

Venda de carros

Prof. MSc. Edgar Luiz de Lima

21/03/2022

Preparação dos dados para as análises exploratórias

Carregando os pacotes necessários.

```
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()

library(ggpubr)
```

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 ...
```

Checando se há dados faltantes.

```
summary(dados)

## DataNotaFiscal Fabricante Estado ValorVenda
## Length:457 Length:457 Length:457 Min. : 22500
```

```
## Class :character   Class :character   Class :character   1st Qu.: 39500
## Mode  :character   Mode  :character   Mode  :character   Median : 44000
##                                     Mean  : 69361
##                                     3rd Qu.:110000
##                                     Max.   :181250
##      ValorCusto      TotalDesconto      CustoEntrega      CustoMaoDeObra
## Min.   : 4500      Min.   : 0      Min.   : -75.0      Min.   : 147.0
## 1st Qu.: 22500      1st Qu.: 150      1st Qu.: 50.0      1st Qu.: 325.0
## Median : 36125      Median : 200      Median : 425.0      Median : 570.0
## Mean   : 44661      Mean   : 493      Mean   : 525.1      Mean   : 679.4
## 3rd Qu.: 67000      3rd Qu.: 750      3rd Qu.: 875.0      3rd Qu.: 987.0
## Max.   :160000      Max.   :5050      Max.   :1750.0      Max.   :1250.0
## NomeCliente      Modelo      Cor      Ano
## Length:457      Length:457      Length:457      Min.   :2016
## Class :character   Class :character   Class :character   1st Qu.:2017
## Mode  :character   Mode  :character   Mode  :character   Median :2018
##                                     Mean   :2018
##                                     3rd Qu.:2019
##                                     Max.   :2019
```

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 o conjunto de dados do custo de entrega de veículos por fabricante.

```
entrega<- dados%>%filter(Ano==2019)%>% select(Fabricante,ValorCusto)%>% group_by(Fabricante)%>%
  summarise(valor_entrega= mean(ValorCusto))
entrega$Fabricante<- as.factor(entrega$Fabricante)
View(entrega)
```

Gerando o conjunto de dados do custo de mão de obra por estado.

```
mobra<- dados%>%filter(Ano==2019)%>%select(Estado,CustoMaoDeObra)%>% group_by(Estado)%>%
  summarise(Media_MaoDeObra= mean(CustoMaoDeObra))
mobra[,1]<-c("BA","ES","MG","PR","RJ","SP" )
View(mobra)
```

Conjunto de dados para avaliar o lucro de cada marca, iremos analisar apenas com dados de 2019.

```
dluc<- dados%>%filter(Ano==2019)%>%
  select(Fabricante,Modelo,ValorVenda, ValorCusto)

gluc<- dluc%>%group_by(Fabricante)%>% summarise(Mvendas= mean(ValorVenda, na.rm= T), Mcusto= mean(ValorCusto, na.rm= T))

lucro<- gluc%>%mutate(Lucro= Mvendas - Mcusto)
lucro<- arrange(lucro,desc(lucro))
luper<- lucro[-1,]%>%mutate(LucroPer= Lucro/sum(lucro$Lucro))
luper$LucroPer<- round(luper$LucroPer, 2)
View(lucro)
View(luper)
```

Análise gráfica dos dados

Análise gráfica do padrão temporal de vendas.

```
st<-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"),panel.border = element_rect(size= 1.2), legend.text = element_text(size= 14),legend.position = c(.1, .9),
    legend.background = element_blank())
st
```

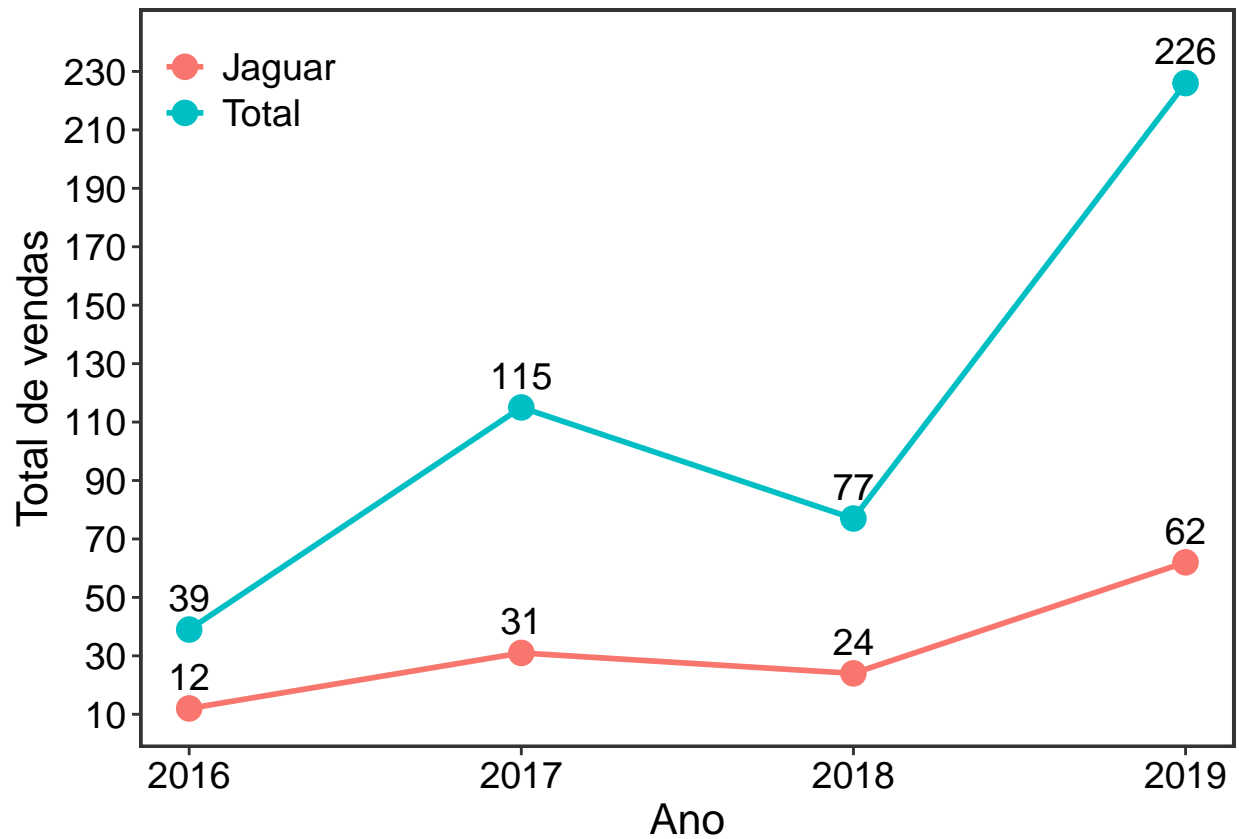


Gráfico do custo de entrega por fabricante

```
ggplot(entrega, aes(Fabricante, valor_entrega))+
  geom_col(fill="black")+
  annotate("text", x =entrega$Fabricante, y =entrega$valor_entrega+3000,
    label = round(entrega$valor_entrega, 2), size= 4)+
  ylab("Valor de entrega")+
  xlab("Fabricante")+
  theme_bw()+
  theme(panel.grid= element_blank(), axis.title = element_text(size= 16),
    axis.text = element_text(size= 14, color = "black"),panel.border = element_rect(size= 1.2), leg
    legend.text = element_text(size = 12))
```

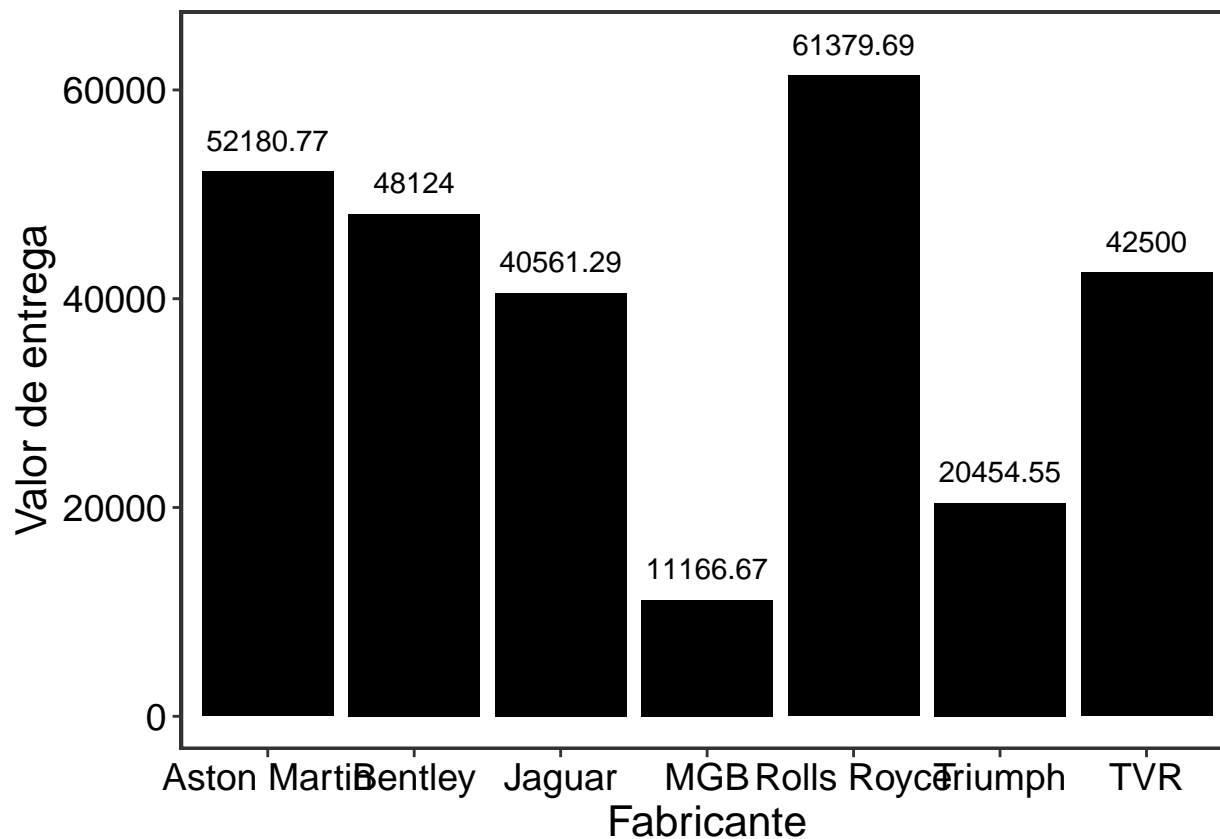


Gráfico da mão de obra por estado.

```
UF<-ggplot(mobra, aes(Estado, Media_MaoDeObra))+
  geom_col(fill="black")+
  coord_flip()+
  annotate("text", x =mobra$Estado, y =mobra$Media_MaoDeObra+75,
           label =round(mobra$Media_MaoDeObra,2), size= 5)+
  scale_y_continuous(limits = c(0,1200), breaks = seq(0,1200,150))+
  ylab("Valor mão de obra")+
  xlab("Fabricante")+
  theme_bw()+
  theme(panel.grid= element_blank(), axis.title = element_text(size= 16),
        axis.text = element_text(size= 14, color = "black"),panel.border = element_rect(size= 1.2), legend.text = element_text(size = 12))
```

UF

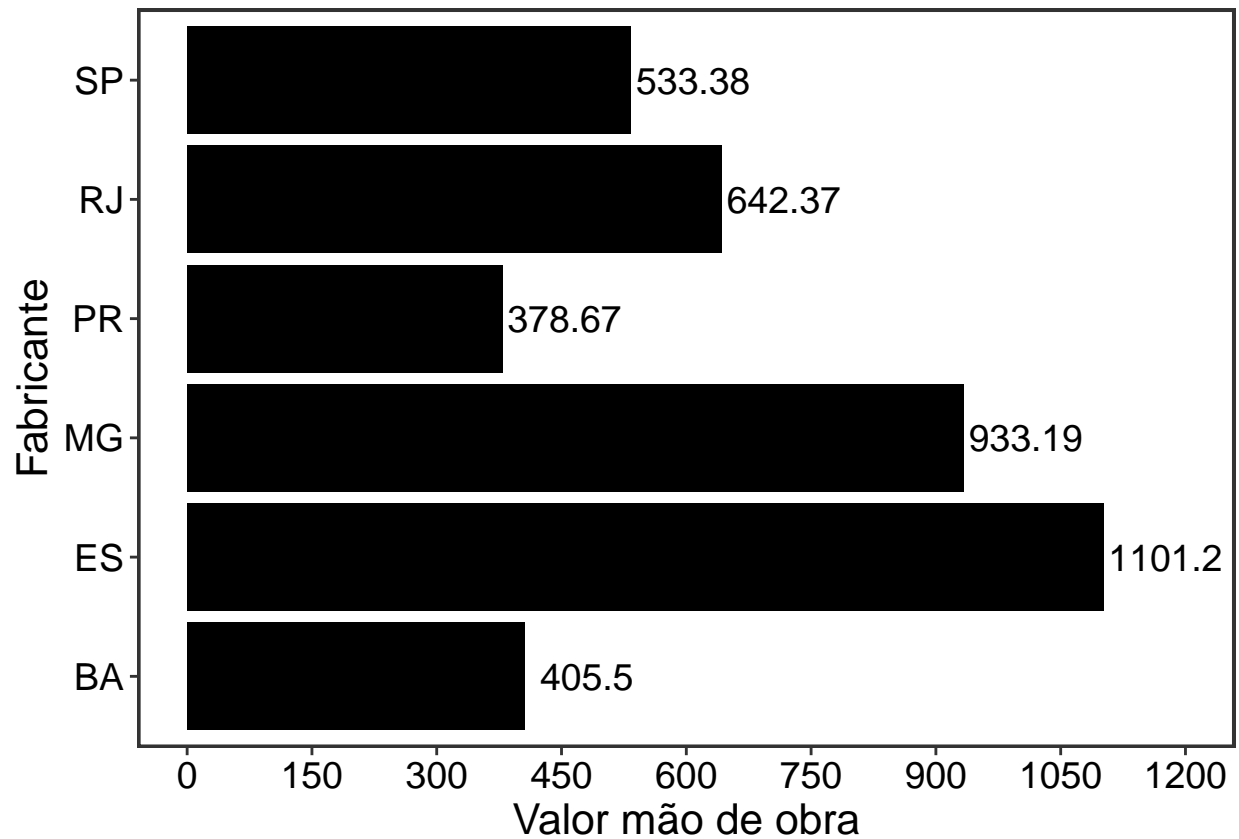


Grafico com o lucro de cada marca.

```
lc<-ggplot( lucro, aes(Fabricante, Lucro,fill= Fabricante))+
  geom_col()+
  coord_flip()+
  scale_fill_manual(values = c("grey", "darkorange", "cyan4","purple",
                                "blue", "darkgreen", "red")) +
  ylab("Lucro")+
  xlab("Fabricante")+
  theme_bw()+
  theme(panel.grid= element_blank(), axis.title = element_text(size= 16),
        axis.text = element_text(size= 14, color = "black"),panel.border = element_rect(size= 1.2), leg
lc
```

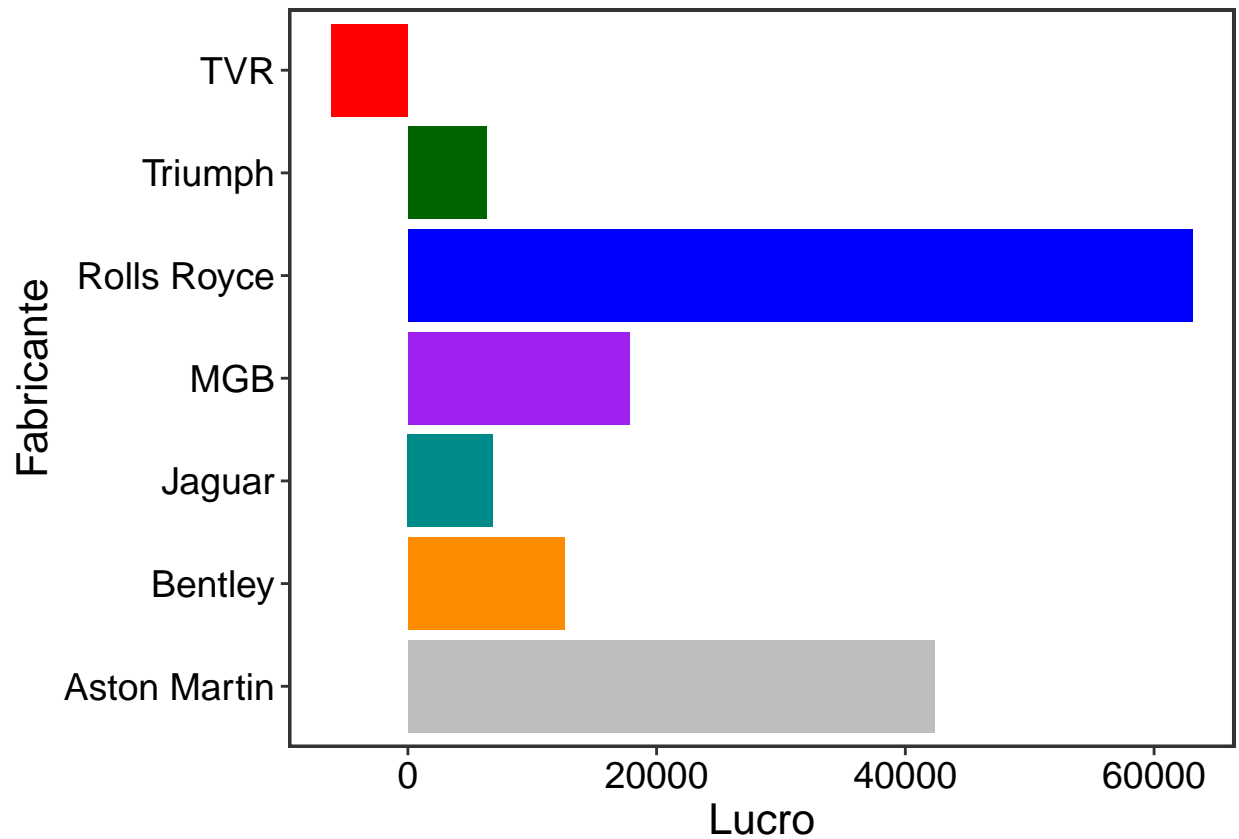
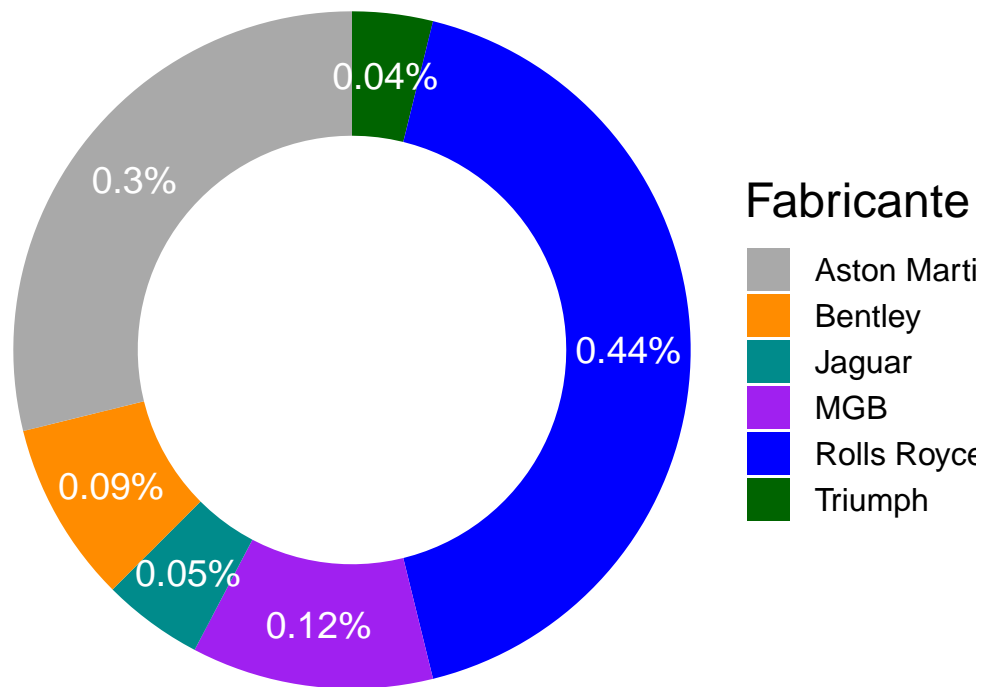


Grafico com a porcentagem de lucro para cada marca.

```
LcPer<-ggplot(luper, aes(2,LucroPer, fill= Fabricante))+
  geom_bar(stat= "identity")+
  geom_text(aes(label= paste0(LucroPer, "%")), color= "white",
            position= position_stack(vjust= .5), size= 5)+
  scale_fill_manual(values = c("darkgrey", "darkorange", "cyan4","purple",
                                "blue", "darkgreen")) +

  xlim(0, 2.5) +
  coord_polar(theta = "y", start = 0) +
  theme_void() +
  theme(legend.position = c(1.1, .5),
        legend.title = element_text(size = 18),
        legend.text = element_text(size = 12))
```

LcPer



Dashboard com os principais resultados.

```
ggarrange(st, UF, lc, LcPer, nrow = 2, ncol = 2)
```