

Nome: **Arthur Henrique de Oliveira Petrolí**

RA: **2408732**

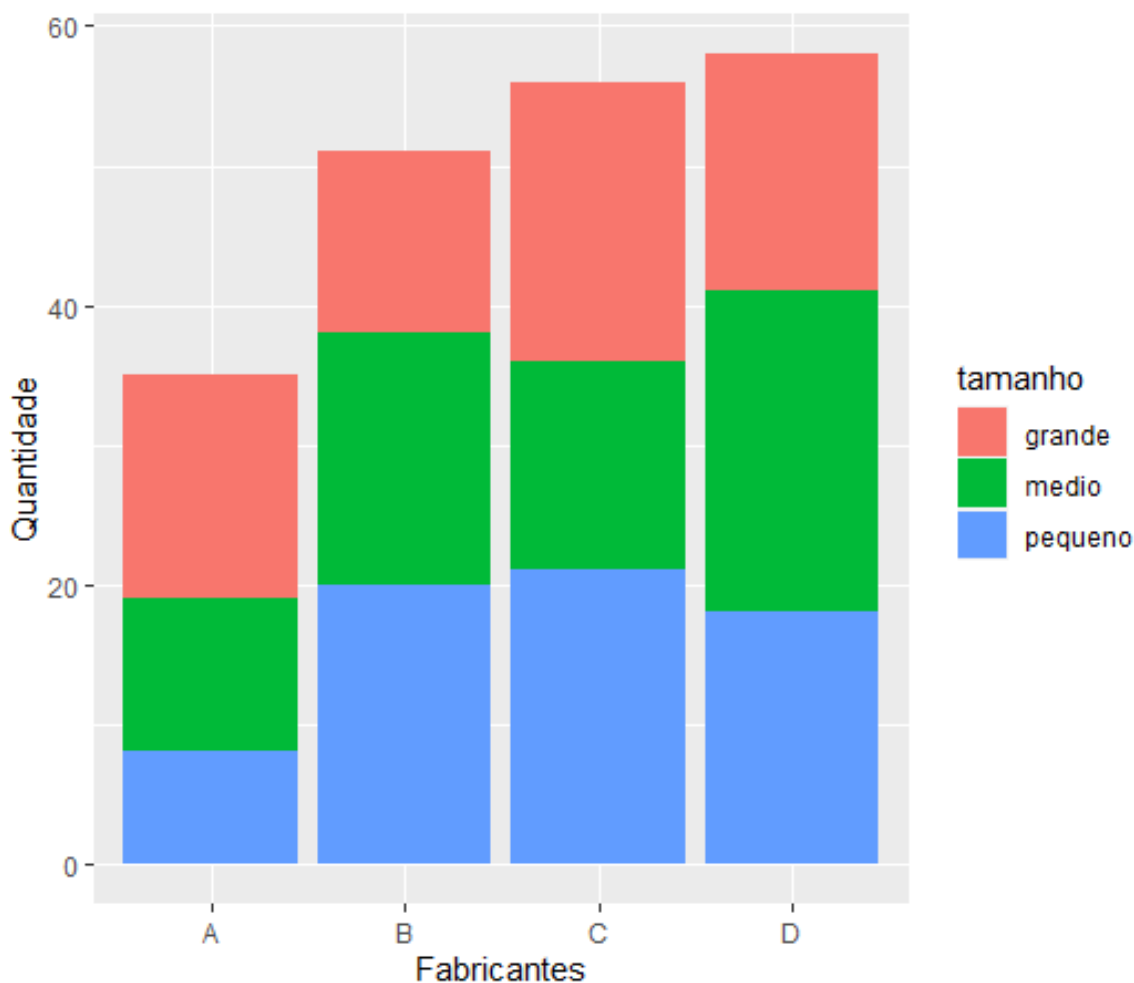
Disciplina matriculado(a): Probabilidade e Estatística – Eng. Comp.

Uma pesquisa foi realizada para identificar a resistência de produtos produzidos por 4 tipos de fabricantes (A, B, C, D). Esses fabricantes também foram classificados considerando 3 tamanhos de empresas (pequeno, médio ou grande). Além das informações das resistências dos produtos, também foram anotados o tempo para produzir os mesmos, bem como o peso.

Para todas as questões abaixo, **interprete os resultados e apresente os códigos**. A base de dados está com o separador decimal em Inglês, ou seja, as decimais estão separadas por ponto.

Pede-se:

ex1) Elabore um gráfico de barras agrupadas, em que cada barra represente a quantidade de fabricantes pesquisados, e que dentro de cada barra apareçam os tamanhos.



Códigos:

```
#carregando a base de dados
```

```
dados=read.csv('dados1.csv', sep = ',',dec='.',header = T)
```

```
#instalando a biblioteca
```

```
install.packages("ggplot2")
```

```
#abrindo a biblioteca
```

```
library(ggplot2)
```

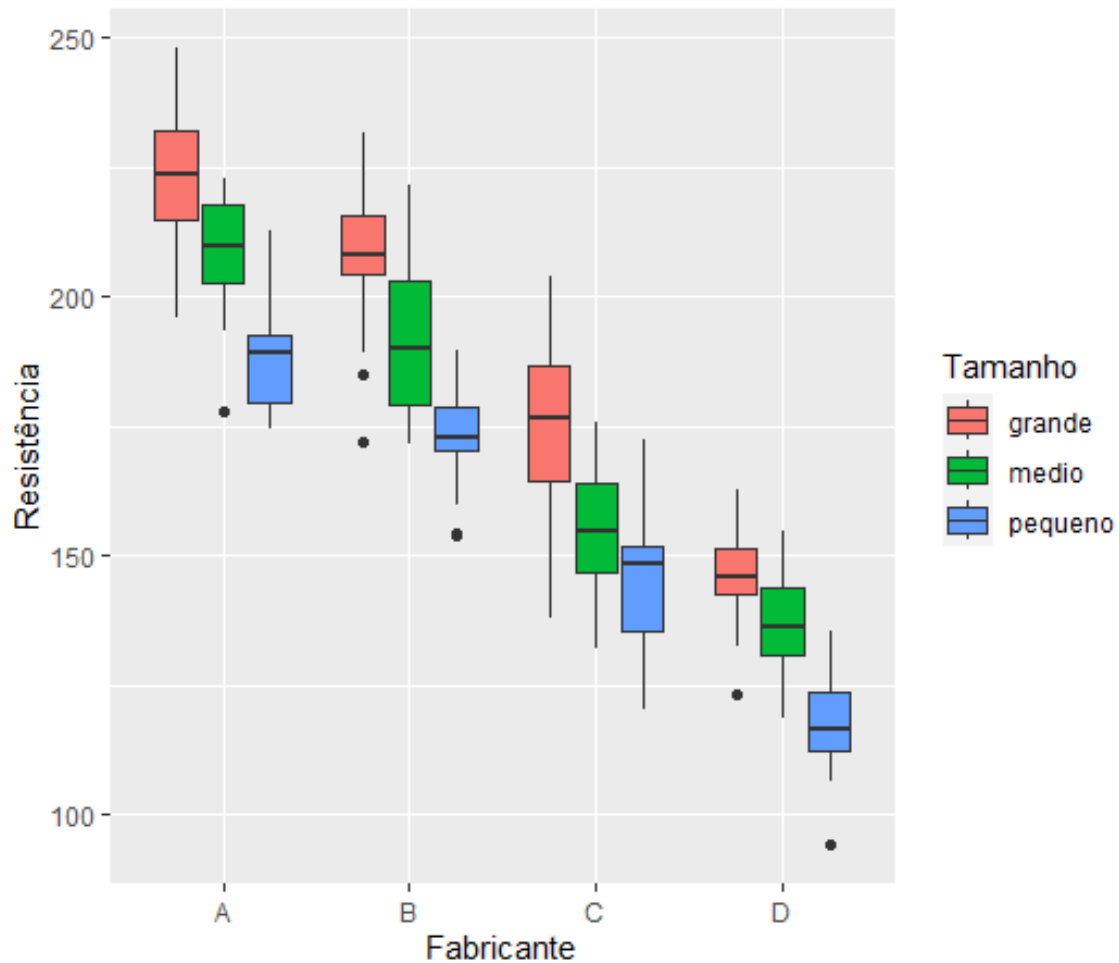
```
####ex1####
```

```
# Criando o gráfico de barras agrupadas
```

```
ggplot(dados,aes(x=fabricante,fill=tamanho))+geom_bar()+labs(x= 'Fabricantes',y=  
'Quantidade')
```

A saída foi um gráfico de barras agrupadas onde, à distribuição da quantidade de fabricantes pelo tamanho para cada tipo de fabricante. As barras do eixo x são divididas em 3 partes representando os tamanhos de empresa(pequeno, médio, grande), e o eixo y representa a quantidade de fabricantes de um determinado tipo.

ex2) Apresente os boxplots das resistências dos produtos, considerando o tipo de fabricante e tamanho.



Códigos:

```
####ex2####
```

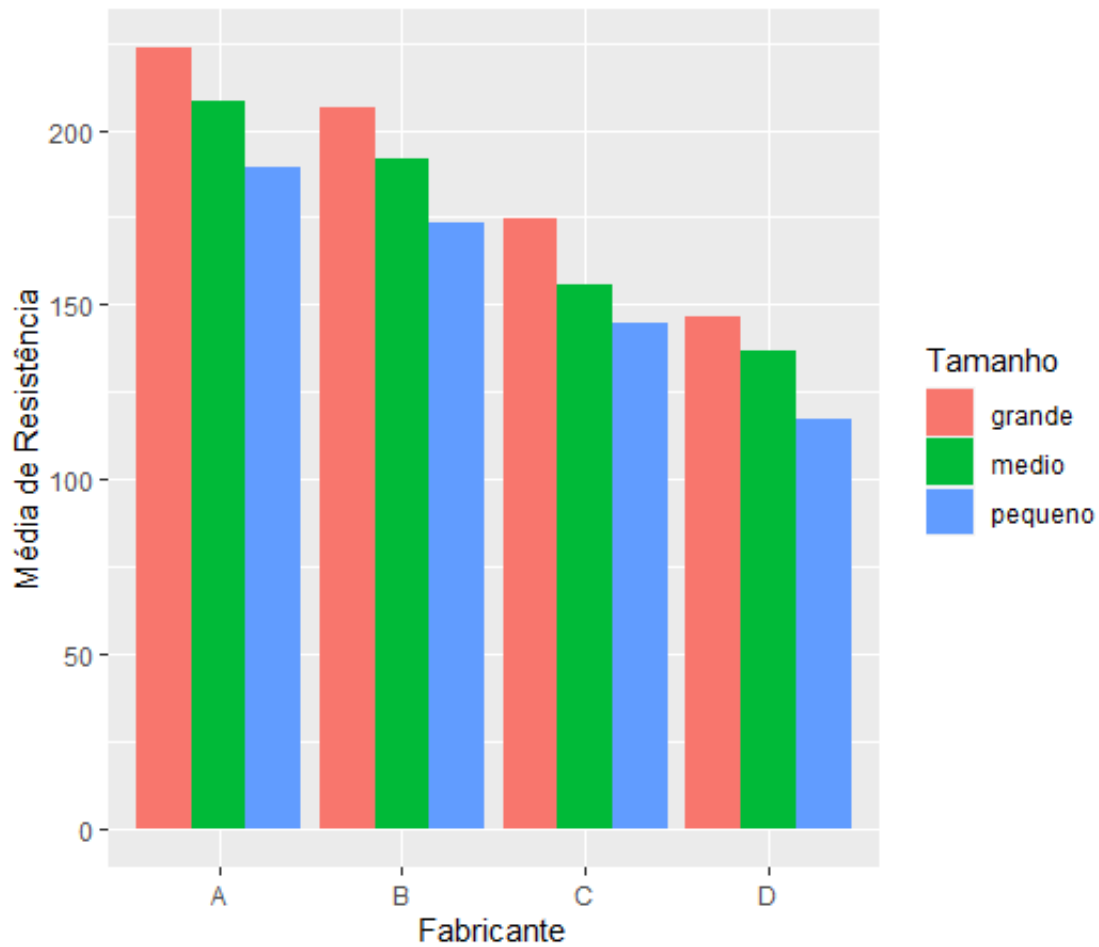
```
# Criando o gráfico de boxplot
```

```
ggplot(dados,aes(y=resistencia,x=fabricante,fill=tamanho))+geom_boxplot()+labs(x =  
"Fabricante", y = "Resistência", fill = "Tamanho")
```

A saída foi um gráfico em forma de diagrama de caixas onde relaciona-se a resistência com o fabricante e seus tipos de tamanho, a linha que divide ao meio os *boxplots* representa a mediana, é possível observar que os fabricantes B e D possuem mais pontos fora da curva que os fabricantes A e C, e além disso os fabricantes A e B possuem as maiores resistências

de produtos.

ex3) Determine a média das resistências por fabricante x tamanho. Apresente um gráfico de sua preferência apresentando os resultados.



Códigos:

```
####ex3####
```

```
# Calculando as médias das resistências por fabricante e tamanho
```

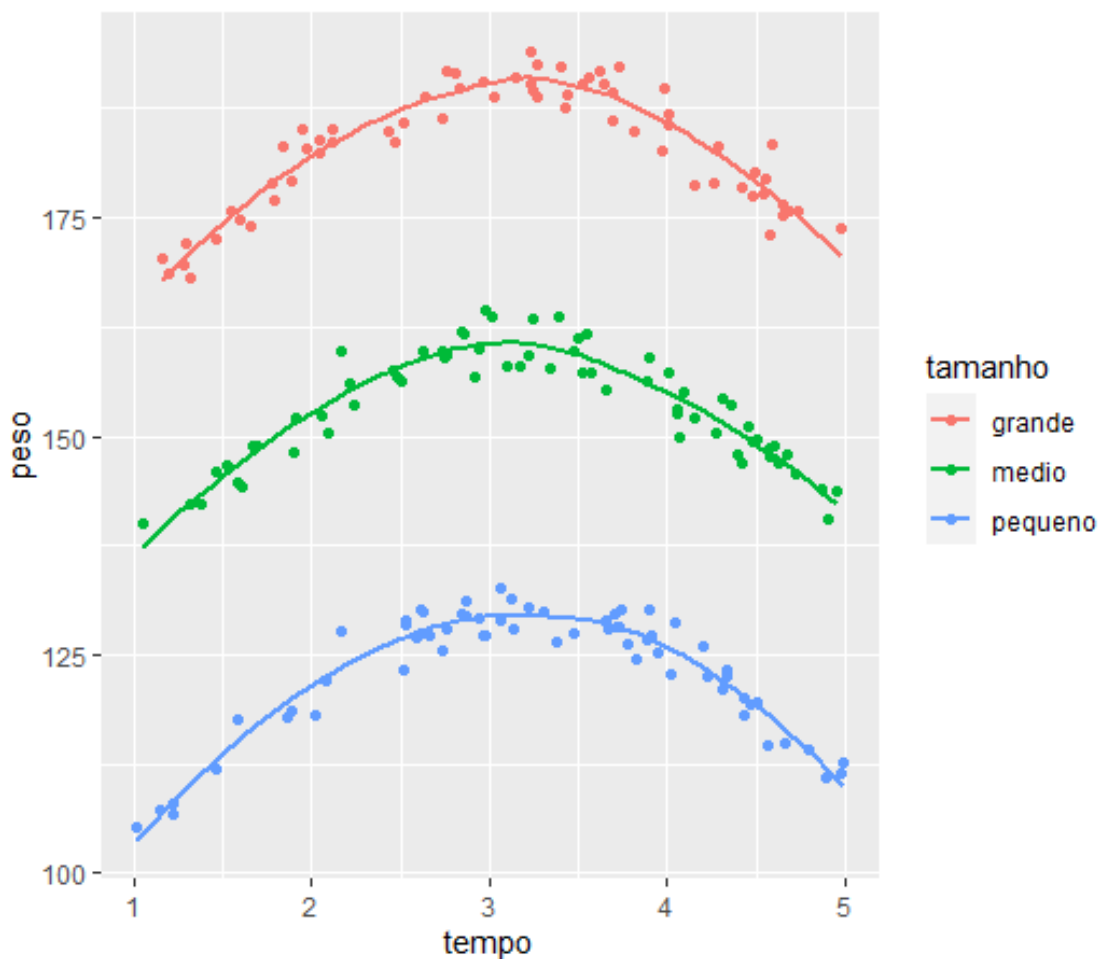
```
medias <- aggregate(resistencia ~ fabricante + tamanho, data = dados, mean)
```

```
# Criando o gráfico de barras agrupadas
```

```
ggplot(medias, aes(x = fabricante, y = resistencia, fill = tamanho)) + geom_bar(stat = "identity",  
position = position_dodge()) + labs(x = "Fabricante", y = "Média de Resistência", fill =  
"Tamanho")
```

A saída foi um gráfico de barras agrupadas por fabricante conforme o tipo onde o eixo y representa a média das resistências e eixo x os fabricantes, é possível notar que o fabricante A apresenta as maiores médias de resistência enquanto que o D apresenta as menores médias.

ex4) Verifique se existe alguma relação entre tempo (eixo x) e peso. Considere para cada tipo de tamanho. Qual o tempo ótimo aproximado para a produção do produto (considere que quanto maior o peso, melhor)?



Códigos:

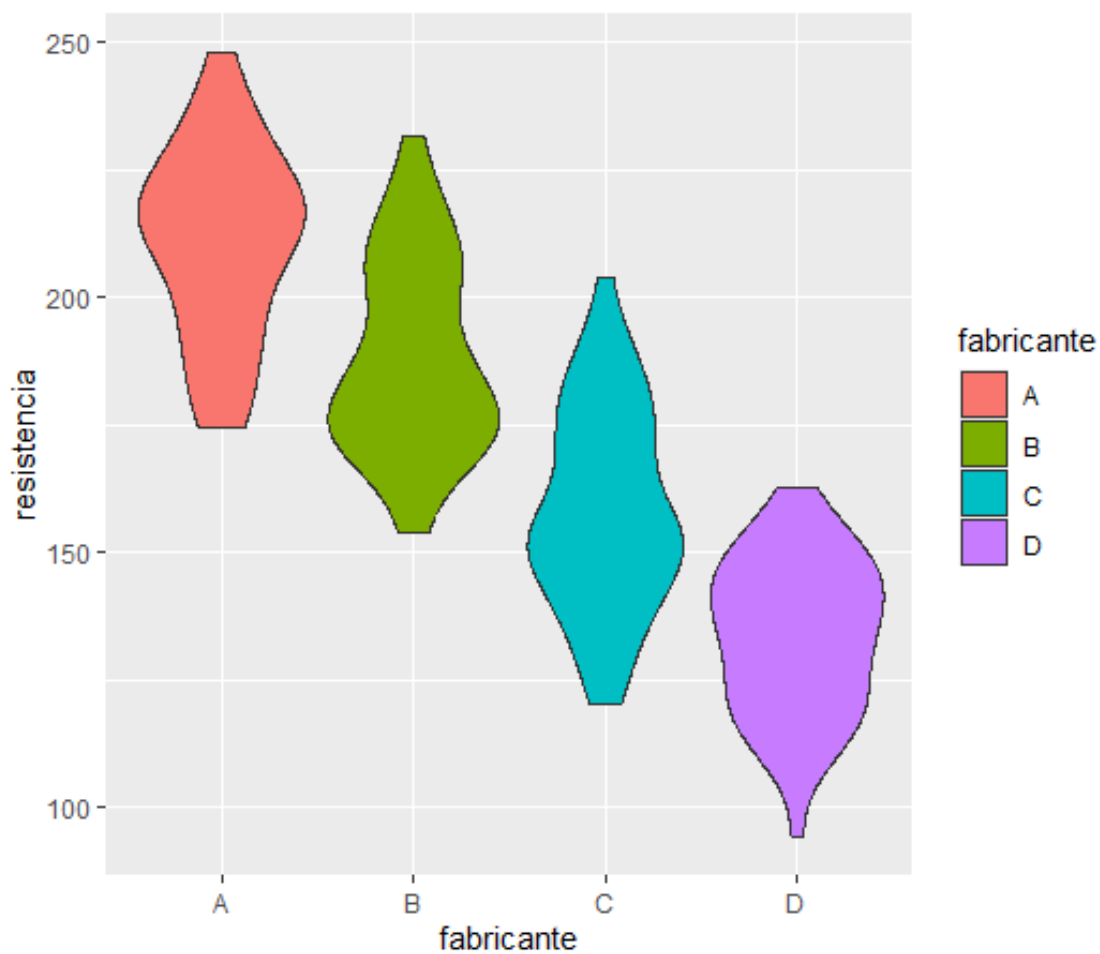
```
####ex4####
```

```
#Criando gráfico de dispersão
```

```
ggplot(dados,aes(x=tempo,y=peso,color=tamanho))+geom_point()+geom_smooth(se = F)
```

A saída é um gráfico de dispersão que mostra o peso por tempo de produção de acordo com o tipo de tamanho, entre os valores 2 e 4 de tempo os 3 tipos apresentam um pico de peso, logo entre esses valores de tempo seria o melhor momento para a produção do produto.

ex5) Faça um gráfico de densidade ou violino das resistências por tipo de fabricante.



Códigos:

```
####ex5####
```

```
#Criando gráfico de violino
```

```
ggplot(dados,aes(fabricante,resistencia,fill=fabricante))+geom_violin()
```

A saída foi um gráfico de violino que relaciona as resistências com o fabricante onde são concentrados a maioria dos valores, assim quanto maior a largura maior a densidade de valores na região.