura dos títulos e resumos dos mesmos. Após este exercício, foram selecionados 28 artigos com interesse para a análise, tendo em conta o âmbito do estudo. Com a leitura dos mesmos, foi possível concluir que apenas 4 seriam incluídos na análise, uma vez que estes são estudos em que são analisados dados do desempenho de estudantes nas unidades curriculares.

2.2. Artigos Analisados

De forma a identificar as semelhanças e diferenças dos estudos selecionados, elaboraram-se tabelas seguintes que permitem a comparação direta dos mesmos. O Anexo A - Tabela 1 apresenta a caracterização a nível de título, autores, ano de publicação e número de citações, enquanto o Anexo A - Tabela 2 se foca no objetivo, metodologia, dados analisados e principais resultados.

2.3. Identificação de Temáticas

Após a análise e classificação dos artigos relevantes, é possível avaliar e discutir os resultados que os diferentes estudos propõem. Neste sentido, surgem dois temas transversais a vários artigos. Adicionalmente, analisa-se a distribuição geográfica dos estudos, abordam-se as questões levantadas pelos estudos analisados, e determinam-se recomendações para investigações futuras.

2.3.1. Temas Transversais

Assinalaram-se as ideias inerentes a diferentes artigos, de forma a reconhecer as semelhanças identificadas nas conclusões dos mesmos.

2.3.1.1. Distribuição Geográfica dos Estudos

Com o intuito de melhor posicionar a análise, analisaram-se as diferenças geográficas entre os vários países em estudo através do Anexo A - Tabela 3.

É possível concluir que, dos 4 estudos analisados, nenhum foi realizado na Europa. Metade dos artigos foram realizados nos Estados Unidos da América, 1 analisou dados de alunos na Austrália, e outro analisou estudantes na China.

Dado que Portugal, o país do presente estudo, ou mesmo a Europa, não fazem parte do conjunto de países, é possível observar os resultados dos 4 artigos e relacioná-los com os resultados obtidos com o mesmo grau de importância.

Os primeiros 2 estudos, realizados nos Estados Unidos da América, procuram analisar a correlação entre as disciplinas lecionadas, sendo que no primeiro, são também analisados os tópicos, com o objetivo de redesenhar o curso de Engenharia Eletrónica. Concluiu-se acerca da importância de fatores como a abordagem do instrutor, que podem ter um maior peso no sucesso dos alunos (Nash et al., 2021). O segundo procura comparar as médias finais dos alunos antes e após a adoção de atividades de integração de conhecimento (KI), de forma a perceber o impacto das mesmas na aprendizagem dos alunos. Foi possível perceber que, para os alunos com acesso às atividade KI, a média final é menos indicativa

do seu desempenho (Robbiano et al., 2020). Ambos os artigos mencionam pré- e pós-requisito/requisito.

Esta distinção traduz-se em disciplinas cujas qualificações os alunos têm de ter para serem admitidos (pré-requisito) e disciplinas cujas qualificações o aluno tem de ter para completar as mesmas (pós-requisito ou requisito).

O terceiro estudo foi realizado na Austrália, procurando analisar a atividade dos alunos de Informática no University Online Learning Management System (LMS). Foi possível observar que o período de exames e tarefas de avaliação com maior peso na nota levam os alunos ao LMS (Wells et al., 2021).

O último estudo debruçou-se sobre a popularidade das disciplinas selecionados pelos alunos. Neste artigo, é possível concluir acerca da importância da opinião que as pessoas com quem os alunos convivem têm, bem como do modo de ensino das disciplinas. Concluiu-se também acerca das disciplinas mais populares e menos populares, sendo que as primeiras são Astronomia, Eletrónica, Tecnologia de Informação e Física, e as segundas Geografia e Dados e Informação (Yungang et al., 2017).

2.3.1.2. Análise aos Alunos

Um objetivo comum aos artigos selecionados é melhorar o desempenho dos estudantes. Para tal, foram utilizadas diferentes abordagens e ferramentas para analisar o comportamento dos alunos, de forma a compreender a melhor forma de atuar.

Uma das abordagens mais comuns é a análise das notas dos alunos. Esta pode ser feita através das notas individuais por disciplina e/ou da média final do curso (Nash et al., 2021; Robbiano et al., 2020; Wells et al., 2021). Este método permite compreender o desempenho do aluno tendo em conta o tema específico da disciplina, ou analisar o seu desempenho ao longo do curso. No entanto, existem certos fatores que podem influenciar estes resultados e que nem sempre são fáceis de identificar, tais como a aptidão de um aluno para um tópico ou um tipo de avaliação, os critérios de avaliação de diferentes professores, aspetos pessoais e externos que acabam por ter influência no aluno, entre outros.

Outra forma de avaliação dos estudantes pode ser a sua admissão para disciplinas (Robbiano et al., 2020). Sendo necessários conhecimentos para a admissão em determinadas disciplinas ou cursos, é interessante comparar o desempenho dos alunos nesta tipologia de disciplina em comparação com as que não necessitam de conhecimentos específicos para a sua admissão. Desta forma, é colocada a questão: será que a seleção de alunos na fase de admissão garante que os resultados sejam mais positivos, ou um aluno com menos conhecimento *à priori* pode ter mais capacidade para aprender e ter sucesso na disciplina?

O terceiro tipo de análise é referente à atividade dos alunos. Neste caso, a atividade pode ser relativa ao acesso e interação com o sistema *online* da universidade, ou pode estar relacionada com a pesquisa e seleção das disciplinas (Wells et al., 2021; Yungang et al., 2017).

2.3.1.3. Análise às Disciplinas

Com o intuito de melhorar o planejamento curricular de forma a concretizar os objetivos de desempenho propostos aos alunos, é necessário analisar e rever as disciplinas que compõe cada curso.

Em dois dos artigos, foi avaliada a correlação entre pares de disciplinas de forma a perceber se existe ligação entre o desempenho das diferentes disciplinas (Nash et al., 2021; Robbiano et al., 2020). Um dos artigos foi mais além e investigou acerca da correlação entre tópicos lecionados (Robbiano et al., 2020).

Metade dos estudos analisa as disciplinas que compõe os dados a partir de diversas métricas, comparando os resultados das mesmas (Wells et al., 2021; Yungang et al., 2017). O quarto artigo explora ainda a popularidade das disciplinas através da pesquisa e escolha das mesmas (Yungang et al., 2017). Para além de ser essencial garantir o sucesso dos estudantes, é também crucial perceber o que os atrai para certos temas e o porquê de decidirem determinada disciplina ou curso.

2.3.2. Questões Levantadas

A análise dos diferentes artigos permitiu assinalar questões levantadas pelos mesmos, relativamente a discrepâncias de informação e necessidade de investigações mais aprofundadas sobre algumas áreas.

O primeiro artigo sugere estudar a relação entre os temas das disciplinas e as respetivas atividades de avaliação, de forma a analisar a performance dos alunos por tópico. Sugere ainda a possibilidade analisar, não só a hierárquica de tópicos do currículo, mas também a influência desta no desempenho dos alunos, com o objetivo de ter informação valiosa para redesenhar o planejamento curricular (Nash et al., 2021).

É também proposta a possível intervenção a alunos com desempenhos mais negativos, de forma a melhorar as suas notas através dos níveis de atividade e, conseqüentemente, impactar a sua retenção no curso (Wells et al., 2021).

Por fim, a hipótese de refazer o mesmo estudo com uma maior quantidade de dados, quer de cursos, semestres ou alunos, pode levar a conclusões mais interessantes (Yungang et al., 2017).

2.3.3. Research Gap

Ao analisar os artigos com o intuito de identificar as lacunas de conhecimento científico, é possível concluir que nem todos artigos utilizaram bases de dados passíveis de serem consideradas *big data*, devido à pequena quantidade de observações. O primeiro artigo identifica algumas variáveis mais difíceis de contabilizar, que podiam ter interesse para a classificação da amostra ou mesmo para a explicação de alguns resultados, tais como o passa-palavra, a forma como o instrutor se relaciona com os estudantes, a confiança que o aluno tem em si, entre outras (Yungang et al., 2017).

CAPÍTULO 3

Metodologia e Tratamento de Dados

Sendo o presente estudo no âmbito da área de Ciência de Dados, com foco em *data mining*, a orientação do mesmo é feita através de três etapas que pretendem assegurar um estudo incremental e cíclico. As etapas incluem a caracterização do problema, através da caracterização dos dados e variáveis relevantes para o mesmo, a preparação dos dados, e a análise dos mesmos. Dentro do âmbito da Ciência de Dados, a metodologia Cross-Industry Standard Process for Data Mining (CRISP-DM) (Shearer, 2000) é bastante utilizada para orientação dos estudos. No entanto, para o caso da presente análise, optou-se por não seguir esta metodologia, pois certas etapas, tais como *Modeling e Deployment*, não são aplicáveis.

De forma a melhorar a interpretação dos resultados obtidos nas diferentes análises, foi importante conhecer as opiniões dos professores que coordenam cada um dos cursos estudados no presente trabalho. Para tal, foram realizadas 3 conversas com os respetivos coordenadores de cada licenciatura para que fossem apresentadas e discutidas as conclusões obtidas.

O Anexo B - Figura 2 apresenta, de forma resumida, o fluxograma das diferentes etapas do estudo.

3.1. Caracterização do Problema

O foco da presente análise é a fornecer informação para a otimização do planeamento curricular, de forma que este garanta o maior sucesso escolar, tanto para os alunos, como para os docentes. É crucial compreender as diferenças entre o desempenho dos alunos dos diferentes cursos, das diferentes unidades curriculares e das diferentes áreas, para que possam ser tomadas decisões que beneficiem todas as partes envolvidas. A organização das disciplinas pode ser diferente ou igual de curso para curso, tendo em conta as necessidades dos estudantes de cada curso.

Para dar início a este processo, foi necessário analisar a base de dados fornecida pelos orientadores com informação retirada da plataforma Fénix, que compreende os dados de alunos do ISCTE-IUL, desde a sua inscrição às notas de cada ano letivo. Para tal, foi utilizada a linguagem SQL através da ferramenta phpMyAdmin e da Linha de Comandos, sendo que a última permite importar, analisar e exportar bases de dados de maior dimensão em menos tempo.

O Anexo B - Tabela 4 descreve as diferentes tabelas que compõe a base de dados. O Anexo B - Figura 3 representa o modelo relacional da base de dados utilizada. De forma a contextualizar o modelo relacional, o Anexo B - Tabela 5 indica os atributos de cada tabela do modelo. O Anexo B - Tabela 6 descreve os conceitos utilizados na tabela anterior para descrever o tipo de dados de cada coluna.

Foram selecionadas as variáveis “studentId” da tabela “Candidacy”, “degreeNamePt”, “yearName” da tabela “executionYear”, “curricularUnitName” e “gradeValue”, de forma a reunir, da melhor forma possível, a informação relativa às notas dos estudantes dos cursos definidos.

Definiu-se que os cursos a analisar seriam da área de Informática, sendo estes a Licenciatura de Engenharia Informática (EI), a Licenciatura de Engenharia de Telecomunicações e Informática (ETI) e a Licenciatura de Informática e Gestão de Empresas (IGE), dado que seria mais fácil aceder a informação qualitativa para auxiliar a compreensão dos resultados. Para cada um destes cursos, foi também analisada a licenciatura correspondente em horário pós-laboral: Licenciatura de Engenharia Informática Pós-Laboral (EI PL), a Licenciatura de Engenharia de Telecomunicações e Informática Pós-Laboral (ETI PL) e a Licenciatura de Informática e Gestão de Empresas Pós-Laboral (IGE PL). No total, foram estudados seis cursos. O Anexo B - Tabela 7 descreve informações de cada curso disponíveis na base de dados.

No *dataset* existem inscrições compreendidas entre os anos letivos de 1997/1998 e 2019/2020, sendo que alguns cursos têm anos letivos que não foram considerados por haver pouca ou nenhuma informação. Os anos letivos antes do Tratado de Bolonha são considerados, havendo UC com a designação “Pré-Bolonha” por se terem realizado antes do ano letivo 2006/2007 (Malva, 2021).

Relativamente às notas atribuídas às quais foi possível ter acesso, estas são relativas à última nota do aluno na UC e no ano letivo, não sendo possível distinguir entre primeira ou segunda fase de avaliação. Raros eram os casos em que as notas de reprovação eram numéricas, ou seja, de 0 a 9 valores. Na sua maioria, as notas indicativas de reprovação seriam “RE” ou “F”. As notas que indicam que o aluno não foi avaliado seriam “NA”, “D” ou um campo vazio, sendo que este último foi alterado para “-“ de forma a garantir que era analisado sem erros. Um aluno não avaliado significa que não realizou nenhuma avaliação, quer seja teste, exame ou trabalho, pelo que o professor responsável pela sua avaliação não teve ferramentas para indicar a nota final. As notas positivas são os valores entre 10 e 20.

Após a compreensão dos dados, foi possível definir as questões às quais o presente estudo responde.

3.2. Preparação dos Dados

Após reunidos os dados necessários para responder às questões propostas, estes passam por um processo de preparação e tratamento que garante maior qualidade nos resultados obtidos.

Foram elaboradas tabelas com o intuito de organizar a informação de forma útil e coerente. A primeira tabela inclui as unidades curriculares, com a indicação dos cursos, horários e semestres nas quais são lecionadas, e a respetiva área à qual pertencem. A tabela compreende, ainda, o número de inscrições, a média, o número de alunos reprovados e o número de alunos não avaliados por UC, por curso e por ano letivo. Das 226 UC lecionadas nas 6 licenciaturas, 8% foram eliminadas por terem 3 ou menos anos letivos com inscrições.

Foram elaboradas tabelas por curso, com a indicação das notas de cada aluno nas respectivas UC, e em cada ano letivo.

Através da análise da média, quantidade de alunos com notas positivas (neste caso, igual ou superior a 10) e desvio padrão, foram eliminadas UC cuja média fosse 0 e/ou cujos alunos inscritos fossem menos de 30 na sua totalidade (Fraenkel et al., 2011). Posteriormente, foram eliminadas UC que não representavam o *core* de cada curso, ou seja, disciplinas de áreas não essenciais a cada licenciatura. O Anexo B - Tabela 8 indica o número de UC em análise para cada curso. Esta decisão foi tomada pela falta de informação relevante a analisar nas UC selecionadas de cada curso.

Às tabelas por curso, foram adicionados os cálculos da média por ano e da média por aluno. A análise destas tabelas foi essencial para dar resposta às diferentes perguntas, sendo que, para a Questão 2, foram necessários passos adicionais.

A Questão 2 tenta compreender se existem dependências entre as UC, fazendo com que a nota de uma disciplina - que pode ser chamada de UC dependente - seja influenciada por uma outra – UC independente - lecionada anteriormente. Para investigar a possível relação entre as UC, foi definido um algoritmo que fosse capaz de calcular a média da UC dependente quando os alunos tinham reprovado à UC independente, e a média da UC dependente quando os alunos tinham tido sucesso, leia-se não reprovado, na UC independente. As duas médias foram calculadas para todos os pares de UC possíveis e, para tal, foi preciso programar na linguagem SQL, na plataforma phpMyAdmin, e na linguagem Python, através das plataformas Google Colab Notebook e Visual Studio Code, de forma a elaborar um algoritmo. Manualmente, foram identificados os pares mais relevantes. Definiu-se como relevante uma diferença de 2 a 10 valores entre as duas médias, de reprovação e de sucesso, dado que a partir desse valor uma das médias seria zero. Com uma diferença de menos de 2 valores entre as duas médias, não seria possível definir os pares de UC mais relevantes, pois estariam incluídos todos os pares. Após esta filtragem, foram selecionados apenas os pares que seguiam a ordem temporal correta, tendo em conta os semestres e os anos letivos em que cada UC é lecionada. Para tal, garantiu-se que, dentro de cada par, a UC independente era lecionada anteriormente à UC dependente, não sendo obrigatório ser no semestre ou ano anterior. Em tom de curiosidade, foram recolhidos 2 pares adicionais em todos os cursos, correspondentes às UC de Introdução à Programação (UC Independente), Programação Concorrente e Distribuída (UC Dependente) e Programação Orientada a Objetos (UC Independente), que foram mencionadas pelos professores orientadores como possíveis casos de interesse. Adicionalmente, os professores coordenadores dos cursos analisados indicaram pares com possível interesse que também foram adicionados posteriormente. O Anexo B - Tabela 9 indica a quantidade de pares relevantes recolhidos, já com a adição dos pares indicados.

3.3. Análise dos Dados

É nesta fase que são avaliados os dados, permitindo extrair informação valiosa dos mesmos. São definidos os modelos que possibilitam dar resposta às questões identificadas inicialmente. Trata-se de um capítulo mais descritivo, sendo que o seguinte terá a componente de análise crítica dos resultados obtidos.

3.3.1. Questão 1

A Questão 1 procura aferir se o desempenho dos estudantes é diferente consoante o seu curso e as áreas das UC que tem ao longo do mesmo.

Foi realizada uma primeira análise geral de cada curso, de forma a compreender se as UC se agrupavam em *clusters* naturalmente. Foi também realizada uma análise por áreas e foram identificadas as áreas que apresentam maior dificuldade aos alunos de cada licenciatura.

3.3.1.1 Engenharia Informática

a) Scree Plot e Clustering

Foram analisadas as UC da Licenciatura de Engenharia Informática com o objetivo de identificar *clusters* às quais podem pertencer, com base nas notas recolhidas dos alunos deste curso. Foi utilizado o método Scree Plot (Anexo C - Figura 4) onde se examina a relação entre a Soma dos Erros Quadrados (SSE) e o número de *clusters*. Neste caso, a regra do cotovelo sugere 3 *clusters*.

O Silhouette Score obtido com 3 *clusters* é de 0.448, o que é uma medida razoável (Apoorva & Sangeetha, 2024). Neste sentido, o Anexo C - Figura 5 apresenta o gráfico de dispersão onde é possível observar que o *cluster* 1 apresenta médias elevadas ao contrário do *cluster* 2, e que o *cluster* 0 se destaca pela variedade de notas atribuídas nas UC correspondentes.

Na Anexo C - Tabela 10, é possível observar quais as UC que pertencem a cada *cluster*, e a área à qual correspondem. Dado o que foi concluído anteriormente, é possível ver que as UC da área de Programação têm, na sua maioria, um maior desvio padrão, ou seja, os alunos têm notas muito dispersas nas disciplinas pertencentes a este *cluster*. No *cluster* 1 encontram-se as UC da área Multimédia, que têm melhores resultados. Já as áreas de Matemática, Arquitetura de Computadores e Redes pertencem maioritariamente ao *cluster* 2, o que indica uma média mais baixa nestas disciplinas.

b) Análise por Área

Dentro da licenciatura de Engenharia Informática, foram destacadas 12 áreas cujas médias podemos comparar no Anexo C - Figura 6. Tal como visto no ponto anterior, a área de Multimédia é a que apresenta maior média, seguida pela área de Informática Aplicada. As áreas com menor média são Redes, Contabilidade e Finanças e Matemática, o que pode indicar que são disciplinas com um maior nível de dificuldade, ou ainda que são áreas para as quais os alunos desta licenciatura terão menos aptidão, como pode ser o caso de Contabilidade e Finanças. É de notar que, após a preparação dos dados, as áreas Informática Aplicada e Contabilidade e Finanças têm apenas uma UC.

c) Áreas Mais Difíceis

Tendo em conta os dados disponíveis, as áreas que apresentam maior dificuldade para os estudantes foram definidas como aquelas com menor média, e com maiores percentagens de alunos reprovados e de alunos não avaliados. Em cada curso, foram identificadas as 5 áreas com menor média, maior percentagem de alunos reprovados e maior percentagem de alunos não avaliados.

No caso da licenciatura de Engenharia de Informática, o Anexo C - Figura 7 indica as 5 áreas com menor média, o Anexo C - Figura 8 as 5 áreas com maior percentagem de

alunos reprovados, e o Anexo C - Figura 9 as 5 áreas com maior percentagem de alunos não avaliados. Ao observar os 3 gráficos, é possível concluir que as áreas de Redes e de Arquitetura de Computadores apresentam uma maior dificuldade para os alunos da licenciatura, sendo que as 5 áreas representam as disciplinas onde os estudantes têm o pior desempenho.

3.3.1.2. Engenharia Informática Pós-Laboral

a) Scree Plot e Clustering

À semelhança do curso anterior, as UC da licenciatura de Engenharia Informática Pós-laboral aparecem distribuídas por 3 *clusters*, tal como é possível concluir através do método do cotovelo no Anexo C - Figura 10.

O Silhouette Score obtido com 3 clusters indica que é uma boa decisão, sendo este de 0.526, e no Anexo C - Figura 11 podemos observar o gráfico de dispersão (Apoorva & Sangeetha, 2024). Mais uma vez, o cluster 0 é o que tem maior variedade de notas, enquanto o cluster 2 tem a melhor média. O cluster 1 tem notas mais baixas e maior densidade.

Tabela 11 permite observar as disciplinas incluídas em cada *cluster* e as respetivas áreas. No *cluster 2*, encontram-se as únicas UC da licenciatura com média superior a 15, sendo estas de diversas áreas. De forma semelhante ao curso de Engenharia Informática em horário diurno, as áreas de Matemática, Redes e Contabilidade e Finanças têm o pior desempenho. As disciplinas que fazem parte do *cluster 0* apresentam uma maior disparidade de notas, sendo que é possível destacar as áreas de Programação e Inteligência Artificial. Dar nota de que existe apenas uma UC nas áreas de Sistemas Operativos e Física e Eletromagnetismo.

b) Análise por Área

Mais uma vez, urge a comparação entre as diferentes áreas da licenciatura de Engenharia Informática Pós-Laboral. No Anexo C - Figura 12 é possível observar um grande destaque positivo na área de Informática Aplicada, que corresponde apenas à UC “Planeamento de Projetos Utilizando Ferramentas Informáticas (Msproject)”. Sendo que esta é uma UC de preparação e organização de projetos, é adequado que as notas sejam maioritariamente positivas. Também em destaque está a área de Multimédia, à semelhança do curso de horário diurno. Com médias menores, observam-se as áreas de Redes e Programação, o que leva à conclusão de que a área de Redes é das áreas que impõe maior dificuldade aos estudantes. Comparando com o curso em horário diurno, a média da maioria das áreas científicas do curso pós-laboral é superior ao curso analisado anteriormente.

c) Áreas Mais Difíceis

Da mesma forma que foram destacadas 5 áreas com resultados menos positivos no curso analisado anteriormente, foram também identificadas as 5 áreas com menor média (Anexo C - Figura 13), as 5 áreas com maior percentagem de alunos reprovados (Anexo C - Figura 14), e as 5 áreas com maior percentagem de alunos não aprovados (Anexo C - Figura 15). Existe uma clara predominância das áreas de Redes e Matemática, sendo estas as áreas que comportam maior dificuldade para os estudantes. Tal como na licenciatura de Engenharia Informática, os alunos parecem ter menos sucesso na área de Redes. A área de Contabilidade e Finanças tem também uma média baixa e a segunda maior percentagem de alunos reprovados, o que pode indicar uma menor aptidão dos alunos de

Informática para áreas mais ligadas à Gestão. Apesar de não ter uma média particularmente alta, os alunos da licenciatura com horário pós-laboral têm um melhor desempenho na área de Arquitetura de Computadores comparativamente aos alunos em horário diurno. A área de Telecomunicações destaca-se, também, pelas percentagens elevadas no que se refere a alunos reprovados e não avaliados.

3.3.1.3. Engenharia de Telecomunicações e Informática

a) Scree Plot e Clustering

Na licenciatura de Engenharia de Telecomunicações e Informática é também necessário fazer a análise de *clusters*, de forma a descobrir possíveis ligações entre as UC lecionadas. Novamente, o Anexo C - Figura 16 indica o método do cotovelo utilizado para a escolha de 3 *clusters*, e o Silhouette Score de 0.4738 confirma a mesma (Apoorva & Sangeetha, 2024).

O Anexo C - Figura 17 representa o gráfico de dispersão com os 3 *clusters*, onde é possível observar uma grande densidade no *cluster* 1, indicando que a maioria das UC tem médias entre 11 e 13 valores. O *cluster* 2 apresenta uma baixa congruência de notas dos seus alunos, e poucas são as disciplinas presentes no *cluster* 0, sendo que estas têm médias entre 14 e 17 valores.

O Anexo C - Tabela 12 representa a distribuição das UC do curso pelos 3 *clusters* selecionados. Existe um claro domínio do *cluster* 1, onde se reúnem as disciplinas em que os alunos desta licenciatura têm maiores dificuldades. Aqui é possível destacar disciplinas como Redes, Matemática, Arquitetura de Computadores e Física e Eletromagnetismo. Contrariamente, no *cluster* 0 encontram-se as UC em que os estudantes têm melhores resultados, sendo que a Multimédia, Telecomunicações e Informática Aplicada compõe a maioria. Urge reforçar que, neste curso, a única UC de Multimédia, após a preparação dos dados, é Computação Gráfica.

b) Análise por Área

Numa análise pelas diferentes áreas que compõe a licenciatura de Engenharia de Telecomunicações e Informática, o Anexo C - Figura 18 permite destacar positivamente as áreas de Multimédia, Informática Aplicada e Telecomunicações. Ao contrário dos dois cursos analisados anteriormente, a área de Contabilidade e Finanças encontra-se no conjunto de áreas com melhor média, o que pode dever-se ao tipo de conteúdos lecionados nesta UC comparativamente à UC lecionada nas licenciaturas de Engenharia Informática, quer em horário diurno, quer em horário pós-laboral. Com menores médias, é possível identificar as áreas de Matemática, Física e Eletromagnetismo, Bases de Dados, Modelação e Sistemas Operativos.

c) Áreas Mais Difíceis

Com o intuito de apurar as áreas que apresentam maior dificuldade aos estudantes da licenciatura de Engenharia de Telecomunicações e Informática, foram destacadas as 5 áreas com menor média (Anexo C - Figura 19), as 5 áreas com maior percentagem de alunos reprovados (Anexo C - Figura 20) e as 5 áreas com maior percentagem de alunos não avaliados (Anexo C - Figura 21). Sabendo que todas as áreas indicadas nas figuras seguintes podem ser consideradas como sendo difíceis, Física e Eletromagnetismo é comum a todas estas análises. Ao contrário dos cursos anteriores, a área de Redes não é uma das áreas com menos casos de sucesso, apesar de ter 18,04% de alunos não avaliados.

Ainda nesta métrica, há um claro destaque na área de Programação, com 35,95% de alunos não avaliados, apesar de não se encontrar na lista das 5 áreas com menor média.

3.3.1.4. Engenharia de Telecomunicações e Informática Pós-Laboral

a) Scree Plot e Clustering

Através da técnica do cotovelo utilizada no Anexo C - Figura 22, é possível concluir que o curso de horário pós-laboral de Engenharia de Telecomunicações e Informática se deve dividir em 3 *clusters*. O Silhouette Score é de 0.511 o que indica que a divisão é interessante (Apoorva & Sangeetha, 2024).

O Anexo C - Figura 23 representa o gráfico de dispersão dos *clusters* selecionados, sendo que é possível observar uma grande separação entre os *clusters* 0 e 2 e o *cluster* 1, sendo que o último tem apenas 4 disciplinas que representam as UC com melhores resultados. O *cluster* mais populoso é o 0, com as UC que têm menor média, entre 11,5 e 12,5 valores (existindo apenas uma UC que ultrapassa ligeiramente este intervalo). O *cluster* 2 tem maior diversidade de notas atribuídas, mas, no geral, as UC não têm uma média muito elevada.

No Anexo C – Tabela 13, é possível concluir que as UC com melhor desempenho por parte dos alunos pertencem maioritariamente à área de Telecomunicações e à disciplina de “Planeamento de Projetos Utilizando Ferramentas Informáticas (Msproject)”. As UC que apresentam piores resultados são, novamente, nas áreas de Redes, Matemática e Física e Eletromagnetismo. As áreas menos coerentes a nível de notas são bastante diversas, podendo destacar as áreas de Programação e Arquitetura de Computadores.

b) Análise por Área

No Anexo C - Figura 24 é possível observar um grande destaque positivo na área de Informática Aplicada, tal como se verificou no curso pós-laboral analisado anteriormente. No entanto, no caso deste curso, esta área não corresponde apenas a uma UC. Também é possível perceber que há uma grande diferença esta área e as restantes, existindo aproximadamente 1 valor de diferença entre esta e a área com média mais alta seguinte, a de Telecomunicações. Com notas menos favoráveis, destacam-se Bases de Dados, Modelação e Redes. Tendo em conta as médias das áreas científicas de ambos os horários deste curso, não existe um que se destaque.

c) Áreas Mais Difíceis

À semelhança do que foi feito nos cursos anteriores, selecionaram-se as 5 áreas com menor média, as 5 áreas com maior percentagem de alunos reprovados, e as 5 áreas com maior percentagem de alunos não avaliados, no Anexo C, nas Figura 25, Figura 26 e Figura 27, respetivamente. Sem surpresa, Matemática está presente nas diferentes métricas. Algumas áreas passíveis de realce são Redes, Física e Eletromagnetismo e Bases de Dados, Modelação. Tal como no curso de Engenharia de Telecomunicações e Informática em horário diurno, a percentagem de alunos não avaliados em Programação está em destaque, neste caso, deve-se a 42,37% de alunos não avaliados.

3.3.1.5. Informática e Gestão de Empresas

a) Scree Plot e Clustering

À semelhança das análises anteriores, foi feita a análise de Clustering na licenciatura de Informática e Gestão de Empresas. O Anexo C - Figura 28 permite utilizar a técnica do cotovelo para a definição de 3 *clusters*. A Silhouette Score indica 0.458, sendo esta uma medida razoável (Apoorva & Sangeetha, 2024).

O Anexo C - Figura 29 indica a distribuição das UC pelos 3 *clusters*. Ao contrário dos gráficos de dispersão observados anteriormente, este apresenta uma maior densidade entre os 3 *clusters*. O *cluster* 1, indicador das UC com pior desempenho por parte dos alunos, reúne a maioria das disciplinas, enquanto o *cluster* 2, que representa as UC com melhor desempenho, tem apenas 6 disciplinas. Tanto o *cluster* 0 como o 1 apresentam uma alta diversidade de notas, possível de concluir através do desvio padrão.

O Anexo C - Tabela 14 indica a distribuição das UC do curso pelos 3 *clusters*. Tal como foi possível observar na figura anterior, o *cluster* 2 é o mais pequeno, onde estão incluídas as UC com melhores médias, leia-se entre 14,28 e 16,34 valores. Neste *cluster*, encontram-se áreas tanto de Informática, como de Gestão. No *cluster* 1, que indica as disciplinas com notas mais baixas, encontram-se áreas de Matemática, Arquitetura de Computadores, Economia, Estatística, Marketing e Redes. Economia, Marketing e Estatística são áreas que os cursos anteriores não tinham no seu planeamento curricular, e que pertencem à grande área de Gestão que diferencia este curso das licenciaturas anteriores. Numa análise geral, os alunos de Informática e Gestão de Empresas têm dificuldades nas duas áreas que compõe o curso: Gestão e Informática.

b) Análise por Área

Tal como foi referido anteriormente, este curso inclui uma maior diversidade de áreas por ter uma componente de Gestão, ao contrário das licenciaturas de Engenharia Informática e Engenharia de Telecomunicações e Informática. A análise por área possível de ser observada no Anexo C - Figura 30 representa essa mesma diferença, com a adição de áreas como Business Intelligence, Economia, Estatística, Gestão Estratégica, Gestão Geral e Operações, Marketing, Recursos Humanos e Sistemas de Informação Gerais. Através da comparação das médias de todas as áreas é possível destacar positivamente as áreas de Sistemas de Informação Gerais e Business Intelligence, e negativamente, sem surpresa, as áreas de Matemática e Redes.

Devido às características deste curso, surge a necessidade de comparação entre as grandes áreas de Gestão e Informática. O Anexo C - Figura 31 é semelhante ao anterior, mas o vermelho representa a área de Gestão e o azul a área de Informática, e as áreas estão ordenadas por ordem decrescente de média, dentro de cada um dos grupos definidos. Neste caso, existem 9 áreas de Gestão e 10 de Informática. Informática tem áreas com médias menores do que Gestão, mas tem uma maior consistência de média ao longo das diferentes áreas. Em Gestão, a área com maior média é Sistemas de Informações Gerais e a área com menor média é Marketing. Já em Informática, Multimédia e Matemática são as áreas com maior e menor média, respetivamente.

c) Áreas Mais Difíceis

Apesar da comparação entre Gestão e Informática, é interessante perceber quais as áreas nas quais os estudantes têm maiores dificuldades, e a que grandes áreas pertencem essas

mesmas disciplinas. As figuras do Anexo C (Figura 32, Figura 33 e Figura 34) representam as 5 áreas com menor média, as 5 áreas com maior percentagem de alunos reprovados, e as 5 áreas com maior percentagem de alunos não avaliados, respetivamente. Numa análise geral, podemos observar que, das 5 áreas com menor média da licenciatura de Informática e Gestão de Empresas, apenas 2 áreas são da grande área de Informática. Já na figura 32, a totalidade das 5 áreas com maior percentagem de alunos reprovados pertencem a Informática e, na figura seguinte, as 2 áreas com menor percentagem de alunos não avaliados das 5 selecionadas são as únicas dessa seleção da área de Gestão. Apesar da área de Programação não ter uma das 5 menores médias do curso, continua a ser a área com maior percentagem de alunos não avaliados, apesar de ter a menor percentagem de alunos reprovados das 5 áreas selecionadas. As áreas de Redes e Matemática não estão presentes nas 5 áreas com maior percentagem de alunos não avaliados, mas têm as duas menores médias e valores elevados de alunos reprovados, com 38,74% e 36,01%, respetivamente. Ao analisar as duas figuras, é possível concluir que as duas áreas com melhor desempenho são da área de Gestão, e as duas áreas com pior desempenho são da área de Informática, o que indica que a área de Informática tem um grau de dificuldade maior e resultados menos bem conseguidos.

3.3.1.6. Informática e Gestão de Empresas Pós-Laboral

a) Scree Plot e Clustering

As UC da licenciatura de Informática e Gestão de Empresas em horário pós-laboral também se organizam em 3 *clusters*, como indicado pelo método do cotovelo possível de observar no Anexo C - Figura 35. A Silhouette Score para este número de *clusters* é de 0.533, pelo que pode ser considerada uma boa divisão (Apoorva & Sangeetha, 2024).

O Anexo C - Figura 36 demonstra a divisão das UC nos 3 *clusters*, em termos de média e desvio padrão. É possível observar uma maior densidade nas disciplinas com menor média, através do *cluster 1*, à semelhança do curso anterior. As UC com melhor média, entre 14,33 e 16 valores, são apenas 6 e indicadas pelo *cluster 2*.

O Anexo C - Tabela 15 permite analisar cada *cluster* com maior pormenor, a nível de UC e de área à qual pertence. No *cluster 2*, é possível concluir que a maioria das disciplinas com maior média pertence à grande área de Gestão. Já no *cluster 1*, reúnem-se as áreas de Matemática, Programação, Recursos Humanos, Marketing, Redes e Sistemas Operativos. É importante dar nota de que, neste curso, existe apenas uma UC nas áreas de Gestão Estratégica, Recursos Humanos, Marketing e Sistemas Operativos. Ainda no *cluster 1*, está a UC Análise de Dados I, que aparenta ter piores resultados do que a UC seguinte, Análise de Dados II, que se encontra no *cluster 0*.

b) Análise por Área

Ao analisar o Anexo C - Figura 37, é possível concluir que as áreas com menor média são Redes e Sistemas Operativos, sabendo que Sistemas Operativos representa apenas a UC de Sistemas Operativos. Em destaque por ter a média mais elevada está a área de Informática Aplicada, que representa apenas a UC de “Planeamento de Projetos Utilizando Ferramentas Informáticas (Msproject)”, e que tem vindo a ter médias elevadas em todos os cursos nas quais é lecionada. Ligeiramente abaixo, mas ainda em destaque positivo, estão as áreas de Business Intelligence e Gestão Estratégica. Comparando com

o horário diurno, o curso pós-laboral tem médias ligeiramente inferiores na maioria das áreas científicas.

O Anexo C - Figura 38 tem como objetivo facilitar a análise e comparação das duas áreas que caracterizam o curso, Gestão e Informática. As áreas estão divididas nestas duas grandes áreas e ordenadas por ordem decrescente de média. Ao contrário da licenciatura anterior, Informática tem uma área que se destaca claramente das restantes, Informática Aplicada, sendo que as restantes áreas têm menos diferença entre elas. Já na área de Gestão, é possível observar 3 grandes destaques das áreas de Gestão Estratégica, Business Intelligence e Sistemas de Informação Gerais, e uma homogeneidade nas restantes áreas. A área com média mais baixa de Gestão é Economia, e a área mais baixa de Informática é Redes. À semelhança do curso anterior, também no horário pós-laboral duas das três áreas com melhores resultados fazem parte da grande área de Gestão, enquanto as áreas com menores médias pertencem à área de Informática. Mais uma vez, a área de Informática parece apresentar maiores dificuldades aos alunos de Informática e Gestão de Empresas, comparativamente à área de Gestão.

c) Áreas Mais Difíceis

À semelhança da análise realizada nos cursos anteriores, são avaliadas as áreas que podem ser consideradas mais difíceis para os alunos da licenciatura de Informática e Gestão de Empresas Pós-Laboral. Assim, reuniram-se as 5 áreas com menor média (Anexo C - Figura 39), as 5 áreas com maior percentagem de alunos reprovados (Anexo C - Figura 40), e as 5 áreas com maior percentagem de alunos não avaliados (Anexo C - Figura 41). Apesar de nas 5 áreas com menor média, apenas 2 serem da área de Informática, não existem áreas de Gestão nas 5 com maior percentagem de alunos reprovados, e existe apenas 1 área de Gestão nas 5 áreas com maior percentagem de alunos não avaliados. Mais uma vez, Programação continua a ser uma área bastante presente no que toca a alunos reprovados e não avaliados. Redes é, sem dúvida, uma área que representa uma grande dificuldade para os alunos deste curso, com a menor média do curso, 38,54% de alunos reprovados e 26,34% de alunos não avaliados. Da área de Gestão, a cadeira que pode ser considerada mais difícil será Recursos Humanos, apesar da área de Economia ter uma média menor.

3.3.1.7 Análise Geral

a) Área de Programação

As análises realizadas ao longo desta primeira questão exploram cada licenciatura. Neste caso, o Anexo C - Figura 42 representa uma avaliação geral em que todos os cursos são examinados relativamente a uma área, a Programação. Esta área é abrangente aos 6 cursos presentes no estudo, e representa uma dificuldade no planeamento curricular identificada pelos professores orientadores. É possível concluir que, tanto em horário diurno como em horário pós-laboral, o curso com médias mais elevadas em Programação é o de Engenharia Informática. No geral, observa-se uma tendência geral para que as licenciaturas em horário pós-laboral tenham uma média superior às correspondentes em horário diurno, sendo que Informática e Gestão de Empresas foge a esta regra.

3.3.2. Questão 2

A Questão 2 tem como objetivo explorar a hipótese de existirem UC que “dependem” de outras, ou seja, se um aluno reprovar numa UC, tem pior nota numa outra UC?

Para conseguir perceber se existem relações entre disciplinas, foram analisados os pares de UC dependente e independente, sendo que a UC dependente tem obrigatoriamente de ser lecionada após a independente. Para cada par, foram calculadas duas médias, uma com as notas dos alunos da UC dependente quando os mesmos alunos reprovavam à UC independente – média de reprovação, e uma média quando os alunos tinham sucesso na UC independente – média de sucesso. Foram selecionados os pares com diferença entre médias superior a 2 valores e inferior a 10 valores, de forma a destacar os pares de UC cuja influência fosse realmente significativa, sem ultrapassar uma diferença tão grande que implicasse que uma das médias fosse zero. A maioria dos pares tinha uma diferença entre as médias de reprovação e sucesso menor que 2 valores, pelo que se decidiu que uma diferença de, pelo menos, 2 valores seria indicativa de uma dependência relevante. Os pares com diferenças entre 3 e 9 valores correspondiam a UC com menos de 30 alunos, pelo que não foram considerados. Ou seja, para que um par fosse analisado, é necessário que existam pelo menos 30 alunos que tiveram sucesso na UC independente e que completaram a UC dependente, bem como 30 alunos que não tiveram sucesso na UC independente e que completaram a UC dependente, de forma a garantir a consistência na análise. Uma diferença superior a 10 valores implica que uma das médias é zero, o que não permite analisar a dependência entre as UC. Para cada curso, foram estudados os pares identificados e as respetivas áreas.

Foram adicionados os pares de disciplinas Programação Concorrente e Distribuída (UC Dependente) e Introdução à Programação (UC Independente), e Programação Concorrente e Distribuída (UC Dependente) e Programação Orientada para Objetos (UC Independente) ao conjunto de pares relevantes de cada curso, por se tratar de um conjunto de UC com importância para o planeamento curricular. Tendo em conta as necessidades específicas de cada licenciatura e as conversas com cada um dos coordenadores das licenciaturas analisadas, foram adicionados pares sobre os quais existia interesse acerca de uma possível relação. Apesar da diferença entre as médias de reprovação e sucesso dos pares indicados ser menor do que 2 valores, e não seguirem obrigatoriamente a ordem temporal utilizada para os restantes, os pares indicados foram adicionados à lista de pares relevantes, com o objetivo de os analisar.

3.3.2.1. Engenharia Informática

O Anexo D - Tabela 16 indica os 30 pares relevantes na licenciatura de Engenharia Informática. Podemos observar que o par de UC com maior diferença entre a média de reprovação e a média de sucesso é entre a UC Inteligência Artificial e Redes Digitais I – Fundamentos (identificada a negrito), o que se traduz na conclusão de que as notas dos alunos em Inteligência Artificial estão dependentes do seu desempenho em Redes Digitais I – Fundamentos, dado que, ao reprovar na UC independente, neste caso Redes Digitais I – Fundamentos, esta terá de ser repetida posteriormente, o que implica um menor desempenho na UC dependente, quer seja porque os conhecimentos deviam ter sido adquiridos na UC independente, ou porque estar a repetir uma UC afeta o desempenho nas restantes. Também é possível observar que as UC Inteligência Artificial, Mecânica e Eletricidade e Sistemas Operativos são bastante dependentes da aprovação em disciplinas

anteriores. No caso dos dois últimos pares, apesar de terem diferenças abaixo do que foi definido como sendo uma diferença relevante, é de notar que existe uma dependência entre estas UC, estando Programação Concorrente e Distribuída dependente do sucesso ou não dos alunos nas duas disciplinas independentes. Talvez pela área de Programação ser tão essencial, é possível concluir que tem uma vasta influência em diversas UC, estando presente em 15 pares do lado da UC independente, através, maioritariamente, das disciplinas de Introdução à Programação, Programação Concorrente e Distribuída e Programação Orientada para Objetos. As áreas de Redes, Arquitetura de Computadores e Bases de Dados, Modelação são também influentes. Para além dos pares adicionados posteriormente, não existem outros pares compostos apenas por UC de uma só área.

No final da tabela, estão os pares adicionados tendo em consideração possível precedentes identificados pelo professor coordenador do curso, sem ter em conta a ordem temporal seguida na identificação dos restantes pares, nem o mínimo de 2 valores de diferença entre as médias de reprovação e sucesso. Os pares adicionados foram Introdução à Programação (independente) e Programação Orientada para Objetos (dependente), Introdução à Programação (independente) e Algoritmos e Estruturas de Dados (dependente), Algoritmos e Estruturas de Dados (independente) e Programação Orientada para Objetos (dependente), e Programação Concorrente e Distribuída (independente) e Inteligência Artificial (dependente). Os pares entre Introdução à Programação e Programação Orientada para Objetos, e Algoritmos e Estruturas de Dados e Programação Orientada para Objetos apresentam uma diferença significativa, com pelo menos 1,5 valores de diferença entre as médias de reprovação e sucesso. Os restantes pares, apresentam uma relação menos significativa, com cerca de meio valor de diferença.

3.3.2.2. Engenharia Informática Pós-Laboral

No Anexo D - Tabela 17 é possível analisar os 10 pares relevantes do curso com horário pós-laboral de Engenharia Informática. É natural que existam menos pares nos cursos neste regime devido ao menor número de alunos e de UC. Neste caso, nota-se um claro domínio da área de Programação, tanto como UC dependente como independente. É possível observar que Sistemas Operativos influenciam e são influenciados pela área de Programação, sendo que a UC de Sistemas Operativos está dependente da UC Programação Orientada para Objetivos, e influencia a UC Teoria da Computação (par com a maior diferença entre as médias de reprovação e sucesso, assinalada a negrito). Esta última está, ainda, dependente da UC Circuitos para Comunicações. O par que junta Programação Concorrente e Distribuída (dependente) e Introdução à Programação (independente) tem uma maior dependência do que no curso anterior, ao contrário do par Programação Concorrente e Distribuída (dependente) e Programação Orientada para Objetos (independente). Para além destes dois últimos pares, o par de UC Teoria da Computação (dependente) e Algoritmos e Estruturas de Dados (independente) é composto apenas pela área de Programação.

Tal como no curso de horário diurno, foram adicionados os pares identificados pelo professor coordenador da licenciatura. Exceto o par entre Introdução à Programação e Algoritmos e Estruturas de Dados, os pares adicionados apresentaram relações relevantes, sendo que tinham uma diferença maior que 1 valor, chegando até 1,81 de diferença entre as médias de reprovação e sucesso. Também à semelhança do horário diurno, o par com

maior relação é entre Introdução à Programação (independente) e Programação Orientada para Objetos (dependente).

3.3.2.3. Engenharia de Telecomunicações e Informática

O Anexo D - Tabela 18 indica os 25 pares relevantes da licenciatura de Engenharia de Telecomunicações e Informática. É possível observar que as UC de Inteligência Artificial e Ondas e Ótica são grandemente influenciadas pelo desempenho dos alunos em disciplinas que sejam lecionadas anteriormente. As áreas de Física e Eletromagnetismo e Arquitetura de Computadores são as mais influentes. É possível destacar as UC Eletromagnetismo e Teoria dos Circuitos, e uma maior influência das disciplinas de Física e Eletromagnetismo na UC Inteligência Artificial, e das disciplinas de Arquitetura de Computadores na UC Ondas e Ótica. Neste sentido, o par com maior diferença entre as médias de reprovação e sucesso é entre Inteligência Artificial (dependente) e Fundamentos de Arquitetura de Computadores (independente). A diferença entre os pares Programação Concorrente e Distribuída (dependente) e Introdução à Programação (independente) e do par Programação Concorrente e Distribuída (dependente) e Programação Orientada para Objetos (independente) é a menor tendo em conta os cursos já analisados. Para além destes últimos pares, existe um par dentro da mesma área de Física e Eletromagnetismo, sendo este com as UC Ondas e Ótica (dependente) e Eletromagnetismo (independente).

No final da tabela, foram ainda adicionados pares identificados como tendo precedentes entre si pelo professor coordenador da licenciatura, pelo que esta adição é feita sem que os pares garantam uma diferença de pelo menos 2 valores entre as médias de reprovação e de sucesso, e um seguimento temporal tendo em conta a informação disponível indicativa do semestre e ano de cada UC. Estes conjuntos de UC incluem, na ordem indicada, a análise entre Introdução à Programação (independente) e Programação Orientada para Objetos (dependente), Redes Digitais I – Fundamentos (independente) e Redes Digitais II - Sistemas, Aplicações e Serviços (dependente), seguida de Redes Digitais II - Sistemas, Aplicações e Serviços (independente) e Redes Digitais III - Segurança, Multimédia e Gestão (dependente), Teoria do Sinal (independente) com Modelação e Codificação (dependente) e Modelação e Codificação (independente) com Sistemas de Telecomunicações Guiados (dependente), Teoria dos Circuitos (independente) com Fundamentos de Eletrónica (dependente) e Fundamentos de Eletrónica (independente) com Eletrónica Programada e Processamento Digital de Sinais (dependente). Por fim, o par entre Eletromagnetismo (independente) e Propagação e Radiação de Ondas Eletromagnéticas (dependente). Dos pares adicionados posteriormente, o par na área de Programação apresenta a maior diferença, sendo que o primeiro par da área de Redes apresenta também uma relação relevante entre si. À semelhança do que acontece com o conjunto de UC seguinte (Teoria do Sinal, Modelação e Codificação, Sistemas de Telecomunicações Guiados), o primeiro par do conjunto de Redes (Redes I – Fundamentos, Redes II - Sistemas, Aplicações e Serviços, Redes III - Segurança, Multimédia e Gestão) apresenta uma maior relação de dependência do que o segundo par. Os restantes conjuntos apresentam ainda uma relação interessante, com quase 1 valor de diferença entre as suas médias de sucesso e reprovação.

3.3.2.4. Engenharia de Telecomunicações e Informática Pós-Laboral

No Anexo D - Tabela 19 é possível avaliar os 15 pares de UC relevantes da licenciatura de Engenharia de Telecomunicações e Informática Pós-Laboral. Neste caso, a área de Matemática tem grande influência e é influenciada, sendo possível destacar a influência da UC Álgebra Linear, Geometria Analítica e Análise Vetorial. O par com maior diferença de valores e , conseqüentemente, maior dependência, inclui as disciplinas Análise Matemática II (dependente) e Teoria dos Circuitos (independente). Os pares da área de Programação têm uma diferença bastante significativa, ao contrário do que se sucede no horário diurno do mesmo curso. Para além deste par, existe outro apenas composto por UC de uma área, neste caso Matemática - Análise Matemática II (dependente) e Álgebra Linear, Geometria Analítica e Análise Vetorial (independente).

À semelhança do curso em horário diurno, foram adicionados os pares identificados anteriormente como tendo uma possível relação de dependência. O par de Programação tem uma diferença superior a 2 valores, o que mostra uma grande dependência entre as UC. Entre o conjunto da área de Redes, a diferença ultrapassa 1 valor em ambos os pares, havendo também uma relação significativa entre as disciplinas. Os restantes conjuntos apresentam dependências interessantes, destacando o par entre Teoria dos Circuitos (independente) e Fundamentos de Eletrónica (dependente), exceto o par de Física e Eletromagnetismo, que tem uma diferença entre as médias de reprovação e sucesso de apenas 0,16 valores.

3.3.2.5. Informática e Gestão de Empresas

O Anexo D - Tabela 20 indica os 24 pares com maior relevância do curso de Informática e Gestão de Empresas. É possível concluir que a UC Fundamentos de Arquitetura de Computadores é bastante dependente do desempenho dos alunos em disciplinas lecionadas anteriormente. A influência da área de Arquitetura de Computadores é possível observar, não na UC referida anteriormente, mas nas áreas de Bases de Dados, Modelação e Inteligência Artificial. Existem, ainda, 2 pares dentro da área de Contabilidade e Finanças – as notas dos alunos na UC de Contabilidade Financeira II influencia as notas dos mesmos alunos nas UC de Gestão Financeira I e Gestão Financeira II. Por outro lado, a UC Inteligência Artificial tem melhor ou pior desempenho de acordo com o desempenho dos estudantes nas UC Fundamentos de Arquitetura de Computadores e na UC Introdução à Programação. O par que reúne Programação Concorrente e Distribuída (dependente) e Introdução à Programação (independente) continua com a mesma diferença de valores que no curso anterior, enquanto o último par da tabela aumentou a medida da sua influência.

No final da tabela, foram ainda adicionados pares indicados pelo professor coordenador da licenciatura, sabendo que estes não necessitam de garantir uma diferença de pelo menos 2 valores entre as médias de reprovação e de sucesso, nem precisam de seguir a ordem temporal definida tendo em conta os dados disponíveis. Estes conjuntos de UC incluem, na ordem indicada, a análise entre Introdução à Programação (independente) e Programação Orientada para Objetos (dependente), Redes Digitais I – Fundamentos (independente) e Redes Digitais II - Sistemas, Aplicações e Serviços (dependente), seguida de Redes Digitais II - Sistemas, Aplicações e Serviços (independente) e Redes Digitais III - Segurança, Multimédia e Gestão (dependente), Fundamentos de Arquitetura de Computadores (independente) com Programação Concorrente e Distribuída

(dependente), com Redes Digitais I – Fundamentos (dependente), e com Sistemas Operativos (dependente), Introdução à Programação (independente) com Sistemas Operativos (dependente) e com Engenharia de Software I (dependente), Contabilidade Financeira I (independente) com Fundamentos de Contabilidade de Gestão (dependente) e com Gestão de Operações (dependente), e Fundamentos de Bases de Dados (independente) com Inteligência Artificial. Dos pares adicionados, o par entre Contabilidade Financeira I e Fundamentos de Contabilidade de Gestão apresenta a maior diferença (2,29 valores), sendo que, exceto os pares entre Introdução à Programação e Engenharia de Software I e Contabilidade Financeira I e Gestão de Operações, todos apresentam uma diferença superior a 1 valor. Os pares com Sistemas Operativos como UC dependente têm uma diferença superior a 1 valor e meio.

3.3.2.6. Informática e Gestão de Empresas Pós-Laboral

O Anexo D - Tabela 21 representa os 12 pares significativos do curso de horário pós-laboral de Informática e Gestão de Empresas. Apenas dois pares apresentam uma diferença superior a 2 valores, sendo os restantes adicionados como indicado anteriormente e pela sugestão do professor coordenador da licenciatura. No caso deste curso, um dos pares sugeridos já tinha sido identificado pela diferença superior a 2 valores, e a UC de Contabilidade Financeira I não tinha dados suficientes para fazer parte da análise, pelo que não foram adicionados os pares que continham esta UC.

Os dois pares de Programação têm uma diferença interessante, sendo que a diferença do primeiro é superior a 1 valor. Os pares da área de Redes apresentam diferenças menores que 1 valor, bem como os pares em que a UC Fundamentos de Arquitetura de Computadores é a UC independente. Os pares com Introdução à Programação como UC independente têm diferenças relativamente altas, bem como o par entre Fundamentos de Bases de Dados (independente) e Inteligência Artificial (dependente).

3.3.3. Questão 3

A Questão 3 debruça-se sobre o comportamento das notas dos alunos dentro de cada área, de forma a perceber se as notas dos alunos são consistentes ou variáveis, analisando as medianas, as notas da maioria dos alunos, as notas máximas e mínimas, e os *outliers* que possam existir.

Com o intuito de avaliar os 6 cursos, foram elaborados gráficos *boxplot* para as áreas lecionadas em cada licenciatura. Como são raras as disciplinas em que as notas dos alunos reprovados estão representadas por valores numéricos (normalmente, estão representadas com “RE” ou “F”), foi definido que o limite mínimo do eixo com os valores da “Nota” correspondesse à nota numérica mais baixa. Na prática, existem *boxplots* que avaliam as notas de 0 a 20, outros de, por exemplo, 2 a 20, e outros que têm apenas as notas de 10 a 20. Neste sentido, é importante ressaltar que esta análise é feita tendo em conta as inconsistências que existem na atribuição das notas em valores numéricos e não numéricos.

Foram analisadas as áreas científicas com maior e menor média de cada curso, tendo em conta os resultados obtidos na Questão 1.

3.3.3.1. Engenharia Informática

O Anexo E - Figura 43 reúne os diferentes *boxplots* que resultam da análise de cada uma das 12 áreas incluídas no curso de Engenharia Informática. Começando por analisar as áreas com as médias mais altas, Multimédia e Informática Aplicada, ambas apresentam uma mediana de 15 valores. Têm características bastante semelhantes, mas a área de Multimédia chega a ter alunos com notas 0 que, apesar de representarem *outliers*, são menores que os 2 valores mínimos que têm os alunos de Informática Aplicada. Quanto às áreas de Contabilidade e Finanças e Redes, as duas áreas com pior desempenho no que refere às médias, a mediana de Redes é cerca de 10,5, um valor e meio abaixo da mediana de Contabilidade e Finanças. Esta última, tem a maioria das suas notas a alternar entre 10 e 14,5 valores, enquanto as notas de Redes vão de 10 a 13 valores. Redes apresenta *outliers*, tanto nos valores mais altos, como nos mais baixos, ao contrário da área comparativa, que apresenta uma maior variabilidade apenas nos valores mais baixos.

3.3.3.2. Engenharia Informática Pós-Laboral

No Anexo E - Figura 44, estão presentes os *boxplots* das 12 áreas que compõe a licenciatura de Engenharia Informática Pós-Laboral. Informática Aplicada e Multimédia são as áreas com melhor desempenho por parte dos alunos. Informática Aplicada tem uma mediana de cerca de 15,5 valores, meio valor abaixo da mediana de Multimédia. Ambas as áreas têm a maioria das notas entre 13 e 17 valores, e ambas apresentam *outliers* nos valores mais baixos. No entanto, os alunos de Informática Aplicada atingem valores mais elevados do que a área em comparação. Relativamente às áreas com menores médias, Matemática e Redes, a primeira tem uma mediana de aproximadamente 12 valores, 1 valor acima da mediana de Matemática. A maioria das notas dos alunos de Matemática está entre 10 e 13 valores, ligeiramente menor que os alunos de Redes (10 a 14,5 valores). A área de Redes apresenta *outliers* nos valores menores, mas os seus estudantes atingem notas de 20 valores. No entanto, em Matemática, existem *outliers*, tanto nos valores mais altos, como nos mais baixos, o que indica uma maior variabilidade nas notas obtidas.

3.3.3.3. Engenharia de Telecomunicações e Informática

A licenciatura de Engenharia de Telecomunicações e Informática contempla 13 áreas e o Anexo E - Figura 45 inclui os *boxplots* para cada uma delas. Iniciou-se a análise pelas áreas nas quais os alunos tiveram melhor desempenho, Informática Aplicada e Multimédia. A primeira apresenta uma mediana ligeiramente menor (13 valores) do que a da área de Multimédia (14 valores). Em Informática Aplicada, as notas mais baixas são representadas por *outliers*, o que indica maior variabilidade e alguns alunos com desempenho abaixo da maioria. Na área de Multimédia, as notas são mais concentradas, com menos variação e sem *outliers*, o que indica um desempenho mais consistente entre os estudantes desta UC. Quanto às duas áreas com médias menores, Física e Eletromagnetismo e Matemática. a mediana da primeira é aproximadamente 12, e de Matemática é cerca de 11 valores. Física e Eletromagnetismo apresenta *outliers* tanto nos valores mais baixo, como nos valores mais altos, apesar da maioria dos alunos ter entre 10 e 13 valores. Relativamente a Matemática, as notas da maior parte dos alunos rondam os 10 e os 14,5 valores, apesar de haver alunos com notas elevadas. Existem também *outliers* nas notas inferiores, mas em menor quantidade do que em Física e Eletromagnetismo.

3.3.3.4. Engenharia de Telecomunicações e Informática Pós-Laboral

As 10 áreas lecionadas no curso de Engenharia de Telecomunicações e Informática Pós-Laboral estão representadas no Anexo E - Figura 46, através de *boxplots* que permitem a análise das notas dos seus alunos. Informática Aplicada e Telecomunicações são as áreas desta licenciatura que se destacaram pela positiva. A mediana de Telecomunicações é 12 valores, menos 2 valores que a mediana de Informática Aplicada. A maioria dos alunos de Informática Aplicada está entre os 12 e os 17 valores, enquanto em Telecomunicações a maioria dos estudantes tem entre 10 e 15 valores. Ambas têm alunos com notas de 20 valores, mas a área de Telecomunicações apresenta mais *outliers* nos valores mais baixos e atinge mesmo valores menores, em comparação com a outra área. As áreas com menor desempenho do curso são Bases de Dados, Modelação e Redes. A mediana de ambas as áreas é aproximadamente 11, apesar de a área de Base de Dados, Modelação apresentar uma maior dispersão. Já no gráfico da área Redes, é possível observar *outliers* nos dois extremos, existindo uma elevada quantidade de *outliers* com notas negativas.

3.3.3.5. Informática e Gestão de Empresas

O Anexo E - Figura 47 inclui os *boxplots* das 19 áreas que compõe o curso de Informática e Gestão de Empresas. As áreas identificadas com tendo melhores resultados foram Sistemas de Informação Gerais e Business Intelligence. A mediana das duas áreas é 14 valores. A maioria dos alunos de Business Intelligence tem notas entre 13 e 15 valores, sendo que os mínimos e máximos são 10 e 18 valores, respetivamente. Já a área de Sistemas de Informação Gerais tem a maioria dos seus estudantes entre 12 e 15 valores, e o mínimo e o máximo entre 8 e 19 valores, respetivamente. Ambas apresentam *outliers* nos valores mais baixos, mas apenas Business Intelligence tem *outliers* nos valores superiores a 18. Por outro lado, as áreas de Matemática e Redes têm os piores resultados do curso no que se refere a médias. A mediana de ambas é aproximadamente 11 valores, apesar da área de Matemática ter uma maior dispersão. Os seus valores máximos estão também 3 valores acima dos valores máximos de Redes e, apesar de ambas as áreas terem *outliers* nos valores mais baixos, apenas Redes tem *outliers* nos valores mais elevados de notas.

3.3.3.6. Informática e Gestão de Empresas Pós-Laboral

Os *boxplots* das 18 áreas da licenciatura Informática e Gestão de Empresas Pós-Laboral estão representados no Anexo E - Figura 48. As áreas destacadas positivamente foram Informática Aplicada, Gestão Estratégica e Business Intelligence. É possível observar que a área de Gestão Estratégica não inclui notas de alunos inferiores a 10, pelo que acaba por ter menor informação possível de ser analisada. Enquanto a mediana de Business Intelligence e Gestão Estratégica é cerca de 14 valores, a mediana de Informática Aplicada é superior por 1 valor. Informática Aplicada tem notas superiores às restantes áreas, e apresenta alguns *outliers* nos valores menores. Gestão Estratégica não tem *outliers*, ao contrário de Business Intelligence, que tem *outliers* significativos nos valores de notas mais baixos. Quanto às áreas com piores resultados, Redes e Sistemas Operativos, a mediana de Redes é 11 valores, 1 valor abaixo da mediana de Sistemas Operativos. Enquanto a maioria das notas atribuídas aos alunos de Redes está entre 8 e 12 valores, para os alunos de Sistemas Operativos está entre 11 e 13 valores. Apesar de ambas as áreas apresentarem *outliers*, tanto nos valores superiores como nos inferiores, Redes tem *outliers* mais significativos.

3.3.4. Questão 4

A Questão 4 procura compreender se os “bons alunos” têm um bom desempenho em todas as UC. Dado que a definição de “bons alunos” é bastante abrangente (Diogo et al., 2018), definiu-se a média final do aluno, calculada ao fim dos anos de curso, seria o indicador principal. Para o cálculo desta média, foram utilizadas apenas as notas positivas, ou seja, iguais ou superiores a 10 valores, por não ser possível garantir todos os valores numéricos das reprovações, caso estas existissem. Para dar resposta à questão, foram analisados os dados estatísticos dos 10 alunos com maior média de cada curso, e com registo de inscrições em 3 ou mais anos de curso (4 ou mais no caso dos cursos de Informática e Gestão de Empresas), de forma a analisar apenas alunos que completaram a licenciatura. Foram, ainda, comparados o desvio padrão, a mediana, e as notas máximas e mínimas. A duração do curso não foi considerada por se poder tratar também de alunos a realizar a licenciatura Pré-Bolonha, que comporta 5 anos de curso, ao contrário dos 3 anos da licenciatura Pós-Bolonha. Dado que os alunos representados na base de dados recolhida têm um ID associado, utilizou-se um número meramente indicativo para facilitar a leitura e para proteger as informações pessoais dos alunos.

Com o objetivo de analisar os resultados de forma mais coerente e contextualizada, foram observadas as médias do último colocado em cada curso, disponibilizadas pela Direção-Geral do Ensino Superior⁴, ao longo dos últimos 7 anos letivos presentes nos dados, de forma a incluir o último do horário pós-laboral da licenciatura de Engenharia de Telecomunicações e Informática, 2013/2014.

3.3.4.1. Engenharia Informática

No curso de Engenharia Informática, foram recolhidos os dados estatísticos dos 10 alunos com maior média, e que estiveram no ISCTE-IUL durante 3 ou mais anos letivos. No Anexo F - Tabela 22, é possível observar que a média mais alta é de 19,167 valores e a mais baixa é de 17,4 valores, sendo que o desvio padrão não ultrapassa os 0,702. As medianas estão entre os 17,3 e os 19, e a nota mais baixa que um destes alunos obteve foi 16,9. No geral, são alunos com um excelente desempenho, e com resultados bastante homogéneos.

No Anexo F - Tabela 23, é possível observar a média do último colocado no horário diurno de Engenharia Informática de 2013 a 2019. É possível concluir que houve um aumento significativo entre 2014/2015 e 2015/2016, mas que, a partir dessa data, as médias se mantiveram constantes.

3.3.4.2. Engenharia Informática Pós-Laboral

O Anexo F - Tabela 24 indica os dados estatísticos dos 10 alunos com maior média do curso de Engenharia Informática Pós-Laboral. As médias estão entre os 15,762 e os 17,830 valores, o que indica notas mais baixas do que o curso em horário diurno. O desvio padrão está entre 0,557 e 1,848, demonstrando que os alunos têm uma maior divergência de notas. A nota mais baixa é 14 e a mais alta 19, enquanto a mediana alterna entre 15,286 e 18,167. Apesar de serem resultados bastante satisfatórios, acabam por ser menos homogéneos e mais baixos do que o curso anterior.

⁴<https://www.dges.gov.pt/pt>

O Anexo F - Tabela 25 apresenta a média do último colocado no curso de Engenharia Informática em horário pós-laboral dos últimos 7 anos disponíveis nos dados analisados. Para além de serem valores menos consistentes que o horário diurno, são também valores mais baixos, com uma diferença de cerca de 2 valores.

3.3.4.3. Engenharia de Telecomunicações e Informática

Os dados estatísticos dos 10 alunos com maior média do curso de Engenharia de Telecomunicações e Informática estão representados no Anexo F - Tabela 26. Comparativamente com a licenciatura de Engenharia Informática, as médias são ligeiramente inferiores. As médias variam entre 16,267 e 17,978 valores, pelo que se pode concluir que são bastante constantes entre si. O desvio padrão vai desde 0,051 a 1,418, havendo uma maior variabilidade do que no curso diurno anterior. As medianas são também menores, variando entre 16,8 e 18,217. A nota mais baixa é 15,2 valores e a mais alta é 18,333.

No Anexo F - Tabela 27 está a média do último colocado no horário diurno de Engenharia de Telecomunicações e Informática de 2013 a 2019, onde é possível observar um aumento constante até ao ano letivo de 2017/2018, com uma diminuição nos anos seguintes. Os valores são menores que os do horário diurno do curso anterior, e semelhantes ao horário pós-laboral.

3.3.4.4. Engenharia de Telecomunicações e Informática Pós-Laboral

No Anexo F - Tabela 28 estão representados os dados dos 10 alunos com maiores médias. É possível observar que as médias são menores que o curso com horário diurno, e também que o curso pós-laboral de Engenharia Informática. O desvio padrão atinge também valores superiores aos cursos analisados anteriormente, atingindo o valor de 3,745. O aluno com este desvio padrão apresenta uma grande variabilidade, com nota mínima de 11 valores, nota máxima de 18, e média de 14,806. A nota mais baixa é de 11 valores, e a mais alta é de 19.

Ao contrário dos restantes cursos, o horário de Engenharia de Telecomunicações e Informática não existe atualmente no ISCTE – IUL. No entanto, no último ano letivo do qual foi possível obter dados sobre a média do último colocado no curso, em 2013/2014, a média era de 106,0 valores. Este valor é diferente dos valores obtidos nos horários diurnos dos restantes cursos, mas semelhante aos valores obtidos nos horários pós-laborais.

3.3.4.5. Informática e Gestão de Empresas

O Anexo F - Tabela 29 representa os dados estatísticos dos 10 alunos com maior média da licenciatura de Informática e Gestão de Empresas. Neste curso, em ambos os horários, foram selecionados apenas alunos com inscrições em 4 ou mais anos letivos, pois são necessários 4 anos para concluir a licenciatura pós-Bolonha. As médias variam entre 16,125 e 17,575 valores, enquanto o desvio padrão está entre 0,096 e 1,382, ligeiramente abaixo dos valores do curso anterior em horário diurno. A nota mais baixa é 14,8 valores e a mais alta 18 valores. As medianas estão entre 15,95 e 17,95. Apesar das médias serem semelhantes às do curso de Engenharia de Telecomunicações e Informática, as medianas são menores.

O Anexo F - Tabela 30 apresenta a média do último colocado no curso de Informática e Gestão de Empresas dos últimos 7 anos disponíveis nos dados analisados. É possível observar um aumento entre 2014 e 2015, sendo que os anos de 2017/2018 e 2019/2020 se destacam positivamente. Apesar das médias serem superiores ao horário diurno de Engenharia de Telecomunicações e Informática, são menores que o horário diurno de Engenharia Informática.

3.3.4.6. Informática e Gestão de Empresas Pós-Laboral

Os dados estatísticos dos 10 alunos com melhores médias do curso de Informática e Gestão de Empresas Pós-Laboral estão representados no Anexo F - Tabela 31. As médias dos alunos variam mais do que as do curso anterior, indo desde 14,395 até 18,867, ou seja, atingindo valores maiores e menores do que o horário diurno. O desvio padrão é semelhante ao anterior, entre 0,163 e 1,372, havendo menor variabilidade do que os restantes cursos em horário pós-laboral. A nota mais baixa é 12,4 valores, e a mais alta 19. As medianas variam entre 14,627 e 18,9 valores.

No Anexo F - Tabela 32, é possível observar a média do último colocado no horário diurno de Engenharia Informática de 2013 a 2019. No ano de 2015, deu-se um aumento significativo, de cerca de 3 valores e, a partir desse ano, os valores mantiveram-se relativamente constantes, com ênfase positivo em 2017/2018 e 2019/2020. Os valores são menores que o horário diurno, e semelhantes aos valores do horário pós-laboral de Engenharia Informática.

CAPÍTULO 4

Resultados e Discussão

Após a análise da investigação científica e a descrição da metodologia utilizada para dar resposta às questões levantadas, é necessário discutir os resultados obtidos à luz da literatura científica, bem como identificar medidas que levem à concretização do objetivo final do presente estudo – transmitir informação de suporte para que seja possível uma melhor adequação do plano curricular e dos mecanismos de avaliação das UC, por parte do corpo docente do ISCTE-IUL.

4.1. Questão 1

A Questão 1 foi definida com o objetivo de comparar as unidades curriculares dos 6 cursos, e as áreas às quais estas pertencem, à semelhança do que foi realizado no artigo Nº 3 (Wells et al., 2021). Foram realizadas diferentes análises, começando pela aplicação do método do cotovelo no Scree Plot, e na visualização do gráfico de dispersão – também utilizado no estudo Nº1 (Yungang et al., 2017), de forma a identificar grupos dentro de cada curso. Após a análise das UC e áreas dentro de cada *cluster*, compararam-se as médias das diferentes áreas com o intuito de encontrar as áreas com melhor e pior desempenho. Para além da média, as percentagens de alunos reprovados e de alunos não avaliados foram utilizadas para identificar as áreas que apresentam maior dificuldade para os alunos de cada curso.

No curso diurno de Engenharia Informática, as áreas com maior média são Multimédia e Informática Aplicada, enquanto Redes e Contabilidade e Finanças têm as piores médias, com Redes e Arquitetura de Computadores sendo as mais difíceis. No curso pós-laboral, as melhores médias são Multimédia e Informática Aplicada, enquanto Redes e Programação tiveram piores médias, com Redes e Telecomunicações identificadas como as áreas mais difíceis.

Em Engenharia de Telecomunicações e Informática, destacou-se positivamente Multimédia, Informática Aplicada e Telecomunicações, com dificuldades em Matemática, Física e Eletromagnetismo, Bases de Dados, Modelação e Sistemas Operativos. No curso com horário pós-laboral, Telecomunicações e Informática Aplicada foram as áreas com melhores médias, enquanto Bases de Dados, Modelação e Redes tiveram piores médias, sendo Matemática a área mais difícil.

Informática e Gestão de Empresas apresentou melhores médias em Business Intelligence e Sistemas de Informação Gerais, com piores médias em Matemática e Redes. As áreas mais difíceis foram Redes e Matemática. No curso pós-laboral, Informática Aplicada, Business Intelligence e Gestão Estratégica tiveram as melhores médias, enquanto Redes e Sistemas Operativos tiveram as piores, sendo Redes a área mais difícil.

Em 5 cursos, excluindo ETI diurno, Redes mostrou-se a área científica que apresenta maiores dificuldades para os alunos. Esta dificuldade parece não depender do que se faz em cada curso, mas sim das características da área. As UC que compõem a área de Redes

parecem apresentar matérias com uma complexidade superior às restantes. No caso do horário diurno da licenciatura de Engenharia de Telecomunicações e Informática, apesar de a área científica ser bastante complexa, é uma das áreas de maior interesse dos alunos deste curso, pelo que podem acabar por ter mais sucesso do que os restantes. Redes não é uma área *core* de IGE, mas acaba por ser lecionada com o mesmo nível de complexidade que as restantes licenciaturas analisadas, pelo que seria de esperar que apresentasse uma grande dificuldade para os alunos deste curso. Pode ser interessante realizar uma análise mais aprofundada nesta área.

No curso de Engenharia Informática, o horário pós-laboral apresenta, na maioria, áreas científicas com médias superiores. No curso de Engenharia de Telecomunicações e Informática, não existe um horário que se destaque, e em Informática e Gestão de Empresas, o curso diurno tem médias superiores na maioria das áreas. Visto que não é possível identificar uma tendência de superioridade das médias num único tipo de horário, esta distinção pode-se dever a fatores mais subjetivos como a motivação dos alunos ou a disponibilidade dos mesmos.

Em ambos os horários da licenciatura de Informática e Gestão de Empresas, foi possível concluir acerca do melhor desempenho em áreas de Gestão, e pior em Informática. Esta diferença deve-se a uma maior complexidade das temáticas de Informática comparativamente às de Gestão. Pode ser interessante investigar mais a fundo esta questão, podendo até perceber se os alunos que escolhem estes cursos têm uma aptidão natural maior para áreas de Gestão.

A área de Programação, de interesse especial, teve as melhores médias em Engenharia Informática, com uma tendência de melhores resultados nos cursos pós-laborais. Na Questão 4, observou-se que os alunos de Engenharia Informática têm notas mais elevadas, pelo que é natural que a área de Programação não seja exceção. Nos horários pós-laborais, é possível os métodos de ensino sejam mais adequados ao perfil destes alunos.

4.2. Questão 2

A Questão investiga a possível dependência entre UC, baseando-se no desempenho dos alunos, ou seja, procura responder à questão: se um aluno reprovar numa UC (independente) influencia negativamente o seu desempenho numa UC (dependente) subsequente? Para dar resposta a esta pergunta, foram analisadas as notas dos alunos em pares de disciplinas, garantindo que estes seguissem a ordem temporal necessária (UC dependente lecionada após UC independente). Foram calculadas médias de reprovação e médias de sucesso, e foram selecionados apenas os pares com diferenças entre 2 e 5 valores entre as médias. Por recomendação dos professores orientadores, foram ainda adicionados 2 pares da área de Programação e, posteriormente, outros pares indicados pelos professores coordenadores de cada licenciatura. Apesar de não se ter aplicado uma correlação, a procura pela relação entre disciplinas remete aos artigos identificados com N°1 e N°2 (Nash et al., 2021; Robbiano et al., 2020).

No curso diurno de Engenharia Informática, identificaram-se 30 pares e a maior dependência foi observada entre Inteligência Artificial (dependente) e Redes Digitais I – Fundamentos (independente). As UC Inteligência Artificial, Mecânica e Eletricidade e Sistemas Operativos apresentam grandes dependências. Os pares entre as UC Mecânica

e Eletricidade (dependente) e Álgebra Linear, Geometria Analítica e Análise Vetorial (independente) e Análise Matemática I (dependente) fazem bastante sentido pelo seguimento de conceitos lecionados. No entanto, o par entre Mecânica e Eletricidade (dependente) e Fundamentos de Bases de Dados (independente) não faz sentido a nível de temáticas, mas pode dever-se ao perfil dos alunos, que têm maior aptidão para estas áreas. Redes, Arquitetura de Computadores e Bases de Dados são influentes. A influência da área de Redes não se deve tanto aos conceitos, mas sim ao tempo que as suas UC exigem dos alunos, que acabam por ter menos disponibilidade para dedicar a outras disciplinas. No caso de Arquitetura de Computadores, o par entre Fundamentos de Arquitetura de Computadores (dependente) e Sistemas Operativos (independente) faz sentido tendo em conta as temáticas das disciplinas. A relação entre Fundamentos de Bases de Dados (independente) e Programação em Rede (dependente) também faz sentido, mas não é uma correlação óbvia. No curso com horário pós-laboral, 10 pares foram selecionados. Programação é altamente dependente e influente, com Sistemas Operativos afetando e sendo afetado por ela. Como se trata de uma área bastante abrangente e transversal a várias competências, faz sentido que Programação tenha este peso no desempenho dos alunos. O par com maior dependência é Teoria da Computação (dependente) e Sistemas Operativos (independente), o que não faz muito sentido, podendo ser apenas uma coincidência. Programação Concorrente e Distribuída (dependente) e Introdução à Programação (independente) também mostram grande dependência, além de outro par na área de Programação: Teoria da Computação (dependente) e Algoritmos e Estruturas de Dados (independente), sendo que ambas as UC são lecionadas pelo mesmo professor, o que pode levar a que os seus métodos de ensino acabem por favorecer naturalmente os alunos que se relacionam mais com os mesmos, ou mesmo a um enviesamento dos alunos para com as disciplinas. Nestes cursos, é possível perceber o peso da área de Programação, que se justifica por serem os cursos com mais UC pertencentes à área científica de Programação. Relativamente aos pares adicionados posteriormente, o par com maior relação de dependência é entre Introdução à Programação (independente) e Programação Orientada para Objetos (dependente), sendo que Algoritmos e Estruturas de Dados (independente) e Programação Orientada para Objetos (dependente) manteve uma diferença relevante em ambos os horários. Os restantes pares não apresentaram uma dependência tão significativa, excetuando o par Programação Concorrente e Distribuída (independente) e Inteligência Artificial (dependente) no horário pós-laboral. A relação entre estes pares faz sentido tendo em conta as temáticas lecionadas, e demonstra a importância de consolidar os conhecimentos em UC como Introdução à Programação, que determinam o sucesso dos estudantes em disciplinas posteriores.

Em Engenharia de Telecomunicações e Informática, foram identificados 25 pares relevantes, e Física e Eletromagnetismo e Arquitetura de Computadores são as áreas mais influentes. Inteligência Artificial e Ondas e Ótica dependem fortemente de outras UC. O par com maior influência é entre Inteligência Artificial (dependente) e Fundamentos de Arquitetura de Computadores (independente). Neste curso, as áreas de Arquitetura de Computadores e Física e Eletromagnetismo são compostas por 7 e 6 UC, respetivamente, o que demonstra a sua importância. Assim, é natural que o seu desempenho seja influente no desempenho de outras UC, não só porque os alunos terão de realizar outras disciplinas enquanto repetem as que reprovaram, mas também porque são lecionados temas cruciais

ao aproveitamento do curso. No curso pós-laboral, identificaram-se 15 pares relevantes, sendo que Matemática é influenciada e depende de várias UC, destacando-se Álgebra Linear, Geometria Analítica e Análise Vetorial. O par com maior relação é Análise Matemática II (dependente) e Teoria dos Circuitos (independente). Os pares adicionados na área de Programação têm diferenças significativas, assim como outro par em Matemática: Análise Matemática II (dependente) e Álgebra Linear, Geometria Analítica e Análise Vetorial (independente). Matemática é uma área transversal aos diferentes cursos, mas no horário pós-laboral de Engenharia de Telecomunicações Informática, os alunos que têm de repetir a UC Álgebra Linear, Geometria Analítica e Análise Vetorial, têm maior dificuldade em ter sucesso na disciplina Análise Matemática II. Esta relação entre UC pode derivar de uma continuidade mais acentuada de temáticas, o que faz com que os alunos que não apreenderam os conteúdos da primeira tenham maior dificuldade na seguinte, apesar de, neste caso, Análise Matemática I ter uma maior relação a nível de conteúdos com Análise Matemática II. Quanto aos pares adicionados posteriormente, é possível reconhecer relações relevantes entre os mesmos, destacando o par entre Introdução à Programação (independente) e Programação Orientada para Objetos (dependente). Os pares com menor relação são Eletromagnetismo (independente) e Propagação e Radiação de Ondas Eletromagnéticas (dependente) no curso pós-laboral, e Redes Digitais II - Sistemas, Aplicações e Serviços (independente) e Redes Digitais III - Segurança, Multimédia e Gestão (dependente) no horário diurno. No geral, as relações entre estes pares são esperadas devido à correlação entre conteúdos lecionados e ao tipo de raciocínio necessário.

Informática e Gestão de Empresas revelou 24 pares. Fundamentos de Arquitetura de Computadores é uma UC dependente, com a área de Arquitetura de Computadores a influenciar Bases de Dados, Modelação e Inteligência Artificial, sendo que estes pares não fazem sentido dentro do contexto do curso. O par com maior dependência é Fundamentos de Arquitetura de Computadores (dependente) e Fundamentos de Bases de Dados (independente), o que não faz sentido tendo em conta as temáticas lecionadas. Na área de Contabilidade e Finanças, destacam-se Contabilidade Financeira II (independente) e Gestão Financeira I e II (dependentes), o que se comprova pelos temas das UC. Esta área é composta pelo maior número de UC, pelo que tem uma grande influência no aproveitamento dos alunos. Através destas relações, é possível observar que os alunos que reprovam a Contabilidade Financeira II têm maiores dificuldades em ter sucesso nas UC de Gestão Financeira I e II, provavelmente pela continuidade das temáticas abordadas. Nos pares adicionados posteriormente, é possível concluir que a UC Fundamentos de Arquitetura de Computadores tem influência em Programação Concorrente e Distribuída, Redes I – Fundamentos e Sistemas Operativos, pela relação dos temas entre as UC. Foi também possível concluir acerca da importância da UC Contabilidade Financeira I para o aproveitamento na UC Fundamentos de Contabilidade de Gestão, bem como da relação entre Fundamentos de Bases de Dados e Inteligência Artificial, devendo-se ao facto de as UC terem temáticas em comum. No curso pós-laboral, identificaram-se 12 pares relevantes. Fundamentos de Arquitetura de Computadores é influente em Programação Concorrente e Distribuída, Redes Digitais I – Fundamentos e Sistemas Operativos, devido à correspondência entre os temas lecionados. Introdução à Programação aparenta ser uma UC essencial no bom desempenho dos alunos em Programação Orientada para Objetos, Programação Concorrente e Distribuída,

Sistemas Operativos e Engenharia de Software I. É possível concluir que o sucesso nesta UC introdutória pode fazer a diferença no futuro dos estudantes. Por fim, tal como indicado pelo professor coordenador do curso, é possível que exista uma dependência entre a área de Base de Dados, Modelação e Inteligência Artificial, pois estas estão interligadas.

As relações entre UC identificadas nesta questão podem dever-se a diferentes fatores. O primeiro é a sorte, podendo haver uma diferença nas médias de reprovação e sucesso apenas por um acaso e não por existir uma razão que a justifique. O segundo é a interligação de conceitos lecionados, o que faz com que os alunos que não tenham consolidados os temas da UC independente tenham dificuldade em ter sucesso na UC dependente. O terceiro fator é a gestão do tempo investido em várias UC, quando é necessário repetir uma à qual que se reprovou. O quarto, e talvez último, é o perfil do aluno, que se pode adequar mais ou menos à matéria, ao professor, e/ou ao método de ensino, o que faz com que tenha maior aptidão para um par de UC específico que, à primeira vista, pouco têm em comum.

4.3. Questão 3

A Questão 3 procura compreender melhor o desempenho das áreas que compõe cada um dos cursos em estudo. Para tal, foram analisadas as notas obtidas em cada uma das áreas, através de *boxplots*. Alguns professores indicam as reprovações e as não avaliações na plataforma Fénix através de caracteres não numéricos (“RE” e “F”, e “NA”, “D” e “ ”, respetivamente), pelo que nem todas as áreas podem ser analisadas de 0 a 20 valores, mas sim de 10 a 20 valores. Neste sentido, as áreas têm de ser estudadas tendo em conta esta inconsistência, sendo que *outliers* nos valores abaixo de 10 podem ser explicadas pela mesma. Para cada curso, a análise foi mais focada nas áreas com melhor e pior média, identificadas na Questão 1.

No curso diurno de Engenharia Informática, as áreas com maior média são Multimédia e Informática Aplicada (mediana de 15 valores). Informática Aplicada apresenta melhores resultados que Multimédia, que tem *outliers* nos valores inferiores. As áreas com menor média são Contabilidade e Finanças e Redes. Redes tem uma mediana de 10,5, inferior à de Contabilidade e Finanças (12 valores), e apresenta *outliers* em ambos os extremos. No curso pós-laboral, as áreas com maior média são Informática Aplicada (mediana de 15,5 valores) e Multimédia (mediana de 15 valores), com notas geralmente entre 13 e 17 valores. Informática Aplicada alcança valores superiores. As áreas com menor média são Matemática (mediana de 11 valores) e Redes (mediana de 12 valores). Redes possui uma variação menor, enquanto Matemática tem *outliers* em ambos extremos.

Em Engenharia de Telecomunicações e Informática, as áreas com maior média são Informática Aplicada (mediana de 13 valores) e Multimédia (mediana de 14 valores), sem *outliers*. As áreas com menor média são Física e Eletromagnetismo (mediana de 12 valores) e Matemática (mediana de 11 valores), com Matemática apresentando maior variabilidade nos resultados. No horário pós-laboral, as áreas com maior média são Informática Aplicada (mediana de 14 valores) e Telecomunicações (mediana de 12 valores). Informática Aplicada tem melhores resultados e menos *outliers*. As áreas com

menor média são Bases de Dados, Modelação e Redes (mediana de 11 valores), com Bases de Dados e Modelação mostrando maior dispersão.

No curso diurno de Informática e Gestão de Empresa, as áreas com maior média são Sistemas de Informação Gerais e Business Intelligence (mediana de 14 valores). Business Intelligence possui *outliers* altos e baixos. As áreas com menor média são Matemática e Redes (mediana de 11 valores). Redes apresenta *outliers*, tanto nos valores mais altos, como nos mais baixos. No horário pós-laboral, as áreas com maior média são Informática Aplicada, Gestão Estratégica e Business Intelligence (mediana de 14 valores). Informática Aplicada tem notas superiores e *outliers* nos valores mais baixos. As áreas com menor média são Redes (mediana de 11 valores) e Sistemas Operativos (mediana de 12 valores). Redes apresenta *outliers* significativos em ambos os extremos.

Os resultados obtidos estão de acordo com o esperado, já que as áreas com melhor desempenho apresentam maior consistência nas notas, e as áreas com menor desempenho apresentam maior variabilidade, havendo mais *outliers* e menos alunos com notas elevadas. Os *outliers* nos valores mais baixos devem-se, maioritariamente, ao facto de que existem menos valores numéricos representativos de reprovações, ou seja, menores que 10 valores.

4.4. Questão 4

A Questão 4 pretende analisar se os “bons alunos” apresentam bom desempenho em todas as UC. A análise considerou os 10 alunos com médias finais mais altas de cada curso, calculadas apenas com notas positivas (iguais ou superiores a 10), e que tivessem notas relativas a 4 ou mais anos letivos para ambos os horários do curso de Informática e Gestão de Empresas, e 3 ou mais anos letivos para os restantes.

Em resposta a esta questão, foi possível concluir que os “bons alunos” apresentam, na sua generalidade, um desempenho elevado em todas as UC, sendo possível observar variações entre os diferentes cursos e horários. As licenciaturas em horário diurno têm médias mais elevadas e as notas mais homogêneas, enquanto os alunos dos cursos pós-laborais apresentam uma maior dispersão das notas obtidas. É, ainda, possível analisar que os horários diurnos apresentam médias superiores no que se refere ao último colocado em cada curso. Assim, a diferença identificada no desempenho dos alunos pode não se dever tanto ao método de ensino, mas sim aos próprios estudantes, que chegam ao ensino superior com melhor ou pior média, e que ingressam no horário que mais se adequa, não só à sua disponibilidade, mas também às suas notas.

Adicionalmente, os alunos com maior média do curso de Engenharia Informática têm médias superiores comparativamente aos alunos com médias superiores dos restantes cursos. Esta superioridade do curso de EI também é visível ao observar a média do último colocado em cada curso, estando menos dependente dos métodos de avaliação ou de ensino, mas sim do perfil dos alunos que ingressam neste curso.

CAPÍTULO 5

Conclusões

Após a análise e discussão dos resultados obtidos, bem como do contributos dos professores coordenadores das licenciaturas aos quais foram realizadas entrevistas, foi possível reunir as conclusões principais do estudo.

Relativamente à Questão 1, é possível concluir acerca das áreas com melhor e pior desempenho por parte dos alunos de Informática, bem como as áreas que implicam mais dificuldades para os estudantes. As áreas de Informática Aplicada, Multimédia e Telecomunicações tendem a ter melhores resultados nos estudantes das licenciaturas de Engenharia Informática e Engenharia de Telecomunicações e Informática. A licenciatura de Informática e Gestão de Empresas tem a particularidade de incluir unidades curriculares da área de Gestão, pelo que as áreas com mais sucesso são Business Intelligence, Sistemas de Informação Gerais, Gestão Estratégica e, novamente, Informática Aplicada. Importa realçar que, por vezes, a área Informática aplicada corresponde apenas a uma unidade curricular. Com menos sucesso, estão as áreas de Redes, Bases de Dados, Modelação, Programação, Matemática, e Sistemas Operativos. Em casos individuais, identificou-se ainda resultados mais baixos em Contabilidade e Finanças, Programação e Física e Eletromagnetismo. Neste sentido, as áreas que se destacam pela maior dificuldade, não só pelas notas, mas pelas reprovações e pelas não avaliações, são Redes, Matemática, Arquitetura de Computadores e, apenas numa situação, Telecomunicações. Assim, as UC que compõe a área científica Redes apresentam dificuldades em todos os cursos e horários, podendo dever-se a temáticas mais complexas. A licenciatura de Informática e Gestão de Empresas, em ambos os horários, mostrou que os alunos têm mais dificuldade nas áreas de Informática, e melhor desempenho nas áreas de Gestão, o que pode dever-se a uma maior complexidade dos temas lecionados em Informática. A área de Programação apresenta as médias mais altas no curso de Engenharia Informática, e tem tendência a ter melhores resultados nos horários pós-laborais, possivelmente devido ao perfil destes alunos. Adicionalmente, verificou-se que no curso de Engenharia Informática, o horário pós-laboral apresenta áreas com maior média em comparação com o horário diurno, ao contrário do curso de Informática e Gestão de Empresas. No curso de Engenharia de Telecomunicações e Informática não existe um horário em destaque. Dado que não é possível observar uma tendência, este acontecimento pode dever-se ao perfil dos alunos de cada horário.

Na Questão 2, concluiu-se acerca da relação entre pares de UC, sendo possível verificar as áreas mais influentes e que mais dependem do desempenho de outras. A licenciatura de Engenharia Informática apresenta grande influência das áreas de Redes, Arquitetura de Computadores e Bases de Dados, Modelação. A importância de Redes não se deve apenas aos conceitos complexos, mas também à exigência das suas UC. As áreas que mais dependem do desempenho de outras são Inteligência Artificial, Mecânica e Eletricidade e Sistemas Operativos. A dependência da UC Mecânica e Eletricidade por UC da área de Matemática faz sentido, bem como o par entre Fundamentos de Arquitetura de Computadores e Sistemas Operativos, pela continuidade dos conceitos. No curso pós-laboral, existem maior dependências das áreas de Programação, e maior influência por

parte de Sistemas Operativos e Programação. Existe um par apenas da área de Programação e, como se trata de uma área abrangente e transversal a várias competências, faz sentido que Programação tenha influência e que seja dependente do desempenho de outras UC. Em ambos os horários, o peso da área de Programação justifica-se por serem os cursos com mais UC pertencentes à área científica de Programação. O par com maior dependência não faz muito sentido, podendo ser um acaso. O par entre Teoria da Computação (dependente) e Algoritmos e Estruturas de Dados (independente) pode dever a sua relação ao facto de ser lecionado pelo mesmo professor. Dos pares adicionados por haver uma possível dependência, o que apresenta maior relação é entre Introdução à Programação e Programação Orientada para Objetos, sendo que Algoritmos e Estruturas de Dados e Programação Orientada para Objetos manteve uma diferença relevante em ambos os horários. A relação entre estes pares pode ser explicada pela continuidade dos conceitos lecionados, e reforça a importância de consolidar os conhecimentos de Programação em UC como Introdução à Programação, de forma a garantir um bom desempenho em disciplinas seguintes.

Em Engenharia de Telecomunicações e Informática, há maior dependência por parte de Inteligência Artificial e Ondas e Ótica, e influência por parte de Física e Eletromagnetismo e Arquitetura de Computadores. Neste curso, as áreas de Arquitetura de Computadores e Física e Eletromagnetismo são compostas por várias UC, pelo que é natural que o seu desempenho seja influente no desempenho de outras disciplinas. No horário pós-laboral, a área de Matemática é influenciada e tem influência, havendo um par só da área de Matemática. A UC Teoria dos Circuitos influencia uma UC de Matemática. Matemática é uma área transversal aos diferentes cursos, mas neste caso, os alunos que têm de repetir a UC Álgebra Linear, Geometria Analítica e Análise Vetorial, têm maior dificuldade em ter sucesso na disciplina Análise Matemática II. Esta relação pode ocorrer devido a existir continuidade dos temas lecionados, levando a que os alunos tenham maior dificuldade na UC seguinte quando não são aprovados na anterior. Quanto aos pares adicionados posteriormente, destaca-se o par entre Introdução à Programação e Programação Orientada para Objetos, apesar de a maioria ter uma dependência significativa. Os pares com menor relação são entre Eletromagnetismo e Propagação e Radiação de Ondas Eletromagnéticas no curso pós-laboral, e Redes Digitais II - Sistemas, Aplicações e Serviços e Redes Digitais III - Segurança, Multimédia e Gestão no horário diurno. É natural que estes pares tenham dependências entre si, devido à correlação entre os temas lecionados.

No curso Informática e Gestão de Empresas, as UC Contabilidade Financeira I e II têm influência nas UC Gestão Financeira I e II, Fundamentos de Contabilidade e Gestão e Gestão de Operações. No curso pós-laboral, observa-se a importância da UC Introdução à Programação em Programação Orientada para Objetos, Programação Concorrente e Distribuída, Sistemas Operativos e Engenharia de Software I. Em ambos os horários, é possível verificar a relação que existe entre Fundamentos de Bases de Dados e Inteligência Artificial, e entre Fundamentos de Arquitetura de Computadores e Programação Concorrente e Distribuída, Redes I – Fundamentos e Sistemas Operativos.

Em todos os cursos, foi possível observar a importância da UC Introdução à Programação no bom desempenho dos alunos nas diversas áreas.

A Questão 3 procura analisar o desempenho dos alunos dentro de cada área. Foram analisados a fundo apenas as áreas com melhor e pior desempenho de cada curso. As áreas com melhor desempenho são Informática Aplicada, Multimédia, Gestão Estratégica e Business Intelligence, que apresentam medianas altas, entre 13 e 15,5 valores, e desempenhos consistentes. As áreas com pior desempenho são Redes, Matemática, Física e Eletromagnetismo, Contabilidade e Finanças e Sistemas Operativos, que tendem a ter menores medianas, entre 10,5 e 12 valores, e apresentam uma maior dispersão das notas obtidas. Os resultados obtidos correspondem às expectativas, já que as áreas com melhor desempenho exibem maior consistência nas notas, enquanto as áreas com pior desempenho mostram maior variabilidade, apresentado mais *outliers* e menos alunos com notas altas. Os *outliers* nas notas mais baixas ocorrem principalmente porque existem menos notas representativas de reprovações, ou seja, inferiores a 10 valores.

A Questão 4 tinha como objetivo perceber se os “bons alunos” teriam um bom desempenho em todas as UC. Foi possível concluir que estes alunos têm, geralmente, um desempenho elevado em todas as UC, com algumas variações entre cursos e horários. As licenciaturas em horário diurno apresentam médias mais altas e notas mais homogêneas, enquanto os cursos pós-laborais mostram maior dispersão nas notas. Estas diferenças podem dever-se aos próprios estudantes, já que a média do último colocado no ensino superior analisada em cada curso mostra que as médias de entrada nas licenciaturas em horário diurno são superiores às médias de entrada nas licenciaturas em horário pós-laboral. O mesmo pode ser concluído acerca das médias superiores observadas no curso de Engenharia Informática, que tem também as médias mais altas de entrada no curso.

5.1. Limitações e Trabalhos Futuros

Relativamente à escolha da metodologia, a *Cross-Industry Standard Process for Data Mining* (CRISP-DM) (Shearer, 2000) não foi selecionada por não se tratar de um projeto que envolvesse modelação.

A análise da base de dados do Fénix levou à identificação de várias incongruências no conteúdo e na forma da mesma. Nomeadamente, o modelo relacional do *dataset* apresentava algumas inconsistências, o que dificultou a análise e a recolha dos dados. Alguns exemplos disto são o facto de não haver uma forma direta de relacionar a informação relativa aos cursos e às suas disciplinas, e o facto de existirem diferentes alunos em diferentes tabelas, que não se relacionavam entre si. Foram tomadas decisões para colmatar os problemas encontrados, tais como a análise apenas dos alunos cuja identificação estaria presente na tabela “Candidacy”.

É de notar que a informação relativa aos anos letivos de 2019/2020 e 2018/2019 é limitada. Alguns alunos cujo primeiro ano escolar era o de 2019/2020 não tinham informação relativa às notas, e os alunos cujo primeiro ano era o de 2018/2019 tinham apenas informação relativas a esse ano. O facto de não haver informação relativa a estes anos ou até mais recente impediu que se fizessem análises ao desempenho escolar relacionadas com o Covid-19.

A informação relativa ao ano e semestre em que cada UC se realiza era incoerente, pelo que se assumiram medidas de modo a contornar a situação. Neste sentido, para indicar o

semestre em que cada UC é lecionada, foram utilizados o Fénix e os dados da base de dados que apenas indicassem um único semestre por UC, e nas disciplinas que compreendiam uma continuidade (e.g. Matemática I e Matemática II, ou Matemática e Matemática Avançada), assumiu-se que a primeira UC correspondia ao primeiro semestre, a segunda ao segundo semestre, e a terceira, caso existisse, ao primeiro semestre. Para as UC sem informação relativa ao semestre, assumiu-se que seriam lecionadas no segundo semestre, de forma a não prejudicar a análise da Questão 2.

Quanto ao ano do curso em que cada UC é lecionada, esta informação seria apenas necessária para resposta à Questão 2 e a informação relativa ao ano letivo estava corretamente associada a cada nota. No caso dos pares de disciplinas, foi verificado caso a caso se estes seguiam a ordem temporal correta a nível de ano do curso, tendo em conta o semestre atribuído e o ano letivo associado a cada nota.

Como a maioria das notas disponíveis tinham valores apenas de 10 a 20 caso fossem positivas, e RE ou F para indicar reprovações, foram apenas considerados os valores iguais ou superiores a 10 para calcular médias. Ainda neste sentido, quando são referidos alunos reprovados ou alunos não avaliados, é importante ter em conta que os mesmos alunos podem repetir a mesma UC várias vezes ao longo do seu percurso no Iscte, garantindo que o ano letivo associado a cada nota é sempre diferente. Assim, um aluno pode reprovar e, no ano seguinte, ter uma nota positiva. Após uma nota positiva, os alunos não repetiam as UC. Assim, um aluno “positivo” é um aluno que tem uma nota positiva, mesmo que já tenha tido várias notas negativas, leia-se reprovações e/ou não avaliações.

Na Questão 1, pode ser interessante comparar também o número de aprovações em cada UC.

Relativamente à Questão 2, também se considerou calcular a diferença entre as médias de Sucesso e Reprovação com a percentagem (Reprovação/Sucesso ao invés de Reprovação – Sucesso), mas ao comparar os dois métodos, foi decidido que a diferença entre as duas médias seria um método mais absoluto e intuitivo de analisar.

Quanto à Questão 3, tal como foi referido, nem todos os professores indicam valores numéricos menores que 10 na atribuição das notas aos alunos. Deste modo, na avaliação da existência de valores negativos, ou seja, menores que 10, é importante considerar esta inconsistência. É importante compreender que os valores de alunos reprovados e não avaliados (“RE”, “F”, “D”, “NA” ou sem qualquer valor de nota) não estão a ser considerados nesta análise. Podia também ser interessante comparar a análise realizada nesta questão com o número de reprovações e de não avaliações, bem como analisar cada UC e não apenas as áreas científicas às quais pertencem.

Na Questão 4, podia acrescentar valor à análise a normalização da média de cada um dos 10 alunos identificados com a respetiva média de entrada no curso, de forma a contextualizar e fundamentar as conclusões obtidas nesta questão.

Todas as questões podem ser analisadas de forma mais interessante e valiosa caso haja acesso aos valores numéricos das notas dos alunos que reprovam.

As situações descritas poderiam ter sido evitadas caso a base de dados tivesse sido recolhida com a análise final em mente e o modelo relacional tivesse sido criado de raiz para o presente estudo.

Para trabalhos futuros, seria interessante ter informação relativa a outras características, tais como demográficas e psicológicas, de forma a analisar a importância destas variáveis no desempenho escolar. Poderia também adicionar valor a análise do sucesso escolar tendo em conta o tipo de avaliação de cada disciplina, ou seja, se é feita através de trabalhos individuais ou de grupo, testes, exames, etc. Adicionalmente, um estudo com dados nos quais todas as notas de reprovação fossem numéricas poderia acrescentar à informação relativa a cada UC.

Referências Bibliográficas

- Apoorva, K. A., & Sangeetha, S. (2024). *Contextual Forensic Analysis of Emails Using Machine Learning Algorithms*. 450–456.
<https://doi.org/10.1109/IC2PCT60090.2024.10486397>
- Azuaje, F., Witten, I., & E, F. (2006). Witten IH, Frank E: Data Mining: Practical Machine Learning Tools and Techniques. *Biomedical Engineering Online - BIOMED ENG ONLINE*, 5, 1–2.
- Baker, R. (2010). Data mining for education. *International Encyclopedia of Education*, 7, 112–118.
- da Silva, L. F., & Rossoni, H. A. V. (2023). Implementation and effectiveness of public policies and rural sanitation programs: An approach through systematic literature review. *Revista Brasileira de Geografia Física*, 16(5), 2325–2344.
<https://doi.org/10.26848/RBGF.V16.5.P2325-2344>
- Delavari, N., Phon-Amnuaisuk, S., & Beikzadeh, M. R. (2008). Data mining application in higher learning institutions. *Informatics in Education*, 7(1), 31–54.
<https://doi.org/10.15388/infedu.2008.03>
- Diogo, A. M., E Melo, B. P., & Ferreira, M. (2018). Exams and logics of producing good pupils in the middle classes. *Cadernos de Pesquisa*, 48(169), 748–775.
<https://doi.org/10.1590/198053145138>
- Fraenkel, J., Wallen, N., & Hyun, H. (2011). How to Design and Evaluate Research in Education. Em *Journal of American Optometric Association* (Vol. 60).
- Gil. (2019). *UNFOLDING THE DRIVERS FOR ACADEMIC SUCCESS: THE CASE OF ISCTE-IUL* (p. 119). <http://hdl.handle.net/10071/20069>
- Goyal, M., & Vohra, R. (2012). Applications of Data Mining in Higher Education. *International Journal of Computer Science Issues*, 9.

- Levine, R. A., Rivera, P. E., He, L., Fan, J., & Bresciani Ludvick, M. J. (2023). A learning analytics case study: On class sizes in undergraduate writing courses. *Stat*, *12*(1). <https://doi.org/10.1002/sta4.527>
- Malva, M. (2021). PROFILE OF STUDENTS AT THE POLYTECHNIC INSTITUTE OF VISEU – AN EVOLUTIONARY STUDY SINCE THE TREATY OF BOLOGNA. *Millenium: Journal of Education, Technologies, and Health*, *2021*(9), 307–321. <https://doi.org/10.29352/mill029e.24879>
- Miguel, H. G., Ramos, P., da Cruz Martins, S., & Costa, J. M. (2020). Provision of academic data for research: A step for academic success. *Education for Information*, *36*(4), 389–407. <https://doi.org/10.3233/EFI-190349>
- Nash, J., Boucheron, L., Stochaj, S., & IEEE. (2021). A Correlative Analysis of Course Grades as Related to Curricular Prerequisite Structure and Inter-Class Topic Dependencies. *New Mexico State University. 2021 IEEE FRONTIERS IN EDUCATION CONFERENCE (FIE 2021)*. <https://doi.org/10.1109/FIE49875.2021.9637401>
- Oliveira, B., Mira, M., Monteiro, S., Elvas, L., Bras-Rosario, L., & Ferreira, J. (2023). *Implementing a Data Integration Infrastructure for Healthcare Data – A Case Study* (pp. 744–753). https://doi.org/10.1007/978-3-031-27499-2_69
- Osmanbegovic, E., & Suljic, M. (2012). Data Mining Approach for Predicting Student Performance. *Economic Review: Journal of Economics and Business*, *10*(1), 3–12.
- Quezada-Ugalde, A. M., García-Hernández, A. M., Maya-Barrios, J. A., & Auza-Benavides, A. (2023). Consequences of the COVID-19 Pandemic on Child Development: An Integrative Literature Review. *Aquichan*, *23*(3). <https://doi.org/10.5294/aqui.2023.23.3.2>

- Robbiano, C., Maciejewski, A., & Chong, E. (2020). Nonparametric Analysis of the Effect of Knowledge Integration Activities on Third-Year Undergraduate Performance. *IEEE TRANSACTIONS ON EDUCATION*, 63(4), 305–313. <https://doi.org/10.1109/TE.2020.2989920>
- Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A Review on Predicting Student's Performance Using Data Mining Techniques. *The Third Information Systems International Conference 2015*, 72, 414–422. <https://doi.org/10.1016/j.procs.2015.12.157>
- Shearer, C. (2000). The CRISP-DM model: The new blueprint for data mining. *Journal of data warehousing*, 5(4), 13–22.
- Trstenjak, B., & Donko, D. (2014). *Determining the impact of demographic features in predicting student success in Croatia*. 1222–1227. <https://doi.org/10.1109/MIPRO.2014.6859754>
- Wells, J., Spence, A., & McKenzie, S. (2021). STUDENT PARTICIPATION IN COMPUTING STUDIES TO UNDERSTAND ENGAGEMENT AND GRADE OUTCOME. *JOURNAL OF INFORMATION TECHNOLOGY EDUCATION-RESEARCH*, 20, 385–403. <https://doi.org/10.28945/4817>
- York, T. T., Gibson, C., & Rankin, S. (2015). Defining and measuring academic success. *Practical Assessment, Research and Evaluation*, 20(5), 1–20.
- Yungang, W., Lina, S., Bo, S., & Xiaoming, Z. (2017). *The course interest analysis of extracurricular scientific practical activity of students in grade seven*. 611–616. <https://doi.org/10.1109/ICCSE.2017.8085565>

Anexos

Anexo A: Análise de Literatura

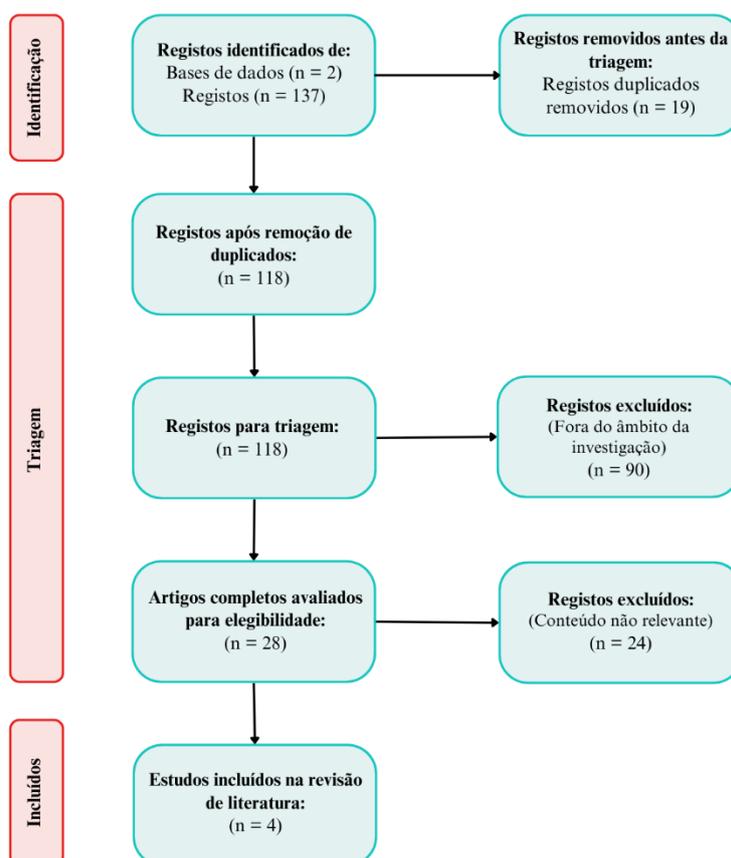


Figura 1. Fases da Revisão Sistemática da Literatura, com Base no Método PRISMA.
Fonte: A autora, adaptado de (Oliveira et al., 2023).

Tabela 1. Caracterização dos Artigos Analisados.

Nº	Título	Autores	Ano	Citações nos Repositórios
1	<i>A Correlative Analysis of Course Grades as Related to Curricular Prerequisite Structure and Inter-Class Topic Dependencies</i>	Nash, J; Boucheron, LE; Stochaj, SJ.	2021	4
2	<i>Nonparametric Analysis of the Effect of Knowledge Integration Activities on Third-Year Undergraduate Performance</i>	Robbiano, C; Maciejewski, AA; Chong, EKP.	2020	0
3	<i>Student Participation in Computing Studies to Understand Engagement and Grade Outcome</i>	Wells, JG; Spence, A; McKenzie, S.	2021	5
4	<i>The Course Interest Analysis of Extracurricular Scientific Practical Activity of Students in Grade Seven</i>	Yungang, W; Lina, S; Bo, S; Xiaoming, Z.	2017	1

Fonte: A autora, adaptado de (da Silva & Rossoni, 2023) e (Quezada-Ugalde et al., 2023), e retirado dos artigos mencionados na tabela.

Tabela 2. Descrição dos Objetivos, Amostra e Resultados dos Artigos Analisados.

Nº	Objetivo	Metodologia	Dados	Resultados Principais
1	Fornecer informação útil para redesenhar o curso de Engenharia Eletrônica (EE).	Análise da correlação entre as notas entre cada par <i>core</i> das disciplinas de EE, e das relações entre os tópicos lecionados.	531 alunos; 3884 notas individuais, reduzidas para 2845.	- Não existe tanta correlação nas notas das disciplinas pré- e pós-requisitadas, levando a acreditar que outros fatores podem ter maior importância no sucesso dos alunos; - Existem disciplinas com relação de pré-requisito, mas com pouca dependência a nível do tópico.
2	Quantificar as mudanças que as atividades KI causaram no desempenho dos estudantes de Engenharia Eletrônica e de Computadores.	Análise comparativa da <i>grade point average</i> (GPA), da quantidade de admissão dos alunos e da correlação das disciplinas, antes e após a implementação das atividades KI.	2800 registos de estudantes, reduzidos para 751.	- Existe uma correlação entre a performance em disciplinas de pré-requisito e requisito para ambos os grupos; - Correlação entre tópicos é maior para os alunos com atividades KI; - O GPA é cada vez menos indicativo da performance individual para os alunos com atividades KI.
3	Análise dos estudantes de Informática a partir da atividade no LMS, e nas notas em 3 disciplinas, com o intuito de analisar a atividade dos estudantes para desenhar as disciplinas.	Abordagem quantitativa para analisar a atividade no LMS e as notas das 3 disciplinas, com especial foco em contagens e médias.	1063 alunos, 280 da disciplina 1, 245 da disciplina 2, 538 da disciplina 3.	- Apesar de receberem resultados mais baixos, os alunos continuam a reportar atividade; - A atividade é baixa durante o período de avaliação, e aumenta nas semanas de exames; - Tarefas com maior peso na nota encorajam atividade.
4	Analisar a popularidade das disciplinas com base na seleção dos mesmos.	Histogramas de frequência da seleção e pesquisa das disciplinas, gráficos de dispersão e análise das disciplinas e da popularidade das mesmas.	848 disciplinas.	- Quanto maior a avaliação da disciplina, maior a popularidade, com importância do passa-palavra e do modo de ensino; - Os distritos com menos opções resultam na escolha de disciplinas menos populares; - Disciplinas oferecidas por museus têm maior popularidade; - As disciplinas mais populares são Astronomia, Eletrônica, Tecnologia de Informação (os estudantes aceitaram bem o novo currículo) e Física; - As menos populares são Geografia e Dados e Informação.

Fonte: A autora, adaptado de (da Silva & Rossoni, 2023) e (Quezada-Ugalde et al., 2023), e retirado dos artigos mencionados na tabela.

Tabela 3. Identificação dos Locais de Estudo dos Artigos Analisados.

Continente	País	Quantidade	Nº Artigos
América	Estados Unidos da América	2	1, 2
Austrália	Austrália	1	3
Ásia	China	1	4

Fonte: A autora, retirado dos artigos mencionados na tabela.

Anexo B: Tratamento dos Dados

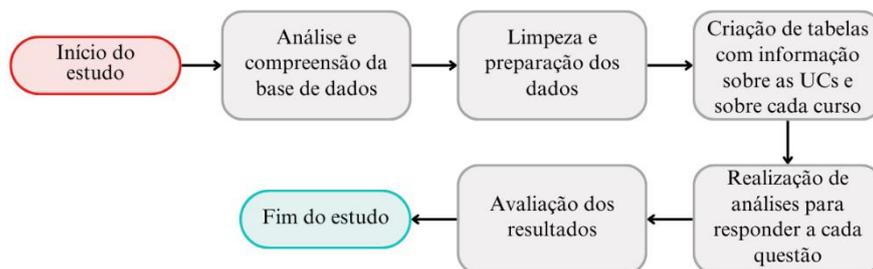


Figura 2. Organização da Metodologia.

Tabela 4. Descrição das Tabelas da Base de Dados.

Tabela	Descrição
Student	Descreve estudantes do Iscte.
StudentSpecialRegime	Inclui um subgrupo de estudantes com regimes especiais.
StudentStatute	Descreve os estatutos dos estudantes.
SocialServices	Representa as condições dos serviços sociais dos estudantes.
Candidacy	Descreve as candidaturas ao ensino superior realizadas no Iscte
CandidacyPreferences	Inclui as preferências das candidaturas efetuadas.
Degree	Representa os cursos do Iscte.
School	Descreve as escolas do Iscte.
Registration	Reune os registos efetuados.
Enrollment	Descreve as inscrições feitas em unidades curriculares.
EnrollmentExecutionCurricularUnit	Permite relacionar as inscrições com as respetivas unidades curriculares.
ExecutionCurricularUnit	Representa as unidades curriculares.
ExecutionPeriod	Inclui os períodos de execução referentes a anos letivos.
ExecutionYear	Inclui os anos letivos de execução.
OutgoingMobility	Inclui informação relacionada com os estudantes em mobilidade.
RegistrationByExecutionYear	Relaciona o registo ao ano de execução.
SchoolDegree	Associa as escolas aos cursos.

Fonte: A autora, adaptado de (Miguel et al., 2020)

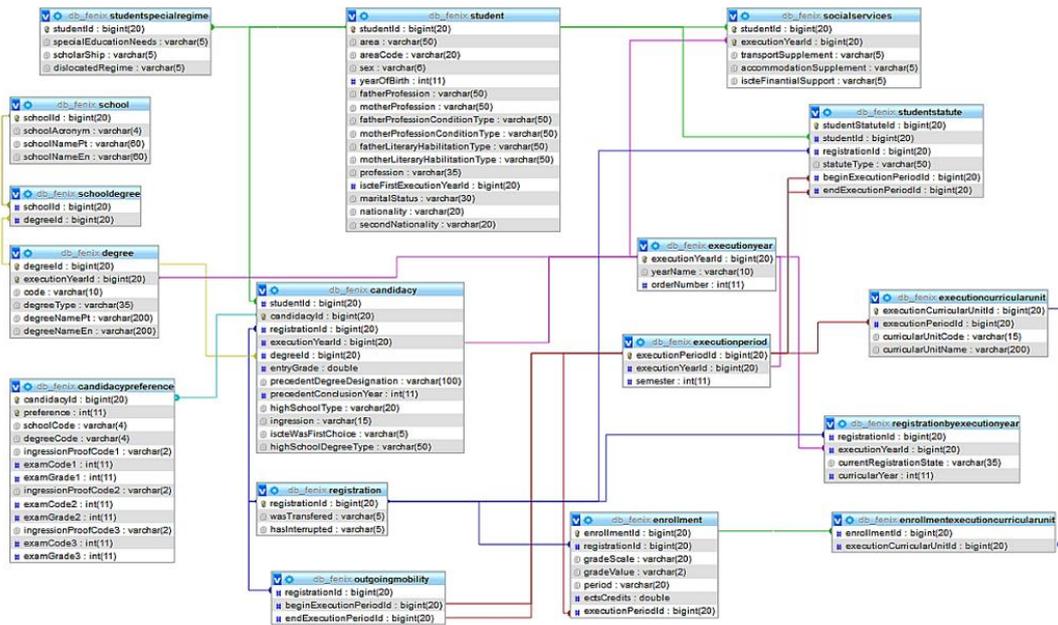


Figura 3. Modelo Relacional
 Fonte: Adaptado de (Miguel et al., 2020).

Tabela 5. Descrição da Base de Dados.

Tabela	Coluna	Tipo	Chave	Descrição	Notas
Student	studentId	bignit	chave primária	Número de identificação do estudante	
	area	varchar(50)		Localidade	
	areaCode	varchar(20)		Código Postal	
	sex	varchar(6)		Sexo	
	yearOfBirth	int		Ano de Nascimento	
	fatherProfession	varchar(50)		Profissão do pai	
	motherProfession	varchar(50)		Profissão da mãe	
	fatherProfessionConditionType	varchar(50)		Condição da profissão do pai	
	motherProfessionConditionType	varchar(50)		Condição da profissão da mãe	
	fatherLiteraryHabilitationType	varchar(50)		Qualificações educacionais do pai	
	motherLiteraryHabilitationType	varchar(50)		Qualificações educacionais do pai	
	profession	varchar(35)		Profissão do estudante	
	iscteFirstExecutionYearId	bignit		O ano em que entrou no Iscte	Na dissertação, está com este nome: iscteFirstExecutionYearExternalId
	maritalStatus	varchar(30)		Estado civil do estudante	
	nationality	varchar(20)		Nacionalidade do estudante	
secondNationality	varchar(20)		Segunda nacionalidade do estudante		
StudentsSpecialRegime	studentId	bignit	chave primária	-	
	specialEducationNeeds	varchar(5)		Necessidades educativas especiais	True, se o estudante tiver o estatuto de necessidades especiais
	scholarShip	varchar(5)		Condição de bolseiro	True, se o estudante tiver algum tipo de bolsa (FCT ou SAS)

	dislocatedRegime	vchar(5)		Regime de deslocação	<i>True</i> , se, na inscrição via DGES, o estudante tiver seleccionado que estava deslocado
StudentStatute	studentStatuteId	bignit	chave primária	Número de identificação do estatuto do estudante	
	studentId	bignit	chave estrangeira 1	-	
	registrationId	bignit	chave estrangeira 2	-	
	statuteType	vchar(50)	-	Designação do tipo de estatuto	Ex: bolsa SAS ou FCT, trabalhador-estudante, estudante a meio termo, estudante internacional, etc
	beginExecutionPeriodId	bignit	chave estrangeira 3	Número que identifica o início do estatuto	
	endExecutionPeriodId	bignit	chave estrangeira 4	Número que identifica o fim do estatuto	
SocialServices	studentId	bignit	chave primária e chave estrangeira	-	
	executionYearId	bignit	chave estrangeira	-	
	transportSupplement	vchar(5)	-	Suplemento de transporte	<i>True</i> , se o estudante tem suplemento de transporte
	accommodationSupplement	vchar(5)		Suplemento de habitação	<i>True</i> , se o estudante tem suplemento de habitação
	iscteFinancialSupport	vchar(5)		Suplemento financeiro do Iscte	<i>True</i> , se o estudante tem suplemento financeiro dado pelo Iscte
Candidacy	studentId	bignit	chave estrangeira 1	-	
	candidacyId	bignit	chave primária	-	
	registrationId	bignit	chave estrangeira 2	-	
	executionYearId	bignit	chave estrangeira 3	-	
	degreeId	bignit	chave estrangeira 4	-	
	entryGrade	double	-	Nota de acesso	
	precedentDegreeDesignation	vchar(100)		Designação do grau anterior	
	precedentConclusionYear	int		Ano de conclusão do grau anterior	
	highSchoolType	vchar(20)		Tipo de escola do ensino secundário	Ex: Pública, privada ou ambas
	ingression	vchar(15)		Tipo de ingresso ou acesso ao ensino superior	Ex: Regime geral de acesso, titular de licenciatura, etc
	iscteWasFirstChoice	vchar(5)		Se o Iscte foi a sua primeira escolha ou não	<i>True</i> , se o Iscte tiver sido a primeira escolha do estudante
	highSchoolDegreeType	vchar(50)	Curso ou área do ensino secundário	Ex: Curso de Ciências e Tecnologias, equivalência, etc	
CandidacyPreference	candidacyId	bignit	chave primária e chave estrangeira	-	
	preference	int	chave primária	Número de ordem de preferência	
	schoolCode	vchar(4)	-	Código do instituto de ensino superior	
	pretendedDegreeCode	vchar(4)		Código do curso pretendido	
	ingressionProofCode1	vchar(2)		Código do exame de ingresso da DGES para o ensino superior (1)	
	examCode1	int		Código do exame da DGES (1)	
	examGrade1	int		Nota do exame (1)	
	ingressionProofCode2	vchar(2)		-	
	examCode2	int		-	
	examGrade2	int		-	
ingressionProofCode3	vchar(2)	-			

	examCode3	int		-	
	examGrade3	int		-	
Degree	degreeId	bignit	chave primária	Número de identificação do curso	
	executionYearId	bignit	chave estrangeira	-	
	degreeCode	varchar(10)	-	Código do curso	
	degreeType	varchar(35)		Tipo de grau	Ex: Licenciatura, mestrado, doutoramento, etc
	degreeNamePt	varchar(200)		Designação do curso em Português	
	degreeNameEn	varchar(200)		Designação do curso em Inglês	
School	schoolId	bignit	chave primária	Número de identificação da escola do Iscte	
	schoolAcronym	varchar(4)	-	Acrónimo da escola	Ex: ISTA, EG, ECSH, etc
	schoolNamePt	varchar(60)		Designação completa da escola em português	
	schoolNameEn	varchar(60)		Designação completa da escola em inglês	
schoolDegree	schoolId	bignit	chave primária e chave estrangeira 1	-	
	degreeId	bignit	chave primária e chave estrangeira 2	-	
registration	registrationId	bignit	chave primária	Número de identificação do registo	
	wasTransferred	varchar(5)	-	Estudante foi transferido	<i>True</i> , se foi requerida a mudança de curso
	hasInterrupted	varchar(5)		Estudante interrompeu estudos	<i>True</i> , se o curso foi interrompido
registrationByExecutionYear	registrationId	bignit	chave primária e chave estrangeira 1	-	
	executionYearId	bignit	chave primária e chave estrangeira 2	-	
	currentRegistrationState	varchar(35)	-	Estado de registo	Ex: Registado, interrompido, desistiu, concluído, etc
	curricularYear	int		Ano do curso no qual está registado	
outgoingMobility	registrationId	bignit	chave primária e chave estrangeira 1	-	
	beginExecutionPeriodId	bignit	chave primária e chave estrangeira 2	Número de identificação do início do período de mobilidade	
	endExecutionPeriodId	bignit	chave primária e chave estrangeira 3	Número de identificação do fim do período de mobilidade	
enrollment	enrollmentId	bignit	chave primária	Número de identificação de matrícula	
	registrationId	bignit	chave estrangeira 1	-	
	gradeScale	varchar(20)	-	Escala de nota de matrícula	
	gradeValue	varchar(2)		Valor da nota de matrícula	
	period	varchar(20)		Período de matrícula	
	ectsCredits	double	Créditos ECTS de matrícula		
executionPeriodId	bignit	chave estrangeira 2	-		
executionCurricularUnit	executionCurricularUnitId	bignit	chave primária	Número de identificação da unidade curricular	
	executionPeriodId	bignit	chave estrangeira	-	

	curricularUnitCode	varchar(15)	-	Código da unidade curricular	
	curricularUnitName	varchar(200)		Nome da unidade curricular	
enrollmentExecutionCurricular Unit	enrollmentId	bignit	chave primária e chave estrangeira 1	-	
	executionCurricularUnitId	bignit	chave primária e chave estrangeira 2	-	
executionPeriod	executionPeriodId	bignit	chave primária	Número de identificação do período	
	executionYearId	bignit	chave primária e chave estrangeira	-	
	semester	int	-	Designação do semestre	Ex: 1 ou 2
executionYear	executionYearId	bignit	chave primária	Número de identificação do ano	
	yearName	varchar(10)	-	Nome do ano	
	orderNumber	int		Valor para organização dos anos	Ex: 2010/2011

Fonte: A autora, adaptado de (Miguel et al., 2020)

Tabela 6. Tipo de Dados

Tipo de Dados	Nome Completo	Significado
int	Número Inteiro	Usado para representar números inteiros.
bignit	Grande Número Inteiro	Usado para representar números inteiros de dimensão superior ao tipo de dados anterior.
double	Duplo	Usado para representar números reais com valores fracionados.
varchar	Campo de Carácter Variável	Usado para representar dados de qualquer tipo.

Fonte: A autora, adaptado de (Miguel et al., 2020)

Tabela 7. Descrição dos Cursos.

Curso	Quantidade de Alunos	Quantidade de Unidades Curriculares
Engenharia Informática	883	137
Engenharia Informática Pós-Laboral	316	82
Engenharia de Telecomunicações e Informática	954	157
Engenharia de Telecomunicações e Informática Pós-Laboral	234	81
Informática e Gestão de Empresas	1002	151
Informática e Gestão de Empresas Pós-Laboral	434	86

Tabela 8. Unidades Curriculares por Curso.

Curso	Quantidade de Unidades Curriculares
Engenharia Informática	38
Engenharia Informática Pós-Laboral	30
Engenharia de Telecomunicações e Informática	52
Engenharia de Telecomunicações e Informática Pós-Laboral	32
Informática e Gestão de Empresas	57
Informática e Gestão de Empresas Pós-Laboral	42

Tabela 9. Pares Relevantes por Curso.

Curso	Quantidade de Pares Relevantes
Engenharia Informática	30
Engenharia Informática Pós-Laboral	10
Engenharia de Telecomunicações e Informática	25
Engenharia de Telecomunicações e Informática Pós-Laboral	15
Informática e Gestão de Empresas	24
Informática e Gestão de Empresas Pós-Laboral	12

Anexo C: Análise de Dados da Questão 1

Engenharia Informática

a) Scree Plot e Clustering

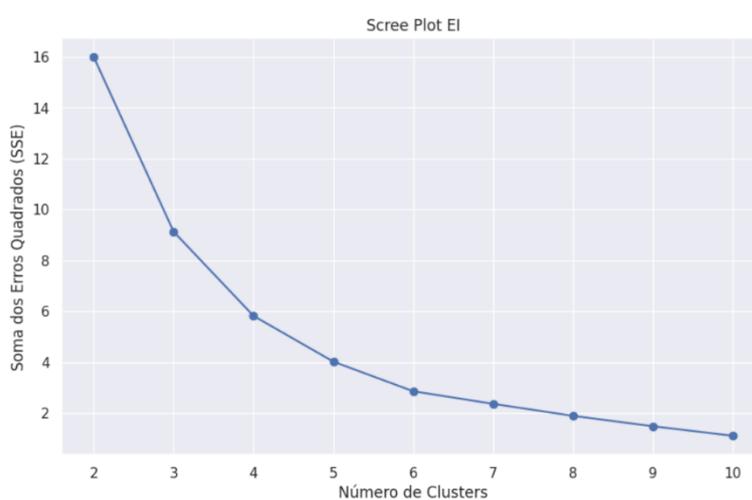


Figura 4. Scree Plot EI.

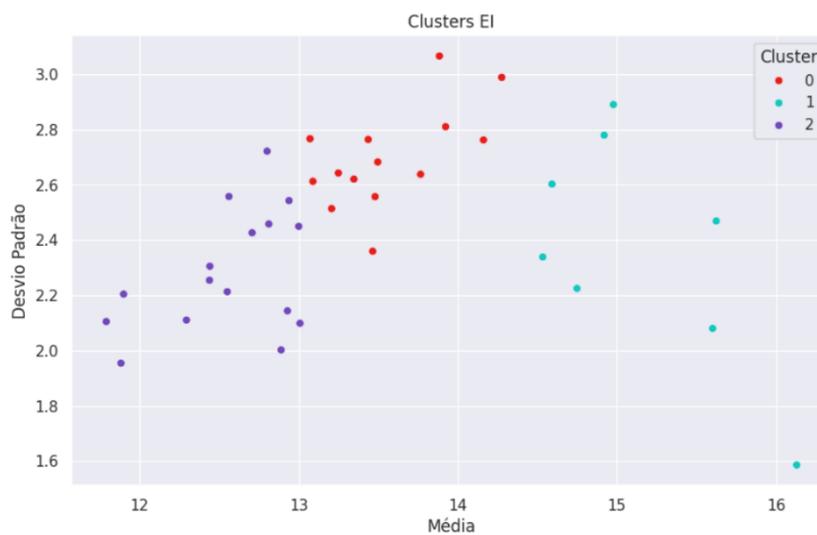


Figura 5. Gráfico de Dispersão EI.

Tabela 10. Distribuição de Unidades Curriculares e Áreas por Cluster EI.

<i>Cluster 0</i>	<i>Área</i>	<i>Cluster 1</i>	<i>Área</i>	<i>Cluster 2</i>	<i>Área</i>
Álgebra Linear, Geometria Analítica e Análise Vetorial	Matemática	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	Análise Matemática I	Matemática
Algoritmos e Estruturas de Dados	Programação	Interação Pessoa-Máquina	Multimídia	Análise Matemática II	Matemática
Bases de Engenharia II	Matemática	Multimídia e Computação Gráfica	Multimídia	Arquitetura de Computadores I	Arquitetura de Computadores
Engenharia de Software I	Programação	Planeamento de Projetos Utilizando Ferramentas Informáticas (Msproject)	Informática Aplicada	Arquitetura de Computadores II	Arquitetura de Computadores
Engenharia de Software II	Programação	Processamento de Informação	Telecomunicações	Bases de Engenharia I	Matemática
Inteligência Artificial	Inteligência Artificial	Programação em Rede	Programação	Circuitos para Comunicações	Telecomunicações
Introdução à Programação	Programação	Sistemas de Informação Distribuídos	Bases de Dados, Modelação	Conceção e Desenvolvimento de Sistemas de Informação	Bases de Dados, Modelação
Mecânica e Eletricidade	Física e Eletromagnetismo	Som e Vídeo para Multimídia	Multimídia	Fluxo de Informação em Sistemas	Telecomunicações
Processamento de Informação II	Telecomunicações			Fundamentos de Bases de Dados	Bases de Dados, Modelação
Programação Orientada para Objetos	Programação			Gestão e Contabilidade Empresarial	Contabilidade e Finanças
Sistemas Operativos	Sistemas Operativos			Microprocessadores	Arquitetura de Computadores
Tecnologias para Sistemas Inteligentes	Inteligência Artificial			Processamento de Informação I	Telecomunicações
Teoria da Computação	Programação			Programação Concorrente e Distribuída	Programação
Teoria do Sinal	Telecomunicações			Redes Digitais I - Fundamentos	Redes
				Redes Digitais II - Sistemas, Aplicações e Serviços	Redes
				Redes Digitais III - Segurança, Multimídia e Gestão	Redes

b) Análise por Área

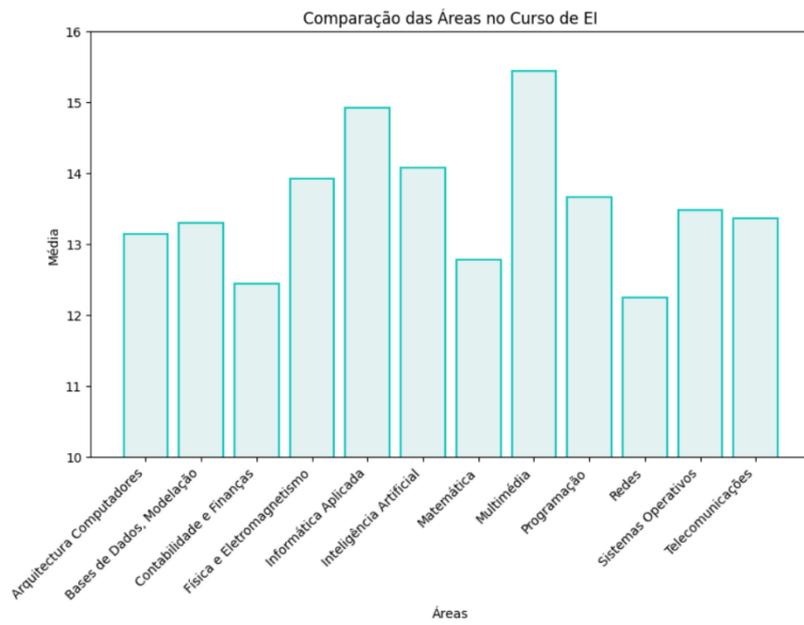


Figura 6. Médias das Áreas EI

c) Áreas Mais Difíceis

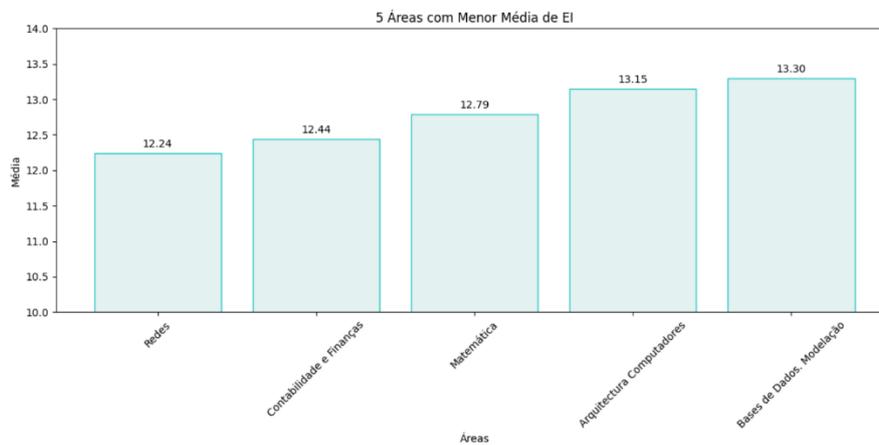


Figura 7. Áreas com Menor Média EI.

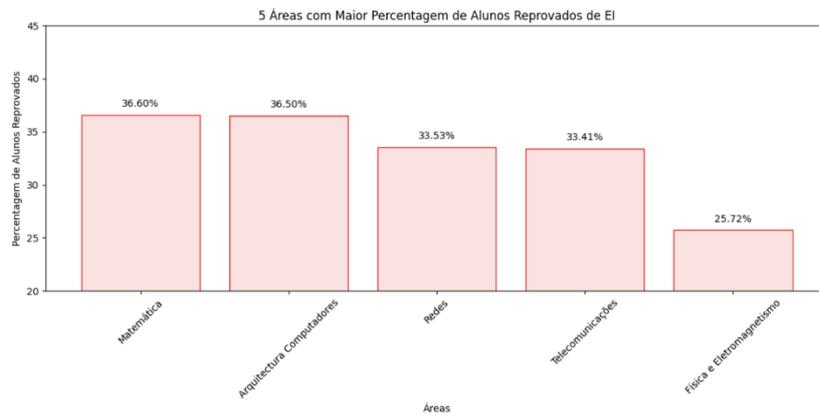


Figura 8. Áreas com Maior Percentagem de Alunos Reprovados EI.

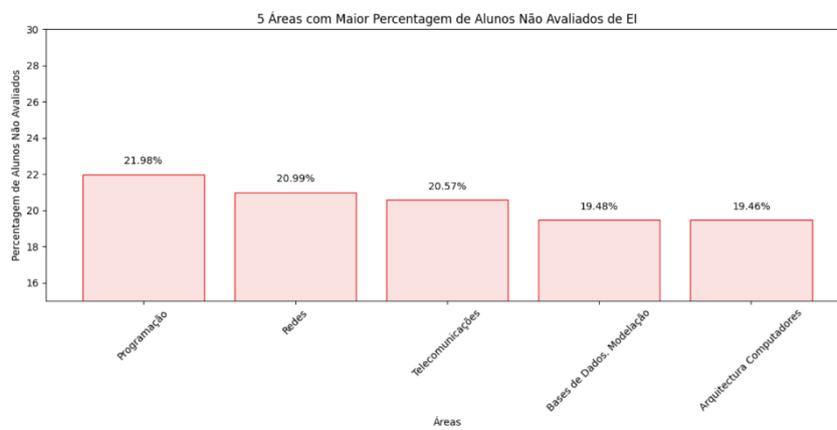


Figura 9. Áreas com Maior Percentagem de Alunos Não Avaliados EI.

Engenharia Informática Pós-Laboral

a) Scree Plot e Clustering

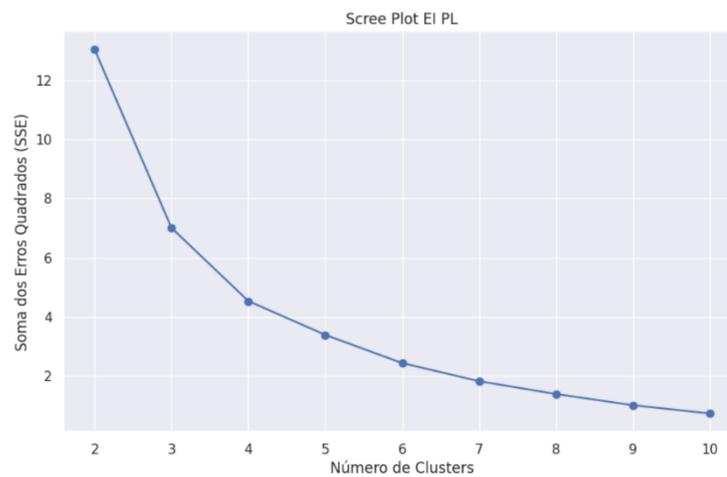


Figura 10. Scree Plot EI PL.

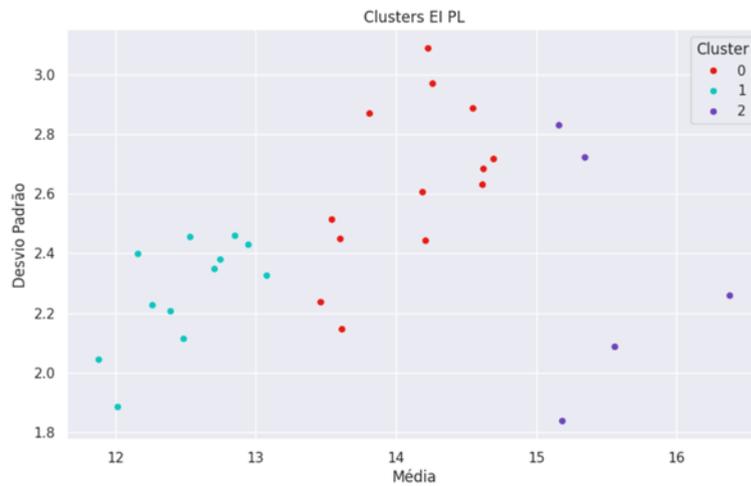


Figura 11. Gráfico de Dispersão EI PL.

Tabela 11. Distribuição de Unidades Curriculares e Áreas por Cluster EI PL.

Cluster 0	Área	Cluster 1	Área	Cluster 2	Área
Conceção e Desenvolvimento de Sistemas de Informação	Bases de Dados, Modelação	Álgebra Linear, Geometria Analítica e Análise Vetorial	Matemática	Interação Pessoa-Máquina	Multimédia
Engenharia de Software I	Programação	Algoritmos e Estruturas de Dados	Programação	Planeamento de Projetos Utilizando Ferramentas Informáticas (Msproject)	Informática Aplicada
Engenharia de Software II	Matemática	Análise Matemática I	Matemática	Processamento de Informação	Telecomunicações
Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	Análise Matemática II	Matemática	Programação em Rede	Programação
Inteligência Artificial	Inteligência Artificial	Circuitos para Comunicações	Telecomunicações	Sistemas de Informação Distribuídos	Bases de Dados, Modelação
Introdução à Programação	Programação	Fundamentos de Bases de Dados	Bases de Dados, Modelação		
Mecânica e Eletricidade	Física e Eletromagnetismo	Gestão e Contabilidade Empresarial	Contabilidade e Finanças		
Multimédia e Computação Gráfica	Telecomunicações	Microprocessadores	Arquitetura de Computadores		
Programação Orientada para Objetos	Programação	Programação Concorrente e Distribuída	Programação		
Sistemas Operativos	Sistemas Operativos	Redes Digitais I - Fundamentos	Redes		
Tecnologias para Sistemas Inteligentes	Inteligência Artificial	Redes Digitais II - Sistemas, Aplicações e Serviços	Redes		
Teoria da Computação	Programação	Redes Digitais III - Segurança, Multimédia e Gestão	Redes		
Teoria do Sinal	Telecomunicações				

b) Análise por Área

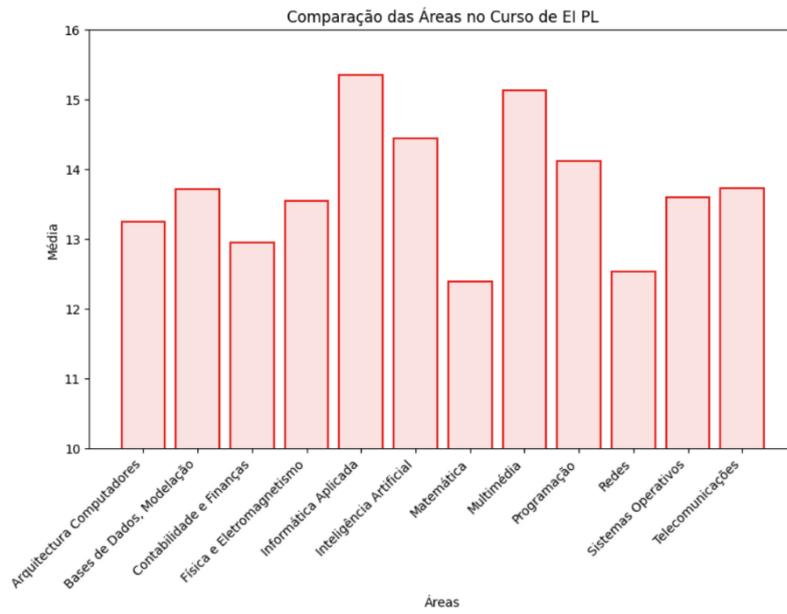


Figura 12. Médias das Áreas EI PL.

c) Áreas Mais Difíceis

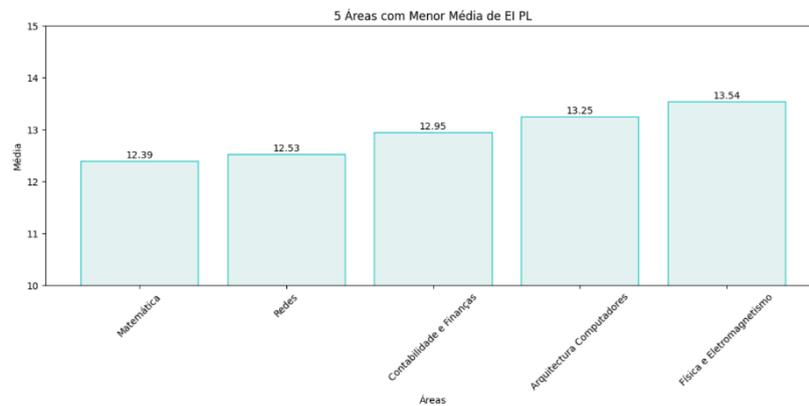


Figura 13. Áreas com Menor Média EI PL.

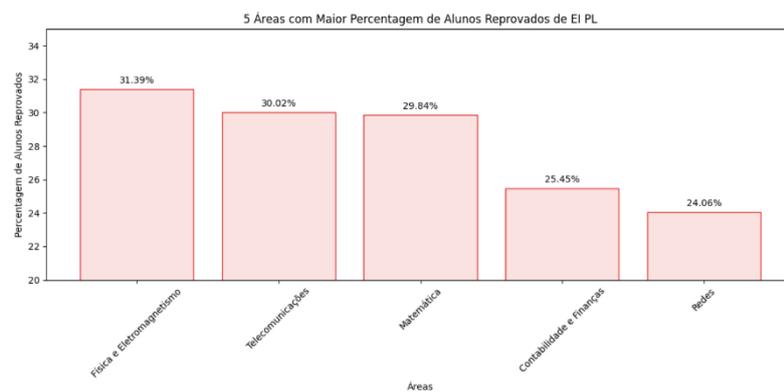


Figura 14. Áreas com Maior Percentagem de Alunos Reprovados EI PL.

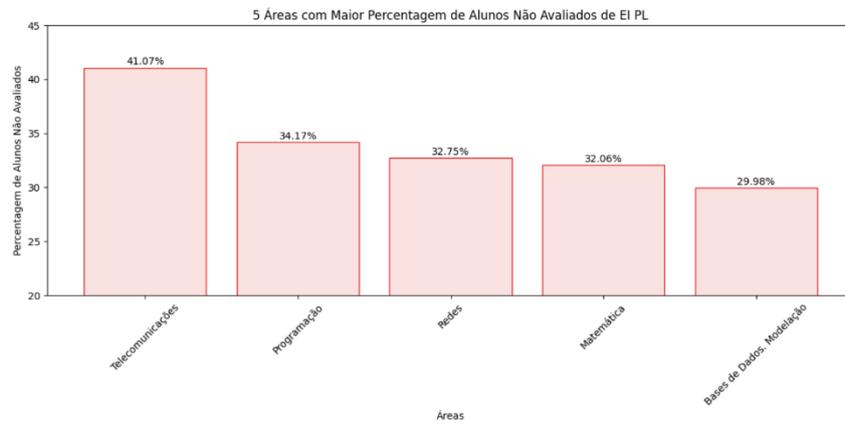


Figura 15. Áreas com Maior Percentagem de Alunos Não Avaliados EI PL.

Engenharia de Telecomunicações e Informática

a) Scree Plot e Clustering

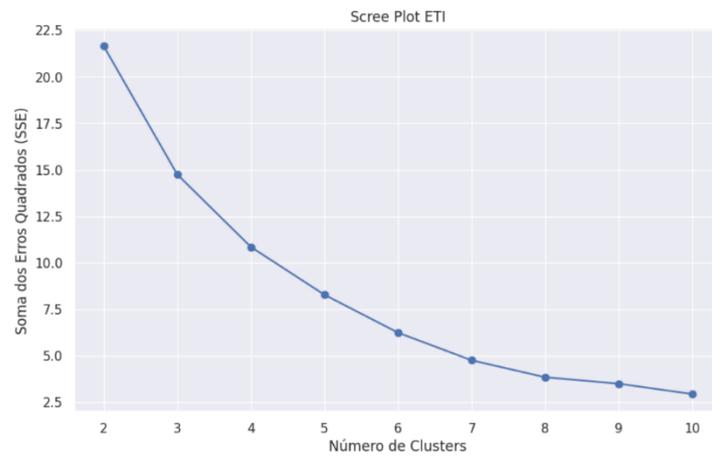


Figura 16. Scree Plot ETI.

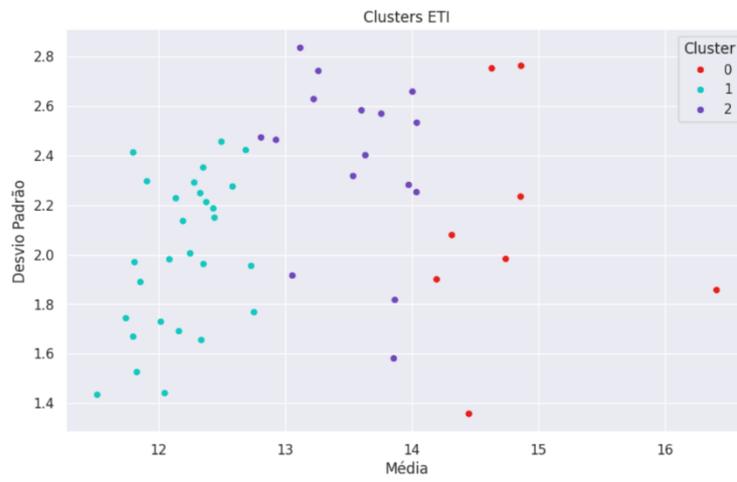


Figura 17. Gráfico de Dispersão ETI.

Tabela 12. Distribuição de Unidades Curriculares e Áreas por Cluster ETI.

Cluster 0	Área	Cluster 1	Área	Cluster 2	Área
Computação Gráfica	Multimédia	Análise Matemática I (Pré Bolonha)	Matemática	Algoritmos e Estruturas de Dados	Programação
Introdução ao Spss	Informática Aplicada	Análise Matemática II (Pré Bolonha)	Matemática	Análise Matemática II	Matemática
Introdução às Redes Sociais	Outros	Arquitetura de Computadores	Arquitetura de Computadores	Eletrónica Programada para Tecnologias da Informação e Comunicação	Informática Aplicada
Planeamento de Projetos Utilizando Ferramentas Informáticas (Msproject)	Informática Aplicada	Arquitetura de Computadores (Pré Bolonha)	Arquitetura de Computadores	Engenharia de Software I	Programação
Processamento de Sinal Multimédia	Telecomunicações	Circuitos e Sistemas Eletrónicos para Tecnologias da Informação e Comunicação	Arquitetura de Computadores	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores
Sistemas de Telecomunicações por Rádio	Telecomunicações	Conceção e Desenvolvimento de Sistemas de Informação	Bases de Dados, Modelação	Gestão Financeira de Empresas e Projetos I	Contabilidade e Finanças
Sistemas e Redes de Comunicação para Móveis	Telecomunicações	Eletromagnetismo	Física e Eletromagnetismo	Gestão Financeira de Empresas e Projetos II	Contabilidade e Finanças
Sistemas Embebidos	Arquitetura de Computadores	Eletrónica Programada e Processamento Digital de Sinais	Informática Aplicada	Inteligência Artificial	Inteligência Artificial
		Física	Física e Eletromagnetismo	Inteligência e Gestão de Redes e Serviços	Redes
		Física (Pré Bolonha)	Física e Eletromagnetismo	Introdução à Programação	Programação
		Fundamentos de Bases de Dados	Bases de Dados, Modelação	Mecânica e Eletricidade	Física e Eletromagnetismo
		Fundamentos de Eletrónica	Física e Eletromagnetismo	Ondas e Ótica	Telecomunicações
		Introdução à Programação (Pré Bolonha)	Programação	Programação Orientada para Objetos	Programação
		Introdução às Bases de Engenharia	Outros	Sistemas e Redes de Comunicação para Móveis Avançados	Redes
		Introdução às Bases de Engenharia (Pré Bolonha)	Outros	Tecnologias para Sistemas Inteligentes	Inteligência Artificial
		Modulação e Codificação	Telecomunicações	Teoria do Sinal	Telecomunicações
		Multiplexagem, Comutação e Integração de Serviços	Telecomunicações		
		Programação Concorrente e Distribuída	Programação		
		Propagação e Radiação de Ondas Eletromagnéticas	Programação		
		Redes Digitais I - Fundamentos	Redes		
		Redes Digitais II - Sistemas, Aplicações e Serviços	Redes		
		Redes Digitais III - Segurança, Multimédia e Gestão	Redes		
		Sistemas de Telecomunicações Guiados	Telecomunicações		
		Sistemas Operativos	Sistemas Operativos		
		Teoria dos Circuitos	Arquitetura de Computadores		
		Teoria dos Circuitos (Pré Bolonha)	Arquitetura de Computadores		

b) Análise por Área

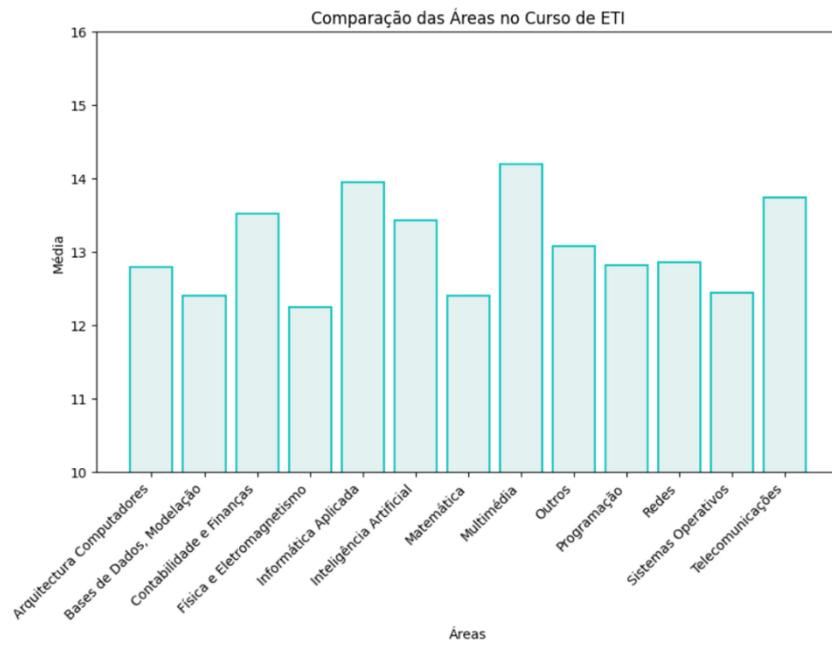


Figura 18. Médias das Áreas ETI.

c) Áreas Mais Difíceis

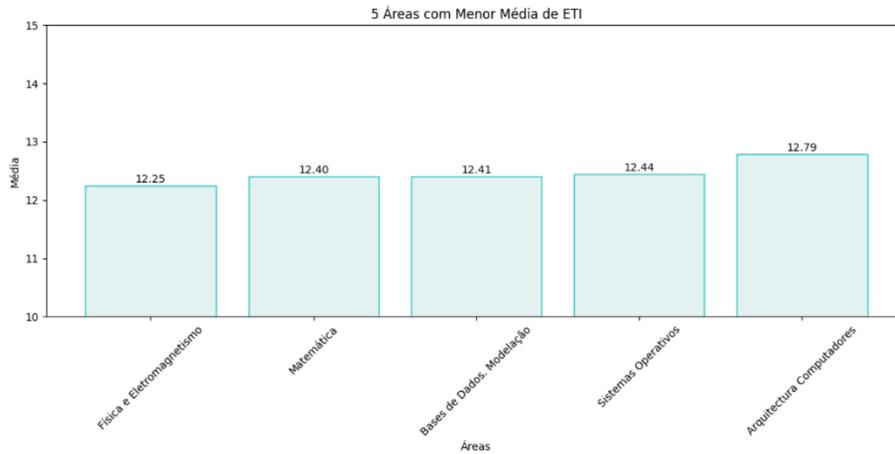


Figura 19. Áreas com Menor Média ETI.

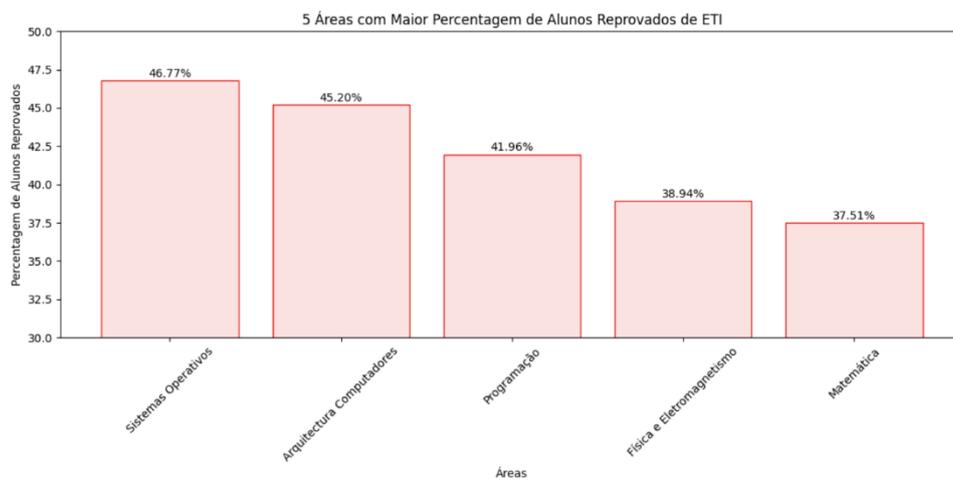


Figura 20. Áreas com Maior Percentagem de Alunos Reprovados ETI.

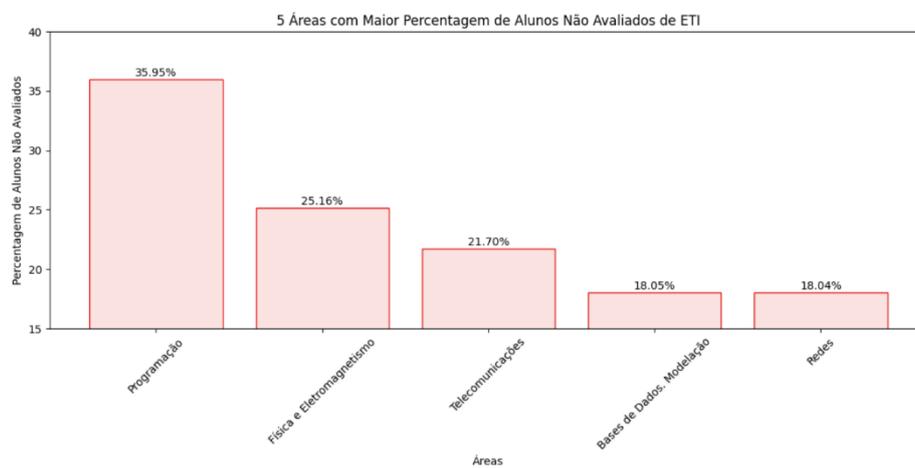


Figura 21. Áreas com Maior Percentagem de Alunos Não Avaliados ETI.

Engenharia de Telecomunicações e Informática Pós-Laboral

a) Scree Plot e Clustering

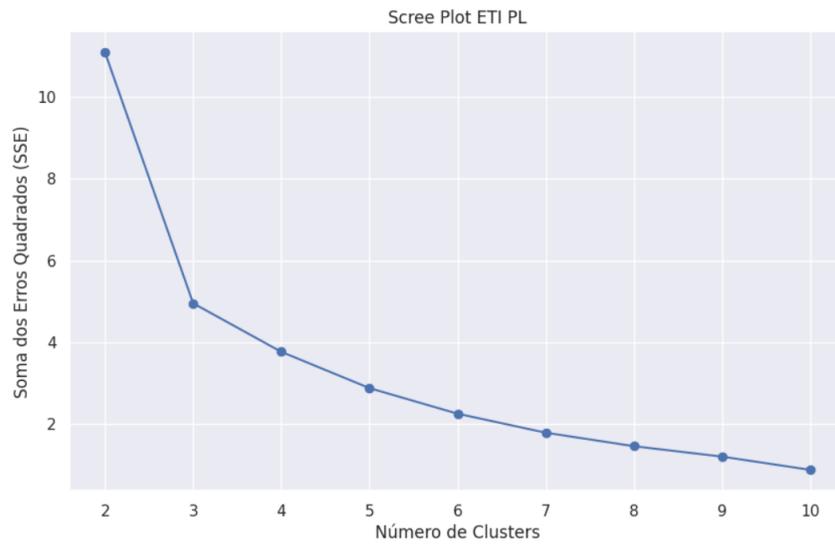


Figura 22. Scree Plot ETI PL.

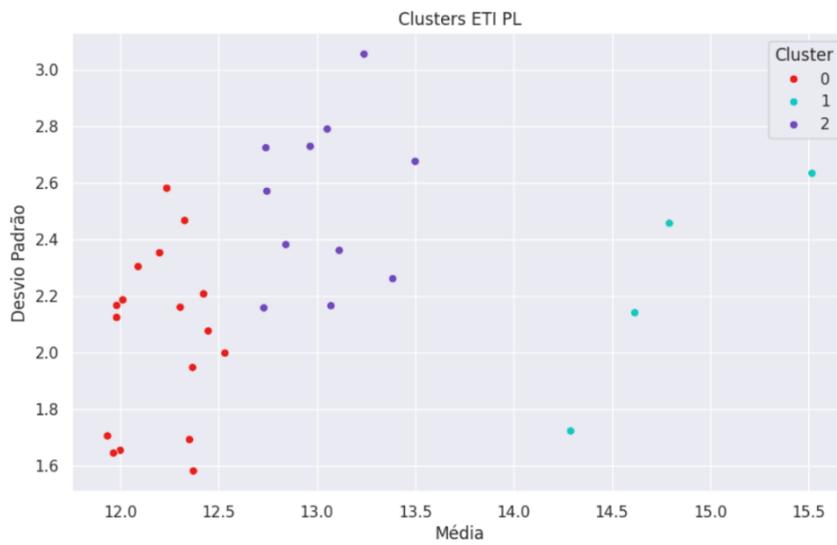


Figura 23. Gráfico de Dispersão ETI PL.

Tabela 13. Distribuição de Unidades Curriculares e Áreas por Cluster ETI PL.

<i>Cluster 0</i>	Área	<i>Cluster 1</i>	Área	<i>Cluster 2</i>	Área
Álgebra Linear, Geometria Analítica e Análise Vetorial	Matemática	Ondas e Ótica	Telecomunicações	Algoritmos e Estruturas de Dados	Programação
Análise Matemática I	Matemática	Planeamento de Projetos Utilizando Ferramentas Informáticas (Msproject)	Informática Aplicada	Análise Matemática II	Matemática
Conceção e Desenvolvimento de Sistemas de Informação	Bases de Dados, Modelação	Sistemas de Telecomunicações por Rádio	Telecomunicações	Arquitetura de Computadores	Arquitetura de Computadores
Eletromagnetismo	Física e Eletromagnetismo	Sistemas e Redes de Comunicação para Móveis	Telecomunicações	Eletrónica Programada e Processamento Digital de Sinais	Informática Aplicada
Física	Física e Eletromagnetismo			Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores
Fundamentos de Bases de Dados	Bases de Dados, Modelação			Introdução à Programação	Programação
Fundamentos de Eletrónica	Física e Eletromagnetismo			Mecânica e Eletricidade	Física e Eletromagnetismo
Inteligência Artificial	Inteligência Artificial			Programação Orientada para Objetos	Programação
Modulação e Codificação	Telecomunicações			Sistemas Operativos	Sistemas Operativos
Multiplexagem, Comutação e Integração de Serviços	Telecomunicações			Tecnologias para Sistemas Inteligentes	Inteligência Artificial
Programação Concorrente e Distribuída	Programação			Teoria do Sinal	Telecomunicações
Propagação e Radiação de Ondas Eletromagnéticas	Programação				
Redes Digitais I - Fundamentos	Redes				
Redes Digitais II - Sistemas, Aplicações e Serviços	Redes				
Redes Digitais III - Segurança, Multimédia e Gestão	Redes				
Sistemas de Telecomunicações Guiados	Telecomunicações				
Teoria dos Circuitos	Arquitetura de Computadores				

b) Análise por Área

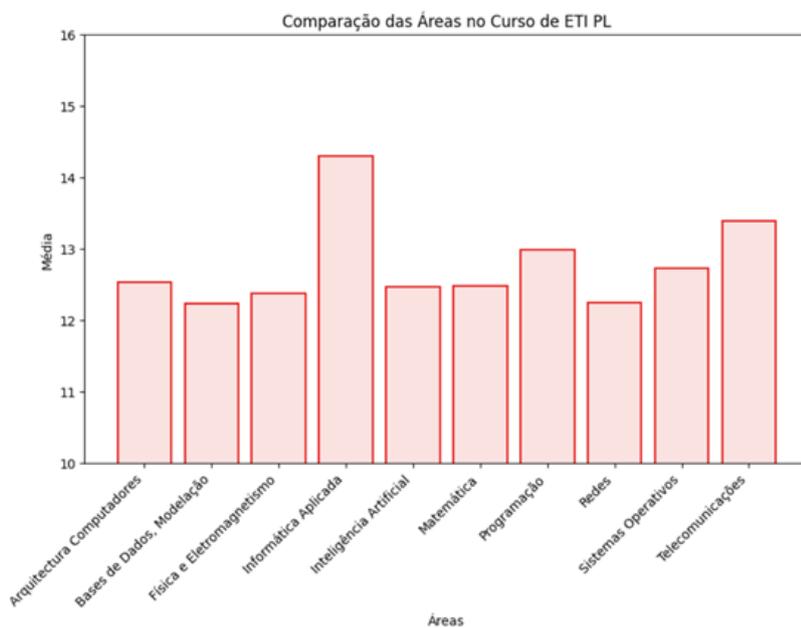


Figura 24. Médias das Áreas ETI PL.

c) Áreas Mais Difíceis

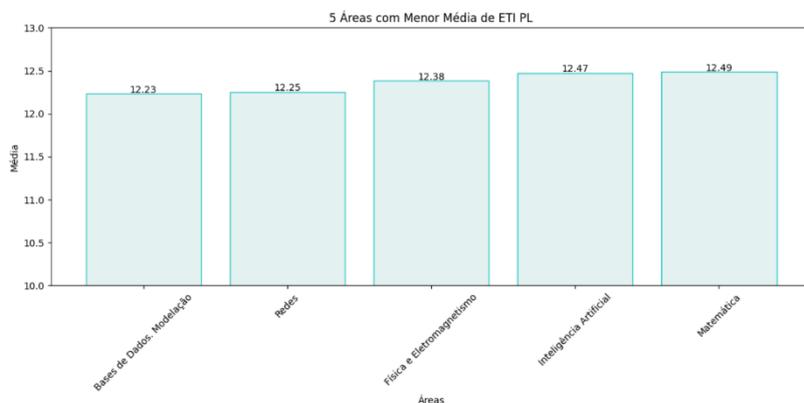


Figura 25. Áreas com Menor Média ETI PL.

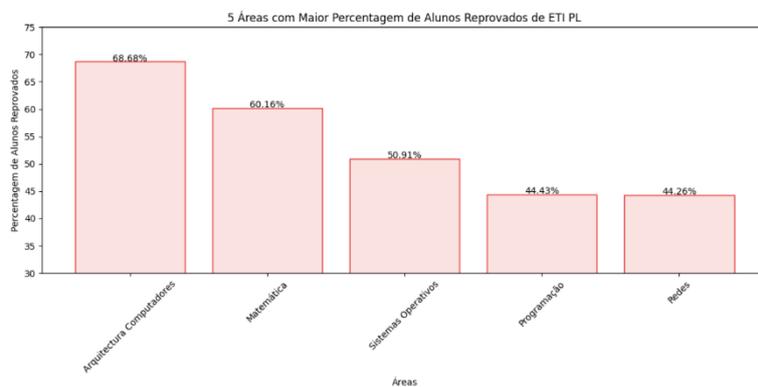


Figura 26. Áreas com Maior Percentagem de Alunos Reprovados ETI PL.

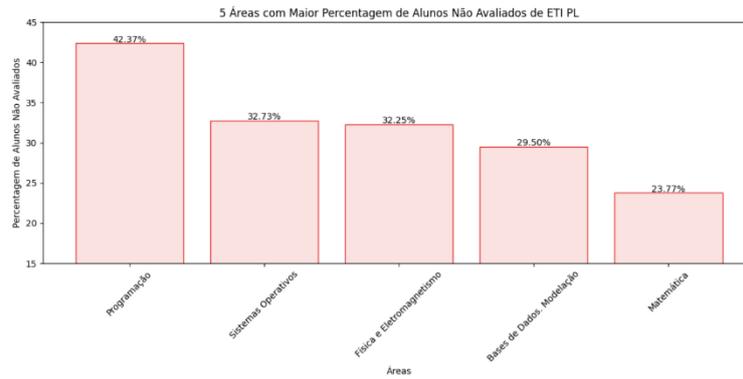


Figura 27. Áreas com Maior Percentagem de Alunos Não Avaliados ETI PL.

Informática e Gestão de Empresas

a) Scree Plot e Clustering

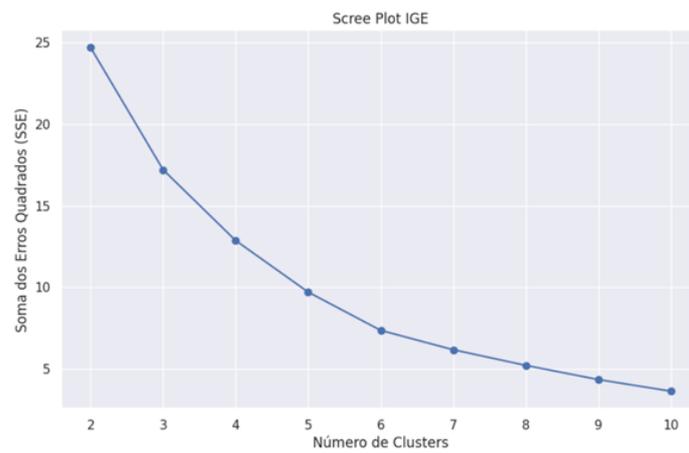


Figura 28. Scree Plot IGE.

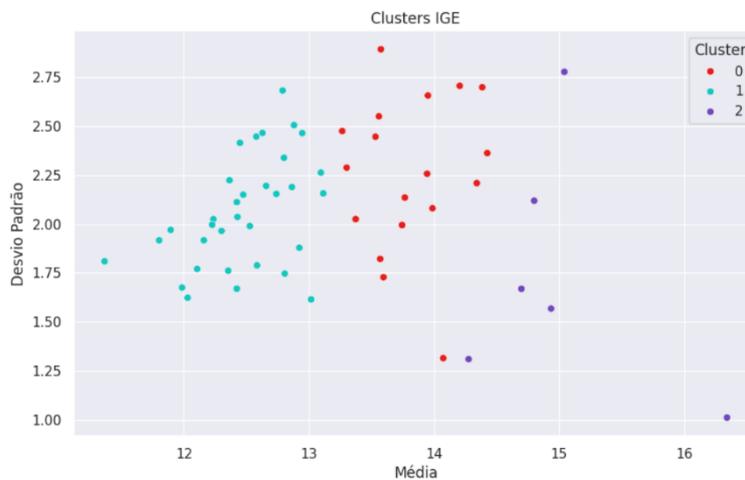


Figura 29. Gráfico de Dispersão IGE.

Tabela 14. Distribuição de Unidades Curriculares e Áreas por Cluster IGE.

Cluster 0	Área	Cluster 1	Área	Cluster 2	Área
Conceção e Desenvolvimento de Sistemas de Informação (Pré Bolonha)	Bases de Dados, Modelação	Álgebra	Matemática	Gestão de Sistemas de Informação	Sistemas de Informação Gerais
Empreendedorismo I - Introdução ao Empreendedorismo e Oportunidades de Negócio	Gestão Geral e Operações	Algoritmos e Estruturas de Dados	Programação	Gestão Integrada de Recursos Humanos	Recursos Humanos
Engenharia de Software I	Programação	Análise de Dados I	Estatística	Planeamento de Projetos Utilizando Ferramentas Informáticas (Msproject)	Informática Aplicada
Engenharia de Software II	Programação	Análise de Dados II	Estatística	Projeto Empresa Digital	Sistemas de Informação Gerais
Estratégia Empresarial	Gestão Estratégica	Análise Matemática	Matemática	Sistemas de Informação Distribuídos	Bases de Dados, Modelação
Estratégia Organizacional	Gestão Estratégica	Arquitetura de Computadores	Arquitetura de Computadores	Sistemas Informáticos de Apoio à Decisão II	Business Intelligence
Fluxo de Informação em Sistemas	Telecomunicações	Arquitetura de Computadores (Pré Bolonha)	Arquitetura de Computadores		
Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	Bases de Engenharia I (Pré Bolonha L0744)	Matemática		
Fundamentos de Contabilidade de Gestão	Contabilidade e Finanças	Complementos de Contabilidade Financeira	Contabilidade e Finanças		
Gestão Financeira I	Contabilidade e Finanças	Conceção e Desenvolvimento de Sistemas de Informação	Bases de Dados, Modelação		
Inteligência Artificial	Inteligência Artificial	Contabilidade de Gestão	Contabilidade e Finanças		
Introdução à Programação	Programação	Contabilidade Financeira I	Contabilidade e Finanças		
Macroeconomia	Economia	Contabilidade Financeira II	Contabilidade e Finanças		
Multimédia e Computação Gráfica	Multimédia	Contabilidade Financeira II (Pré Bolonha)	Contabilidade e Finanças		
Sistemas de Controlo de Gestão	Gestão Geral e Operações	Controlo de Gestão	Gestão Geral e Operações		
Sistemas de Informação em Estruturas Organizacionais	Sistemas de Informação Gerais	Economia I	Economia		
Sistemas Informáticos de Apoio à Decisão I	Business Intelligence	Economia II	Economia		
Tecnologias para Sistemas Inteligentes	Inteligência Artificial	Fundamentos de Bases de Dados	Bases de Dados, Modelação		
		Fundamentos de Contabilidade Financeira	Contabilidade e Finanças		
		Gestão de Operações	Gestão Geral e Operações		
		Gestão de Projetos de Tecnologia e Sistemas de Informação	Sistemas de Informação Gerais		
		Gestão de Recursos Humanos	Recursos Humanos		
		Gestão do Marketing	Marketing		
		Gestão Financeira II	Contabilidade e Finanças		
		Introdução à Informática (Pré-Bolonha)	Informática Aplicada		
		Marketing para as Tecnologias	Marketing		
		Microeconomia	Economia		
		Programação Concorrente e Distribuída	Programação		

		Programação Orientada para Objetos	Programação		
		Redes Digitais I - Fundamentos	Redes		
		Redes Digitais II - Sistemas, Aplicações e Serviços	Redes		
		Redes Digitais III - Segurança, Multimédia e Gestão	Redes		
		Sistemas Operativos	Sistemas Operativos		

b) Análise por Área

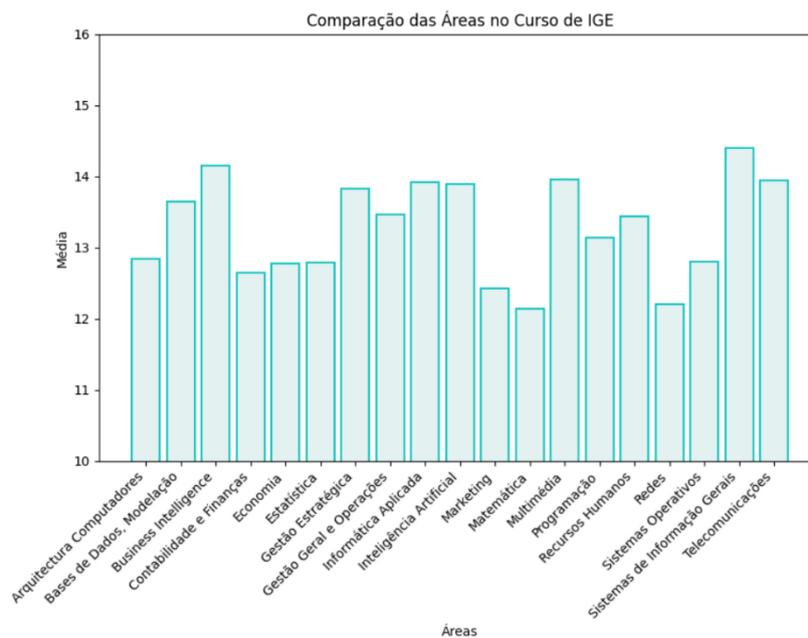


Figura 30. Médias das Áreas IGE.

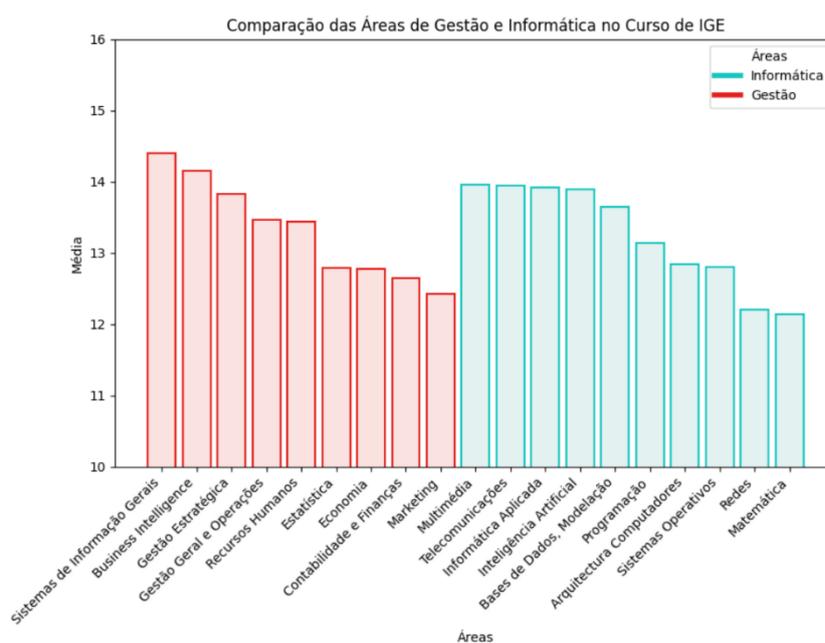


Figura 31. Comparação das Médias de Gestão e Informática IGE.

c) Áreas Mais Difíceis

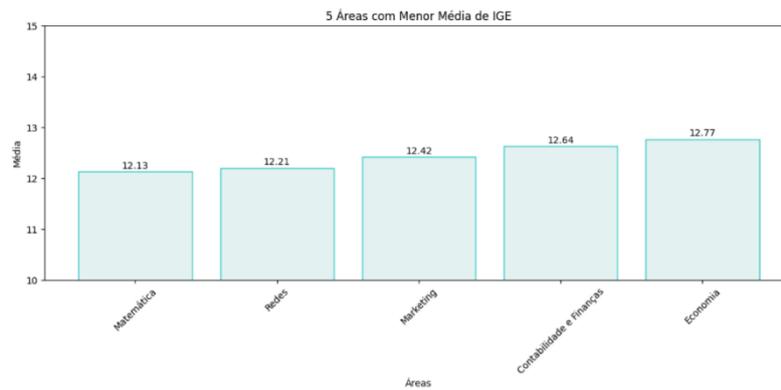


Figura 32. Áreas com Menor Média IGE.

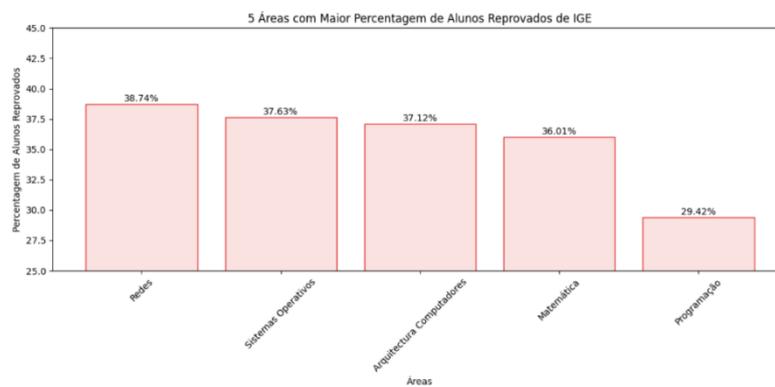


Figura 33. Áreas com Maior Percentagem de Alunos Reprovados IGE.

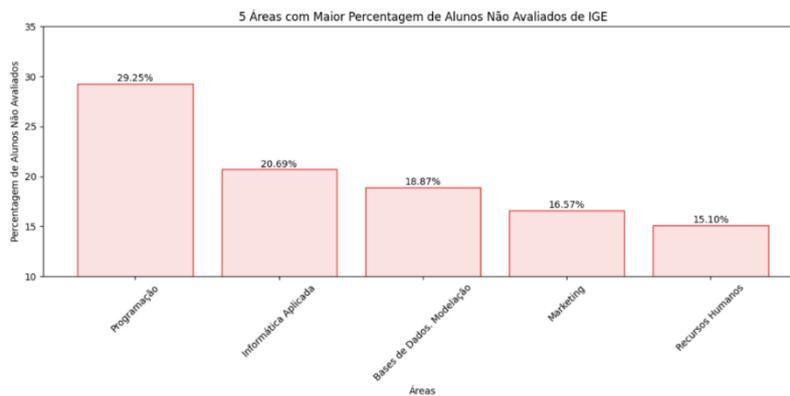


Figura 34. Áreas com Maior Percentagem de Alunos Não Avaliados IGE.

Informática e Gestão de Empresas Pós-Laboral

a) Scree Plot e Clustering

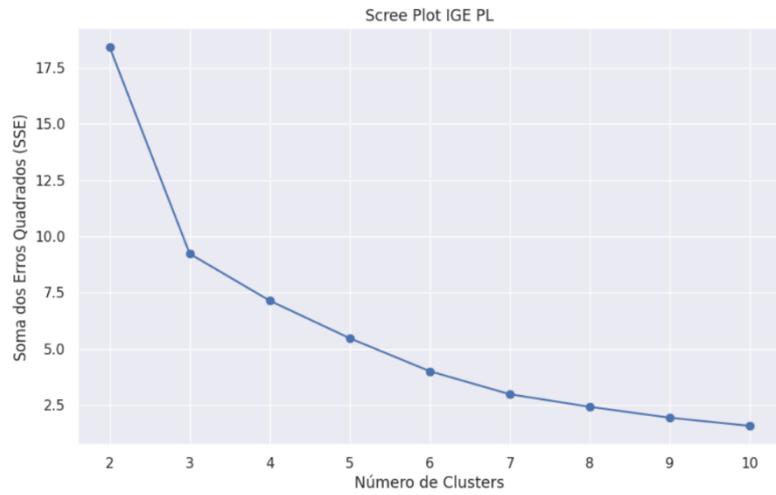


Figura 35. Scree Plot IGE PL.

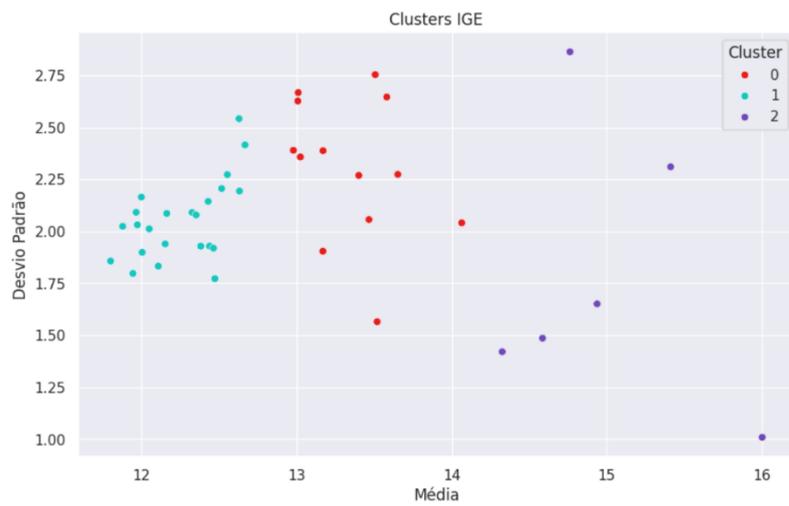


Figura 36. Gráfico de Dispersão IGE PL.

Tabela 15. Distribuição de Unidades Curriculares e Áreas por Cluster IGE PL.

Cluster 0	Área	Cluster 1	Área	Cluster 2	Área
Análise de Dados II	Estatística	Álgebra	Matemática	Estratégia Organizacional	Gestão Estratégica
Controlo de Gestão	Gestão Geral e Operações	Algoritmos e Estruturas de Dados	Programação	Gestão de Sistemas de Informação	Sistemas de Informação Gerais
Engenharia de Software I	Programação	Análise de Dados I	Estatística	Planeamento de Projetos Utilizando Ferramentas Informáticas (Msproject)	Informática Aplicada
Engenharia de Software II	Programação	Análise Matemática	Matemática	Projeto Empresa Digital	Sistemas de Informação Gerais
Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	Arquitetura de Computadores	Arquitetura de Computadores	Sistemas de Informação Distribuídos	Bases de Dados, Modelação
Fundamentos de Contabilidade de Gestão	Contabilidade e Finanças	Complementos de Contabilidade Financeira	Contabilidade e Finanças	Sistemas Informáticos de Apoio à Decisão II	Business Intelligence
Gestão Financeira I	Contabilidade e Finanças	Conceção e Desenvolvimento de Sistemas de Informação	Bases de Dados, Modelação		
Inteligência Artificial	Inteligência Artificial	Fundamentos de Bases de Dados	Bases de Dados, Modelação		
Introdução à Programação	Programação	Fundamentos de Contabilidade Financeira	Contabilidade e Finanças		
Multimédia e Computação Gráfica	Multimédia	Gestão de Operações	Gestão Geral e Operações		
Sistemas de Informação em Estruturas Organizacionais	Sistemas de Informação Gerais	Gestão de Projetos de Tecnologia e Sistemas de Informação	Sistemas de Informação Gerais		
Sistemas Informáticos de Apoio à Decisão I	Business Intelligence	Gestão de Recursos Humanos	Recursos Humanos		
Tecnologias para Sistemas Inteligentes	Inteligência Artificial	Gestão do Marketing	Marketing		
		Gestão Financeira II	Contabilidade e Finanças		
		Macroeconomia	Economia		
		Marketing para as Tecnologias	Marketing		
		Microeconomia	Economia		
		Programação Concorrente e Distribuída	Programação		
		Programação Orientada para Objetos	Programação		
		Redes Digitais I - Fundamentos	Redes		
		Redes Digitais II - Sistemas, Aplicações e Serviços	Redes		
		Redes Digitais III - Segurança, Multimédia e Gestão	Redes		
		Sistemas Operativos	Sistemas Operativos		

b) Análise por Área

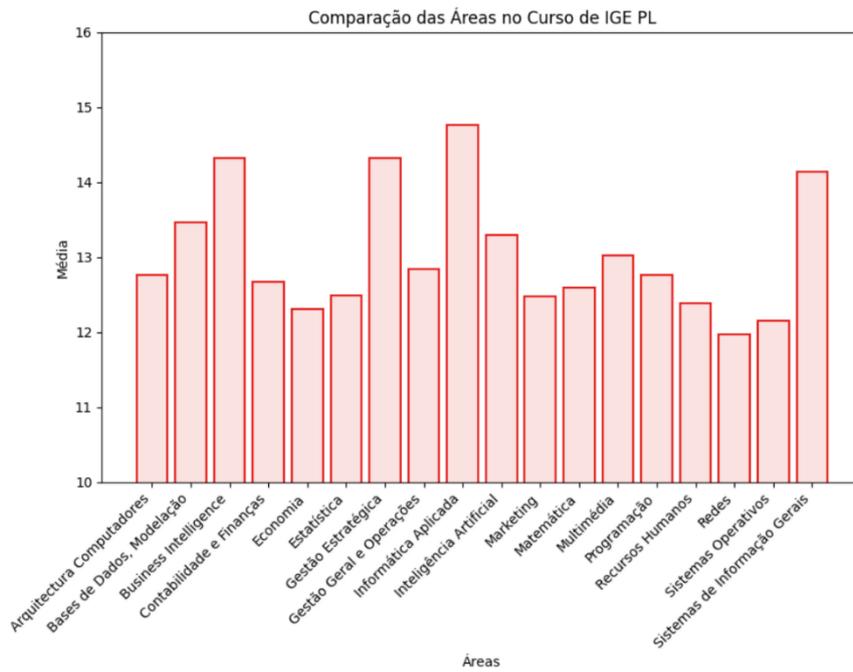


Figura 37. Médias das Áreas IGE PL.

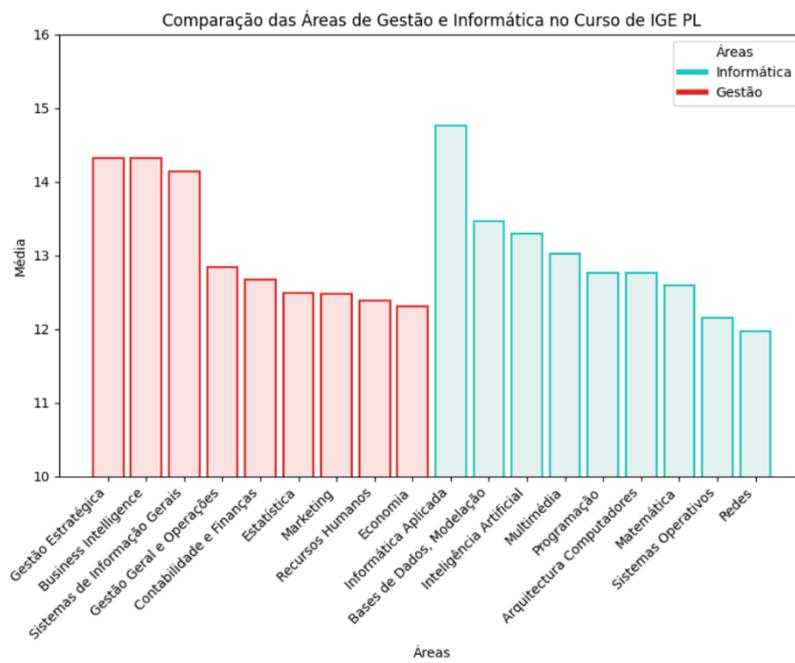


Figura 38. Comparação das Médias de Gestão e Informática IGE PL.

c) Áreas Mais Difíceis

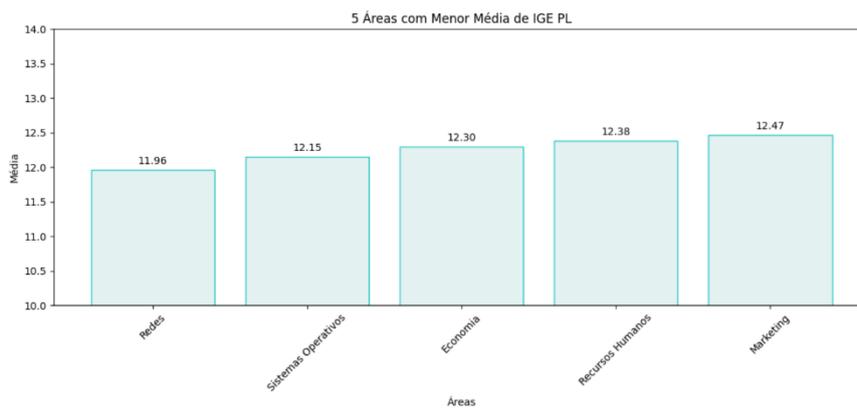


Figura 39. Áreas com Menor Média IGE PL.

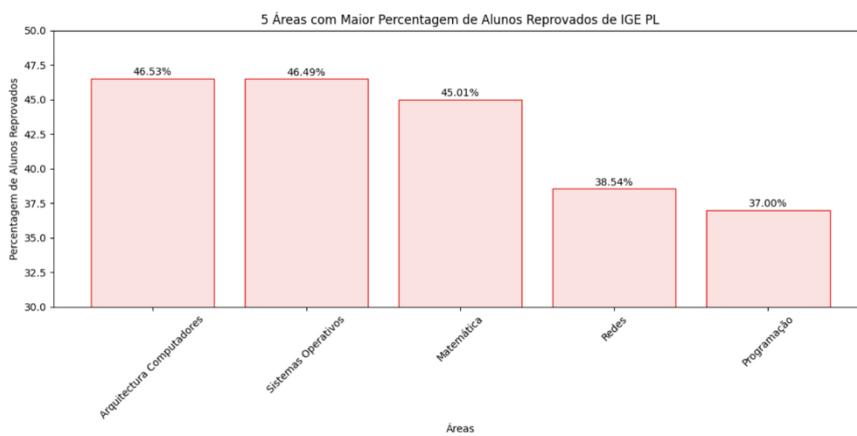


Figura 40. Áreas com Maior Percentagem de Alunos Reprovados IGE PL.

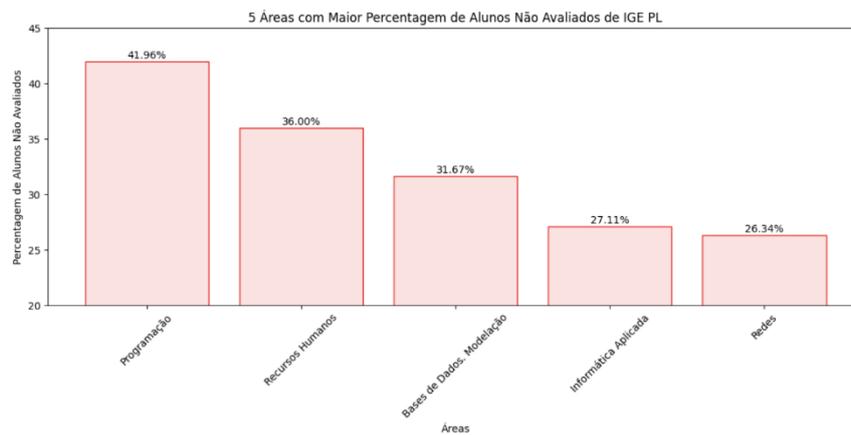


Figura 41. Áreas com Maior Percentagem de Alunos Não Avaliados IGE PL.

Análise Geral de Programação

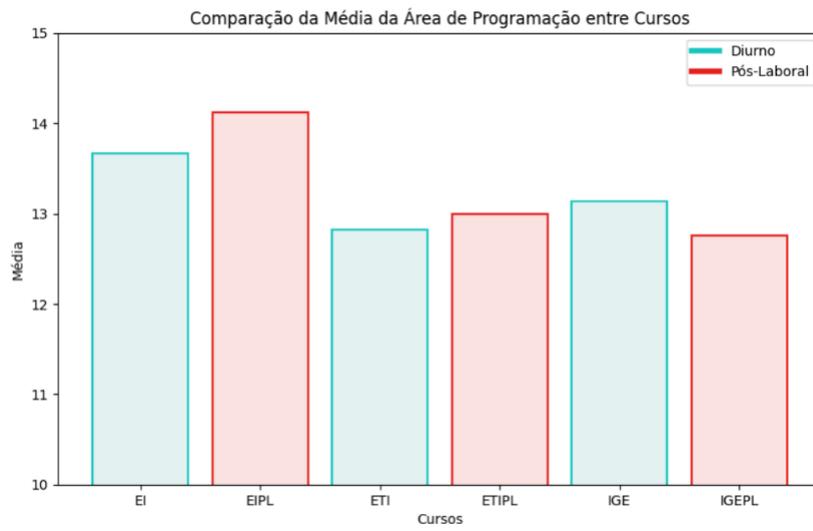


Figura 42. Comparação das Médias de Programação nos Diferentes Cursos

Anexo D: Análise de Dados da Questão 2

Engenharia Informática

Tabela 16. Pares Relevantes, Áreas e Valores de Diferença EI.

UC Dependente	Área	UC Independente	Área	Valores de Diferença
Circuitos para Comunicações	Telecomunicações	Programação Concorrente e Distribuída	Programação	2,17
Engenharia de Software II	Programação	Microprocessadores	Arquitetura de Computadores	2,26
Engenharia de Software II	Programação	Redes Digitais I - Fundamentos	Redes	2,05
Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	Introdução à Programação	Programação	2,39
Inteligência Artificial	Inteligência Artificial	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	2,11
Inteligência Artificial	Inteligência Artificial	Introdução à Programação	Programação	2,00
Inteligência Artificial	Inteligência Artificial	Programação Concorrente e Distribuída	Programação	2,38
Inteligência Artificial	Inteligência Artificial	Programação Orientada para Objetos	Programação	2,33
Inteligência Artificial	Inteligência Artificial	Redes Digitais I - Fundamentos	Redes	2,86
Inteligência Artificial	Inteligência Artificial	Sistemas Operativos	Sistemas Operativos	2,15
Mecânica e Eletricidade	Física e Eletromagnetismo	Álgebra Linear, Geometria Analítica e Análise Vetorial	Matemática	2,54
Mecânica e Eletricidade	Física e Eletromagnetismo	Análise Matemática I	Matemática	2,16
Mecânica e Eletricidade	Física e Eletromagnetismo	Fundamentos de Bases de Dados	Bases de Dados, Modelação	2,70
Mecânica e Eletricidade	Física e Eletromagnetismo	Introdução à Programação	Programação	2,52
Multimédia e Computação Gráfica	Multimédia	Engenharia de Software I	Programação	2,04
Processamento de Informação II	Telecomunicações	Redes Digitais I - Fundamentos	Redes	2,02
Programação em Rede	Programação	Fundamentos de Bases de Dados	Bases de Dados, Modelação	2,08
Programação em Rede	Programação	Sistemas Operativos	Sistemas Operativos	2,17
Sistemas Operativos	Sistemas Operativos	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	2,37
Sistemas Operativos	Sistemas Operativos	Fundamentos de Bases de Dados	Bases de Dados, Modelação	2,09
Sistemas Operativos	Sistemas Operativos	Programação Concorrente e Distribuída	Programação	2,67
Sistemas Operativos	Sistemas Operativos	Programação Orientada para Objetos	Programação	2,32
Tecnologias para Sistemas Inteligentes	Inteligência Artificial	Redes Digitais I - Fundamentos	Redes	2,50
Tecnologias para Sistemas Inteligentes	Inteligência Artificial	Redes Digitais III - Segurança, Multimédia e Gestão	Redes	2,21
Programação Concorrente e Distribuída	Programação	Introdução à Programação	Programação	1,21
Programação Concorrente e Distribuída	Programação	Programação Orientada para Objetos	Programação	1,51
Programação Orientada para Objetos	Programação	Introdução à Programação	Programação	1,71
Algoritmos e Estruturas de Dados	Programação	Introdução à Programação	Programação	0,52
Programação Orientada para Objetos	Programação	Algoritmos e Estruturas de Dados	Programação	1,48
Inteligência Artificial	Inteligência Artificial	Programação Concorrente e Distribuída	Programação	0,43

Engenharia Informática Pós-Laboral

Tabela 17. Pares Relevantes, Áreas e Valores de Diferença EI PL.

UC Dependente	Área	UC Independente	Área	Valores de Diferença
Sistemas Operativos	Sistemas Operativos	Programação Orientada para Objetos	Programação	2,24
Teoria da Computação	Programação	Algoritmos e Estruturas de Dados	Programação	2,09
Teoria da Computação	Programação	Circuitos para Comunicações	Telecomunicações	2,31
Teoria da Computação	Programação	Sistemas Operativos	Sistemas Operativos	2,42
Programação Concorrente e Distribuída	Programação	Introdução à Programação	Programação	1,68
Programação Concorrente e Distribuída	Programação	Programação Orientada para Objetos	Programação	0,91
Programação Orientada para Objetos	Programação	Introdução à Programação	Programação	1,81
Algoritmos e Estruturas de Dados	Programação	Introdução à Programação	Programação	0,52
Programação Orientada para Objetos	Programação	Algoritmos e Estruturas de Dados	Programação	1,52
Inteligência Artificial	Inteligência Artificial	Programação Concorrente e Distribuída	Programação	1,27

Engenharia de Telecomunicações e Informática

Tabela 18. Pares Relevantes, Áreas e Valores de Diferença ETI.

UC Dependente	Área	UC Independente	Área	Valores de Diferença
Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	Eletromagnetismo	Física e Eletromagnetismo	2,38
Inteligência Artificial	Inteligência Artificial	Eletromagnetismo	Física e Eletromagnetismo	2,36
Inteligência Artificial	Inteligência Artificial	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	2,66
Inteligência Artificial	Inteligência Artificial	Introdução à Programação	Programação	2,03
Inteligência Artificial	Inteligência Artificial	Modulação e Codificação	Telecomunicações	2,00
Inteligência Artificial	Inteligência Artificial	Programação Concorrente e Distribuída	Programação	2,33
Inteligência Artificial	Inteligência Artificial	Propagação e Radiação de Ondas Eletromagnéticas	Física e Eletromagnetismo	2,17
Inteligência Artificial	Inteligência Artificial	Redes Digitais I - Fundamentos	Redes	2,15
Mecânica e Eletricidade	Inteligência Artificial	Teoria dos Circuitos	Arquitetura de Computadores	2,22
Ondas e Ótica	Inteligência Artificial	Álgebra Linear, Geometria Analítica e Análise Vetorial	Matemática	2,19
Ondas e Ótica	Física e Eletromagnetismo	Análise Matemática I	Matemática	2,19
Ondas e Ótica	Física e Eletromagnetismo	Eletromagnetismo	Física e Eletromagnetismo	2,59
Ondas e Ótica	Física e Eletromagnetismo	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	2,01
Ondas e Ótica	Física e Eletromagnetismo	Sistemas Operativos	Sistemas Operativos	2,08
Ondas e Ótica	Multimédia	Teoria dos Circuitos	Arquitetura de Computadores	2,49
Programação Concorrente e Distribuída	Programação	Introdução à Programação	Programação	1,19
Programação Concorrente e Distribuída	Programação	Programação Orientada para Objetos	Programação	0,74
Programação Orientada para Objetos	Programação	Introdução à Programação	Programação	1,54
Redes Digitais II - Sistemas, Aplicações e Serviços	Redes	Redes Digitais I - Fundamentos	Redes	1,22
Redes Digitais III - Segurança, Multimédia e Gestão	Redes	Redes Digitais II - Sistemas, Aplicações e Serviços	Redes	0,28
Modelação e Codificação	Telecomunicações	Teoria do Sinal	Telecomunicações	1,35
Sistemas de Telecomunicações Guiados	Telecomunicações	Modelação e Codificação	Telecomunicações	0,43
Fundamentos de Eletrónica	Física e Eletromagnetismo	Teoria dos Circuitos	Arquitetura Computadores	0,72
Eletrónica Programada e Processamento Digital de Sinais	Informática Aplicada	Fundamentos de Eletrónica	Física e Eletromagnetismo	0,94
Propagação e Radiação de Ondas Eletromagnéticas	Física e Eletromagnetismo	Eletromagnetismo	Física e Eletromagnetismo	0,82

Engenharia de Telecomunicações e Informática Pós-Laboral

Tabela 19. Pares Relevantes, Áreas e Valores de Diferença ETI PL.

UC Dependente	Área	UC Independente	Área	Valores de Diferença
Análise Matemática II	Matemática	Álgebra Linear, Geometria Analítica e Análise Vetorial	Matemática	2,51
Análise Matemática II	Matemática	Teoria dos Circuitos	Arquitetura de Computadores	2,52
Mecânica e Eletricidade	Física e Eletromagnetismo	Álgebra Linear, Geometria Analítica e Análise Vetorial	Matemática	2,29
Ondas e Ótica	Telecomunicações	Álgebra Linear, Geometria Analítica e Análise Vetorial	Matemática	2,15
Sistemas Operativos	Sistemas Operativos	Programação Orientada para Objetos	Programação	2,21
Programação Concorrente e Distribuída	Programação	Introdução à Programação	Programação	1,23
Programação Concorrente e Distribuída	Programação	Programação Orientada para Objetos	Programação	1,24
Programação Orientada para Objetos	Programação	Introdução à Programação	Programação	2,06
Redes Digitais II - Sistemas, Aplicações e Serviços	Redes	Redes Digitais I - Fundamentos	Redes	1,40
Redes Digitais III - Segurança, Multimédia e Gestão	Redes	Redes Digitais II - Sistemas, Aplicações e Serviços	Redes	1,19
Modelação e Codificação	Telecomunicações	Teoria do Sinal	Telecomunicações	0,87
Sistemas de Telecomunicações Guiados	Telecomunicações	Modelação e Codificação	Telecomunicações	0,56
Fundamentos de Eletrónica	Física e Eletromagnetismo	Teoria dos Circuitos	Arquitetura Computadores	1,29
Eletrónica Programada e Processamento Digital de Sinais	Informática Aplicada	Fundamentos de Eletrónica	Física e Eletromagnetismo	0,70
Propagação e Radiação de Ondas Eletromagnéticas	Física e Eletromagnetismo	Eletromagnetismo	Física e Eletromagnetismo	0,16

Informática e Gestão de Empresas

Tabela 20. Pares Relevantes, Áreas e Valores de Diferença IGE.

UC Dependente	Área	UC Independente	Área	Valores de Diferença
Conceção e Desenvolvimento de Sistemas de Informação	Bases de Dados, Modelação	Arquitetura de Computadores	Arquitetura de Computadores	2,09
Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	Análise Matemática	Matemática	2,42
Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	Fundamentos de Bases de Dados	Bases de Dados, Modelação	2,97
Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	Introdução à Programação	Programação	2,41
Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	Programação Concorrente e Distribuída	Programação	2,72
Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	Sistemas Operativos	Sistemas Operativos	2,44
Gestão Financeira I	Contabilidade e Finanças	Contabilidade Financeira II	Contabilidade e Finanças	2,11
Gestão Financeira II	Contabilidade e Finanças	Contabilidade Financeira II	Contabilidade e Finanças	2,37
Inteligência Artificial	Inteligência Artificial	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	2,47
Inteligência Artificial	Inteligência Artificial	Introdução à Programação	Programação	2,07
Sistemas de Informação Distribuídos	Bases de Dados, Modelação	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	2,08
Programação Concorrente e Distribuída	Programação	Introdução à Programação	Programação	1,23
Programação Concorrente e Distribuída	Programação	Programação Orientada para Objetos	Programação	1,43
Programação Orientada para Objetos	Programação	Introdução à Programação	Programação	0,95
Redes Digitais II - Sistemas, Aplicações e Serviços	Redes	Redes Digitais I - Fundamentos	Redes	1,12
Redes Digitais III - Segurança, Multimédia e Gestão	Redes	Redes Digitais II - Sistemas, Aplicações e Serviços	Redes	1,11
Programação Concorrente e Distribuída	Programação	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	1,06
Redes Digitais I - Fundamentos	Redes	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	1,25
Sistemas Operativos	Sistemas Operativos	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	1,69
Sistemas Operativos	Sistemas Operativos	Introdução à Programação	Programação	1,73
Engenharia de Software I	Programação	Introdução à Programação	Programação	0,94
Fundamentos de Contabilidade de Gestão	Contabilidade e Finanças	Contabilidade Financeira I	Contabilidade e Finanças	2,29
Gestão de Operações	Gestão Geral e Operações	Contabilidade Financeira I	Contabilidade e Finanças	0,71
Inteligência Artificial	Inteligência Artificial	Fundamentos de Bases de Dados	Bases de Dados, Modelação	1,12

Informática e Gestão de Empresas Pós-Laboral

Tabela 21. Pares Relevantes, Áreas e Valores de Diferença IGE PL.

UC Dependente	Área	UC Independente	Área	Valores de Diferença
Programação Orientada para Objetos	Programação	Introdução à Programação	Programação	2,13
Redes Digitais II - Sistemas, Aplicações e Serviços	Redes	Programação Concorrente e Distribuída	Programação	2,19
Programação Concorrente e Distribuída	Programação	Introdução à Programação	Programação	1,27
Programação Concorrente e Distribuída	Programação	Programação Orientada para Objetos	Programação	0,75
Redes Digitais II - Sistemas, Aplicações e Serviços	Redes	Redes Digitais I - Fundamentos	Redes	0,49
Redes Digitais III - Segurança, Multimédia e Gestão	Redes	Redes Digitais II - Sistemas, Aplicações e Serviços	Redes	0,58
Programação Concorrente e Distribuída	Programação	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	0,81
Redes Digitais I - Fundamentos	Redes	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	0,82
Sistemas Operativos	Sistemas Operativos	Fundamentos de Arquitetura de Computadores	Arquitetura de Computadores	0,75
Sistemas Operativos	Sistemas Operativos	Introdução à Programação	Programação	1,73
Engenharia de Software I	Programação	Introdução à Programação	Programação	0,97
Inteligência Artificial	Inteligência Artificial	Fundamentos de Bases de Dados	Bases de Dados, Modelação	0,96

Anexo E: Análise de Dados da Questão 3

Engenharia Informática

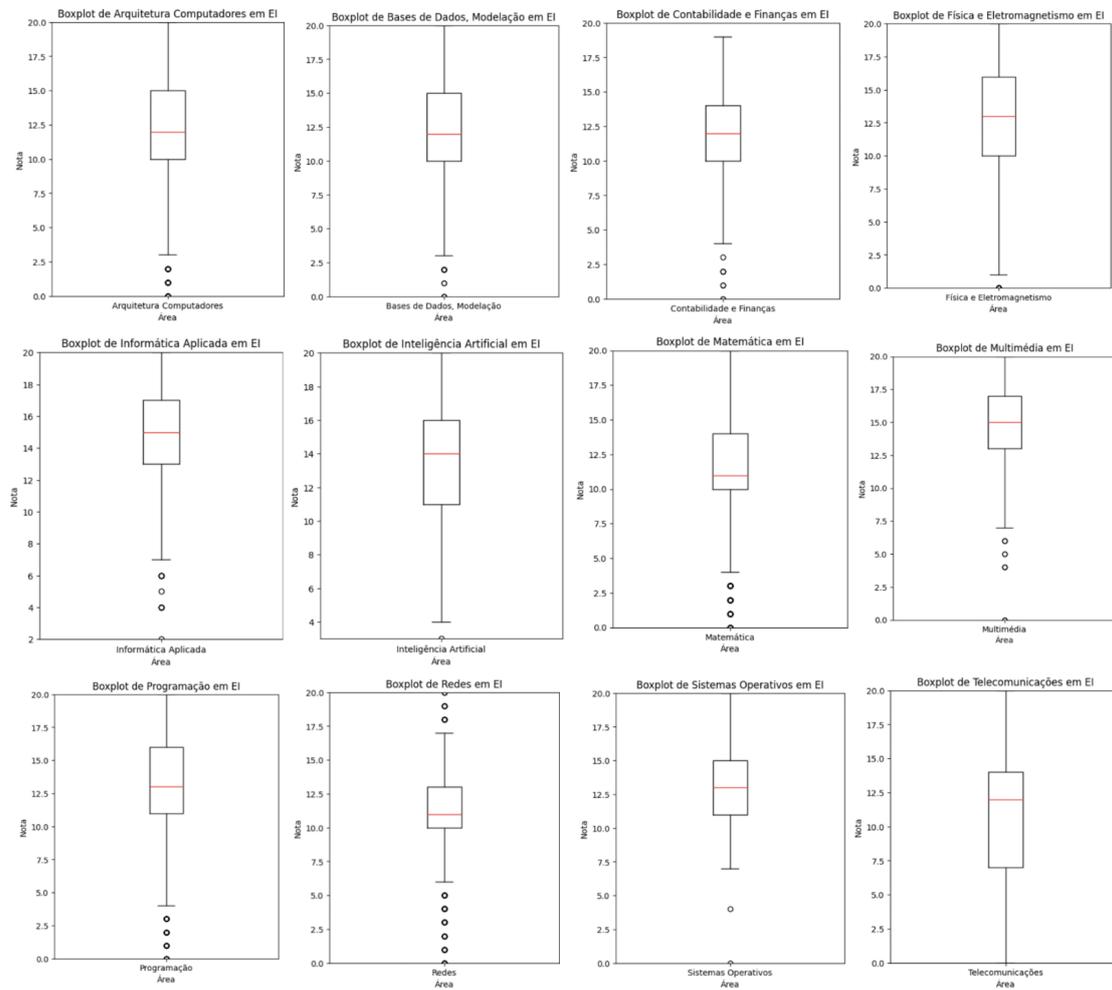


Figura 43. Boxplots por Área EI.

Engenharia Informática Pós-Laboral

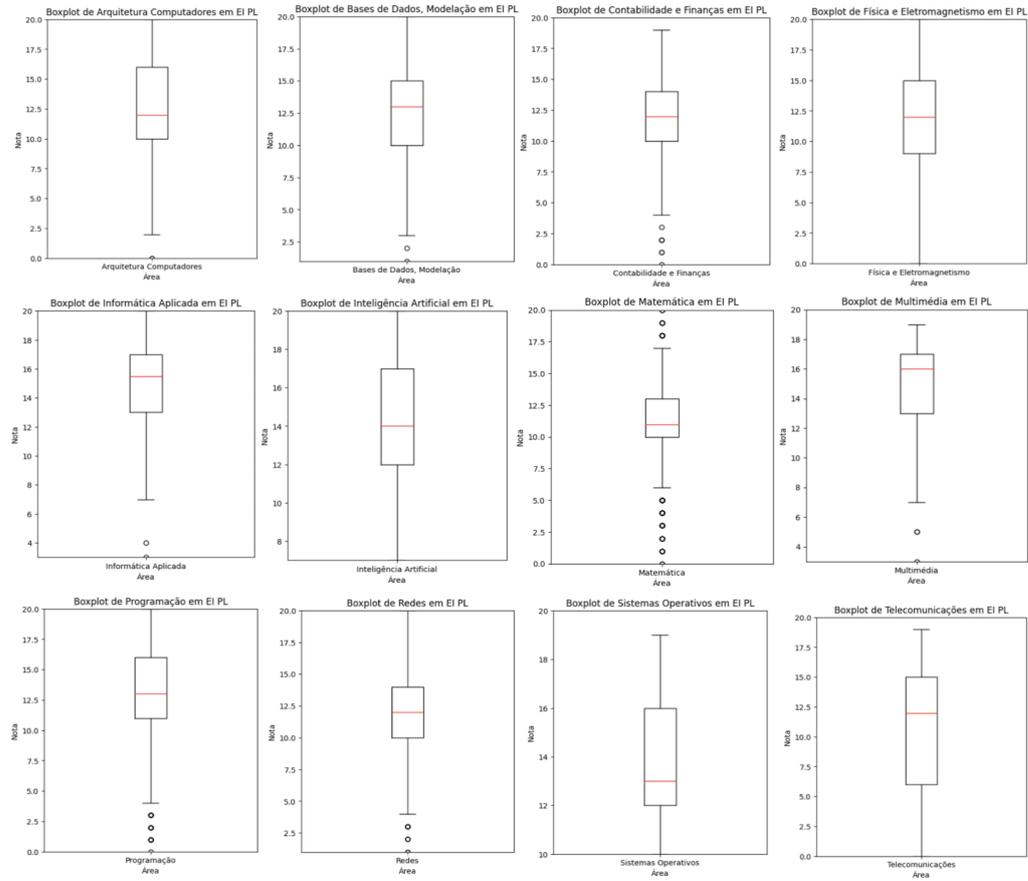


Figura 44. Boxplots por Área EI PL.

Engenharia de Telecomunicações e Informática

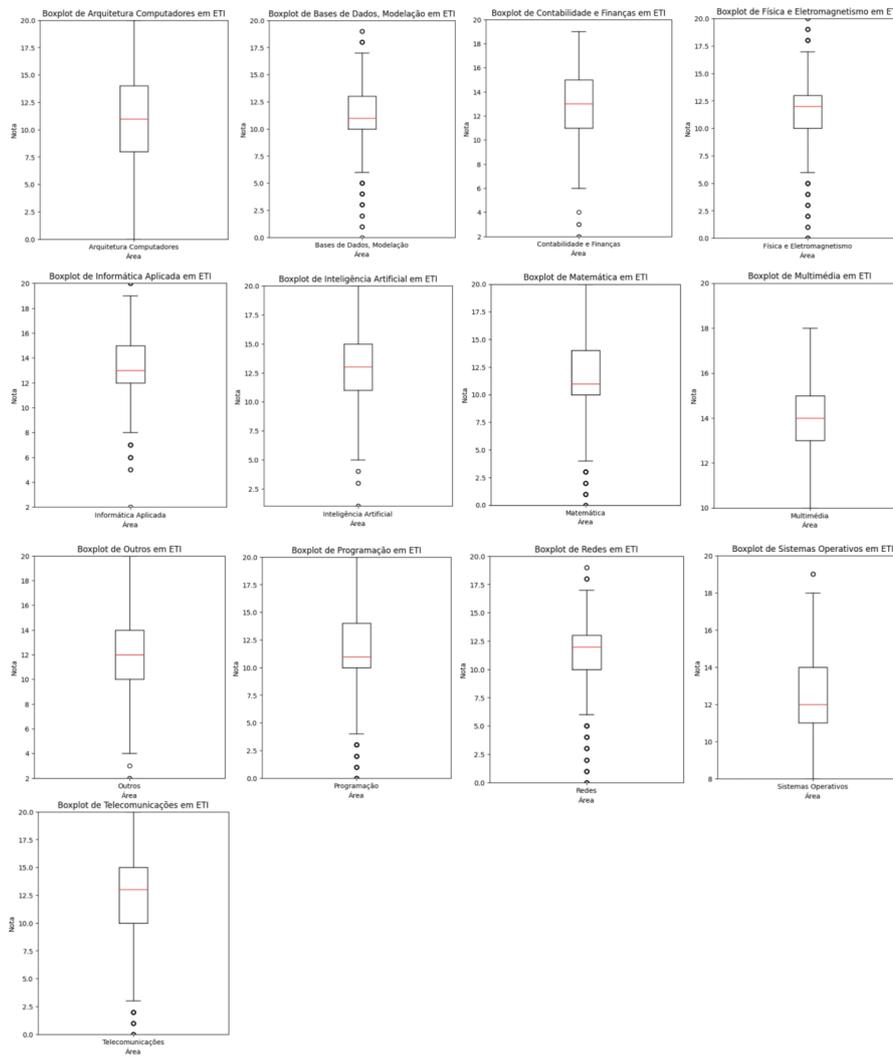


Figura 45. Boxplots por Área ETI.

Engenharia de Telecomunicações e Informática Pós-Laboral

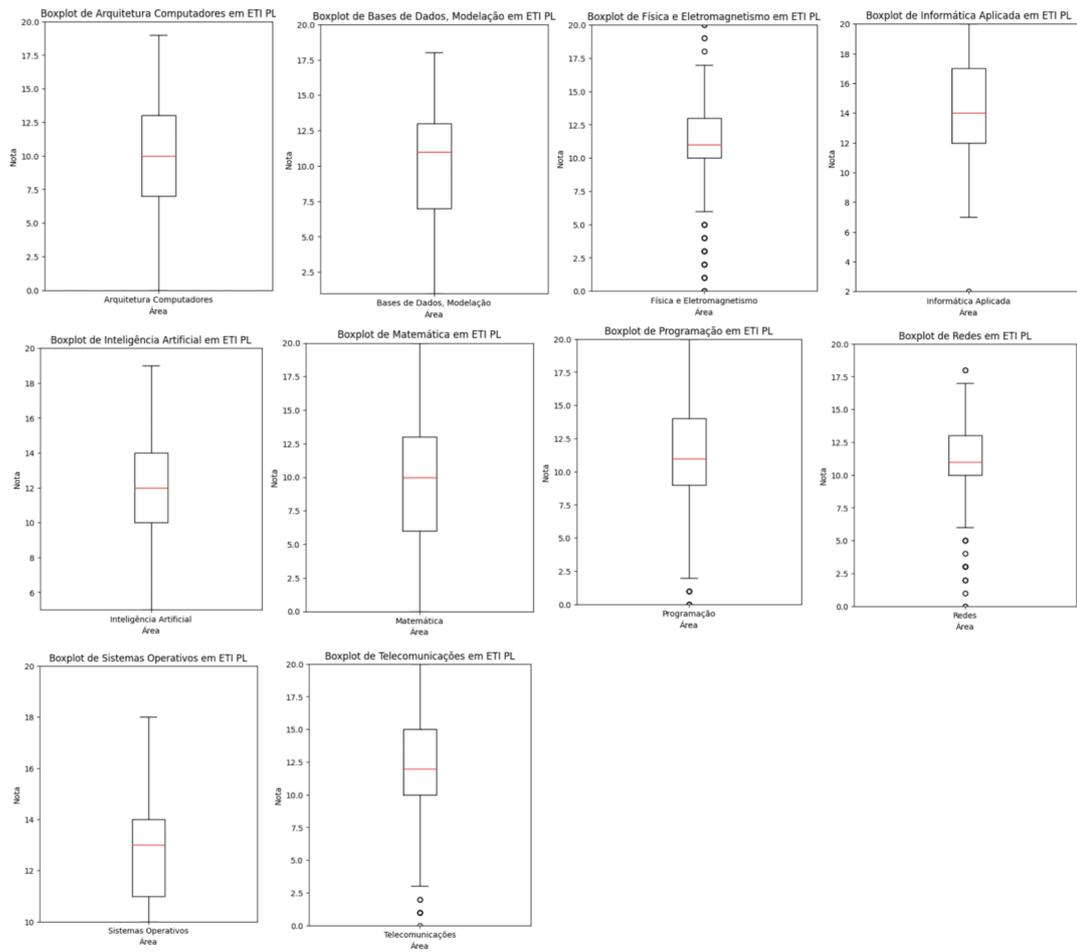


Figura 46. Boxplots por Área ETI PL.

Informática e Gestão de Empresas

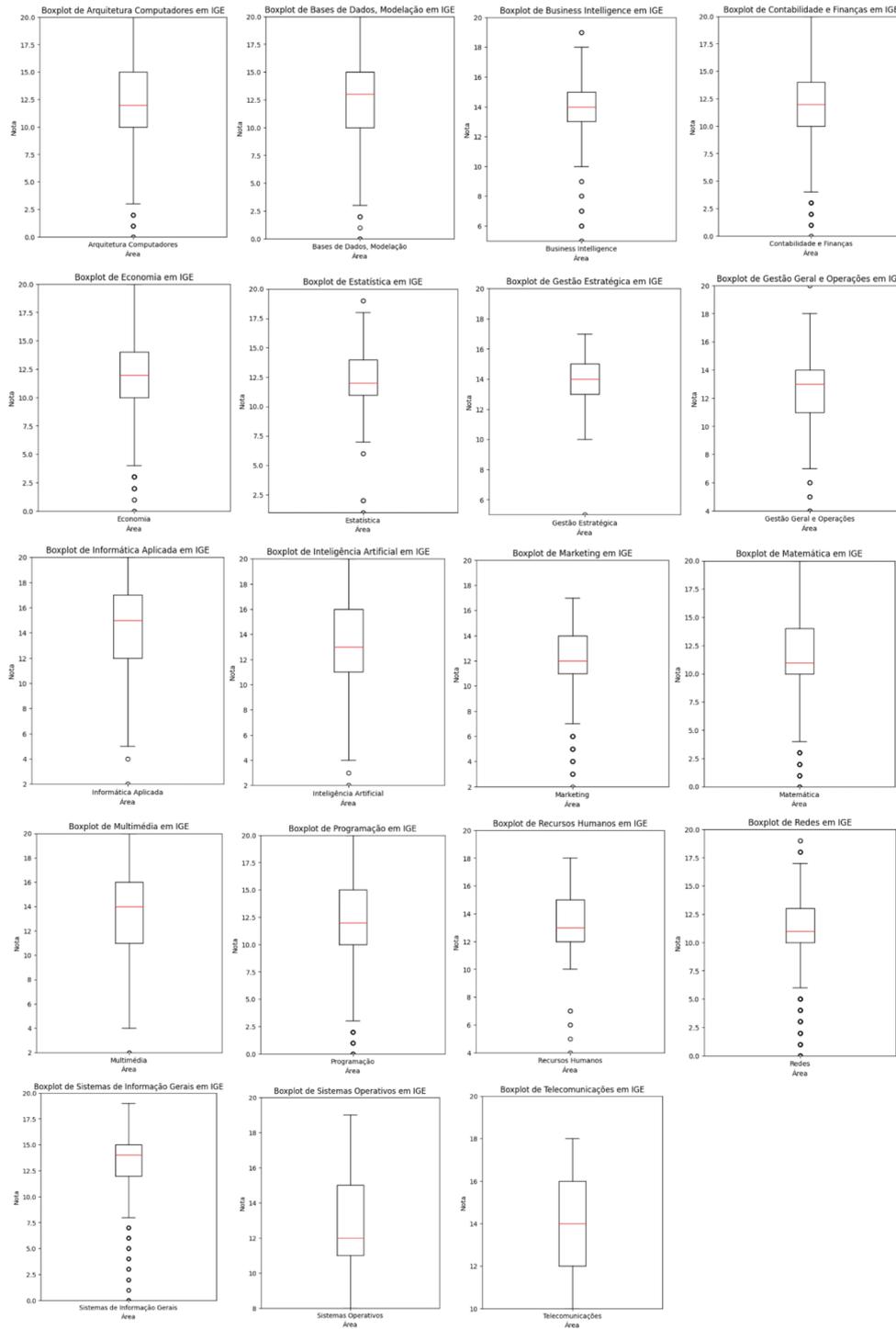


Figura 47. Boxplots por Área IGE.

Informática e Gestão de Empresas Pós-Laboral

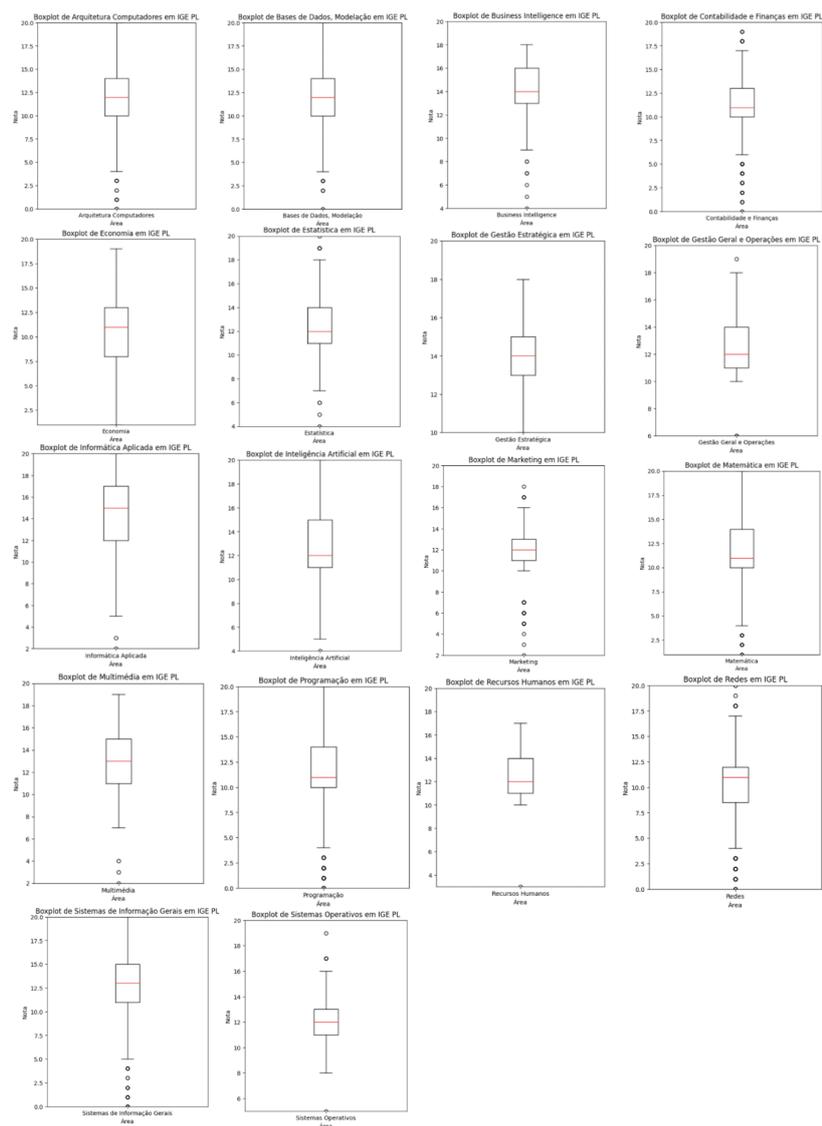


Figura 48. Boxplots por Área IGE PL.

Anexo F: Análise de Dados da Questão 4

Engenharia Informática

Tabela 22. Análise Estatística dos 10 Alunos com Maior Média EI.

Aluno	Média	Desvio Padrão	Nota Mínima	Mediana	Nota Máxima
1	19,167	0,289	19	19	19,2
2	18,467	0,321	18,1	18,6	18,7
3	18,367	0,568	17,9	18,2	19
4	18,3	0,265	18	18,4	18,5
5	18,267	0,611	17,6	18,4	18,8
6	17,933	0,416	17,6	17,8	18,4
7	17,567	0,702	16,9	17,5	18,3
8	17,567	0,208	17,4	17,5	17,8
9	17,5	0,265	17,35	17,4	17,8
10	17,4	0,265	17,2	17,3	17,7

Tabela 23. Média do Último Colocado em EI ao Longo dos Anos Letivos.

Ano Letivo	2013/2014	2014/2015	2015/2016	2016/2017	2017/2018	2018/2019	2019/2020
Média do Último Colocado	122,5	126,0	151,4	154,5	156,8	151,0	155,5

Engenharia Informática Pós-Laboral

Tabela 24. Análise Estatística dos 10 Alunos com Maior Média EI PL.

Aluno	Média	Desvio Padrão	Nota Mínima	Mediana	Nota Máxima
1	17,830	1,062	16,5	17,912	19
2	17,633	0,808	16,9	17,5	18,5
3	17,467	1,193	16,167	18,167	18,5
4	17,2	0,917	16,2	17,4	18
5	16,710	0,900	15,7	17	17,429
6	16,6	0,557	16,1	16,5	17,2
7	16,233	1,332	15,1	15,9	17,7
8	16,067	0,643	15,6	15,8	16,8
9	16,025	1,848	14,4	15,85	18
10	15,762	2,042	14	15,286	18

Tabela 25. Média do Último Colocado em EI PL ao Longo dos Anos Letivos.

Ano Letivo	2013/2014	2014/2015	2015/2016	2016/2017	2017/2018	2018/2019	2019/2020
Média do Último Colocado	109,5	102,0	127,3	135,1	144,8	133,0	136,5

Engenharia de Telecomunicações e Informática

Tabela 26. Análise Estatística dos 10 Alunos com Maior Média ETI.

Aluno	Média	Desvio Padrão	Nota Mínima	Mediana	Nota Máxima
1	17,978	0,430	17,5	18,217	18,333
2	17,859	0,051	17,8	17,889	17,889
3	16,985	0,494	16,7	17,128	17,556
4	16,7	0,781	16,2	16,95	17,6
5	16,630	1,353	15,2	17,344	17,889
6	16,6	0,529	16	16,9	17
7	16,507	0,621	16,1	16,711	17,222
8	16,5	0,7	15,7	16,9	17
9	16,283	1,418	15,25	16,8	17,9
10	16,267	1,007	15,2	16,8	17,2

Tabela 27. Média do Último Colocado em ETI ao Longo dos Anos Letivos.

Ano Letivo	2013/2014	2014/2015	2015/2016	2016/2017	2017/2018	2018/2019	2019/2020
Média do Último Colocado	116,5	117,0	125,3	135,4	140,8	126,5	128,5

Engenharia de Telecomunicações e Informática Pós-Laboral

Tabela 28. Análise Estatística dos 10 Alunos com Maior Média ETI PL.

Aluno	Média	Desvio Padrão	Nota Mínima	Mediana	Nota Máxima
1	17,356	0,336	17	17,4	17,667
2	17,167	0,802	16,4	17,1	18
3	15,967	1,501	15,1	15,1	17,7
4	15,940	1,573	14,125	16,778	16,917
5	15,663	0,807	14,889	15,6	16,5
6	15,536	0,503	15,25	15,304	16,286
7	14,852	1,223	13	15,155	16
8	14,806	3,745	11	14,861	18,5
9	14,722	2,810	11,5	16	16,667
10	14,512	2,315	11,75	14,5	19

Informática e Gestão de Empresas

Tabela 29. Análise Estatística dos 10 Alunos com Maior Média IGE.

Aluno	Média	Desvio Padrão	Nota Mínima	Mediana	Nota Máxima
1	17,575	1,164	15,9	17,95	18,5
2	17,169	0,513	16,7	17,175	17,625
3	16,7	0,716	16	16,55	17,7
4	16,595	1,077	15,5	16,541	17,8
5	16,554	0,715	15,5	16,857	17
6	16,396	0,680	15,9	16,141	17,4
7	16,375	1,382	14,8	16,35	18
8	16,225	0,096	16,1	16,25	16,3
9	16,2	1,219	15	15,95	17,9
10	16,125	0,472	15,7	16	16,8

Tabela 30. Média do Último Colocado em IGE ao Longo dos Anos Letivos.

Ano Letivo	2013/2014	2014/2015	2015/2016	2016/2017	2017/2018	2018/2019	2019/2020
Média do Último Colocado	124,5	128,0	147,1	149,6	152,2	144,5	151,5

Informática e Gestão de Empresas Pós-Laboral

Tabela 31. Análise Estatística dos 10 Alunos com Maior Média IGE PL.

Aluno	Média	Desvio Padrão	Nota Mínima	Mediana	Nota Máxima
1	18,867	0,163	18,667	18,9	19
2	16,73	0,803	15,8	16,6	17,8
3	16,575	0,299	16,3	16,5	17
4	16,224	0,791	15,429	16,233	17
5	16,075	1,372	15,1	15,55	18,1
6	15,869	1,357	14,4	15,788	17,5
7	15,756	0,395	15,2	15,85	16,125
8	14,862	1,157	13,625	14,811	16,2
9	14,553	0,553	13,833	14,627	15,125
10	14,395	1,336	12,4	14,990	15,2

Tabela 32. Média do Último Colocado em IGE PL ao Longo dos Anos Letivos.

Ano Letivo	2013/2014	2014/2015	2015/2016	2016/2017	2017/2018	2018/2019	2019/2020
Média do Último Colocado	105,0	106,5	134,1	137,4	143,9	134,0	142,0