ANP Brasil

Evolução do Preço dos combustíveis em 2012

Luciano Teixeira

11 de setembro de 2018

# Comandos R de preparação e transformação de dados

Resolva os exercícios a seguir utilizando os comandos no software RStudio através do RMarkdown. Gere um relatório em .doc com os comandos/código e as respectivas saídas e seus comentários.

* Exercício 1 - Crie um dataframe com 10 colunas e 40 linhas e imprima na tela as primeiras 6 linhas;
* Exercício 2 - Mostre a que classe pertence cada uma das 10 colunas do seu dataframe;
* Exercício 3 - Utilize um comando que faça uma avaliação exploratória das variáveis do seu dataframe e imprima na tela;
* Exercício 4 - Adicione mais uma coluna gerada a partir da transformação de uma das 10 originais;
* Exercício 5 - Faça um gráÒco que mostre a distribuição dos dados dessa nova variável do exercício 4.

# Dados da ANP - Agência Nacional do Petróleo

## Evolução de Preço dos Combustíveis no último Bimestre de 2012

### Resolução dos Exercícios

1.Crie um dataframe com 10 colunas e 40+ linhas

Carregando a biblioteca readr para leitura do arquivo em csv

library(readr)

Definindo a diretório de trabalho com a função setwd().

setwd("~/GitHub/GeneralRepositoriesUnisinos/PosUnisinosIntroducaoPythonR/Tarefa")

Lendo o arquivo “brasil.csv”.

anpbrasil <- read\_delim("brasil.csv", ";", escape\_double = FALSE, trim\_ws = TRUE)

## Parsed with column specification:  
## cols(  
## DATA\_INICIAL = col\_character(),  
## DATA\_FINAL = col\_character(),  
## PRODUTO = col\_character(),  
## POSTOS\_PESQUISADOS = col\_integer(),  
## UNIDADE\_MEDIDA = col\_character(),  
## PRECO\_MEDIO\_VENDA = col\_number(),  
## DESVIO\_PADRAO\_REVENDA = col\_character(),  
## PRECO\_MINIMO\_REVENDA = col\_number(),  
## PRECO\_MAXIMO\_REVENDA = col\_number(),  
## MARGEM\_MEDIA\_REVENDA = col\_character(),  
## COEF\_DE\_VARIACAO\_REVENDA = col\_character(),  
## PRECO\_MEDIO\_DISTRIBUICAO = col\_number(),  
## DESVIO\_PADRAO\_DISTRIBUICAO = col\_character(),  
## PRECO\_MINIMO\_DISTRIBUICAO = col\_character(),  
## PRECO\_MAXIMO\_DISTRIBUICAO = col\_number(),  
## COEF\_DE\_VARIACAO\_DISTRIBUICAO = col\_character()  
## )

Imprimindo na tela as primeiras 6 linhas.

head(anpbrasil)

## # A tibble: 6 x 16  
## DATA\_INICIAL DATA\_FINAL PRODUTO POSTOS\_PESQUISAD~ UNIDADE\_MEDIDA  
## <chr> <chr> <chr> <int> <chr>   
## 1 04/11/2012 10/11/2012 ETANOL HIDRATA~ 8195 R$/l   
## 2 11/11/2012 17/11/2012 ETANOL HIDRATA~ 8184 R$/l   
## 3 18/11/2012 24/11/2012 ETANOL HIDRATA~ 8176 R$/l   
## 4 25/11/2012 01/12/2012 ETANOL HIDRATA~ 8164 R$/l   
## 5 02/12/2012 08/12/2012 ETANOL HIDRATA~ 8168 R$/l   
## 6 09/12/2012 15/12/2012 ETANOL HIDRATA~ 8155 R$/l   
## # ... with 11 more variables: PRECO\_MEDIO\_VENDA <dbl>,  
## # DESVIO\_PADRAO\_REVENDA <chr>, PRECO\_MINIMO\_REVENDA <dbl>,  
## # PRECO\_MAXIMO\_REVENDA <dbl>, MARGEM\_MEDIA\_REVENDA <chr>,  
## # COEF\_DE\_VARIACAO\_REVENDA <chr>, PRECO\_MEDIO\_DISTRIBUICAO <dbl>,  
## # DESVIO\_PADRAO\_DISTRIBUICAO <chr>, PRECO\_MINIMO\_DISTRIBUICAO <chr>,  
## # PRECO\_MAXIMO\_DISTRIBUICAO <dbl>, COEF\_DE\_VARIACAO\_DISTRIBUICAO <chr>

2.Mostre a que classe pertence cada uma das 10 colunas do seu dataframe

## [1] "DATA\_INICIAL = character"

## [1] "DATA\_INICIAL = character"

## [1] "DATA\_FINAL = character"

## [1] "PRODUTO = character"

## [1] "POSTOS\_PESQUISADOS = integer"

## [1] "UNIDADE\_MEDIDA = character"

## [1] "PRECO\_MEDIO\_VENDA = numeric"

## [1] "DESVIO\_PADRAO\_REVENDA = character"

## [1] "PRECO\_MINIMO\_REVENDA = numeric"

## [1] "PRECO\_MAXIMO\_REVENDA\t= numeric"

## [1] "MARGEM\_MEDIA\_REVENDA\t= character"

## [1] "COEF\_DE\_VARIACAO\_REVENDA = character"

## [1] "PRECO\_MEDIO\_DISTRIBUICAO\t= numeric"

## [1] "DESVIO\_PADRAO\_DISTRIBUICAO = character"

## [1] "PRECO\_MINIMO\_DISTRIBUICAO = character"

## [1] "PRECO\_MAXIMO\_DISTRIBUICAO = numeric"

## [1] "COEF\_DE\_VARIACAO\_DISTRIBUICAO = character"

3.Utilize um comando que faça uma avaliação exploratória das variáveis do seu dataframe e imprima na tela.

summary(anpbrasil)

## DATA\_INICIAL DATA\_FINAL PRODUTO   
## Length:40 Length:40 Length:40   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##   
## POSTOS\_PESQUISADOS UNIDADE\_MEDIDA PRECO\_MEDIO\_VENDA  
## Min. : 491 Length:40 Min. : 40   
## 1st Qu.:7078 Class :character 1st Qu.: 1758   
## Median :8052 Mode :character Median : 2146   
## Mean :6499 Mean : 8612   
## 3rd Qu.:8181 3rd Qu.: 2749   
## Max. :8685 Max. :40151   
## DESVIO\_PADRAO\_REVENDA PRECO\_MINIMO\_REVENDA PRECO\_MAXIMO\_REVENDA  
## Length:40 Min. : 18 Min. : 62   
## Class :character 1st Qu.:1021 1st Qu.: 295   
## Mode :character Median :1399 Median : 312   
## Mean :1298 Mean :1378   
## 3rd Qu.:1869 3rd Qu.:2449   
## Max. :2279 Max. :3799   
## MARGEM\_MEDIA\_REVENDA COEF\_DE\_VARIACAO\_REVENDA PRECO\_MEDIO\_DISTRIBUICAO  
## Length:40 Length:40 Min. : 157   
## Class :character Class :character 1st Qu.: 1599   
## Mode :character Mode :character Median : 1927   
## Mean : 7366   
## 3rd Qu.: 2365   
## Max. :29857   
## DESVIO\_PADRAO\_DISTRIBUICAO PRECO\_MINIMO\_DISTRIBUICAO  
## Length:40 Length:40   
## Class :character Class :character   
## Mode :character Mode :character   
##   
##   
##   
## PRECO\_MAXIMO\_DISTRIBUICAO COEF\_DE\_VARIACAO\_DISTRIBUICAO  
## Min. : 44.0 Length:40   
## 1st Qu.: 242.0 Class :character   
## Median : 289.0 Mode :character   
## Mean : 938.5   
## 3rd Qu.:1846.0   
## Max. :2927.0

4.Adicione mais uma coluna gerada a partir da transformação de uma das 10 originais.

anpbrasil$TOTAL\_POSTOS<-(anpbrasil$POSTOS\_PESQUISADOS/sum(anpbrasil$POSTOS\_PESQUISADOS))/100

5.Faça um gráfico que mostre a distribuição dos dados dessa nova variável do exercício 4.