

# Trabalho Prático II

Visualização de Dados - 2018/02  
Departamento de Ciência da Computação  
Universidade Federal de Minas Gerais

Jonatas Cavalcante

## Introdução

O objetivo deste trabalho é colocar em prática os conhecimentos aprendidos na disciplina para propor e desenvolver visualizações interativas de dados. Espera-se que sejam propostas análises relevantes para um determinado conjunto de dados e desenvolver visualizações e técnicas de interação analítica que proporcionem estas análises.

## Dados

O conjunto de dados selecionados para esse trabalho contém informações a respeito do Campeonato Brasileiro de Futebol de 2018. Essa base foi escolhida pois esse campeonato possui grande relevância e enfoque no campo esportivo no Brasil. Esses dados foram disponibilizados pela monitora da disciplina no formato JSON e CSV.

Nesse conjunto de dados que vamos utilizar estão contidas informações sobre cada uma das equipes que compõem a série A do campeonato brasileiro, a saber:

- *position*: classificação do time no momento da extração dos dados
- *team*: nome do clube
- *P*: pontos conquistados pelo time
- *J*: quantidade de jogos que o time participou
- *V*: quantidade de vitórias
- *E*: quantidade de empates
- *D*: quantidade de derrotas
- *GP*: quantidade de gols feitos pela equipe
- *GC*: quantidade de gol sofridos pela equipe
- *SG*: saldo de gols da equipe (gols feitos - gols sofridos)
- *percent*: porcentagem de aproveitamento do time no campeonato

## Tarefa

A tarefa consiste em desenvolver uma ou mais análises utilizando a base de dados escolhida. É necessário levantar duas perguntas que podem ser respondidas através da visualização de dados e implementar visualizações interativas para responder essas questões. Após isso, deve-se realizar uma análise dos resultados e aplicar conceitos e boas práticas aprendidos na disciplina nas implementações realizadas.

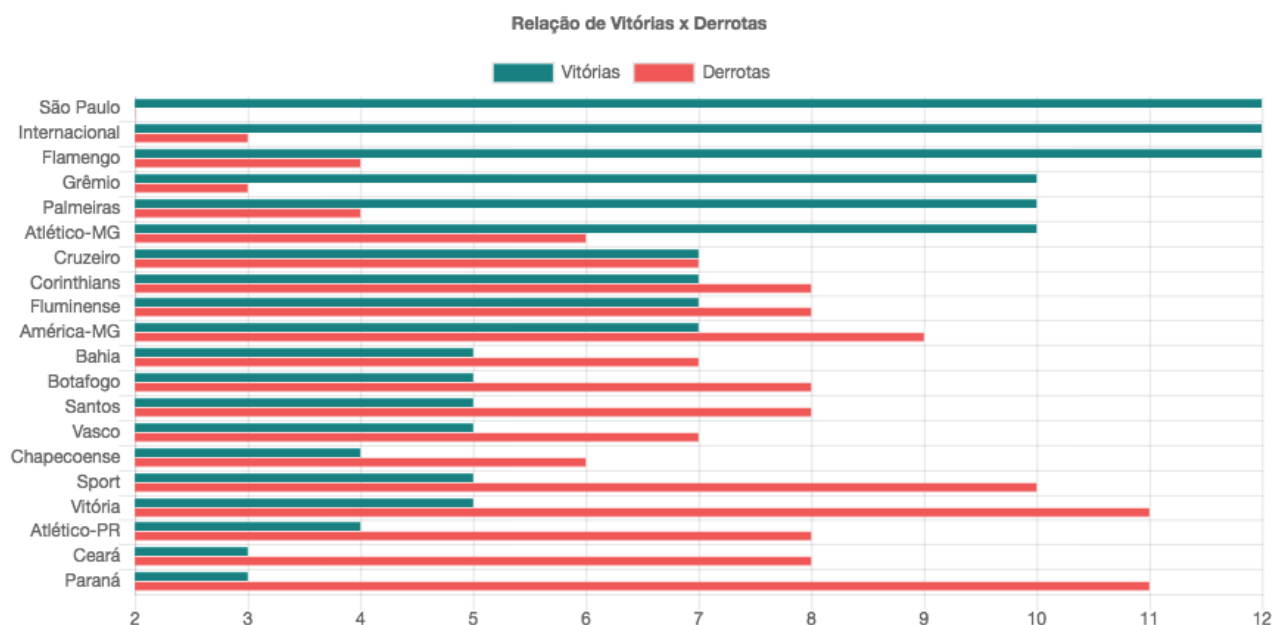
## Questões Levantadas e Análise dos Dados

Com base nos dados coletados, temos informações sobre os 20 times que compõem a série A do Brasileirão. Tradicionalmente, os dados do campeonato são expostos em uma tabela de classificação, onde cada um dos itens presentes na base são colunas dessa tabela. É uma forma bem sucinta e prática de acomodar essa quantidade de dados para visualização rápida.

Porém, não é possível ter uma escala boa de comparação entre alguns dos dados, sendo que a visualização por barras horizontais (escolhida para esse trabalho) permite uma compreensão melhor por meio da comparação entre os times. Apesar dessa visualização ser interessante, não é prático trazer todos os dados para esse gráfico, pois torna-se um grande amontoado de dados, tornando a visualização complexa e confusa.

Assim sendo, foram elaborados três visualizações que buscam elucidar alguns dados interessantes a respeito dos times do campeonato brasileiro. São elas:

### I. Relação de vitórias e derrotas de cada um dos times de acordo com a sua posição no campeonato

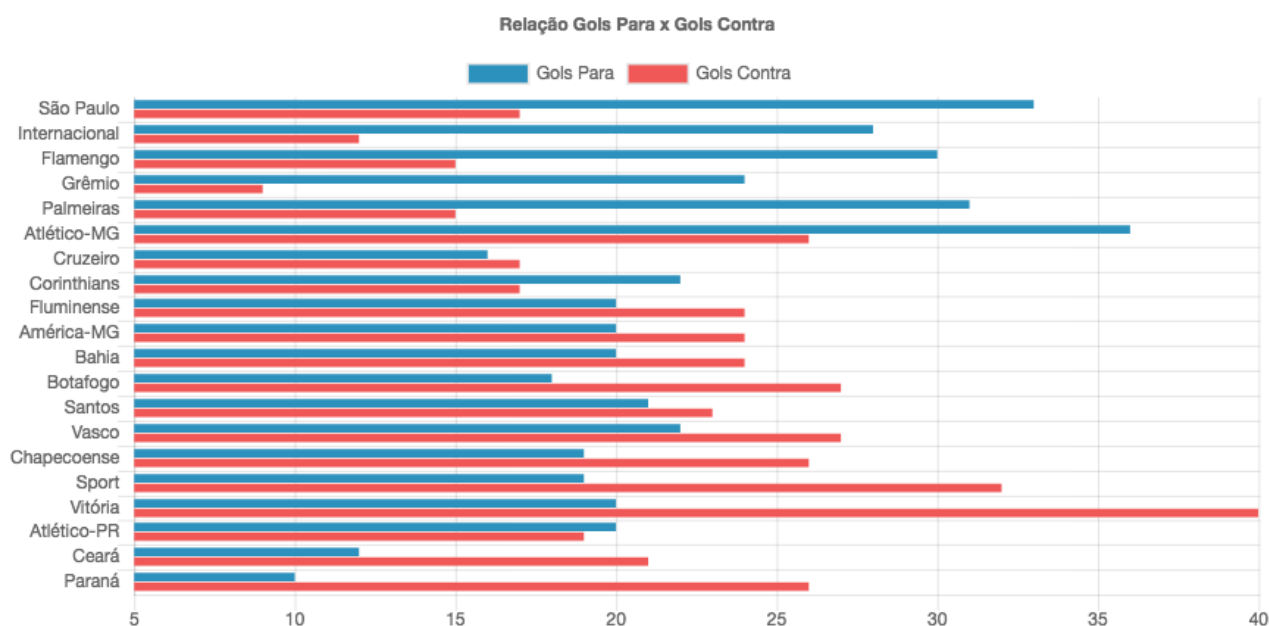


Nessa visualização, cada time possui duas barras: uma das vitórias e uma para as derrotas, sendo que a listagem segue a ordem da classificação da tabela. É possível visualizar e compreender a disparidade das vitórias e derrotas dos variados times do campeonato. Pode-se analisar o quanto que os primeiros colocados possuem uma grande quantidade de vitórias e poucas derrotas, enquanto os times mais mal classificados apresentam a situação reversa: muitas derrotas e poucas vitórias.

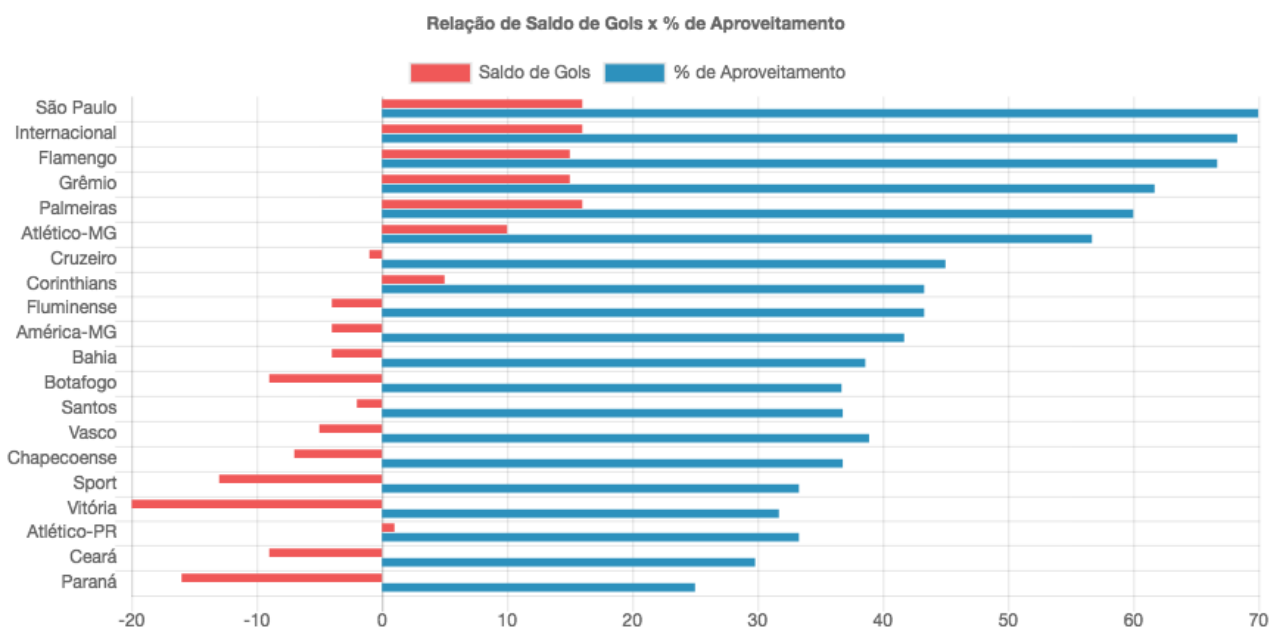
### II. Relação de Gols Para x Gols Contra

Seguindo a mesma ideia da visualização anterior, os times são ordenados segundo a classificação com a tabela. O interessante dessa visualização é que é possível buscar compreender como funciona a questão do ataque e defesa de cada um dos times. É possível verificar que os mais bem colocados sofrem poucos gols, mas não necessariamente fazem muitos

gols. Pode-se inferir que se trata mais de uma questão de equilíbrio, ou seja, ter o melhor ataque da competição ou a melhor defesa não significa a primeira posição do campeonato, é necessário um balanço entre essas duas características, uma vez que o campeonato é longo e a constância é uma aspecto muito relevante.



### III. Relação de Saldo de Gols x % de Aproveitamento



Essa visualização busca estabelecer uma relação entre o saldo de gols (gols feitos - gols sofridos) e o aproveitamento da equipe no campeonato. O que se pode entender desses dados é que os times que possuem os melhores aproveitamentos e, portanto, as melhores posições na tabela do campeonato possuem um saldo positivo de gols e que, até o sexto colocado, é acima de 10 gols. Podemos inferir que, como apontado na visualização anterior, esse equilíbrio de fazer

gols e sofrer poucos resulta numa boa colocação no campeonato, reforçando a ideia de que a constância tem papel fundamental para o sucesso da equipe no Brasileirão.

## Ferramentas

Nesse trabalho, foi utilizada a ferramenta *Chart.js*[1] para desenvolver as visualizações acima propostas. O código estrutura-se com um arquivo *index.html* que utiliza CSS para estabelecer os estilos da página e *javascript* para realizar a leitura da base dos dados, assim como filtrar cada um deles para os gráficos exibidos. Afim de permitir uma boa compreensão dos dados, a página estrutura-se com um menu lateral esquerdo com a tabela do campeonato brasileiro e, abaixo dele, uma listagem com links para as três visualizações.

O usuário assim, clica na visualização desejada e o gráfico será exibido na aba direita da página. Ao clicar em outra visualização, essa substitui a corrente. Com a tabela fixa, simplifica-se o entendimento dos dados do campeonato, caso o usuário deseje fazer outras comparações de outros dados não presentes no gráfico.

Foi atentado a questão das cores, a colocação da legenda e do título, visando aplicar os conceitos aprendidos na disciplina, buscando trazer uma visualização simples e prática. Como recurso do *Chart.js*, o usuário, ao clicar na legenda, omite a aparição daquele dado. Assim sendo, como as visualizações tratam de dois dados contrastando, caso o usuário queira apenas ver um deles, basta clicar no elemento não desejado na legenda.

Além disso, temos o recurso de, quando o cursor do mouse passa encima do elemento do gráfico, um *tooltip* aparece com o valor de cada um dos dados pra aquele time, afim de elucidar os dados, caso não fique tão claro na escala das barras horizontais.

## Estrutura do Projeto e Acesso às Visualizações

O projeto estrutura-se da seguinte forma: o arquivo *index.html* possui toda a codificação necessária para a visualização, englobando o CSS, o *javascript* e o HTML da página. Além dele, na pasta principal do projeto há a especificação desse trabalho e essa documentação.

Para acessar as visualizações, basta acessar a referência [2] ou abrir o arquivo *index.html* disponível no pacote, utilizando o navegador da sua preferência. O código também está disponível no repositório[3] do *GitHub*.

## Conclusão

Através do desenvolvimento deste trabalho, foi possível ter um contato mais próximo com ferramentas/bibliotecas para o desenvolvimento de visualizações de dados bem como a exploração e análise visual. Apesar de algumas dificuldades encontradas no desenvolvimento, foi possível colocar em prática conceitos e aprendizados até então teóricos da disciplina, além de aprimorar os conhecimentos adquiridos do último trabalho prático, além do projeto final.

## Referências

[1] <http://www.chartjs.org/>

[2] <https://jonatascavalcante.github.io/brasileirao/>

[3] <https://github.com/jonatascavalcante/brasileirao>