

Projeto 1 - Visualization as communication

Juliana Curty | 2018145670

Sara Costa | 2016234640

Introdução

Avizualização de informação tem como objetivo facilitar a organização e a compreensão de grandes volumes de dados, de modo a otimizar a comunicação desta informação através de uma experiência mais agradável e prática para o utilizador.

Para este Projeto optamos por demonstrar e apresentar a análise de dados, do Ranking de Felicidade do Mundo. Estes dados são úteis para perceber a base estrutural de um país nos vários sectores de dados recolhidos, e deste modo podemos organizar esses dados e criar uma ferramenta de visualização sobre este tema, dando a oportunidade de perceber as informações obtidas de fácil visualização.

Pretende-se então mostrar um gráfico, no formato de stacked bar, desta felicidade nos vários tópicos recolhidos, como a nível da sua Economia, Família, Liberdade, Saúde, Confiança, Generosidade e Distopia Residual, assim como uma comparação da evolução no período de dois anos (2015 e 2016) dentro de um mesmo país.



Objetivos

Avizualização de informação tem como objetivo facilitar a organização e a compreensão de grandes volumes de dados, de modo a otimizar a comunicação desta informação através de uma experiência mais agradável e prática para o utilizador.

Para este Projeto optamos por demonstrar e apresentar a análise de dados, do Ranking de Felicidade do Mundo. Estes dados são úteis para perceber a base estrutural de um país nos vários sectores de dados recolhidos, e deste modo podemos organizar esses dados e criar uma ferramenta de visualização sobre este tema, dando a oportunidade de perceber as informações obtidas de fácil visualização.

Pretende-se então mostrar um gráfico, no formato de stacked bar, desta felicidade nos vários tópicos recolhidos, como a nível da sua Economia, Família, Liberdade, Saúde, Confiança, Generosidade e Distopia Residual, assim como uma comparação da evolução no período de dois anos (2015 e 2016) dentro de um mesmo país.



Contextualização

O dataset que escolhemos para esse projeto consiste em um ficheiro 'csv' com um ranking de felicidade de 158 países ao redor do mundo. Os valores e rankings dessa tabela foram baseados em perguntas feitas aos cidadãos de cada um dos países. Essas pessoas tinham que atribuir um valor (0-10) a uma série de categorias, nomeadamente, produção econômica, suporte social, expectativa de vida, liberdade, ausência de corrupção e generosidade. A escala correspondia a 0 para um cenário distópico, ou seja, o pior cenário possível para determinada categoria e 10 para um cenário utópico.

Por fim, tinha também uma categoria chamada distopia, que corresponde a um país imaginário que reúne o menor dos valores de cada uma das categorias, ou seja, ao considerar os 158 países, pega-se o valor mais baixo de cada categoria e atribui-se a esse país imaginário. Esse país não tem impacto no valor final da felicidade, mas pode explicar o porquê de algum país estar em uma posição mais elevada que outros. As amostras que nos foram disponibilizadas correspondem aos anos de 2015 e 2016, no entanto, a pesquisa reúne informações de 2013-2017.

Como não tínhamos a informação de como foi obtido cada um dos atributos da tabela e os mesmos não estavam no intervalo de 0-10, assume-se aqui que houve algum tipo de normalização. Portanto, para não perdermos informação importante, não fizemos mais nenhum tipo de manipulação de dados. A única alteração que fizemos no dataset original foi adicionar uma coluna extra para cada atributo com aquele valor rankeado, ou seja, o país com o maior valor de felicidade ficou rankeado em primeiro nesta categoria. No entanto, essas colunas adicionais não foram utilizadas no nosso projeto.

Em seguida, focamos em como iríamos explorar a parte de visualização dos dados que tínhamos adquirido, quais os detalhes mais interessantes e importantes que o utilizador iria querer ver. Concluímos que para nós esses dados seriam uma comparação dentro de um país, em anos diferentes dentro das variáveis que formam o Ranking de Felicidade.

Table 2.1: Regressions to Explain Average Happiness across Countries (Pooled OLS)

Independent Variable	Dependent Variable			
	Cantril Ladder (0-10)	Positive Affect (0-1)	Negative Affect (0-1)	Cantril Ladder (0-10)
Log GDP per capita	0.31 (0.066)***	-0.009 (0.01)	0.008 (0.008)	0.324 (0.065)***
Social support	2.362 (0.363)***	0.247 (0.048)***	-0.336 (0.052)***	2.011 (0.389)***
Healthy life expectancy at birth	0.036 (0.01)***	0.001 (0.001)	0.002 (0.001)	0.033 (0.009)***
Freedom to make life choices	1.199 (0.298)***	0.367 (0.041)***	-0.084 (0.04)**	0.522 (0.287)*
Generosity	0.661 (0.275)**	0.135 (0.03)***	0.024 (0.028)	0.39 (0.273)
Perceptions of corruption	-0.646 (0.297)**	0.02 (0.027)	0.097 (0.024)***	-0.720 (0.294)**
Positive affect				1.944 (0.355)***
Negative affect				0.379 (0.425)
Year fixed effects	Included	Included	Included	Included
Number of countries	156	156	156	156
Number of obs.	1627	1624	1626	1623
Adjusted R-squared	0.751	0.475	0.3	0.768

Notes: This is a pooled OLS regression for a tapered panel explaining annual national average Cantril ladder responses from all available surveys from 2005 to 2019. See Technical Box I for detailed information about each of the predictors. Coefficients are reported with robust standard errors clustered by country in parentheses. ***, **, and * indicate significance at the 1, 5 and 10 percent levels respectively.

- Tabela de explicativa dos valores do Ranking -

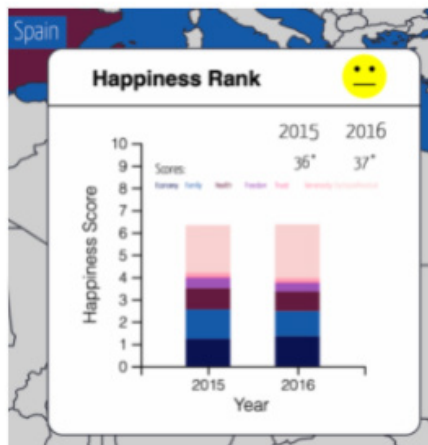
Contextualização

Em seguida, focamos em como iríamos explorar a parte de visualização dos dados que tínhamos adquirido, quais os detalhes mais interessantes e importantes que o utilizador iria querer ver. Concluímos que para nós esses dados seriam uma comparação dentro de um país, em anos diferentes dentro das variáveis que formam o Ranking de Felicidade.

Optamos então por mostrar uma comparação de anos dentro de cada país, através de um gráfico de barras onde identifica cada categoria com cores, mostrando a evolução e diferença nos anos de 2015 e 2016, indo buscar cada valor aos dados fornecidos, estando presente o valor numeral total da pontuação do país naqueles anos.

Ao gráfico juntamos uma carinha para identificar o quão satisfatório é aquele país em geral, tendo 3 caras identificativas, uma cara feliz, uma cara triste e outra neutra, uma maneira diferente de abordar este tema, sugerido pelo professor, chamado de “Chernoff Faces”.

Esta filtragem acontece na `<function functionName(selected)>`, no caso das caras, vai buscar o valor do primeiro ano e, se o valor for menor que 30 irá criar uma cara feliz, se o valor se encontrar entre 30 e 70 uma cara neutra e maior que 70 uma cara triste.



- Gráfico de barras -



- Caras de acordo com a Felicidade de cada País -

Para o gráfico das caras

```
<if (parseInt(HappinessRank[0]) < 30) {  
  happyFace(box, 200, 25, 15);  
}  
else if (parseInt(HappinessRank[0]) < 70) {  
  neutralFace(box, 200, 25, 15);  
}  
else {  
  sadFace(box, 200, 25, 15);  
}>
```

Para o Gráfico de Barras

```
<data = data.map(function (d) {  
  return { Time: parseTime(d.Time).getUTCFullYear(), Economy: +d.Economy,  
    Family: +d.Family, Health: +d.Health, Freedom: +d.Freedom, Trust: +d.Trust,  
    Generosity: +d.Generosity, DystopiaResidual: +d.DystopiaResidual };  
});  
stack = d3.stack().keys(labels)  
stack.value(function (d, key) {  
  return d[key];  
});  
dataset = stack(data);  
dataset.forEach(level => {  
  var cont = 0;  
  level.forEach(year => {  
    if (cont == 0) {  
      year.push('2015');  
      cont = 1;  
    }  
    else if (cont == 1) {  
      year.push('2016');  
      cont = 1;  
    }  
  })  
})  
console.log(dataset);>
```

Conceptualização

Funções

No código começamos pela inicialização e definição das variáveis usadas, definindo a projeção do mapa, do seu caminho, do grupo de países e o zoom aplicado, assim como o seu comportamento. Usando um rácio de 12:5 para aumentar o mapa, cabendo no ecrã da forma desejada.

Inicializamos também algumas funções para calcular os limites e características do zoom e para mostrar o mapa, assim como para criar a caixa onde vai estar a informação exposta, tendo em consideração as suas áreas, onde vai aparecer e ter a certeza que esta aparece sempre como um todo e com os devidos espaçamentos. Também as funções que vão dar origem à criação de cada uma das caras são desenvolvidas neste seguimento.

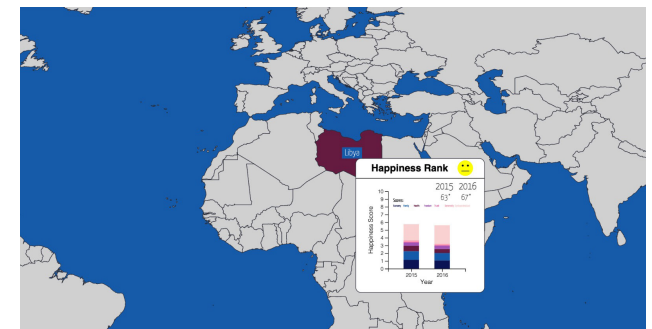
Em seguida temos a função para o gráfico de barras, no qual se estabelece as datas e os labels que vão estar representados no gráfico, com uma cor. Criando uma coluna para cada um dos anos (2015 e 2016), de forma a ir buscar as informações de cada label para o respectivo ano.

Numa próxima etapa encontramos a criação do gráfico, em valores de escala e identificação de dados. E onde vamos buscar os valores recolhidos e fornecidos, colocando-os visualmente, com variáveis de texto, estilo e localização.

Concluindo com a função de nome do país selecionado, em que inserimos o array de objetos para o Ranking de Felicidade, sendo aqui que filtramos os valores para colocar no gráfico de barras segundo cada uma das variáveis que dão origem à pontuação final do Ranking e nas caras representadas em cada país usando o valor total.



- Imagem inicial ao abrir a página -



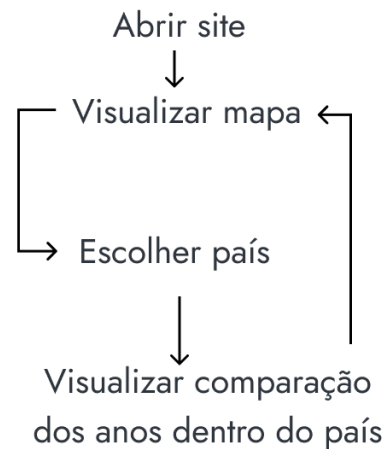
- Imagem ao escolher um país -

Conceptualização

Interações

No decorrer desta visualização de informação, existem algumas interações, na página Inicial, assim que entramos na visualização, temos o mapa mundo, no qual ao passar o rato por cima, aparece o nome que o representa, caso contrário este não é visível, dado o caso de clicarmos num país à escolha, qualquer que seja este país, dá-se um zoom, que cria um caminho e aproxima-nos desse país, aparecendo o label com o nome, e a caixa de informação com os dados apresentados.

Para cada um dos países recorremos ao push e return dos valores a utilizar para o gráfico, a cara desenhada, o nome e identificação de cada país. Na imagem seguinte podemos ver um esquema simples de interação da parte do utilizador nesta ferramenta.



Descrição do Processo

O processo de trabalho para chegar ao nosso objetivo começou pela escolha do tema, seguido de uma sobre o mesmo de modo a perceber as suas características, tendo em consideração as variáveis e como se definiam os valores fornecidos, e como iríamos trabalhar, ou manipular estes dados.

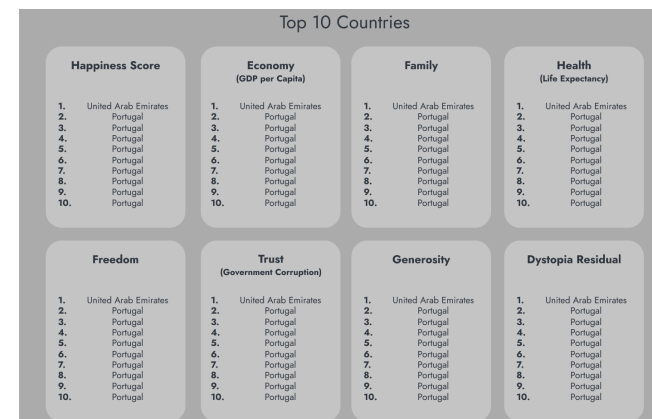
Entramos na parte visual, onde optamos por escolher interações simples, mas que desse uma sensação de interatividade e proximidade perante a informação que se iria obter, para além de dar um aspeto mais interessante e menos estático a uma pesquisa de informação que muitas vezes pode acabar por se tornar aborrecida.

No final achamos que alcançamos os nossos objetivos propostos inicialmente, porém existem sempre melhorias que se poderiam fazer uma proposta futura, como poder visualizar um ranking discriminado, para cada variável, pois um país pode ter uma boa classificação geral, mas cada um de nós valoriza mais certos aspectos no seu dia-a-dia, como exemplo, uma maior apreciação de um bom sistema de saúde, podendo assim ver quais os países que se encontram melhor classificados nesta categoria, representado na imagem ao lado.

De notar as referências utilizadas para a conclusão deste projeto:

-> Helliwell, John F., Richard Layard, Jeffrey Sachs, and Jan-Emmanuel De Neve, eds. 2020. World Happiness Report 2020. New York: Sustainable Development Solutions Network, para a pesquisa

-> <https://codepen.io/ssz360/pen/jOPMwme>, para o código



Obrigada!