

# Gastos Diretos do Governo Federal

Um trabalho prático de Introdução à Ciência dos Dados - UFMG

Bruno Maletta Monteiro	2017015150
Francisco Galuppo Azevedo	2017014960
Thais Soares Lamas	2017014910

# Nosso Trabalho

## Os dados

O nosso trabalho parte da ideia de estudar os gastos diretos do governo federal, isto é, as obras e serviços que o governo contrata, não engloba os gastos com pessoal. Optamos pelo período de 2011 a 2017.

## Objetivos

Gostaríamos de saber como esses gastos são divididos, como eles variaram com o tempo, o que eles dizem sobre os ministérios etc.

## Fonte

Dados disponibilizados pelos governos passados através do portal da transparência.

Usamos o projeto [brasil.io](https://brasil.io), que busca facilitar o acesso da população aos dados públicos.

# Nossas Perguntas

## Pergunta 1

### **Gastos uniformes?**

Os gastos são uniformemente distribuídos ou há meses de maior gasto?

## Pergunta 2

### **Gastos políticos?**

E quanto aos anos eleitorais, os gastos aumentam ou são comparáveis com os demais?

## Pergunta 3

### **Classificação?**

Conseguimos classificar os ministérios apenas pela quantia e data dos gastos?

## Pergunta 4

### **Regressão?**

Quais são as tendências desses gastos? De alta? De baixa?

# Caracterização

Um pouco dos dados

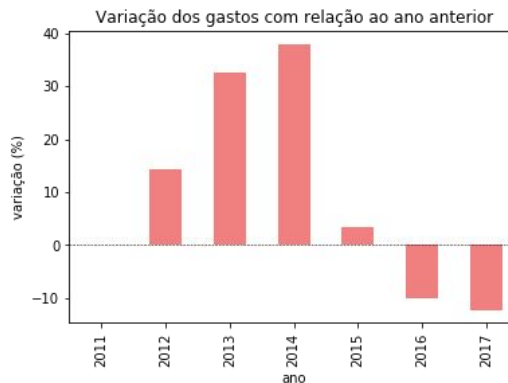
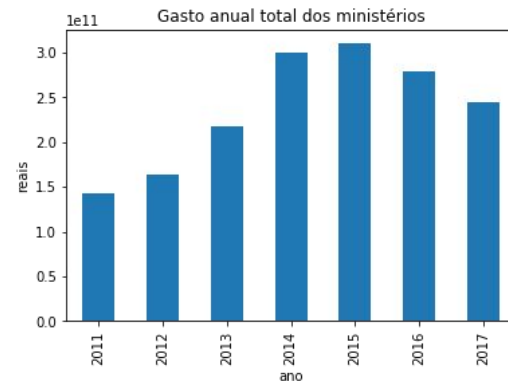
Nesta seção, mostraremos algumas informações e gráficos que conseguimos extrair dos dados.

---

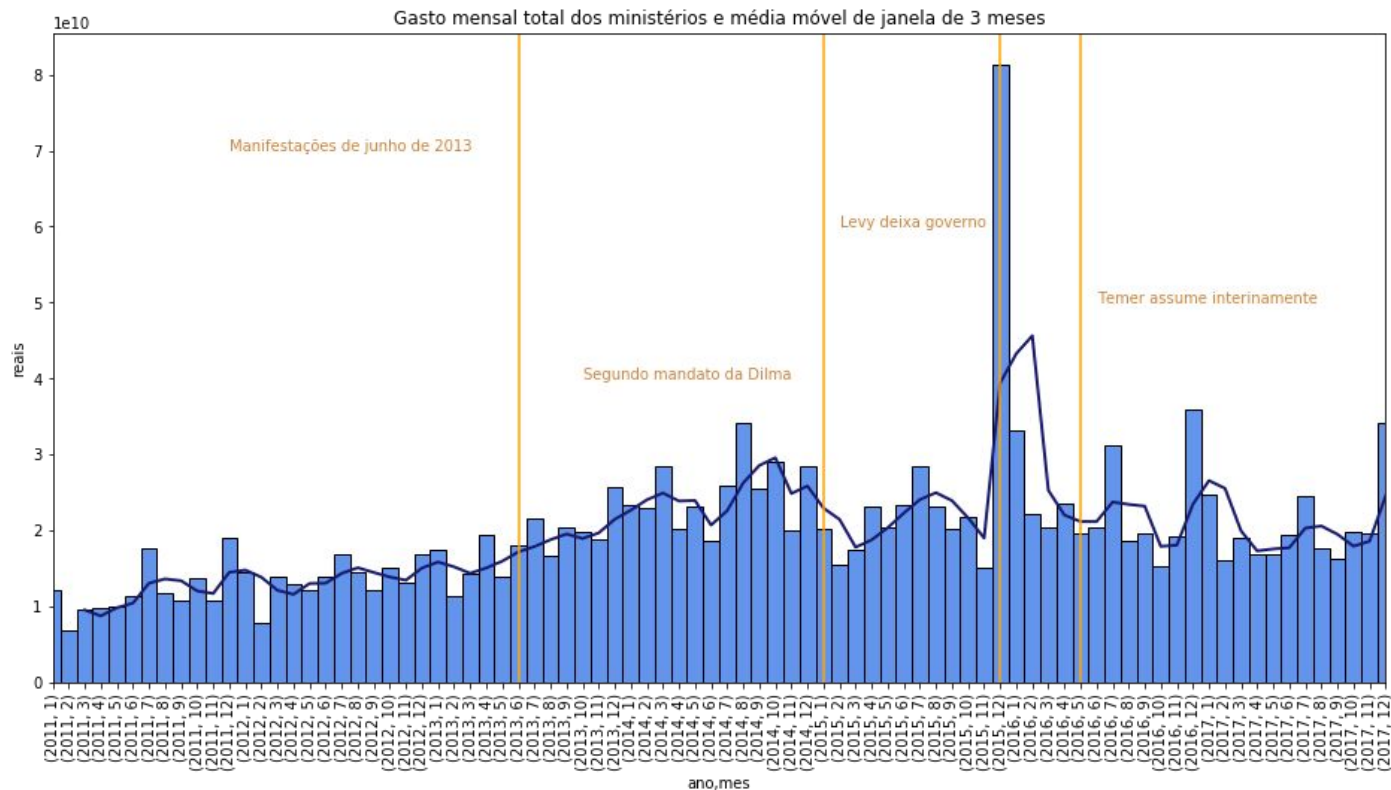
# Gastos Totais Anuais

Os gráficos ao lado mostram como foi o total de gastos ao longo do período

- os gastos mais do que dobraram entre 2011 e 2015
- a variação foi positiva e crescente de 2012 e 2014
- O ano de 2015 parece mostrar uma desaceleração
- 2016 e 2017 já estão em uma fase de queda dos gastos diretos.



# Gastos Totais Mensais



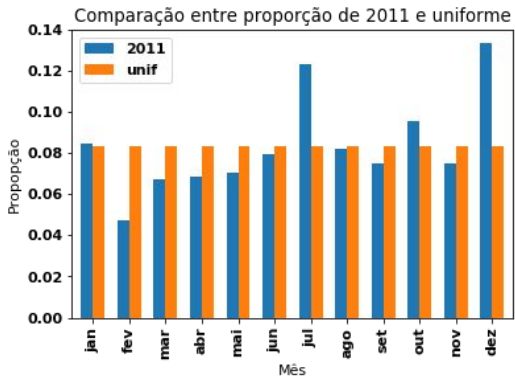
# Testes de Hipótese

Respondendo algumas perguntas

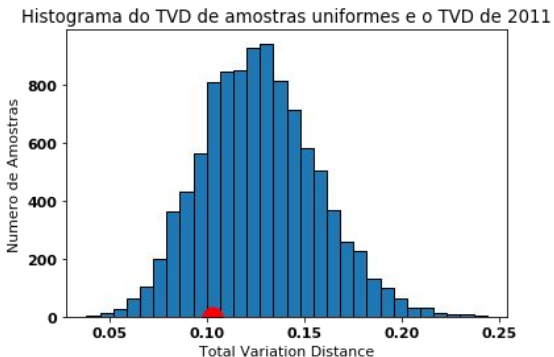
Nesta seção, investigaremos as duas primeiras perguntas, se os gastos são uniformes ao longo do ano e se existe alguma anomalia nos anos eleitorais.

---

# Gastos Sazonais?



Usando simulações para calcular a **Total Variance Distance**, estudamos os gastos totais mensais de cada ano para verificar se conseguimos descartar que eles vem de uma distribuição uniforme.



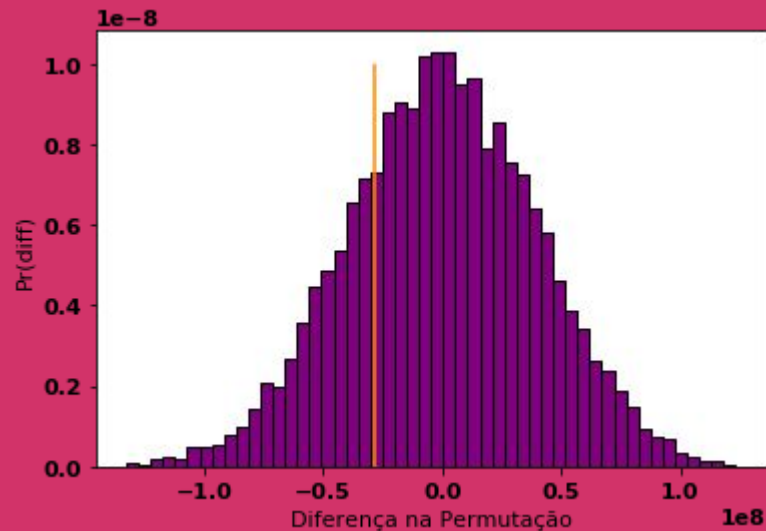
Com exceção de um dos anos do período (2015), não conseguimos descartar que os gastos são de fato uniformes.



# Gastos Políticos?

Em seguida, avaliamos se os gastos da Presidência da República (órgão cuja secretaria de Comunicação está sujeita) variam entre os anos eleitorais e os não eleitorais usando um Teste de Permutação.

O teste não descarta que as diferenças venham apenas do acaso.



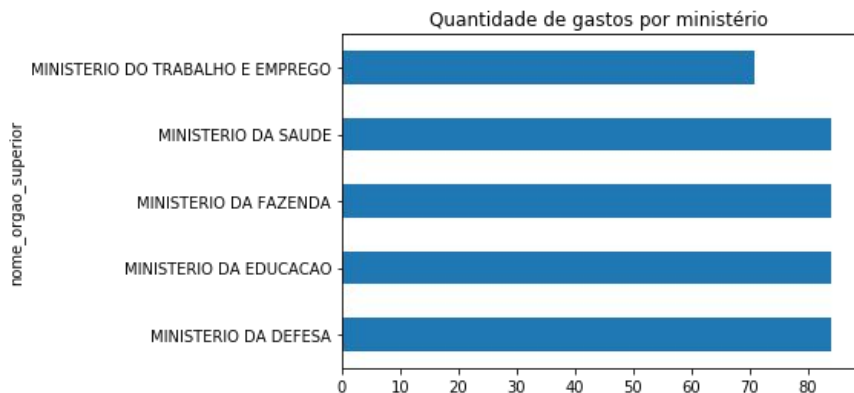
# Classificação

Classificamos os gastos por  
ministério

Nesta seção, vamos tentar  
classificar os gastos por ministério,  
com base em quando ele foi  
efetuado e quanto foi gasto.

---

# Tratamento dos dados

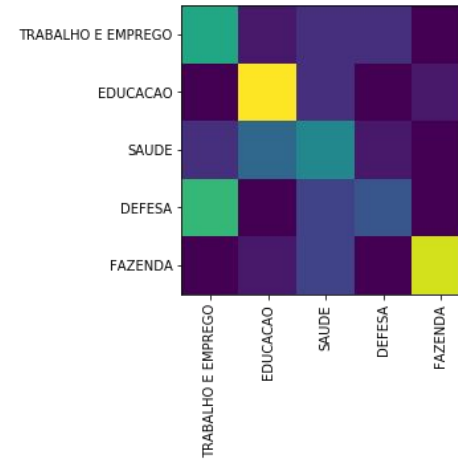


Como existem ministérios demais, reduzimos os dados de forma a conter apenas 5 dos mais importantes ministérios:

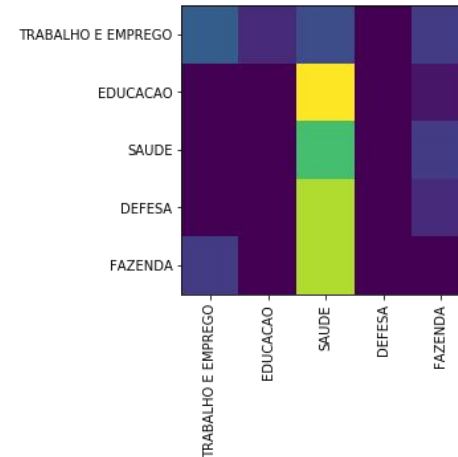
Da Educação, Fazenda, Saúde, Defesa e Trabalho e Emprego.

Vimos que o kNN teve um desempenho muito melhor que a regressão. Pudemos verificar também suas respectivas matrizes de confusão.

Algoritmo	Acurácia
kNN	0.70
Regressão	0.23



Matriz de confusão do kNN



Matriz de confusão da regressão

# Regressão

Verificando a mudança dos gastos ao longo do tempo

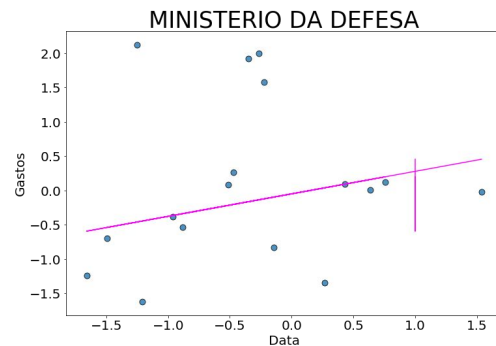
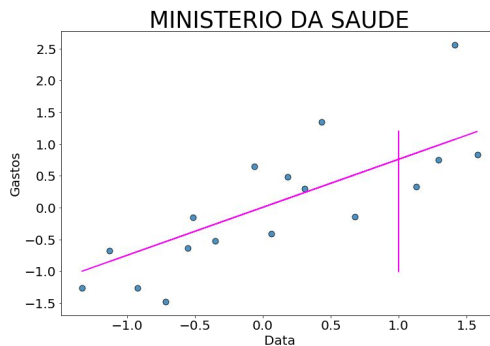
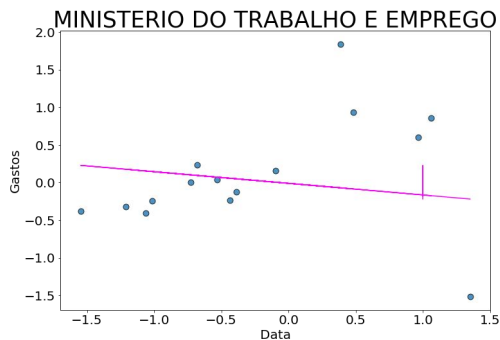
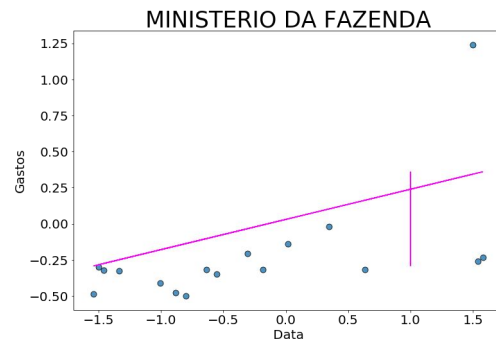
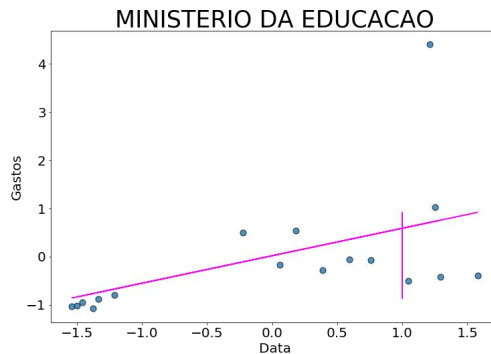
Nesta seção, vamos verificar se os gastos de alguns ministérios e do governo vem aumentando ao longo do tempo.

---

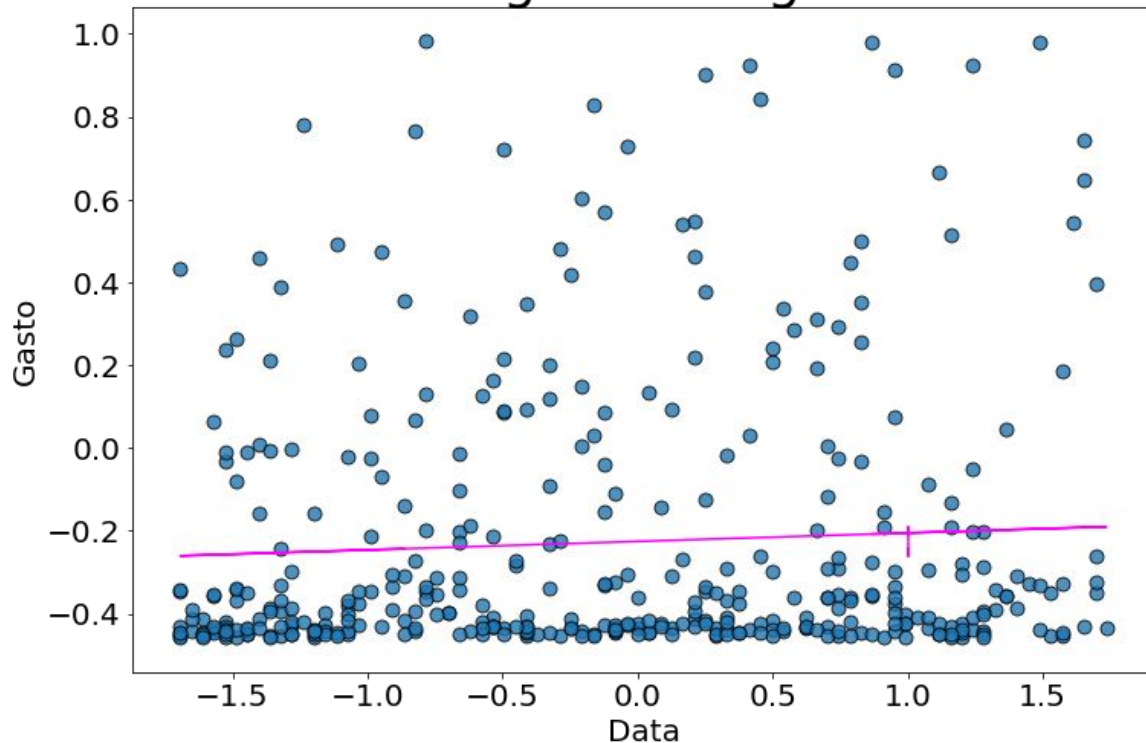
# Por ministérios

Como existem ministérios demais, analisamos os dados dos 5 ministérios mais importantes:

Da Educação, Fazenda, Saúde, Defesa e Trabalho e Emprego.



## Gastos gerais do governo



Como alguns gastos têm valores muito dispersos, pegamos uma região de maior concentração de pontos para analisar.

E é perceptível um leve aumento nos gastos ao longo do tempo.

# Conclusão



# Conclusão

Por fim, façamos uma análise de nossas perguntas:

1. Não conseguimos mostrar que os gastos não são uniformemente distribuídos ao longo do ano.
  2. Não conseguimos mostrar que os anos eleitorais possuem gastos diferentes dos demais anos.
  3. Conseguimos classificar razoavelmente bem os ministérios apenas pelos gastos e datas.
  4. As regressões capturaram parte do comportamento dos gastos.
- 