Universidade de Aveiro

2015

Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro

João Nelson Ferreira Rocha

Visualização da Evolução Temporal de Rankings

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Sistemas de Informação, realizada sob a orientação científica do Dr.-Ing. Joaquim João Estrela Ribeiro Silvestre Madeira, Professor Auxiliar, e da Doutora Maria Beatriz Alves de Sousa Santos, Professora Associada com Agregação, ambos do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro.

o júri

presidente	Professor Doutor Paulo Miguel de Jesus Dias Professor Auxiliar da Universidade de Aveiro
arguente principal	Professora Doutora Maria Beatriz Duarte Pereira do Carmo Professora Auxiliar da Faculdade de Ciências da Universidade de Lisboa
orientador	Professor Dr.-Ing. Joaquim João Estrela Ribeiro Silvestre Madeira Professor Auxiliar da Universidade de Aveiro

agradecimentos

Agradeço ao Dr.-Ing. Joaquim João Estrela Ribeiro Silvestre Madeira, Professor Auxiliar, e à Doutora Maria Beatriz Alves de Sousa Santos, Professora Associada com Agregação, ambos do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro, pela orientação e disponibilidade, partilha de conhecimentos, opiniões e críticas, e um grande apoio no solucionar de dúvidas e dificuldades que foram surgindo no desenvolvimento da dissertação.

A todos os que, direta ou indiretamente contribuíram na elaboração deste projeto e fizeram parte da minha formação.

Agradeço também aos meus pais, irmão e familiares pela confiança transmitida, pelo apoio, incentivo, amizade e paciência durante o processo.

palavras-chave visualização de informação, rankings, evolução temporal.

resumo

As tabelas classificativas agregam, num dado instante temporal, a classificação ordenada de um conjunto de entidades e alguns dos seus atributos. Esta dissertação tem como objetivo desenvolver visualizações interativas de rankings (classificações) e integra-las num *dashboard* apropriado, onde a sua aplicação é testada com diferentes casos de estudo. As visualizações desenvolvidas foram validadas recorrendo a duas metodologias de avaliação diferentes, avaliação heurística e testes de usabilidade, para avaliar a sua utilidade.

keywords information visualization, rankings, time evolution.

abstract

Ranking tables present, at a given time, the sorted ranking values of a group of entities and some of their attributes.
This thesis aims to design and develop interactive visualizations for time-varying rankings and integrate them in a dashboard. The visualizations were tested using different case studies. The developed visualizations were tested using two different evaluation methodologies, heuristic evaluation and usability testing, to assess their usability.

Índice

1. Introdução ... 1
   1.1. Contexto e Motivação ... 1
   1.2. Objetivos ... 1
   1.3. Organização da Dissertação ... 2
2. Estado da Arte ... 3
   2.1. Visualização de Informação ... 3
      2.1.1. Pipeline de Visualização de Informação ... 3
      2.1.2. Visualizações de Dados ... 4
      Tabelas ... 5
      Gráficos ... 5
      Georreferenciação ... 9
      Dados Multivariados ... 9
      Dashboards ... 10
      2.1.3. Técnicas de Interação ... 11
      2.1.4. Ferramentas para Visualização de Informação ... 12
      Prefuse Visualization Toolkit ... 12
      Protovis ... 13
      JavaScript InfoVis Toolkit (Jit) ... 14
      D3.js: Data-Driven Documents ... 15
      Tableau ... 16
      Google Charts ... 17
      2.1.5. Testes de Usabilidade ... 18
   2.2. Visualização da Evolução de Rankings ... 18
      2.2.1. Técnicas de Visualização Convencionais ... 18
      2.2.2. Representações Visuais inovadoras ... 20
      RankExplorer: Visualization of Ranking Changes in Large Time Series Data ... 20
      LineUp: Visual Analysis of Multi-Attribute Rankings. ... 22
      À Table! Improving Temporal Navigation in Soccer Ranking Tables ... 23
      Investigating the Direct Manipulation of Ranking Tables ... 25
      2.2.3. Requisitos Gerais ... 26
   2.3. Conclusões ... 27

3. Um Modelo para a Visualização de Rankings 29
3.1. Requisitos para análise de Rankings 29
3.2. Casos de Estudo 29
Rankings de Futebol 30
Rankings de Escolas Secundárias 30
3.3. Tecnologia selecionada 32
Tecnologias Complementares 32
3.4. Arquitetura desenvolvida 33
3.4.1. Camada de dados 33
3.4.2. DashBoard e Fluxo de Dados 33
3.4.3. Rankings de Futebol 34
3.4.4. Rankings de Escolas 39
Evolução Médias Escolares 41
3.4.5. Conclusões 45
4. Testes e Avaliação 47
4.1. Avaliação Heurística 47
4.1.1. Resultados 47
4.2. Testes de Usabilidade 51
4.2.1. Metodologia 51
4.2.2. Participantes 52
4.2.3. Resultados 52
4.2.4. Conclusões 58
5. Conclusão 59
5.1. Conclusão 59
5.2. Trabalho Futuro 60
6. Referências 61
Apêndice A Ficheiros de Dados 63
Apêndice A.1 Dados de Rankings Desportivos 63
Apêndice A.2 Dados de Rankings Escolares 64
Apêndice B Testes de Usabilidade 65
Apêndice B.1 Tarefas de Utilizadores 65
Apêndice B.2 Ficha do Observador 71
Anexos 73
Anexo 1: Avaliação de João Rocha 74
Anexo 2: Avaliação de Pedro Santos 84

Índice de Figuras

Figura 2.1. - InfoVis Pipeline baseado em (Liu et al. 2014) 3
Figura 2.2 – Representação em Tabela 5
Figura 2.3 - Gráfico de Barras 6
Figura 2.4 - Gráfico de Linhas 6
Figura 2.5 - Pie Chart 6
Figura 2.6 - Histograma 6
Figura 2.7 - Bubble Chart 7
Figura 2.8 - Tree Map 7
Figura 2.9 - Area Chart 8
Figura 2.10 – Stream Graph 8
Figura 2.11 - Heat Map 8
Figura 2.12 - Representação de dados geográficos 9
Figura 2.13 - Exemplo de representações de dados multivariados 9
Figura 2.14- Exemplo de *Dashboards* 10
Figura 2.15 - Visualização desenvolvida com Prefuse. 12
Figura 2.16 - Visualização Index Chart com Protovis 13
Figura 2.17 - Visualização HyperTree com Jit. 14
Figura 2.18 – Dashboard desenvolvido com D3.js. 15
Figura 2.19 - Dashboard desenvolvido em Tableau 16
Figura 2.20 - Composição de Visualizações em Google Charts 17
Figura 2.21- Tabelas de Rankings 19
Figura 2.22 - Gráficos de Linhas 19
Figura 2.23 - Gráficos de Coordenadas Paralelas 20
Figura 2.24 – RankExplorer visualization, como apresentado no tutorial 21
Figura 2.25 – LineUp, classificação das universidades de topo de acordo com "QS World University Ranking 2012", com atributos e pesos alterados, em comparação com os dados oficiais. Retirado de (Gratzl et al. 2013) 22
Figura 2.26 – Àtable Web Page, retirado de *(Perin et al. 2014)* 24

Figura 2.27- Exemplo de utilização de manipulação em tabelas de Rankings. Tabela de *Rankings* de equipas de futebol onde, ao arrastar o cursor, é apresentado um gráfico de linhas para manipular os dados da tabela em relação ao tempo, apresentando a evolução do *Ranking* para os elementos selecionados. 25
Figura 3.1 - Exemplo do Dashboard desenvolvido 33
Figura 3.2 - *Slider* Temporal para Jornadas de Futebol 34
Figura 3.3 - Tabela de Rankings de Futebol 35
Figura 3.4 - Gráfico de Coordenadas Paralelas para Rankings de Futebol 35
Figura 3.5 - Gráfico de Coordenadas paralelas com seleção de Equipas 36
Figura 3.6 - Gráfico de Linhas que apresenta evolução de classificações ao longo da época 36
Figura 3.7 - Gráfico de Linhas para a Evolução dos Pontos com todas as Equipas 37
Figura 3.8 - Gráfico de Linhas para a evolução de pontos com seleção de equipas 37
Figura 3.9 - Gráfico de Barras que mostra a Diferença de Golos por equipa 38
Figura 3.10 - Gráfico que mostra a performance das equipas até a jornada atual 38
Figura 3.11 – Tabela de Rankings Escolares 40
Figura 3.12 – Coordenadas Paralelas, as 10 melhores escolas 40
Figura 3.13 - *Dashboard* de Evolução de Médias e Rankings Escolares 41
Figura 3.14 - Tipos de Legenda do *HeatMap* para o Mapa de Portugal 42
Figura 3.15 - Exemplo da representação Geográfica da diferença entre a Média Distrital e a Média Nacional 42
Figura 3.16 - Comparação entre os mapas resultantes para a relação entre Médias Distritais com Nacional e Médias Positivas/Negativas 43
Figura 3.17 - Exemplo de Escolha de Concelhos para o Distrito de Aveiro, e de Legendas para Concelhos Diferentes 44
Figura 3.18 - Gráfico de Estrela a apresentar Médias de Escolas de Concelhos Diferentes, Escolas de Aveiro (Esquerda) e Escolas de Ílhavo (Direita) 44
Figura 3.19 - Gráfico de Barras a apresentar Médias de Escolas de Concelhos Diferentes, Escolas de Aveiro (Esquerda) e Escolas de Ílhavo (Direita) com exemplo de *Tooltip* 45
Figura 4.1 - Ordenação da Tabela 48
Figura 4.2 - Exemplo de sobreposição de texto 48
Figura 4.3 - Nova posição de texto, sem sobreposição 48
Figura 4.4 - Problema com contraste entre cores das barras e texto 49
Figura 4.5 - Alteração da cor do texto para as barras do gráfico 49
Figura 4.6 - Linha da Média Distrital por cima dos seus valores 50
Figura 4.7 - Resultados de Respostas às Tarefas 53
Figura 4.8 - Facilidade das Tarefas por Utilizador 54
Figura 4.9 - Duração média das tarefas 55
Figura 4.10 - Escala de Usabilidade e interpretação de resultados (Bangor et al. 2009) 57

1. Introdução

1.1. Contexto e Motivação

As tabelas classificativas (rankings) agregam, num dado instante temporal, a classificação ordenada de um conjunto de entidades e alguns dos seus atributos. Exemplos típicos são as tabelas classificativas de campeonatos desportivos e os rankings de escolas secundárias ou universidades.

A comparação com tabelas anteriores e a perceção (visual) da evolução temporal e de tendências são fundamentais para a análise da evolução de um dado ranking. O eventual grande número de entidades e atributos representados, bem como eventuais lacunas nos dados, são dificuldades acrescidas que podem prejudicar uma correta análise.

Muitas vezes, estas tabelas classificativas não são disponibilizadas com o intuito de efetuar comparações de rankings entre diferentes instantes de tempo, apesar de, regularmente, ser esta a análise pretendida por quem procura estes tipos de dados. Para além disto, também são encontradas visualizações que não oferecem qualquer tipo de interatividade com os utilizadores, o que facilitaria a exploração dos dados. (Perin et al. 2014)

1.2. Objetivos

O trabalho apresentado nesta dissertação teve como objetivo geral o desenho, desenvolvimento, implementação e avaliação de representações interativas para visualização da evolução temporal de rankings. Essas visualizações devem suportar a seleção do instante de tempo e permitir avaliar a evolução temporal dos dados classificativos que são apresentados de forma gráfica e interativa aos utilizadores, facilitando a perceção dos dados e de eventuais padrões.

Esta dissertação tem também a finalidade de validar a aplicação das visualizações desenvolvidas com diferentes casos de estudo.

As visualizações devem ser disponibilizadas numa plataforma Web, facilitando a execução de testes de usabilidade com utilizadores. Com esses testes é possível avaliar e validar a usabilidade das visualizações desenvolvidas. Desta forma é fácil concluir se facilitam a perceção dos dados, ou se é necessário fazer alterações ou adicionar novas funcionalidades ao *dashboard* e às visualizações desenvolvidas, de acordo com o indicado pelos utilizadores ou concluído pela forma como estes utilizam as visualizações.

1.3. Organização da Dissertação

Esta dissertação está dividida em cinco capítulos, incluindo o capítulo atual.

No Capitulo 2 encontra-se descrito o Estado da Arte, dividido em duas secções principais: Visualização de Informação e Visualização da Evolução de *Rankings*. Na primeira, são apresentados os conceitos fundamentais para Visualização de Informação, assim como uma breve lista de visualizações comummente utilizadas e encontradas na Web. Também ainda nesta secção, é apresentada uma lista de ferramentas a ponderar para o desenvolvimento do trabalho, assim como uma breve introdução à avaliação de usabilidade. Na segunda secção, são referenciados artigos que apresentam ferramentas desenvolvidas com o intuito de representar informação relacionada com rankings e a sua evolução temporal.

O Capítulo 3 descreve o processo de desenvolvimento do trabalho, iniciando com uma listagem de requisitos para visualização de Rankings e com a apresentação dos casos de estudo e dados utilizados. O capítulo prossegue com a apresentação das visualizações desenvolvidas para cada caso de estudo, descrevendo a sua utilidade e explicando a interatividade associada.

O Capitulo 4 apresenta o processo de avaliação das visualizações, expondo os resultados obtidos com os testes efetuados e conclusões relativas à sua utilização.

Por último, no Capítulo 5, são apresentadas as principais conclusões deste trabalho e são indicadas algumas sugestões para o trabalho futuro.

2. Estado da Arte

Este capítulo apresenta os conceitos fundamentais e as publicações mais recentes relacionadas com o tema desta dissertação. O capítulo está dividido em duas secções principais: Visualização de Informação e Visualização da Evolução de *Rankings*. Apresenta-se, de uma forma geral, o que é a Visualização de Informação e algumas das suas características, e é feita uma análise mais focada no tema da evolução temporal de *Rankings*.

2.1. Visualização de Informação

Visualização de Informação (InfoVis – Information Visualization) abrange o conjunto de métodos e técnicas, que permitem a geração computacional de representações visuais de informação variada, dados abstratos e não-espaciais. O seu objetivo principal é auxiliar os utilizadores a melhor analisar e compreender a informação, através da exploração progressiva e interativa dessas representações.

Um aspeto importante é manter uma boa relação entre a qualidade e a quantidade de informação apresentada, o que implica que sejam desenvolvidos métodos para melhorar as representações dos dados e manter as visualizações compatíveis com grandes quantidades de informação.

2.1.1. Pipeline de Visualização de Informação

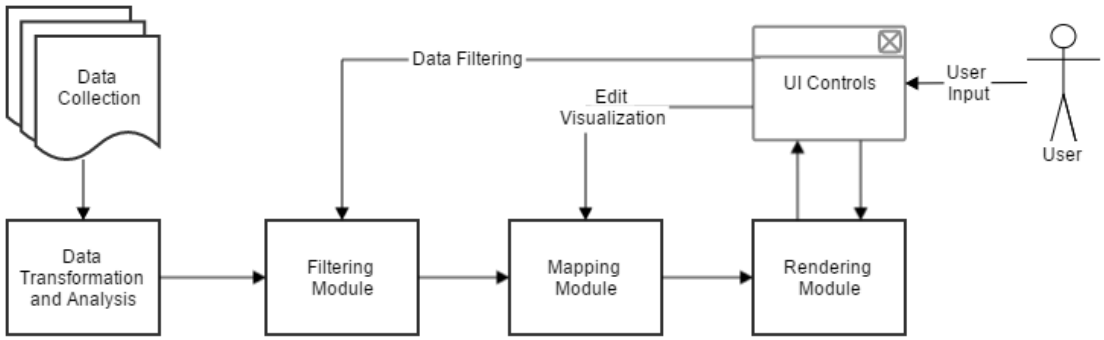

Figura 2.1. - InfoVis Pipeline baseado em **(Liu et al. 2014)**

A sequência de operações habituais em Visualização de Informação está apresentada no *pipeline* da Fig. 2.1, que se encontra estruturado em quatro módulos, sendo também apresentada a interface com utilizador (Liu et al. 2014):

- **Transformação e análise de dados**: Os dados originais podem estar armazenados em formato textual simples, por exemplo ficheiros CSV, ou numa base de dados estruturada, por exemplo SQL. O objetivo do módulo de transformação e análise de dados é receber os dados, analisar a sua estrutura inicial para, por exemplo, corrigir valores em falta e reestrutura-los de acordo com a informação pretendida;
- **Filtragem:** O módulo de filtragem permite determinar que conjuntos de dados (informação relevante) serão representados na visualização;
- **Mapeamento de informação:** Os dados filtrados são depois mapeados em primitivas de visualização e seus atributos. Entendem-se por primitivas de visualização pontos, segmentos de reta, polígonos, etc., e por atributos a sua posição, tamanho e cor, por exemplo;
- **Representação de Informação**: Finalmente, com o módulo de representação, é apresentada a visualização construída, que o utilizador pode explorar com a utilização de técnicas de interação apropriadas;
- **Interface com o utilizador**: Disponibiliza funcionalidades de interação que possibilitam aos utilizadores escolher novas visualizações, filtrar novos dados a visualizar, ou alterar o tipo de mapeamento (escalas, cores, etc.).

2.1.2. Visualizações de Dados

Existem várias técnicas de representação de dados que, de acordo com (Liu et al. 2014), foram categorizados em quatro grupos distintos, tendo por base as características dos dados tratados. Nesta secção serão apresentadas e descritas os tipos de representações habitualmente aplicadas.

Tabelas

Ranking	Equipa	Pontos	Jogos	Vitorias	Empates	Derrotas	Golos Marcados	Golos Sofridos
1	Benfica	65	34	21	2	3	63	13
2	FC Porto	62	34	19	5	2	58	11
3	Sporting	56	34	16	8	2	50	23
4	SC Braga	47	34	14	5	7	39	17
5	V. Guimarães	40	34	11	7	8	38	28
6	Paços Ferreira	37	34	10	7	9	31	35
7	Rio Ave	36	34	9	9	8	33	32
8	Belenenses	36	34	9	9	8	25	27
9	Nacional	33	34	9	6	11	32	36
10	Moreirense	32	34	8	8	10	23	29
11	Marítimo	31	34	9	4	13	32	36
12	Estoril Praia	28	34	6	10	10	30	45
13	Boavista	28	34	8	4	14	21	40
14	Académica	26	34	4	14	8	19	30
15	V. Setúbal	24	34	6	6	14	18	40
16	Arouca	23	34	6	5	15	20	40
17	Gil Vicente	19	34	3	10	13	20	43
18	Penafiel	17	34	4	5	17	25	52

Figura 2.2 – Representação em Tabela

A informação pode ser disponibilizada e analisada recorrendo a tabelas estruturadas. Esta é uma forma muito comum de apresentar informação, assim como a mais fácil de desenvolver e, em muitos casos, é a forma mais simples de permitir a análise dos dados.

Assim, embora possa não ser fácil verificar tendências nos dados, podem analisar-se os atributos de um elemento, ou a diferença de valores dos atributos entre diferentes elementos, recorrendo à reordenação de valores.

Gráficos

Os vários tipos de gráficos mais comuns são um formato fácil de representação de informação de forma visual, consistindo na representação de um ou mais atributos dos elementos a analisar, permitindo uma melhor comparação entre eles. Estas representações podem ser apresentadas de uma forma estática, ou podem conter técnicas de interação, possibilitando a manipulação e remapeamento dos dados de acordo com ações dos utilizadores. (Heer et al. 2010), (Hom et al. 2014), (Tableau 2014).

Há quatro tipos de representações gráficas mais comuns:

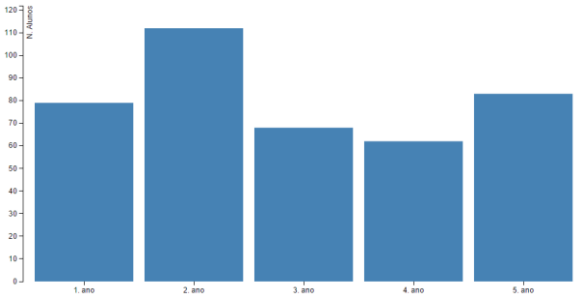

Figura 2.3 - Gráfico de Barras

Gráfico de Barras, gráfico com barras retangulares com dimensões proporcionais aos valores que representam. As barras podem ser desenhadas vertical ou horizontalmente, e providenciam uma forma fácil de comparar informação. Estes gráficos são usados para comparar o mesmo atributo de diferentes elementos.

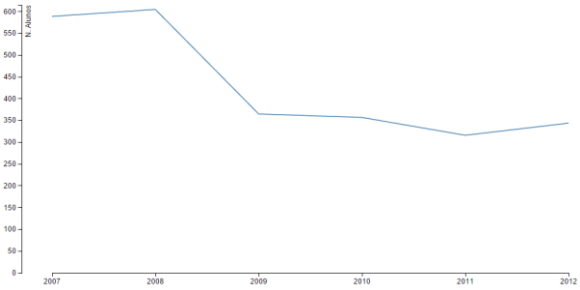

Figura 2.4 - Gráfico de Linhas

Gráfico de Linhas, gráfico que conecta pontos representando os dados, relativamente a dois eixos (vertical e horizontal). Resulta numa visualização com o objetivo de apresentar tendências de valores num período de tempo.

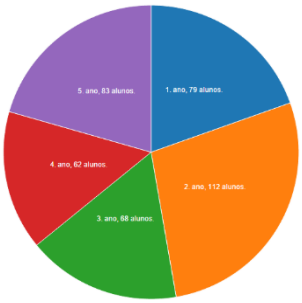

Figura 2.5 - Pie Chart

Pie Chart, gráfico que mostra valores relativos ou percentagens como fatias de uma tarte. Este gráfico é para ser usado quando se pretende representar proporções de um todo. Se a finalidade for a comparação de valores quantitativos, é melhor optar por outro tipo de gráfico.

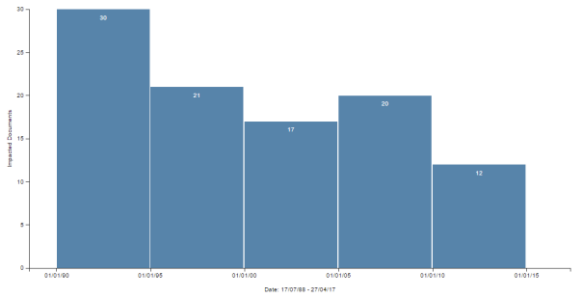

Figura 2.6 - Histograma

Histograma, gráfico de barras adjacentes, com dimensões relativas à frequência do atributo em relação ao intervalo de valores, o seu objetivo é a análise da distribuição dos dados por grupos. Por exemplo, as notas de exames para uma determinada disciplina, podem ser distribuídas por intervalos de notas.

Outras representações gráficas também utilizadas, mas menos comuns, são:

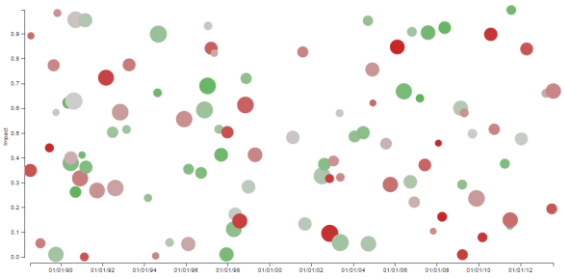

Figura 2.7 - Bubble Chart

Scatter Plot, gráfico para representar os dados como um conjunto de pontos, cada um com atributos que determinam a sua posição horizontal e vertical em relação a dois eixos, que facilita a análise da relação entre diferentes atributos, representando tendências e concentrações nos dados. Por exemplo, alunos do género masculino e feminino que obtiveram aprovação ao longo dos anos.

Bubble Chart, um *scatter plot* bidimensional com um terceiro atributo que define o tamanho dos pontos no gráfico. Este terceiro atributo (diâmetro dos círculos) dá mais informação aos utilizadores. Por exemplo, o diâmetro dos círculos indica quando houve um maior ou menor número de alunos aprovados.

Figura 2.8 - Tree Map

Tree Map, gráfico que representa dados na forma de uma árvore estruturada, com um conjunto de retângulos aninhados, cuja área corresponde ao valor quantitativo de um dos seus atributos. Outras dimensões (atributos) podem ser representadas com cor, por exemplo. O seu objetivo principal é representar os dados de forma hierárquica e a sua proporção em relação a todo o conjunto de dados (Hom et al. 2014). Por exemplo, o sistema de ficheiros e o espaço em disco que estes ocupam.

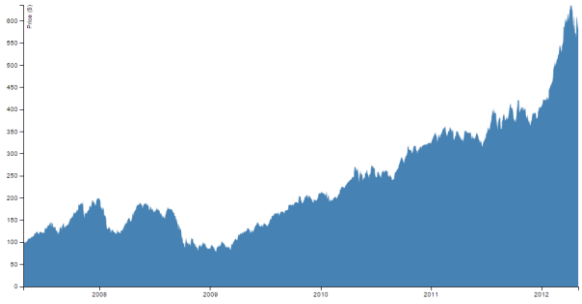

Figura 2.9 - Area Chart

Area Chart apresenta valores quantitativos dos dados e é baseado no gráfico de linhas. A diferença é que, nesta representação, a área entre a linha e o eixo horizontal está preenchida com cores ou texturas, para indicarem o volume dos dados.

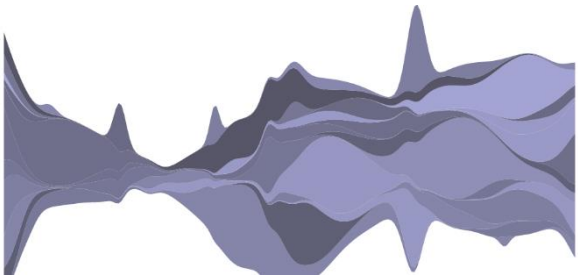

Figura 2.10 – Stream Graph

Streamgraph, um gráfico de áreas empilhado, onde estas são desenhadas em relação a um eixo central, ao contrário de um eixo inferior.

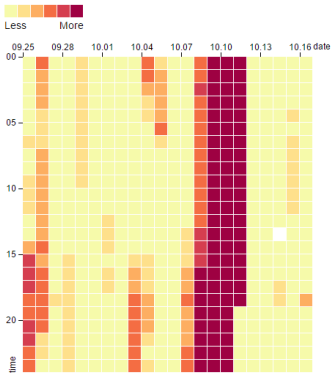

Figura 2.11 - Heat Map

Heat Map, gráfico onde os valores dos atributos são apresentados ao longo de uma matriz, representando cada intervalo de valores com cores diferentes. São uma boa forma de visualizar onde o cruzamento de dois atributos é mais forte e mais fraco (Hom et al. 2014), ou simplesmente, quais dos atributos são superiores ou inferiores a um dado valor.

Georreferenciação

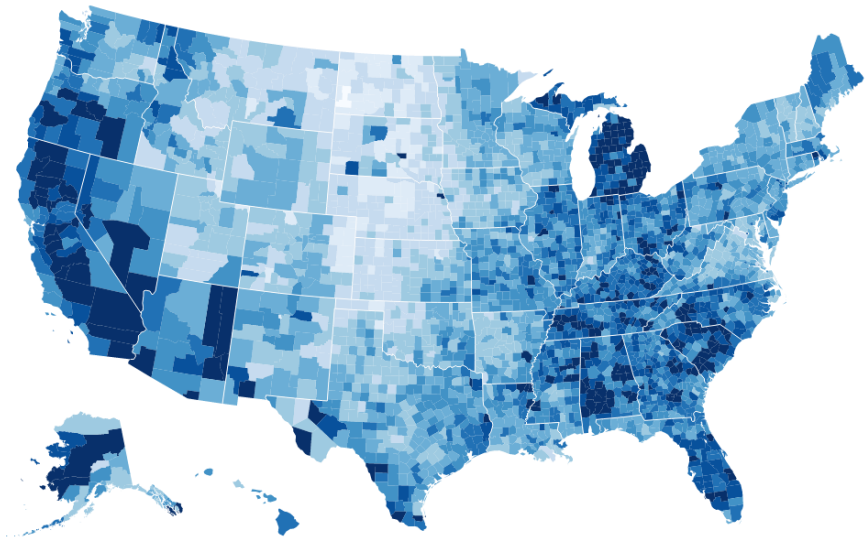

Figura 2.12 - Representação de dados geográficos[1]

Os Sistemas de Informação Geográfica (SIG) permitem interagir com as visualizações e ver as mudanças imediatas nos mapas. Isto não só proporciona uma diferença quantitativa no número de resultados que se podem visualizar, mas também uma diferença qualitativa na forma como os dados podem ser analisados. Mapas são agora uma interface de visualização de dados geográficos que permitem, para além da exploração da informação associada, a sua manipulação (Liu et al. 2014).

Dados Multivariados

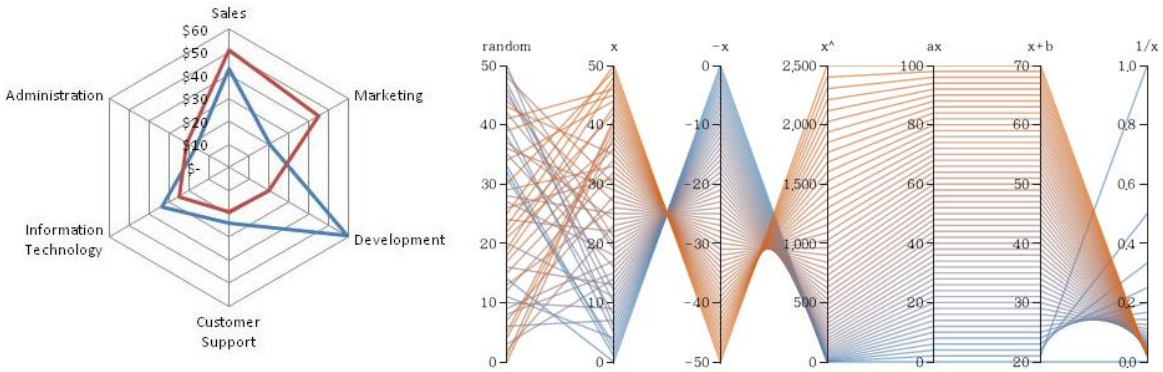

Figura 2.13 - Exemplo de representações de dados multivariados

Técnicas de representação que suportam a análise de dados multivariados, permitem analisar o comportamento de mais de que um atributo por elemento, em simultâneo.

[1] http://bl.ocks.org/mbostock/4060606

Para este tipo de dados, é normalmente utilizada uma representação que mapeie os diferentes atributos em paralelo. Dois exemplos comuns deste tipo de visualizações são:

- **Radar chart**, ou "**Spider Chart**" (fig 2.13.) é um gráfico bidimensional com três ou mais variáveis quantitativas. Os valores destas variáveis são representados ao longo de eixos que iniciam no centro do gráfico, normalmente terminando com um anel externo. São utilizados para facilmente apresentar tendências nos dados relativamente ao valor dos atributos, por exemplo, no futebol para representar os pontos fortes e fracos dos jogadores, podendo haver comparação entre vários.
- **Gráfico de Coordenadas Paralelas** (fig 2.13.) é utilizado para apresentar vários atributos de um ou mais elementos. Esta representação consiste em n eixos paralelos, tipicamente verticais, onde os atributos são mapeados. Os elementos são representados como linhas poligonais com vértices nos eixos paralelos. Esta assemelha-se a representações de séries temporais, mas, neste caso, os dados não correspondem a instantes de tempo e, por isso, podem ter diferentes ordenações e escalas (Heer et al. 2010).

Dashboards

Figura 2.14- Exemplo de *Dashboards*

Um *Dashboard* é uma coleção de representações gráficas e tabelas, apresentadas no mesmo espaço e interligadas, para comparação e visualização de dados em simultâneo. Os *dashboards* são uma forma normalmente utilizada para manter os gráficos e tabelas considerados mais importantes rapidamente ao alcance do utilizador.

Para que um *dashboard* seja funcional, é importante ter em conta mais que as visualizações utilizadas. É importante a utilização do tipo de letra e estilos para chamar a atenção para determinados detalhes, assim como os padrões de cores a aplicar nas páginas e visualizações em geral. Também é importante que a

informação siga um determinado fluxo, que o utilizador consiga acompanhar. (Juice 2009) e (Tableau 2014) contêm, com mais detalhe, boas práticas no *design* de *Dashboards* e na estruturação da sua informação.

Com uma composição de visualizações, as técnicas anteriormente apresentadas podem ser utilizadas em conjunto como forma de simplificar a análise e compreensão da informação. Estas composições são vulgarmente utilizadas para adaptar a visualização da informação a um determinado público ou dados.

2.1.3. Técnicas de Interação

É importante a possibilidade de interagir com as visualizações, o que pode facilitar bastante a análise de informação e é uma boa forma de captar a atenção dos utilizadores. De uma forma geral, a interatividade ajuda os utilizadores a seguir alterações nos dados e a associar informação entre representações.

De acordo com (Yi et al. 2007), as técnicas de interação estão divididas em sete categorias:

- **Selecionar (*select*)**, "marcar" um elemento como importante. Possibilita aos utilizadores selecionar um elemento, ou mais, da coleção de dados como importante, para acompanhar as suas alterações;
- **Explorar (*explore*)**, mostrar novos dados. Possibilita ao utilizador examinar diferentes coleções de dados. Este tipo de interação não altera os dados da visualização, mas sim os dados que estão a ser visualizados, por exemplo, são representados novos elementos enquanto outros são retirados. Um exemplo de interação é o *panning* dos dados;
- **Reconfigurar (*reconfigure*)**, mostrar diferentes disposições. Disponibiliza diferentes perspetivas dos dados aos utilizadores, alterando o arranjo espacial das visualizações, possibilitando aos utilizadores alterarem a forma como os elementos estão organizados ou desenhados. Exemplo desta interação é a reorganização de colunas numa tabela;
- **Codificar (*encode*)**, mostrar novas visualizações. Possibilita ao utilizador alterar a representação visual dos dados ou de cada elemento (por exemplo, cor, tamanho). Alterar como os dados são apresentados (por exemplo, alterar de um *pie chart* para um *bar chart*) é um exemplo desta interação;
- **Resumir/Elaborar (*abstract/elaborate*)**, mostrar mais ou menos informação. Possibilita aos utilizadores alterarem o nível de abstração dos dados representados. Este tipo de interação possibilita aos utilizadores a utilização de diferentes níveis de detalhe. Um exemplo é a expansão de elementos (nós) de uma visualização *treemap*;
- **Filtrar (*filter*)**, mostrar dados de interesse (condicionalmente). Possibilita aos utilizadores alterarem o conjunto de dados a ser apresentados com base em condições ou intervalos de valores específicos. Apenas os dados que cumpram os critérios escolhidos são apresentados, os que não cumprem são escondidos ou apresentados de forma diferente, mas os dados originais não são alterados, para o caso de reversão de filtros;
- **Conectar (*connect*)**, mostrar elementos relacionados. Destacamento de associações e de relações entre os elementos representados e apresentação de elementos escondidos que sejam relevantes para

um elemento em específico. Um exemplo geral é, ao selecionar um elemento (nó) num *TreeMap*, são destacados os nós que estão diretamente relacionados.

2.1.4. Ferramentas para Visualização de Informação

Apresentam-se as ferramentas mais comummente utilizadas para o desenvolvimento de visualizações para a Web. Foi com este levantamento que se selecionou a ferramenta a utilizar para o desenvolvimento do trabalho descrito nesta dissertação.

Prefuse Visualization Toolkit

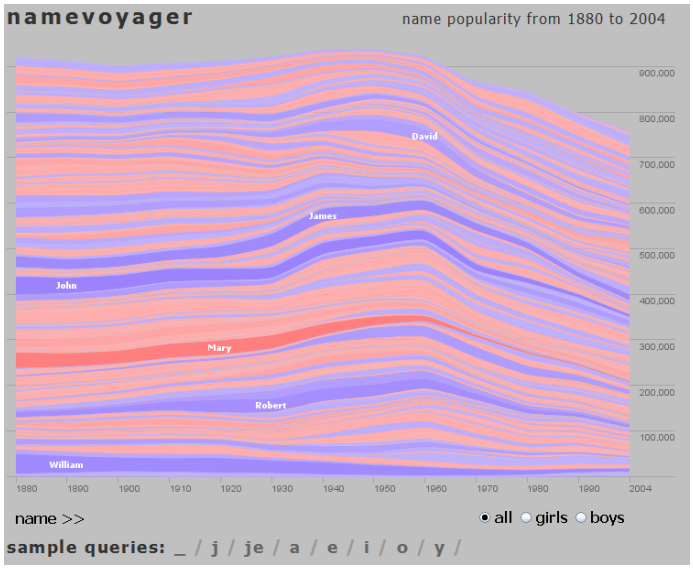

Figura 2.15 - Visualização desenvolvida com Prefuse.[2]

Prefuse[3] (Heer et al. 2005) é uma *framework* para Java, que contém uma biblioteca para criar visualizações de dados interativas. Prefuse Flare é uma *toolkit* para criar visualizações e animações para *ActionScript* e *Adobe Flash Player*.

O Prefuse utiliza estruturas de dados otimizadas para a construção de tabelas, gráficos e árvores de diferentes tipos, técnicas de codificação visuais, suporta animações, conexão com bases de dados e métodos de pesquisa.

[2] http://prefuse.org/gallery/namevoyager/ (Consultado em 20/10/2015)
[3] http://prefuse.org/ (Consultado em 20/10/2015)

Protovis

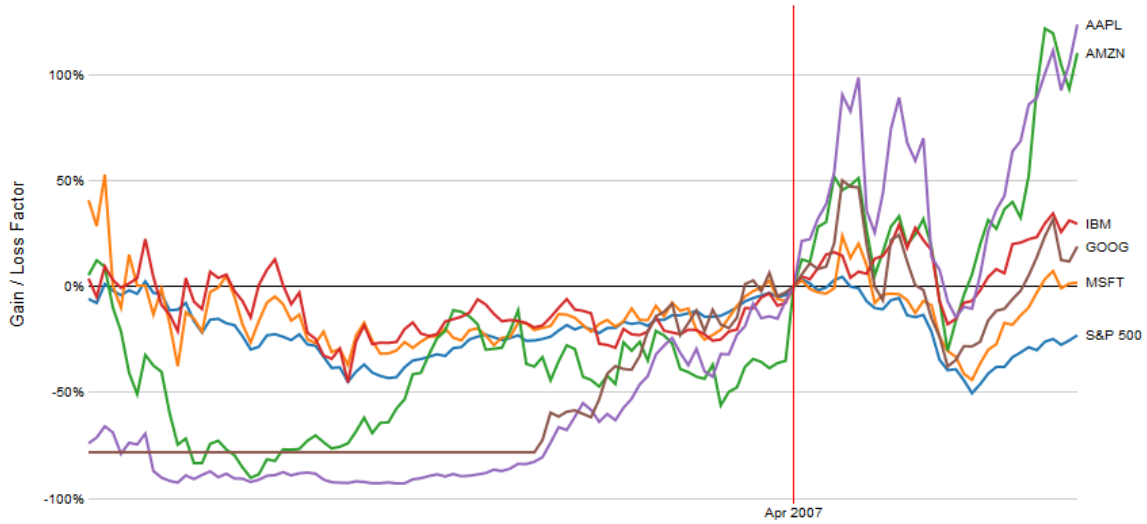

Figura 2.16 - Visualização Index Chart com Protovis[4]

Protovis[5] (Bostock & Heer 2009) é uma biblioteca *open-source*, sob a licença BSD. Usa *Javascript* e SVG (*Scalable Vector Garphics*) para criar visualizações Web interativas.

Protovis deixou de estar sob desenvolvimento ativo (atualizações ativas) porque a equipa migrou o seu desenvolvimento para uma nova biblioteca, D3.js. Isto verifica-se na similaridade entre as duas bibliotecas, onde o aspeto das visualizações criadas é muito semelhante, mas o D3 oferece técnicas de interação e de desenho mais fáceis e eficazes de utilizar e desenvolver.

[4] http://mbostock.github.io/protovis/ex/index-chart.html (Consultado em 20/10/2015)
[5] http://mbostock.github.io/protovis/ (Consultado em 20/10/2015)

JavaScript InfoVis Toolkit (Jit)

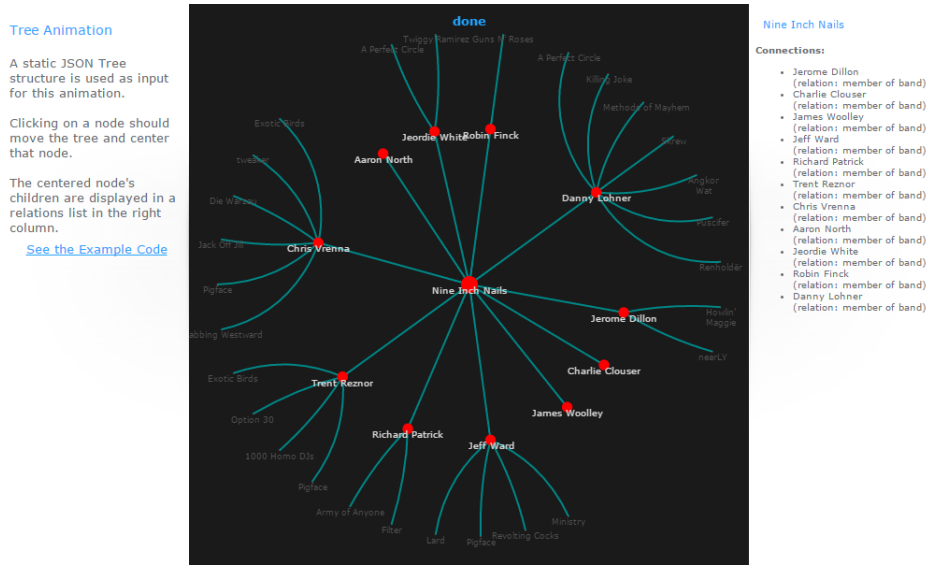

Figura 2.17 - Visualização HyperTree com Jit[6].

Jit (*JavaScript InfoVis Toolkit*)[7] é uma biblioteca de Visualização de Informação baseada em Java, com estruturas de dados e algoritmos de visualização, para construir representações de dados interativas para a Web.

Jit implementa recursos avançados de visualização de informação como *TreeMaps*, *SpaceTrees*, técnicas de foco-contexto para desenhar *Hyperbolic Trees* (*Hypertrees*), uma disposição radial das árvores com animações avançadas (*RGraph*) e outras visualizações.

O artigo de (Fekete 2004) contem uma análise extensiva sobre esta *toolkit*, e pode ser consultado para se perceber melhor o seu funcionamento.

[6] http://philogb.github.io/jit/static/v20/Jit/Examples/Hypertree/example1.html (Consultado em 20/10/2015)
[7] http://philogb.github.io/jit/ (Consultado em 20/10/2015)

D3.js: Data-Driven Documents

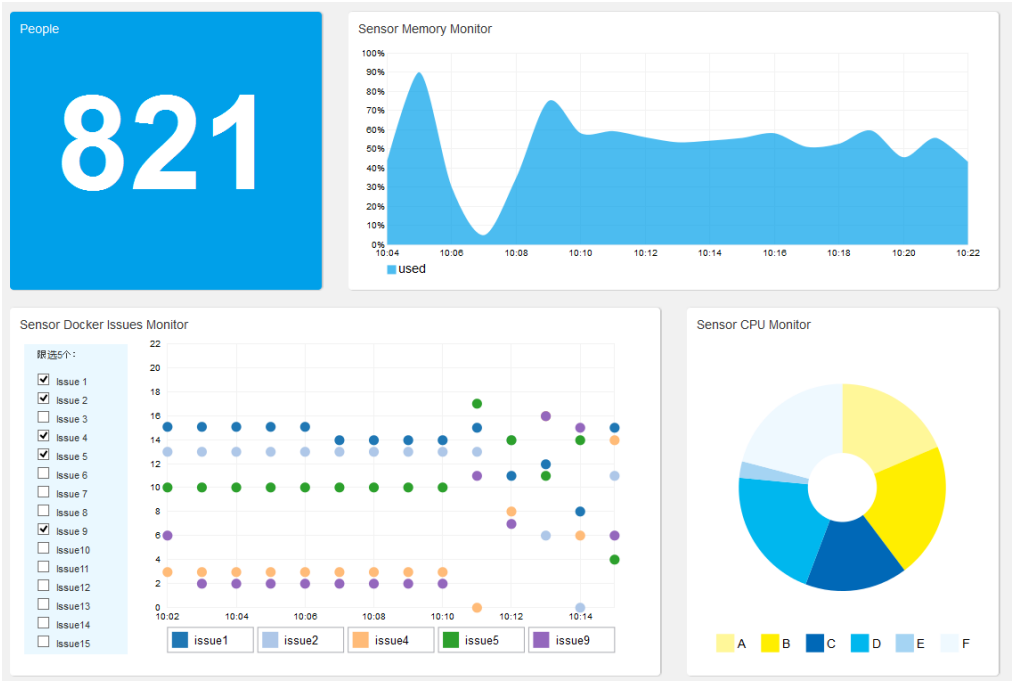

Figura 2.18 – Dashboard desenvolvido com D3.js[8].

A D3.js[9] (Bostock et al. 2011) é uma biblioteca em *JavaScript* desenvolvida por Mike Bostock, utilizada para a construção de visualizações avançadas e compatíveis com ambientes *Web*.

Esta biblioteca para visualização de informação utiliza linguagens e protocolos conhecidos para a criação das visualizações e manipulação dos dados, nomeadamente, HTML, CSS, *JavaScript* e elementos SVG. Isto requer que haja, pelo menos, um mínimo de conhecimento prévio para se utilizar o D3.js. Mas, para os utilizadores, isto facilita o seu acesso e utilização, porque há uma grande compatibilidade entre dispositivos e *software*.

D3.js possibilita a manipulação de dados, oferecendo funções especializadas na sua leitura e processamento. Para além disto, a D3.js manipula diretamente os elementos dos documentos *Web* com base nos dados, combinando componentes de visualização e uma abordagem orientada aos dados para manipulação de elementos DOM (*Document Object Model*).

Para mais informações podem ser consultados (Dewar 2012), (Murray 2013), (Malcolm Maclean 2015).

[8] http://hijiangtao.github.io/ss-vis-component/ (Consultado em 20/10/2015)
[9] http://d3js.org/ (Consultado em 20/10/2015)

Tableau

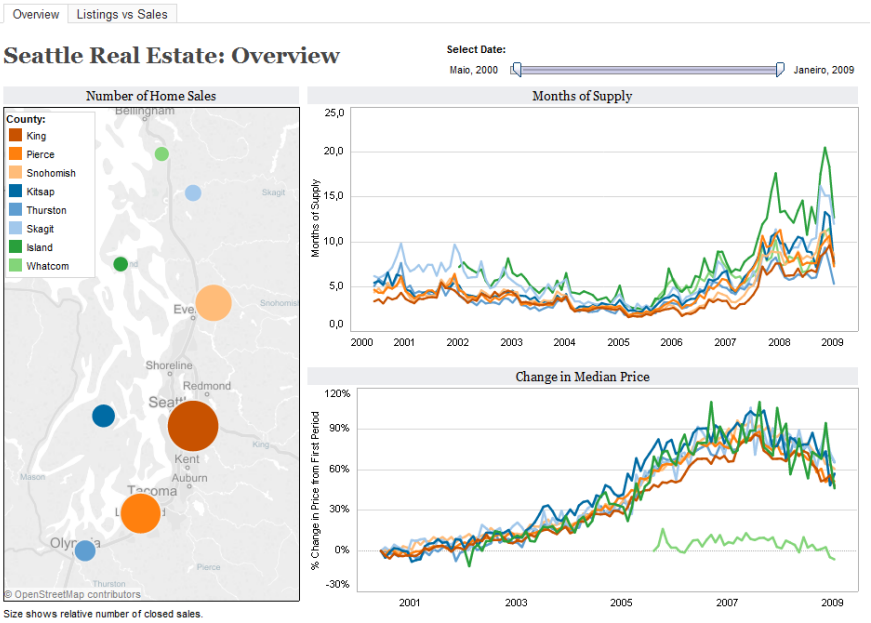

Figura 2.19 - Dashboard desenvolvido em Tableau[10]

Tableau[11] é um *software* de BI (*Business Itelligence*) fácil de utilizar, que permite aos utilizadores a criação de *Dashboards* interativos e compartilháveis na Web, a partir de Navegadores Web ou Tablet (mobile). Estes *Dashboards* são criados com o processo de "*Drag and Drop*" de visualizações, para apresentar graficamente os dados, ou através de tabelas.

O Tabelau contem três produtos para análise de informação (Rueter & Fields 2012), (Fields & Sheppard 2013):

- Tableau *Desktop*;
- Tableau *Server*;
- Tableau *Online*.

Tabelau *Desktop* é um *software* de descoberta e exploração de dados que permite, como indicado anteriormente, a construção de *Dashboards*, para a exploração de dados com diferentes visualizações e diferentes contextos, com a capacidade de combinar dados de múltiplas fontes. Os dados podem ser acedidos diretamente em Bases de Dados ou armazenados em memória, para aumentar a velocidade de acesso. Os *Dashboards* podem ser compartilhados facilmente na Web, podendo ser publicados no Tableau *Server*; para então serem acedidos pelos membros de uma instituição; ou no Tableau *Online*; para ser acedidos na *Web*.

Tableau *Server* permite que os utilizadores de uma empresa ou instituição tenham acesso aos *Dashboards* a partir de Navegadores Web.

[10] http://public.tableau.com/profile/technical.product.marketing#!/vizhome/Real-Estate-Seattle_24/Overview (Consultado em 20/10/2015)
[11] http://www.tableau.com/ (Consultado em 20/10/2015)

Tableau *Online* é um SaaS (*Software as a Service*), afastando o processo de configuração e instalação das mãos do utilizador e oferecendo a possibilidade de aceder aos dados através da Web.

O sítio oficial do Tableau disponibiliza o acesso a cursos, tutoriais, *Whitepapers* e *Webinars* para ensinar e ajudar os utilizadores a utilizar as suas ferramentas.

Google Charts

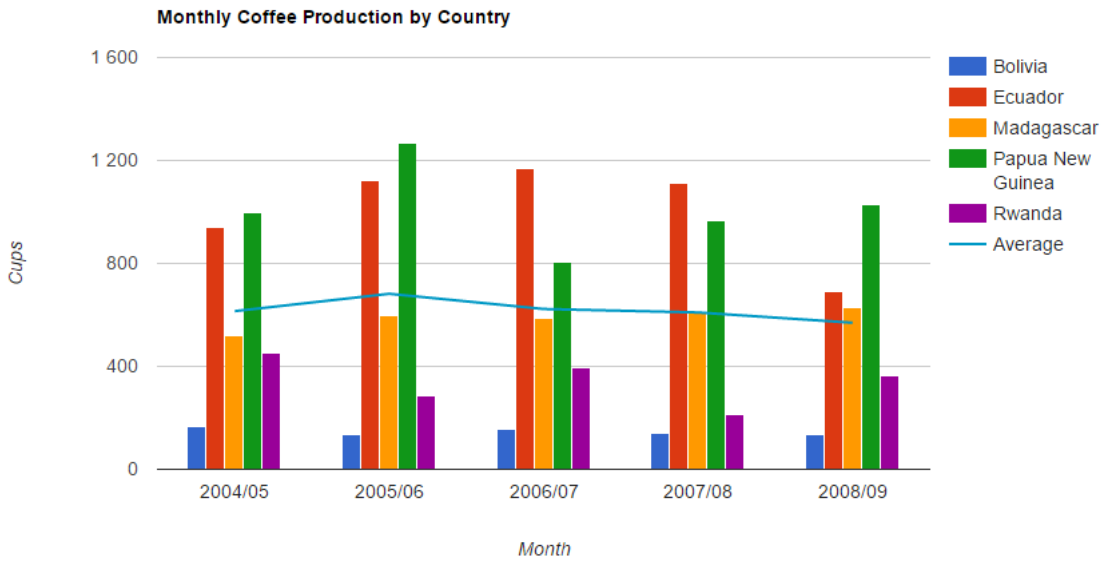

Figura 2.20 - Composição de Visualizações em Google Charts[12]

Google *Charts*[13] é uma junção de bibliotecas que possibilitam o mapeamento de informação e a apresentação de visualizações em Websites. Esta ferramenta possibilita a criação de visualizações simples, como gráficos de barras, e de visualizações mais complexas, como *tree maps* hierárquicas e composições de visualizações.

As visualizações são desenvolvidas com *JavaScript*, tratando os dados, que podem ser ordenados, filtrados e modificados; preparando as visualizações; definindo escalas, cores, título, etc. Finalmente as visualizações são apresentadas em HTML/SVG. É possível alterar a aparência dos gráficos, adicionar interações e eventos, possibilitando a criação de *DashBoards* com várias visualizações interligadas. Os dados a mapear podem ser acedidos diretamente pela *WebPage* ou em Bases de Dados SQL.

O Google *Developers* contem documentação extensiva em como utilizar esta ferramenta, com uma grande variedade de exemplos para ajudar na explicação e no desenvolvimento de visualizações.

[12] https://google-developers.appspot.com/chart/interactive/docs/gallery/combochart (Consultado em 20/10/2015)
[13] https://developers.google.com/chart/ (Consultado em 20/10/2015)

2.1.5. Testes de Usabilidade

É importante avaliar as visualizações desenvolvidas, determinando se estas necessitam de melhorias. Há uma variedade de formas de o efetuar, mas convém sempre utilizar uma amostra dos utilizadores alvo, para efetuar as avaliações e apontar problemas encontrados. Assim, os resultados obtidos serão próximos do esperado na utilização final.

Há quatro tipos de testes que se podem utilizar para avaliar a usabilidade:

- Avaliação heurística;
- Questionários informais;
- Estudos laboratoriais;
- *Crowdsourcing*.

As heurísticas são utilizadas, podendo ser apenas uma pessoa a efetuar a avaliação (mas com alguma experiência), para avaliar a técnica de visualização desenvolvida, encontrando problemas de usabilidade da interface, podendo também oferecer feedback em como as melhorar ou resolver os problemas.

Crowdsourcing e questionários tornam mais fácil o processo de recrutamento para a avaliação da visualização, uma vez que o *software* a analisar pode ser disponibilizado online, assim como o questionário, sendo o processo de avaliação efetuado remotamente por qualquer pessoa. Isto limita os utilizadores às opções apresentadas no questionário.

Estudos laboratoriais são formas de avaliação onde a recolha de dados é feita localmente, com uma seleção de avaliadores e um ou mais observadores, que seguem e auxiliam os avaliadores durante o processo. Isto dificulta o recrutamento de utilizadores, mas os resultados são mais concretos e não há tanta restrição nas respostas, já que os observadores podem tirar notas pessoais.

A grande diferença entre estas técnicas está no número e no tipo de resultados que são obtidos. Estudos laboratoriais estão mais limitados no número de utilizadores que recrutam em comparação com *Crowdsourcing*.

2.2. Visualização da Evolução de Rankings

A Classificação ou *Ranking* de uma entidade deriva da relação entre os seus atributos e os atributos de outras entidades. A forma como estes atributos estão definidos, ou a sua importância, pode alterar a classificação de uma entidade.

Nesta secção serão apresentados os artigos referenciados sobre técnicas de visualização para evolução de rankings.

2.2.1. Técnicas de Visualização Convencionais

Durante a pesquisa de métodos de representação de *rankings* ao longo do tempo, foi encontrado uma grande variedade de visualizações que recorriam a tabelas e gráficos de linhas.

			P	J	V	E	D	GM	GS	DG	
1	=	Benfica	71	28	23	2	3	71	15	+56	Jogos
2	=	FC Porto	68	28	21	5	2	66	12	+54	Jogos
3	=	Sporting	60	28	17	9	2	53	25	+28	Jogos
4	=	SC Braga	53	28	16	5	7	45	17	+28	Jogos
5	=	V. Guimarães	43	28	12	7	9	40	30	+10	Jogos
6	=	Belenenses	42	28	11	9	8	28	27	+1	Jogos
7	=	Paços Ferreira	38	28	10	8	10	32	37	-5	Jogos
8	=	Rio Ave	37	28	9	10	9	34	35	-1	Jogos
9	=	Nacional	36	28	10	6	12	36	41	-5	Jogos
10	=	Moreirense	35	28	9	8	11	25	32	-7	Jogos
11	=	Marítimo	35	28	10	5	13	35	37	-2	Jogos
12	↑	Estoril Praia	31	28	7	10	11	31	50	-19	Jogos
13	↓	Boavista	29	28	8	5	15	23	44	-21	Jogos
14	=	Académica	27	28	4	15	9	20	35	-15	Jogos
15	=	V. Setúbal	25	28	6	7	15	20	43	-23	Jogos
16	=	Arouca	23	28	6	5	17	20	42	-22	Jogos
17	=	Gil Vicente	19	28	3	10	15	22	48	-26	Jogos
18	=	Penafiel	18	28	4	6	18	27	58	-31	Jogos

Figura 2.21- Tabelas de Rankings[14]

As tabelas utilizadas (exemplo Fig.2.21) apresentam, de uma forma estruturada, os dados das entidades relativamente a um instante de tempo escolhido pelo utilizador. Tendo como exemplo os dados futebolísticos, as tabelas apresentam as classificações das equipas para uma determinada data (jornada), e caso se pretenda visualizar valores de jornadas diferentes, é necessário escolher a jornada em, por exemplo, menus *dropdown*, recarregando a página com os novos dados.

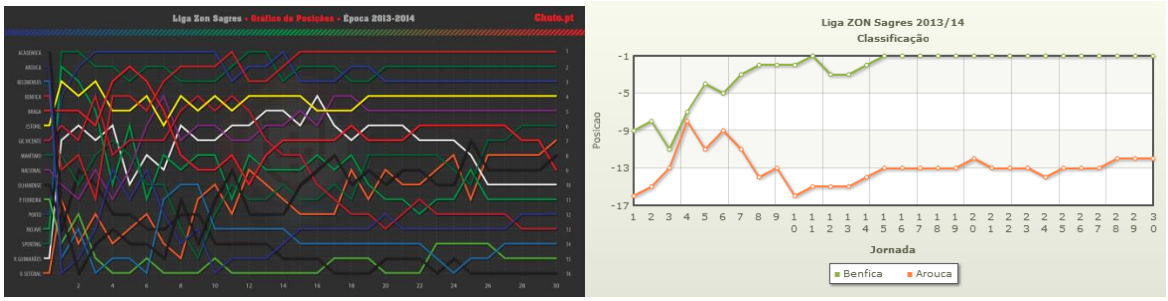

Figura 2.22 - Gráficos de Linhas[15]

Gráficos de linhas (exemplo Fig.2.22), são normalmente utilizados para apresentar a evolução de determinados atributos das entidades ao longo de um período de tempo. Tomando o mesmo exemplo anterior, estas visualizações apresentam a evolução das classificações das equipas, desde o início (jornada 1) até à última jornada. No caso do exemplo da figura à direita, o gráfico é construído com animações *flash* e limitado a duas equipas, não possui qualquer tipo de interação com as tabelas da página *Web*, sendo necessário especificar quais as equipas que se querem visualizar para recarregar a página e reconstruir o gráfico. Outros exemplos são constituídos por imagens fixas, que não oferecem qualquer tipo de feedback ou funcionalidades de interação com o utilizador.

[14] http://www.zerozero.pt/edition.php?jornada_in=28&id_edicao=70079&fase=70545 (Consultado em 20/10/2015)
[15] http://chuto.pt/infografia/ (Consultado em 20/10/2015)

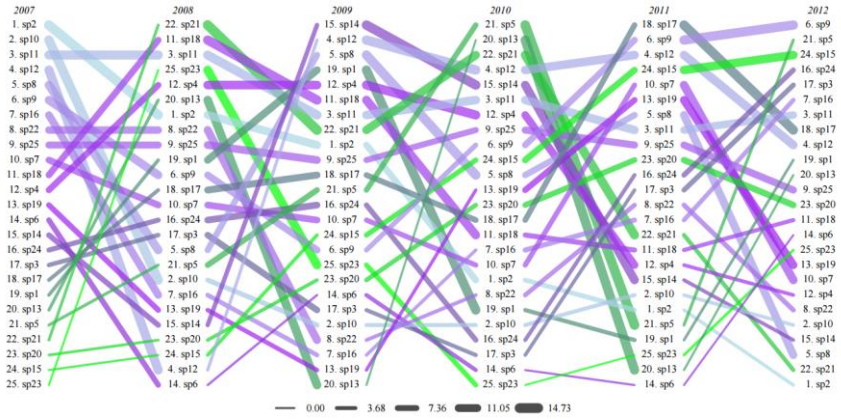

Figura 2.23 - Gráficos de Coordenadas Paralelas[16]

Em casos particulares, foram também encontrados exemplos que utilizam coordenadas paralelas (exemplo, Fig.2.23), onde é apresentada a evolução das características dos itens por data. Nestas visualizações, os eixos verticais representam as datas do conjunto de dados, e ao longo desses eixos, estão ordenados os elementos, as linhas interligam os elementos entre dois anos, indicando a sua evolução de um ano para o outro.

2.2.2. Representações Visuais inovadoras

Para o tema específico de visualização de *rankings*, foram encontrados apenas quatro referências, um número reduzido que pode indicar que ainda não houve muitos estudos sobre este tema. São apresentadas as técnicas desenvolvidas pelos autores dos artigos encontrados e a sua motivação.

RankExplorer: Visualization of Ranking Changes in Large Time Series Data

Este artigo, (Shi et al. 2012), aborda a melhor forma de apresentar visualmente uma grande quantidade de informação, e a sua evolução temporal. Defende técnicas de visualização temporal de classificações, com interações dinâmicas, de forma a facilitar a visualização dos dados com diferentes níveis de detalhe.

O seu objetivo principal é a compreensão de dados de séries temporais e a alteração na sua classificação ao longo do tempo, recorrendo para isso a técnicas de visualização que utilizam gráficos de linhas.

As visualizações convencionais existentes têm limitações em relação à apresentação de séries temporais. Nomeadamente, os gráficos em pilha (*Stacked Graphs*) não representam de forma intuitiva a alteração da classificação de itens ao longo do tempo. Gráficos de linhas são limitados, relativamente à quantidade de dados que podem apresentar sem ficarem sobrecarregados.

Foi optado pela utilização da representação *ThemeRiver*, que usa áreas para representar o valor quantitativo de um dado item num determinado instante de tempo. O *ThemeRiver* tem duas desvantagens: primeiro, pode ser difícil visualizar os dados se estes tiverem muitos elementos; segundo, cada área apenas representa a evolução do valor de um atributo por elemento.

16 https://beckmw.wordpress.com/2013/04/01/a-nifty-line-plot-to-visualize-multivariate-time-series/ (Consultado em 20/10/2015)

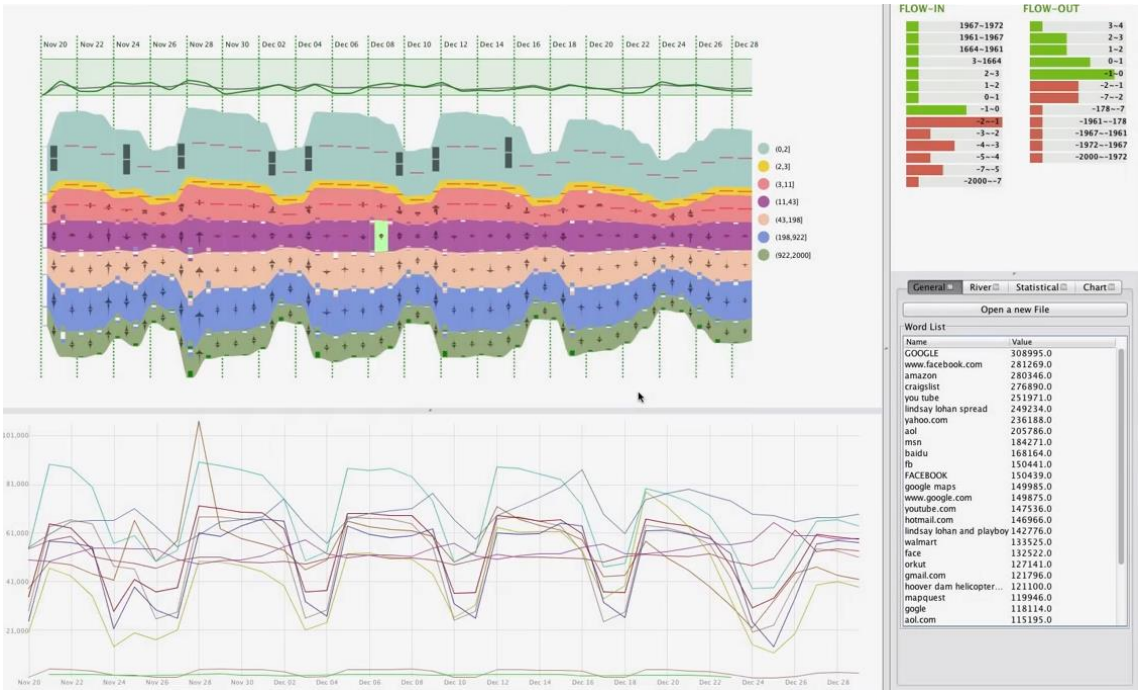

Figura 2.24 – RankExplorer visualization, como apresentado no tutorial[17]

Para evitar este problema, foi otimizada esta representação. As otimizações foram, para o primeiro problema, segmentar os dados, de forma a codificar as classificações por categorias (por exemplo, rankings 1-10). Isto fez com que não haja tanta "obstrução" visual na representação das áreas, mostrando a informação adicional dos elementos quando uma área é selecionada. Para o segundo, adicionaram barras de cores e glifos, para ajudar a identificar mudanças dos elementos entre categorias ao longo do tempo, por exemplo, quando um elemento sobe ou desce de categoria.

O RankExplorer estaá desenvolvido para que todas as visualizações estejam ligadas entre si, cada uma complementando a informação da anterior, à semelhança de um Dashboard:

- O **ThemeRiver** apresenta os rankings ao longo do tempo de uma forma geral, onde cada área representa um intervalo de classificações. Apresenta também as alterações que as categorias sofreram, com glifos, indicando a entrada e/ou saída de elementos;
- Os **Gráficos de Barras** apresentam valores individuais dos elementos que, em conjunto, fazem parte de um intervalo de rankings, formando a área de uma categoria no ThemeRiver.
- O **Gráfico de Linhas** desenvolvido apresenta a evolução individual do elemento, podendo efetuar comparações entre vários, quantos os possíveis selecionar. Dependendo da quantidade de elementos escolhidos, esta visualização pode ficar demasiado preenchida e ser difícil analisar os dados.

A técnica de visualização desenvolvida ajuda os utilizadores a avaliarem as evoluções de *rankings* dos itens com diferentes níveis de detalhe. Várias visualizações complementares é uma boa forma de manter todos os dados importantes acessíveis sem se sobrecarregar uma visualização.

[17] tutorial - https://www.youtube.com/watch?v=rdgn1qcZ2A4

LineUp: Visual Analysis of Multi-Attribute Rankings.

De acordo com (Gratzl et al. 2013), a motivação que levou ao desenvolvimento desta técnica foi a visualização e análise de classificações de entidades com vários atributos.

O objetivo é oferecer uma forma interativa de analisar classificações, com diferentes escalas e/ou semânticas. Estas novas técnicas possibilitam que os utilizadores combinem atributos e alterem parâmetros para explorar as alterações nas classificações.

Na análise de classificações é importante saber o porquê dessas classificações e, muitas vezes, pode-se pretender alterar os atributos para se concluir qual das entidades se classifica melhor de acordo com as expectativas dos utilizadores. As visualizações interativas fornecem formas de visualizar a classificação de entidades de acordo com os seus atributos, mas as visualizações de classificações atuais são estáticas, ou limitadas no que diz respeito à alteração dos atributos a aplicar na classificação.

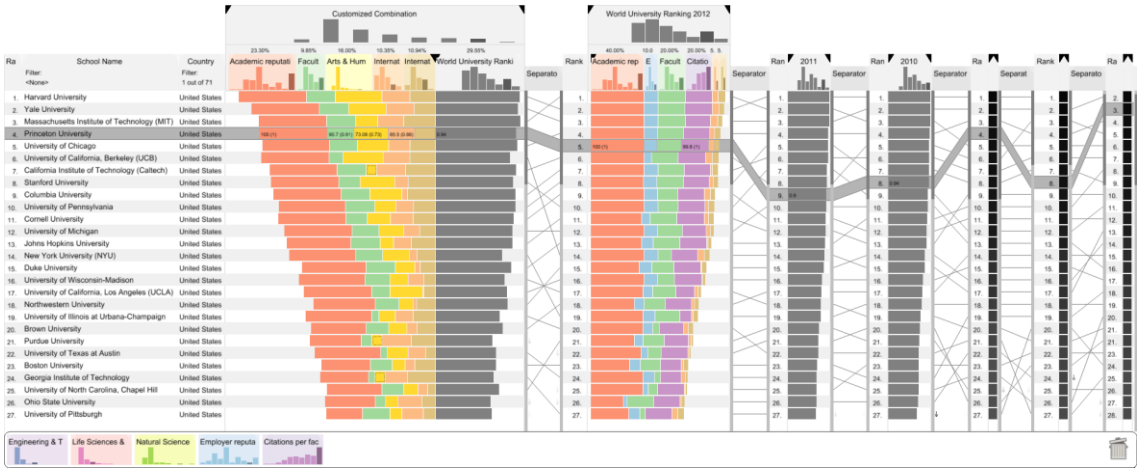

Figura 2.25 – LineUp, classificação das universidades de topo de acordo com "QS World University Ranking 2012", com atributos e pesos alterados, em comparação com os dados oficiais. Retirado de (Gratzl et al. 2013)

Para o desenvolvimento desta técnica de visualização, foi utilizada a *framework* de visualização *Caleydo*. Esta é uma ferramenta *open-source* implementada em Java, utilizando *OpenGL* para a representação das visualizações.

A técnica de visualização desenvolvida (*LineUp*) é uma visualização escalável que utiliza quase exclusivamente gráficos de barras. Com estes gráficos representa a classificação de entidades baseada em múltiplos atributos, com diferentes escalas e significados. *LineUp* oferece formas de combinar atributos e alterar a sua importância no cálculo das classificações de objetos. Com este processo, mostra a classificação de uma entidade de acordo com as espectativas de um utilizador, onde, por exemplo, se podem retirar os atributos que não se consideram importantes. *LineUp* pode também ser utilizado para comparar várias variações da classificação da mesma entidade, com combinações de atributos diferentes, com a utilização de *slope graphs*; por exemplo, a evolução da classificação de um objeto entre anos diferentes.

De uma forma geral, esta *framework* utiliza uma combinação de três tipos de representação de dados diferentes:

- **Tabela**, com a finalidade de representar os *rankings* de uma forma organizada e fácil de perceber.
- **Gráficos de Barras,** apresentam os valores quantitativos dos atributos. As barras ajudam na visualização da importância dos atributos e quais deles favorecem mais ou menos o *ranking* de cada entidade.
- **Slope graphs,** linhas que interligam as entidades entre variações de *rankings* diferentes, estas linhas servem para indicar ao utilizador como o *ranking* das entidades se altera ao modificar os seus atributos.

De acordo com os resultados relatados, está é uma técnica de visualização fácil de utilizar e percetível, que facilmente seria aplicável em vários cenários, que oferece várias formas de analisar as classificações de entidades e a evolução da classificação de uma entidade, quando os seus atributos são alterados. Possibilita assim a visualização da evolução temporal de *Rankings*, mesmo que este não seja o seu principal foco.

À Table! Improving Temporal Navigation in Soccer Ranking Tables

O artigo, elaborado por (Perin et al. 2014), trata a visualização de classificações de equipas de futebol, com suporte a avaliação temporal dinâmica. De acordo com os autores, o que motivou este trabalho foi a falta de escalabilidade, interação e comparação temporal de classificações em tabelas e gráficos para épocas de futebol.

O objetivo principal é encontrar formas dinâmicas de comparar as classificações de clubes de futebol ao longo da época, suportando funções temporais, para efetuar as comparações e a avaliação ao longo do tempo.

As tabelas atuais, que se encontram na *Web*, para a análise dos resultados das épocas de futebol, são "imagens" estáticas, gráficos ou tabelas onde não se podem aplicar filtros para reorganizar os dados. Aqueles que o permitem, muitas vezes não o fazem com suporte temporal, onde se possa ver a evolução da classificação de uma determinada equipa ao longo de toda a época. Nos casos que têm suporte a funções temporais, são disponibilizadas as datas em listas, onde se escolhe de forma predefinida a época a visualizar, apresentando imediatamente a resposta na visualização, sem *feedback* na alteração das classificações.

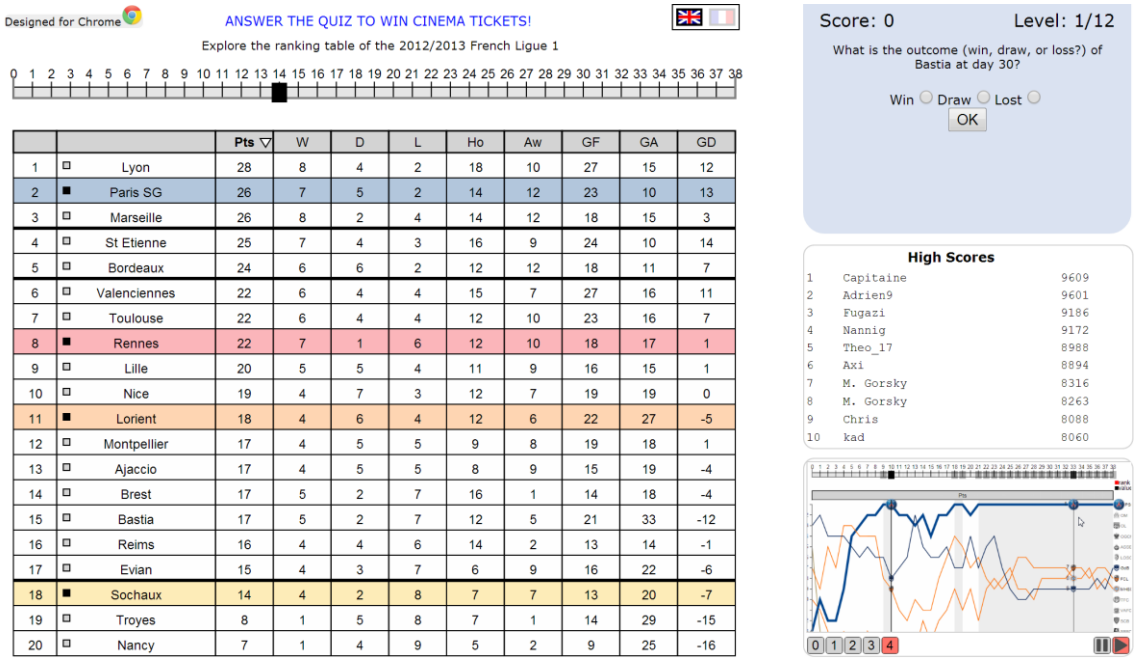

Figura 2.26 – Àtable Web Page, retirado de *(Perin et al. 2014)*

Para o desenvolvimento da sua visualização, os autores decidiram recorrer a *JavaScript* e D3.js. Desenvolveram uma tabela de dados, e duas formas inovadoras de representação e interação com os dados, uma barra temporal que nomeiam "*Temporal Slider*" e um gráfico de linhas, que nomeiam "*VIZ-RANK*":

- ***Temporal Slider*** oferece uma forma intuitiva de se alterar o instante de tempo dos dados. Ao alterar a jornada nesta barra, os dados na tabela são alterados, correspondendo à jornada aqui selecionada.
- **Tabela** é uma forma simples de apresentar *rankings* de equipas e os atributos que o geraram. Contem glifos que indicam se o ranking da equipa melhorou, piorou ou se manteve em relação à jornada anterior e feedback em relação à sua alteração quando se altera a jornada. Ao selecionar elementos e clicar na coluna das classificações, uma animação altera a tabela para o VIZ-RANK. Há também a possibilidade de arrastar um item na tabela para cima ou baixo, mostrando as alterações que este sofre, mas sem as efetuar até que o cursor do rato não seja mais pressionado.
- ***VIZ-RANK*** apresenta a evolução do ranking durante toda a época para os elementos selecionados na tabela. Aqui os elementos nunca se sobrepõem por não haver classificações iguais; em conjunto com um número de elementos reduzido, dificilmente esta se torna uma visualização confusa. Único senão é esta visualização não ser apresentada simultaneamente com a tabela.

Àtable foi um grande avanço em relação aos gráficos de classificações de futebol vulgarmente encontrados. Oferece várias técnicas para interagir com os elementos da tabela assim como suporte de funções temporais dinâmicas, onde se pode verificar a evolução das classificações ao longo do tempo.

Durante os testes com utilizadores, foi observado que as novas técnicas foram pouco utilizadas ou descobertas por engano. A grande dificuldade com que os autores se debatem é tornar as suas técnicas inovadoras mais óbvias para os utilizadores.

Investigating the Direct Manipulation of Ranking Tables

Este artigo, (Vuillemot & Perin 2015), introduz uma nova técnica, muito semelhante às técnicas introduzidas em Àtable, que permite a análise e manipulação de dados temporais (por exemplo, evolução de *Rankings*).

O objetivo no seu desenvolvimento foi ajudar os utilizadores a controlar a forma como os dados são atualizados nas tabelas e, ao mesmo tempo, ter uma visão global dos dados em relação ao que foi alterado. É explorada a usabilidade e o *design* das técnicas para tabelas de diferentes tamanhos, períodos de tempo e variedade de dados.

Esta técnica foi desenvolvida utilizando *JavaScript* e a biblioteca D3.js. Foram desenvolvidas representações para dados sobre citações académicas (autores com mais publicações no IEEE), evolução de rankings económicos de países ("*World Bank Indicators and an Economic Complexity Index (ECI)*") e campeonatos de futebol.

Para estes diferentes tipos de dados, foi necessário adaptar o gráfico de linhas, já que a evolução e informação a apresentar se alterava de acordo com a necessidade da visualização. Ou seja, nos dados dos campeonatos de futebol o gráfico de linhas apresenta a alteração do *ranking* de uma equipa ao longo das jornadas, assim como se o que levou a essa mudança de *ranking* foi uma vitória, derrota ou empate. Já não é o mesmo para as citações ou a economia de países, isto porque as alterações podem resultar de diferentes motivos, mas também em diferentes períodos de tempo.

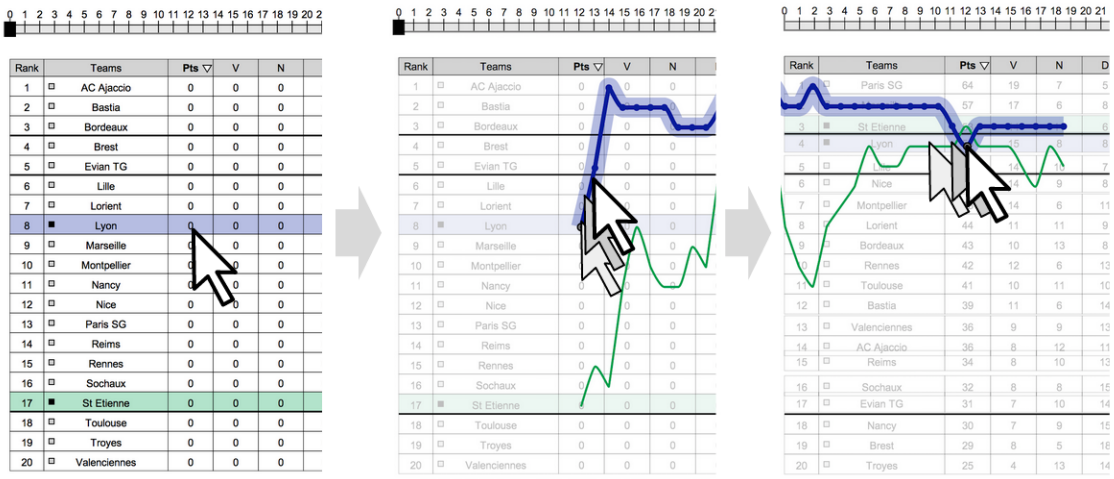

Figura 2.27- Exemplo de utilização de manipulação em tabelas de Rankings. Tabela de *Rankings* de equipas de futebol onde, ao arrastar o cursor, é apresentado um gráfico de linhas para manipular os dados da tabela em relação ao tempo, apresentando a evolução do *Ranking* para os elementos selecionados[18].

À semelhança da visualização Àtable, esta contem as mesmas visualizações e utiliza técnicas semelhantes. Contem uma *Temporal Slider*, uma tabela de dados, e um gráfico de linhas:

[18] http://romain.vuillemot.net/ manipulation-ranking-tables publications/chi15-direct- /

- ***Temporal Slider*** oferece uma forma de se alterar o instante de tempo. Ao alterar o valor do ano ou jornada nesta barra, os dados na tabela são alterados, correspondendo ao instante de tempo selecionado.
- **Tabela** apresenta *rankings* de elementos e os seus atributos. Permite ao utilizador alterar o instante de tempo dos dados: ao "arrastar" as células da tabela para a esquerda ou direita, é apresentado um gráfico de linhas, mostrando as alterações que os dados sofrem, mas sem as efetuar até que o cursor do rato não seja mais pressionado.
- **Line Chart** apresenta a evolução do *ranking* durante todo o período de tempo para os elementos selecionados na tabela. Este gráfico é apresentado enquanto se arrasta o cursor para a esquerda ou direita, alterando o instante de tempo dos dados enquanto se mantem uma visão geral da alteração do seu *ranking*. Quando se termina de arrastar o cursor, esta visualização desaparece e as alterações permanecem efetuadas na tabela. Aqui a interação tem que ser sobre a linha e não sobre a área de desenho.

A técnica introduzida não é óbvia para os utilizadores, e os autores tentaram chegar a uma forma de a anunciar ao utilizador, com um *popup* quando o cursor se encontra sobre uma célula da tabela.

Esta visualização, assim como Àtable, contém uma linha temporal, mas neste caso esta fica "desalinhada" com o gráfico de linhas (porque, como o utilizador tem que arrastar o cursor ao longo da linha, literalmente, é difícil manter a atenção na linha temporal e no gráfico de linhas em simultâneo), tentaram acrescentar o instante temporal ao gráfico, o que tornava a visualização muito confusa, devido ao excesso de informação. No final este não foi considerado um problema muito importante, porque não foi apontado pelos utilizadores, foi apenas uma suposição prévia dos autores à avaliação de usabilidade efetuada.

2.2.3. Requisitos Gerais

Durante o estudo de aplicações que apresentavam o *Ranking* de entidades (Gratzl et al. 2013) efetuaram um levantamento de requisitos que, embora se tenha iniciado como base na sua pesquisa para o seu *software*, pode ser generalizado e aplicado a diferentes domínios:

R I: Codificar a classificação: Utilizadores devem visualizar e compreender as classificações de elementos individuais, bem como empates nas classificações.

R II: Codificar atributos dos elementos, o que causou a classificação*:* Utilizadores devem poder avaliar os atributos dos elementos a partir do qual a classificação deriva e como estes se relacionam entre si.

R III: Suportar múltiplos atributos: Utilizadores devem poder combinar múltiplos atributos para produzir uma única classificação abrangente.

R IV: Suportar filtros: Utilizadores podem preferir excluir elementos ou alguns atributos do cálculo da classificação.

R V: Disponibilizar mapeamento flexível de valores de atributos para pontuações: Atributos podem ser de diferentes tipos e escalas. A visualização deve disponibilizar aos utilizadores uma forma de normalizar os valores dos atributos.

R VI: Adaptar escalabilidade para a visualização: Embora seja possível apresentar um grande número de elementos na visualização, é necessário um equilíbrio entre o nível de detalhe (LOD) e a escalabilidade.

R VII: Tratar valores em falta: Como muitas vezes os dados de fontes reais estão incompletos, uma visualização deve ser capaz de "reparar" ou recalcular os valores em falta nos ficheiros fonte.

R VIII: Interações e feedback visual: A visualização deve disponibilizar formas de o utilizador, dinamicamente, adicionar e remover atributos, modificar as suas combinações, e alterar a sua importância para o cálculo das classificações.

R IX: Otimizar atributos relacionados com a Classificação: Ao contrário de analisar como a classificação se altera de acordo com os atributos de um elemento, deve ser possível visualizar a melhor classificação possível de um elemento com a otimização dos seus atributos. Por exemplo, em vez de verificar para uma data qual o ranking de um dado elemento, verificar quando é que este elemento terá a sua melhor classificação, otimizando os seus valores ou a importância para o cálculo da classificação dos seus atributos.

R X: Comparar múltiplos Rankings: Em algumas situações, é importante relacionar várias classificações. Por exemplo, verificar se um elemento mantém a mesma classificação se trocarmos os seus atributos. Um bom exemplo seria verificar a evolução das classificações por ano.

2.3. Conclusões

De seguida apresenta-se um breve resumo das características principais das visualizações apresentadas nos artigos analisados, justificando as escolhas para o desenvolvimento do trabalho descrito nesta dissertação.

A tabela seguinte apresenta os pontos fortes das representações em relação a fatores essenciais para a visualização da Evolução de *Rankings*. Para esta tabela foi efetuada uma seleção de requisitos, relativamente aos requisitos gerais, que se enquadrassem mais com a evolução de *Rankings*, sem manipulação de valores de atributos e classificações.

Para a construção da tabela foi considerada a representação visual Àtable (2014) semelhante à analisada no artigo "Investigating the Direct Manipulation of Ranking Tables" (2015), visto que ambas as representações foram desenvolvidas pelos mesmos autores, contêm as mesmas visualizações e utilizam técnicas de interação semelhantes, que alteram pouco a forma como os dados são apresentados.

Tabela 2.1 - Comparação de características

Características	*RankExplorer (2012)*	*LineUp (2013)*	*Àtable (2014/2015)*
Mapear Rankings	+	++	++
Mapear atributos dos elementos (o que causou a classificação)	++	++	+
Filtros	++	++	+/-
Interações e feedback	++	++	++
Evolução Temporal de Rankings	++	+/-	++
Adaptabilidade a outros tipos de dados	++	++	-

Para "Filtros", foi considerado que a Linha temporal de Àtable é uma forma simples de filtrar os dados que são apresentados, porque, de acordo com o valor selecionado, apenas os dados correspondentes são exibidos.

Para "Evolução Temporal de *Rankings"*, LineUp suporta a construção de visualizações para a evolução e comparação de *Rankings*, mas, não sendo este o seu principal objetivo, a única forma de representar esta evolução é com *Slope Graphs* que interligam os diferentes *Rankings*.

Para "Adaptabilidade a diferentes tipos de dados", a representação desenvolvida para Àtable é muito limitada, sendo necessário a adaptação das representações para diferentes estruturas de dados, assim como no artigo "Investigating the Direct Manipulation of Ranking Tables", onde foi necessário adaptar o gráfico de linhas para a avaliação de Rankings de futebol.

Em relação a ferramentas para desenvolvimento, foi optado neste trabalho pela utilização da Biblioteca D3.js. É uma biblioteca relativamente recente com compatibilidade entre plataformas. Em comparação com as ferramentas analisadas, esta oferece métodos e funções mais abrangentes e diretamente mais relacionados com a manipulação dos dados que o Tableau; e oferece maior variedade e liberdade na criação e manipulação das visualizações e funções interativas que o Google Charts.

3. Um Modelo para a Visualização de Rankings

Neste capítulo são apresentadas as opções tomadas para o desenvolvimento do trabalho desta dissertação, com base no estudo efetuado no capítulo 2. O capítulo está dividido em quatro secções chave: Requisitos para Análise de Rankings, Dados Utilizados, Tecnologia Selecionada e Arquitetura Desenvolvida.

3.1. Requisitos para análise de Rankings

Com base nos Requisitos gerais apontados no capítulo 2, foi estabelecido um conjunto de requisitos considerados essenciais para a representação e análise de *Rankings*. Para esta lista não se consideraram requisitos que indiquem a manipulação de valores de atributos e classificações, visto este não ser um objetivo do trabalho.

A lista de Requisitos de Usabilidade estabelecida foi:

R I: Codificação de Rankings;
R II: Codificação de Atributos;
R III: Suporte de Múltiplos Atributos;
R IV: Suporte de Filtros;
R V: Adaptar Escalabilidade para Representações;
R VI: Tratar Valores em falta;
R VII: Interações;
R VIII: Feedback Visual;
R X: Comparar múltiplos Rankings.

Esta lista será utilizada para efetuar uma auto avaliação ao *Dashboard* desenvolvido e às suas representações, indicando se satisfazem os requisitos aqui indicados.

3.2. Casos de Estudo

Foram desenvolvidos dois casos de estudo diferentes, cada um utilizando dados com estruturas distintas, testando a hipótese de se poderem utilizar as mesmas visualizações, sem alterações adicionais, para a apresentação de dados diferentes. Os dados selecionados para o desenvolvimento das representações visuais foram:

- *Rankings* de épocas de futebol;
- *Rankings* de escolas secundárias em relação às médias de notas nos exames.

Estes dados encontram-se armazenados em ficheiros de formato CSV (Comma-separated values), sendo o primeiro caso de estudo os *Rankings* de Futebol. Os *Rankings* de Escolas Secundárias surgem mais tarde com o objetivo de comparar as visualizações e a sua possível reutilização. Os dados relativos às Médias Escolares e Disciplinas completam a informação relativa aos *Rankings* Escolares, detalhando o modo como as escolas obtiveram as suas classificações.

Rankings de Futebol

O primeiro caso de estudo, *Rankings* de Futebol, com os dados retirados do *site* zerozero.pt[19], permite a utilização de representações visuais para apresentar a evolução da classificação das equipas da Primeira Liga Portuguesa, ao longo da época de 2012/2013 e 2013/2014. Estes dados seguem a seguinte estrutura, para cada equipa:

- Ranking
- Nome
- Pontos
- Jogos
- Vitórias
- Empates
- Derrotas
- Golos Marcados
- Golos Sofridos.

Pode ser visualizado um exemplo dos dados em um dos ficheiros utilizados para este caso no Apêndice A.1.

Ficheiros com esta estruturação contêm, normalmente, um baixo número de elementos (equipas) e um baixo número de atributos. Caso se queira aprofundar os atributos, o seu número cresce para aproximadamente o dobro, diferenciando os atributos por jogos efetuados no campo da equipa ou no campo da equipa adversária.

Com esta análise dos dados, podemos concluir que, estas representações podem ser utilizadas para apresentar dados com estruturas semelhantes, por exemplo, dados de outro desporto ou outra liga.

Rankings de Escolas Secundárias

O segundo caso de estudo, *Rankings* de Escolas Secundárias, surgiu da necessidade de testar as visualizações utilizadas anteriormente, para apresentar dados estruturados de forma distinta. Para este caso de estudo, foram utilizados os dados disponíveis na plataforma Web BESP (Benchmarking de Escolas Secundárias Portuguesas) desenvolvido pela Faculdade de Economia e Gestão da Universidade Católica, no Porto[20]. Neste caso as visualizações têm que apresentar a evolução nos *Rankings* de Escolas Secundárias ao longo de

[19] http://www.zerozero.pt/edition.php?id_edicao=70079 (Consultado em 20/10/2015)
[20] http://besp.feg.porto.ucp.pt/pagina.php (Consultado em 20/10/2012)

anos diferentes, isto tendo em conta as médias obtidas em exames nacionais. Os dados relativos aos *Rankings* das escolas contêm os seguintes atributos, para cada Estabelecimento de Ensino:

- Ranking
- Nome
- Localidade
- Ano Letivo
- Média de Exames
- Número de Exames efetuados
- Média CIF (Classificação Interna Final)
- Média CFD (Classificação Final Distribuída).
- Média CIF - Média Exame (Diferença entre média exames e média CIF)
- Média Relativa

Como exemplo, pode ser visualizado um excerto de um ficheiro utilizado para este caso no Apêndice A.2.

Também para este caso de estudo, foram selecionados dados que representam a evolução das médias de exames por Concelho, para comparar a evolução das médias também entre Concelhos e Distritos.

Para este caso, foram utilizados dados de fontes distintas:

- Ficheiro de dados com informação sobre as escolas do País e as Médias dos exames por disciplina
- Ficheiro de dados para associar os Concelhos ao respetivo Distrito.
- Ficheiros Shapefile com informação geo-espacial relativa a Portugal Continental e Ilhas

O ficheiro *Shapefile*, retirado do sítio da Direção-Geral do Território[21], com todos os Concelhos de Portugal, foi trabalhado de forma a condensar a informação, para manter apenas a necessária, por exemplo, a informação geo-espacial dos Distritos e o nome, facilitando a futura visualização do mapa em ambiente Web. O ficheiro para as médias de exames contem a informação relativa à escola:

- Nome da Escola
- Distrito
- Média das notas por disciplina.

Foram utilizadas as notas das disciplinas de Português, Biologia/Geologia, História A, Economia A, Física e Química A, Geografia A e Matemática A para calcular as médias por Distrito e por Concelho.

Ficheiros com estas estruturas de dados contêm um grande número de elementos (escolas do país) e um número reduzido de atributos por elemento, dependendo dos dados que se pretende apresentar. Para apresentar estes dados, foi necessário otimizar a representação, para suportar a grande diferença no número de elementos e diferentes processos de filtragem.

[21] http://www.dgterritorio.pt/cartografia_e_geodesia/cartografia/ (Consultado em 20/10/2012)

3.3. Tecnologia selecionada

Para o desenvolvimento das visualizações, como já indicado, foi optado pela utilização da D3.js, pois possibilita a criação de *frameworks* de visualizações, sem limitações técnicas em relação às visualizações e estrutura dos dados utilizados (CSV/TSV/JSON).

Sendo os dados mapeados associados a elementos DOM, e como a D3.js pode manipular diretamente estes elementos, há uma facilidade na manipulação posterior das visualizações, podendo alterar escalas e até o mapeamento de novos dados para visualizações já existentes. A biblioteca contem funções e métodos padrão para adicionar animações quando existem alterações nos dados, e é relativamente fácil programar interações entre as visualizações e com o utilizador.

Como foi indicado, a dificuldade de utilização e desenvolvimento com esta biblioteca depende de conhecimentos prévios. Mas existe uma grande variedade de conhecimento partilhado *online*, relativo à utilização de D3.js. Como exemplo pode ser utilizado o sítio oficial da D3.js, com uma lista extensiva de exemplos de utilização e um tutorial desenvolvido pelo próprio autor da biblioteca. Também há uma grande variedade de tutoriais *online*, com o intuito de reduzir o tempo e a curva de aprendizagem.

Tecnologias Complementares

Como, para o segundo caso de Estudo (*Rankings* de Escolas Secundárias), foi decidido utilizar dados geográficos para apresentar informação adicional ao utilizador, houve a necessidade de utilizar uma outra forma de estruturação de dados, o TopoJSON[22], para além de CSV.

Este formato, desenvolvido pelo mesmo autor que a D3.js, foi desenvolvido como uma extensão de GeoJSON[23], um formato que codifica e armazena estruturas de dados geográficos, como por exemplo, pontos e polígonos. A grande diferença entre estes dois formatos é a forma como os dados estão estruturados. No TopoJSON, os dados são codificados de forma a otimizar a sua dimensão e facilitar a sua manipulação, removendo redundância de informação. O TopoJSON, ao contrário de armazenar a estrutura geométrica de um objeto, armazena referências por objeto, que apontam para um índice de sequências de pontos, onde as coordenadas estão armazenadas, denominados "*arcs*", para que alguns objetos compartilhem sequências de pontos entre si, por exemplo, fronteiras.

Os ficheiros *Shapefile* utilizados, com dados geo-espaciais de Portugal, foram pré-processados em QuantumGis, para retirar informação desnecessária para o tratamento dos dados na Web. Os ficheiros gerados foram então convertidos para GeoJSON com a ferramenta ogr2ogr e, finalmente, foram todos eles convertidos para um único ficheiro TopoJSON com toda a informação geo-espacial de Portugal Continental e Ilhas.

[22] https://github.com/mbostock/topojson (Consultado em 20/10/2012)
[23] http://geojson.org/ (Consultado em 20/10/2012)

3.4. Arquitetura desenvolvida

Nesta secção são apresentadas as opções tomadas para o desenvolvimento do *Dashboard*, tendo sido considerados os requisitos referenciados no início deste capítulo, e efetuada uma análise para avaliar se estes eram suportados pelas representações desenvolvidas.

São também descritos os métodos de interação que foram desenvolvidos e como estes afetam as representações.

3.4.1. Camada de dados

Os dados são lidos e alterados para um formato compreendido pelo D3.js, para poderem ser mapeados e apresentados. Neste caso, os dados são formatados para um *array* de objetos, que é tratado de acordo com as necessidades de cada visualização, formando uma nova estrutura de dados sempre que necessário.

Durante o mapeamento da informação para a construção de visualizações, é feito um *parse*, certificando que os dados não contêm caracteres especiais ou desconhecidos, para apresentar dados corretos e compreensíveis pelo utilizador e também facilitar a interação entre visualizações.

3.4.2. DashBoard e Fluxo de Dados

Figura 3.1 - Exemplo do Dashboard desenvolvido

É através da exploração e análise das visualizações que se podem identificar padrões nos dados. Por isso, foi optado integrar todas as representações num *Dasboard*, para manter as visualizações no mesmo local e interligadas entre si.

A ideia chave aplicada para a interação com as visualizações é permitir aos utilizadores a manipulação do instante temporal dos dados nas representações e na tabela, para analisarem a evolução de *rankings*. A tabela e o mapa são as representações que disponibilizam as principais formas de interação, seleção e filtragem de elementos, mas a maioria das representações apresentam *feedback* e interatividade mútua.

De início, o *Dashboard* apresenta uma página com a finalidade de introduzir os diferentes casos de estudo, *Rankings* de futebol ou *Rankings* escolares e as Médias Escolares.

No caso dos *Rankings* de Futebol, toda a informação é apresentada na mesma página para facilitar a sua visualização. Nos casos em que as visualizações não sejam visíveis em simultâneo, há formas de apresentar a informação necessária para o entendimento dos dados a analisar.

Para o caso dos *Rankings* Escolares foi seguido outro método: como a quantidade de dados é muito superior, estes foram separados por temas. Aqui, para não sobrecarregar a página e as visualizações, foram criadas três páginas diferentes. A página principal apresenta a informação relativa aos *Rankings* das Escolas e a sua evolução ao longo dos anos, a segunda página apresenta a informação relativa às Médias das Escolas por Distrito e, finalmente, a terceira página apresenta as Médias das Escolas por Disciplina.

3.4.3. Rankings de Futebol

Para o *Ranking* de Futebol pode ser escolhida a que época correspondem os conjuntos de dados a apresentar, caso contrário é apresentado o último ano registado.

Este *Dashboard* contem um total de quatro representações, sem contar com *slider* temporal:

- Tabela;
- Gráfico de Coordenadas Paralelas;
- Gráfico de Linhas para a evolução de *Rankings* ou de Pontos;
- Gráficos de Barras para diferença de golos e com resultados de Jogos.

Figura 3.2 - *Slider* Temporal para Jornadas de Futebol

Os dados a apresentar nas visualizações contêm valores que se alteram ao longo do tempo, desta forma foi optado, à semelhança da visualização de (Perin et al. 2014), pela utilização de um elemento que consiste em uma barra horizontal com um objeto SVG que pode ser arrastado para alterar a jornada dos dados a apresentar. Por omissão, os dados apresentados são sempre os mais atuais (última jornada registada).

A tabela apresenta os dados em formato textual, onde são apresentados os atributos das equipas, para uma dada jornada.

Ranking	Equipa	Pontos	Vitorias	Empates	Derrotas	Golos Marcados	Golos Sofridos
1	Benfica	74	23	5	2	58	18
2	Sporting	67	20	7	3	54	20
3	FC Porto	61	19	4	7	57	25
4	Estoril Praia	54	15	9	6	42	26
5	Nacional	45	11	12	7	43	33
6	Marítimo	41	11	8	11	40	44
7	V. Setúbal	39	10	9	11	41	41
8	Académica	37	9	10	11	25	35
9	SC Braga	37	10	7	13	39	37
10	V. Guimarães	35	10	5	15	30	35
11	Rio Ave	32	8	8	14	21	35
12	Arouca	31	8	7	15	28	42
13	Gil Vicente	31	8	7	15	23	37
14	Belenenses	28	6	10	14	19	33
15	P. Ferreira	24	6	6	18	28	59
16	Olhanense	24	6	6	18	21	49

Figura 3.3 - Tabela de Rankings de Futebol

Como forma de interagir com os dados, a tabela possibilita a seleção de elementos para os destacar nas restantes visualizações, ou para desenhar novas representações para comparar informação, por exemplo, o gráfico de barras empilhadas com o desempenho das equipas.

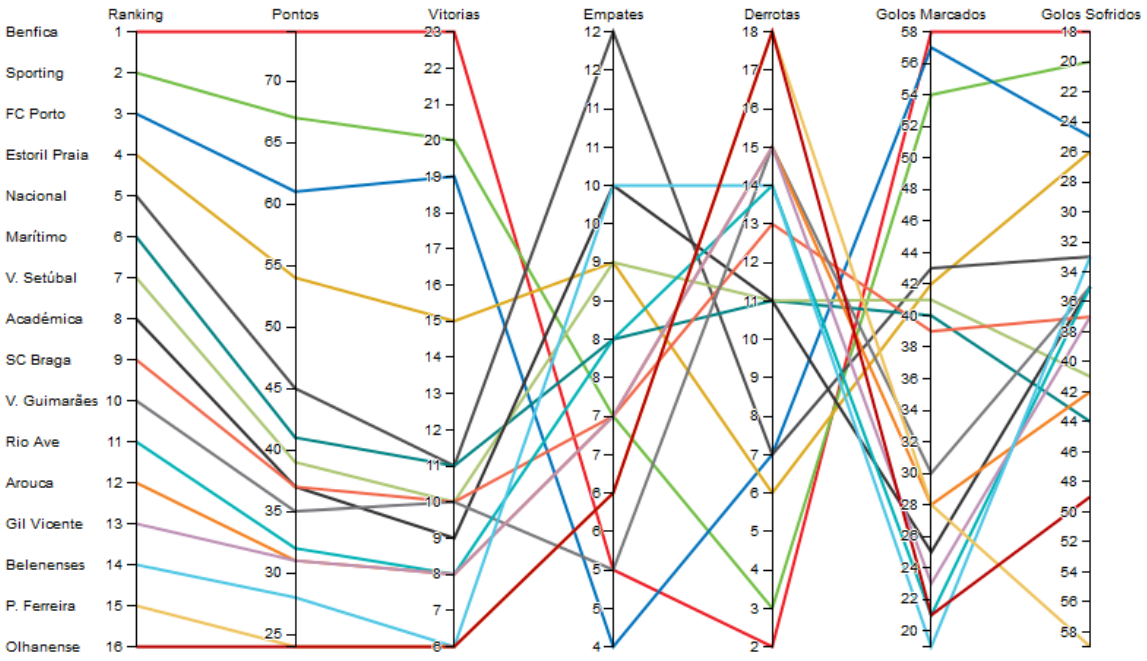

Figura 3.4 - Gráfico de Coordenadas Paralelas para Rankings de Futebol

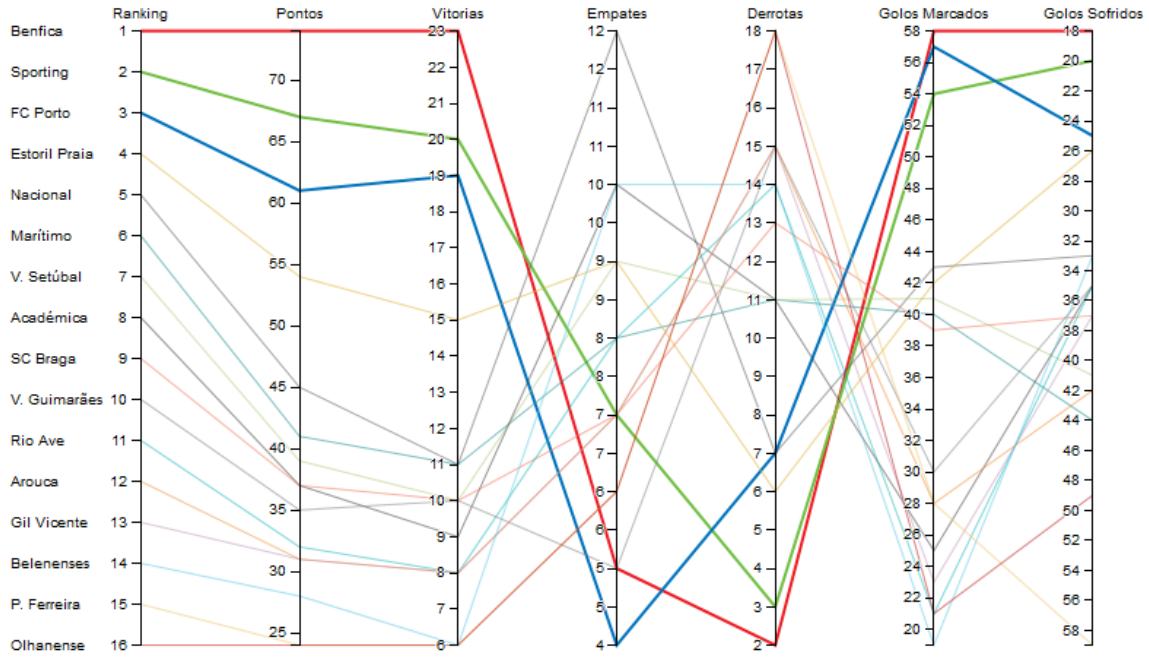

Figura 3.5 - Gráfico de Coordenadas paralelas com seleção de Equipas

Para complementar a informação apresentada na tabela, é utilizado um gráfico de coordenadas paralelas, apresentando, mais eficazmente, tendências entre os elementos, por exemplo, se a equipa com mais pontos, tem também mais golos marcados. Este gráfico destaca os elementos que estão selecionados na tabela e apresenta *tooltips* para facilitar a visualização dos valores de atributos. Para este gráfico, houve uma reordenação nas escalas, onde o eixo que representa o valor dos golos sofridos foi invertido, impedindo a ocorrência da sobreposição de linhas e facilitando a análise.

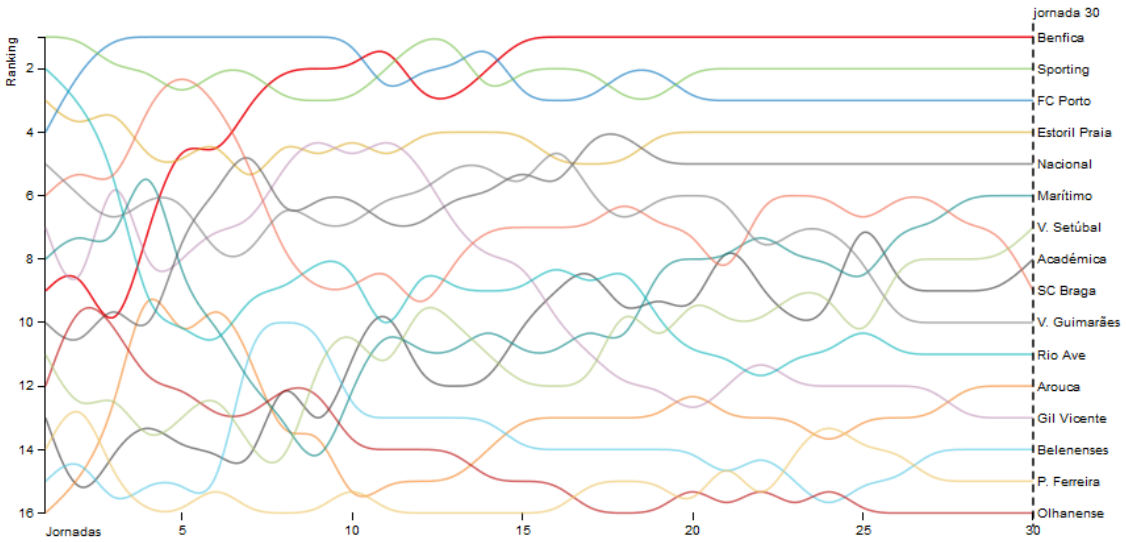

Figura 3.6 - Gráfico de Linhas que apresenta evolução de classificações ao longo da época

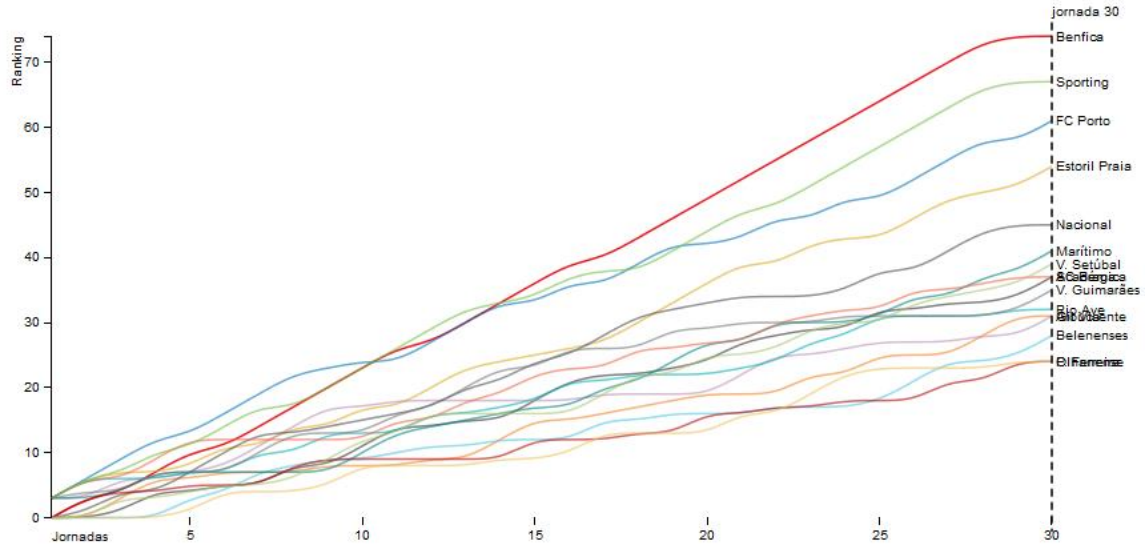

Figura 3.7 - Gráfico de Linhas para a Evolução dos Pontos com todas as Equipas

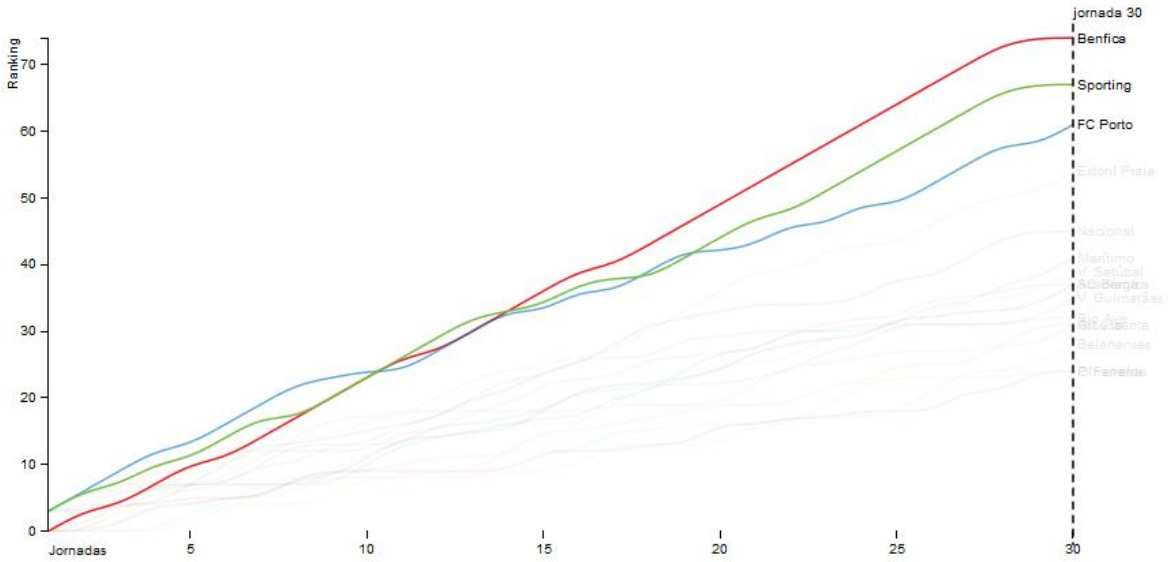

Figura 3.8 - Gráfico de Linhas para a evolução de pontos com seleção de equipas

Para facilitar a visualização de tendências diretamente relacionadas com a evolução dos dados ao longo de toda a época, foi optado pela utilização de um gráfico de linhas, que apresenta, de acordo com o *input* do utilizador, a evolução no *Ranking* ou dos Pontos das equipas. Este gráfico contém uma linha vertical que indica a jornada a ser visualizada, facilitando o reconhecimento dos dados e destaca os elementos que estão selecionados na tabela.

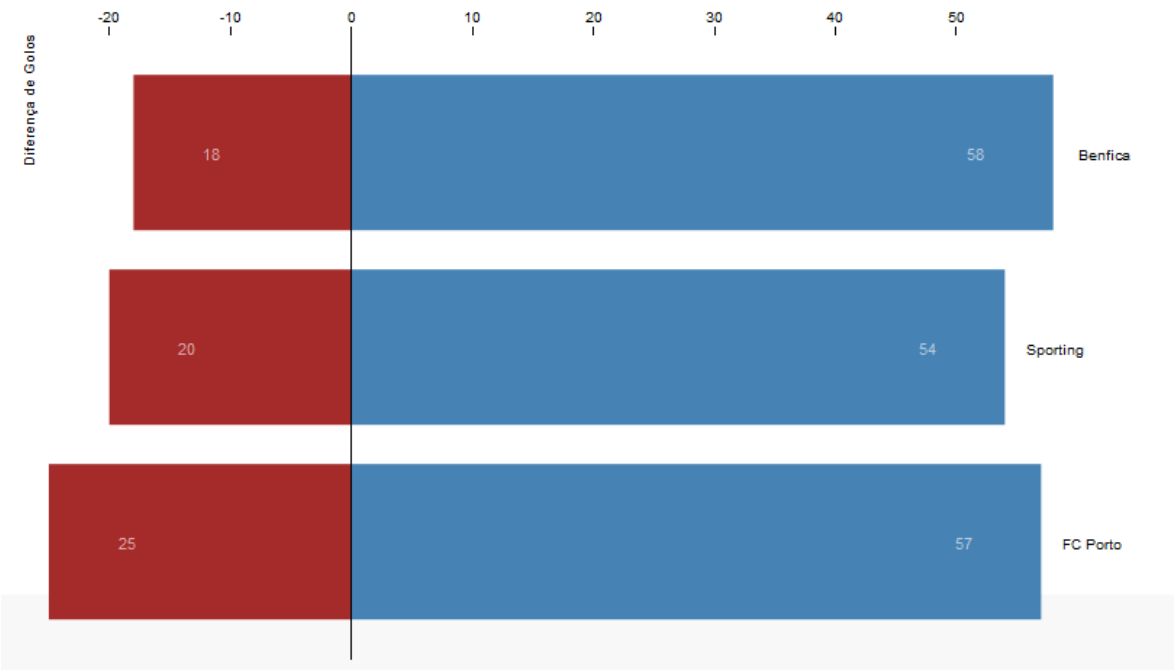

Figura 3.9 - Gráfico de Barras que mostra a Diferença de Golos por equipa

Ao selecionar uma ou mais equipas, é apresentado um gráfico de barras horizontais. Estas barras contêm valores negativos e positivos para representar os golos sofridos e os golos marcados, respetivamente, para a jornada selecionada. Ou seja, contém um eixo vertical (x) que representa o valor 0 (zero), e todos os valores à esquerda (negativos) representados pela cor Vermelha, indicam os golos sofridos pela equipa, e os valores à direita (positivos) representados pela cor Azul, indicam os golos marcados. Foi escolhida esta orientação porque facilita a análise destes dados.

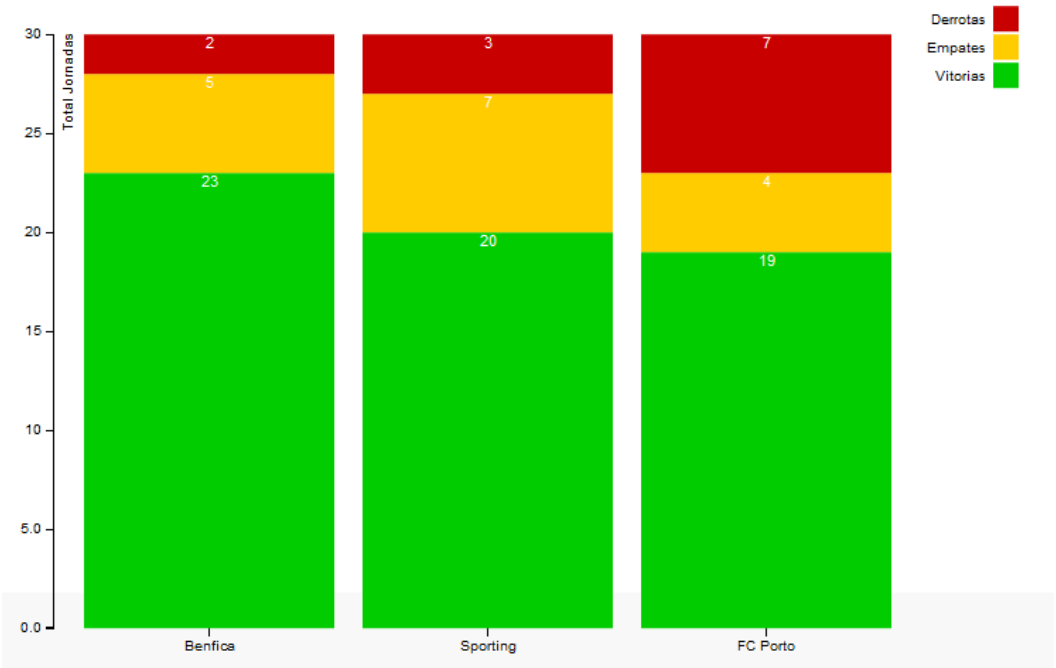

Figura 3.10 - Gráfico que mostra a performance das equipas até a jornada atual

Em conjunção com o gráfico anterior, é apresentado um gráfico que apresenta a performance das equipas em relação aos seus jogos até à jornada selecionada. Esta informação está apresentada num gráfico de barras empilhadas, onde o eixo vertical representa o número de jogos das equipas até à data, e as cores verde, amarelo e vermelho representam as vitórias, empates e derrotas, respetivamente. Com esta visualização pode-se analisar a performance da equipa, verificando qual tem mais vitorias ou derrotas.

3.4.4. Rankings de Escolas

Para a análise da evolução de *Rankings* escolares, foram adaptadas as visualizações anteriores, com a exceção dos gráficos de barras.

Com estas alterações, os dados podem ser analisados recorrendo às seguintes visualizações:

- Tabela;
- Gráfico de Coordenadas Paralelas;
- Gráfico de Linhas para evolução de *Rankings* ou de Médias.

A tabela possibilita a análise dos Rankings das Escolas e seus atributos, para um determinado ano letivo. À semelhança da Tabela de *Rankings* de Futebol, é possível selecionar estabelecimentos de ensino para os filtrar nas restantes visualizações. Mas, como estes dados contêm uma grande quantidade de elementos, houve a necessidade de integrar mecanismos de filtragem para a própria Tabela, para além da variável "Tempo".

Ranking	Estabelecimento	Localidade	Media Exame	NProvas	Media CIF	Media CFD	ExameCIF	Media relativa
1	Colégio Nossa Senhora do Rosário	Porto	14.36	460	16.71	16.09	-2.35	1.4
2	Colégio Moderno	Lisboa	14.07	301	15.42	15.06	-1.34	1.37
3	Colégio Valsassina	Lisboa	14.03	174	14.2	14.16	-0.17	1.38
4	Externato Horizonte	Porto	13.99	10	17.5	16.7	-3.51	1.38
5	Escola Técnica e Liceal Salesiana Santo António	Estoril (Cascais)	13.98	366	14.49	14.43	-0.51	1.38
6	Colégio Oficinas de São José	Lisboa	13.95	484	15.05	14.76	-1.09	1.37
7	Colégio Mira Rio	Lisboa	13.88	59	15.81	15.32	-1.94	1.33

Figura 3.11 – Tabela de Rankings Escolares

Além disso, a tabela foi desenvolvida com suporte de *scrollbars*, para manter toda a informação acessível, sem tornar a sua leitura impossível ou a página demasiado grande. É importante notar que, para este suporte ser eficaz, é necessário separar o *Hearder* e o *Body* da Tabela em elementos diferentes na formatação em CSS, isto faz com que haja mais trabalho para manter as dimensões corretas e os títulos do Cabeçalho alinhados com as colunas do Corpo da Tabela. Além deste suporte, houve a necessidade de desenvolver uma função de pesquisa, para pesquisar e filtrar a informação da Tabela de acordo com o *input* do utilizador, podendo este ser efetuado em relação ao nome das Escolas ou Concelho.

Com os gráficos de Coordenadas Paralelas e Linhas, houve a mesma preocupação. Como os dados são muito extensos, é impossível apresentar toda a informação e em simultâneo manter a informação limpa e percetível. Por este motivo, foi optado limitar a informação apresentada às 5, 10 e 100 escolas do topo da classificação, de acordo com o *input* do utilizador.

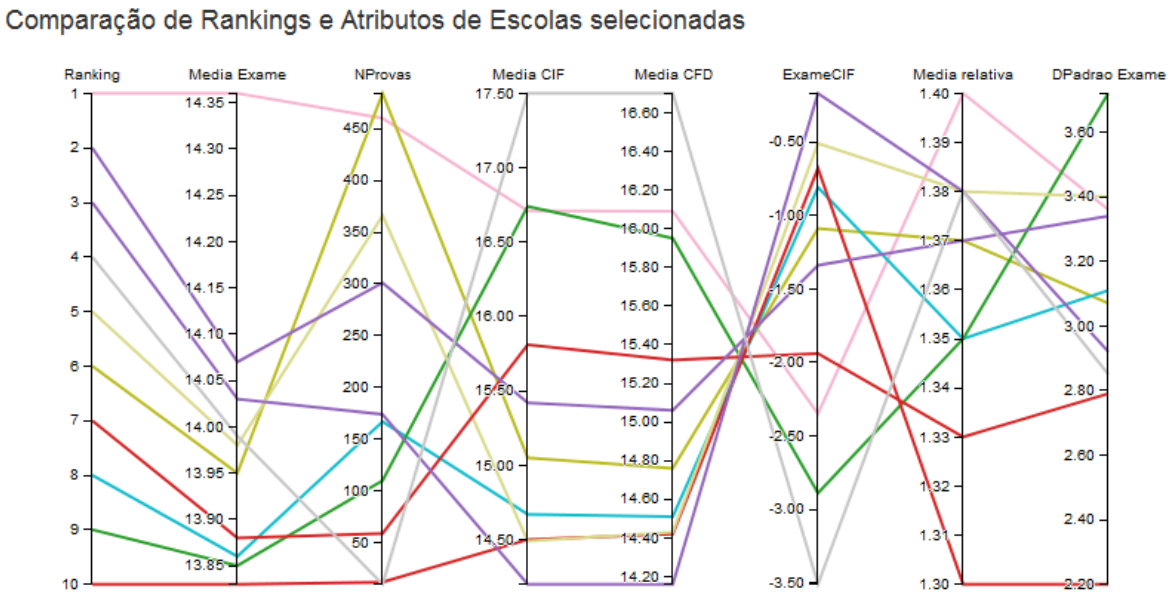

Figura 3.12 – Coordenadas Paralelas, as 10 melhores escolas

Assim limita-se bastante o emaranhado de linhas e facilita-se a visualização dos dados. Este método serve apenas caso o utilizador não tenha selecionada nenhuma linha da tabela, caso isto aconteça, o gráfico apresenta quantas escolas forem selecionadas, à exceção de ser apenas uma linha. Isto acontece por não haver a necessidade de redesenhar o gráfico apenas com dados de uma Escola, visto não ser necessária a comparação de dados.

Para além das visualizações apresentadas, há a possibilidade de aceder ao *Dashboard* para análise de Evolução de Médias Escolares, apresentando informações relativas às Médias das Escolas por Distrito e Disciplinas.

Evolução Médias Escolares

Para a Evolução de Médias Escolares, foram criadas novas visualizações, para melhor suportar os dados a apresentar. Para estes dados, foi desenhada uma tabela e um mapa: ao selecionar-se um Distrito no mapa são filtrados os dados da tabela, possibilitando a sua análise com mais detalhe.

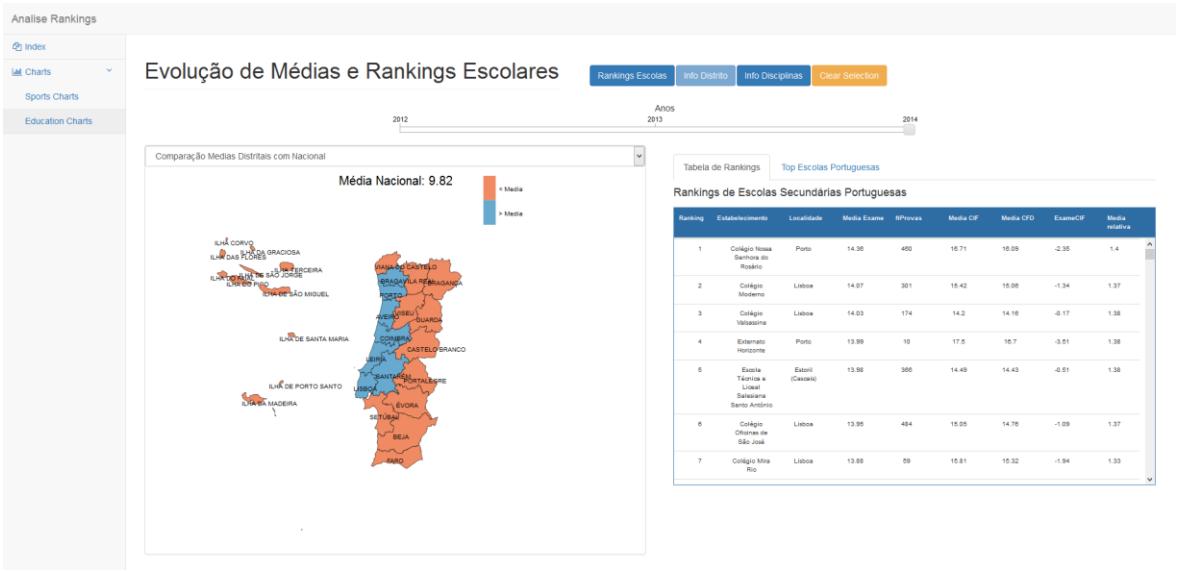

Figura 3.13 - *Dashboard* de Evolução de Médias e Rankings Escolares

O *Dashboard* para as Médias Escolares contém funcionalidades diferentes dos anteriores. Aqui manteve-se a ideia de utilizar uma tabela para representar a informação textualmente, também com a opção de filtrar os dados apresentados. Mas desta vez, o Mapa é a visualização principal e onde se centram a maior parte das interações. Esta visualização apresenta os atributos de um Distrito relativamente a uma escala de cores, tal como sucede num *HeatMap*.

O tipo de informação que pode aqui ser visualizado é:

- Médias de Escolas por Distrito;
- Média Nacional - Relação entre Médias Distritais e Nacional;
- Médias Negativas (menores que 9,5) e Médias Positivas (maiores que 9,5).

O Mapa utiliza uma legenda para indicar as cores dos valores usados nas áreas do mapa. Esta é composta por uma escala de cores, com valores numéricos que estão associados a cada intervalo. O intervalo de valores da escala é sempre calculado para se adaptar ao conjunto de dados que está a ser visualizado.

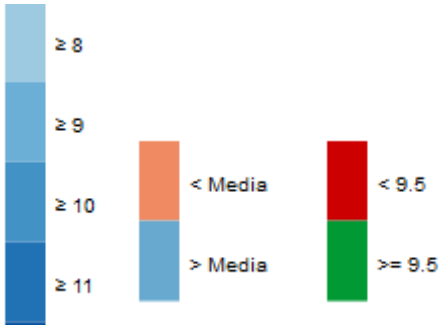

Figura 3.14 - Tipos de Legenda do *HeatMap* para o Mapa de Portugal

Podem ser visualizados dois tipos diferentes de escalas, dependendo dos dados. O primeiro tipo de escala contém um conjunto de cores, onde cada uma abrange um intervalo de valores. A cor apresentada no Distrito é mais saturada dependendo do valor lá representado, ou seja, Distritos com Médias mais elevadas, são representados por cores mais escuras/fortes. O segundo tipo de escala funciona como um indicador de desempenho, onde a cor apresentada no Distrito depende de o valor da Média ser inferior ou superior a um determinado valor, sendo o desempenho negativo ou positivo, respetivamente. Este tipo de escala é utilizado para comparar as Médias Distritais à Nacional, ou se são negativas/positivas.

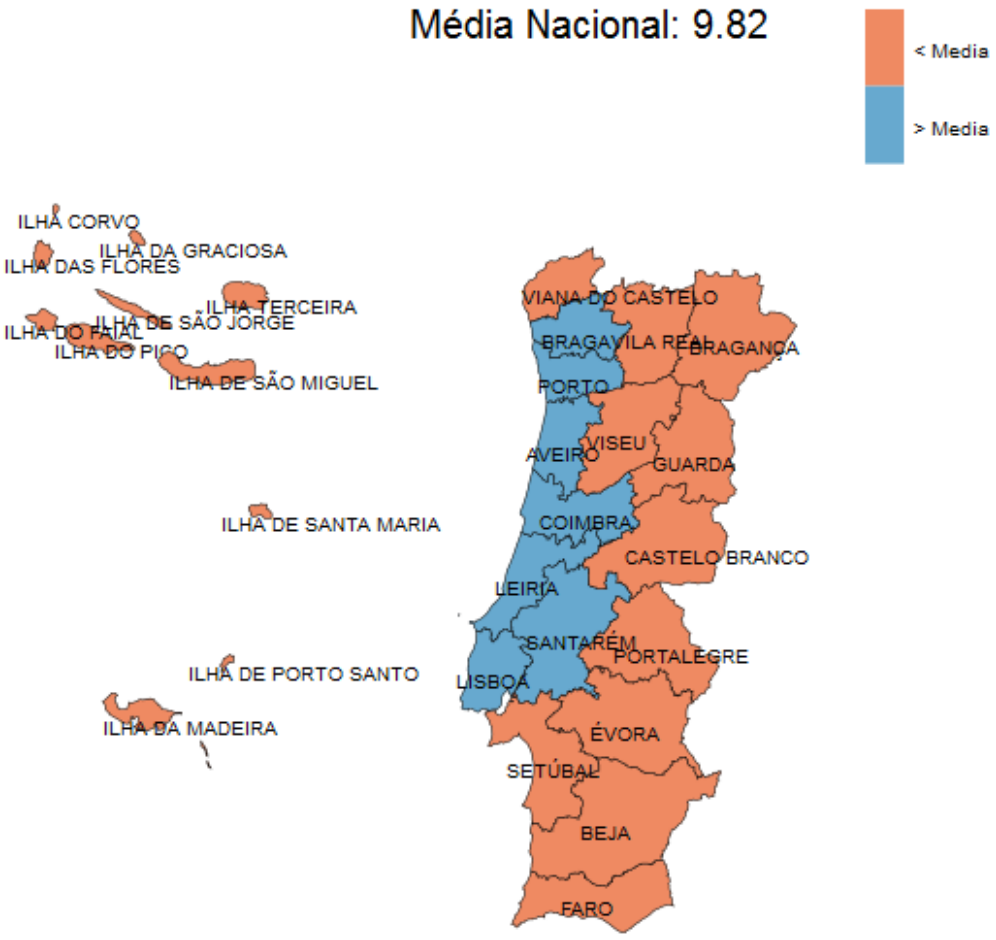

Figura 3.15 - Exemplo da representação Geográfica da diferença entre a Média Distrital e a Média Nacional

Para indicar se o valor se encontra abaixo ou acima da Média Nacional, são usadas as cores Laranja e Azul, respetivamente. Isto porque, uma Média Distrital estar abaixo da Média Nacional, não significa que esta seja negativa e vice-versa, podendo mostrar uma discrepância nos dados. Um exemplo para esta afirmação pode ser visualizado na figura 3.15, aqui os dados apresentados correspondem ao mesmo ano, e pode-se verificar que alguns dos Distritos que têm médias inferiores à média Nacional, mesmo esta sendo baixa, obtiveram Médias Positivas.

Por isso foi utilizada a cor Vermelha e Verde exclusivamente para a escala que indica se a Média Distrital é negativa (inferior a 9.5) ou positiva (superior a 9.5).

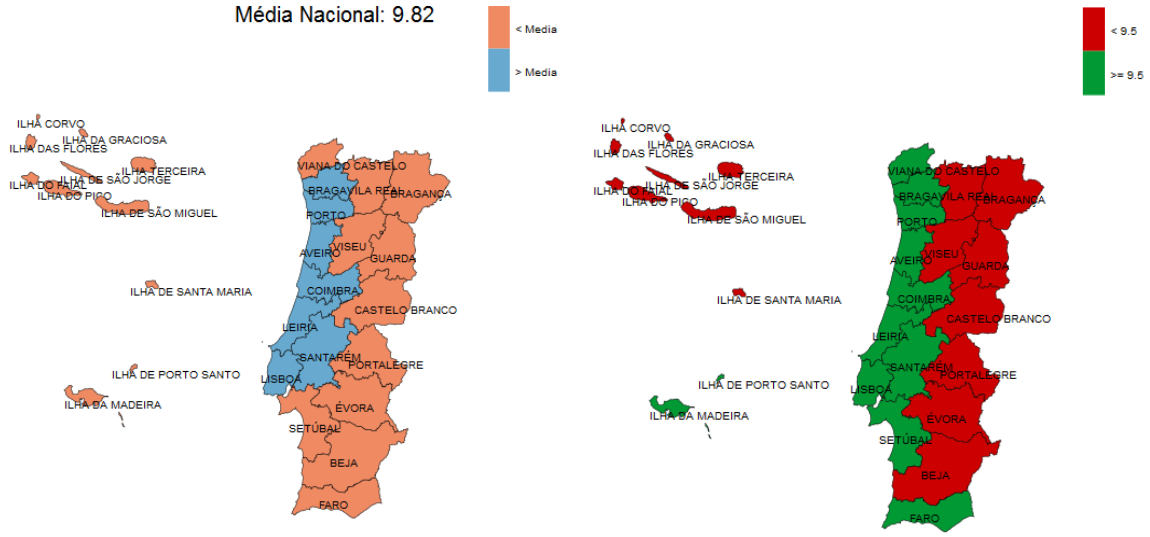

Figura 3.16 - Comparação entre os mapas resultantes para a relação entre Médias Distritais com Nacional e Médias Positivas/Negativas

Ao selecionar um Distrito do Mapa há também a possibilidade de abrir uma nova janela, que apresenta informação das escolas desse Distrito com mais detalhe, apresentando novas visualizações:

- Tabela;
- Gráfico em Estrela;
- Gráfico de barras agrupadas, com Linha de Médias.

A finalidade desta Janela é apresentar o que levou a que o Distrito ou Escola tivesse a média indicada anteriormente, apresentando, para cada escola, os valores das médias por disciplina (Português, Biologia/Geologia, História A, Economia A, Física e Química A, Geografia A e Matemática A). Aqui é possível escolher qual Concelho, do Distrito selecionado, se pretende analisar.

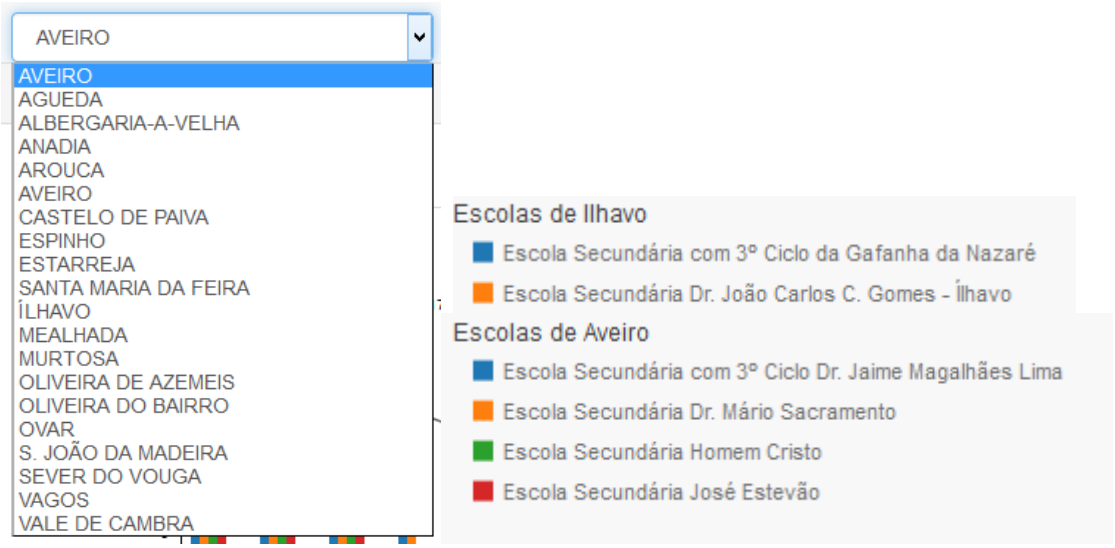

Figura 3.17 - Exemplo de Escolha de Concelhos para o Distrito de Aveiro, e de Legendas para Concelhos Diferentes

A tabela dá ao utilizador a possibilidade de selecionar que escolas quer apresentar. Caso não haja nenhuma escola selecionada na tabela, as visualizações apresentam apenas até às 5 melhores escolas do Concelho, para manter as visualizações limpas e fáceis de analisar.

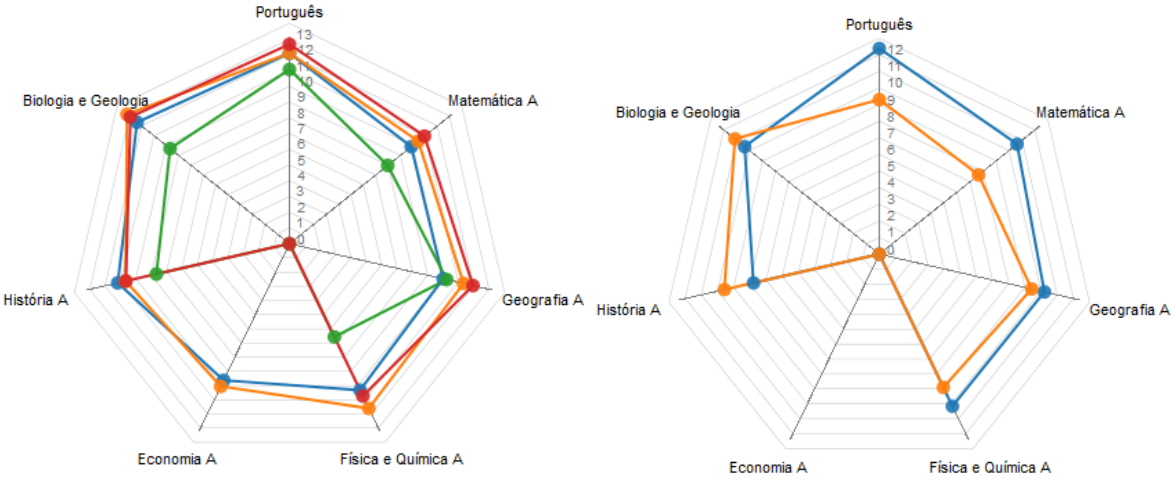

Figura 3.18 - Gráfico de Estrela a apresentar Médias de Escolas de Concelhos Diferentes, Escolas de Aveiro (Esquerda) e Escolas de Ílhavo (Direita)

O Gráfico de Estrela apresenta os valores das médias das disciplinas para as Escolas selecionadas. Nesta visualização é fácil concluir qual das escolas teve uma maior média geral, relativamente à sua área, ou em determinadas disciplinas.

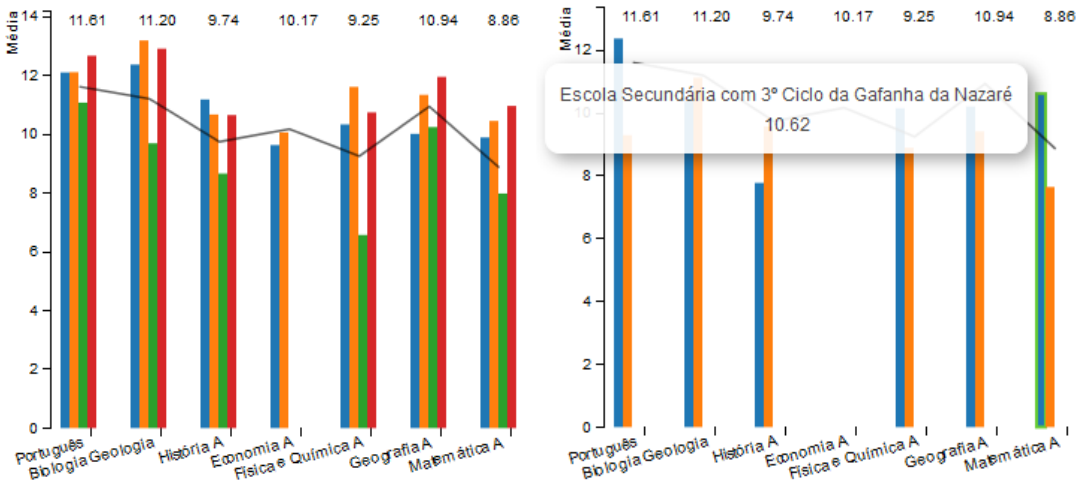

Figura 3.19 - Gráfico de Barras a apresentar Médias de Escolas de Concelhos Diferentes, Escolas de Aveiro (Esquerda) e Escolas de Ílhavo (Direita) com exemplo de *Tooltip*

O Gráfico de Barras Agrupado apresenta as médias das Escolas por disciplina. Neste gráfico cada grupo corresponde a uma disciplina, a cor de cada barra representa uma escola, onde o valor da média é a sua altura. Neste gráfico pode também ser visualizada uma linha horizontal, que representa a Média Distrital, para se poder comparar, para além das escolas que tiverem maior média que outras, se estas se encontram abaixo ou acima da média Distrital.

3.4.5. Conclusões

De seguida são discutidas e apresentadas opções tomadas no desenvolvimento das visualizações, e uma análise do *dashboard*, indicando se este satisfaz os requisitos estipulados.

Foi testado também a utilização de um gráfico de estrela para representar os dados relativos aos *Rankings*. O problema com este gráfico é que os atributos destes dados contêm escalas muito distintas umas das outras e não justificava a utilização de um gráfico tão característico como este. Podia então ser utilizado para apresentar as vitórias, empates e derrotas de uma equipa, mas concluiu-se que, no final, esta representação se tornava mais confusa que os gráficos de barras empilhadas.

No DashBoard de *Rankings* Escolares houve uma grande dificuldade em construir um gráfico de linhas onde estas não se tornassem difíceis de diferenciar. Isto porque, como foi indicado, há uma grande variedade de elementos, mas também, muitas vezes, escolas sobem ou descem bastante a sua posição de *Ranking* no intervalo de um ano, fazendo com que as restantes escolas, onde não houve uma discrepância tão grande no seu *Ranking*, fiquem muito juntas.

De seguida é apresentada uma análise do *Dashboard* e das suas visualizações relativamente aos requisitos expostos no início deste capítulo. Para esta análise foi construída uma tabela onde são apresentados os requisitos e se o *Dashboard* desenvolvido os satisfaz.

Tabela 3.1 - Análise de requisitos para o *dashboard* desenvolvido

Requisitos de Usabilidade	*DashBoard* desenvolvido
R I: Codificar Rankings	+
R II: Codificar atributos dos elementos	++
R III: Suportar múltiplos atributos	++
R IV: Suportar filtros	++
R V: Adaptar escalabilidade para a visualização	+/-
R VI: Tratar valores em falta	-
R VII: Interações	++
R VIII: Feedback visual	+/-
R X: Comparar múltiplos Rankings	++

Com a análise da tabela anterior, pode-se verificar que o *dashboard* e visualizações desenvolvidas cumprem a maior parte dos requisitos estipulados. Foi atribuída uma menor nota ao **RI** pelo simples facto de não ter sido implementada nenhuma forma de calcular o ranking dos elementos baseado nos seus atributos, apenas é indicado de acordo com o estipulado nos dados obtidos. No caso do **RVIII**, foi atribuída esta nota porque, em alguns dos casos, as visualizações não oferecem feedback ao utilizador, no sentido em que não destacam as ações do utilizador, por exemplo o *hover* sobre as barras do gráfico. Para **RV**, algumas das visualizações não lidam muito bem com e excesso de informação e tiveram que ser escolhidas formas alternativas de apresentar os dados. O requisito **RVI** não é cumprido porque não foi desenvolvida nenhuma forma de tratar os valores em falta nos dados, foram apenas aplicadas formas de estes serem suportados em alguns dos casos. O exemplo mais aparente é no caso das médias por disciplina, onde escolas que não fizeram o exame à disciplina são apresentadas com o valor 0 (zero).

4. Testes e Avaliação

Neste capítulo são apresentados e descritos as avaliações heurísticas e os testes de usabilidade efetuados para perceber se as visualizações têm utilidade e se foram implementadas da melhor forma. Este capítulo continua indicando os resultados obtidos nestes testes, pelos analistas, utilizadores e observadores, descrevendo erros, limitações e observações com a utilização e funcionalidade das visualizações.

Neste capítulo é também apresentada a usabilidade da plataforma, avaliada pelos participantes com a utilização da escala padrão SUS – *System Usability Scale* (Brooke 1996), e explicado o seu significado.

4.1. Avaliação Heurística

A avaliação heurística é um método analítico utilizado para analisar a usabilidade de uma interface em avaliação (Forsell 2012). No caso desta Dissertação, o objetivo desta análise foi encontrar problemas de usabilidade nas características e funcionalidades das visualizações.

Para realizar os testes, é necessário que se selecione um número de avaliadores suficientes para encontrar a maior parte dos problemas de usabilidade com as visualizações, por exemplo, entre 3 a 5 utilizadores (Nielsen 2000). Com a utilização de um avaliador, é encontrado um número reduzido de erros, mas quando se seleciona um número elevado de avaliadores, os resultados podem deixar de ser relevantes, começando a haver redundância nos erros encontrados.

Foram efetuadas duas avaliações distintas, uma pelo autor da Dissertação e a segunda por um colega de Mestrado, Pedro Santos. Cada avaliador utilizou duas listas de heurísticas, as Heurísticas de Nielsen (Nielsen 2005) e as Heurísticas de Zuk e Carpendale (Zuk et al. 2006), avaliando a severidade dos problemas de usabilidade encontrados de acordo com uma escala de zero a quatro definida por Nielsen[24].

4.1.1. Resultados

Os resultados das avaliações podem ser analisados no anexo 1 (João Rocha) e no anexo 2 (Pedro Santos). Esta secção serve apenas para apontar conclusões resultantes das avaliações e possíveis alterações que surgiram como seu resultado.

[24] http://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/ (Consultado em 28/10/2015)

Falta de indicação que tabelas podem ser ordenadas

Figura 4.1 - Ordenação da Tabela

Os valores das tabelas podem ser ordenados por ordem crescente, decrescente ou alfabética, mas esta funcionalidade não é indicada, nem há feedback com o ponteiro (alterar aspeto do ponteiro). Uma forma para resolver este problema seria adicionar símbolos ou *tooltips* que indicam que valores podem ser ordenados, e a sua ordenação. Como forma temporária de contornar este problema, foi adicionado um *tooltip* com o *hover* do ponteiro do rato, indicando que os valores podem ser ordenados (Figura 4.1).

Textos do eixo horizontal para gráfico de performance sobrepõem-se

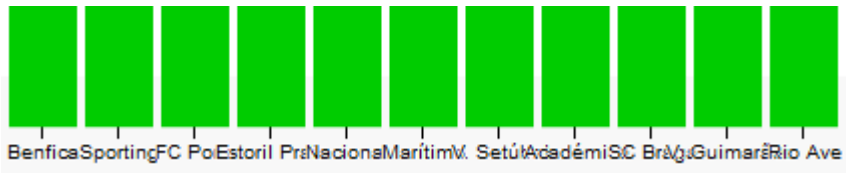

Figura 4.2 - Exemplo de sobreposição de texto

Figura 4.3 - Nova posição de texto, sem sobreposição

Por vezes, dependendo da quantidade de equipas a visualizar, os seus nomes ficam sobrepostos ao longo do eixo horizontal, tornando a sua leitura impossível (Figura 4.2). Este problema agrava-se pelo facto de a visualização não disponibilizar outra forma de verificar que equipas estão a ser apresentadas para além da tabela. Para este problema, os nomes das equipas são apresentados com um certo ângulo de escrita, evitando que se sobreponham (Figura 4.3).

Problema no contraste entre as cores das barras dos gráficos e dos valores apresentados.

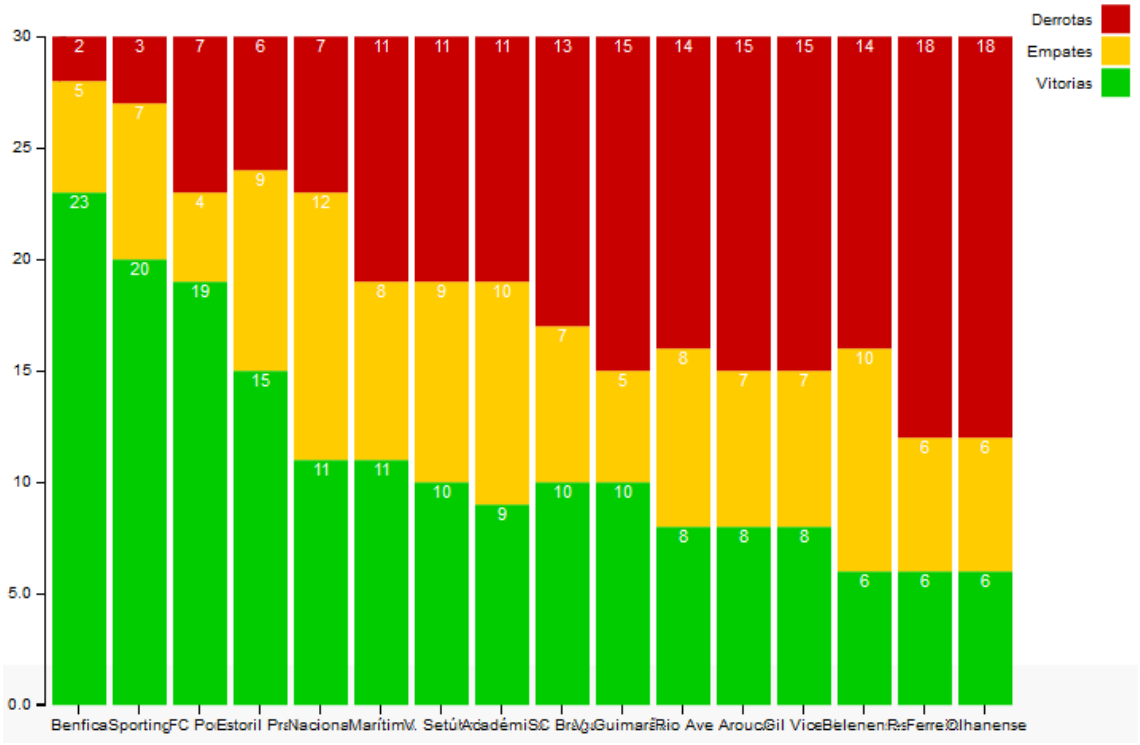

Figura 4.4 - Problema com contraste entre cores das barras e texto

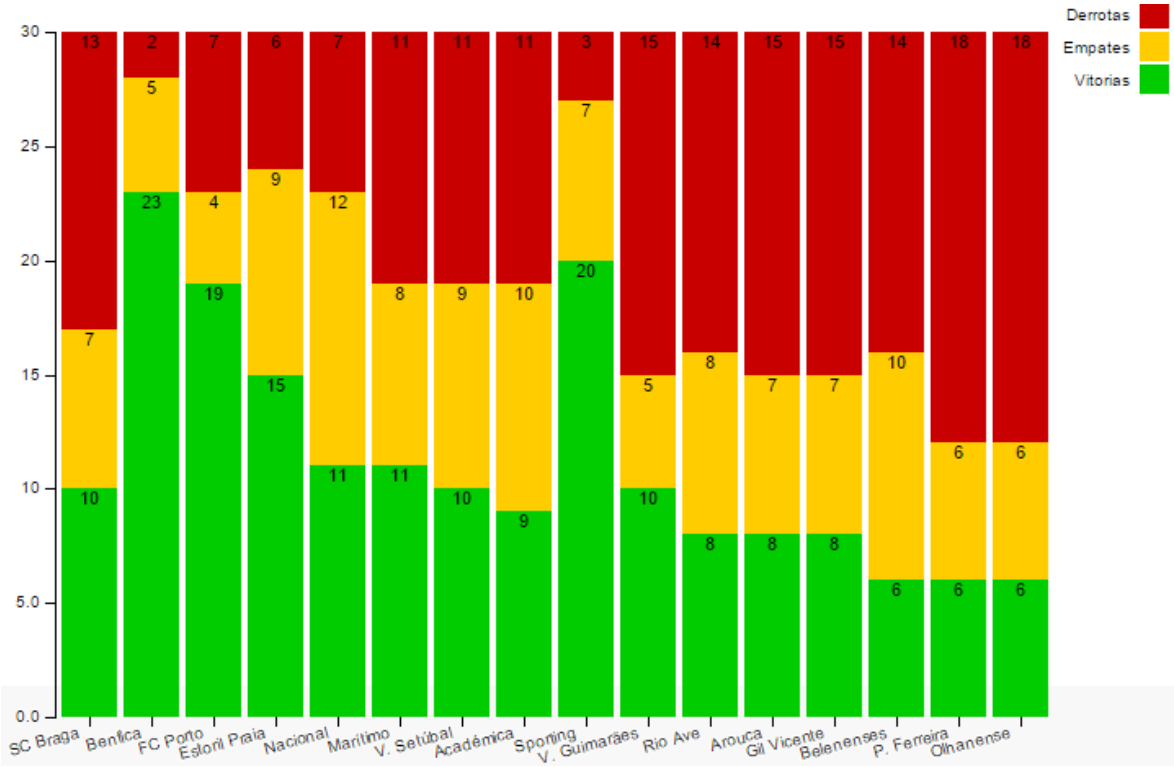

Figura 4.5 - Alteração da cor do texto para as barras do gráfico

Em alguns casos, o texto escrito nas barras, para indicar o valor que representam, pode ser difícil de distinguir, devido ao contraste entre as cores, principalmente texto branco sobre uma barra amarela (Figura 4.4). Como forma de reduzir temporariamente este problema, foi alterada a cor do texto para se visualizar mais facilmente o que se encontra escrito (Figura 4.5). Com esta alteração, também se verifica que as barras verdes e amarelas são mais facilmente distinguíveis, porque a cor do texto altera a forma como as cores das barras são percetíveis. No entanto, as cores das barras também têm que ser modificadas para ser mais fácil a sua distinção em casos de daltonismo.

Uma ferramenta importante utilizada para a seleção de cores e que merece ser referenciada é o ColorBrewer[25]. Esta ferramenta ajuda na seleção de cores, oferecendo uma variedade de opções de escolha e indicando se a combinação de cores é percetível por daltónicos. Esta foi utilizada para definir as cores do mapa, mas é uma boa ajuda para redefinir as cores das restantes visualizações.

Linha da média fica sobre os seus valores na visualização

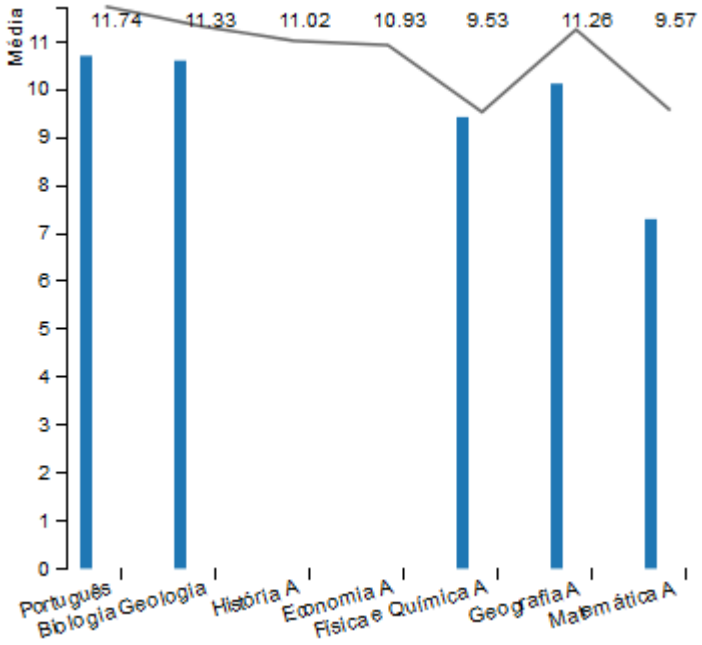

Figura 4.6 - Linha da Média Distrital por cima dos seus valores

Em algumas ocasiões, ao visualizar as médias das escolas de um Distrito, quando as médias das escolas de um Concelho são muito mais baixas que as médias Distritais, a linha da média do Distrito fica sobreposta aos seus valores (Figura 4.6). Este erro vem ao encontro de outro erro indicado, que os valores da linha deveriam ser apresentados de uma forma diferente, para não criar confusão aos utilizadores. Desta forma os valores podem aparecer na forma de uma *tooltip* com o *hover* do rato sobre a linha; apresentando os valores de todas as médias distritais das disciplinas; ou *hover* sobre a barra de uma escola; apresentando o valor da média distrital apenas dessa disciplina. Esta função deveria ser aplicada sem entrar em conflito com as *tooltip* apresentadas sobre as médias das escolas.

25 http://colorbrewer2.org/ (Consultado em 28/0/2015)

Não existe documentação que descreva o funcionamento da plataforma e ajudas à utilização.

Não existe nenhum tipo de opções de ajuda para operar e interagir com as visualizações e *Dashboard*. Mas principalmente, não há descrições sobre o funcionamento das visualizações ou explicação sobre termos técnicos, o que foi também apontado nos testes de usabilidade (por exemplo, o significado de "Médias CIF"). Para contornar este problema, deviam ser implementados controlos de ajuda em todas as páginas para explicar o seu contexto e funcionamento, e também definir termos técnicos.

4.2. Testes de Usabilidade

Os testes de usabilidade servem habitualmente para determinar até que ponto um interface desenvolvido facilita a navegação e a execução de tarefas. Normalmente, estes testes são efetuados por possíveis utilizadores, onde lhes é pedido que efetuem determinadas tarefas de rotina sobre o interface a analisar. As sessões de teste com os utilizadores são registadas e analisadas para identificar possíveis áreas ou funcionalidades do interface que necessitem ser melhoradas ou alteradas.

Esta secção apresenta o feedback dos participantes, uma análise de erros e tarefas corretas, a dificuldade atribuída, tempo das tarefas e uma análise de usabilidade. A documentação utilizada para os testes, incluindo o questionário e a ficha de observador estão incluídos no Apêndice B.

Para desenvolver estes testes foi utilizada a informação disponível no sítio usability.gov[26]. Este disponibiliza uma variedade de documentos e métodos de ajuda à preparação, execução e análise de processos de testes de usabilidade.

4.2.1. Metodologia

O questionário foi dividido por casos de estudo, dando um total de duas tarefas a realizar, com um total de quatro secções de aproximadamente sete questões cada. Uma tarefa para analisar os Rankings de Futebol, e a segunda para analisar os Rankings de Escolas Secundárias. A segunda tarefa encontra-se dividida em três secções, uma secção por cada tipo de dados provenientes dos Rankings de Escolas Secundárias; os Rankings Escolares, as Médias Distritais e as Médias Escolares. Para cada tarefa, o utilizador teria de encontrar a informação no *Dashboard* em estudo, utilizando a visualização que considerasse mais eficaz e fácil de entender.

Cada sessão foi composta por duas fases de trabalhos, na primeira fase é explicado o procedimento do questionário e da resolução das tarefas (ver Apêndice B.1), e apresentada a ferramenta a analisar, explicando funcionalidades e o contexto da informação. Na segunda fase, procedesse a resolução das tarefas do questionário e à sua análise.

No final de cada tarefa, o utilizador teria de avaliar a sua dificuldade de resolução, numa escala de *Likert* de 1 a 5 pontos. Esta escala variava entre "Nada Fácil" (1) e "Muito Fácil" (5). Com as tarefas concluídas, era

[26] http://www.usability.gov/how-to-and-tools/methods/reporting-usability-test-results.html (Consultado em 30/10/2015)

apresentado um questionário aos utilizadores para avaliarem a usabilidade do sítio web. Para este questionário foi utilizada a escala padrão SUS (Satisfaction Usability Scale) (Brooke 1996). Onde são apresentados 10 itens que são respondidos numa escala *Likert* de 1 a 5, variando entre "Discordo Completamente" e "Concordo Completamente". Em conjunção com esta escala, foram acrescentados quatro itens adicionais para aprofundar os resultados obtidos.

Cada utilizador foi supervisionado por um observador (ver Apêndice B.2), que registou o tempo necessário para cada tarefa ser concluída, a forma como esta foi executada, dúvidas, dificuldades e observações do utilizador.

4.2.2. Participantes

Foram contactados doze participantes para efetuar os testes de usabilidade. Dez dos participantes compareceram para realizar os testes, onde cada sessão durou, em média, cerca de trinta minutos. As sessões foram efetuadas em três datas, quatro participantes efetuaram os testes no dia 19 de Outubro, cinco no dia 21 de Outubro e um no dia 25 de Outubro.

Dos dez participantes, cinco eram do género masculino e cinco do género feminino, com idades que variaram entre os 15 e os 49 (Tabela 4.1).

Tabela 4.1 – Intervalos etários dos Participantes

17 ou menor	18 -20	21 - 29	30 - 39	40 - 49
1	1	5	1	2

Foi pedido a cada participante que se autoavaliasse com respeito à experiência que tem com a utilização de computadores usando uma escala *Likert* de 5 pontos (Tabela 4.2). Dos dez participantes, 30% indicou que tinham experiencia média, com a utilização diária da Web (e-mail, *facebook*, etc.), 20% tinha alguma experiencia e 50% tinha muita experiencia na utilização de computadores e navegação na Web.

Tabela 4.2 - Experiência com computadores e navegação na Web

Participante	1	2	3	4	5	6	7	8	9	10
Experiência	3	5	3	4	5	5	5	3	4	5

4.2.3. Resultados

Nesta secção são apresentados os resultados obtidos com a análise do questionário efetuado.

De acordo com o gráfico seguinte (fig.4.7), pode-se verificar que os utilizadores tiveram mais dificuldade a efetuar tarefas relacionadas com a análise de rankings. De acordo com as observações relatadas nas sessões, isto deve-se à má associação entre "classificação" e "Rankings" (Rankings de Futebol, tarefa 1), ou a

confusão e desconhecimento do significado de alguns termos (Rankings de Escolas, tarefa 2.1) e, em muitos casos, lapso de ordenação de valores na tabela (nas duas tarefas). Em algumas situações também houve a confusão entre "pontos" e "classificação" no caso de Rankings de Futebol.

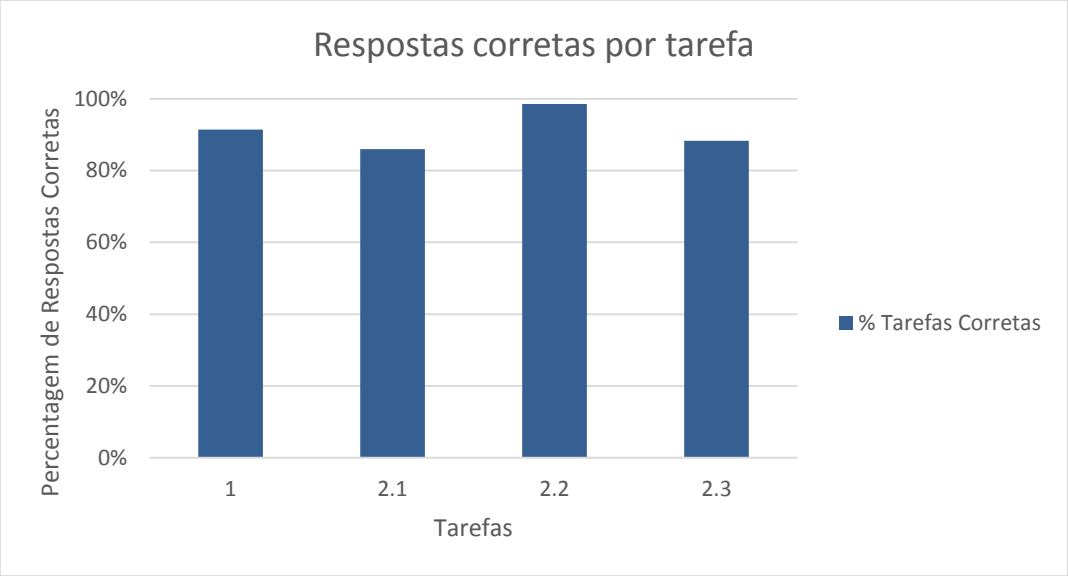

Figura 4.7 - Resultados de Respostas às Tarefas

Na tarefa 1, Rankings de Futebol, em ocasiões, os utilizadores partiram do princípio que o 1º classificado era o que tinha os valores mais altos nos restantes atributos, e não compararam ou ordenaram valores. Para a tarefa 2.1 e 2.2, numa das questões os utilizadores tiveram alguns problemas em entender o que era pretendido, devido ao vocabulário técnico (onde era pedido a comparação entre Média CIF e Média de Exame). Isto gerou confusão porque os utilizadores não sabiam o significado dessa Média. Daí a quantidade de respostas erradas para 2.1 e a diferença positiva no ponto 2.2, que continha uma questão idêntica. Com a tarefa 2.3, houve uma questão onde se perguntava a opinião ao utilizador: "Qual a Escola que considera ter uma melhor média geral?". Nesta questão, quatro utilizadores erraram porque não tiveram atenção à nota de cabeçalho do questionário, explicando os dados e a sua apresentação, isto significa que é necessária uma nova indicação na página.

No final de cada tarefa, os utilizadores avaliavam a facilidade da execução das questões. Esta avaliação era feita numa escala de 5 níveis. Na figura 4.8 é apresentada a facilidade das tarefas, indicada por cada utilizador. Com este gráfico, pode-se também verificar que a facilidade atribuída a cada tarefa está relacionada com a experiencia indicada pelo utilizador, onde utilizadores que indicaram ser mais experientes, avaliaram as tarefas como sendo mais fáceis de resolver.

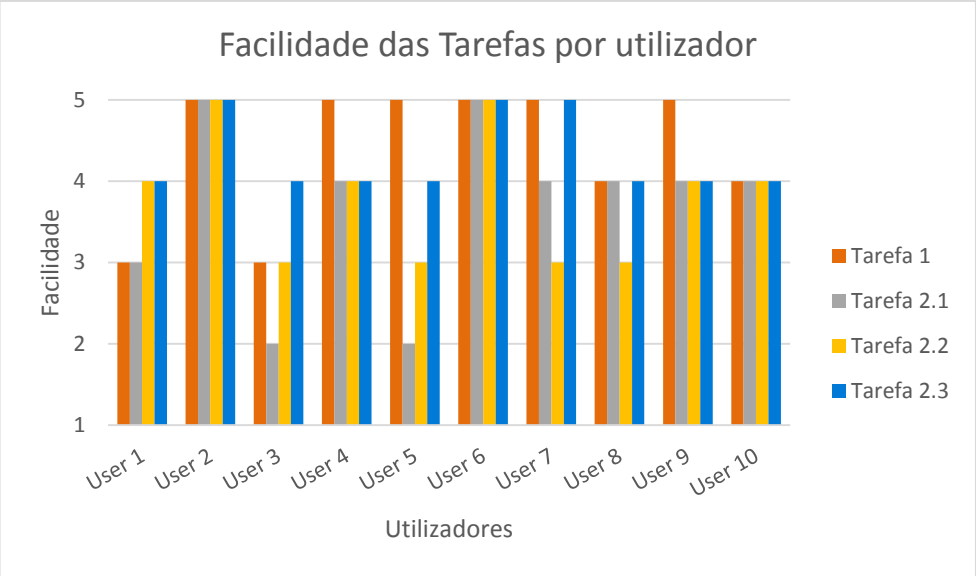

Figura 4.8 - Facilidade das Tarefas por Utilizador

Atendendo ao indicado pelos utilizadores em relação à facilidade das tarefas (Tabela 4.3), verifica-se que a tarefa considerada mais fácil foi a relativa à avaliação de Rankings de Futebol, talvez por estas visualizações serem as que apresentam os dados mais simples e fáceis de avaliar.

Tabela 4.3 - Avaliação da facilidade das Tarefas

Tarefa	Nada Fácil	Complicado	Média	Fácil	Muito Fácil
1 – Ranking Futebol	0 (0%)	0 (0%)	2 (20%)	2 (20%)	6 (60%)
2.1 – Ranking Escolas	0 (0%)	2 (20%)	1 (10%)	5 (50%)	2 (20%)
2.2 – Médias Distritais	0 (0%)	0 (0%)	4 (40%)	4 (40%)	2 (20%)
2.3 – Médias Escolares	0 (0%)	0 (0%)	0 (0%)	7 (70%)	3 (30%)

A tabela 4.4 mostra o tempo que cada utilizador demorou a completar cada tarefa. Nota-se que algumas tarefas foram mais fáceis de resolver que outras e esse feito é refletido pelo tempo que estas demoraram a ser concluídas. Aqui destaca-se um dos utilizadores que demorou 1 hora a efetuar todas as tarefas, mas, entre os dez participantes, foi o único que visualizou os dados em mais que uma visualização e se certificou que estes estavam corretos múltiplas vezes, resultando numa avaliação sem erros.

Tabela 4.4 - Tempo de Execução de Tarefas por utilizador, em horas e minutos

Tarefa	Duração das Tarefas por Utilizador (h:mm)										Tempo Médio
	1	2	3	4	5	6	7	8	9	10	
1	0:07	0:05	0:08	0:05	0:05	0:03	0:06	0:08	0:15	0:05	***0:06***
2.1	0:09	0:04	0:14	0:06	0:08	0:03	0:06	0:06	0:14	0:06	***0:07***
2.2	0:08	0:08	0:12	0:07	0:05	0:05	0:07	0:09	0:14	0:08	***0:08***
2.3	0:05	0:09	0:08	0:07	0:04	0:02	0:06	0:04	0:17	0:04	***0:06***
Tempo Total	**0:29**	**0:26**	**0:42**	**0:25**	**0:22**	**0:13**	**0:25**	**0:27**	**1:00**	**0:23**	***0:29***

As tarefas 2.1 e 2.2 foram as que demoraram mais tempo a serem efetuadas (Médias = 7 minutos e 8 minutos, respetivamente). Nestas tarefas os utilizadores tinham que avaliar os Rankings Escolares (2.1) e as Médias Distritais (2.2). Todas as tarefas originam um tempo médio entre os 6 e os 8 minutos para serem concluídas.

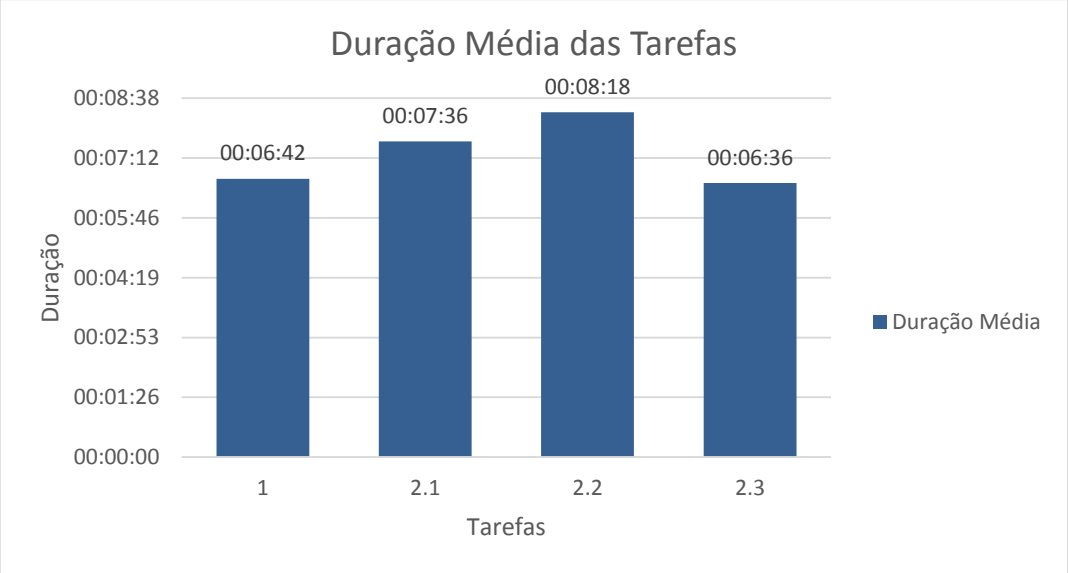

Figura 4.9 - Duração média das tarefas

Na tabela 4.5 podem ser verificados o número de problemas que os utilizadores encontraram ou cometeram durante as tarefas. A maior parte dos erros cometidos deriva das dificuldades com que se depararam, e não conseguiram resolver ou não solicitaram ajuda.

Foram considerados erros críticos, erros que alteravam a perceção dos dados, e originaram respostas erradas.

Para a tarefa 1, as dificuldades que houveram, estavam todas relacionadas com a alteração do gráfico de linhas, como alterar do gráfico que apresenta a evolução da classificação para o gráfico de evolução do total de pontos. Um dos utilizadores não conseguiu fazer a associação entre os termos "Ranking" e "Classificação"

(não foi considerado erro crítico porque não afetou o seu raciocínio e respondeu à questão, observando os pontos e a posição relativa na tabela), e dois utilizadores não ordenaram a tabela, resultando em respostas erradas às questões.

Na tarefa 2.1, a dificuldade foi a definição dos termos técnicos, específicos para quem esta dentro do tema (Médias CIF e CFD).

Para a tarefa 2.2, quatro utilizadores tiveram dificuldades para encontrar onde se alterava as representações do mapa entre médias Distritais, Nacional e Positivas/Negativas. Um utilizador teve problemas a encontrar o valor da Média Nacional na visualização e outro teve dificuldade a perceber onde se alterava do gráfico de barras para a tabela. A dificuldade em alterar a representação do mapa foi considerado um erro crítico, mesmo que o obstáculo tenha sido ultrapassado, porque este é um método muito importante desta visualização.

Na tarefa 2.3 quatro utilizadores tiveram problemas, no gráfico barras, por não perceberem que a linha poligonal representa a média distrital das disciplinas, o que levou a que dois deles, no final, respondessem errado à questão.

Tabela 4.5- Número de Erros cometidos por tarefa e dificuldades encontradas

Tarefa	Dificuldades	Erros	Erros Críticos
1 – Ranking Futebol	4	1	2
2.1 – Ranking Escolas	3	0	0
2.2 – Médias Distritais	6	0	4
2.3 – Médias Escolares	4	0	2

Avaliação de Usabilidade de Sistema

Para avaliar a usabilidade do sistema, como indicado no início desta secção, foi utilizada a escala padrão SUS (Brooke 1996). Esta escala consiste numa lista com 10 afirmações, onde os utilizadores têm que responder utilizando uma escala de *Likert* de 5 níveis (Discordo Completamente - 1; Discordo - 2; Indiferente - 3; Concordo - 4; Concordo Completamente - 5). Para esta escala, a pontuação individual de cada item não é importante, mas sim o cálculo da média dos pontos, que indica a usabilidade do sistema. Para este cálculo, cada item contribui com valores entre os 0 e 4 pontos. Ou seja, para os itens 1,3,5,7 e 9 é calculado o valor do item menos 1. Para os itens 2,4,6,8 e 10 é calculado 5 menos o valor do item. Finalmente, para obter a pontuação da usabilidade do sistema, é multiplicado o resultado da soma dos valores anteriores por 2.5.

Isto resulta num valor que varia entre 0 e 100 (Tabela 4.6).

Tabela 4.6 - Pontuação de Usabilidade de Sistema por utilizador.

User 1	User 2	User 3	User 4	User 5	User 6	User 7	User 8	User 9	User 10	Valor Médio
65	100	72,5	90	75	65	72,5	67,5	92,5	80	**78**

Para determinar o significado dos valores obtidos, considera-se que valores superiores a 68 são acima da média, e inferiores abaixo da média[27]. Na figura 4.10, pode-se verificar com mais detalhe a forma como se podem interpretar os resultados desta escala e avaliar a usabilidade da plataforma.

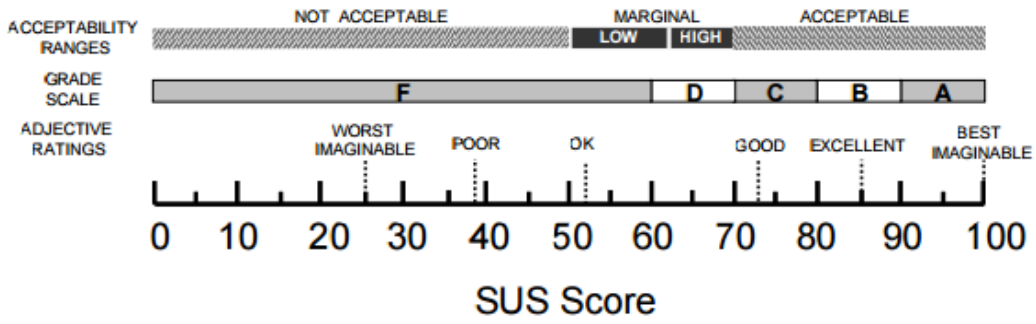

Figura 4.10 - Escala de Usabilidade e interpretação de resultados (Bangor et al. 2009)

Os valores obtidos com esta avaliação oscilam entre o 65 e 100, com média de 78. Este resultado permite concluir que, de acordo com a Escala de Usabilidade do Sistema (SUS), esta plataforma apresenta uma boa usabilidade para os seus utilizadores.

Foram ainda apresentadas quatro afirmações para os utilizadores pontuarem, utilizando a mesma escala *Likert*, que, em vez de avaliar a usabilidade do sistema, permite obter uma ideia geral do que o utilizador achou da aparência das visualizações e da forma como os dados eram apresentados. Para esta análise, 70% dos utilizadores concordou com a afirmação "o conteúdo é claro e compreensível" e os resultados variaram entre a pontuação 4 e 5. Para a segunda afirmação, "Aparência é limpa e organizada", os resultados variaram entre 4 e 5, sendo que 60% concordou completamente com a afirmação. O terceiro ponto, "É difícil encontrar a informação pretendida", obteve resultados entre 1 e 2, onde 60% discordou com a afirmação. Finalmente, o último ponto, "A ferramenta tem utilidade", foi avaliado entre 4 e 5, 90% dos resultados concordando completamente com a afirmação apresentada.

Recomendações e observações

Finalizado o processo de análise, os utilizadores recomendaram alterações e comentaram a sua utilização da plataforma. Este é um passo importante, porque permite que se tome nota, em primeira mão, do que possíveis

[27] http://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html (Consultado em 29/10/2015)

utilizadores esperam da ferramenta e o que acham dela, sem limitações do questionário. Os utilizadores recomendaram as seguintes alterações:

- Ter dois "quadrados" no *slider* para selecionar dois valores em simultâneo, para se compararem dois instantes de tempo nas visualizações;
- Adicionar um eixo vertical com valores de *rankings* no último ano (à direita) do gráfico de linhas, idêntica ao eixo de *rankings*, para facilitar a perceção da evolução da linha entre o primeiro e o último ano*;*
- Indicar a jornada da época futebolística atualmente a ser visualizada nos gráficos de barras, para manter toda a informação necessária visível;
- Melhorar a aplicação de filtros, para suportar variações de minúsculas, maiúsculas, e similaridade de texto/palavras;
- Indicar na página de Médias Escolares que o valor 0 (zero) indica disciplinas a que não foram efetuados exames, não tendo nada a ver com a nota, ou tentar retirar este valor por completo.
- Ter mais informação sobre os dados representados no Mapa, por exemplo, acrescentar um gráfico com os valores das médias distritais;

4.2.4. Conclusões

Com a informação obtida dos diferentes tipos de testes e questionários efetuados, é fácil ter uma ideia geral do estado atual do trabalho efetuado. Também é fácil saber os aspetos que têm que ser alterados, porque foram relatados por analistas durante as heurísticas e pelos utilizadores durante os seus testes de usabilidade, assim como observados durante as sessões.

Estes testes de usabilidade são difíceis de preparar e executar, em comparação aos testes e questionários efetuados pela Web e *crowdsourcing*, tendo em conta o recrutamento de participantes. Embora os resultados não sejam tão quantitativos, considera-se que, qualitativamente, são resultados mais eficazes e abrangentes.

5. Conclusão

Nesta dissertação foram desenvolvidas visualizações interativas e avaliada a sua aplicabilidade para a representação da Evolução temporal de Rankings, usando como casos de estudo Rankings de Futebol e Rankings de Escolas Secundárias. Neste capítulo é apresentado um resumo do trabalho, algumas conclusões e sugestões de trabalho futuro sobre o desenvolvimento da ferramenta, com o intuito de melhorar a sua usabilidade.

5.1. Conclusão

Este trabalho teve como principal objetivo o desenvolvimento de visualizações interativas, capazes de representar a evolução de rankings, com diferente profundidade de atributos. Foram analisadas publicações e técnicas desenvolvidas com intuitos similares, para avaliar a melhor forma de prosseguir com o desenvolvimento deste trabalho.

Para o trabalho foi utilizada a biblioteca D3.js para o desenvolvimento de visualizações suportadas em ambiente Web. Para os casos de estudo e as visualizações desenvolvidas, foi criado um *dashboard* que apresenta dois tipos de dados estruturalmente diferentes:

- A evolução das classificações de Equipas da 1ª Liga Portuguesa de futebol;
- A evolução das classificações de Escolas Secundárias Portuguesas em relação às notas em Exames, apresentando também a distribuição das médias por distrito.

As visualizações e *dashboards* desenvolvidos foram submetidos a dois tipos de avaliações distintas: avaliação heurística; e testes de usabilidade, efetuados por dez participantes.

O trabalho oferece formas de analisar a evolução de rankings de elementos ao longo de um período de tempo, e a possibilidade de interagir com o instante de tempo. De acordo com o recolhido durante os testes com utilizadores, está implementado de forma eficaz, não causando nenhuma confusão ou dificuldade com a sua utilização.

Os resultados da avaliação indicam que os utilizadores consideram que as visualizações desenvolvidas têm boa usabilidade, sendo fácil navegar e perceber padrões nos dados, assim sendo úteis para a avaliação da evolução de rankings.

Uma das principais vantagens da elaboração deste trabalho reside no facto de se proporcionar um melhor conhecimento sobre os aspetos fundamentais à criação de visualizações para ambiente Web.

5.2. Trabalho Futuro

Foram apontadas e descobertas limitações do trabalho desenvolvido que poderão ser tópicos a melhorar no futuro, melhorando a usabilidade permitindo atingir um maior público-alvo. De seguida são apresentados tópicos com possíveis melhorias:

- Resolução de problemas de usabilidade mais graves, relatados nas análises heurísticas;
- Resolução de problemas e limitações apontadas pelos utilizadores;
- Implementação de sugestões indicadas pelos utilizadores;
- Análise de outros casos de estudo (por exemplo, Rankings de Áreas Cientificas com mais publicações; Rankings de Autores, dado o número de publicações, citações, etc.);
- Alterar a técnica de mapeamento dos dados, tornando as visualizações independentes dos dados utilizados.

6. Referências

Bangor, A., Kortum, P. & Miller, J., 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3), pp.114–123.

Bostock, M. & Heer, J., 2009. Protovis: A graphical toolkit for visualization. *IEEE Transactions on Visualization and Computer Graphics*, 15(6), pp.1121–1128.

Bostock, M., Ogievetsky, V. & Heer, J., 2011. D3; Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12), pp.2301–2309.

Brooke, J., 1996. SUS - A quick and dirty usability scale. *Usability evaluation in industry*, 189(194), pp.4–7.

Dewar, M., 2012. *Getting Started with D3*,

Fekete, J.D., 2004. The InfoVis Toolkit. In *Proceedings - IEEE Symposium on Information Visualization, INFO VIS*. pp. 167–174.

Fields, E. & Sheppard, B., 2013. A New Approach to Business Intelligence: Rapid-fire BI. , (February).

Forsell, C., 2012. Evaluation in Information Visualization: Heuristic Evaluation. *Information Visualisation (IV), 2012 16th International Conference on*, pp.136–142.

Gratzl, S. et al., 2013. LineUp: Visual analysis of multi-attribute rankings. *IEEE Transactions on Visualization and Computer Graphics*, 19(12), pp.2277–2286.

Heer, J., Bostock, M. & Ogievetsky, V., 2010. A Tour Through the Visualization Zoo. *Commun. ACM*, pp.59–67.

Heer, J., Card, S.K. & Landay, J. a., 2005. Prefuse: a Toolkit for Interactive Information Visualization. *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '05*, p.421.

Hom, D., Perez, R. & Williams, L., 2014. Which chart or graph is right for you ? *Tableau*.

Juice, 2009. *A Guide to Creating Dashboards People Love to Use*,

Liu, S. et al., 2014. A Survey on Information Visualization: Recent Advances and Challenges. *The Visual Computer*, pp.1373–1393.

Malcolm Maclean, 2015. *D3 Tips and Tricks Interactive Data Visualization in a Web Browser*,

Murray, S., 2013. *Interactive Data Visualization for the Web*,

Nielsen, J., 2005. Ten Usability Heuristics. *Communications of the ACM*, 3(1990), pp.1–2.

Nielsen, J., 2000. Why You Only Need to Test with 5 Users. *Jakob Nielsens Alertbox*, 19(September 23), pp.1–4.

Perin, C., Vuillemot, R. & Fekete, J., 2014. A table! In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. New York, New York, USA: ACM Press, pp. 887–896.

Rueter, M. & Fields, E., 2012. Tableau for the Enterprise : An Overview for IT. , (May).

Shi, C. et al., 2012. Rankexplorer: Visualization of ranking changes in large time series data. *IEEE Transactions on Visualization and Computer Graphics*, 18(12), pp.2669–2678.

Tableau, 2014. Visual analysis best practices - Simple techniques for making every data viz beautiful. *Tableau*.

Vuillemot, R. & Perin, C., 2015. Investigating the Direct Manipulation of Ranking Tables for Time Navigation. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. pp. 2703–2706.

Yi, J.S. et al., 2007. Toward a deeper understanding of the role of interaction in information visualization. *IEEE transactions on visualization and computer graphics*, 13, pp.1224–31.

Zuk, T. et al., 2006. Heuristics for Information Visualization Evaluation. *Beliv'06*, pp.1–6.

Apêndice A Ficheiros de Dados

Apêndice A.1 Dados de Rankings Desportivos

A tabela seguinte apresenta a forma como os dados estão estruturados no ficheiro, listando apenas as primeiras 5 equipas nas primeiras 2 jornadas. No documento os dados são apresentados em formato de tabela para facilitar a sua compreensão e análise, mas na realidade estão armazenados e são lidos em formato CSV. Nas apresentações e na Tabela é utilizado o atributo "Jogos" para separar os dados por Jornadas, para as apresentar individualmente.

Ranking	Equipa	Pontos	Jogos	Vitorias	Empates	Derrotas	Golos Marcados	Golos Sofridos
1	Sporting	3	1	1	0	0	5	1
2	Rio Ave	3	1	1	0	0	3	0
3	Estoril Praia	3	1	1	0	0	3	1
4	FC Porto	3	1	1	0	0	3	1
5	V. Guimarães	3	1	1	0	0	2	0
1	Sporting	6	2	2	0	0	9	1
2	FC Porto	6	2	2	0	0	6	1
3	Rio Ave	6	2	2	0	0	5	0
4	Estoril Praia	6	2	2	0	0	5	2
5	SC Braga	6	2	2	0	0	4	1

Apêndice A.2 Dados de Rankings Escolares

Esta tabela apresenta um excerto dos dados que representam os Rankings Escolares. Tal como os dados desportivos, estes encontram-se armazenados no formato CSV.

Ranking	Escola	Localidade	Media Exame	NProvas	Media CIF	Media CFD	CIF - Exame	Media relativa	Ano Letivo
1	Colégio Nossa Senhora do Rosário	Porto	14.36	460	16.71	16.09	-2.35	1.40	2014
2	Colégio Moderno	Lisboa	14.07	301	15.42	15.06	-1.34	1.37	2014
3	Colégio Valsassina	Lisboa	14.03	174	14.20	14.16	-0.17	1.38	2014
4	Externato Horizonte	Porto	13.99	10	17.50	16.70	-3.51	1.38	2014
5	Escola Técnica e Liceal Salesiana Santo António	Estoril (Cascais)	13.98	366	14.49	14.43	-0.51	1.38	2014
6	Colégio Oficinas de São José	Lisboa	13.95	484	15.05	14.76	-1.09	1.37	2014
7	Colégio Mira Rio	Lisboa	13.88	59	15.81	15.32	-1.94	1.33	2014
8	Colégio de Manuel Bernardes	Lisboa	13.86	167	14.67	14.51	-0.81	1.35	2014
9	Externato Casa Mãe	Paredes	13.85	110	16.74	15.95	-2.89	1.35	2014
10	Colégio Arautos do Evangelho	Guimarães	13.83	12	14.50	14.42	-0.68	1.30	2014

Apêndice B Testes de Usabilidade

Apêndice B.1 Tarefas de Utilizadores

Visualização da Evolução Temporal de Rankings

Sessão de Avaliação

Este documento apresenta uma **lista de tarefas** a serem executadas sobre as visualizações do sítio web apresentado e um questionário sobre a sua experiencia enquanto as efetuava. Este teste de usabilidade destina-se a avaliar o sistema e não o utilizador. Como tal, não se sinta pressionado pelo tempo nem pela necessidade em completar as tarefas com sucesso. Se sentir dificuldades, pode pedir ajuda ou desistir livremente de uma ou mais tarefas.

Género: | M | F |

Idade: _____

Nível de Escolaridade: _____

Experiencia com Computadores e Navegação na Internet**:**

Nenhuma						Muita
	1	2	3	4	5	

Tarefa 1	Questionário para os **Rankings Desportivos**, disponível através do menu de navegação. Responda às seguintes questões observando e interagindo com os *rankings* das equipas e respetivas visualizações, para os dados de **2012/2013**. Para este processo pede-se que, entre as Equipas apresentadas na tabela, **selecione duas equipas** para efetuar as tarefas apresentadas: **Equipa A:**_____ **Equipa B:**_____ **A)** Qual a Classificação da equipa A na Jornada 15? **Resposta:** _____ **B)** Que equipa venceu a Liga? **Resposta:** _____ **C)** Qual a equipa com mais Golos Marcados na Jornada 10? **Resposta:** _____ **D)** Em relação aos Golos, para as equipas A e B, qual tem o melhor saldo de golos marcados-sofridos até à Jornada 15? **Resposta:** _____ **E)** Que equipa manteve uma posição mais consistente durante toda a Época? **Resposta:** _____ **F)** Entre as equipas A e B, qual obteve mais pontos entre as Jornadas 5 e 15? **Resposta:** _____ **G)** Das equipas A e B, qual obteve mais vitórias até à Jornada 15? **Resposta:** _____ Facilidade da Tarefa 1) Nada Fácil | 1 | 2 | 3 | 4 | 5 | Muito Fácil

Tarefa 2.1	Questionário para os **Rankings de Escolas Secundárias Portuguesas**, disponível através do menu de navegação **Rankings Escolares**. Responda às seguintes questões observando e interagindo com os *rankings* das Escolas e respetivas visualizações. **A)** Que Escola se encontra em 1º nos Rankings do ano 2014? **Resposta:** _____ **B)** Em que posição se encontra a Escola Secundária José Estevão no ano 2014? **Resposta:** _____ **C**) Qual a Escola com maior Média CFD (Classificação Final da Disciplina) no ano 2012? **Resposta:** _____ **D)** Que Escola obteve uma maior diferença entre as Médias CIF (Classificação Interna Final) e Média de Exames (ExameCIF) para o ano 2012? **Resposta:** _____ **E)** Entre as Escolas de Aveiro, qual obteve uma melhor evolução de classificação entre os anos 2013 e 2014? **Resposta:** _____ Facilidade da Tarefa 2.1) Nada Fácil | 1 | 2 | 3 | 4 | 5 | Muito Fácil

Tarefa 2.2	Questionário para as **Médias Distritais de Escolas Secundárias Portuguesas**, disponível através do submenu de navegação **Médias Distritais**, acessível por **Rankings Escolares**. **A)** Qual a Escola com maior Média de Exames no ano 2012? **Resposta:** _____ **B)** Qual a Escola de Aveiro que obteve a maior Média de Exames em 2012? **Resposta:** _____ **C)** Qual a Escola de Aveiro com maior diferença entre as Médias CIF e Média de Exames (ExameCIF) para o ano 2012? **Resposta:** _____ **D)** Que Distrito obteve a melhor média Distrital no ano 2014? **Resposta:** _____ **E)** Em que ano foi menor a Média Nacional? **Resposta:** _____ **F)** Em que ano houve mais Escolas com Média Distrital inferior à Nacional? **Resposta:** _____ **G)** Em que ano ocorreu um maior número de Médias Negativas? **Resposta:** _____ Facilidade da Tarefa 2.2) Nada Fácil | 1 | 2 | 3 | 4 | 5 | Muito Fácil

Tarefa 2.3	Questionário para as **Médias de Escolas Secundárias Portuguesas**, disponível através da página de **Médias Distritais**, com a seleção de um Distrito. Para estas tarefas, pede-se que **selecione o ano 2014 e o Distrito de Aveiro**. Clique no botão "**Médias Escolares**" disponível no topo da página. ***Nota:*** *Valores a 0 (Zero), não indicam necessariamente que a escola tem média 0 para a disciplina, indica que não foram efetuados exames a essa disciplina na escola.* **A)** Qual a Disciplina com Menor Média Distrital? **Resposta:** _____ **B)** Que Escola obteve a melhor média a Português? **Resposta:** _____ **C)** Qual a média de Português da Escola Secundária José Estevão? **Resposta:** _____ **D)** Para o Concelho de Ílhavo, como se compara a nota de Português da Escola Secundária com 3º Ciclo da Gafanha da Nazaré com a média Distrital? **Resposta:** _____ **E)** Para o Concelho de Aveiro, que Escola considera que obteve uma melhor Média Geral? **Resposta:** _____ **F)** Entre a Escola Secundária com 3º Ciclo Dr. Jaime Magalhães Lima e a Escola Secundária Dr. Mário Sacramento, qual obteve uma melhor média a Português? **Resposta:** _____ Facilidade da Tarefa 2.3) Nada Fácil | 1 | 2 | 3 | 4 | 5 | Muito Fácil

Escala de Usabilidade de Sistema

Este questionário serve para apreciar a usabilidade do Sistema que avaliou. Preencha os quadrados com uma cruz de acordo com o que sentiu enquanto utilizava o sistema.

1. Eu gostaria de usar este sistema com frequência.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
2. O sistema é desnecessariamente complexo.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
3. Acho o sistema fácil de usar.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
4. Preciso de ajuda para operar o sistema.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
5. As diversas funções deste sistema foram bem integradas.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
6. Existem muitas inconsistências no sistema.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
7. A maior parte das pessoas aprenderia a usar o sistema rapidamente.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
8. O sistema é muito complicado de usar.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
9. Eu me senti muito confiante com a utilização do sistema.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
10. É preciso aprender muitas coisas antes de usar o sistema.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
A. O conteúdo é claro e compreensível.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
B. Apresenta uma aparência limpa e organizada.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
C. É difícil encontrar a informação pretendida.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	
D. O Sistema tem utilidade.	Discordo Completamente						Concordo Completamente
		1	2	3	4	5	

Apêndice B.2 Ficha do Observador

Tarefa	Completou a tarefa	Sentiu-se perdido / Cometeu erros?	Solicitou ajuda?
1	☐Sim ☐Não	☐Sim:_____ ☐Não	☐Sim: _____ ☐Não
	Observações:		
	Hora Início / Fim ___:___/___:___		

Anexos

Anexo 1: Avaliação de João Rocha

Avaliação Heurística

Uma avaliação heurística é um método utilizado por analistas para avaliar a usabilidade das interfaces de ferramentas. Com esta análise são determinados os aspetos considerados problemáticos da interface e atribuídas severidades, dentro de uma determinada escala, de forma a separar esforços para corrigir os problemas mais danosos.

Esta avaliação surgiu devido à necessidade de efetuar uma avaliação heurística à usabilidade do sítio web disponível em http://sweet.ua.pt/jmadeira/Vis_Rankings/pages/index.html. Este contem visualizações desenvolvidas para avaliar a evolução temporal de rankings, desenvolvido como tema de dissertação de mestrado em Sistemas de Informação, do departamento de Eletrónica, Tecnologias e Informação da Universidade de Aveiro.

Para esta avaliação foram utilizadas duas listas de heurísticas distintas, as heurísticas de Nielsen[1] [1] e as heurísticas de Zuk [2].

Lista dos 10 princípios Heurísticos de Nielsen

1. **Visibilidade do *status* do sistema.** Se o sistema apresenta feedback ao utilizador, indicando a sua atual "posição" e tarefas. Também com atenção ao tempo da resposta.
2. **Equivalência entre o sistema e o mundo real.** "Vocabulário" ou símbolos utilizados para descrever ou representar uma função não se diferem do "mundo real". Informação deve ser apresentada de uma forma natural e lógica.
3. **Controle do usuário e liberdade.** Utilizadores têm a opção de sair rapidamente de qualquer janela/separador, nunca ficando presos a uma visualização ou página.
4. **Consistência e padrões.** O mesmo símbolo, texto ou ações em diferentes páginas ou visualizações deve efetuar as mesmas funções.
5. **Prevenção de erro.** A interface deve prevenir que o utilizador cometa erros durante a utilização da ferramenta.
6. **Reconhecer ao invés de relembrar.** O utilizador não deve ser obrigado a lembrar-se de informação que tenha visto em uma janela anterior para poder perceber os dados expostos na atual. Informação importante de ser sempre visível ou facilmente acessível.
7. **Flexibilidade e eficiência de uso.** A ferramenta pode deixar que utilizadores mais experientes utilizem (por ex.:) atalhos quando utilizam esta ferramenta, aumentando o seu desempenho.

[1] http://www.nngroup.com/articles/ten-usability-heuristics/

8. **Estética e *design* minimalista.** A interface não deve ter informação irrelevante para o utilizador, ou que seja raramente utilizada. Cada excerto de informação irrelevante afeta diretamente a informação relevante, dificultando a sua visibilidade.
9. **Auxílio ao usuário para reconhecer, diagnosticar e recuperar- se de erros.** Mensagens de erro devem ser fáceis de entender, utilizando diálogo que o utilizador consiga interpretar e não códigos.
10. **Ajuda e documentação.** A ferramenta deve ter documentação de ajuda, mesmo que seja de fácil utilização. Esta ajuda tem que ser focada na ação que o utilizador que efetuar, com todos os paços e não ser demasiado extensa. Pode ser utilizadas *tooltips* ou páginas apropriadas para esta função.

Lista dos princípios Heurísticos de Zuk e Carpendale

1. Garantir que os elementos visuais têm dimensão suficiente
2. A cor não deve transmitir qualquer tipo de ordem
3. Perceção de cores varia com o tamanho do item colorido
4. Conforme o ambiente/contexto a perceção da cor varia
5. Tem em consideração pessoas Daltónicas
6. A perceção subconsciente melhora com o aumento do campo de visão
7. Avaliação quantitativa requer variação do tamanho ou posição
8. Preservar posição dos dados para dimensão dos gráficos
9. Visualizações relacionadas com a dimensão de dados - *Put the most data in the least space*
10. Não existência de informação adicional que distraia o utilizador do que é relevante ou dificulte a realização de comparações - *Remove the extraneous (ink)*
11. Utilização de propriedades como proximidade, similaridade, continuidade, simetria, fecho, conexidade e tamanho relative de modo a facilitar a perceção de padrões
12. Oferece múltiplos níveis de detalhe
13. Inclusão de texto sempre que possa ser útil; *tooltips*, por exemplo

Escala de Severidade dos Erros

Para a avaliação, foi utilizada uma escala "recomendada" por Nielsen, a escala utilizada para exemplificar os seus artigos. Ao serem atribuídos estes graus de severidade, é mais fácil distribuir esforços para a correção de problemas mais relevantes, com valores mais altos na escala.

A severidade de um problema de usabilidade de uma interface é calculado tendo em conta os seguintes fatores:

- A frequência com que o problema ocorre. (comum ou raro)
- O impacto do problema caso ocorra. (fácil ou não de contornar)
- A persistência do problema. (acontece uma vez ou é recorrente)

Ao se analisarem estes fatores, pode-se então passar à análise da severidade destes problemas, atribuindo valor a cada um de acordo com uma escala acordada, escala utilizada por Nielsen nos seus artigos, como indicado anteriormente.

- 0: Não se considera um problema.
- 1: Problema cosmético. Resolver se houver tempo.
- 2: Problema menor. Baixa prioridade
- 3: Problema grave. Alta prioridade
- 4: Problema catastrófico. Tem que ser resolvido.

O valor 0 (zero) da escala não é utilizado em casos que a análise dos problemas é feita por pessoas diferentes, ou seja, houvesse utilizadores para identificar as falhas de usabilidade e utilizadores diferentes para atribuir a severidade.

Erros de Usabilidade

Nesta secção serão apresentados os aspetos que foram considerados problemas de usabilidade da interface da ferramenta analisada. Primeiro será indicada a heurística e de seguida os problemas que as violam, assim como a gravidade ponderada.

Heurísticas de Nielsen

Visibility of system status

Geral:

- Não há indicação que a tabela pode ser ordenada a clicar nos *headers*. Pode ser resolvido pela adição de glifos que indicam que as linhas podem ser ordenadas por valor. **Severidade 3.**
- Feedback para filtro da tabela, a indicar se filtro foi aplicado ou limpo. Indicar valor que está a ser filtrado e estado de filtragem ("Concluído", "A Filtrar", etc.). **Severidade 3.**

Rankings Equipas:

- Não há interação *MouseOver* com os gráficos de Barras. Adicionar destaque quando se passa o ponteiro do rato sobre as barras, podendo também existir *tooltips* com informação sobre os dados a visualizar. **Severidade 1.**

Rankings Escolares:

- Indicar que os gráficos apresentam as n top escolas. Gráficos não indicam que apresentam apenas um determinado número de escolas. **Severidade 2.**
- Interação incorreta com gráfico de coordenadas paralelas. Ao selecionar uma linha da tabela que não esteja desenhada no gráfico, a linha não é desenhada, mas faz fade a todas as existentes. **Severidade 4.**

Médias Escolares:

- *Mouseover* com barras do gráfico pouco visível em alguns casos. Alterar a forma como o destaque da barra é efetuado, para facilitar a sua perceção. **Severidade 3.**
- Não é apresentado feedback ao utilizador quando um conselho escolhido não tem escolas associadas. Quando um concelho não tem escolas, podem ser limpas as visualizações e apresentado um alerta. Ou simplesmente um alerta sobre as visualizações em sombreado. **Severidade 4.**
- Não mantem a seleção entre o Gráfico e o mapa quando altera o tempo dos dados. **Severidade 4.**
- Escolas selecionadas na tabela permanecem selecionadas, mesmo quando se altera o Concelho. Pode ser aplicado, neste caso, uma função para limpar todas as seleções e filtros. **Severidade 4.**

Match between system and the real world

Rankings Equipas:

- Valores que indicam golos sofridos são apresentados com números negativos. Mas os valores da direita representam valores com definições diretamente opostas aos da direita. **Severidade 1.**

User control and freedom

Geral:

- Faltam formas de limpar filtros e seleções em grande parte das páginas. **Severidade 3.**

Consistency and standards

Rankings Equipas:

- Textos do eixo x para gráfico de performance sobrepõem-se. Aplicar angulo de escrita ao texto, para tentar contrariar o acontecimento. **Severidade 3.**

Figura 1 - Exemplo de sobreposição de texto

Rankings Escolares:

- Eixo x do gráfico de linhas contem valores com casas decimais em vez de valores inteiros que representem o ano. Ex. 2,014.0 = 2014. **Severidade 4.**

Figura 2 - Formatação errada dos valores

Médias Escolares:

- Menus que não apresentam alteração. **Severidade 4.**

Error prevention

Rankings Escolares:

- Forma incorreta de filtragem de valores. Apenas apresenta resultados se dados introduzidos forem exatamente iguais. **Severidade 3.**

Recognition rather than recall

Rankings Equipas:

- Não há opção de verificar jornada atual nos gráficos de Barras, e como estão distantes do slider, é difícil manter os dois visíveis em simultâneo. **Severidade 1.**

Rankings Escolares:

- Não há uma linha que indique o ano a ser visualizado no gráfico de linhas, ao contrário da visualização utilizada para Rankings de Futebol. **Severidade 2.**
- Não há correspondência entre linhas do gráfico de coordenadas paralelas e Escolas que estas representam. Adicionar Legendas ou Títulos. **Severidade 4.**

Médias Escolares:

- Indicar que o *DropDown* é para escolher o Concelho. **Severidade 1.**
- Indicar a que ano correspondem os dados. **Severidade 2.**

Aesthetic and minimalist design

Rankings Equipas:

- O contraste entre as cores das barras dos gráficos e dos valores lá apresentados podem causar com que o valor seja difícil de se ler. Principalmente cor branca sobre uma cor amarela. Por exemplo, gráfico de performance. **Severidade 2.**

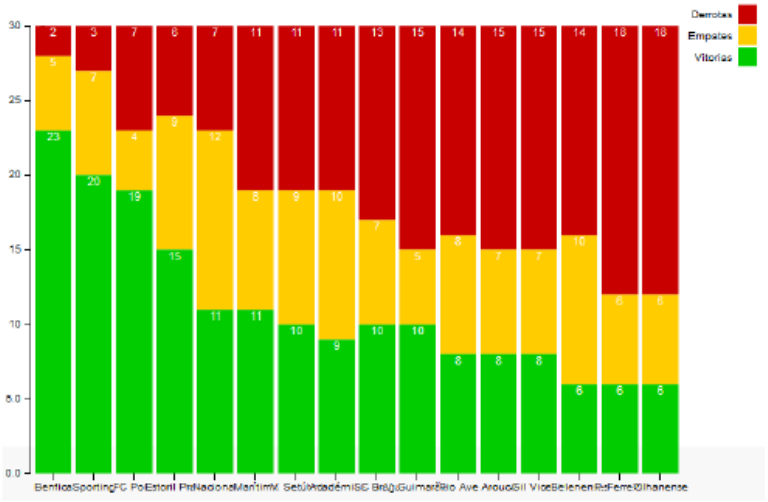

Figura 3 - Problema com contraste entre cores das barras e Texto

Rankings Escolares:

- Alguns eixos do gráfico de coordenadas paralelas contem valores repetidos. **Severidade 3.**

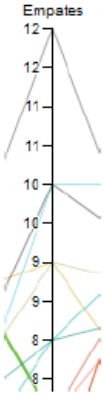

Figura 4 - Valores Repetidos no Eixo de Empates

Médias Escolares:

- Contraste entre cores do texto com nome nos Distritos e a cor da área, ou a cor do texto e a cor co contorno da área do Distrito, pode tornar difícil a leitura do texto. **Severidade 2.**
- *Tooltips* do gráfico de estrela com valores não formatados e em falta, assim como uma *tooltip* repetida. **Severidade 1.**

Help and documentation

Geral:

- Indicar que tabela pode ser ordenada. **Severidade 3.**

Rankings Equipas:

- Indicar que ao selecionar linhas, são apresentadas novas visualizações. **Severidade 1.**

Médias Escolares:

- Indicar a possibilidade de selecionar áreas do mapa. **Severidade 2.**
- Indicar que há mais informação acerca das escolas do distrito noutra página. **Severidade 2.**
- Botão para abrir nova janela com detalhes do distrito não indica o porquê de estar desativada até se selecionar um distrito do mapa. **Severidade 1.**
- Não há forma de indicar que, valores 0 (zero) significa que não foram efetuados exames daquela disciplina na escola em causa. **Severidade 1.**

Heurísticas de Zuk e Carpendale

Ensure visual variable has sufficient length

- Por vezes a linha da média do gráfico de barras para médias de disciplinas, fica sobre os seus valores na visualização. **Severidade 2.**

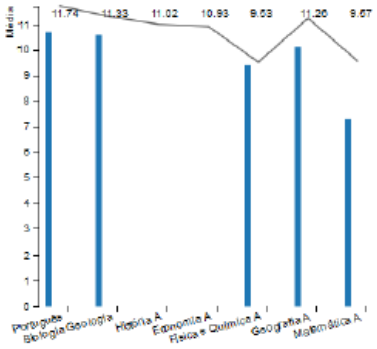

Figura 5 - Linha da Média Distrital por cima dos seus valores

Local contrast affects color & gray perception

- Contraste entre as cores das barras, para gráfico de performance de jogos, e do texto afeta a perceção do texto e as cores das barras entre si. **Severidade 3.**

Consider people with color blindness

- Apenas o *HeatMap* do mapa que apresenta a Diferença entre a Média Distrital e a Média Nacional tem esta preocupação. **Severidade 3.**

Preserve data to graphic dimensionality

- Gráficos de Rankings de Futebol não se encontram com a mesma formatação e nas mesmas posições relativas que os Gráficos de Rankings Escolares. **Severidade 1.**

Remove the extraneous (ink)

- *Zoom* do mapa pode ser considerado um problema de acordo com esta heurística, porque ao fazer zoom a um distrito, esconde a informação dos restantes. **Severidade 1.**

Consider Gestalt Laws

- Filtro das Tabelas não está bem aplicado, o texto introduzido tem que ser 100% igual ao valor armazenado nos dados para serem apresentados resultados. Não há funções de similaridade. **Severidade 3.**

Integrate text wherever relevant

- Indicar que valores superiores ao gráfico, para as médias das disciplinas, correspondem aos valores da Linha da Média Distrital. **Severidade 2.**
- Indicar que a linha, no gráfico de médias de disciplinas, representa a Média Distrital. **Severidade 3.**

Referências

[1] J. Nielsen, "Ten Usability Heuristics," *Commun. ACM*, vol. 3, no. 1990, pp. 1–2, 2005.

[2] T. Zuk, L. Schlesier, P. Neumann, M. S. Hancock, and S. Carpendale, "Heuristics for Information Visualization Evaluation," *Beliv'06*, pp. 1–6, 2006.

Anexo 2: Avaliação de Pedro Santos

João Pedro Santos
2015

Mestrado em Sistemas de Informação

Avaliação Heurística

Esta avaliação surgiu com a necessidade de realizar uma avaliação heurística à interface da plataforma disponível em http://sweet.ua.pt/jmadeira/Vis_Rankings/pages/index.html. Esta plataforma apresenta dados classificativos de elementos, sendo estes Clubes de Futebol e Escolas Secundárias de Portugal.

Esta plataforma foi desenvolvida por um colega de mestrado como tema de Dissertação para o mestrado de Sistemas de Informação, do departamento de Eletrónica, Tecnologias e informática.

Escala de Severidade dos Erros

Segundo Nielsen, a severidade de um problema de usabilidade de uma interface é calculado tendo em conta os seguintes fatores:

- A frequência com que o problema pode ocorrer ocorre.
- O impacto gerado pelo problema caso ocorra. E se o utilizador se consegue librar dele.
- A persistência do problema. Se ocorre demasiadas vezes.

Com isto é possível pontuar a severidade do problema de usabilidade, o que também permite priorizar a sua resolução. A escala que se apresenta de seguida foi a optada para o efeito:

- 0: Não concordo que seja um problema de usabilidade.
- 1: Apenas um problema cosmético, não há necessidade de ser corrigido, apenas se houver disponibilidade.
- 2: Problema de usabilidade menor. Baixa prioridade de resolução deste problema.
- 3: Problema de usabilidade grave. A sua resolução é importante e tem alta prioridade.
- 4: Problema de usabilidade crítico. Tem que ser resolvido de imediato.

Heurísticas utilizadas

Para esta avaliação foram utilizadas duas listas de heurísticas distintas, as heurísticas de Nielsen e as heurísticas de Zuk.

Lista dos 10 princípios Heurísticos de Nielsen

1. Visibility of system status
2. Match between system and the real world
3. User control and freedom

4. Consistency and standards
5. Error prevention
6. Recognition rather than recall.
7. Flexibility and efficiency of use - Accelerators - unseen by the novice user.
8. Aesthetic and minimalist design.
9. Help users recognize, diagnose, and recover from errors.
10. Help and documentation.

Lista dos princípios Heurísticos de Zuk e Carpendale

1. Ensure visual variable has sufficient length
2. Don't expect a reading order from color
3. Color perception varies with size of colored item
4. Local contrast affects color & gray perception
5. Consider people with color blindness
6. Preattentive benefits increase with field of view
7. Quantitative assessment requires position or size variation
8. Preserve data to graphic dimensionality
9. Put the most data in the least space
10. Remove the extraneous
11. Consider Gestalt Laws
12. Provide multiple levels of detail
13. Integrate text wherever relevant

Erros de Usabilidade

Heurísticas de Nielsen

1. Visibility of system status

A interação no gráfico apresenta os valores sobrepostos às linhas gerando confusão na perceção dos valores. **Detetado em Ranking Desportivo e Ranking Escolar**. Atribuída severidade 2.

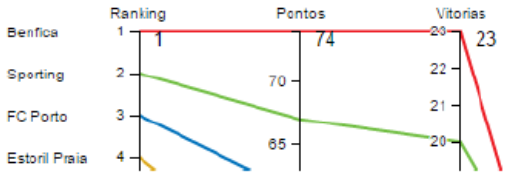

Figura 1 - Sobreposição de valores com linhas

2. Aesthetic and minimalist design

Informação no Ranking de escolas sobrepõe-se, gerando informação não percetível e com desorganização visual. **Detetado em Ranking Escolares**. Com severidade 4.

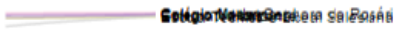

Figura 2 - Sobreposição de valores do gráfico no Ranking Escolar

3. Help users recognize, diagnose, and recover from errors

Não foi encontrado qualquer tipo de diagnóstico de erros. Severidade 2.

4. Help and documentation

Não existe documentação que descreva o funcionamento da plataforma. Severidade 4.

Heurísticos de Zuk e Carpendale

1. Consider people with color blindness

Amarelos e Verdes muitos semelhantes e que poderão gerar confusão. **Detetado em Rankings de Desporto e Médias Disciplinas**. Com severidade 3.

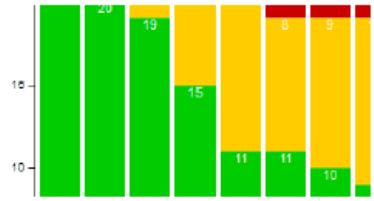

Figura 3 - Contraste de cores para daltónicos

2. Put the most data in the least space

Excesso de informação no gráfico de Rankings **em Ranking Escolares** rebenta a linha. Atribuída severidade 2.

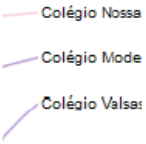

Figura 4 - Exemplo de falta de área de desenho para informação apresentada

3. Provide multiple levels of detail

Níveis de detalhe muito superficiais. Falta profundidade na plataforma. Severidade 2.