

# Introdução ao Tratamento e Análise de Dados em R

## Tabelas em R

Sérgio Rivero

PPGE-UFPA

19 de maio de 2019



# Sumário

- 1 Objetivo da Aula
- 2 Gerando Tabelas com Xtable e Stargazer
- 3 Análise
- 4 Exercícios



# Objetivo da Aula

- 1 Neste capítulo veremos como é possível gerar um grande conjunto de tabelas em R
- 2 vamos utilizar o *Markdown* para isso
- 3 bem como discutiremos a apresentação de resultados de modelos com estas bibliotecas.



# Gerando Tabelas com Xtable e Stargazer

- A geração de tabelas no R pode ser feita utilizando alguns pacotes específicos adequados para tal.
- Estes pacotes leem dataframes ou matrizes e geram tabelas a partir destes.
- Muitos destes pacotes têm também a capacidade de:
  - 1 ler objetos específicos do R (como resultados de regressões, análise de cluster, etc)
  - 2 e gerar tabelas correspondentes às saídas normais de apresentação destes resultados.



# xtable

Aqui temos o nosso velho conhecido banco de dados MTCARS. Utilizando a opção `auto`, o `xtable` define o tipo de saída em termos de casas decimais, automaticamente.

```
library(xtable)
```

```
mtc2 <- xtable(head(mtcars),  
auto = TRUE,  
caption = 'Base mtcars, primeiras linhas')
```

```
print.xtable(mtc2, caption.placement='top')
```



# Tabelas com xtable

**Tabela 1. Base mtcars, primeiras linhas**

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225	105	2.76	3.460	20.22	1	0	3	1

A tabela apresentando as primeiras 6 linhas do dataframe **mtcars**



# Gerando uma tabela sintética com dplyr

```
library(dplyr)

data(mtcars)

sintesePot <- as.data.frame(mtcars %>% group_by(cyl) %>%
  summarize(potenciaMedia = mean(hp, na.rm = TRUE),
    sigmaPotencia = sd(hp, na.rm = TRUE),
    n = n()
)
)

tabelaPot <- xtable(sintesePot,
  caption = 'Média e Desvio Padrão ...')
print(tabelaPot, type = "html",
  caption.placement='bottom')
```



## Tabelas com xtable

	cyl	potenciaMedia	sigmaPotencia	n
1	4.00	82.64	20.93	11
2	6.00	122.29	24.26	7
3	8.00	209.21	50.98	14

**Tabela 2. Média e Desvio Padrão da potência por no. de cilindros**





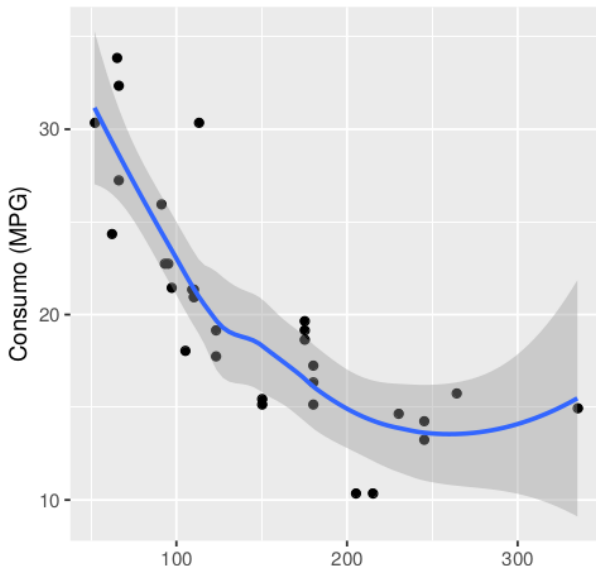
# Olhando para os dados

```
library(ggplot2)

ggplot(mtcars, aes(x=hp, y=mpg)) +
  geom_point() +
  geom_smooth() +
  labs(x='Potência (HP)',
       y = 'Consumo (MPG)')
```



# O gráfico



# Análise

Podemos então pensar em uma regressão que olhe para a relação entre potência e consumo

```
hpmpg <- lm(mpg~hp, data = mtcars)
```

```
summary(hpmpg)
```



# O resultado da regressão pelo *summary*

Call:

```
lm(formula = mpg ~ hp, data = mtcars)
```

Residuals:

Min	1Q	Median	3Q	Max
-5.7121	-2.1122	-0.8854	1.5819	8.2360

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	30.09886	1.63392	18.421	< 2e-16 ***
hp	-0.06823	0.01012	-6.742	1.79e-07 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.863 on 30 degrees of freedom

Multiple R-squared: 0.6024, Adjusted R-squared: 0.5892

F-statistic: 45.46 on 1 and 30 DF, p-value: 1.788e-07



# O resultado da regressão pelo *xtable*

```
lm.hpmpg.tab <- xtable(hpmpg,  
  caption = 'Resultado do Modelo Consumo versus HP',  
  label   = 'tabreg:consumohp'  
)  
  
print.xtable(lm.hpmpg.tab, caption.placement = 'top')
```

**Tabela 3. Resultado do Modelo Consumo versus HP**

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	30.0989	1.6339	18.42	0.0000
hp	-0.0682	0.0101	-6.74	0.0000



# Utilizando o **stargazer**

```
library(stargazer)

stargazer(hmpmg, type='html',
          title = 'Resultado do Modelo Consumo versus HP'
)
```



# A saída do **stargazer**

Tabela 4. Resultado do Modelo Consumo versus HP

<i>Dependent variable:</i>	
	mpg
hp	-0.068*** (0.010)
Constant	30.099*** (1.634)
Observations	32
R <sup>2</sup>	0.602
Adjusted R <sup>2</sup>	0.589
Residual Std. Error	3.863 (df = 30)
F Statistic	45.460*** (df = 1; 30)

*Note:*

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01



# Exercícios

## Exercícios

