



**Fábio Xavier Leite
Ferreira**

**FICAVIS: Visualização de Dados para Combate ao
Abandono Escolar na Universidade de Aveiro**



**Fábio Xavier Leite
Ferreira**

**FICAVIS: Visualização de Dados para Combate ao
Abandono Escolar na Universidade de Aveiro**

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Informática, realizada sob a orientação científica da Doutora Maria Beatriz Alves de Sousa Santos, Professora Associada com Agregação do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro, e do Doutor Paulo Miguel de Jesus Dias, Professor Auxiliar do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro.

o júri / the jury

presidente / president

Professor Doutor Joaquim João Estrela Ribeiro Silvestre Madeira
Professor Auxiliar, Universidade de Aveiro

vogais / examiners committee

Professora Doutora Ana Isabel Barreto Furtado Franco de Albuquerque
Veloso
Professora Auxiliar com Agregação, Universidade de Aveiro

Professora Doutora Maria Beatriz Alves de Sousa Santos
Professora Associada com Agregação, Universidade de Aveiro

agradecimentos / acknowledgements

Quero começar por agradecer à professora Beatriz Sousa Santos e ao professor Paulo Dias, orientadora e coorientador desta dissertação, por toda a disponibilidade e acompanhamento que sempre me prestaram ao longo deste período.

A todos os professores que mostraram interesse no meu trabalho e tiveram a gentileza de abdicar de um pouco das suas agendas ocupadas para participar nas entrevistas e testes realizados, nomeadamente ao professor Joaquim Madeira, ao professor António Neves, ao professor Augusto Silva e à professora Leonor Teixeira. De igual modo agradeço também à Carolina Costa por ter participado nestas sessões e pelas várias sugestões oferecidas.

Ao Bernardo Marques que mostrou enorme disponibilidade e me ajudou durante toda a fase inicial deste projeto, ao estar presente em todas as entrevistas que realizei.

À equipa dos sTic com quem partilhei a sala durante algumas semanas e me acolheu, em especial ao Telmo Rocha e à Dra. Loraine Nazaré, pelo apoio e dúvidas que me foram esclarecendo.

Agradeço ainda à Mariana Ferreira, à Filipa Tavares, ao Tiago Faria, ao Tiago Oliveira, ao Tiago Santos e ao Dinis Cruz pela contribuição na avaliação da solução que desenvolvi.

Por fim, e não menos importante, um especial agradecimento à minha família e amigos por sempre me apoiarem durante todo o meu percurso académico e acreditarem em mim.

A todos, o meu mais sincero obrigado.

Palavras Chave

visualização de informação, abandono do ensino superior, power bi, focus group, participatory design.

Resumo

O projeto FICA - Ferramentas de Identificação e Combate ao Abandono - criado pela Universidade de Aveiro com o objetivo de contribuir para a diminuição e prevenção do abandono e a melhoria do sucesso académico dos estudantes, gera uma grande quantidade de dados que podem fornecer informação relevante para diferentes utilizadores da Universidade, no entanto atualmente estes dados são-lhes fornecidos num formato difícil de analisar. Esta dissertação teve como objetivos fazer um estudo destes dados, dos seus utilizadores e contextos de utilização e propor uma solução que permita a sua exploração visual e interativa ajudando os utilizadores a obter as respostas necessárias para monitorizar o percurso dos estudantes e identificar indicadores de risco e fatores de insucesso que possam levar a situações mais críticas, como o abandono do ensino superior. Nesta dissertação, apresentam-se trabalhos e ferramentas de visualização relevantes e detalham-se as etapas da proposta de uma solução de exploração visual de dados do FICA adotando-se a metodologia do *participatory design*, desde a elicitação de requisitos, neste contexto realizada através de entrevistas do tipo *focus group* com futuros utilizadores, *design* de uma solução, protótipo desenvolvido em Power BI e avaliação da solução. Finalmente, apresentam-se conclusões e propostas de trabalho futuro.

Keywords

information visualization, higher education dropout, power bi, focus group, participatory design.

Abstract

The FICA project - Tools for Identifying and Combating Abandonment - created by the University of Aveiro with the aim of contributing to the reduction and prevention of abandonment and improving the academic success of students, generates a large amount of data that can provide relevant information for different University stakeholders, however these data are currently provided in a format difficult to analyze. This dissertation had as objectives to make a study of this data, of its users and contexts of use and to propose a solution that allows the visual and interactive exploration of this data and helps to analyze and obtain answers that the users may need to monitor the path of the students and identify risk indicators and failure factors that can lead to more critical situations, such as dropping out of higher education. In this dissertation, relevant works and visualization tools are presented and the stages of the proposal for a visual data exploration solution for data exploration adopting the methodology of participatory design, from the elicitation of requirements, in this context carried out through focus group interviews with target users, design of a solution, prototype developed in Power BI and evaluation of the solution. Finally, conclusions are presented as well as avenues for future work.

Índice

Índice	i
Lista de Figuras	iii
Lista de Tabelas	v
Glossário	vii
1 Introdução	1
1.1 Motivação	1
1.2 Objetivos	2
1.3 Estrutura	2
2 Visualização de Dados Acadêmicos	5
2.1 Exploração Visual de Dados	5
2.2 Aplicações de Exploração Visual de Dados Acadêmicos	7
2.3 Bibliotecas e Plataformas de Visualização	14
3 Visualização de Dados do Projeto FICA	21
3.1 Os Dados do FICA	21
3.2 Elicitação de Requisitos	26
3.2.1 <i>Focus Group</i>	26
3.2.2 Definição de <i>Personas</i>	30
4 Solução Proposta: FICAviz	35
4.1 Pré-processamento dos Dados	35
4.2 Desenvolvimento do Protótipo	37
4.2.1 Estrutura das Páginas	37
4.2.2 Perfil de Diretor de Curso	39
4.2.3 Perfil de Diretor de Departamento	48
4.2.4 Perfil de Assessor da Reitoria	51

4.2.5	Questões de Desenvolvimento	51
5	Avaliação	55
5.1	Avaliação Heurística	55
5.2	Testes de Usabilidade	58
5.2.1	Protocolo Experimental	58
5.2.2	Resultados e Principais Alterações	59
5.3	Discussão de Resultados	62
6	Conclusões e Trabalho Futuro	65
6.1	Conclusões	65
6.2	Trabalho Futuro	66
	Referências	67
	Apêndices	71
	Apêndice A - <i>Script</i> das Sessões de <i>Focus Group</i>	72
	Apêndice B - Resultados das Sessões de <i>Focus Group</i>	74
	Apêndice C - Avaliação Heurística da Primeira Versão da Solução Proposta	80
	Apêndice D - Procedimentos e Resultados das Avaliações Heurísticas com Utilizadores . . .	82
	Apêndice E - Análise de Resultados dos Testes de Usabilidade e Sugestões dos Utilizadores	95

Lista de Figuras

2.1	Fases do ciclo de desenvolvimento de uma aplicação de visualização segundo (Kirk, 2016)	6
2.2	Página principal do TECNICOVIS com todas as visualizações: seleção de uma instância de um curso (Martins et al., 2018)	8
2.3	Distribuição dos acessos aos vários recursos por parte dos estudantes entre Julho de 2013 e Fevereiro de 2014 na Faculdade Cenecista de Osório (Weiland & Manssour, 2015)	9
2.4	Alternativas de resposta para a <i>query</i> com o número de professores com mestrado no sistema desenvolvido pela Universidade de Washington (Desai et al., 2014)	10
2.5	Diferentes visualizações de interesse para o departamento de investigação geradas pelo SAPI (Cruz-Guzmán & Rosas, 2020)	11
2.6	Gráfico com a pontuação geral obtida por cada investigador gerado pelo SAPI (Cruz-Guzmán & Rosas, 2020)	12
2.7	Fluxo dos Estudantes Inscritos numa Faculdade da BME em 2012 (Horváth et al., 2018)	13
2.8	Situação atual dos alunos de uma faculdade da BME no final do ano letivo 2016/17 (Horváth et al., 2018)	13
2.9	Principais funcionalidades presentes numa solução desenvolvida em Power BI	16
3.2	Amostra de dados do FICA	22
3.3	Mapeamento visual das células correspondentes aos indicadores	24
3.4	Fontes dos dados utilizados pelo FICA	25
3.5	Protótipo apresentado relativamente à relação entre assiduidade e desempenho	30
3.6	Diferentes graus de acesso aos dados do FICA pelas várias <i>personas</i> definidas	31
4.1	Expressão DAX utilizada para a criação da tabela "Data"	36
4.2	Estrutura das páginas do relatório	37
4.3	<i>Tooltip</i> com informação sobre um elemento	38
4.4	Lista de alunos para uma determinada interação e seleção da ação <i>Pormenorizar</i>	39
4.5	Página Inicial do relatório para o perfil de Diretor de Curso	40
4.6	Página do Desempenho do Curso	40
4.7	Interação entre as visualizações A e B: distribuição das médias dos alunos que tiveram desempenho entre [50%, 100%]	41

4.8	Página de Análise Mensal de Indicadores	42
4.9	Página da Evolução do Curso	44
4.10	Página de Comparação do Desempenho (por fase)	45
4.11	Diferentes tipos de comparação do desempenho dos alunos	46
4.12	<i>Tooltip</i> com detalhes dos alunos da primeira fase com desempenho [50%,100%[.	46
4.13	Página Pessoal do Estudante	47
4.14	<i>Tooltip</i> com informação adicional do aluno	48
4.15	Estrutura das barras lateral e superior nas páginas do diretor de departamento	49
4.16	Visão Global do Desempenho dos alunos do Departamento	49
4.17	Escala de cor utilizada no mapeamento do desempenho	50
4.18	Ação de desagregar: habilitação → curso	50
4.19	Visão Global do Desempenho dos alunos na UA	51
4.20	Relação entre a tabela criada e a tabela principal	53
4.21	Conjunto de expressões DAX utilizados nos campos de informação dinâmicos	53
5.1	Página <i>Desempenho</i> antes e depois da Avaliação Heurística	56
5.2	Opção Ver Mais Informação do Aluno antes e depois da avaliação heurística	57
5.3	Subtítulo da página resultado da navegação antes e depois da avaliação heurística	58
5.4	Página Inicial para o perfil de Assessor da Reitoria (Nova disposição das páginas)	60
5.5	Menu flutuante com múltiplas opções de filtragem dos dados da página	61
5.6	Substituição do título das tabelas por informação dinâmica	61
5.7	Comparação Mensal de alunos em risco	62
5.8	Matriz de evolução de indicadores do aluno dividida pela granularidade de atualização	62

Lista de Tabelas

2.1	Comparação das bibliotecas e plataformas de visualização consideradas segundo alguns atributos relevantes no contexto do projeto FICA	18
2.2	Comparação de preços e tipos de licença disponíveis	19
3.1	Descrição dos atributos dos estudantes	23
3.2	Fórmulas utilizadas para calcular os principais indicadores do FICA	25
5.1	Principais problemas de usabilidade detetados nas avaliações heurísticas	57

Glossário

FICA	Ferramentas de Identificacao e Combate ao Abandono
MOOCs	Massive Open Online Courses
API	Application Programming Interface
SQL	Structured Query Language
BI	Business Intelligence
D3	Data-Driven Documents
DAX	Data analysis expressions
ECTS	European Credit Transfer System
sTIC	Serviços de Tecnologias de Informação e Comunicação
CNAEF	Classificação Nacional de Áreas de Educação e Formação
SWLS	Satisfaction With Life Scale
SPANE	Scale of Positive and Negative Experience
CTeSP	Curso Técnico Superior Profissional
IHC	Interação Humano-Computador

Introdução

1.1 MOTIVAÇÃO

Atualmente entrar no ensino superior deixou de ser um privilégio único das classes sociais mais altas e é agora uma possibilidade para a maioria dos estudantes que, após terminarem o secundário, pretendem continuar a desenvolver os seus conhecimentos em meio académico. Entre outros aspetos, as mudanças sociais e a democratização do ensino têm contribuído para que nas últimas décadas o número de alunos inscritos no ensino superior tenha disparado (Gao, 2010).

Ao elevado número de matrículas está também frequentemente associado um número significativo de alunos que, nomeadamente durante o primeiro ano, abandonam o ensino superior (Albuquerque, 2008). Estas desistências podem derivar de fatores como o insucesso académico, regime de entrada, acesso a bolsas de estudo, entre outros.

"As taxas de insucesso e abandono no ensino superior em Portugal atingem uma expressão considerável de alunos."

(DGES, 2018)

Nas instituições de ensino superior, os sistemas de informação têm um papel central relativamente à obtenção de dados dos alunos, no entanto encontram-se geralmente focados em armazenar apenas dados brutos, sem qualquer tipo de processamento, o que dificulta a deteção de padrões e consequentemente obtenção de conclusões (Barata et al., 2014).

Neste contexto, a Universidade de Aveiro desenvolveu o projeto FICA - Ferramentas de Identificação e Combate ao Abandono – que, desde o ano letivo de 2015/16, monitoriza uma série de indicadores de modo a detetar alunos da Universidade que se encontrem em risco de abandono escolar e, com isso, contribuir para a prevenção e diminuição deste abandono, especialmente pelos alunos do primeiro ano. Um dos problemas deste projeto reside na forma

como os dados obtidos são apresentados: extenso conjunto de indicadores no formato de tabelas Excel de difícil interpretação. Uma solução para contrariar este aspeto passa pelo desenvolvimento de visualizações que permitam aos vários *stakeholders* detetar padrões e obter conhecimento a partir desses mesmos dados.

1.2 OBJETIVOS

O objetivo desta dissertação consiste no estudo do problema em causa no projeto FICA, dos dados gerados, utilizadores e contextos de utilização, tendo como foco a proposta de uma solução que permita a exploração visual desses dados e ajude a obter as respostas que os utilizadores necessitem a partir deles. As visualizações propostas devem poder ser usadas por várias entidades (professores, Diretores de Curso, membros da Reitoria, etc.) de modo a identificar alunos em situações de risco através do cruzamento e integração de vários indicadores.

Um dos desafios deste trabalho é a sua interdisciplinaridade e a necessidade de interagir com vários intervenientes e por isso é seguida a metodologia do *participatory design* que promove o envolvimento ativo de todas as partes interessadas garantindo que o resultado final encontra-se de acordo com as suas expectativas. A solução proposta segue também uma abordagem *user-centered design* (UCD) que garante a acessibilidade a todos os seus futuros utilizadores, independentemente do seu perfil (Dix et al., 2003).

Este trabalho poderá contribuir de forma significativa para o combate ao abandono escolar através da sua futura integração nos sistemas de informação da Universidade de Aveiro permitindo aos vários *stakeholders* a exploração visual dos dados gerados pelo projeto FICA e consequentemente tornar mais eficaz o processo de extração de informação relevante.

1.3 ESTRUTURA

Esta dissertação encontra-se dividida em seis capítulos.

No atual capítulo é apresentado um breve enquadramento do tema, as motivações e os principais objetivos do trabalho.

O capítulo 2, para além de uma introdução à exploração visual de dados, é composto pela análise de vários trabalhos de visualização realizados com dados académicos e pela comparação de múltiplas bibliotecas e plataformas de visualização.

No capítulo 3 é apresentado em detalhe o projeto FICA e descritos todos os passos adotados durante a fase de elicitação de requisitos.

No capítulo 4 é detalhado todo o processo de desenvolvimento da solução proposta para a exploração visual dos dados, desde o pré-processamento dos dados até ao desenvolvimento das visualizações.

No capítulo 5 é efetuada a avaliação da solução proposta nomeadamente através de testes de usabilidade e avaliações heurísticas.

Por último, no capítulo 6 são apresentadas as principais conclusões obtidas com o trabalho realizado e os possíveis passos futuros, nomeadamente ao nível da integração da solução proposta nos Serviços de Tecnologias de informação e Comunicação (sTIC) da Universidade de Aveiro.

Visualização de Dados Acadêmicos

A Visualização de Informação tem vindo a ganhar relevância recentemente apesar das primeiras visualizações datarem já de há muitos de anos. O principal objetivo de qualquer visualização é permitir extrair informação a partir de um grande volume de dados abstratos de uma forma rápida, clara e precisa, sem necessidade de efetuar um grande esforço cognitivo (Kosara, 2016).

Neste capítulo são descritas as principais etapas no desenvolvimento de uma aplicação de exploração visual de dados e analisados alguns trabalhos de visualização realizados com dados académicos. Por fim é realizado um estudo de *benchmarking* relativamente a algumas das bibliotecas e plataformas mais populares e conceituados no atual mercado de visualização e relevantes no contexto deste trabalho.

2.1 EXPLORAÇÃO VISUAL DE DADOS

A comunicação de informação através de estímulos visuais pode ser estabelecida a partir de várias técnicas de visualização baseadas em diferentes metáforas visuais (Keim, 2002). A decisão final acerca das técnicas a utilizar só deve ser tomada após serem considerados fatores como o tipo de dados a representar, as perguntas a que os utilizadores pretendem responder, a tarefa a realizar e as suas características cognitivas, entre outros. Esta forma de transmitir informação é utilizada nas mais diversas áreas do conhecimento, desde a Informática, até à Medicina, Economia, Física, etc. Seja qual for o tipo de visualização a criar, a sua principal função é oferecer uma maior facilidade ao utilizador na execução de determinada tarefa e por isso, de modo a serem cumpridos todos os requisitos por si especificados, a adoção de uma metodologia *user-centered design* é essencial (Munzner, 2014).

Por sua vez, uma aplicação de exploração visual de dados tem como principal objetivo, a partir do uso de visualizações, permitir aos seus utilizadores analisar grandes *datasets* e extrair conhecimento dos dados em estudo. O desenvolvimento de uma aplicação deste tipo deve seguir uma abordagem *user-centered design* centrada nos utilizadores, nas suas necessidades e nas suas expectativas em relação ao produto final. Mais concretamente, deve seguir a

metodologia do *participatory design* onde o utilizador participa ativamente durante todo o processo de desenvolvimento (Dix et al., 2003).

Desde a contextualização e definição dos fatores chave da informação a representar até ao lançamento da solução final são várias as etapas a percorrer. A Figura 2.1 mostra os diferentes passos presentes no desenvolvimento de uma aplicação de visualização segundo (Kirk, 2016).

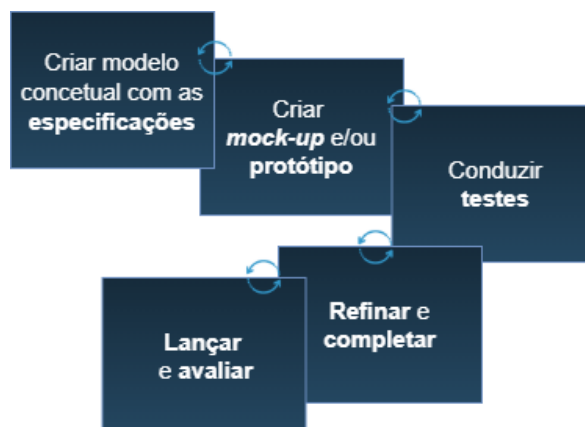


Figura 2.1: Fases do ciclo de desenvolvimento de uma aplicação de visualização segundo (Kirk, 2016)

O primeiro passo consiste em perceber o que irá ser visualizado, por quem, e sobretudo porquê. É na fase de *design* que estas questões são respondidas. Segue-se a criação de *mock-ups* e/ou protótipos que dão origem a uma primeira versão da solução e que devem já demonstrar algumas das principais funcionalidades da versão final. De modo a avançar para a fase seguinte é necessário efetuar testes sobre os protótipos desenvolvidos, primeiro pelos próprios programadores e posteriormente por utilizadores. De acordo com os resultados obtidos, se necessário, procede-se ao refinamento da solução e à condução de novos testes. Este processo é repetido até que a solução cumpra todos os requisitos definidos. Aqui são corrigidos possíveis erros e problemas identificados durante os testes, removidos conteúdos desnecessários, melhorada a consistência e coesão das escolhas inicialmente tomadas, verificada a eficácia de cada componente e revisitados os requisitos iniciais. Finalmente é lançada a versão final da aplicação que, embora já tenha sido avaliada durante a fase de desenvolvimento devido à metodologia do *participatory design*, é submetida a uma avaliação definitiva (Kirk, 2016).

Nos últimos anos, o termo *Big Data*, que descreve o grande volume de dados estruturados e não estruturados gerados a cada segundo, é presença assídua nas mais diversas instituições, sejam elas académicas, companhias de engenharia ou governamentais. Este fenómeno tem provocado a saturação dos limites das visualizações tradicionais e levado à procura de novos métodos que não tornem a representação da informação demasiado densa e confusa (Ali et al., 2016).

Por forma a apresentar uma grande quantidade de dados rapidamente e num só lugar é cada vez mais frequente a utilização de *dashboards*, que por sua vez pressupõem um conjunto de aspetos visuais essenciais para a existência de uma comunicação clara, precisa e eficaz

com os seus utilizadores. (Few & Edge, 2007) definem *dashboard* como a exibição visual da informação mais importante necessária para atingir um ou mais objetivos, consolidada e organizada num só ecrã para que possa ser facilmente monitorizada. O conteúdo exibido deve não só fornecer uma visão geral sobre os dados, como também possuir mecanismos que permitam explorá-los com maior nível de detalhe e assim oferecer um maior suporte na tomada de decisões.

Quando comparado ao processo de criação de gráficos individuais, o desenho de um *dashboard* é uma tarefa bem mais complexa que exige especial cuidado pois o facto de reunir uma enorme quantidade de informação num só lugar pode levar ao surgimento de problemas como a ultrapassagem dos limites do ecrã, falta ou excesso de precisão na representação dos dados, escolha inoportuna das visualizações, ineficiência ao realçar a informação mais importante, escassez ou utilização em abundância de uma determinada cor, decorações pouco relevantes, entre outros. É essencial que a informação apresentada permita ao utilizador obter conhecimento a partir dos dados de forma clara sem causar quaisquer desvios no seu foco de interesse (Few, 2006).

Tendo em conta as características dos *dashboards*, na solução proposta no Capítulo 4 é utilizado este conceito para mostrar a cada tipo de utilizador um resumo da informação mais relevante para o seu perfil, dando acesso à exploração dos dados em detalhe através de diferentes visualizações.

2.2 APLICAÇÕES DE EXPLORAÇÃO VISUAL DE DADOS ACADÉMICOS

O aumento do número de alunos no ensino superior e o aparecimento e rápida expansão dos MOOCs (*Massive Open Online Courses*) (Vardi, 2012) tem dado origem a vários estudos e trabalhos que têm como principal foco permitir a análise e interpretação da grande quantidade de dados gerados pelas instituições de ensino. Neste âmbito destacam-se sobretudo investigações realizadas em áreas como Psicologia ou Ciências da Educação (Albuquerque, 2008), no entanto, também projetos de visualização em campos como Finanças (Sorenson & Brath, 2013) ou Medicina (Harle et al., 2012) podem ser considerados de alguma forma análogos ao contexto do ensino facultando assim sugestões úteis para a solução a desenvolver.

Nesta secção são apresentados alguns dos trabalhos analisados que utilizam dados obtidos a partir dos sistemas de informação de diferentes instituições de ensino.

TECNICOVIS

O TECNICOVIS (Martins et al., 2018) é uma solução de visualização de informação que, com base na extração e respetivo pré-processamento de dados recolhidos da API do FenixEdu (Barata et al., 2014), permite analisar o percurso académico dos estudantes do Instituto Superior Técnico de Lisboa e estudar a sua evolução nas diferentes disciplinas e cursos ao longo dos anos.

Considerando indicadores como o número de matrículas em cada curso, notas dos alunos, qualidade das unidades curriculares ou avaliação dos professores, foram criadas várias visualizações interativas que têm um papel fundamental na identificação de potenciais problemas nos diferentes cursos e disciplinas contribuindo assim para a melhoria do ensino praticado.

Na Figura 2.2 são apresentadas as visualizações utilizadas nesta solução, dispostas numa única página, das quais fazem parte uma matriz, que permite representar múltiplos cursos e unidades curriculares simultaneamente, onde cada célula a representa a percentagem de aprovações dos alunos a uma determinada disciplina, um gráfico de dispersão que permite estabelecer a correlação entre os vários atributos e um gráfico de linhas que representa a avaliação da qualidade da unidade curricular por parte dos professores. É possível observar constante interconectividade entre todas as visualizações através do destaque de certos elementos de acordo com a seleção efetuada.

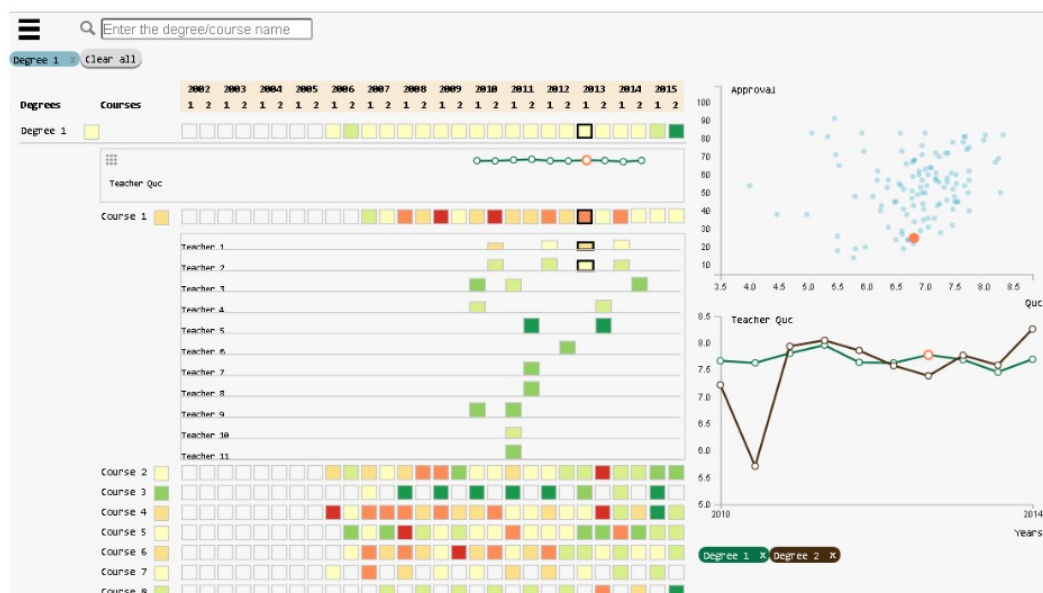


Figura 2.2: Página principal do TECNICOVIS com todas as visualizações: seleção de uma instância de um curso (Martins et al., 2018)

Aplicação de Visualização da Faculdade Cenecista de Osório

Os ambientes de aprendizagem utilizados nos cursos à distância geram habitualmente grandes quantidades de dados relativos aos seus alunos que permitem retirar diversas conclusões em relação ao seu sucesso académico. A partir desta premissa foi desenvolvido um sistema de visualização que, utilizando dados referentes a três anos da Faculdade Cenecista de Osório, localizada no Brasil, relativos aos recursos acedidos pelos alunos, permite prever se estes irão finalizar ou não o curso (Weiland & Manssour, 2015). Após recolhida toda a amostra, são aplicados algoritmos de *data mining* sobre os dados que por sua vez seguem para o módulo de visualização, representado na Figura 2.3, onde é possível observar duas vistas: no topo, sob a forma de *table lens*, um resumo dos dados carregados e, no restante espaço da página, através

de um *pie chart* e de um gráfico de barras sobrepostas, informação sobre a distribuição dos acessos aos vários recursos por parte dos estudantes.



Figura 2.3: Distribuição dos acessos aos vários recursos por parte dos estudantes entre Julho de 2013 e Fevereiro de 2014 na Faculdade Cenecista de Osório (Weiland & Manssour, 2015)

A disposição dos elementos visuais e as diferentes cores associadas ao tipo de recurso são algumas das características que promovem a fácil interpretação dos dados por parte dos utilizadores. A aplicação oferece ainda várias opções de interação e mecanismos de filtragem que permitem manipular as visualizações criadas e explorar os dados com maior nível de detalhe contribuindo para uma eficaz deteção de padrões.

Visualização de *datasets* educacionais utilizando um sistema de *queries*

O Sistema de Educação Pública do Estado de Washington, nos Estados Unidos da América, publica anualmente vários relatórios com diversos dados académicos que permitem obter informação sobre a performance dos seus estudantes, taxas de desistência/graduação, credenciais dos docentes, dados demográficos, entre outros. Considerando este *dataset* foi desenvolvido um sistema web que, a partir de uma *query* em linguagem SQL definida pelo utilizador, consegue inferir a visualização mais apropriada a usar (Desai et al., 2014). Após carregada a *query* no sistema, é realizada uma verificação do tipo de dados (numéricos ou categóricos) que por sua vez são representados através do elemento visual mais adequado. Para responder às *queries* do utilizador, o sistema utiliza um dos três diferentes tipos de gráficos: *doughnut chart*, gráfico de barras ou mapa coroplético.

Quando o resultado de uma determinada *query* pode ser representado através de mais do que um tipo de gráfico, o sistema opta por aquele que permite analisar os dados de forma

mais eficaz. Por exemplo, para uma mesma *query* questionando o número de professores com Mestrado nas várias escolas públicas do Estado de Washington, como é possível verificar pela Figura 2.4, poderiam resultar dois diferentes *outputs*, no entanto, devido à quantidade de informação representada, o *doughnut chart* 2.4a, onde cada setor corresponde a uma escola e o seu tamanho é diretamente proporcional ao número de professores com mestrado, é de difícil leitura sobretudo devido à sobreposição de legendas verificada. Por sua vez, o gráfico de barras 2.4b, ao utilizar uma barra de deslocamento que permite deslizar entre as várias instituições de ensino, permite uma análise muito mais rápida e clara dos resultados sendo por isso a visualização preferível para responder à *query* em questão.

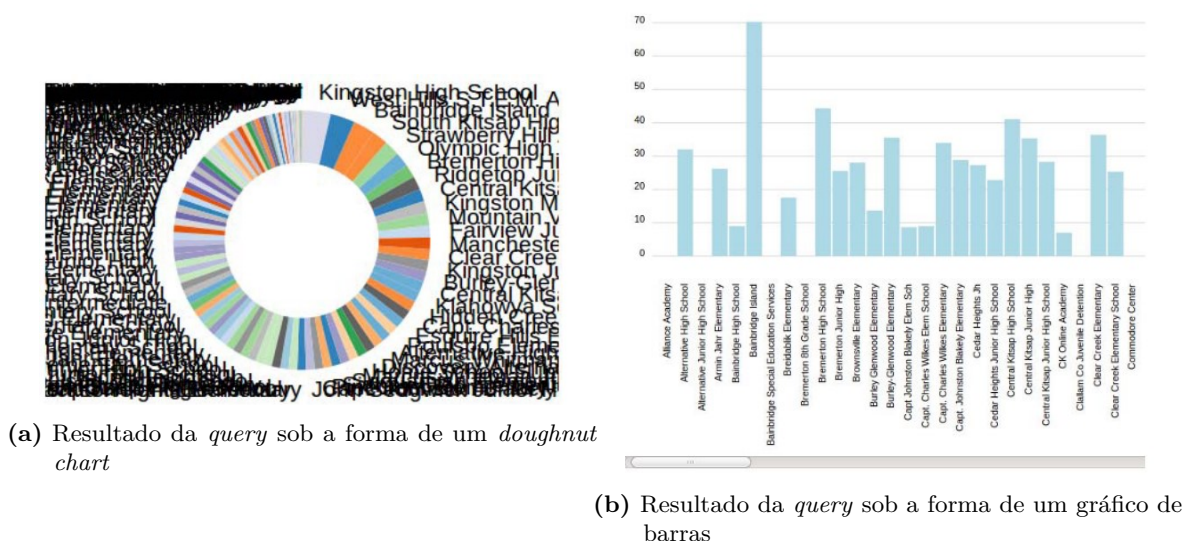


Figura 2.4: Alternativas de resposta para a *query* com o número de professores com mestrado no sistema desenvolvido pela Universidade de Washington (Desai et al., 2014)

SAPI: Sistema de Gestão de Publicações para investigadores da UPAEP

Todos os anos são realizados e publicados vários trabalhos de investigação na Universidade Popular do Estado de Puebla (UPAEP), localizada no México, que levam à geração de grandes quantidades de dados. Com o objetivo de obter mais facilmente informação acerca do que já foi realizado até à data na instituição, foi criado um sistema web que permite acompanhar e ajudar no suporte à tomada de decisões relativamente a futuras publicações (Cruz-Guzmán & Rosas, 2020).

Existe na UPAEP uma área que se dedica exclusivamente a promover e apoiar professores e estudantes interessados em desenvolver novas investigações. Os incentivos oferecidos, quer económicos quer a nível de recursos, são decididos tendo em consideração vários projetos de anos anteriores, dos quais fazem parte pesquisas, publicações de artigos, traduções, congressos organizados, prémios conquistados, entre outros, e apenas são atribuídos após a proposta ser analisada e validada pelo departamento de investigação. O sistema de gestão de publicações para investigações - SAPI - desenvolvido surge com o propósito de agilizar todo este processo.

Este sistema permite a comunicação entre três tipos de utilizadores com diferentes responsabilidades: os investigadores que criam/editam as publicações, o departamento de investigação que as aprova/rejeita e o departamento organizacional e de avaliação cujo foco principal é observar e analisar os resultados obtidos. Todos os utilizadores têm acesso a relatórios compostos por vários gráficos interativos que permitem rapidamente obter resposta ao pedido efetuado. Em poucos cliques é possível adicionar ou remover campos nos gráficos gerados permitindo uma análise personalizada dos resultados obtidos no pedido atual.

As visualizações são geradas sobretudo tendo em conta as necessidades do departamento de investigação. Na Figura 2.5 é possível observar algumas das representações presentes nos relatórios, mais especificamente, um gráfico de barras que mostra a quantidade de publicações aprovadas e rejeitadas nos últimos anos, um *pie chart* que apresenta o número total de publicações por departamento e um gráfico de linhas que mostra o número de publicações aprovadas por categoria.

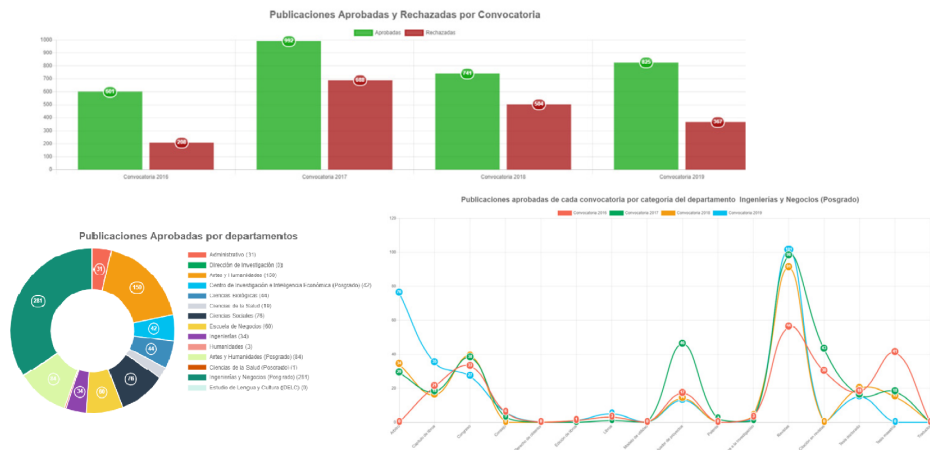


Figura 2.5: Diferentes visualizações de interesse para o departamento de investigação geradas pelo SAPI (Cruz-Guzmán & Rosas, 2020)

Estes gráficos permitem, por exemplo, obter conclusões relativamente ao investimento efetuado nos diversos projetos com base nos resultados obtidos.

O gráfico de barras sobrepostas representado na Figura 2.6 mostra a contribuição de todos os investigadores da instituição, baseada num sistema de pontos. Esta visualização permite à direção e Reitoria a análise do perfil de cada investigador e, no caso de ser detetada baixa produção, investigar as possíveis causas do insucesso por forma a adotar estratégias para melhorar a sua performance.

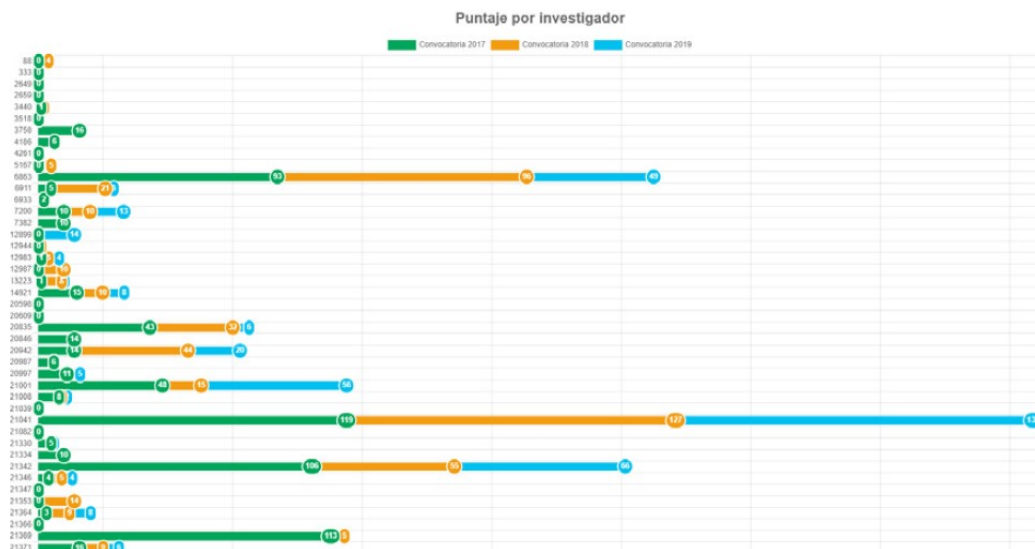


Figura 2.6: Gráfico com a pontuação geral obtida por cada investigador gerado pelo SAPI (Cruz-Guzmán & Rosas, 2020)

Monitorização de Taxas de Graduação e Retenção dos Estudantes

A retenção e o abandono prematuro no ensino superior são dois problemas centrais que têm recentemente atraído especial atenção por parte de diversos trabalhos de investigação que se focam sobretudo na performance e progresso académico dos estudantes. Neste contexto, foi desenvolvida na Universidade de Tecnologia e Economia de Budapeste (BME) uma ferramenta de visualização que permite monitorizar o percurso dos seus mais de 30.000 alunos através de diagramas de Sankey criados no *Google Charts*¹ (Horváth et al., 2018). A partir da análise desses diagramas o principal objetivo deste trabalho é detetar e entender alterações temporais no fluxo dos estudantes e assim encontrar justificação para as consideráveis taxas de abandono verificadas.

Para o desenvolvimento desta ferramenta, foram disponibilizados pelo Gabinete Central Académico da BME dados de mais de 30.000 alunos, devidamente anonimizados por um código de seis dígitos, relativos aos anos letivos entre 2010/11 e 2016/17, perfazendo um total de 14 semestres. Para além do código identificativo, o registo de cada estudante contém informação sobre a faculdade que frequenta, o estado da sua matrícula no semestre atual (ativo ou passivo), ano e razão de admissão e nota de ingresso. À semelhança do que acontece no programa FICA da Universidade de Aveiro, os dados encontram-se também armazenados na forma de extensos ficheiros Excel com milhares de entradas.

Neste trabalho, antes de correr as visualizações, é efetuado o processamento dos dados através de um algoritmo desenvolvido em *python* que permite a configuração de múltiplos filtros relativos à faculdade que o aluno frequenta, ao seu ano de inscrição, se é a primeira vez que está no ensino superior e à sua nota de entrada. Também de modo a possibilitar

¹<https://developers.google.com/chart/interactive/docs/gallery/sankey> (consultado em 26-03-2020)

uma análise mais clara da informação extraída a partir dos dados representados, é possível selecionar diferentes versões dos diagramas.

Na Figura 2.7 é possível acompanhar o percurso de todos os estudantes que se inscreveram numa determinada faculdade da BME no ano letivo 2012/13. Os diagramas são compostos por vários nós com um nome único identificado pelo semestre e estado atual - ativo se está a frequentar o plano curricular ou inativo no caso de ter os estudos em pausa - e vários arcos correspondentes ao caminho dos vários estudantes.

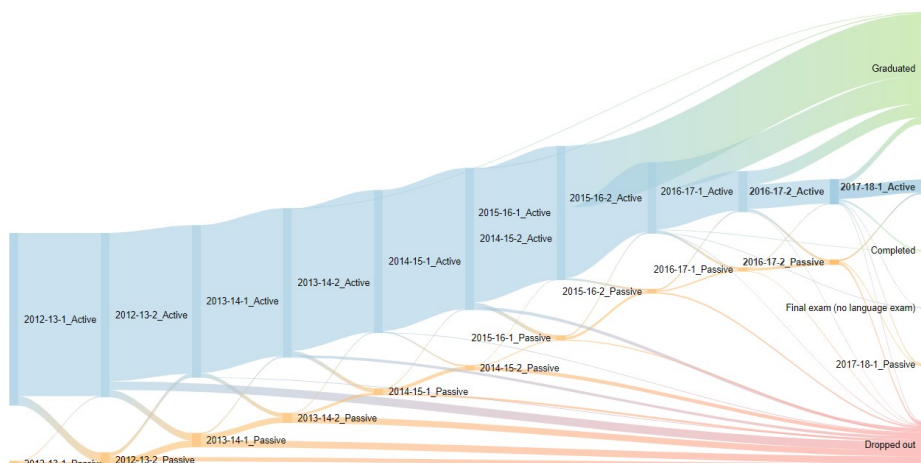


Figura 2.7: Fluxo dos Estudantes Inscritos numa Faculdade da BME em 2012 (Horváth et al., 2018)

Para além de fazer o acompanhamento dos estudantes ao longo de vários anos, é também possível analisar a sua situação num momento específico. O diagrama da Figura 2.8 mostra o estado atual dos alunos de uma faculdade da BME no final do ano letivo 2016/17.

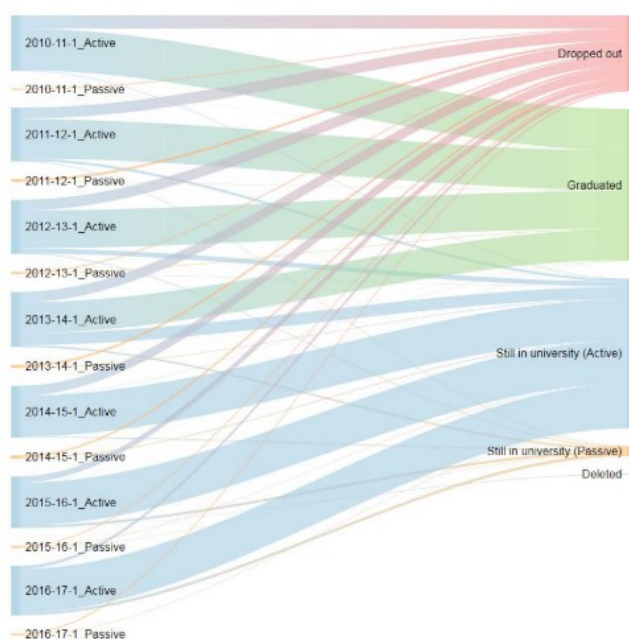


Figura 2.8: Situação atual dos alunos de uma faculdade da BME no final do ano letivo 2016/17 (Horváth et al., 2018)

As visualizações seguem uma orientação da esquerda para a direita e a largura dos arcos é proporcional ao número de estudantes que percorreu um determinado caminho permitindo assim detetar possíveis padrões. Também a cor tem um papel fundamental na identificação destes padrões: inicialmente os nós/arcos ativos estão representados a azul e os passivos a laranja no entanto, com o avançar da dimensão temporal, convergem para verde no caso do aluno ter sucesso e graduar-se, ou para vermelho caso abandone os estudos.

Principais Conclusões

Apesar dos trabalhos analisados utilizarem diferentes conjuntos de dados e apresentarem objetivos distintos foi possível detetar vários padrões, nomeadamente ao nível das técnicas de visualização utilizadas.

Todas as soluções desenvolvidas apresentam mecanismos de seleção e filtragem dos dados que permitem ao utilizador interagir com as várias visualizações de modo a destacar certos elementos ou eliminar informação adicional que possa desviar o seu foco de análise. A possibilidade de estudo dos dados com uma dimensão temporal é outro dos traços que, à exceção do sistema de inferência (Desai et al., 2014), é partilhado pelas várias soluções desenvolvidas.

A cor assume igualmente um papel fundamental na maioria das visualizações. A utilização de cores habitualmente associadas a aspetos positivos ou negativos faz com que o utilizador consiga rapidamente obter conclusões sobre os dados mesmo antes de iniciar uma análise mais cuidada. Por exemplo nos trabalhos analisados é frequente a utilização de cores como o verde ou vermelho para indicar respetivamente sucesso ou insucesso académico.

Por fim, as soluções que apresentam múltiplas visualizações numa mesma página promovem a interconectividade entre elas, refletindo a seleção de um elemento nas restantes, fazendo com que a informação apresentada seja constantemente coerente.

2.3 BIBLIOTECAS E PLATAFORMAS DE VISUALIZAÇÃO

Atualmente existe uma enorme variedade de recursos para a criação de visualizações. Esta lista de tecnologias vai desde as bibliotecas direcionadas para desenvolvimento web, como a D3.js², Vega.js³ ou p5.js⁴, que por serem de baixo nível oferecem bastante liberdade na criação de novas visualizações, até às plataformas de *Business Intelligence* (BI), utilizadas no desenvolvimento de *dashboards* e relatórios que contribuem para a gestão e planeamento do futuro dentro de uma organização através da visualização, análise e deteção de padrões em grandes quantidades de dados (Gounder et al., 2016). Apesar de as plataformas de BI serem normalmente caracterizadas pela sua facilidade de utilização, as bibliotecas *javascript* referidas, por se tratarem de linguagens de baixo nível, são mais flexíveis, no entanto requerem boas capacidades de programação que reduzem a sua acessibilidade.

²<https://d3js.org/> (consultado em 25-03-2020)

³<https://vega.github.io/vega/> (consultado em 25-03-2020)

⁴<https://p5js.org/> (consultado em 25-03-2020)

A seguir são apresentadas de forma resumida algumas das mais populares e conceituadas bibliotecas e plataformas de visualização que possuem características importantes no âmbito do projeto FICA.

D3 - Data-Driven Documents. D3, também conhecida por D3.js, é uma biblioteca JavaScript utilizada para manipular documentos com base em dados. Combina poderosos componentes de visualização de informação com uma abordagem orientada a dados para manipulação do DOM (*Document Object Model*) que possibilita a sua integração com os *browsers* modernos garantindo acesso a todas as suas capacidades.

Essencialmente, a D3 facilita a geração e manipulação de documentos web ao permitir controlar os processos de *loading* e *binding* de dados a elementos do documento que posteriormente sofrem transformação através do mapeamento das respetivas propriedades visuais. Por fim os elementos alteram o seu estado em resposta aos *inputs* do utilizador (Murray, 2017). O facto de ser baseada na web permite que esta biblioteca utilize *standards* como HTML, SVG e CSS para controlar a camada de apresentação e com isso a criação de visualizações dinâmicas e interativas. Para além de possuir vários *layouts* e técnicas de visualização pré-definidas, a flexibilidade da D3 permite uma fácil integração de novos *plugins* e com isso o desenvolvimento das mais diversas representações gráficas (Blasilli et al., 2019). Possui também uma comunidade bastante forte e ativa.

O principal aspeto negativo reside no facto da sua API ser de baixo nível exigindo assim um maior esforço de aprendizagem por parte de novos utilizadores. A utilização de tecnologias como HTML, JavaScript ou CSS, que exigem algum conhecimento prévio, contribui ainda mais para o aumento desta curva de aprendizagem.

Tableau. O Tableau⁵ é uma solução de *Business Intelligence* de visualização de dados que permite aos utilizadores criar visualizações interativas na forma de *dashboards* ou *worksheets*. A sua natureza *drag & drop* permite que em pouco tempo qualquer utilizador consiga representar os seus dados através de múltiplas opções, que incluem *pie charts*, gráficos de barras ou linhas, *heatmaps*, gráficos de dispersão, entre outras, que utilizam algumas das técnicas de visualização mais comuns, otimizando a análise de dados e influenciando o processo de tomada de decisão por parte das organizações (Keim, 2002). Para além do *Tableau Desktop*, existem também o *Tableau Server* e o *Tableau Online* que permitem respetivamente ligar os elementos visuais a um vasto número de fontes e a partilha de dados na web.

Apesar de ser uma ferramenta simples pelo facto de não exigir conhecimentos de programação e não obrigar os utilizadores a escrever código, o Tableau apresenta preços de subscrição bastante superiores a outras soluções de *Business Intelligence* como o Power BI⁶ ou o Spotfire⁷.

⁵<https://www.tableau.com/> (consultado em 25-03-2020)

⁶<https://powerbi.microsoft.com/pt-pt/> (consultado em 25-03-2020)

⁷<https://tibco.com/products/tibco-spotfire> (consultado em 25-03-2020)

Power BI. O Microsoft Power BI é uma plataforma de *Business Intelligence* que oferece aos seus utilizadores um pacote de ferramentas para a agregação, análise, visualização e partilha de dados.

As principais três componentes do Power BI são a *Power Query*, onde é efetuada a transformação dos dados, a *Power Pivot*, onde é criada uma modelação tabular e por fim a *Power View* que apresenta os vários elementos visuais criados. Os dados podem ser carregados a partir das mais variadas fontes, desde ficheiros Excel, PDF, XML, bases de dados SQL, páginas web, entre outras. A Figura 2.9 mostra as principais funcionalidades presentes no desenvolvimento de uma solução em Power BI.



Figura 2.9: Principais funcionalidades presentes numa solução desenvolvida em Power BI

O facto do Power BI ser um serviço baseado na cloud (SaaS⁸) permite que os *dashboards* interativos exibidos possam ser criados e atualizados a partir de várias fontes de dados e publicados online em locais como blogues, redes sociais, ou e-mails (Negrut, 2014).

Apesar de ser uma solução intuitiva, poderão surgir algumas dificuldades na manipulação das tabelas e consultas na *Power Query* e na sua ligação com as páginas compostas pelas visualizações uma vez que são utilizadas duas linguagens próprias, M e Data Analysis Expressions (DAX), que exigem algum conhecimento prévio.

Existe uma forte integração do Power BI com o Excel que permite, entre outros aspetos, permutar e manipular dados entre ambas as ferramentas. Este software é ainda dotado de uma comunidade ativa e atualizações mensais, baseadas na opinião dos seus utilizadores, que ajudam a consolidá-lo como uma das melhores opções de mercado no ramo de BI ⁹.

Microsoft Excel. O Microsoft Excel é um software produzido pela Microsoft que permite aos seus utilizadores organizar, formatar e aplicar fórmulas a dados através de um sistema de folhas de cálculo. Estruturado por uma coleção de células organizadas em linhas e colunas, o Excel torna possível o processo de transformar dados em informação. Possibilita também a criação de vários tipos de visualizações.

Atualmente o Excel é o software de análise e modelação de dados mais popular do mercado, no entanto, o facto de as visualizações que disponibiliza derivarem de *templates* já definidos torna limitada a representação de dados por parte dos utilizadores que nem sempre encontram

⁸<https://azure.microsoft.com/pt-pt/overview/what-is-saas/> (consultado em 25-03-2020)

⁹<https://info.microsoft.com/ww-landing-2020-gartner-magic-quadrant-for-analytics-and-business-intelligence.html> (consultado em 18-06-2020)

a visualização adequada (Wun et al., 2016). Trata-se também de uma solução pouco interativa, nomeadamente para utilizadores pouco experientes.

TIBCO Spotfire. O TIBCO Spotfire é uma ferramenta inteligente, segura, flexível e escalável de análise de dados que permite impulsionar o valor de um projeto através da visualização de informação e capacidades preditivas. O Spotfire apresenta um *dashboard* eficiente que permite aos seus utilizadores organizar várias fontes de dados, relatórios e objetos numa só localização central que permanece atualizada em tempo real. O seu *design* intuitivo e o assistente de recomendação que sugere as melhores práticas e visualizações a utilizar baseadas nos dados em questão são algumas das características que tornam possível a utilização desta plataforma por utilizadores menos experientes (Choo & Saeger, 2010).

O Spotfire faz ainda uso de alguns conceitos e técnicas de *machine learning*¹⁰ que, a partir do histórico dos dados, permitem a antecipação de tendências e impulsionam a tomada de ações preventivas.

Gephi. O Gephi é um software de visualização utilizado na representação de grafos e redes. É uma ferramenta gratuita e *open source* que possui inúmeras aplicações em áreas como as Ciências Biológicas ou Economia. Permite efetuar visualizações interativas e em tempo real que possibilitam a análise e manipulação de grandes quantidades de dados.

Uma vez que adota um modelo multitarefa, consegue lidar com redes que podem conter mais de 20.000 nós (Bastian et al., 2009).

Comparação de Bibliotecas e Plataforma de Visualização

Tendo em conta a grande quantidade de recursos disponíveis na área de visualização de informação, torna-se imprescindível a análise e a comparação entre as várias opções antes da escolha final. Este processo de comparação é essencial para definir as principais metas a atingir em cada projeto.

A facilidade de acesso por parte de utilizadores que não sejam profissionais na área das tecnologias, rapidez na adaptabilidade às condições de mudança do negócio e o encorajamento à exploração de dados são alguns dos aspetos mais importantes nesta tomada de decisão.

Na Tabela 2.1 é possível observar uma comparação entre as várias tecnologias e ferramentas analisadas segundo alguns atributos essenciais dentro do contexto do projeto FICA:

- **Open Source:** a ferramenta é *open source*, ou seja, é possível aceder ao seu código fonte, inspecioná-lo, modificá-lo e contribuir com novas visualizações;
- **Interatividade:** as visualizações criadas permitem interação com o utilizador;
- **Cliente Desktop:** existe uma versão *desktop* do software onde podem ser criadas as visualizações;

¹⁰<https://tibco.com/blog/2018/11/28/ai-and-machine-learning-inside-and-out-in-spotfire-x/> (consultado em 25-03-2020)

- **Cliente Online:** existe a possibilidade de criar visualizações online e publicá-las na Web de modo a poder partilhá-las com outros utilizadores;
- **API:** existência e disponibilidade de uma API;
- **Modelação de dados:** capacidade de manipular e conectar dados de múltiplas fontes; combinar dados a partir do relacionamento entre tabelas;
- **Dashboard:** possibilidade de juntar várias visualizações num só lugar;
- **Comunidade:** existência de uma comunidade ativa e tutoriais/cursos online.

	D3	Gephi	Spotfire	Tableau	Excel	Power BI
Open Source	✓	✓	✗	✗	✗	✓
Interatividade	✓	✗	✓	✓	✓	✓
Cliente Desktop	✗	✓	✓	✓	✓	✓
Cliente Web	✗	✗	✓	✓	✓	✓
API	✗	✓	✓	✓	✓	✓
Modelação de Dados	✗	✗	✓	✓	✓	✓
Dashboard	✗	✗	✓	✓	✓	✓
Comunidade	✓	✓	✓	✓	✓	✓

Tabela 2.1: Comparação das bibliotecas e plataformas de visualização consideradas segundo alguns atributos relevantes no contexto do projeto FICA

Os resultados obtidos na comparação das várias medidas, presentes na Tabela 2.1, mostram uma diferença significativa entre os resultados obtidos para a D3 e Gephi relativamente às restantes ferramentas, principalmente justificada por serem ambas dirigidas a utilizadores com conhecimentos de programação. Por atuarem a baixo nível, podem ser consideradas ferramentas mais flexíveis na criação de visualizações, no entanto, entre outros aspetos, o facto de não oferecerem suporte à criação de *dashboards*, levou a que o leque de escolhas fosse reduzido às restantes quatro opções, todas plataformas de BI.

Entre estas plataformas, o Excel, apesar de respeitar praticamente todos os atributos das restantes, apresenta-se como uma solução mais orientada para a análise em profundidade dos dados no seu estado bruto e não representados através de uma vertente mais visual. Por sua vez, o Power BI é uma ferramenta poderosa que combina a maioria dos recursos do Excel e possui amplas capacidades de visualização, fazendo dele um candidato mais forte no contexto desta dissertação.

O *Spotfire*, *Tableau* e Power BI têm funcionalidades semelhantes e constituem todos soluções de BI intuitivas com sistemas *drag & drop* que possibilitam a sua utilização por parte de qualquer utilizador com conhecimentos básicos na área da visualização, contudo apenas o Power BI possui uma componente *opensource* destacando-se como a mais completa destas *frameworks*.

O Power BI oferece aos seus utilizadores a possibilidade de acederem ao código fonte dos elementos visuais já existentes e contribuir com novas visualizações, sendo estas desenvolvidas a partir da biblioteca D3.

Complementando a comparação de atributos efetuada, foram analisados os diferentes preços e tipos de licença das tecnologias destacadas:

Tecnologia	Licença	Preço (mensal)
D3	Gratuita	-
Gephi	Gratuita	-
Spotfire	Trial (30 dias)	25-125\$
Tableau	Trial (14 dias)	70\$
Power BI	Versão Desktop (Func. limitadas)	10\$ (Versão Pro)

Tabela 2.2: Comparação de preços e tipos de licença disponíveis

A partir da Tabela 2.2 é possível constatar que o Power BI sai uma vez mais vencedor sendo o serviço que apresenta o preço mais baixo. Também o facto de a Universidade de Aveiro estar de momento a investir nesta ferramenta e oferecer a todos os seus membros acesso à sua versão *pro* afirma-se como um aspeto importante para a possível integração de algum trabalho futuro nos serviços de informação da instituição. Por estas razões, a solução proposta nesta dissertação foi desenvolvida em Power BI.

Visualização de Dados do Projeto FICA

No capítulo anterior foram abordados alguns casos de estudo de aplicações de exploração visual de dados académicos e analisadas algumas das principais plataformas e bibliotecas de visualização atualmente existentes tendo sido destacadas as suas principais qualidades e limitações. Uma componente imprescindível para o desenvolvimento de uma aplicação deste género consiste nos dados que são usados pelas visualizações criadas. No âmbito desta dissertação é utilizada uma amostra de dados, devidamente anonimizados, dos alunos monitorizados pelo projeto FICA.

A primeira etapa presente no desenvolvimento de uma qualquer aplicação de visualização consiste na contextualização do problema. Através da elicitação de requisitos levada a cabo com peritos do domínio é possível encontrar as principais questões a serem respondidas e estabelecer o que será importante analisar. Definidas as questões, o passo seguinte consiste em perceber qual a melhor forma de obter as respetivas respostas e verificar se existem dados suficientes para cumprir esse objetivo.

Neste capítulo é apresentado em detalhe o projeto FICA, os dados por si gerados e todo o processo de elicitação de requisitos a partir do qual se definiram as principais metas a atingir com este trabalho.

3.1 OS DADOS DO FICA

O Que é o FICA?

Como já referido, o FICA – Ferramentas de Identificação e Combate ao Abandono – é um projeto desenvolvido pela Universidade de Aveiro que monitoriza uma série de atributos dos seus estudantes com vista a detetar possíveis situações de risco de abandono e prestar apoio na tomada de decisões que permitam combater estes casos e com isso melhorar o seu

aproveitamento. Este projeto iniciou-se no ano letivo de 2015/16 e é direcionado sobretudo para os Diretores de Curso, Diretores de Departamento e Assessores da Reitoria a quem são enviados mensalmente ficheiros Excel complexos e de difícil análise, preenchidos por milhares de entradas compostas por um vasto conjunto de atributos dos alunos.

O FICA surge como o principal foco num conjunto de estratégias adotado pela Universidade com o objetivo de monitorizar o percurso dos alunos nomeadamente durante o primeiro ano de estudos, o período que geralmente apresenta maiores taxas de insucesso e abandono. Deste projeto fazem parte o acolhimento dos estudantes aquando a sua chegada à universidade, o programa tutoria que visa o seu acompanhamento por estudantes de anos mais avançados, as iniciativas pedagógicas compostas por sessões de formação, *workshops* ou partilhas de experiência e o observatório do percurso dos estudantes que tem como objetivo identificar possíveis situações de risco e consequente abandono (Moreira et al., 2017). Embora continue a atribuir uma importância significativa à análise dos dados dos alunos do primeiro ano, o projeto FICA sofreu recentemente um alargamento e monitoriza atualmente os mais de 13.000 estudantes matriculados na Universidade. A Figura 3.1 exibe uma síntese dos principais objetivos deste projeto.

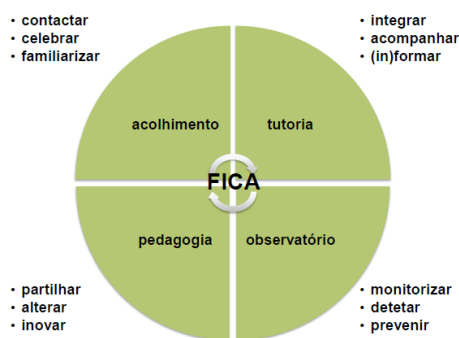


Figura 3.1: Síntese do projeto FICA (Moreira et al., 2017)

Atributos e Indicadores

Os dados disponibilizados através dos ficheiros Excel são compostos por três folhas de cálculo que incluem respetivamente uma lista dos estudantes da Universidade e os seus atributos, correspondentes a informação pessoal ou dados relativos ao seu contexto académico, uma descrição desses atributos e uma lista dos alunos que anularam a matrícula, complementada pelo motivo e data da anulação. A Figura 3.2 mostra um *sample* de um desses ficheiros.

	B	C	D	E	F	G	H	I	J
	Numero	Nome	Email	País/Nacionalidade	Programa/Tutoria	TempoParcial	SituacaoPrescricao	Habilitacao	Curso
1									
2	2/A2			PORTUGAL	Não	Sim	Não Prescrito	Licenciatura - 1º Ciclo	19
3	3/A3			PORTUGAL	Sim	Sim	Não Prescrito	Licenciatura - 1º Ciclo	22
4	4/A4			PORTUGAL	Não	Sim	Não Prescrito	Licenciatura - 1º Ciclo	52
5	9/A9			PORTUGAL	Não	Sim	Não Prescrito	Mestrado Integrado	18
6	15/A15			PORTUGAL	Não	Sim	Não Prescrito	Mestrado Integrado	32
7	17/A17			PORTUGAL	Não	Sim	Não Prescrito	Licenciatura - 1º Ciclo	25
8	26/A26			PORTUGAL	Não	Não	Não Prescrito	Licenciatura - 1º Ciclo	71
9	28/A28			PORTUGAL	Não	Sim	Não Prescrito	Licenciatura - 1º Ciclo	56
10	29/A29			PORTUGAL	Não	Não	Não Prescrito	Licenciatura - 1º Ciclo	28
11	31/A31			PORTUGAL	Não	Não	Não Prescrito	Mestrado Integrado	18

Figura 3.2: Amostra de dados do FICA

A cada aluno, representado por uma linha da folha de cálculo, estão associados vários atributos, simbolizados por colunas, que são devidamente preenchidos a partir dos dados recolhidos. Os atributos inicialmente disponibilizados encontram-se descritos na Tabela 3.1.

Atributo	Descrição
AnoCurricular	Ano mais avançado ao qual o estudante está matriculado numa UC
AnoLetivo	Ano letivo ao qual correspondem os dados do ficheiro
AnoLetivoMatricula	Ano letivo em que o estudante se matriculou no curso
AreaCNAEF	Área de Educação e Formação (CNAEF) do curso do estudante
BolsaSAS	Informação relativa a bolsa dos Serviços de Ação Social: data de candidatura e despacho, estado/situação e valor
Curso e CursoNome	Código e nome do curso do estudante
DataAmostra	Data referente à coleta dos dados
DataNascimento	Data de nascimento do estudante
ECTS_Feitos	Informação do número total de ECTS feitos pelo estudante no ano atual, no 1º semestre e da totalidade do plano
ECTS_Inscritos	Número de ECTS inscritos no 1º semestre e na totalidade do ano atual
Email	Endereço de email da UA do estudante
Escala SWLS	Resultado das questões respondidas pelo estudante relativas à escala de satisfação com a vida (SWLS)
Escala SPANE	Resultado das questões respondidas pelo estudante relativas à escala de experiência positiva e negativa (SPANE)
EstadoMatricula	Estado da matrícula do estudante à data da amostra
Habilitacao	Habilitação relativa ao curso (Mestrado, Licenciatura, CTeSP)
IngressoRG	Informação da nota de ingresso, fase do concurso nacional e ordem de preferência do estudante no curso
Media	Nota média das Unidades Curriculares feitas pelo estudante no curso
MotivoAnulacao	Se a matrícula do estudante se encontra anulada, indica o motivo da anulação
NivelEscol_Pai Mae	Nível de escolaridade do pai e da mãe do estudante
Nome e Numero	Nome e Numero do estudante
NumUCinscitas	Número de unidades curriculares às quais o estudante está inscrito
PaisNacionalidade	Pais de origem do estudante
PrimeiraVezCurso	Indica se o estudante está inscrito pela primeira vez no curso
ProgramaTutoria	Indica se o estudante está associado a um programa de tutoria
Propina_TotalEmDivida	Valor total em dívida relativo a propinas do estudante
Regime	Regime de frequência do estudante no curso
Sexo	Sexo do estudante
SituacaoPrescricao	Situação do estudante relativamente à prescrição da sua matrícula
Tempo Parcial	Indica se o estudante frequenta o curso a tempo parcial
Tipo Acesso	Código com o tipo de acesso do estudante no curso
UC	Dados de todas as unidades curriculares inscritas: código da UC, taxa de assiduidade e avaliação parcial

Tabela 3.1: Descrição dos atributos dos estudantes

O FICA atribui especial foco a um conjunto de sete indicadores, a partir dos quais se pretende monitorizar e detetar comportamentos geralmente associados a situações de risco, que são identificados por:

- **IND2_1:** Estudantes com taxa de sucesso académico abaixo de 50 % ECTS a que estão inscritos (por curso);
- **IND2_2:** Estudantes com propinas em atraso (por curso);
- **IND2_3:** Estudantes que pediram e não obtiveram bolsa ou aguardam resposta
- **IND2_4:** Estudantes com nota de ingresso abaixo de 120 pontos (1ª fase + 2ª fase);
- **IND2_5:** Número de aulas frequentadas por estudante face ao número de aulas dadas;
- **IND2_6:** Estudantes com nível de satisfação com a vida baixa (escala SWLS¹);
- **IND2_7:** Experiência positiva e negativa dos estudantes baixa (escala SPANE²).

Nas folhas de cálculo, estes indicadores localizam-se lado a lado com os restantes atributos, apresentando a ligeira diferença das respetivas células serem mapeadas por cores como se pode verificar na Figura 3.3.

IND2_1	IND2_2	IND2_3	IND2_4	IND2_5	IND2_6	IND2_7
10	10	30	30	20	10	10
10	30		10	20		
10	10	10	10	20	30	10
10	10	30	10	30	10	10
10	10		30	20	10	10
30	10		10	20	10	10
10	30	10		20	10	10
10	10	10	10	10	10	10

Figura 3.3: Mapeamento visual das células correspondentes aos indicadores

As fórmulas utilizadas para calcular os indicadores e consequentemente determinar a codificação a utilizar são descritas na Tabela 3.2.

Fontes e carregamento dos dados

A recolha dos dados do projeto FICA é efetuada pelos Serviços de Tecnologias de Informação e Comunicação (sTIC) da Universidade de Aveiro a partir de diferentes origens. Desde o Sistema Integrado de Gestão Académica (SIGAcad)³ que faz o acompanhamento de todas as atividades letivas da Universidade fornecendo dados relativamente às matrículas, número de alunos equivalentes a tempo integral (AETI), aproveitamento, prescrições e propinas, o Portal da Universidade de Aveiro (PACO), que contém informação acerca das unidades curriculares, assiduidade, etc, a Reitoria que emite resultados do programa tutoria e avaliação parcial, aos inquéritos de bem estar (SWLS e SPANE) realizados no início do ano letivo, são várias as

¹<http://labs.psychology.illinois.edu/~ediener/SWLS.html> (consultado em 08-06-2020)

²<http://labs.psychology.illinois.edu/~ediener/SPANE.html> (consultado em 08-06-2020)

³<https://www.ua.pt/stic/PageText.aspx?id=11645> consultado em 08-06-2020

fontes utilizadas para a obtenção dos dados necessários para calcular os diferentes indicadores do FICA.

Indicador	Fórmula	Cor
IND2_1	$ECTS_InscritosAnoAtual = 0$	-
	$ECTS_FeitosAnoAtual > (ECTS_InscritosAnoAtual/2)$	Verde
	$ECTS_FeitosAnoAtual = 0 \vee$	Vermelho
	$ECTS_FeitosAnoAtual < (ECTS_InscritosAnoAtual/2)$	
IND2_2	$Propina_TotalEmDivida = 0$	Verde
	$Propina_TotalEmDivida > 0$	Vermelho
IND2_3	$BolsaSAS_EstadoBolsa = 'Deferida'$	Verde
	$BolsaSAS_EstadoBolsa = 'Em análise'$	Amarelo
	$BolsaSAS_EstadoBolsa = 'Indeferida'$	Vermelho
IND2_4	$IngressoRG_nota \geq 120$	Verde
	$IngressoRG_nota \geq 0 \wedge < 120$	Vermelho
IND2_5	$UC1_CodUC \neq null \vee \dots \vee UC8_CodUC \neq null$	Verde
	$UC1_Assiduidade \in [-1, -2] \vee \dots \vee UC8_Assiduidade \in [-1, -2]$	Amarelo
	$0 < UC1_Assiduidade < 80 \vee \dots \vee$	Vermelho
	$0 < UC8_Assiduidade < 80$	
IND2_6	Média das 5 respostas da questão 6 do inquérito SWLS = <i>null</i>	-
	Média das 5 respostas da questão 6 do inquérito SWLS ≥ 3	Verde
	Média das 5 respostas da questão 6 do inquérito SWLS < 3	Vermelho
IND2_7	SPANE Positivo / SPANE Negativo (questão 7) = <i>null</i>	-
	SPANE Positivo / SPANE Negativo (questão 7) ≥ 0.75	Verde
	SPANE Positivo / SPANE Negativo (questão 7) < 0.75	Vermelho

Tabela 3.2: Fórmulas utilizadas para calcular os principais indicadores do FICA

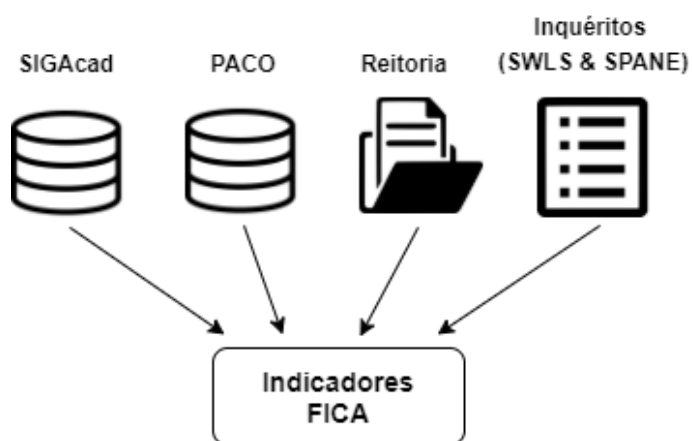


Figura 3.4: Fontes dos dados utilizados pelo FICA

A recolha e respetivo carregamento dos dados é efetuada mensalmente durante todo o ano letivo, desde o mês de Outubro até Julho, no entanto, nem todos os indicadores são atualizados com a mesma granularidade temporal. Por exemplo o IND2_1 correspondente ao aproveitamento do aluno só é calculado a partir do mês de Fevereiro (final do primeiro semestre) e respetivamente atualizado em Julho (final do ano letivo). O mesmo sucede com os

resultados obtidos nos inquéritos de bem estar, que por serem respondidos pelos estudantes apenas no início do primeiro semestre, não sofrem qualquer alteração até ao final do ano. Já os dados referentes ao número de matrículas anuladas, prescrições, bolsa, propinas, entre outros, são constantemente atualizados e carregados nos ficheiros do FICA dos meses seguintes.

3.2 ELICITAÇÃO DE REQUISITOS

Considerando o tempo e a quantidade de recursos que o desenvolvimento de qualquer projeto pode consumir, é imprescindível que exista um cuidado levantamento de requisitos antes de se proceder à sua implementação. A primeira fase de todo o ciclo de vida de um sistema deve focar-se na definição das necessidades e expectativas dos utilizadores que por sua vez irão guiar todo o desenvolvimento. Neste trabalho usou-se uma abordagem centrada no utilizador onde a fase de análise de requisitos começa precisamente pelo estudo dos utilizadores e das suas necessidades (Mayhew, 1999). A falta de sucesso de muitos projetos está frequentemente associada a uma fraca elicitação de requisitos (Goguen & Linde, 1993).

Desde introspeção, entrevistas do tipo questionário ou resposta aberta, análise de protocolos ou *focus groups*, são diversas as técnicas utilizadas no levantamento de requisitos de um projeto. Nesta dissertação, tendo em conta a natureza exploratória característica dos sistemas de visualização de informação e a cada vez mais frequente utilização de *focus groups* na avaliação de protótipos de visualização (Mazza & Berre, 2007), foi seguida esta metodologia na realização desta etapa.

3.2.1 *Focus Group*

O *focus group* é uma técnica de pesquisa qualitativa que consiste na realização de entrevistas dentro de pequenos grupos, formados por pessoas com algum tipo de semelhança e interesse, guiadas por questões abertas que privilegia a interação entre os vários participantes e tem como principal finalidade estudar as suas opiniões e perceções relativamente a um produto ou serviço permitindo determinar as reações esperadas pela população em geral. Uma característica distinguível desta abordagem relativamente aos métodos de pesquisa mais tradicionais é a existência de um moderador, frequentemente interpretado pelo próprio investigador, que tem um papel ativo na dinamização da discussão do grupo e na consequente recolha de dados (Silva et al., 2014).

O *focus group* é um método extremamente útil para explorar os pensamentos, ideias, atitudes e experiências das pessoas em relação a um tópico em particular. Um comentário de um determinado participante pode despoletar toda uma corrente de respostas dos restantes levando à aquisição de novos conhecimentos não só por parte dos participantes como também pelo próprio investigador. Não existe um número de entrevistados fixo definido, podendo os grupos variar, por exemplo, entre 3 a 12 membros. No entanto, de acordo com vários autores o número ideal recomendado é de 6 a 8 participantes (Plummer, 2017)

No decorrer desta dissertação foram realizadas no total duas sessões de *focus group* e uma entrevista por videoconferência, com alguns peritos do domínio com o objetivo de definir

de forma clara as principais questões que os futuros utilizadores do sistema pretendem ver respondidas. Estaria ainda programado mais um *focus group* que acabou por não se realizar. As entrevistas foram essencialmente estruturadas segundo as várias etapas recomendadas por (Simon, 1999) para a realização de um *focus group* de sucesso:

1. **Definir o propósito:** Essencialmente é necessário perceber o porquê da realização do *focus group*. O propósito da entrevista deve ser claro e específico de modo a definirem-se as questões mais adequadas ao tópico em estudo. É uma das fases mais importantes desta metodologia uma vez que irá guiar todo o restante planeamento. O grande objetivo das sessões realizadas foi perceber aquilo que os utilizadores procuram habitualmente saber a partir dos dados do FICA.
2. **Definir uma *timeline*:** Um *focus group* deve ser planeado com bastante antecedência – algumas semanas antes da sessão – de modo a reunir todos os elementos necessários: identificar, convidar e agendar a entrevista com os potenciais participantes, desenvolver e testar as questões e definir um local. Para cada uma das tarefas devem ser estabelecidos intervalos de tempo específicos que garantam a realização de todo o processo dentro do prazo desejado.
3. **Identificar e convidar os participantes:** Uma vez definidos os principais objetivos a alcançar, mais fácil se torna identificar os participantes a convidar. Desde a definição do número de membros do *focus group* até às características e contextos dos potenciais interessados são vários os pontos a ter em conta antes de lhes ser efetuado o convite formal. Nas sessões realizadas foi possível contar com a participação de quatro professores, do Departamento de Eletrónica e Telecomunicações e Informática (DETI) e do Departamento de Economia, Gestão, Engenharia Industrial e Turismo (DEGEIT), dos quais faziam parte dois Diretores e um Vice Diretor de Curso e um professor de uma cadeira importante do primeiro ano, com elevado número de alunos e uma percentagem de insucesso alta, e com um Assessor da Reitoria da Universidade de Aveiro.
4. **Formular as questões:** Uma vez que o *focus group* decorre dentro de um intervalo de tempo limitado – normalmente entre uma a duas horas – e como a maioria das questões são de resposta aberta não existe espaço para um grande número de perguntas. A divisão das questões é geralmente efetuada em dois tipos: introdutórias/lançamento do tema e focadas no núcleo do tema. As questões elaboradas seguiram uma lógica de mais gerais para específicas e foram formuladas de acordo com o perfil dos participantes.
5. **Gerar um *script*:** A existência de um *script* num *focus group* garante que as questões são colocadas de acordo com o contexto dos participantes ajudando o moderador a manter a sessão de acordo com o planeado e dentro do tempo previsto. Essencialmente, um *script* pode ser dividido em três partes: A abertura da sessão, onde o moderador recebe o grupo e explica o objetivo da pesquisa, a secção de questões e uma sequência final onde são discutidas as principais conclusões obtidas e prestados os devidos agradecimentos aos entrevistados. Nas entrevistas realizadas, tendo em conta o diferente conhecimento dos dados por parte dos vários participantes, começou-se por fazer uma breve apresentação do projeto FICA, orientada ao contexto de cada um, seguindo-se uma discussão geral e

a colocação de questões mais específicas acerca do tema, suportadas pela demonstração de alguns protótipos de visualizações. O *script* utilizado pode ser observado em detalhe no Apêndice A.

6. **Escolher o moderador:** O principal papel do moderador é garantir que a discussão não se dispersa e que todos os participantes são ouvidos. Deve existir especial cuidado por parte dos moderadores, com interesse profissional/académico no tópico em estudo, para não influenciar os entrevistados com a sua opinião pessoal e com isso enviesar os resultados (Plummer, 2017). As entrevistas foram moderadas pelo autor desta dissertação que contou ainda com o apoio de um dos seus orientadores e de um aluno de doutoramento com experiência em *focus groups* na monitorização das sessões.
7. **Escolher a localização:** O lugar escolhido para a realização do *focus group* deve ser de fácil acesso e oferecer aos seus participantes um ambiente de conforto e liberdade que os incentive a expressar e discutir as suas opiniões da forma mais natural possível. Tendo em conta a familiaridade da maioria dos participantes com o local, ambas as sessões presenciais decorreram numa sala de reuniões do IEETA (Instituto de Eletrónica e Telemática de Aveiro) com disposição dos lugares em estilo mesa-redonda, facilitando assim a comunicação e troca de ideias entre os intervenientes.
8. **Guiar o *Focus Group*:** Definidos os principais aspetos referidos nos pontos anteriores e reunido todo o material necessário, procede-se então ao início da sessão. Após apresentados os vários participantes, o moderador deve guiar o *focus group* de acordo com o *script* criado e gerir as várias dinâmicas de grupo envolvidas anotando/gravando todas as respostas dadas. Quando este verifica que o tópico em estudo está prestes a atingir o ponto de saturação indica ao final da sessão. Os *focus groups* realizados tiveram aproximadamente uma hora de duração.
9. **Interpretar e reportar resultados** Após a realização da sessão do *focus group* devem ser sumarizados os vários tópicos abordados e as respetivas respostas dadas. De modo a evitar lapsos de memória, esta recolha e sumarização de resultados foi efetuada logo a seguir às entrevistas e pode ser vista em detalhe no Apêndice B. Posteriormente foram analisados os sumários elaborados de modo a cruzar os dados das várias sessões, detetar padrões ou informações inesperadas e obter as respetivas conclusões definindo-se assim as principais questões às quais a solução proposta pretende dar resposta.
10. **Traduzir resultados em ação** Analisados os resultados é necessário aplicá-los. Uma boa prática dos *focus group* é manter os participantes a par de futuras ações baseadas nos resultados da sessão. Este acompanhamento vai ajudar a manter as várias pessoas interessadas no tópico e levar inclusive à sua contribuição em futuras fases do desenvolvimento/implementação do sistema em questão.

O *script* criado para as várias sessões de *focus group* foi construído essencialmente tendo em conta os principais atributos destacados nos trabalhos relacionados analisados e a informação disponibilizada pelo projeto FICA.

Após a apresentação dos participantes e uma breve contextualização do tema da dissertação, no início de cada *focus group* procurou-se saber qual a experiência de cada um dos entrevistados com os dados fornecidos pelo FICA.

Embora com granularidade diferente, ficou claro que todos os elementos presentes tinham algum conhecimento sobre os dados e, no caso dos Diretores de Curso, contactado diretamente com eles, apontando desde logo como principais problemas a sua complexidade, que desincentiva a análise, e a existência de vários campos com informação semelhante que dificultam a associação entre um aluno específico e o seu respetivo *background*. Outro fator referenciado por todos foi a quantidade de tempo despendida, muitas vezes exagerada, na realização de algumas tarefas.

No caso dos Diretores de Curso que já tinham trabalhado com os dados, os seus principais objetivos passaram sobretudo por verificar o desempenho dos alunos no curso, particularmente no final de cada semestre, detetar potenciais situações de risco a partir da identificação do incumprimento de indicadores, correlacionar atributos como assiduidade ou nota de entrada com o aproveitamento dos alunos e obter informação relativamente ao número de desistências/anulações de matrículas. Daqui surgiu desde logo um conjunto inicial de questões composto por:

- Qual foi a média das classificações médias dos alunos do curso?
- Qual foi a distribuição do aproveitamento dos alunos no curso?
- Existe alguma diferença entre o aproveitamento dos alunos no 1º e 2º semestre?
- Como foi o percurso de um determinado aluno ao longo do ano letivo?
- Quais alunos não satisfazem determinado indicador a uma dada altura?
- De que forma determinado atributo influencia o desempenho dos alunos de um curso?
- Quantos alunos anularam ou têm a matrícula suspensa num determinado mês?

Seguiu-se uma pequena discussão acerca das expectativas em relação ao que seria importante visualizar numa aplicação de exploração visual de dados académicos. Nesta fase, para além de debatidas as questões anteriormente referidas, ficou saliente a necessidade da existência de mecanismos que permitam efetuar a filtragem dos dados por atributos como ano curricular, regime, tipo de acesso, nota de entrada, etc, no processo de obtenção das suas respostas de uma forma rápida e interativa, duas características dificilmente atingíveis através do formato em que os dados do FICA são atualmente disponibilizados.

Foi ainda dedicada uma parte do *focus group* à apresentação de alguns protótipos de baixa fidelidade, como o representado na Figura 3.5, desenvolvidos em Power BI utilizando dados do FICA, com o objetivo de validar as questões anteriormente propostas e a importância dos atributos referenciados bem como avaliar a utilidade de algumas possíveis representações visuais. Daqui uma vez mais foi destacada a importância de fatores como desempenho, assiduidade, acumulação de indicadores de risco, informação sobre estado de matrícula e a possibilidade de filtragem dos dados por ano curricular, regime de estudo, entre outros, assim como a existência de uma boa legenda e ligação entre as várias visualizações de uma página.

Assiduidade vs. Desempenho

Dados Julho 2018

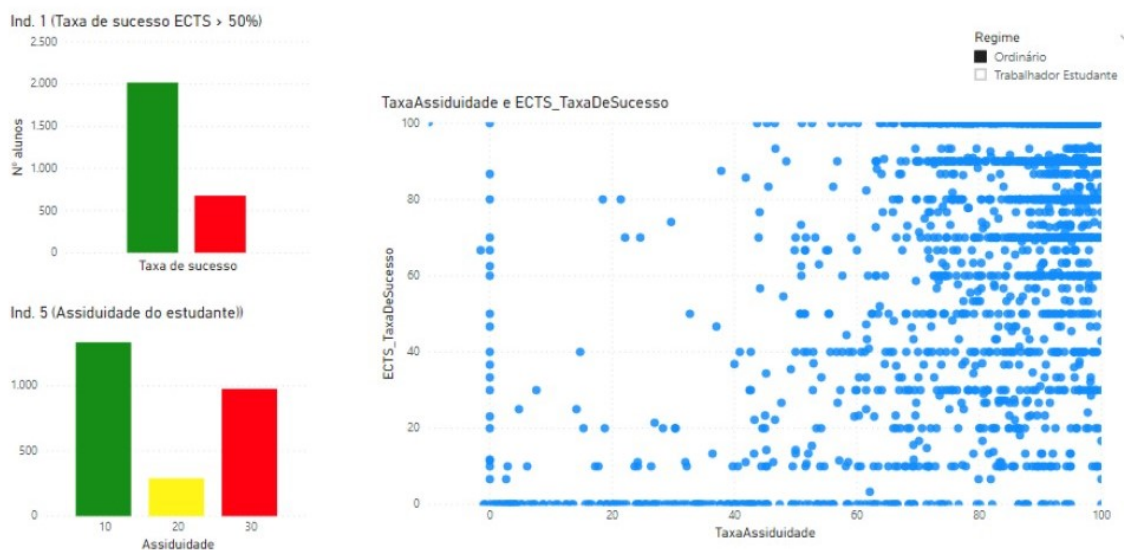


Figura 3.5: Protótipo apresentado relativamente à relação entre assiduidade e desempenho

No final das várias sessões, durante um pequeno debate sobre as conclusões obtidas, surgiram ainda algumas sugestões, que embora não sejam atualmente suportadas pelos dados disponibilizados pelo FICA, seriam interessantes considerar, entre as quais:

- Qualificar a entropia da gestão de tempo;
- Quantificar o acesso às plataformas digitais como email, *e-learning* ou PACO;
- Dependências entre unidades curriculares;
- Influência da residência atual do estudante, isto é, se está deslocado da residência habitual e da família, no seu desempenho;
- Influência da qualidade do descanso/horas de sono no aproveitamento do estudante;
- De que forma a mudança de regente de uma cadeira influencia o seu desempenho.

Um outro aspeto concluído pelas entrevistas realizadas prende-se com o facto de existirem diferentes tipos de utilizadores que poderão ter interesses dispares sobre os dados. Por exemplo, enquanto os Diretores de Curso pretendem analisar sobretudo os dados dos seus alunos, para os Diretores de Departamento ou Assessores da Reitoria será mais fácil iniciar esta observação ao nível do curso. Como tal, foram definidos três perfis de utilizador: Diretor de Curso, Diretor de Departamento e Assessor da Reitoria, sendo que cada um conta com uma personalização diferente das visualizações apresentadas, configurada de acordo com os seus interesses e informação a que tem acesso. Para uma melhor definição dos respetivos cenários e casos de utilização foi utilizado o conceito de *personas*.

3.2.2 Definição de *Personas*

O termo *persona* foi proposto por (Cooper, 1999) e é utilizado para descrever personagens fictícios que incorporam características essenciais para o *design* de um dado sistema interativo

com diferentes tipos de futuros utilizadores. A sua criação permite não só entender as suas necessidades e motivações como também os seus comportamentos e objetivos durante a utilização de um determinado produto ou serviço. A utilização desta metodologia pressupõe que já tenha sido realizada alguma investigação acerca dos utilizadores da aplicação a desenvolver de modo a tornar a experiência de interação das *personas* com o produto o mais natural e próxima da realidade possível. Todo este processo torna a fase de *design* menos complexa (Dam & Siang, 2019).

Cada *persona* deve ter a si associada um cenário que por sua vez descreve o seu contexto de utilização do sistema. Saber onde e quando ocorre determinada ação são dois dos pontos essenciais nesta caracterização. Até fazerem parte de um cenário, as *personas* não apresentam qualquer valor real pelo que devem ser considerados cenários de utilização típicos para cada *persona*.

No âmbito desta dissertação, como ponto inicial do desenvolvimento da aplicação de exploração dos dados do FICA - FICAvis - e tendo em conta os vários testemunhos e questões propostas nas sessões de *focus group* realizadas foram definidas três *personas* que têm diferentes tipos de acesso aos dados e objetivos: Diretor de Curso, Diretor de Departamento e Assessor da Reitoria. Enquanto o primeiro tem acesso apenas aos dados dos alunos do seu curso, o seguinte tem acesso à informação dos vários cursos do departamento e por fim o Assessor da Reitoria tem acesso aos dados de todos os estudantes da Universidade. A Figura 3.6 esquematiza os diferentes graus de privilégio de acesso aos dados do FICA pelas várias *personas* definidas.

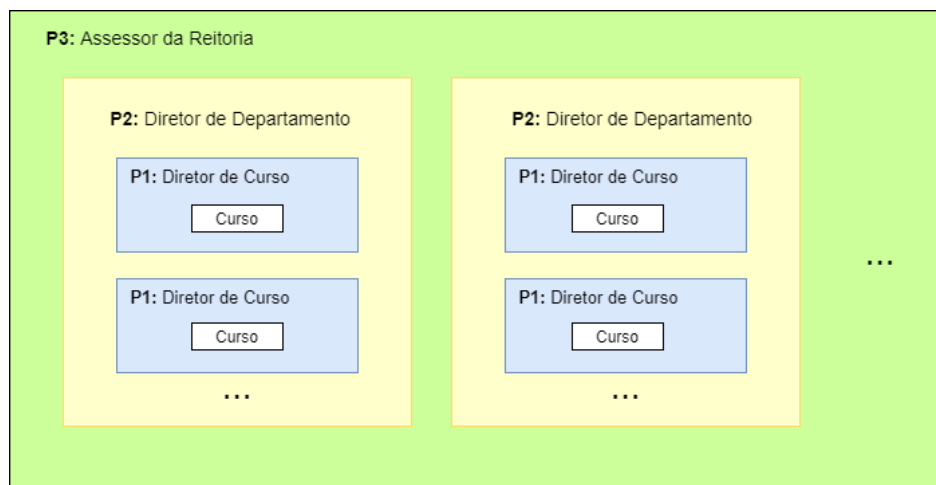


Figura 3.6: Diferentes graus de acesso aos dados do FICA pelas várias *personas* definidas

Diretor de Curso

O Roberto é professor num Departamento da Universidade de Aveiro há vários anos. Leciona uma unidade curricular de primeiro ano com muitos alunos de um curso em Engenharia. É também diretor desse mesmo curso.

Como tal, no final de cada mês recebe por email um ficheiro Excel contendo dados recolhidos pelo projeto FICA referentes a todos os alunos do curso que administra, altura em que, através

da sua análise, procura obter feedback da situação atual do curso, desde o desempenho dos alunos ao funcionamento das unidades curriculares. No entanto, devido à complexidade destes ficheiros e o facto de cada um corresponder a apenas um mês, torna-se difícil acompanhar a evolução dos alunos ao longo de todo o ano letivo. Este fator leva a que o Roberto realize apenas um estudo mais profundo em momentos chave do ano curricular, isto é, no final de cada um dos semestres. Aqui surge um novo problema, pois uma vez que esta é uma tarefa que exige um enorme dispêndio de tempo e, por ser também professor, o Roberto encontra-se sempre bastante ocupado nestes períodos devido às avaliações das cadeiras que leciona, e por isso acaba por dar prioridade unicamente à análise de alguns fatores como o aproveitamento dos alunos e a deteção de incumprimento de indicadores ignorando outras informações que lhe seriam úteis conhecer. Outro aspeto negativo desta organização é a difícil análise global do curso, por exemplo tendo em vista fins estatísticos.

Com uma nova apresentação dos dados, de forma interativa e intuitiva, para além de realizar as tarefas já referidas mais rapidamente, o Diretor de Curso poderia usufruir de um nível de detalhe dos dados muito superior ao atual conseguindo assim acompanhar a evolução do curso e o trajeto de todos os alunos ao longo de todo ano e deste modo conseguir responder a várias questões como:

- Qual a percentagem de alunos que obtiveram aprovação no curso?
- Qual é a média das classificações médias dos alunos do curso?
- De que forma a classificação de acesso ao ensino superior influencia o desempenho dos alunos?
- Quantos alunos anularam a sua matrícula durante o 1º semestre?
- Qual é a evolução de um determinado aluno que registou baixo aproveitamento?
- De que forma varia o desempenho dos estudantes relativamente a certos atributos?
- Num determinado mês:
 - Quantos alunos têm propinas em atraso?
 - Quantos alunos se encontram em situação de risco?
 - Quais são os alunos que não cumprem mais de 3 indicadores?
- Como variou a acumulação de indicadores de risco ao longo do ano?
- Em que mês se verificaram mais anulações de matrículas?
- O desempenho do curso melhorou relativamente ao último ano?

Diretor de Departamento

O Lucas é atualmente professor num Departamento de Engenharia da Universidade de Aveiro do qual é também Diretor.

A sua principal função é garantir o bom funcionamento do departamento do qual é o principal responsável. Entre outros fatores, um forte indicador de vitalidade do departamento é o sucesso dos estudantes e dos cursos que estão sob a sua responsabilidade.

Todos os meses o Lucas recebe via email um ficheiro Excel que contém dados de todos os alunos do departamento recolhidos pelo projeto FICA a partir do qual tenta estudar a sua

situação atual. O facto de os dados surgirem em ficheiros extensos que contêm dados referente a todos os alunos implica a existência de centenas de entradas o que dificulta ainda mais a análise de conjuntos específicos de alunos, como por exemplo o agrupamento destes por curso ou tipo de habilitação. Também devido aos dados de meses diferentes se encontrarem em múltiplos ficheiros, a sua observação tendo em conta uma variável contínua temporal afirma-se como uma tarefa complexa e dispendiosa a nível de tempo. A elevada quantidade de tarefas que o cargo de Diretor de Departamento implica fazem com que a análise dos dados do FICA não seja tão frequente como gostaria.

Com uma representação mais clara e interativa dos dados, o Diretor de Departamento, para além de agilizar a sua análise, poderia ter um nível de detalhe muito superior destes e conseguir responder rapidamente a questões como:

- Quais são os cursos do departamento onde se verificou melhor desempenho?
- Existe alguma relação entre o desempenho dos alunos de um curso e o tipo de ciclo de estudos no qual este está contido?
- De que forma o desempenho do curso é influenciado pelo regime de frequência dos seus alunos?
- Como evoluiu a acumulação de indicadores de risco num determinado curso ao longo do ano letivo?
- Quantas anulações de matrícula se registaram nos cursos de licenciatura do departamento? Qual foi o mês mais crítico?
- Que cursos do departamento merecem especial atenção?

Assessor da Reitoria

A Laura é Doutorada numa Engenharia e Gestão Industrial e trabalha atualmente nos Serviços da Reitoria da Universidade de Aveiro. É uma das pessoas responsáveis por analisar ao pormenor os dados do FICA estando em constante comunicação com os STIC, responsáveis pela recolha e envio dos dados.

Embora esteja bastante familiarizada com a formato como os dados são armazenados, a existência de uma solução que agregue toda a informação e apresente resultados para questões às quais geralmente tenta responder seria de enorme utilidade para a Laura, que no final de cada ano letivo elabora um relatório com base na análise em detalhe dos dados do FICA.

Como trabalha na Reitoria, os dados do FICA recebidos pelo Assessor da Reitoria são referentes a todos os alunos e portanto a grande maioria das questões para as quais procura resposta são colocadas tendo em consideração todos os estudantes da Universidade:

- De que forma variou o desempenho global dos alunos?
- Quais são os ciclos de estudo com pior aproveitamento?
- De que forma o regime de acesso ao ensino superior dos estudantes influencia o seu desempenho?
- Como evoluiu a acumulação de indicadores de risco dos estudantes ao longo do ano?

- Quais os meses mais críticos quanto ao nível de anulações de matrículas?
- Quais foram as áreas científicas/departamentos onde se registou pior desempenho?

Dando continuidade ao processo de desenvolvimento de um sistema de exploração visual de dados, após estudado o conjunto de dados a utilizar e definidas as principais questões a responder pelos diferentes tipos de utilizador, o passo seguinte consiste na proposta de uma solução que permita satisfazer os requisitos estabelecidos nesta fase.

Solução Proposta: FICAvís

O grande foco do FICAvís - Visualização de Dados para Combate ao Abandono Escolar na Universidade de Aveiro - passa por permitir aos seus futuros utilizadores obter respostas para as questões levantadas nos *focus groups*, apresentados no capítulo anterior, de uma forma rápida e interativa. Pretende-se com estas visualizações não só agilizar as consultas atualmente efetuadas sobre os dados do FICA como também permitir uma análise mais detalhada dos mesmos.

Neste capítulo são apresentadas e justificadas as principais opções tomadas na implementação da solução proposta, desde o pré-processamento dos dados, caracterização das tecnologias utilizadas, até à escolha das visualizações.

O FICAvís, desenvolvido em Power BI, consiste num conjunto de diferentes versões de relatórios destinados aos três perfis de utilizador anteriormente definidos. Um relatório Power BI é composto por múltiplas páginas que por sua vez contêm vários elementos visuais relacionados entre si permitindo facilmente extrair informação a partir dos dados (Ferrari & Russo, 2016).

Na secção seguinte é descrito todo o processo de criação das diversas páginas presentes nos relatórios desenvolvidos. Os dados utilizados no desenvolvimento desta solução encontram-se devidamente anonimizados por questões de privacidade.

4.1 PRÉ-PROCESSAMENTO DOS DADOS

Como referido no final do estudo de *benchmarking* realizado na secção 2.3, o Power BI foi a plataforma que mais se ajustava no contexto deste trabalho. O facto de ser uma *framework* que oferece componentes para suportar as diversas etapas de um projeto torna possível a sua utilização ao longo de todo o processo de desenvolvimento.

O primeiro passo para poder explorar um conjunto de dados consiste na sua aquisição. Os dados do programa FICA são enviados pelos sTIC para os diferentes utilizadores mensalmente

sendo que a informação disponibilizada é filtrada de acordo com o perfil e funções de cada um.

Uma vez que os dados chegam aos seus utilizadores com uma estrutura idêntica (mesmas colunas e atributos), com o Power BI, através do editor da *Power Query*, é possível combinar todos os ficheiros Excel e agregá-los numa única tabela lógica. Apenas é necessário especificar a localização dos ficheiros e definir qual das páginas do primeiro Excel contém os dados necessários para a análise. Logo de seguida o Power BI automaticamente cria uma consulta que efetua os restantes passos necessários para a extração dos dados numa única tabela. Os passos aplicados nesta transformação são todos registados facilitando a possível reformulação ou adição de novas etapas intermédias em qualquer altura sem a necessidade de revisitar o processo desde o início. Também qualquer alteração relativamente à fonte dos dados, por exemplo a adição de um ficheiro do FICA com dados de um novo mês à localização dos restantes, é imediatamente considerada pelo Power BI que garante a sincronização constante do conjunto de dados.

Encontrando-se toda a informação dos ficheiros reunida numa só tabela, a partir da *Power Query* e das várias tarefas de consulta comuns disponibilizadas, foram realizados alguns passos de pré-processamento sobre os dados que, entre outros aspetos, envolveram:

- Remoção e filtragem de colunas com dados em excesso;
- Alteração do tipo de dados de algumas colunas;
- Anulação da dinamização de colunas (*unpivot*): criação de pares atributo-valor ao nível dos dados das colunas dos sete principais indicadores do FICA;
- Alteração de valores com vista à normalização de alguns dados, etc.

Para além das operações realizadas ao nível da *Power Query*, antes de avançar para o desenvolvimento das diferentes páginas de relatório, com recurso a algumas fórmulas da coleção de funções DAX, mostradas na Figura 4.1, foi criada uma tabela calculada do tipo calendário.

```
Data =  
VAR BaseCalendar = CALENDARAUTO()  
RETURN  
    GENERATE (  
        BaseCalendar,  
        VAR BaseDate = [Date]  
        VAR YearDate = YEAR ( BaseDate )  
        VAR MonthNumber = MONTH ( BaseDate )  
        VAR MonthName = FORMAT ( BaseDate, "mmm" )  
        VAR MonthShort = FORMAT ( BaseDate, "MMM" )  
        VAR YearMonthName = FORMAT ( BaseDate, "MMM yyyy" )  
        RETURN ROW (  
            "Dia", BaseDate,  
            "Ano", YearDate,  
            "Mês Número", MonthNumber,  
            "Mês", MonthName,  
            "Mês Short", MonthShort,  
            "Ano Mês", YearMonthName  
        )  
    )
```

Figura 4.1: Expressão DAX utilizada para a criação da tabela "Data"

Contendo colunas com informação relativa ao dia, mês e ano e estabelecendo uma relação de um para muitos com a tabela global, a tabela *Data* permite a análise dos dados do FICA tendo em conta uma dimensão temporal.

4.2 DESENVOLVIMENTO DO PROTÓTIPO

Na sequência das entrevistas realizadas e com base nos resultados obtidos foi criado, também em Power BI, um primeiro protótipo de um relatório de baixa fidelidade que foi apresentado a alguns representantes dos diferentes perfis de utilizador e alvo de uma avaliação heurística por parte de um avaliador experiente de modo a recolher *feedback* para o desenvolvimento da solução proposta. Os principais aspetos alterados como resultado deste processo são detalhados na secção 5.1 do próximo capítulo.

Com o objetivo de responder às questões dos vários tipos de utilizadores, foi criado para cada perfil uma versão de um relatório em Power BI com diferentes páginas compostas por diferentes componentes e visualizações orientadas aos requisitos e necessidades de cada um.

4.2.1 Estrutura das Páginas

De modo a facilitar a análise dos dados tornando-a num processo rápido e interativo, todas as páginas desenvolvidas exibem uma estrutura idêntica composta por três principais divisões como se ilustra na Figura 4.2:



Figura 4.2: Estrutura das páginas do relatório

- **A: Barra superior** - Na parte superior de cada página, para além do título está também presente informação relativa ao curso que está a ser visualizado, como nome, habilitação, média e número total de alunos. Enquanto os primeiros dois atributos são

estáticos, tanto o valor da média como o número de alunos é dinamicamente atualizado de acordo com as interações efetuadas sobre os restantes elementos da página.

- **B: Barra lateral** - Dado o destaque atribuído à necessidade de mecanismos de filtragem dos dados durante o levantamento de requisitos, foi incluído no lado esquerdo de cada uma das páginas um painel com vários filtros, adaptados ao contexto dos dados em questão, que permite segmentar a informação existente e assim realçar determinadas características. Desde a escolha do período, seleção do ano letivo, ano curricular ou regime, são vários os ajustes que o utilizador pode fazer nos elementos de cada página. O botão *Limpar Seleção* anula todos os filtros configurados e qualquer seleção efetuada sobre as visualizações restabelecendo o estado inicial da página.
- **C: Conteúdo da página** - O espaço deixado livre por **A** e **B** é preenchido por diferentes visualizações adequadas ao contexto de cada página que estão correlacionadas entre si. Através do uso do rato, o utilizador pode clicar sobre partes específicas das representações, fazendo-as interagir entre si, ou obter informação adicional na forma de um *tooltip* ao passar o cursor sobre elas. De igual modo, ao passar o rato sobre o símbolo ? presente no canto superior direito de cada uma das visualizações, é possível obter, também na forma de um *tooltip*, informação mais detalhada acerca de cada elemento como é exemplificado na Figura 4.3.

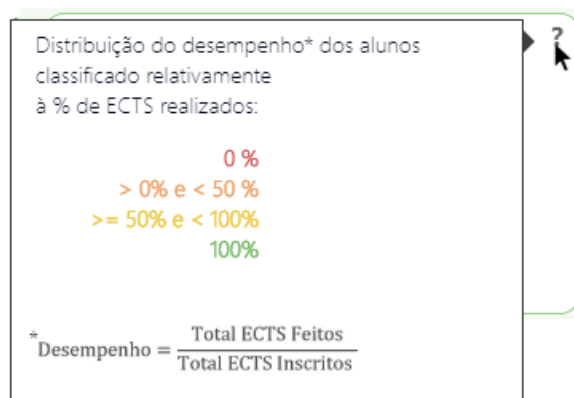


Figura 4.3: *Tooltip* com informação sobre um elemento

- **C1: Lista de Alunos** - Em todas as páginas que agrupam dados de vários alunos, geralmente localizada no canto inferior direito, existe uma tabela com a lista de todos os estudantes que estão a ser considerados pelos elementos visuais da página e que fazem parte do conjunto de dados em exibição, resultante dos mecanismos de filtragem usados permitindo ao utilizador identificar grupos específicos de alunos. A partir desta lista, ao selecionar um dado estudante com o botão direito do rato, é possível escolher a opção *Pormenorizar* que por sua vez encaminha o utilizador para a página pessoal do aluno onde pode consultar em detalhe os seus dados, evolução de indicadores e progresso académico.

Foi adotado um esquema de cores que tem o verde como predominante pelo facto de esta ser normalmente a cor associada ao universo da Universidade de Aveiro.

Nome	ECTS Inscritos	Desempenho	Média
Aluno 1087	30,00	80%	15,75
Aluno 138	24,00	75%	11,00
Aluno 1394	30,00	100%	11,80
Aluno 1395	30,00	80%	12,75
Aluno 14	24,00	0%	13,25
Aluno 1400	30,00	80%	11,25
Média	29,17	86%	12,54

Figura 4.4: Lista de alunos para uma determinada interação e seleção da ação *Pormenorizar*

4.2.2 Perfil de Diretor de Curso

Os dados recebidos pelo Diretor de Curso dizem respeito apenas aos alunos do curso que administra.

A versão do relatório desenvolvida para os utilizadores com este perfil é composta por seis páginas:

- **Página Inicial:** Identificação do perfil do utilizador e visão geral das páginas existentes no relatório, com a possibilidade de navegação até cada uma delas.
- **Desempenho:** Perspetiva global sobre o desempenho dos alunos do curso, nomeadamente ao nível de ECTS realizados.
- **Indicadores:** Acumulação mensal dos 7 principais indicadores do FICA descritos na secção 3.1 e de anulações de matrícula.
- **Evolução:** Evolução ao longo do ano de indicadores de risco e comparação do desempenho médio dos estudantes no curso ao longo dos anos.
- **Comparação:** Comparação do desempenho dos alunos por diferentes atributos.
- **Página pessoal:** Informação pessoal e percurso académico de cada um dos estudantes.

De seguida, são descritas em detalhe todas as páginas que fazem parte desta versão do relatório.

Página Inicial

A página inicial representada na Figura 4.5 tem como principal objetivo oferecer ao utilizador uma contextualização do que é possível analisar a partir do relatório.

Para além da sigla, e respetivo significado, do programa FICA, é apresentado ao utilizador informação acerca do seu perfil e dos dados aos quais tem acesso e uma lista das páginas disponíveis. Para cada uma das páginas é dada uma breve descrição do seu conteúdo e finalidade, sendo possível navegar diretamente até à mesma. As páginas encontram-se ordenadas segundo o seu grau de importância, deduzido a partir dos resultados obtidos nos *focus groups* realizados. As setas, que contêm as hiperligações para navegar pelo relatório, seguem a orientação da esquerda para a direita de modo a destacar esta ordem.

O utilizador pode também, em qualquer altura, permutar entre as várias páginas a partir da barra de separadores do próprio Power BI.



Figura 4.5: Página Inicial do relatório para o perfil de Diretor de Curso

Desempenho

A página *Desempenho*, exibida na Figura 4.1, oferece uma perspetiva global do aproveitamento dos estudantes do curso.

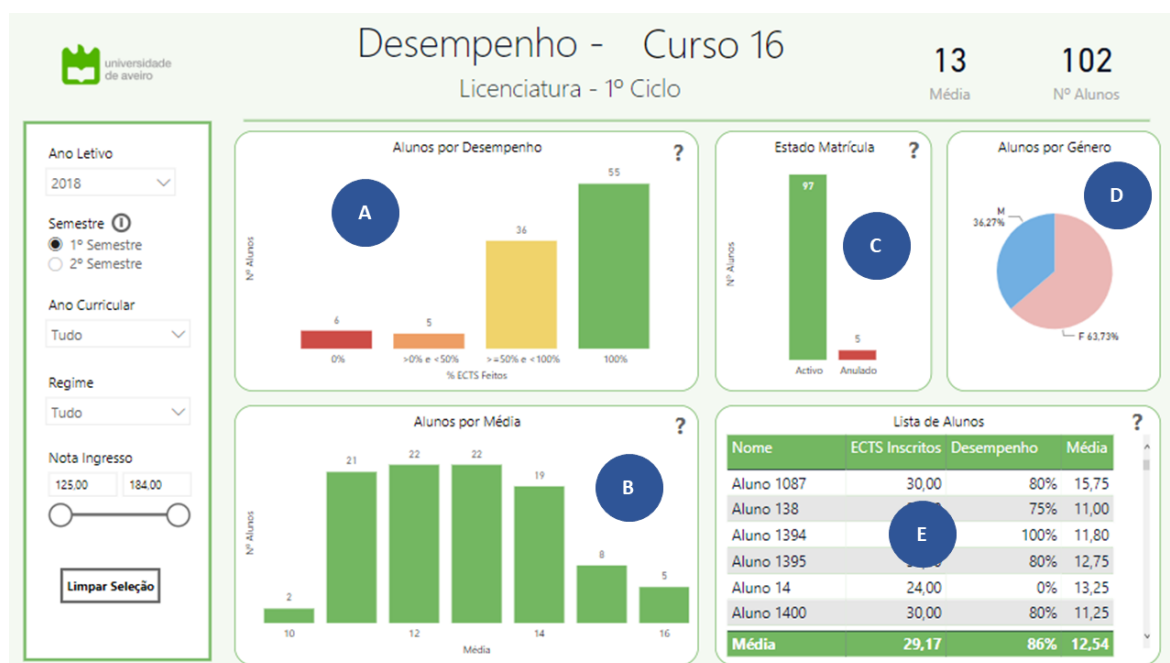


Figura 4.6: Página do Desempenho do Curso

A barra superior permite desde logo perceber qual é a média e o número total de alunos do curso e a barra lateral oferece ao utilizador a possibilidade de definir a dimensão temporal dos

dados, através da escolha do ano letivo e período (final do 1º ou 2º semestre) correspondente, e filtrá-los por atributos específicos dos alunos, como ano curricular, regime ou nota de ingresso. Uma vez que a configuração dos filtros é feita com base nos dados existentes, o *slicer* relativo à nota de ingresso oferece, ainda que de forma indireta, informação sobre a classificação do último colocado e a nota máxima verificada entre os alunos que ingressaram no curso pelo regime geral.

O gráfico de barras **A** mostra a distribuição do número de alunos do curso pelo seu desempenho, que corresponde à taxa de ECTS por si realizados e é calculado pela fórmula:

$$Desempenho = \frac{ECTSFeitosAnoAtual}{ECTSInscritosAnoAtual} \quad (4.1)$$

De forma a tornar a sua representação mais percetível e uniforme, os valores referentes ao desempenho dos estudantes foram agrupados e classificados em quatro diferentes taxas de sucesso, de acordo com a percentagem de ECTS feitos equivalente: 0%,]0%, 50%[, [50%, 100%[e 100%. Para além de ordenadas pelo desempenho, as barras do gráfico estão mapeadas segundo uma escala de cores, que varia entre vermelho e verde, associada ao risco de cada um dos casos, que facilita a interpretação por parte do utilizador através do realce de informação.

$$Cor = \begin{cases} \text{Vermelho} & \text{se } 0\% \\ \text{Laranja} & \text{se }]0\%, 50\%[\\ \text{Amarelo} & \text{se } [50\%, 100\%[\\ \text{Verde} & \text{se } 100\% \end{cases} \quad (4.2)$$

Adotando a lógica da visualização anterior, o gráfico de barras **B** permite ver a distribuição das médias dos alunos no curso. A média de um estudante corresponde à nota média das unidades curriculares por si concluídas, não sendo por isso considerados neste elemento os estudantes com média igual a zero.

Esta representação pode, por exemplo, explicar a média do curso através da moda ou, a partir da interação com a visualização anterior, a correlação da taxa de ECTS realizados com a média verificada.

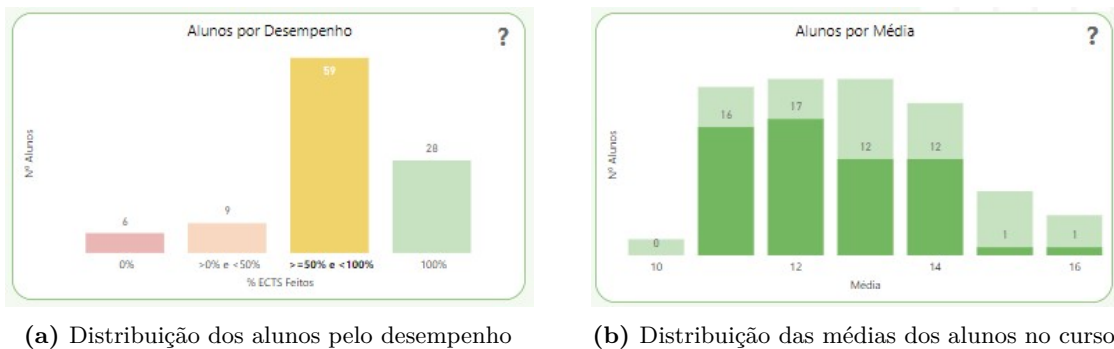


Figura 4.7: Interação entre as visualizações A e B: distribuição das médias dos alunos que tiveram desempenho entre [50%, 100%[

A Figura 4.7 mostra um destes exemplos de interação ao efetuar a análise da distribuição das médias dos alunos que tiveram desempenho dentro do intervalo [50%, 100%] .

Os elementos **C** e **D** contemplam uma vertente mais estatística dos dados. Enquanto o primeiro mostra uma visão geral do comportamento dos alunos do curso a partir do estado atual da sua matrícula, codificado pela cor, verde, amarelo ou vermelho, o segundo apresenta uma distribuição dos estudantes por género na forma de um *pie chart*, também codificado pelas cores habitualmente a si associadas.

Por fim, a tabela **E** com a lista de alunos constantemente adaptada à configuração atual da página, exhibe atributos relevantes para o contexto da mesma como o número de ECTS inscritos, desempenho e média.

Indicadores

A página de análise mensal de indicadores do curso permite explorar o comportamento dos estudantes em relação aos sete principais indicadores do FICA - sucesso, propinas, bolsa, nota de ingresso, assiduidade e inquéritos SWLS e SPANE - e situações de risco geralmente a si associadas.

As barras superior e lateral da página são semelhantes à anterior, sendo que a principal diferença consiste na possibilidade de filtragem dos dados por mês em alternativa ao final de cada semestre. Uma vez que os atributos em causa são atualizados mensalmente, a diminuição desta janela temporal torna possível a deteção de padrões de alunos em situação de risco e a consequente atuação e tomada de medidas em qualquer fase do ano letivo e não só nos períodos de avaliação.

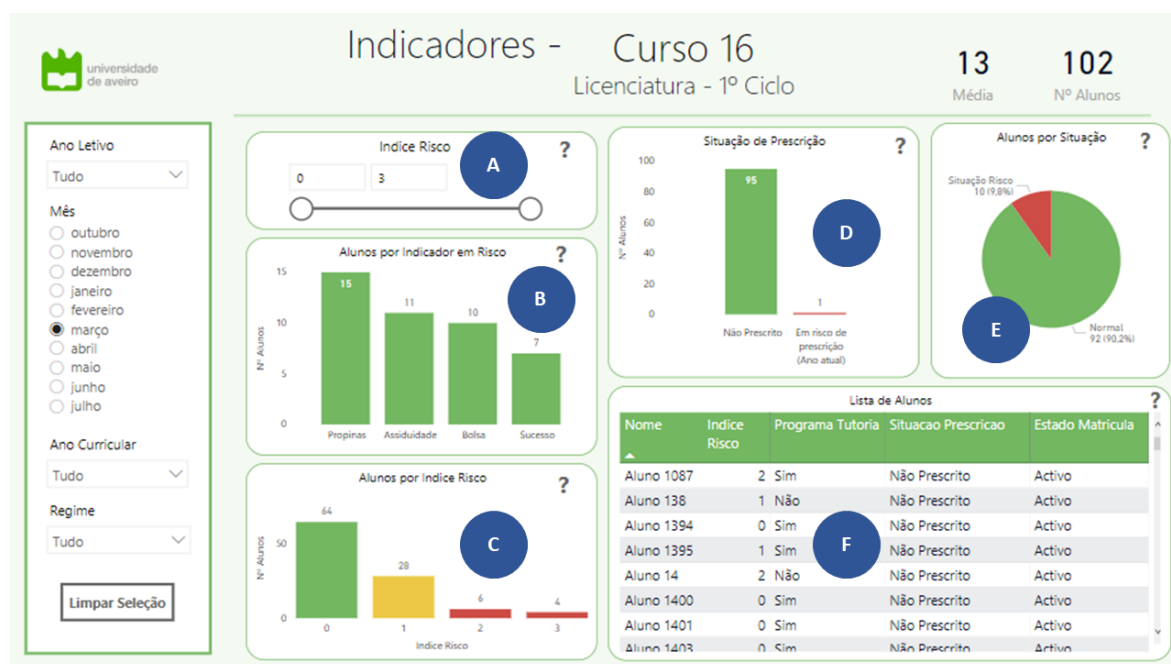


Figura 4.8: Página de Análise Mensal de Indicadores

Adjacente ao painel de filtros, o elemento **A** é um *slicer* que permite segmentar todo o conteúdo da página pelo índice de risco dos alunos. Considerando os sete principais indicadores do FICA, este atributo corresponde ao número máximo de indicadores que um estudante não satisfaz. Os valores do filtro encontram-se compreendidos entre o valor mínimo e máximo do índice de risco verificado nos dados.

Diretamente relacionada com esta medida, a visualização **C** mostra a distribuição do número alunos do curso pelo índice de risco ajudando o utilizador a detetar possíveis situações críticas de alunos que não cumprem mais do que um dos sete indicadores. De modo a realçar os casos mais críticos, as barras do gráfico são mapeadas pela seguinte escala:

$$Cor = \begin{cases} \text{Verde} & \text{se índice de risco} = 0 \\ \text{Amarelo} & \text{se índice de risco} = 1 \\ \text{Vermelho} & \text{se índice de risco} \geq 2 \end{cases} \quad (4.3)$$

O gráfico de barras **B** representa a acumulação de indicadores de risco por parte dos estudantes do curso no mês selecionado. A partir daqui é possível responder a questões como quantos alunos não satisfazem determinado indicador, qual é o indicador mais crítico em determinado período, entre outras. Apenas são representados indicadores para os quais existem alunos que não os satisfazem e o mesmo aluno pode ser contabilizado em mais do que uma das barras já que pode não cumprir ao mesmo tempo múltiplos indicadores.

Ainda referente ao número de indicadores, o *pie chart* **E** exhibe a comparação entre o número de alunos que estão em possível situação de risco, isto é, que possuem mais de um dos sete indicadores FICA por satisfazer, e os restantes.

O gráfico de barras **D**, embora não represente diretamente nenhum dos indicadores, apresenta a distribuição dos alunos por um atributo geralmente associado ao seu incumprimento: situação de prescrição. Um estudante prescrito fica impedido de continuar a frequentar o curso, em que estava matriculado e inscrito, durante dois semestres consecutivos.

A tabela **F** que completa a da página e permite pormenorizar em detalhe cada um dos alunos da lista, para além do seu índice de risco, contém informação acerca da sua presença ou não no programa tutoria, a sua situação de prescrição e o estado da sua matrícula.

Evolução

A página da evolução do curso, mostrada na Figura 4.9, permite efetuar uma comparação mensal dos sete principais indicadores do FICA, estudar a acumulação de matrículas anuladas e observar a evolução da média de desempenho dos estudantes do curso ao longo dos anos.

Os dados apresentados são referentes a todo o período de aulas do ano letivo selecionado na barra lateral. No caso de consulta da página para um ano atualmente em curso, é exibida apenas informação até ao último mês para o qual existem registos.

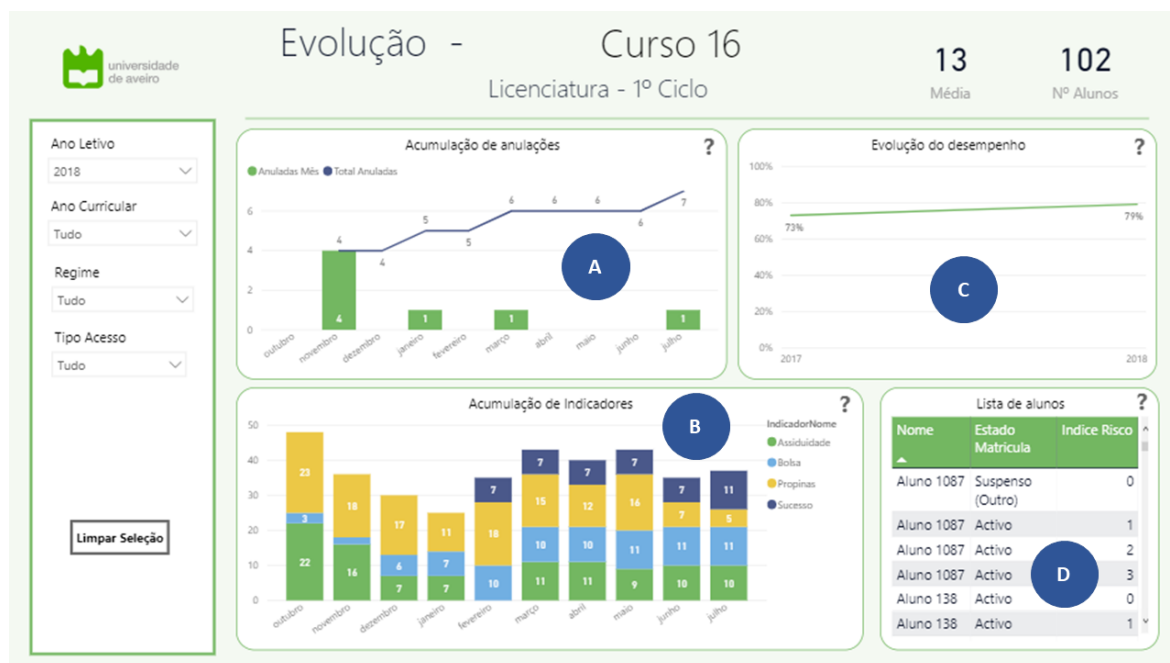


Figura 4.9: Página da Evolução do Curso

A visualização **A** consiste num gráfico combinado de barras e linha que mostra a acumulação de anulações de matrícula ao longo de todo o ano letivo. Enquanto as barras correspondem ao número de anulações efetuadas num determinado mês, a linha indica o número total acumulado de matrículas anuladas. A partir desta representação é possível, por exemplo, perceber qual o mês onde se verificam mais anulações de matrícula ou quantos alunos abandonaram o curso no decorrer do ano letivo.

O gráfico de barras sobrepostas **B** mostra a evolução dos sete principais indicadores do FICA ao longo dos vários meses do ano letivo. À semelhança da página anterior, aqui é também possível perceber quais são os períodos mais críticos para cada um dos indicadores levando à ponderação e adoção de novas medidas para combater estes números.

Por fim, o gráfico de linha **C** apresenta a média do desempenho dos alunos do curso ao longo dos vários anos para os quais existe informação nos dados.

A tabela **D** que completa o conteúdo da página contém como atributos o estado da matrícula e o índice de risco de cada um dos estudantes considerado pela configuração e seleção atual.

Comparação

Um dos pontos destacados durante os *focus groups* realizados foi a influência de certos atributos no desempenho dos estudantes. Destaca-se sobretudo informação relativa ao seu contexto de ingresso no ensino superior e regime de frequência. Este facto é aliás confirmado por alguns estudos já efetuados a partir dos dados do projeto FICA (Moreira & Costa, 2018).

A página de comparação é dedicada essencialmente à comparação do desempenho dos

estudantes do curso por quatro diferentes atributos: fase e opção de ingresso, tipo de acesso e regime de frequência. A Figura 4.10 mostra a configuração da página que permite comparar o desempenho dos alunos do curso relativamente à sua fase de ingresso.

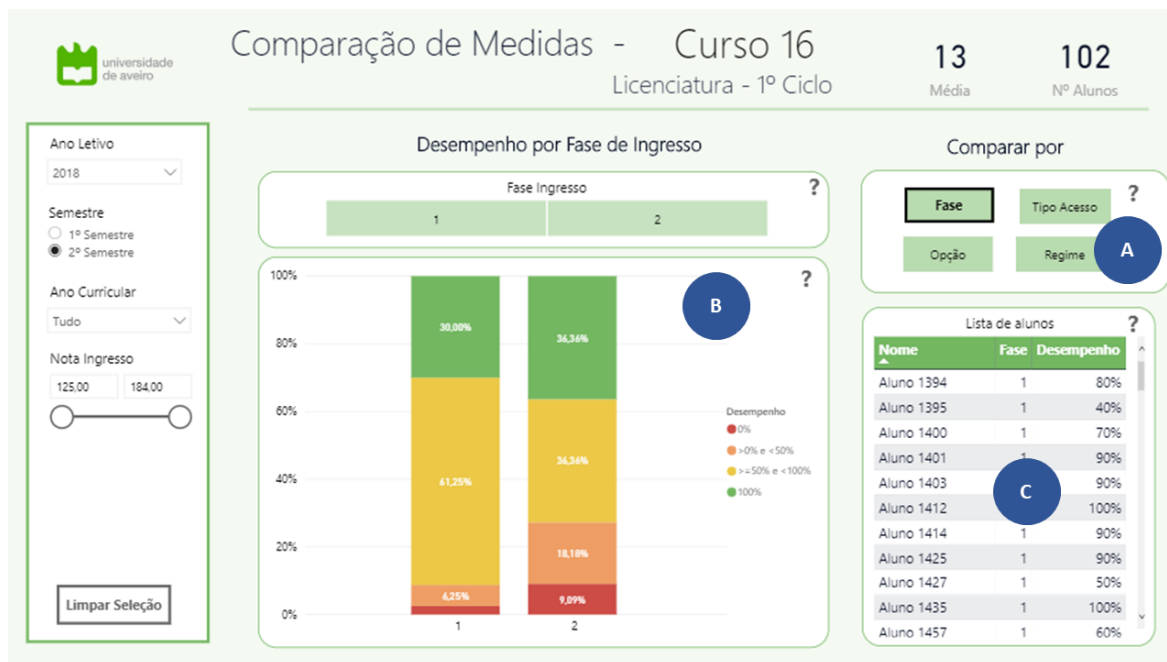


Figura 4.10: Página de Comparação do Desempenho (por fase)

Os botões agrupados em **A** permitem ao utilizador selecionar o atributo relativamente ao qual pretende comparar o desempenho dos alunos. Logo de seguida, todo o conteúdo da página é atualizado segundo esta seleção.

O gráfico de colunas 100% empilhadas **B** permite comparar o desempenho dos alunos pelas várias opções existentes no atributo atualmente selecionado.

Inicialmente são exibidas no gráfico todas as opções existentes, no entanto, a partir do menu localizado acima da visualização, o utilizador pode optar por filtrá-la de forma por valores específicos. Na Figura 4.11 estão representadas as quatro possíveis versões deste elemento visual.

Por forma a tornar a comparação mais intuitiva, os valores apresentados são relativos, sendo que o total cumulativo de cada coluna empilhada é sempre igual a 100%, contudo, como se pode verificar na Figura 4.12, ao passar o cursor do rato sobre cada uma das barras, é possível obter na forma de um *tooltip* o número absoluto de alunos correspondente e a sua percentagem relativamente ao total de estudantes do curso.

Para cada uma das comparações efetuadas, são tidos em conta apenas os alunos que têm informação relativa ao atributo escolhido. Por exemplo, se o objetivo é comparar os alunos por fase de ingresso, apenas são considerados os estudantes que ingressaram no ensino superior pelo regime geral.

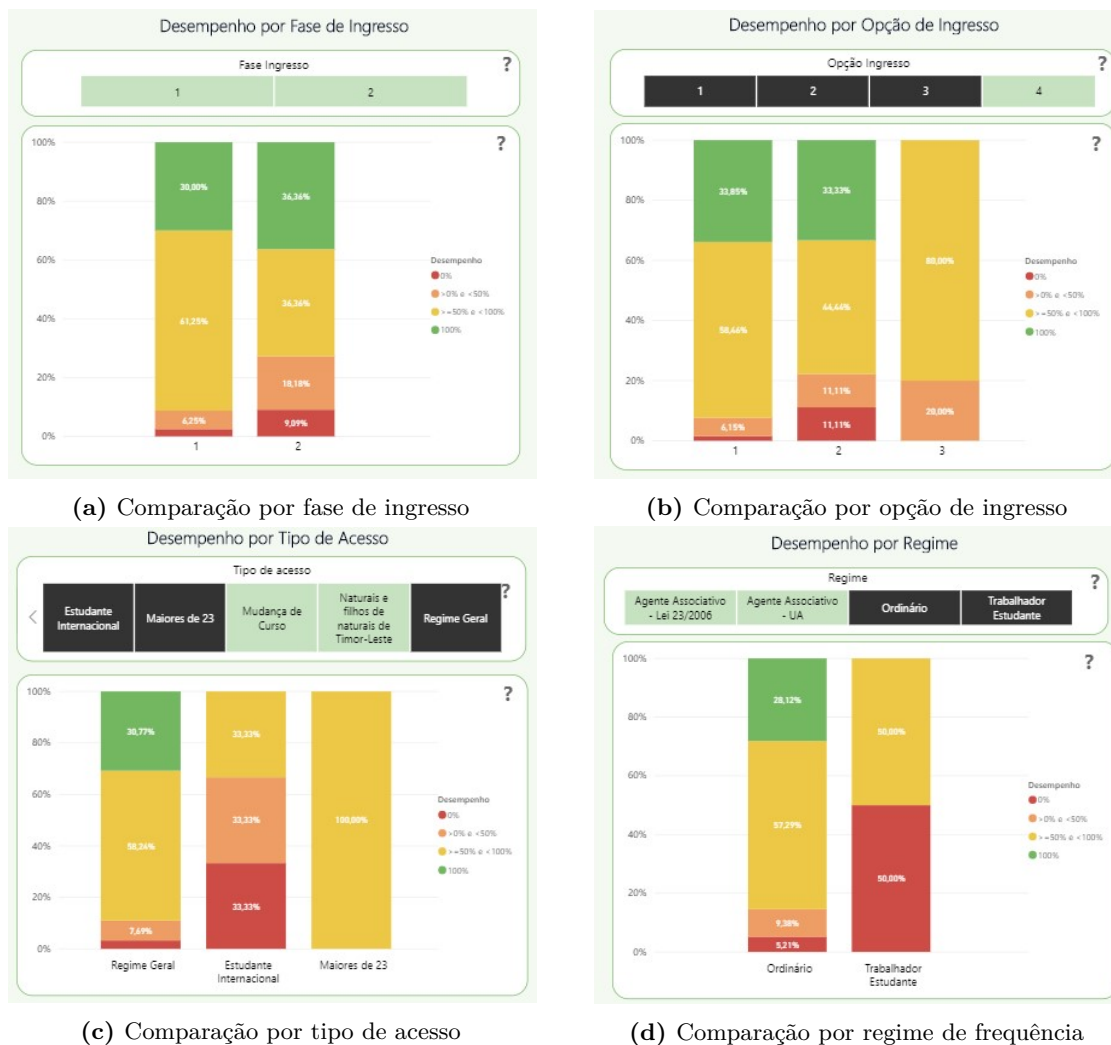


Figura 4.11: Diferentes tipos de comparação do desempenho dos alunos

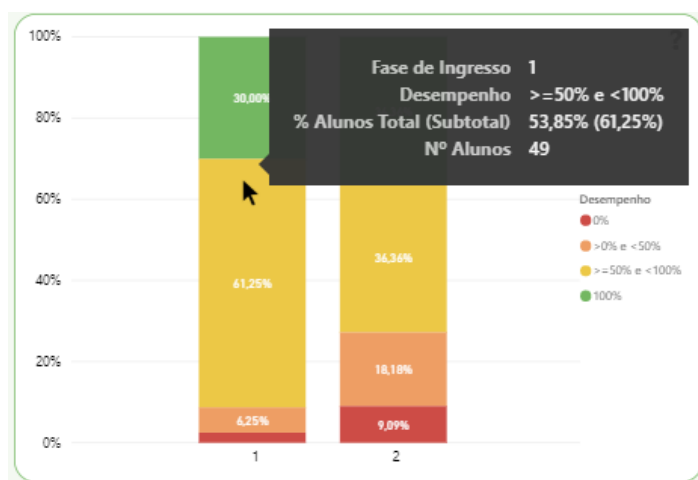


Figura 4.12: *Tooltip* com detalhes dos alunos da primeira fase com desempenho [50%,100%]

O cálculo do desempenho e o mapeamento de cores são os mesmos utilizados na página do desempenho do curso referidos respetivamente nas expressões 4.1 e 4.2 e as barras encontram-se ordenadas pelo número de alunos presente em cada uma das opções.

Uma vez mais, a tabela **C** é adaptada à informação atualmente exibida e tem como colunas para além do nome do aluno e do seu desempenho, informação relativa ao atributo atualmente em análise.

Página Pessoal

A página pessoal, representada na Figura 4.13, apresenta em detalhe os dados pessoais e o percurso académico de um determinado estudante do curso.



Figura 4.13: Página Pessoal do Estudante

Embora com a mesma finalidade, existem duas versões desta página: a primeira surge como resultado da ação *pormenorizar* efetuada sobre qualquer uma das listas de alunos das páginas referidas anteriormente e a segunda, acessível em qualquer altura, permite efetuar a pesquisa e respetiva análise da informação pessoal de um determinado aluno a partir do seu número mecanográfico.

Ao contrário das restantes páginas, na barra superior encontra-se agora presente o nome do aluno selecionado, a sua média e o estado atual da matrícula.

Os elementos **A** e **B** apresentam informação relativa às unidades curriculares a que o aluno está inscrito e estão diretamente relacionados. Enquanto a tabela identifica as disciplinas inscritas em cada semestre pelo seu respetivo código, o gráfico de linhas mostra a assiduidade do estudante (em percentagem) ao longo do ano letivo nestas disciplinas.

A matriz **C** monitoriza o comportamento do aluno perante os sete principais indicadores do FICA ao longo do ano letivo. É utilizado o mesmo mapeamento de cores utilizado nos ficheiros Excel originais descrito na secção 3.2. Esta visualização permite ao utilizador facilmente perceber quais foram os meses em que o aluno não cumpriu determinado indicador e, quer pela análise isolada da matriz, quer pelo cruzamento com outros elementos da página, detetar possíveis padrões de risco no seu percurso.

Do lado direito da página, para além do aproveitamento do aluno no final do ano letivo, calculado pelo número de ECTS feitos em relação ao número de ECTS inscritos, o utilizador pode também verificar pelo gráfico de linhas **D** como foi a sua evolução ao longo dos semestres.

Por fim, ao passar o cursor sobre o elemento **E**, é possível obter na forma de um *tooltip*, exemplificado na Figura 4.14, informação adicional sobre o aluno contendo dados pessoais, dados de ingresso e evolução da sua presença no programa tutoria e estado de matrícula ao longo do ano.



Figura 4.14: *Tooltip* com informação adicional do aluno

4.2.3 Perfil de Diretor de Departamento

O Diretor de Departamento tem acesso aos dados de todos os estudantes do departamento que é responsável.

A principal diferença da versão do relatório desenvolvida para utilizadores com este tipo de perfil em relação ao Diretor de Curso é a possibilidade de análise de múltiplos cursos.

A estrutura das páginas do relatório é semelhante à descrita na secção 4.2.1 sendo que agora existe a possibilidade de pesquisa e seleção de múltiplos cursos através da barra lateral e a informação relativa ao nome do curso e respetiva habilitação presente na barra superior é atualizada de acordo com o número de cursos escolhidos.

As tabelas com a lista de alunos das várias páginas têm agora uma coluna adicional com indicação do curso de cada estudante.

Ano Letivo
2018

Ano Curricular
Tudo

Semestre
☒ 1º Semestre
☐ 2º Semestre

Cursos selecionados

Procurar

☐ Curso 31
☒ Curso 32
☐ Curso 33
☒ Curso 34
☐ Curso 35
☒ Curso 36
☒ Curso 37
☐ Curso 38

Limpar Seleção

Comparação de Medidas - Vários Cursos
Total de 4 cursos selecionados

(b) Título da página atualizado de acordo com a seleção

(a) Seleção de múltiplos cursos

Figura 4.15: Estrutura das barras lateral e superior nas páginas do diretor de departamento

Visão Global do Desempenho

Por forma a oferecer ao Diretor de Curso uma perspetiva global do desempenho de todos os alunos do departamento foi criada a página representada na Figura 4.16.



Figura 4.16: Visão Global do Desempenho dos alunos do Departamento

Dada a necessidade de apresentar numa só visualização múltiplos atributos de uma grande quantidade de dados hierárquicos, o elemento **A**, representado na forma de um *treemap*, agrega toda a informação dos alunos do departamento em três diferentes níveis de detalhe: habilitação, curso e aluno.

Em cada nível, os dados de cada hierarquia são agregados em retângulos, organizados por tamanho, desde a parte superior esquerda até à parte inferior direita e cuja dimensão corresponde ao número total de alunos. A cor de cada um é mapeada pela média do desempenho dos alunos no curso, calculado pela expressão 4.1, através da escala:

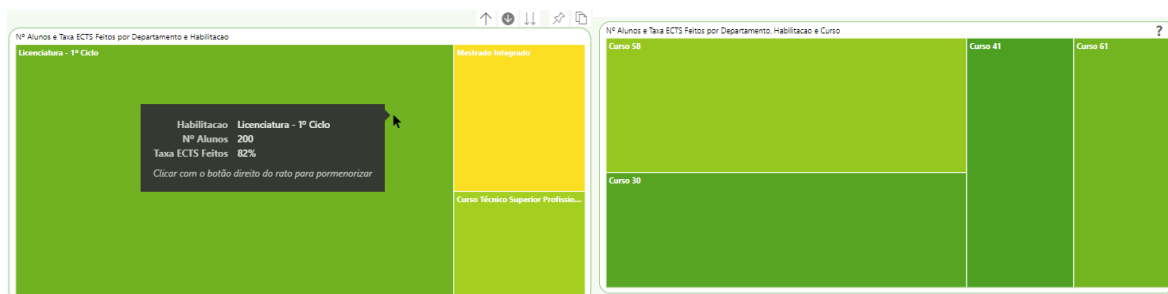


Figura 4.17: Escala de cor utilizada no mapeamento do desempenho

Ao passar o cursor sobre cada um dos campos é possível obter na forma de um *tooltip* o resultado exato destas duas medidas.

O cabeçalho do *treemap* oferece ao utilizador a possibilidade de realizar uma das seguintes três ações:

- **Desagregar:** ver detalhes adicionais da hierarquia selecionada. Como é exemplificado na Figura 4.18, ao escolher uma determinada habilitação, o utilizador pode ter uma visão global do desempenho dos alunos nos diferentes cursos que a constituem.



(a) Diferentes habilitações do departamento

(b) Cursos do tipo Licenciatura - 1º ciclo

Figura 4.18: Ação de desagregar: habilitação → curso

- **Agregar:** agrupar todos os valores atualmente visíveis no *treemap* em hierarquias mais genéricas (voltar ao nível anterior).
- **Ir para o nível seguinte:** ver todos os valores do nível seguinte. Por exemplo, passar da vista das habilitações, para uma visão global de todos os cursos, sem distinção do ciclo de estudos.

Esta visualização oferece ainda a possibilidade de navegar entre as diferentes páginas do relatório. Ao escolher um dos campos do *treemap*, os botões que compõem o menu **B**, inicialmente desativados, passam a estar ativos, e permitem ao utilizador dirigir-se até à respetiva página, devidamente configurada pelo agrupamento de dados da hierarquia

selecionada. Esta navegação pode também ser realizada através do clique com botão direito do rato sobre o conjunto de dados desejado, seguido da escolha da ação *Pormenorizar*.

4.2.4 Perfil de Assessor da Reitoria

Os utilizadores que trabalham na Reitoria têm acesso aos dados de todos os estudantes da universidade.

Embora envolva uma amostra mais abrangente dos dados, as principais questões a responder pelos serviços de reitoria são semelhantes às dos restantes perfis. A versão do relatório desenvolvida para os utilizadores com este perfil é por isso idêntica à destinada aos Diretores de Departamento. A diferença mais significativa encontra-se na página com a visão global do desempenho dos alunos que agora inclui no *treemap* um nível adicional de detalhe que permite ter uma perspetiva global do desempenho dos alunos da UA pela sua área CNAEF correspondente, como se pode verificar na Figura 4.19.

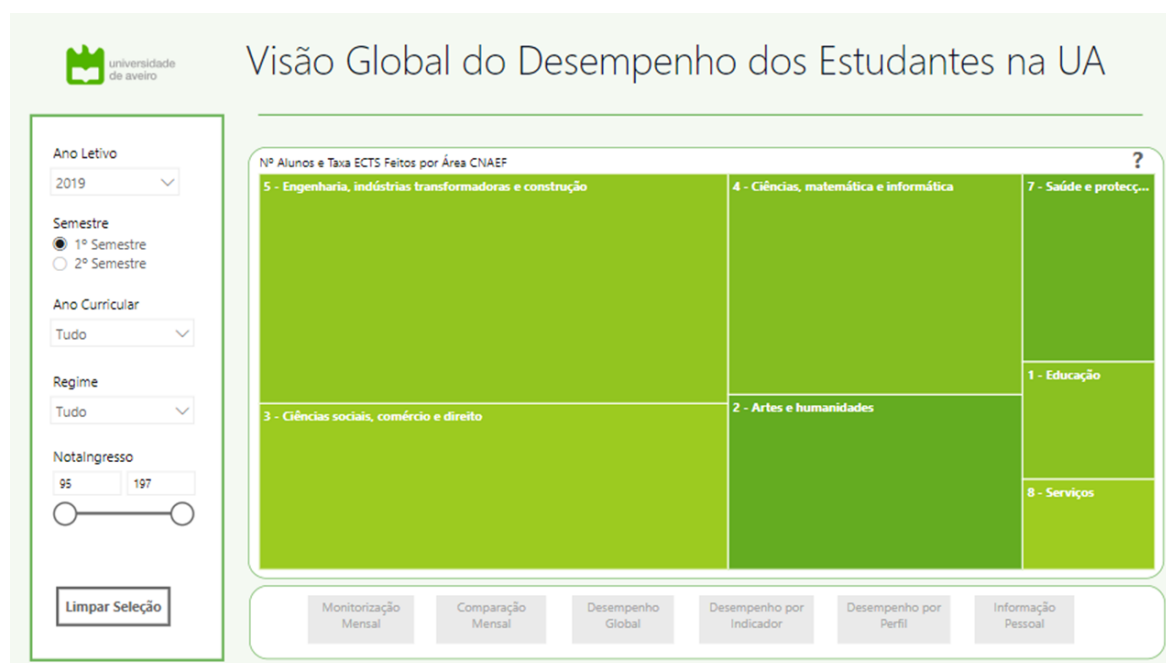


Figura 4.19: Visão Global do Desempenho dos alunos na UA

Uma vez que até ao ano letivo 2019/20 o projeto FICA não incluía nos dados informação da área CNAEF de um curso, para anos anteriores, em vez da área CNAEF o protótipo desenvolvido incluía na hierarquia mais alta do *treemap* uma lista de todos departamentos da Universidade.

4.2.5 Questões de Desenvolvimento

O grande foco da solução proposta passa por agilizar a análise dos dados do FICA e ajudar a responder às principais questões levantadas pelos utilizadores. Ao longo de todo o processo de criação das diferentes páginas, foram adotadas várias estratégias que tornam estes objetivos atingíveis pela maioria dos utilizadores, independentemente da sua posição ou experiência.

Apesar de ser uma das ferramentas mais completas no mercado atual de BI, o Power BI é um serviço relativamente recente e possui por isso algumas limitações a nível de recursos que podem dificultar a criação de soluções interativas. De seguida são descritas as técnicas que foram utilizadas para contornar algumas destas situações.

Alterar os atributos dos gráficos dinamicamente. O Power BI não oferece ainda recursos que permitam alterar de forma direta o atributo de um gráfico após o relatório ser publicado. Uma alternativa para fazer esta substituição dinamicamente é através do uso de marcadores. Os marcadores capturam a vista atual de uma página guardando todos os filtros e estados dos elementos visuais. O Power BI possui também um painel de visualização que permite ver uma lista de todos os elementos da página e controlar se estão ou não visíveis. A combinação destes dois recursos, complementada pela sua ligação a um botão, torna possível a alteração dos atributos de uma visualização de forma interativa mesmo após a publicação do relatório.

Na solução proposta, este método é utilizado, por exemplo na página de comparação do desempenho dos alunos por diferentes atributos, detalhada na secção 4.2.2. No início, são criadas na mesma página todas as versões da visualização com os diferentes atributos a comparar. De seguida, através do painel de seleção, são colocados visíveis apenas os elementos referentes a um atributo e ocultados os restantes. Após esta ação é guardado o estado da página num marcador que é depois associado a um botão. Por fim é repetido o mesmo processo para os restantes atributos.

Os marcadores são também utilizados nos botões *Limpar Secção* presentes nas várias páginas, onde cada um tem a si associado o estado inicial da respetiva página.

Ordenação de Campos. Na maioria das visualizações, a ordem pela qual os dados são apresentados é um indicador fundamental para uma análise clara por parte do utilizador.

Embora o Power BI ofereça a opção de ordenação, ascendente ou descendente, para a maioria dos seus visuais, esta apenas funciona para dados numéricos ou alfabéticos. Por exemplo na primeira visualização da página do desempenho, descrita na secção 4.2.2, a classificação utilizada não se enquadra em nenhum destes dois conjuntos.

Para tornar estes itens ordenáveis, foi necessário efetuar algumas alterações ao nível do modelo de dados. Inicialmente foi criada uma nova tabela, onde é definido o índice de cada elemento, que por sua vez tem uma relação de 1 para muitos com a tabela principal, representada na Figura 4.20. De seguida, na tabela principal é criada uma nova coluna calculada que associa a cada aluno da tabela original o índice da nova tabela através da expressão DAX

$$DesempOrdem = RELATED(Desempenho[Ordem]) \quad (4.4)$$

Por fim, a coluna que contém a classificação do desempenho do aluno é ordenada pela coluna com o índice do desempenho resultado da expressão 4.4.

Intervalo	Ordem	Desempenho	DesempenhoClass	DesempOrdem
0%	1	100%	100%	4
>0% e <50%	2	0%	0%	1
>=50% e <100%	3	90%	>=50% e <100%	3
100%	4	80%	>=50% e <100%	3
		90%	>=50% e <100%	3
		90%	>=50% e <100%	3
		100%	100%	4
		50%	>=50% e <100%	3
		100%	100%	4
		50%	>=50% e <100%	3
		0%	0%	1

Figura 4.20: Relação entre a tabela criada e a tabela principal

Campos de informação dinâmicos. Nas versões do relatório dirigidas aos perfis de Diretor de Departamento e Assessor da Reitoria, com o objetivo de tornar clara a origem da informação mostrada, foram criados campos de informação dinâmicos.

O Power BI permite a criação de campos calculados, denominados por medidas, que devolvem valores resultado de expressões DAX.

O campo situado abaixo do título de cada página é assim dinâmico resultado das duas medidas presentes na Figura 4.21. Enquanto a primeira devolve o número de elementos filtrados, a segunda referenciando-a, utiliza uma função do tipo *SWITCH* para determinar que informação apresentar: no caso de seleção única, o ciclo de estudos correspondente ao curso e nos restantes, o número de cursos selecionado.

`CursosSelecionados = IF(ISFILTERED('Global'[CursoNome]), DISTINCTCOUNT('Global'[CursoNome]), -1)`

(a) Expressão DAX que retorna o número de cursos selecionado

`Descricao =
SWITCH(
TRUE(),
[CursosSelecionados] = 1, FIRSTNONBLANK('Global'[Habilitacao], 'Global'[Habilitacao]),
[CursosSelecionados] > 1, "Total de " & [CursosSelecionados] & " cursos selecionados",
""
)`

(b) Expressão DAX que retorna a informação a apresentar

Figura 4.21: Conjunto de expressões DAX utilizados nos campos de informação dinâmicos

Avaliação

Existem vários métodos para avaliar uma aplicação de exploração visual de dados, no entanto, nem todos oferecem garantias de resultados que permitam realçar as suas potencialidades e limitações. Para obter conclusões úteis relativamente a uma solução de visualização é essencial que as avaliações efetuadas decorram dentro de um contexto real e com utilizadores que dominem o tema (Plaisant, 2004).

Neste capítulo descreve-se a avaliação de usabilidade feita ao protótipo apresentado anteriormente, através da utilização de um método de avaliação analítico, avaliação heurística, e de um método de avaliação empírico, testes de usabilidade (Dix et al., 2004).

5.1 AVALIAÇÃO HEURÍSTICA

A avaliação heurística é um método de avaliação que tem como principal objetivo detetar possíveis problemas de usabilidade num sistema, nomeadamente ao nível da sua interface de utilizador. É normalmente realizada de forma individual e por avaliadores com experiência na área da visualização que percorrem a interface múltiplas vezes de modo a avaliar os seus elementos segundo princípios de usabilidade previamente definidos: heurísticas. As avaliações heurísticas realizadas foram baseadas nas dez heurísticas de usabilidade propostas por (Nielsen, 1995).

Uma das principais vantagens deste modelo de avaliação é a possibilidade da sua utilização nas várias fases de um projeto. No contexto desta dissertação, foram realizadas avaliações heurísticas em dois momentos: antes do início do desenvolvimento da solução proposta no capítulo anterior e após terminada a sua primeira versão.

A avaliação heurística inicial foi realizada por um utilizador experiente na área de Visualização de Informação sobre o protótipo de um relatório composto por várias visualizações de modo a encontrar possíveis problemas de desenho ou usabilidade e prevenir a sua propagação durante o processo de desenvolvimento da nova versão.

Desta primeira avaliação, que pode ser consultada em detalhada no Apêndice C, foram desde logo destacados e corrigidos, entre outros, aspetos como:

- Consistência dos títulos das páginas com o nome dos respetivos separadores;
- Inexistência de uma estrutura nos elementos das páginas: muita informação sem uma organização específica;
- Falta de redundância na navegação entre páginas, nomeadamente em relação à *Página Pessoal*;
- Pouca utilização de cores com caráter subjetivo;
- Falta de informação sobre a origem de algumas fórmulas de cálculo;
- Inexistência de um botão para limpar as filtrações e seleções;
- Títulos de elementos complexos e pouco explicativos.

A Figura 5.1 mostra a página referente ao desempenho dos alunos de um curso antes e depois das alterações aplicadas resultado desta avaliação heurística.

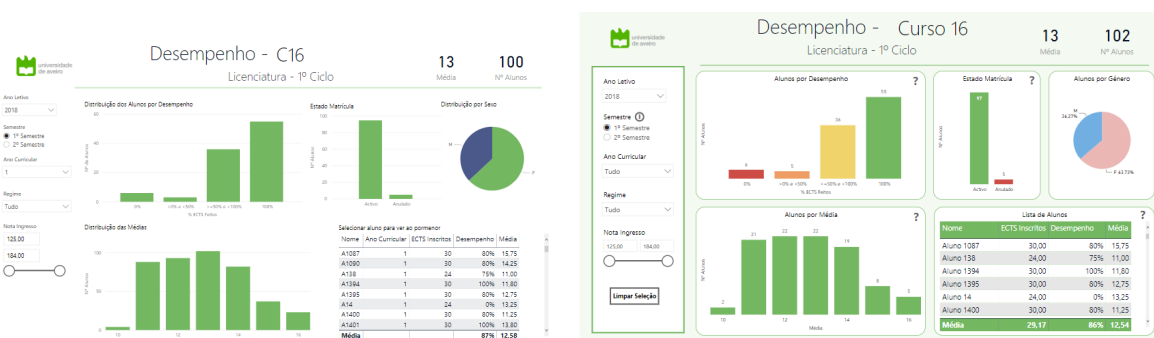


Figura 5.1: Página *Desempenho* antes e depois da Avaliação Heurística

A segunda fase de avaliações heurísticas ocorreu já depois do desenvolvimento do primeiro protótipo do FICAvi e foi realizada na totalidade por seis diferentes utilizadores experientes na área de Interação Humano-Computador (IHC), individualmente e sem contato entre si de maneira a garantir resultados independentes e imparciais.

Considerando as 10 heurísticas de Nielsen e aplicando uma escala de severidade de 0 a 4, onde 0 corresponde à inexistência de problemas de usabilidade e 4 a problemas críticos, foram adotadas duas diferentes metodologias: na primeira, realizada por quatro dos utilizadores, o avaliador estudava cada uma das páginas do relatório segundo as 10 heurísticas de Nielsen atribuindo-lhe o respetivo grau de severidade e na segunda, efetuada pelos restantes, o avaliador após explorar o protótipo da solução proposta, indicava os problemas de usabilidade encontrados e o seu respetivo grau de severidade. No Apêndice D é possível consultar em pormenor as heurísticas aplicadas, a escala utilizada e os resultados obtidos nas várias avaliações heurísticas realizadas.

De uma forma global as diferentes avaliações mostraram resultados construtivos e positivos em relação à solução proposta confirmando a utilidade das várias páginas criadas e a consistência entre as visualizações nelas utilizadas.

A Tabela 5.1 mostra os principais problemas de usabilidade da interface detetados em múltiplas avaliações, classificados pelos avaliadores com um grau de severidade de valor 3 ou superior, juntamente com a informação de qual a heurística de Nielsen em causa.

Problema	Heurística	Severidade					
		Aval.1	Aval.2	Aval.3	Aval.4	Aval.5	Aval.6
Prob. 1: Opção 'Ver Mais Informação do Aluno'	N1: Visibilidade do estado do sistema; N3: Controlo e liberdade do utilizador	1	0	4	0	3	3
Prob. 2: Botão de retroceder	N3: Controlo e liberdade do utilizador	-	1	3	-	3	-
Prob. 3: Falta de informação após navegar entre páginas	N6: Reconhecimento em vez de recordação	-	-	3	-	-	-

Tabela 5.1: Principais problemas de usabilidade detetados nas avaliações heurísticas

A seguir são descritas as alterações efetuadas de modo a corrigir estes problemas:

- **Problema 1:** Na *Página Pessoal* do aluno, a maioria dos avaliadores entendeu o elemento *mouse over* utilizado para mostrar mais dados do estudante como sendo um botão tendo por isso tentado clicar em vez de passar apenas o cursor sobre este ficando assim com a ideia errada de que a opção não funcionava. De modo a tornar esta funcionalidade mais clara foi adicionada uma pequena legenda abaixo do elemento que indica ao utilizador o que fazer, como é ilustrado na Figura 5.2.



Figura 5.2: Opção Ver Mais Informação do Aluno antes e depois da avaliação heurística

- **Problema 2:** Antes das avaliações heurísticas nem todas as páginas possuíam um botão para voltar à página anterior. Este foi um dos problemas de usabilidade mais destacado pelos vários avaliadores. Tendo em conta este fator, o FICAvi possui agora, localizado no canto superior esquerdo de cada página, um botão de retroceder que permite ao utilizador voltar à página anterior. Também ao clicar no logotipo da Universidade de Aveiro presente em todas as páginas, é possível em qualquer momento voltar à *Página Inicial*.
- **Problema 3:** Anteriormente, no processo de navegação entre páginas pelo *treemap* presente na página da *Visão Global do Desempenho* apenas era mostrada na nova página informação do nível da hierarquia escolhida pelo utilizador no *treemap*. De forma a identificar com maior detalhe o *drill down* realizado, a barra superior da página inclui agora informação de todas as seleções anteriormente efetuadas sobre o *treemap*, como exemplificado na Figura 5.3, oferecendo um contexto mais específico dos dados apresentados na página.

Desempenho - Todos os Cursos
do ciclo de estudos selecionado

Desempenho - Todos os Cursos
Dept. Ciências; Licenciatura - 1º Ciclo

Figura 5.3: Subtítulo da página resultado da navegação antes e depois da avaliação heurística

Foram ainda efetuadas algumas modificações de modo a corrigir alguns problemas de usabilidade identificados de acordo com a heurística *N10: Ajuda e documentação* que consistiram nomeadamente na adição de informação nos botões de navegação e nos *tooltips* de ajuda das várias visualizações contribuindo assim para uma melhor usabilidade da interface da solução proposta.

Apesar da avaliação heurística permitir detetar alguns problemas de usabilidade mais simples, é necessário a adoção de outras técnicas de avaliação, sobretudo com futuros utilizadores reais, que permitam complementar esta avaliação e antecipar as suas possíveis dificuldades. Para este efeito podem ser realizados, por exemplo, testes de usabilidade.

5.2 TESTES DE USABILIDADE

De modo a avaliar a capacidade de resposta do FICAVIS às questões levantadas nos *focus groups*, foram realizados dois testes de usabilidade com utilizadores peritos do domínio: um Diretor de Curso e um Assessor da Reitoria, ambos com conhecimento e experiência na análise dos dados do FICA, embora que com privilégios de acesso diferentes.

Nas próximas secções são apresentados os procedimentos e resultados obtidos com os testes e as consequentes alterações efetuadas sobre o protótipo da solução anteriormente proposta.

5.2.1 Protocolo Experimental

Os testes foram realizados com dois participantes representantes de dois diferentes grupos de futuros utilizadores, um Diretor de Curso e um Assessor da Reitoria, por videoconferência com recurso à partilha de ecrã e tiveram aproximadamente duração de uma hora cada. Por terem participado ambos nas sessões de *focus group* realizadas, os utilizadores já conheciam o contexto da dissertação e da solução proposta levando apenas a uma breve introdução das etapas e objetivos do teste.

Inicialmente, e com o utilizador na página principal do FICAVIS, foi-lhe feita uma pequena apresentação do seu conteúdo e dada liberdade para, durante alguns minutos, explorar as várias páginas e funcionalidades.

Após esta fase inicial, para cada página do protótipo, foram colocadas aos utilizadores algumas questões sobre a página atual de modo a avaliar a sua utilidade, detetar possíveis elementos em falta, obter sugestões de melhorias e validar as suas funcionalidades. De

seguida foi-lhes solicitada a realização de duas tarefas que implicavam a utilização de várias visualizações e interações executáveis a partir da página em questão.

Durante a execução das tarefas, utilizou-se o método *think aloud*, com os utilizadores a relatar todos os passos que seguiam, permitindo assim avaliar o seu desempenho e identificar potenciais problemas de usabilidade (Dix et al., 2004).

Os testes foram realizados na versão do relatório desenvolvida para o perfil de Diretor de Curso, com informação de apenas um curso, e envolveram as seguintes tarefas:

Considerando o ano letivo mais recente para o qual existem registos nos dados (2018):

- **Desempenho**

1. Quantos alunos não fizeram quaisquer ECTS no primeiro semestre? Quais foram esses estudantes?
2. No final do ano letivo, quantos alunos do primeiro ano curricular matriculados em regime ordinário fizeram todos os ECTS a que estavam inscritos?

- **Indicadores**

3. No mês de Março, qual foi o maior índice de risco verificado entre os estudantes?
4. No mês de Maio, quantos alunos do primeiro ano tinham as propinas em atraso?

- **Evolução:**

5. Qual foi o mês para o qual se registaram mais atrasos no pagamento de propinas?
6. Quantas anulações de matrículas se contabilizaram no final do ano letivo? Qual foi o mês mais crítico?

- **Comparação**

7. No primeiro semestre, como variou o desempenho dos alunos em relação à sua opção de ingresso? Qual opção registou melhor percentagem de aproveitamento?

- **Página Pessoal**

8. Considerando o estudante com o número mecanográfico 3096:
 9. Quantos indicadores não satisfazia no mês de Julho?
 10. Esteve presente no programa tutoria? Se sim, durante que meses?

5.2.2 Resultados e Principais Alterações

Todas as dez questões foram executadas com sucesso por parte dos utilizadores e num intervalo de tempo reduzido espaço de tempo, não tendo a duração total excedido os cinco minutos, confirmando assim a facilidade de utilização da solução proposta e a sua acessibilidade para novos utilizadores.

Em relação às observações feitas pelos utilizadores, antes e durante as tarefas, foram registados alguns comentários e sugestões, que podem ser consultados em pormenor no Apêndice E, com o objetivo de tornar o relatório ainda mais claro e intuitivo. De seguida são descritas e justificadas algumas das principais alterações efetuadas de acordo com estas apreciações.

A primeira alteração consistiu na mudança dos títulos das páginas e respetivas descrições, em concordância com as sugestões dos utilizadores:

- Desempenho → Desempenho Global
- Indicadores → Monitorização Mensal
- Evolução → Comparação Mensal
- Comparação de Medidas → Desempenho por Perfil
- Página Pessoal → Informação Pessoal

Dada a importância atribuída durante os testes à informação apresentada em cada uma das páginas foi também adotada uma nova ordem de disposição das páginas, como se pode verificar pela Figura 5.4.



Figura 5.4: Página Inicial para o perfil de Assessor da Reitoria (Nova disposição das páginas)

Um aspeto em falta evidenciado pelos utilizadores foi a inexistência de uma forma de comparação do desempenho dos estudantes por indicadores como bolsa ou propinas. De modo a cumprir este requisito, foi criada uma nova página, *Desempenho por Indicador*, que, adotando a mesma estrutura da já existente *Desempenho por Perfil*, permite a comparação do desempenho dos alunos em relação a quatro dos principais indicadores do FICA: propinas, bolsa, nota de ingresso e assiduidade.

Para além dos mecanismos de filtragem já presentes na barra lateral das diferentes páginas, foram sugeridos outros atributos de segmentação dos dados, no entanto, devido ao espaço reduzido desta secção, optou-se por manter na barra apenas os filtros relativos ao período da informação e ano curricular e acrescentou-se um botão que permite aceder ao menu flutuante, composto por outras múltiplas opções de filtragem, representado na Figura 5.5.

Painel de Seleção

Contexto Acadêmico

Regime
Tudo

Tempo Parcial
Tudo

Primeira Vez Curso
Tudo

Programa Tutoria
Tudo

Dados Ingresso

Tipo Acesso
Tudo

Fase
Tudo

Opção
Tudo

Nota Ingresso
125 184

Limpar Seleção

Figura 5.5: Menu flutuante com múltiplas opções de filtragem dos dados da página

Apesar de ambos os utilizadores terem conseguido realizar com sucesso a tarefa 1, que implicava identificar um conjunto específico de estudantes a partir da tabela da página, verificou-se que a atualização da lista de alunos não foi imediatamente percebida. Para corrigir este problema, tornou-se o título das tabelas dinâmico, incluindo constantemente a informação de quantos alunos estão a ser considerados como se pode verificar na Figura 5.6.

Lista de alunos

Nome	Estado Matricula	Indice Risco
Aluno 14	Activo	2
Aluno 1771	Activo	1
Aluno 2014	Activo	3
Aluno 256	Suspenso (Outro)	3
Aluno 2988	Anulado	1
Aluno 686	Activo	1

→

Total de 7 Estudantes

Nome	Estado Matricula	Indice Risco
Aluno 14	Activo	2
Aluno 1771	Activo	1
Aluno 2014	Activo	3
Aluno 256	Suspenso (Outro)	3
Aluno 2988	Anulado	1
Aluno 686	Activo	1
Aluno 766	Activo	?

Figura 5.6: Substituição do título das tabelas por informação dinâmica

A página de *Comparação Mensal*, descrita na secção 4.9, foi a que sofreu mais alterações relativamente à versão do relatório anterior aos testes. Com base nas sugestões dadas pelos utilizadores, para além da reordenação e alteração do título dos elementos, a visualização referente à acumulação mensal de indicadores, anteriormente representada por um gráfico de barras sobrepostas, foi substituída por um gráfico de linhas que mostra a mesma informação de forma mais clara, existindo ainda a possibilidade de a visualizar através de uma tabela. As duas opções de representação destes dados estão exemplificadas na Figura 5.7.

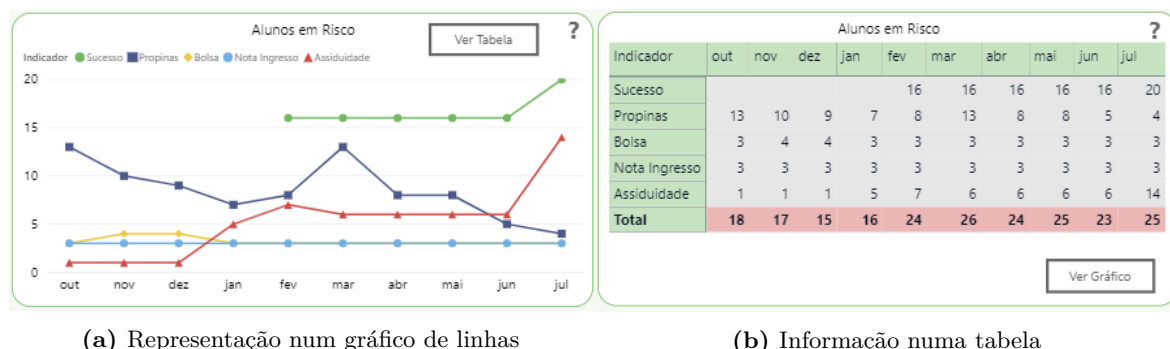


Figura 5.7: Comparação Mensal de alunos em risco

Por fim, na página *Informação Pessoal*, os utilizadores alertaram para o facto de nem todos os indicadores serem atualizados com a mesma granularidade. De modo a prevenir possíveis leituras erradas da acumulação de indicadores do aluno, a matriz que mostra a sua evolução ao longo do ano foi dividida em três categorias: indicadores mensais, semestrais ou anuais como se pode verificar pela Figura 5.8.

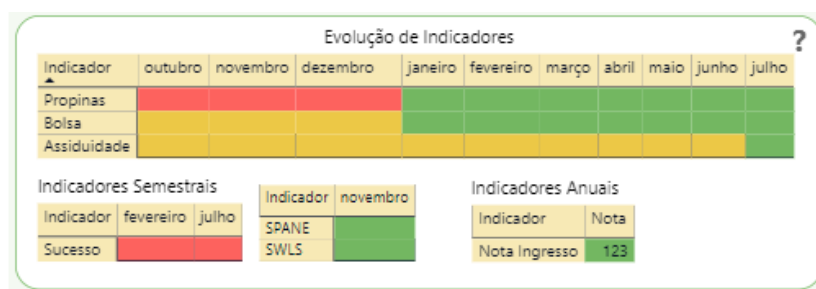


Figura 5.8: Matriz de evolução de indicadores do aluno dividida pela granularidade de atualização

5.3 DISCUSSÃO DE RESULTADOS

Os vários métodos aplicados nesta fase mostraram-se bastante úteis quer ao nível da avaliação da solução proposta quer pela contribuição dada para a melhoria da mesma.

A avaliação heurística assumiu um papel essencial na deteção de alguns problemas de usabilidade que, após corrigidos, para além de melhorar significativamente o contacto do utilizador com a interface do sistema, tornou os testes de usabilidade com os futuros utilizadores do FICAvis ainda mais eficientes. Aqui, a funcionalidade do Power BI de partilha online de relatórios foi fundamental para conseguir realizar as diferentes avaliações com utilizadores.

Os testes de usabilidade por terem contado com a presença de utilizadores com experiência sobre os dados do FICA, e por isso bem esclarecidos sobre as questões mais frequentes ao nível da sua utilização, permitiram obter resultados muito consistentes com base em casos reais. Daqui surgiram também sugestões úteis que possibilitaram melhorar vários aspetos na solução inicialmente proposta.

Os principais potenciais utilizadores do sistema, que contactam directamente com os dados do FICA, revelaram-se satisfeitos com a solução proposta e mostraram interesse na sua utilização com frequência no futuro.

Conclusões e Trabalho Futuro

6.1 CONCLUSÕES

Os dados disponibilizados pelo projeto FICA têm um importante papel na monitorização do percurso dos estudantes da Universidade de Aveiro assumindo-se como essenciais para deteção e consequente prevenção de situações de risco que acabam muitas vezes por levar ao abandono do ensino superior, no entanto, a forma complexa como chegam aos seus utilizadores constitui um desafio extra aquando a sua análise.

O principal objetivo desta dissertação consistiu no estudo das necessidades dos diferentes tipos de utilizador do FICA e na proposta de uma solução que, através dos dados disponibilizados por este projeto, permite responder às suas principais questões.

Inicialmente foram analisados vários trabalhos de visualização realizados com dados académicos de modo a detetar possíveis padrões nas questões habitualmente colocadas pelos utilizadores e nas visualizações usadas. Seguiu-se uma breve apresentação e comparação de algumas ferramentas de visualização a partir da qual se elegeu o Power BI como o serviço a utilizar no desenvolvimento da solução proposta, o FICAvis.

Após realizado um estudo detalhado do projeto FICA e dos dados por ele gerados, deu-se início à fase de elicitação de requisitos com peritos do domínio, representantes dos vários tipos de futuros utilizadores. Aplicando o *focus group* como técnica para as entrevistas, foram definidas as principais questões a responder a partir dos dados.

A solução proposta foi elaborada com base em *personas*, definidas a partir dos resultados obtidos nos vários *focus groups*, e implementada em Power BI. O FICAvis consiste num relatório Power BI composto por vários elementos visuais que, ao contrário dos ficheiros Excel atualmente disponibilizados, permite aos diferentes utilizadores do FICA analisar visualmente os seus dados de forma simples e interativa.

Por fim, para avaliar a solução desenvolvida foram aplicados vários métodos de avaliação.

Os resultados da avaliação heurística realizada por avaliadores com experiência em Visualização e dos testes de usabilidade com peritos do domínio do FICA permitiram validar as visualizações e mecanismos aplicados e destacar a sua acessibilidade para os diferentes perfis de utilizador.

Todo o projeto contou com a participação ativa de peritos do domínio, com diferentes objetivos e necessidades perante os dados, que foram acompanhando as várias etapas do desenvolvimento do protótipo proposto oferecendo constante *feedback* acerca das decisões tomadas, privilegiando uma abordagem centrada no utilizador e promovendo *participatory design*.

6.2 TRABALHO FUTURO

O futuro do FICAvis passa pela sua integração nos Serviços de Tecnologias de Informação e Comunicação sTIC da Universidade de Aveiro permitindo a exploração visual dos dados do FICA e a possível antecipação na deteção e atuação perante situações de risco.

Neste âmbito, os próximos passos consistirão na ligação do FICAvis à base de dados do projeto FICA, seguida da definição da camada de segurança dos dados de acordo com os diferentes perfis e privilégios dos utilizadores. Os dados passariam assim a ser disponibilizados aos Diretores de Curso, Diretores de Departamento e Assessores da Reitoria sob a forma de um relatório Power BI composto por múltiplas visualizações de interesse promovendo a sua exploração visual de forma clara e interativa.

Outra adição importante ao trabalho desenvolvido seria a implementação de mecanismos de alerta automáticos que permitissem chamar a atenção para situações de risco específicas mesmo antes da análise mais cuidada dos dados por parte dos utilizadores.

Nomeadamente durante os *focus groups*, foram também sugeridos vários dados dos estudantes que seriam interessantes recolher, entre os quais, o acesso às plataformas digitais, dependência entre unidades curriculares, informação de residência ou horas de descanso. A representação destes atributos aliada ao cruzamento com outros dados já existentes poderia levar à identificação de novos traços no comportamento dos alunos permitindo melhorar todo o processo de acompanhamento atualmente realizado.

Referências

- Albuquerque, T. (2008). Do abandono à permanência num curso de ensino superior. *Sísifo / Revista de Ciências de Educação*, (2008), 19–28. <http://sisifo.ie.ulisboa.pt/index.php/sisifo/article/viewFile/115/189>
- Ali, S. M., Gupta, N., Nayak, G. K. & Lenka, R. K. (2016). Big data visualization: Tools and challenges. *Proceedings of the 2016 2nd International Conference on Contemporary Computing and Informatics, IC3I 2016*, 656–660. <https://ieeexplore.ieee.org/document/7918044>
- Barata, R., Silva, S., Martinho, D., Cruz, L. & Guerra e Silva, L. (2014). Open APIs in Information Systems for Higher Education, Em *20th EUNIS Congress, Johan Bergström (Ed.)*. Umeå, Sweden. https://www.eunis.org/download/2014/papers/eunis2014_submission_71.pdf
- Bastian, M., Heymann, S. & Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. *Icwm*, 361–362. <https://www.aaai.org/ocs/index.php/ICWSM/09/paper/viewPaper/154>
- Blasilli, G., Lenti, S., Palleschi, A., Allain, K., Turkay, C., Dykes, J., Landesberger, T., Wunderlich, M., Baumgartl, T., Höhn, M. Et al. (2019). CrossWidget: a D3.js Plugin to Support Multiple Scented Cross Filtering Activities, Em *EuroVis*. <https://diglib.eg.org/bitstream/handle/10.2312/eurp20191131/001-003.pdf>
- Choo, A. & Saeger, T. (2010). Data analysis for yield improvement using TIBCO’s spotfire data analysis software. *2010 International Conference on Compound Semiconductor Manufacturing Technology, CS MANTECH 2010*. <https://csmantech.org/OldSite/Digests/2010/Papers/18.3.096.pdf>
- Cooper, A. (1999). The inmates are running the asylum. Indianapolis, IA: SAMS. *Macmillan*.
- Cruz-Guzmán, G. & Rosas, L. V. (2020). Visualization of Data for Decision Making in a University, Em *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 3: IVAPP*, SciTePress. INSTICC. <https://www.scitepress.org/Papers/2020/89847/89847.pdf>
- Dam, R. & Siang, T. (2019). Personas – A Simple Introduction. *Interaction Design Foundation*, 1–15. <https://www.interaction-design.org/literature/article/personas-why-and-how-you-should-use-them>
- Desai, A., Mian, M., Hazel, D., Teredesai, A. & Benner, G. (2014). Data visualization in educational datasets using a rule-based inference system. *Proceedings - 2014 IEEE International Congress on Big Data, BigData Congress 2014*, 462–469. <https://ieeexplore.ieee.org/document/6906816>
- Dix, A., Finlay, J., Abowd, G. & Beale, R. (2003). *Human-Computer Interaction (3rd Edition)*. Upper Saddle River, NJ, USA, Prentice-Hall, Inc. <https://dl.acm.org/citation.cfm?id=1203012>
- Dix, A., Finlay, J., Abowd, G. & Beale, R. (2004). *Human-Computer Interaction*.
- Ferrari, A. & Russo, M. (2016). *Introducing Microsoft Power BI*. Microsoft Press.
- Few, S. (2006). Common pitfalls in dashboard design. *Perceptual Edge*. https://www.perceptualedge.com/articles/Whitepapers/Common_Pitfalls.pdf
- Few, S. & Edge, P. (2007). Dashboard confusion revisited. *Perceptual Edge*, 1–6. https://www.perceptualedge.com/articles/visual_business_intelligence/dboard_confusion_revisited.pdf

- Gao, Y. (2010). A study on mass higher education in the world-based on comparative perspectives, Em *2010 International Conference on Education and Management Technology*. IEEE. <https://ieeexplore.ieee.org/document/5657601>
- Goguen, J. A. & Linde, C. (1993). Techniques for requirements elicitation, Em *[1993] Proceedings of the IEEE International Symposium on Requirements Engineering*. IEEE. <https://ieeexplore.ieee.org/document/324822>
- Gounder, M., Iyer, V. & Mazyad, A. (2016). A survey on business intelligence tools for university dashboard development. *2016 3rd MEC International Conference on Big Data and Smart City, ICBDS 2016*, 85–91. <https://ieeexplore.ieee.org/document/7460347>
- Harle, C., Neill, D. & Padman, R. (2012). Information visualization for chronic disease risk assessment. *IEEE Intelligent Systems*, 27(6), 81–85. <https://ieeexplore.ieee.org/document/6365201>
- Horváth, D. M., Molontay, R. & Szabó, M. (2018). Visualizing student flows to track retention and graduation rates, Em *2018 22nd International Conference Information Visualisation (IV)*. IEEE. <https://ieeexplore.ieee.org/document/8564182>
- Keim, D. (2002). Information Visualization and Visual Data Mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1), 1–8. <https://doi.org/10.1109/2945.981847>
- Kirk, A. (2016). *Data Visualisation: A Handbook for Data Driven Design*. SAGE Publications.
- Kosara, R. (2016). Presentation-Oriented Visualization Techniques. *IEEE Computer Graphics and Applications*, 36(1), 80–85. <https://doi.org/10.1109/MCG.2016.2>
- Martins, T., Gonçalves, D. & Gama, S. (2018). Visualizing Historical Patterns in Large Educational Datasets. *International Journal of Creative Interfaces and Computer Graphics*, 9(1), 32–48. <https://www.igi-global.com/article/visualizing-historical-patterns-in-large-educational-datasets/210549>
- Mayhew, D. (1999). *The Usability Engineering Lifecycle: A Practitioner's Handbook for User Interface Design*. Elsevier Science.
- Mazza, R. & Berre, A. (2007). Focus group methodology for evaluating information visualization techniques and tools, Em *2007 11th International Conference Information Visualization (IV'07)*. IEEE. <https://ieeexplore.ieee.org/document/4271964>
- Moreira, G. & Costa, C. (2018). Percursos dos estudantes de CTESP: a situação dos estudantes inscritos na UA em 2015'16, Em *CNaPPES.18 - 5º Congresso Nacional de Práticas Pedagógicas no Ensino Superior*.
- Moreira, G., Costa, C., Martins, A. V. & Temudo, E. (2017). FICA@UA – integração e monitorização dos estudantes do 1.º ano, Em *Conferência Insucesso e Abandono no Ensino Superior*, Coimbra.
- Munzner, T. (2014). *Visualization Analysis and Design*. A K Peters/CRC Press.
- Murray, S. (2017). *Interactive data visualization for the web: an introduction to designing with*. O'Reilly Media, Inc.
- Negrut, V. (2014). Power BI: Effective Data Aggregation. *Quaestus*, (5), 146–153. <http://www.quaestus.ro/en/wp-content/uploads/2012/02/Viorel-NEGRUT.pdf>
- Nielsen, J. (1995). How to conduct a heuristic evaluation. *Nielsen Norman Group*, 1, 1–8. <https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/>
- Plaisant, C. (2004). The challenge of information visualization evaluation. *Proceedings of the Workshop on Advanced Visual Interfaces AVI*, 109–116. <https://doi.org/10.1145/989863.989880>
- Plummer, P. (2017). Focus group methodology. Part 1: Design considerations. *International Journal of Therapy and Rehabilitation*, 24(7), 297–301. <https://doi.org/10.12968/ijtr.2017.24.7.297>
- Silva, I. S., Veloso, A. L. & Keating, J. B. (2014). Focus group: Considerações teóricas e metodológicas. *Revista Lusófona de Educação*, (26), 175–190. <http://www.scielo.mec.pt/pdf/rle/n26/n26a12.pdf>
- Simon, J. S. (1999). How to conduct focus groups. *Nonprofit World*, 17, 40–43.

- Sorenson, E. & Brath, R. (2013). Financial visualization case study: Correlating financial timeseries and discrete events to support investment decisions. *Proceedings of the International Conference on Information Visualisation*, 232–238. <https://ieeexplore.ieee.org/document/6676568>
- Vardi, M. (2012). Will MOOCs destroy academia? *Communications of the ACM*, 55(11), 5. <https://dl.acm.org/doi/10.1145/2366316.2366317>
- Weiland, A. & Manssour, I. H. (2015). Towards visual analysis techniques for monitoring students of distance education courses, In *Proceedings of the 6th Workshop on Visual Analytics, Information Visualization and Scientific Visualization, 2015, Brasil*.
- Wun, T., Payne, J., Huron, S. & Carpendale, S. (2016). Comparing Bar Chart Authoring with Microsoft Excel and Tangible Tiles. *Computer Graphics Forum*, 35(3), 111–120. <https://onlinelibrary.wiley.com/doi/full/10.1111/cgf.12887>

Apêndices

Script Focus Group

Fábio Ferreira

Universidade de Aveiro, IEETA

ABSTRACT

Neste documento encontram-se descritos os principais passos seguidos e as questões efetuadas durante as sessões de *focus group* realizadas durante a fase de levantamento de requisitos.

Boas Vindas

Boa tarde. Em primeiro lugar obrigado por estarem presentes e despenderem um pouco do vosso tempo. O meu nome é Fábio Ferreira, sou aluno do mestrado em informática e estou a realizar esta pequena entrevista no âmbito da minha dissertação que tem como tema o FICA.

Como já devem estar familiarizados, o FICA – Ferramentas de Identificação e Combate ao Abandono – é um programa desenvolvido pela UA que monitoriza uma série de dados dos alunos com o objetivo de acompanhar o seu processo académico e detetar possíveis situações de risco.

Nesta sessão estão presentes sobretudo alguns diretores de curso de diferentes áreas uma vez que são habitualmente estas as pessoas que têm acesso aos dados do FICA e por isso os potenciais interessados nas visualizações a serem criadas.

Não existem respostas certas ou erradas: a ideia é discutir os vários pontos de vista de cada um dos participantes e com isso entender quais são as principais questões que se pretendem ver respondidas com as visualizações a desenvolver. São igualmente bem-vindos comentários positivos e negativos.

Se não houver problema da vossa parte, vou tirar notas/registar o que for discutido mas de forma totalmente anónima, ou seja, nos relatórios da sessão não irá constar qualquer identificação acerca de quem deu as respostas.

Alguma dúvida? Vamos então dar início à sessão.

Introdução

Qual é o vosso conhecimento acerca do FICA?

Para além de outros atributos, os principais indicadores recolhidos pelo FICA são:

- **IND. 1:** Estudantes com taxa de sucesso académico abaixo de 50 % ECTS a que estão inscritos (por curso);
- **IND.2:** Estudantes com propinas em atraso;
- **IND.3:** Estudantes que pediram bolsa e respetiva resposta (deferida, em análise, indeferida)
- **IND. 4:** Estudantes com nota de ingresso abaixo de 120 pontos;
- **IND.5:** Número de aulas frequentadas pelo estudante face ao número de aulas dadas;
- **IND.6:** Resultados do questionário relativamente à escala de satisfação com a vida (SWLS);
- **IND.7:** Resultado obtido pelo estudante segundo a escala de experiência positiva e negativa.

O objetivo desta dissertação passa por estudar e desenvolver visualizações, intuitivas e interativas, que permitam essencialmente a interpretação dos dados do FICA. A realização destas sessões com os principais *stakeholders* é um dos pontos fulcrais para a realização de um trabalho que possa vir a contribuir significativamente para a Universidade de Aveiro através da sua futura integração nos sistemas de informação da instituição.

Questões de Abertura

- Já trabalharam alguma vez com os dados?
 - Em que situação/com que objetivo?
 - Com que opinião ficaram?
 - O que acham que poderia ser feito para melhorar?
- O que gostariam de saber a partir dos dados?
 - Alunos em dificuldade, lista de alunos segundo algum critério, ...
- O que estariam à espera de encontrar num sistema para exploração visual de dados centrado no FICA?
- O que gostariam de poder ver e qual informação obter numa aplicação deste género?

Questões exploratórias

- Quais são os indicadores que acham ter maior importância/relevância?
- Que tipo de visualizações seriam úteis?
 - Que atributos representar?
 - Que tipo de comparações fazer?

Apresentação de protótipos

A seguir serão apresentados alguns protótipos de visualizações criados a partir de dados processados pelo FICA. Para cada uma das visualizações iremos discutir os seus pontos positivos e negativos.

- Que tipo de informação conseguem obter desta visualização?
- Acham esta representação útil?
 - Devia ser representada de outra forma? Como?

Tendo em conta as técnicas de visualização descritas abaixo, qual é a vossa perceção de importância da utilização de cada uma delas numa representação visual?

- **Localizar e Identificar:** Consegue localizar e identificar um objeto particular na visualização ao olhar para esta pela primeira vez?
- **Distinguir:** Consegue distinguir os diferentes objetos do *dataset*? Existem itens visuais que não estão claros nesta visualização?
- **Categorizar e Classificar:** Esta representação permite classificar os dados?
- **Comparar:** Acha que esta visualização oferece oportunidade para comparar de forma clara atributos relacionados?

Encerramento da sessão

- Têm mais alguma sugestão que não tenha surgido durante a sessão?

Uma vez mais agradeço a vossa presença nesta pequena reunião. Após reunir e analisar toda a informação do que aqui foi falado (tal como referi no início, de forma anónima) vou então ver de que forma a posso aplicar na dissertação. Se pretenderem continuar a seguir este projeto, por favor indiquem-me (deixem-me os vossos mails) e claro, como isto se trata de um projeto que implica constante troca de ideias em todo o desenvolvimento (*participatory design*), sintam-se à vontade para dar sugestões. Obrigado.

Resultados dos *Focus Groups*

Sessão 04/03/2020 (IEETA)

- Elementos:
 - Diretor de Curso (DETI)
 - Vice diretor de curso (DEGEIT)

Introdução

- Ambos os participantes têm conhecimento do programa FICA
- “Dados chegam c/ algum atraso” aos destinatários
- Já trabalhou com os dados
- Procurou alunos que não satisfaziam determinado indicador
- Alunos que não satisfazem (pelo menos) 2 indicadores -> considerados de risco pela reitoria
- Com os dados do final do semestre, agrupar alunos pelas cadeiras a que foram bem sucedidos
 - Saber a % de alunos com aprovação (fins estatísticos)
 - Histograma com a % ou valor absoluto (desempenho)
 - Seleção total -> ordenação manual
 - Especificar as cadeiras feitas (número de ECTS, qual a cadeira)

Objetivo: saber quais os alunos e as cadeiras que precisam de especial atenção

Alunos devem ser classificados em 2 grupos: reprovados e não reprovados

- Tratar de forma diferente (filtrar) dados da 1ª matrícula e das outras
 - Saber se os ECTS são do 1º/2º ano, etc.
 - Saber taxas/ratios de desempenho

Mostrar detalhes <-> comparações de indicadores:

(Indicadores com granularidade mensal)

- Faltas
- Média de entrada -> evolução nota
- Pedido de bolsa

Os dados do excel obrigam à filtragem de colunas

- Vários campos semelhantes: curso, código curso, nome curso
- “É difícil associar rapidamente um curso/cadeira pelo seu código”

O cruzamento de dados relativamente às propinas (e bolsas) é importante para identificar situações de risco.

Apresentação protótipo

Estatísticas gerais

- Possível filtragem por curso

Assiduidade Vs. Desempenho

- Filtrar por 1ª matrícula
- Adicionar legenda: “bem sucedidos”...
- Identificar código dos indicadores (10,20,30) de outra forma
- % de alunos em vez do valor absoluto
- Ter em atenção se os dados são só do 1º ano
- Nem todas as cadeiras fazem o registo das faltas (!)

Indicadores de Risco e Estado Matrícula

- Acessos diferente mediante privilégios
 - o Reitoria: informação por área/curso
 - o Diretores: informação específica alunos

Estudante em Risco

- Dados apresentados são importantes

Evolução Anulação Matrícula

- É possível perceber quem anulou?
- Como saber qual o motivo da desistência?
 - o Os alunos que anulam pessoalmente especificam o motivo (MotivoAnulacao)
 - o Os alunos que “deixam de aparecer” não têm motivo associado

Distribuição aproveitamento alunos

- Pie chart com a % -> possibilidade de ter um histograma
- Desempenho de curso X -> desempenho dos alunos matriculados no curso X

Outras visualizações ? (sugestões)

Intervenção Tutoria:

- Cruzar propinas com informação bolsa
- Saber se é a 1ª vez no curso do aluno

Filtro (permitir filtrar por)

- Alunos em tempo parcial
- Máx. 50% ECTS
- Devem ser tratados à parte
- > 23 / candidaturas especiais – diferenciar estatutos

Seria importante saber a residência atual do aluno (para já não existe essa informação)

Diferentes estatutos, diferentes visualizações

Possibilidade de diferentes dashboards para pessoas com privilégios diferentes.

Personas a definir

- Diretor de curso
- Diretor de departamento
- Reitoria

(por exemplo, o diretor de curso tem acesso à informação dos alunos, o diretor de departamento à informação das cadeiras e a reitoria informação dos cursos)

Trabalho futuro

Configurar dashboard (?)

- Pode ser complicado para algumas pessoas (com pouca experiência)
- “Útil, mas não é algo necessário”

❖ No DEGEET dedicam período tempo (~2h) análise dados do FICA

Ideia

Fazer “*focus group* online” com um questionário pré-definido (google forms)

(para os diretores de curso que não conseguirmos entrevistar)

(reduzir tempo para obter dados mais fiéis)

Possíveis perguntas

- **2 blocos**
 - Tipo de dados que são importantes para análise
 - Identificar dados cruzados para obter casos críticos
 - Tipo de indicadores que gostariam de ter
- Duração/frequência/objetivo/tipos de dados
- Pergunta aberta
- Sugerir utilização e pedir “favoritos”
- Indicadores mais importantes (avaliados segundo uma escala)

Sessão 11/03/2020 (IEETA)

- Elementos:
 - Professor de uma cadeira importante de 1º ano (DETI)
 - Diretor de Curso (DETI)

Diretor de Curso (DETI)

- Utilizador do FICA
- Análise estatística
- Nota entrada / nota obtida => estudo correlação
- Número de faltas
- Notas
- Mudanças / desistências

Prof. de uma cadeira importante de 1º ano (DETI)

- Conhece, mas nunca utilizou o FICA
- Essencialmente informação do 1º ano / 1º semestre
- Dashboard web
- “A partir de quando existe o trigger” – passagem do verde -> vermelho
- Bons alunos são afetados pela gestão do tempo
- Qualificar entropia da gestão de tempo
- Análise 1º ano: 1º semestre -> 2º semestre

(os 2 últimos indicadores são apenas questionados 2 vezes durante o ano letivo?)

ECTS

- Qualificar distribuição por UC ao longo do semestre
 - Pode-se tentar qualificar a gestão do tempo por aqui
- Identificação situações dentro dos *vermelhos*
 - “Existem coisas feitas pelos docentes, mas sem a utilização de ferramentas concretas”
 - Correlação com a área científica
- Progressão do aluno (“dependência de UC”): Aluno que fez sistemas digitais deve inscrever-se em Arquitetura de Computadores (AC1)? Mesma situação em P2 -> P3

Custo envolvido na formação de um aluno

- Mão de obra
- Recursos humanos e não humanos
- Etc.
- Pedido de bolsa (e consequente resposta) pode estar ligado ao abandono?
- Quais os casos de abandono ou de mudança de curso / instituição !
- Considerar assiduidade => regime de faltas é diferente nas teóricas e práticas
- Em que fase entrou no curso? 1ª/2ª/3ª (influência no desempenho)
- **Acesso às plataformas digitais:** email, e-learning, ...

- Correlação com a nota obtida
- Correlação entre métricas do meio e final do semestre
 - Permitir detetar de imediato situações *vermelho*.
 - Despoletar alertas
- Configurar a correlação de notas UC, nota entrada, etc.
- Determinar média dos casos concretos de aproveitamento

Outros possíveis indicadores

- Forma de estudo – contínuo (semanal) / apenas época de exame
- Dificuldade cadeira
- Autonomia do aluno (~~diminuir~~ horas de contacto)
- Horas de sono / qualidade descanso
- De que forma a mudança de regente (da cadeira) afeta % casos *vermelhos*?

Diferentes *personas*, diferentes motivações (1..., 2..., 3...)

Requisitos:

- para executar
- desejados

Inquérito *online* onde se enquadra o trabalho futuro.

Sessão 17/03/2020 (Vídeoconferência)

- Elementos:
 - Membro da Reitoria

O que geralmente interessa?

Desempenho dos estudantes durante e no final do ano letivo.

. Grupo de estudantes em risco

. Curso e tipo de acesso

FICA: Para já os dados são referentes apenas ao 1º ano, no entanto a ideia é alargar o programa para os restantes anos

(os dois novos indicadores relacionados com a escala de satisfação ainda não são enviados para os diretores de curso)

Sugestões de indicadores

- Nº acessos às plataformas e a sua relação com o aproveitamento dos alunos (possível?)

- Residência atual dos alunos (possível obter pelos questionários de satisfação?)

Grupos de risco

Geralmente alunos que não entram pelo regime geral

- Alunos internacionais
- Maiores de 23 anos

Os principais indicadores que têm impacto no comportamento dos alunos são:

- Fase de entrada
- Opção
- Nota-ingresso

Deteção de Padrões

Analisar dados mensalmente e comparar com meses/anos homólogos

(pessoas envolvidas: vice reitor da educação, vice reitor de qualidade, ensino pedagógico,...)

Objetivo: Dados em bruto + dashboard

FICA

Tipos de teste: modelos de regressão múltipla (referentes ao desempenho)

Monitorização dos ficheiros ano anterior

- Caracterizar casos de abandono:

Formal vs. desistência

APÊNDICE C - AVALIAÇÃO HEURÍSTICA DA PRIMEIRA VERSÃO DA SOLUÇÃO PROPOSTA

Análise heurística – Visualização FICA

11/05/2020

#1: Visibility of system status

- Título da página e da tab não é sempre igual - Confuso
- Títulos do gráficos com uma cor mais marcada e mais destaque (maior ou bold)
- A seleção (desempenho do curso) é interessante, mas seria fixe realçar o que foi selecionado (se for fácil) por exemplo um border a border do que for selecionado e (talvez complicado) um título/nome a indicar - Em particular na lista de alunos fica se na dúvida se foi atualizada ou não - Pouco importante
- Talvez uma linha a indicar seleção (Se tenho só 1º ano 1º semestre, indicar no título) - pouco importante e pode complicar também
- Página pessoal do aluno: Título das tabelas um pouco maiores e/ou bold para evidenciar. Talvez uma caixa a volta das visualizações. Muita informação pode ser confusa perceber a organização ainda mais dado o alinhamento que não é sempre igual

#2: Match between system and the real world

- A utilização dos códigos (e não nomes reais) dificulta um pouco a analise - mas não depende de ti (talvez uma versão com nomes reais ou colocar Curso XPTO - Aluno XYZ)
- Quando falta uma categoria isso ser claro (por exemplo nos alunos por desempenho se nenhum aluno tem 100% deixar a barra a indicar isso com o valor 0). Depois é confuso porque as vezes são 4 categorias, as vezes 3...

#3: User control and freedom

- A ida para a página do aluno deveria ter outra alternativa para além do menu contextual (muito nunca vão encontrar) - Redundância
- Na página pessoal permitir de alguma formar ver a lista de alunos e seleciona um sem ter de ir a outra página

#4: Consistency and standards

- Talvez em alguns casos usar cores para chamar atenção (por exemplo ativo = verde | Inativo = vermelho) - ECTS _ Abaixo de 50% vermelho, acima Verde, etc... Mas com cuidado!
- A tabela dos alunos não apresenta sempre a mesma informação em função da página onde esta (por exemplo no curso: Anos, Ects, desempenho..., mas na evolução so o estado da matricula)? Qual a lógica ou apresentar sempre tudo ou pelo menos informação que tem a

ver com o que é apresentado (na evolução do curso, falta os indicadores de risco por exemplo)

#5: Error prevention

- Alguns gráficos por exemplo na evolução do curso atualiza o eixo em função dos valores. Por exemplo ao ver Curso C#0 ano curricular 1, a evolução de desempenho parece descer muito, mas o eixo vai de 74 a 74.5. Se for assim engana (na escala de 0 a 100 seria quase uma linha horizontal).

#6: Recognition rather than recall

#7: Flexibility and efficiency of use

- [Desempenho mas deve ser preciso nas outras] falta um botão para cancelar as seleções feitas (por exemplo tenho 1º semestre e notas entre 12 e 16) e quero cancelar todas as filtrações
- Na nova comparação por medidas, não me parece muito útil mostrar só 1 fase - Mostrava sempre as 2 - igual para os outros - Só interessa ver o desempenho em função das medidas não haverá outros indicadores interessantes de mostrar?

#8: Aesthetic and minimalist design

- Não gosto muito do slicer (por exemplo para a nota no desempenho) - 2 casa decimais e valores acima, Talvez ir de 0.5 em 0.5 e colocar valores das notas mais perto do slicer (a esquerda e direita)
- Na análise mensal como os valores estão lado a lado, já gostei um pouco mais
- Talvez centrar sempre o título da página.

#9: Help users recognize, diagnose, and recover from errors

- Ter nomes que ajudem a perceber o que se vê / Simplificar
 - Exemplo "Distribuição dos alunos por desempenho" = "Desempenho por aluno"
- Na página de evolução de curso - será sucesso ou insucesso - Penso que os alunos que aparecem são os que chumbaram = insucesso (tooltip ou documentação poderia ajudar)

#10: Help and documentation

- Falta Informação na página sobre as médias, métricas, etc... Dar acesso a informação sobre os dados do FICA
- Informação sobre a origem e forma de cálculo de alguns indicadores
- Se possível tooltips quando necessário a dar informação adicional.

Outras coisas:

- Desempenho de curso: Não seria possível ver a informação anual compilada e não semestre a semestre. Talvez ver também da duração de todo o curso?
- Ao nível dos indicadores, não seria possível (e útil) fazer algumas visualizações a longo do tempo a mostrar a evolução dos indicadores de um aluno ou de alunos de um curso ao longo do ano?
- Nos indicadores, A situação de prescrição do aluno aparece sem nenhuma informação?

FICAvís

Visualização de Dados para Combate ao Abandono Escolar na Universidade de Aveiro

(Fábio Ferreira)

Avaliação Heurística

O FICA – Ferramentas de Identificação e Combate ao Abandono – é um projeto desenvolvido pela Universidade de Aveiro que monitoriza uma série de atributos dos seus estudantes com vista a detetar possíveis situações de risco de abandono e prestar apoio na tomada de decisões que permitam combater estes casos e com isso melhorar o aproveitamento verificado.

Desenvolvida em Power BI, o FICAvís consiste numa solução de visualização dirigida a três diferentes tipos de utilizadores – Diretor de Curso, Diretor de Departamento e Assessor da Reitoria – compostos por múltiplas páginas que por sua vez contêm vários elementos visuais relacionados entre si e permitem facilmente extrair informação a partir dos dados.

A avaliação heurística é um método de avaliação que tem como principal objetivo detetar possíveis problemas de usabilidade num sistema, nomeadamente ao nível da sua interface.

Com este documento pretende-se realizar uma avaliação heurística sobre a solução proposta tendo como referência as 10 heurísticas propostas por Nielsen:

- N1 Visibilidade do estado do sistema:** o sistema deve manter os utilizadores informados de tudo o que está a acontecer numa determinada interação oferecendo *feedback* de forma clara e dentro de um intervalo de tempo razoável.
- N2 Correspondência entre o sistema e o mundo real:** o sistema deve disponibilizar uma linguagem e conceitos familiares ao utilizador; por exemplo a apresentação da informação por uma ordem lógica, ou a utilização de cores específicas pode invocar experiências passadas dos utilizadores tornando a utilização do sistema mais fácil e intuitiva.
- N3 Controlo e liberdade do utilizador:** o sistema não deve decidir pelo utilizador; o utilizador deve ser livre na escolha das suas ações e devem existir mecanismos que lhe permitam voltar atrás ou anular uma decisão sempre que desejar.
- N4 Consistência e padronização:** o sistema deve manter consistência entre todos os seus elementos, como ícones ou expressões; os padrões de interação devem ser mantidos nos diversos contextos do sistema permitindo ao utilizador facilmente identificar as suas funcionalidades;
- N5 Prevenção de erros:** o sistema deve sempre que possível prever a ocorrência de potenciais erros sem a necessidade de intervenção do utilizador, no entanto, quando não o é possível, o utilizador deve ser alertado das consequências das suas ações.

- N6 Reconhecimento em vez de recordação:** o sistema deve minimizar a carga de memória imposta ao utilizador sempre que possível permitindo-lhe, por exemplo, reconhecer possíveis interações anteriores com o sistema; os diferentes elementos e funcionalidades de um sistema devem manter-se visíveis sempre que possível;
- N7 Flexibilidade e eficiência:** o sistema deve ser acessível tanto a utilizadores experientes como novos utilizadores; deve ser possível configurar a interface de acordo com as necessidades do utilizador

N8 Estética e *design* minimalista: o sistema deve apenas apresentar informação relevante de forma a não desviar o foco do utilizador.

N9 Ajuda no reconhecimento, diagnóstico e recuperação de erros: o sistema deve informar os seus utilizadores da ocorrência de erros e, nesses casos, de possíveis formas de os corrigir, por exemplo através de mensagens de erro.

N10 Ajuda e documentação: uma interface clara e intuitiva é o primeiro passo para permitir ao utilizador navegar por um sistema sem qualquer ajuda adicional, no entanto, é importante que existam formas de prestar auxílio ou informação acerca das ações possíveis sempre que necessário.

Nielsen definiu também uma escala de severidade que permite classificar as várias heurísticas:

0 = Não concordo que seja de todo um problema de usabilidade

1 = Apenas um problema de estética: não há necessidade de corrigir a não ser que sobre tempo no final do projeto

2 = Problema mínimo de usabilidade: baixa prioridade de resolução

3 = Problema grave de usabilidade: alta prioridade de resolução

4 = Usabilidade catastrófica: imperativo corrigir antes e lançar o sistema

Considerando as 10 heurísticas de Nielsen foram seguidas duas diferentes metodologias:

Na primeira, o avaliador classifica cada página do FICAVIS segundo as 10 heurísticas atribuindo-lhe o grau de severidade correspondente, e na segunda esta classificação é feita tendo em consideração todo o relatório e os respetivos problemas de usabilidade encontrados. Os resultados são detalhados a seguir.

Resultados das avaliações heurísticas

Classificação da usabilidade de cada página do FICAVIS	3
Avaliador 1	3
Avaliador 2	5
Avaliador 3	7
Avaliador 4	10
Problemas de usabilidade do FICAVIS, avaliado como um todo	12
Avaliador 5	12
Avaliador 6	13

Classificação da usabilidade de cada página do FICAVIS

→ Avaliador 1

Página Inicial

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	0	

Visão Global do Desempenho

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	1	Na barra lateral que contém os filtros na parte do subtítulo “Regime” quando se clica para visualizar as opções de filtragem aparecem duas opções “Agente Associati...” o que não permite reconhecer com clareza que regime é dado que a opção aparece truncada. Talvez uma ajuda a indicar para passar o rato em cima dado que o texto da opção assim já aparece na sua totalidade.
N6	0	
N7	0	
N8	0	
N9	0	
N10	0	

Monitorização Mensal

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	

N10	0	
-----	---	--

Comparação Mensal

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	0	

Desempenho Global

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	0	

Desempenho por Indicador / Perfil

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	0	

Informação Pessoal

Heurística	Severidade	Comentários
N1	1	No canto inferior direito ("Ver mais informação do aluno") um ícone de ajuda talvez a indicar que a funcionalidade é ativada com mouse over, pois no meu caso ela demorou a ativar por isso assumi que fosse clique mas não tinha feedback.
N2	0	

N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	0	

→ **Avaliador 2**

Página Inicial

Heurística	Severidade	Comentários
N1	0	Cada botão diz para clicar ao passar por cima.
N2	0	
N3	1	Ao clicar deveria de haver um botão para retroceder.
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	Informação clara e bem estruturada.
N9	0	
N10	0	

Visão Global do Desempenho

Heurística	Severidade	Comentários
N1	0	
N2	0	Fácil compreensão, as cores diferenciam a parte selecionada.
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	0	

Monitorização Mensal

Heurística	Severidade	Comentários
N1	0	Os dados são apresentados de forma clara e bem estruturados a nível visual.
N2	0	Os meses encontram-se por a ordem uma vez que começam no mês em que começa o ano.
N3	0	Existe o botão de Limpar Seleção.
N4	0	
N5	0	
N6	0	
N7	0	Bastante acessível.
N8	0	
N9	0	
N10	0	Informação bem estruturada, ao passar por cima de cada gráfico é possível obter mais informações do mesmo.

Comparação Mensal

Heurística	Severidade	Comentários
N1	0	
N2	0	Fácil compreensão, as cores diferenciam a parte selecionada.
N3	0	É possível clicar para voltar a ver o gráfico e a tabela.
N4	0	
N5	0	
N6	1	No painel de seleção uma vez que só existe uma opção poderia estar a opção em vez de ver tudo.
N7	0	
N8	0	
N9	0	
N10	0	

Desempenho Global

Heurística	Severidade	Comentários
N1	0	
N2	0	No gráfico é possível uma fácil compreensão, as cores diferenciam a parte selecionada.
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	1	A barra de mover a tabela para cima e para baixo deveria ser um pouco maior.
N9	0	
N10	0	

Desempenho por Indicador / Perfil

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	1	A barra de mover a tabela para cima e para baixo deveria ser um pouco maior.
N9	0	
N10	0	

Informação Pessoal

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	Informação clara e bem estruturada.
N9	0	
N10	0	

→ Avaliador 3

Página Inicial

Heurística	Severidade	Comentários
N1	2	Nas setas de acesso às páginas, talvez fosse melhor substituir o “Clique aqui para seguir a ligação” por uma mensagem que informe qual vai ser a página de destino
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	1	Dada a existência de várias páginas de navegação, talvez fosse interessante uma barra de pesquisa de acesso rápido às páginas. Conforme o utilizador for introduzindo os caracteres, o sistema poderia ir dando uma sugestão das páginas existentes.
N8	3	A utilização de setas como elemento de navegação para as mais diversas páginas não é muito apelativo. Talvez dividir o ecrã por áreas/divisões para cada tema fosse mais perceptível.
N9	0	
N10	1	Cada título de cada tema poderia ter um icon de ajuda onde

		pudesse ser mostrada a sua descrição, invés de ser apresentada no <i>display</i> da página inicial.
--	--	---

Visão Global do Desempenho

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	3	A página não tem uma seta de navegação que permita voltar à Página Inicial
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	3	Após ter sido selecionado um departamento, os <i>links</i> de acesso para as páginas da visualização de dados (Monitorização Mensal, Composição Mensal, etc) poderiam ser mais informativas do que apenas dizer “Clique para explorar até à Monitorização Mensal”

Monitorização Mensal

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	3	No título da página não menciona qual o Departamento que foi anteriormente selecionado
N7	0	
N8	0	
N9	0	
N10	0	

Comparação Mensal

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	3	No título da página não menciona qual o Departamento que foi anteriormente selecionado
N7	0	
N8	0	
N9	0	

N10	0	
-----	---	--

Desempenho Global

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	3	No título da página não menciona qual o Departamento que foi anteriormente selecionado
N7	0	
N8	0	
N9	0	
N10	0	

Desempenho por Indicador / Perfil

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	2	Para filtrar os dados de visualização é primeiramente necessário escolher uma categoria (Fase, Tipo de Acesso, Opção e Regime), seguida da opção desejada. Dada esta ordem, talvez fosse mais intuitivo que a escolha das categorias fosse à esquerda das opções.
N5	0	
N6	3	No título da página não menciona qual o Departamento que foi anteriormente selecionado
N7	0	
N8	0	
N9	0	
N10	0	

Informação Pessoal

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	4	A opção “Ver Mais Informações do Aluno” não funciona
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	1	Seria interessante ter um ícone de ajuda nas secções “Unidades Curriculares Inscritas” e “Informações ECTS Ano Atual”

→ **Avaliador 4**

Página Inicial

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	0	

Visão Global do Desempenho

Heurística	Severidade	Comentários
N1	2	Algumas cores tem baixo contraste com o texto, o que dificulta a sua visualização
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	0	

Monitorização Mensal

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	0	

Comparação Mensal

Heurística	Severidade	Comentários
N1	1	O tamanho de letra é pequeno em alguns locais o que pode

		dificultar a sua correta visualização
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	0	

Desempenho Global

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	0	

Desempenho por Indicador / Perfil

Heurística	Severidade	Comentários
N1	0	
N2	0	
N3	0	
N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	2	Deveria ser mais explícito que o utilizador pode expandir informações em cada bloco no gráfico, não é evidente ao abrir a página

Informação Pessoal

Heurística	Severidade	Comentários
N1	2	Na secção "Evolução de Indicadores" a informação sobre significado das cores deveria estar mais explícito e não deveria ser necessário utilizar o ícone de ajuda, pois isso faz o utilizador abrir e fechar várias vezes a janela de ajuda desnecessariamente
N2	0	
N3	0	

N4	0	
N5	0	
N6	0	
N7	0	
N8	0	
N9	0	
N10	2	Correlacionado com o problema mencionado anteriormente, elementos indispensáveis à análise do gráfico deveriam ser mais visíveis/explicitos

Problemas de usabilidade do FICAvís, avaliado como um todo

→ Avaliador 5

1: Na visão global do desempenho dos estudantes nos quadrados dos departamentos, não se consegue ler o quadrado do lado direito no canto inferior direito que corresponde ao departamento de desporto.

Heurística: Visibility

Severidade: 2

2: Na visão global do desempenho dos estudantes ao seleccionarmos ano letivo 2017 e 1º semestre o ano curricular e o regime não dão para alterar e a grelha com os quadrados dos departamentos não carrega. Não apresenta dados nem nenhum aviso de que não existem dados sobre aquele particular semestre.

Heurística: Recover

Severidade: 2

3: No painel de seleção de filtros avançados na monitorização mensal, quando seleccionamos algum filtro as dropbox não se fecham automaticamente ao serem seleccionadas opções ao contrário das outras dropboxes nos filtros de cada página.

Heurística: Recognition

Severidade: 1

4: Ao entrarmos no acesso aos dados de todos os estudantes na página principal, não temos depois como voltar à página inicial.

Heurística: User control

Severidade: 3

5: Na visão global do desempenho dos estudantes, se seleccionarmos um departamento e entrarmos na opção monitorização mensal e abrirmos o menu de filtros avançados, se não o fecharmos e voltarmos para a página anterior, ao entrar de novo na página o menu continua aberto.

Heurística: Aesthetic and minimalist

Severidade: 1

6: Na visão global do desempenho dos estudantes ao seleccionarmos um departamento e entrarmos em qualquer uma das opções disponíveis debaixo dos departamentos, tudo o que utilizarmos de filtros ficará gravado ao voltarmos a entrar noutra opção mas num diferente departamento.

Heurística: Recognition

Severidade: 2

7: A página inicial podia ter um pouco mais de cor em vez de ser tudo a preto para apelar a atenção visual.

Heurística: Visibility

Severidade: 1

8: Na página de informação pessoal o botão de limpar os filtros não funciona.

Heurística: Aesthetic and minimalist

Severidade: 3

9: Na página de informação pessoal o botão de mais informações do aluno não funciona.

Heurística: Aesthetic and minimalist

Severidade: 3

→ **Avaliador 6**

1: Indicação de *loading* não é perceptível

Severidade: 2

2: Não clicando em informação o site atualiza na mesma (Monitorização mensal)

Severidade: 1

3: Nas mais opções de filtragem o pop-up que é aberto fica *glitchy* (Comparação Mensal)

Severidade: 3

4: Nas mais opções de filtragem o pop-up na opção primeira vez no curso True/Falso utiliza duas linguagens diferentes (Comparação Mensal, Desempenho Global)

Severidade: 1

5: No Tipo de acesso seria de esperar poder escolher mais do que 1

Severidade: 1

6: Botão ver mais informação do aluno não funciona

Severidade: 3

Nota: Quanto a questão nas mais opções de filtragem eu não sei muito bem como replicar o problema mas algumas vezes clicando fora das mesma depois quando o rato voltava a entrar as opções desapareciam ou ficavam inacessíveis

APÊNDICE E - ANÁLISE DE RESULTADOS DOS TESTES DE USABILIDADE E SUGESTÕES DOS UTILIZADORES

Testes de Usabilidade – Resultados e Sugestões

22/05/2020 (Vídeoconferência)

Diretor de Curso (DETI)

Página Inicial

- Para além do nome do curso, mostrar também o código ✓

Desempenho

- Nota de ingresso e ECTS: mudar o número de casas decimais (atenção escala 0-20 vs. 0-100) ✓
- Ano curricular: ano mais avançado ao qual o aluno está matriculado numa UC (manter) ✓
- Na parte superior, manter média global e acrescentar a **selecionada** (com a média de todos os alunos numa cor diferente + transparente) (informação adicionada ao título da tabela)
- Sidebar: acrescentar tipo de acesso, tempo parcial (**sugerir provas de acesso do aluno**) ✓
- Lista de alunos: mostrar o número de alunos representados (no título por exemplo) ✓

Página pessoal

- Mudar título da página: informação do aluno (de acordo com o “mais info.” em baixo) ✓
- Matriz: ordem dos indicadores de acordo com o FICA ✓
 - Diferenciar indicadores fixos dos que variam ao longo do mês:
 - Assiduidade, Propinas
 - Bolsa (vale a pena representar tudo ou só o último valor?)
 - Sucesso (só Fevereiro e Só Julho)
 - Inqueritos, nota de ingresso (mostrar só um valor - Outubro)
- ECTS Feitos: especificar que são referentes ao ano atual/corrente; mostrar o total de ECTS feitos do plano ✓
- (Código das cadeiras mostrar o nome) (sem essa informação no FICA)

Indicadores

- Para além do mês, uma espécie de slicer para filtrar por períodos (ex. de Fevereiro a Maio)

Evolução

- Mudar títulos:
 - Acumulação de anulações -> “anulações acumuladas” ✓
 - Acumulação de indicadores -> “alunos em risco” ✓
- Alterar cores dos casos negativos: por exemplo anulações a vermelho ✓

- Alertar nos “alunos em risco” que o mesmo aluno pode ser contado mais do que uma vez (info. adicionada no tooltip)
- Repensar e alterar posições dos gráficos: primeiro alunos em risco e em baixo anulações ✓

Comparação

- Ordenar as opções pelo número de alunos ✓
- Faz sentido visualizar para além dos dados normalizados, o valor absoluto em gráfico?
- Mais Informação no título da lista de alunos (nº de alunos) ✓

Membro da Reitoria

Desempenho

- Na lista mudar “desempenho” -> taxa ECTS ✓
- 1ª + 2ª semestre em alternativa a 2ª semestre ✓
- Cruzar propinas (e/ou bolsa) com o desempenho ✓
- Média com duas casas decimais ✓

Página pessoal

- Alterar nome página para “informação pessoal” ✓
- Matriz de indicadores: fixos vs. variáveis ao longo do ano ✓
- Mais campos de pesquisa (nome para além de nMec) (o filtro de pesquisa do Power BI só aceita um campo de pesquisa)

Indicadores

Evolução

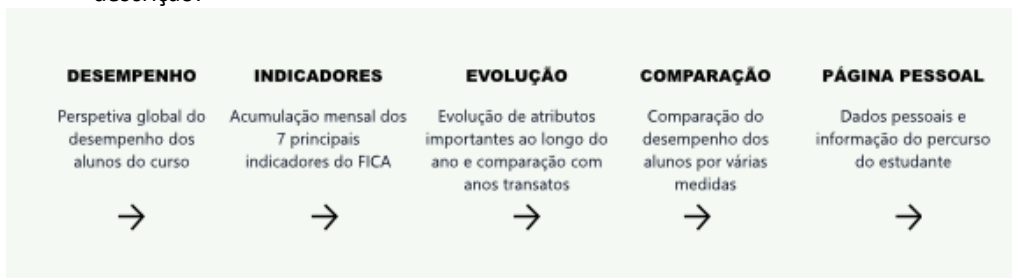
- Acumulação de indicadores: alterar tipo de gráfico ✓
- A informação é importante
- Ativar múltipla seleção ✓

Comparação

- Mudar título “comparar desempenho” ✓
- Mudar legenda: 1ª, 2ª, 3ª, 4ª ... -> 1ª fase, 2ª fase... ✓
- Comparar pela nota de ingresso (definir intervalos) ✓
- Importante adicionar subconjuntos (na barra de filtros) (os “novos” dados do FICA não têm esse indicador)

Algumas sugestões:

- 1) FICA é o projeto onde envolve várias ações, em que uma delas é os ficheiros FICA, mas o nome formal é Observatório do Percurso dos Estudantes. ✓
- 2) Eu colocaria a ordem das páginas de forma diferente e com outro nome e outra descrição: ✓



Página principal para os DC são os indicadores, seguindo-se a Evolução, o Desempenho, a Comparação e a Página Pessoal.

Sugestões de nomes:

Indicadores → monitorização mensal (Apresentação das situações em risco relativo aos 7 principais indicadores mensalmente)

Evolução → Comparação mensal (Comparação mensal dos principais indicadores)

Desempenho → Desempenho global

Comparação → Desempenho por perfil (Desempenho dos alunos (ou estudantes) por pelos diversos perfis.

- 3) Ver melhor a legenda de acordo com o ficheiro FICA. ✓

Considerando as 7 principais medidas do FICA:

- Sucesso** - Taxa de ECTS feitos < 50%
- Propinas** - Propinas em atraso
- Bolsa** - Pedido de bolsa rejeitado
- Nota de entrada** - Abaixo de 120 valores
- Assiduidade** - Pelo menos uma UC com presença < 80%
- SWLS** - Resultados negativos no questionário
- SPANE** - Resultados negativos no questionário,

apresenta a contagem do número de alunos que não satisfaz cada um dos indicadores
(para os quais existe registo nos dados)

Uma coisa é a definição do indicador e outra é a situação em risco (ou vermelho)

Sugestão onde está medidas ser indicadores

Cuidado não é 120 valores mas 120 pontos (valores varia entre 0 a 20 e pontos de 0 a 200);

Na legenda do ficheiro está:

Indicador 2.1 - Estudantes com taxa de sucesso académico abaixo de 50% dos ECTS a que estão inscritos (por curso)

Indicador 2.2 - Estudantes com propinas em atraso (verde = 0; vermelho >0; branco = NA)

Indicador 2.3 - Estudantes que pediram bolsa (verde = Deferida; amarelo = Em análise; vermelho = Indeferida; branco = NA)

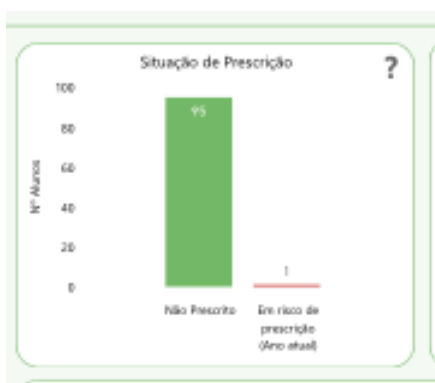
Indicador 2.4 - Estudantes com nota de ingresso abaixo de 120 pontos (verde ≥ 120 ; vermelho <120; branco = NA)

Indicador 2.5 - Número de aulas frequentadas pelo estudante face ao número de aulas dadas (vermelho = assiduidade a pelo menos 1 UC < 80%; amarelo = situações de erro administrativo (-2 e -1); branco = NA; verde as restantes situações). As situações de erro administrativo podem possuir o valor -2 (UC com turmas criadas para as componentes existentes, mas sem registo de aulas dadas, logo sem assiduidade) ou o valor -1 (UC sem turmas criadas para as componentes existentes).

Indicador 2.6 - Estudantes com nível de satisfação com a vida baixa (verde ≥ 3 ; vermelho <3; branco=NA)

Indicador 2.7 - Estudantes com nível de experiência positiva e negativa baixa (verde $\geq 0,75$; vermelho <0,75; branco=NA)

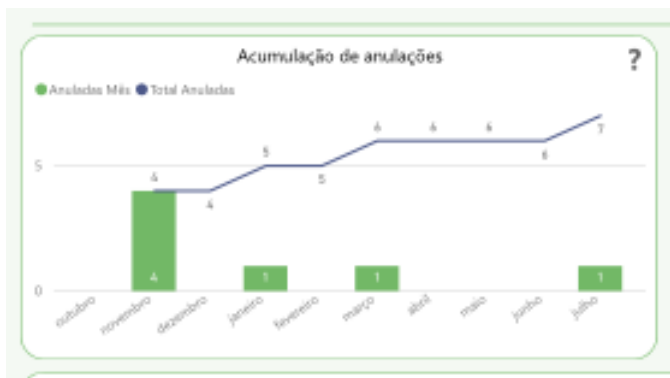
4) Definição de prescrito: ✓



Prescrito - Estudante que, devido ao insucesso escolar repetido, fica impedido de continuar a frequentar o curso, em que estava matriculado e inscrito, durante dois semestres consecutivos, conforme previsto na lei 37/2003 e no regulamento de prescrições da Universidade de Aveiro. O ano letivo associado ao estudante prescrito corresponde ao ano em que o estudante fica impedido de frequentar o curso em que estava matriculado e inscrito.

5) Sobre as acumulações

Não temos a informação do motivo da anulação (é importante). No Ficheiro FICA está numa das folhas, no FICA-alargamento estará numa coluna. (só disponível nos dados 2019/20)

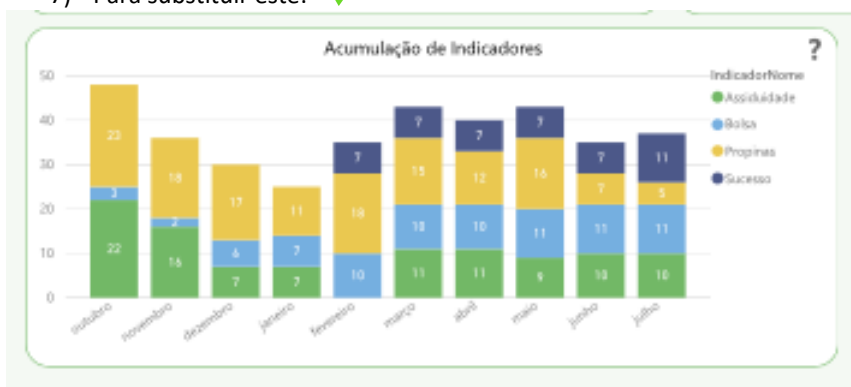


6) Aqui: ✓

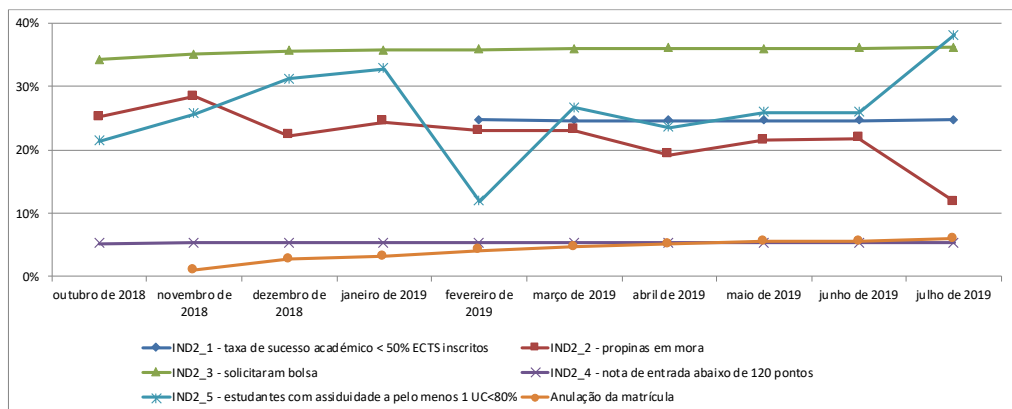


Acho que falta se é 1º semestre ou final.

7) Para substituir este: ✓



Um gráfico de linhas, tipo este, mas com os valores absolutos e não os relativos (%):



8) No Ponto ? do gráfico: ✓



Está errado:

