

# Venda de carros

Prof. MSc. Edgar Luiz de Lima

21/03/2022

## Preparação dos dados para as análises exploratórias

Definindo diretório e carregando os pacotes necessários.

```
setwd("C:\\Users\\Edgar\\OneDrive - unb.br\\Venda_de_carros\\Venda_de_Carros")
library(openxlsx)
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --

## v ggplot2 3.3.5      v purrr 0.3.4
## v tibble 3.1.4       v dplyr 1.0.7
## v tidyr 1.1.3        v stringr 1.4.0
## v readr 2.0.1        v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

Carregando e visualizando os dados.

```
dados<- read.xlsx("DadosVendaCarros.xlsx")
View(dados)
str(dados)

## 'data.frame': 457 obs. of 12 variables:
## $ DataNotaFiscal: chr "04-10-2016" "01-01-2016" "02-02-2016" "03-03-2016" ...
## $ Fabricante : chr "Rolls Royce" "Aston Martin" "Rolls Royce" "Rolls Royce" ...
## $ Estado : chr "São Paulo" "São Paulo" "São Paulo" "São Paulo" ...
## $ ValorVenda : num 95000 120000 88000 89000 92000 ...
## $ ValorCusto : num 50000 75000 75000 88000 62000 ...
## $ TotalDesconto : num 500 0 750 0 0 0 750 2500 0 0 ...
## $ CustoEntrega : num 750 1500 1000 1000 1500 1000 500 1000 1000 1000 ...
## $ CustoMaoDeObra: num 750 550 550 550 550 550 550 550 750 750 ...
## $ NomeCliente : chr "Aldo Motors" "Honest John" "Bright Orange" "Honest John" ...
## $ Modelo : chr "Camargue" "DBS" "Prata Ghost" "Prata Ghost" ...
## $ Cor : chr "Vermelho" "Azul" "Verde" "Azul" ...
## $ Ano : num 2016 2016 2016 2016 2016 ...
```

Construindo o conjunto de dados para avaliar o total de vendas por ano. Irei preparar também um conjunto de dados do total de vendas por ano para a marca Jaguar.

```
# Criando a matriz com o número de vendas por ano para todas as marcas.
TVanos<- data.frame(table(dados$Ano))
names(TVanos)<- c("Ano", "Total")
```

```
# Criando a matriz do número de vendas por ano do carros da marca Jaguar.
TVAjaguar<- dados%>%filter(Fabricante=="Jaguar")%>%
  select(Ano)%>%table()%>%data.frame()
names(TVAjaguar)<- c("Ano","Total")
```

```
#Juntando em um mesmo dataframe
vanuais<-rbind.data.frame(TVanos,TVAjaguar)
vanuais<- rep(c("Total","Jaguar"), each= 4)%>%data.frame(vanuais)
names(vanuais)<- c("Escala","Ano","Total")
```

Checando a estrutura dos dados

```
str(vanuais)

## 'data.frame': 8 obs. of 3 variables:
## $ Escala: chr "Total" "Total" "Total" "Total" ...
## $ Ano : Factor w/ 4 levels "2016","2017",...: 1 2 3 4 1 2 3 4
## $ Total : int 39 115 77 226 12 31 24 62
```

Transformando Escala em fator e ano em numérico.

```
vanuais$Escala<- as.factor(vanuais$Escala)
vanuais$Ano<- as.numeric(as.character(vanuais$Ano))
str(vanuais)
```

```
## 'data.frame': 8 obs. of 3 variables:
## $ Escala: Factor w/ 2 levels "Jaguar","Total": 2 2 2 2 1 1 1 1
## $ Ano : num 2016 2017 2018 2019 2016 ...
## $ Total : int 39 115 77 226 12 31 24 62
```

Gerando a planilha de custo de entrega de veículos por fabricante.

```
entrega<- dados%>% select(Fabricante,ValorCusto)%>% group_by(Fabricante)%>%
  summarise(valor_entrega= mean(ValorCusto))
```

## Análise gráfica dos dados

Análise gráfica do padrão temporal de vendas.

```
ggplot(vanuais, aes(Ano, Total, colour= Escala))+
  geom_line(size= 1)+
  geom_point(size= 4)+
  scale_y_continuous(limits = c(10,240), breaks = seq(10,240,20))+
  annotate("text", x = vanuais$Ano, y =vanuais$Total+11,
    label = vanuais$Total, size= 5)+
  ylab("Total de vendas")+
  theme_bw()+
  theme(panel.grid= element_blank(), axis.title = element_text(size= 16),
    axis.text = element_text(size= 14, color = "black"), legend.title= element_blank(),
    legend.text = element_text(size = 14))
```

