**ATIVIDADE 2**

Realizar a atividade abaixo e postar a resolução no local indicado para envio de arquivo no APRENDER.

**EXERCÍCIOS**

1. Para um estudo sobre o perfil dos idosos que moram no Asilo Santa Marcelina, foi coletada a idade de uma amostra de 10 idosos:

67 69 102 72 83 77 69 82 74 83

1. Calcule as medidas de posição para o conjunto acima.

dados <- c(67, 69, 102, 72, 83, 77, 69, 82, 74, 83)

quantile(dados)

1. Calcule o coeficiente de variação e verifique se o conjunto de dados podem ser considerados homogêneos. Justifique sua resposta.

# Calcular a média

media <- mean(dados)

# Calcular o desvio padrão

desvio\_padrao <- sd(dados)

# Calcular o coeficiente de variação

coefficiente\_variacao <- (desvio\_padrao / media) \* 100

Resultado: 13.39614

Os dados podem ser considerados homogêneos pois a variabilidade dos dados em relação à média é baixa, o que sugere que os valores estão mais próximos uns dos outros..

1. Verifique se a média é uma medida adequada para representar esse conjunto de dados e justifique sua resposta.

**Não, pois a presença de valores extremos, como o 102 no conjunto de dados, pode distorcer a média, tornando-a menos representativa da maioria dos valores.**

1. Suponha que as notas finais de uma disciplina de programação foram:

3, 7, 4, 5, 1, 8, 4, 6, 5, 6, 2, 4, 6, 9, 8, 4, 5, 5, 6

Separe os dados em dois grupos:

“Aprovados”, com nota maior ou igual que 5;

aprovados <- c(5, 5, 5, 5, 6, 6, 6, 6, 7, 8, 8, 9)

“Reprovados”, com nota menor que cinco.

reprovados <- c(1, 2, 3, 4, 4, 4, 4)

a) Calcule a amplitude de cada grupo.

range(aprovados)

# Resposta = 5 9

range(reprovados)

# Resposta = 1 4

a) Analise se é possível utilizar o desvio-padrão para comparar a variabilidade dos 2 grupos. Explique.

**Ao comparar a variabilidade entre dois grupos com tamanhos diferentes, como no caso dos grupos de aprovados e reprovados, é importante ter em mente que o desvio-padrão pode não ser a medida mais apropriada por si só.**

b) Determine, através da medida adequada, qual é o grupo com menor variação.

aprovados # 21.64246

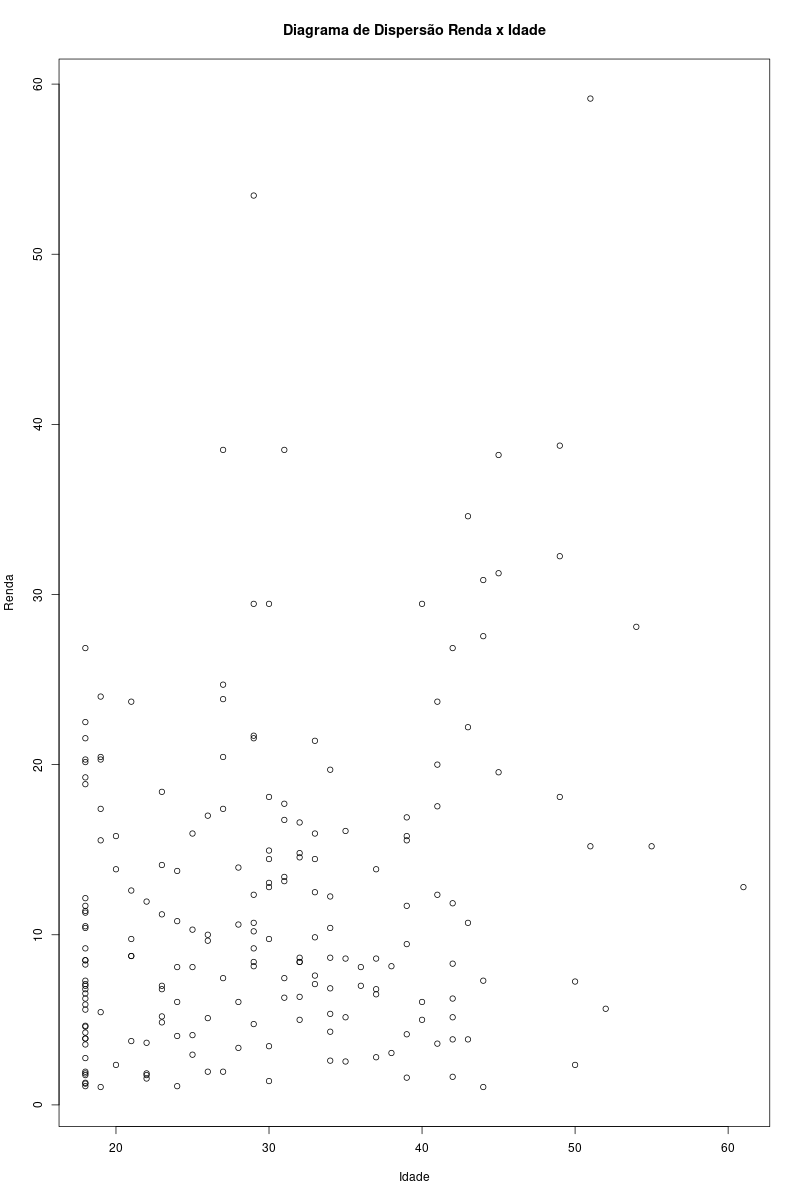
c) Verifique se os 2 grupos podem ser considerados homogêneos.

**Como o coeficiente de variação para ambos os grupos é igual a zero, isso indica que não há variabilidade nos dados, o que não é um resultado comum em situações reais. Portanto, com base nos dados fornecidos, não é possível determinar se os dois grupos são homogêneos a partir do coeficiente de variação.**

1. Importe o conjunto de dados “amostra.csv” e realize uma Análise de Correlação entre as variáveis Renda e Idade, cumprindo as 3 etapas:

* Construção do Diagrama de Dispersão

plot(dados$Idade, dados$Renda, xlab = "Idade", ylab = "Renda", main = "Diagrama de Dispersão Renda x Idade")



* Teste de Pearson e Cálculo do Coeficiente de Correlação de Pearson
* **resultado**
* **Pearson's product-moment correlation**
* **data: dados$Renda and dados$Idade**
* **t = 3.7473, df = 198, p-value = 0.0002345**
* **alternative hypothesis: true correlation is not equal to 0**
* **95 percent confidence interval:**
* **0.1229895 0.3824264**
* **sample estimates:**
* **cor**
* **0.2573394**
* Ajuste do modelo de Regressão

modelo <- lm(Renda ~ Idade, data = dados)

summary(modelo)

Residuals:

Min 1Q Median 3Q Max

-14.728 -6.374 -2.213 3.727 41.822

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.61088 2.08604 2.210 0.028227 \*

Idade 0.24935 0.06654 3.747 0.000234 \*\*\*

---

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

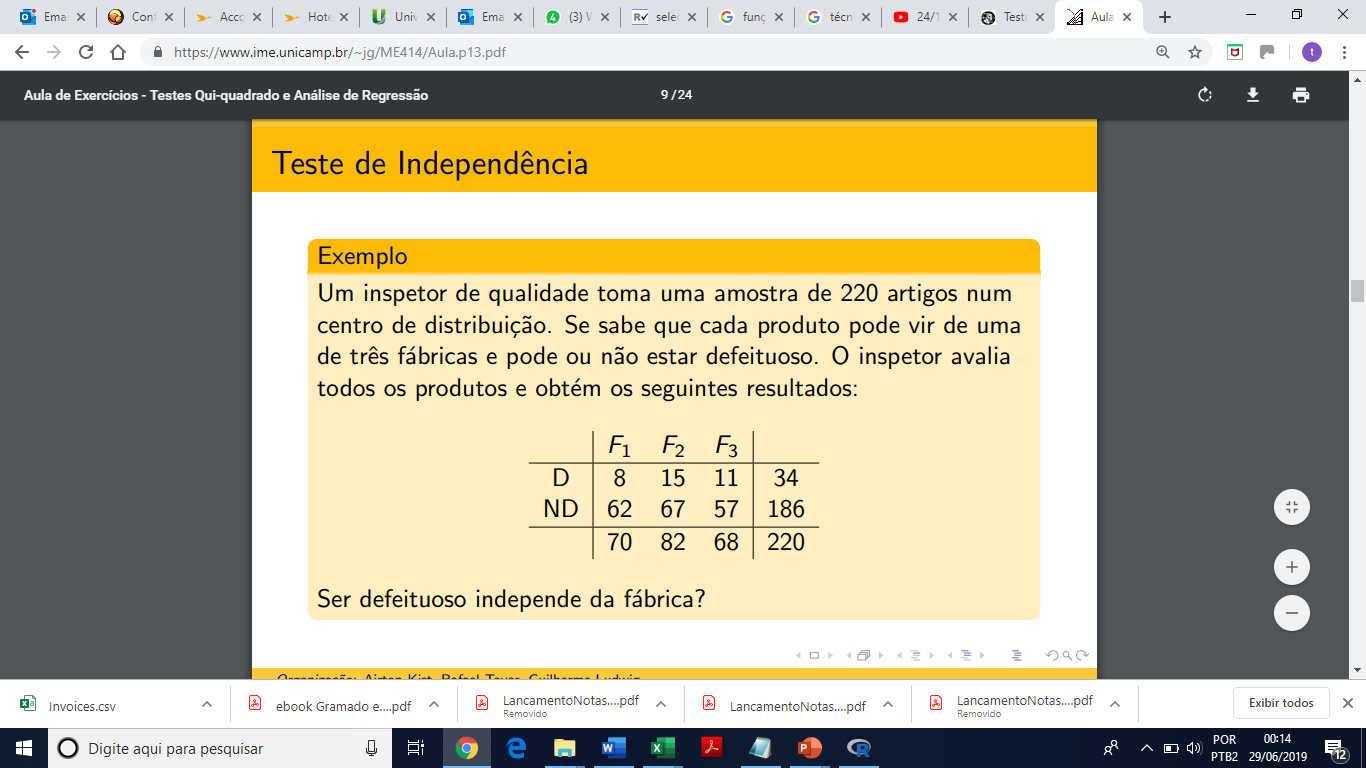
Residual standard error: 9.189 on 198 degrees of freedom

Multiple R-squared: 0.06622, Adjusted R-squared: 0.06151

F-statistic: 14.04 on 1 and 198 DF, p-value: 0.0002345

Interprete os resultados de cada etapa e insira a reta do modelo no Diagrama de Dispersão utilizando a função *abline.*

**1. Ajuste do Modelo de Regressão: - O primeiro passo foi ajustar um modelo de regressão linear simples entre as variáveis "Renda" e "Idade" com o comando `lm`. - O resumo do modelo foi obtido com o comando `summary(modelo)` para visualizar os coeficientes de regressão, estatísticas de ajuste e significância dos coeficientes. 2. Interpretação dos Resultados: - No resumo do modelo, observe os coeficientes de regressão, em particular o coeficiente associado à variável "Idade", que indicará a variação na variável "Renda" para cada unidade de variação na "Idade". - Verifique o valor de R² (coeficiente de determinação), que indica a proporção da variabilidade da variável resposta explicada pelas variáveis explicativas. 3. Inserção da Reta do Modelo no Diagrama de Dispersão: - Utilize a função `abline` para adicionar a reta de regressão ao Diagrama de Dispersão. - O argumento `abline(modelo)` ajusta a reta do modelo de regressão ao gráfico criado com a função `plot`.**



Aplique o teste adequado e verifique se o fato de o produto ser defeituoso depende da fábrica de origem ou não.

**Pearson's Chi-squared test**

**data: dados**

**X-squared = 1.401, df = 2, p-value = 0.4963**

**Pelo resultado do p–value é rejeitado o H0 portanto não tem relação**