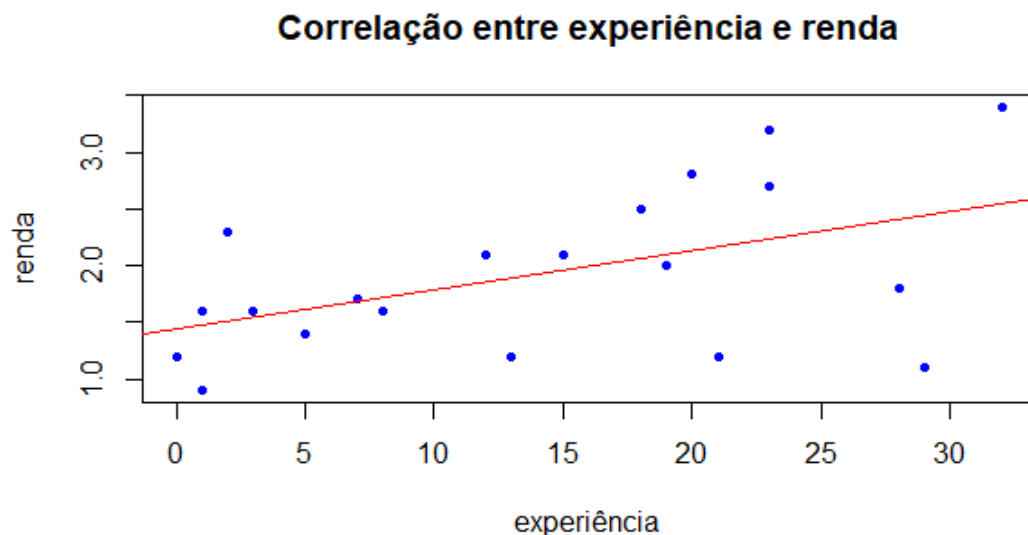


## QUESTÃO 2

### Item 3.

- Construindo um scatter plot em R:



Nota-se que não há um padrão referente a posição dos pontos. Isso indica uma considerável dispersão entre os dados em questão.

- Cálculo do coeficiente de correlação de Pearson

$$P = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

**Saída em R:** 0.4977672

O coeficiente de Pearson nos indica uma correlação positiva entre os dados, ou seja, eles tendem a crescer juntos.

No entanto, essa correlação é moderada, tendo em vista que o coeficiente de Pearson varia de -1 a 1. E como indicado no gráfico, a dispersão entre os dados é bastante considerável.

## Code R:

```
experiencia <- c(2, 8, 21, 1, 15, 3, 28, 5, 13, 20, 32, 23, 1, 0, 29,
18, 19, 7, 12, 23)
renda <- c(2.3, 1.6, 1.2, 0.9, 2.1, 1.6, 1.8, 1.4, 1.2, 2.8, 3.4, 2.7,
1.6, 1.2, 1.1, 2.5, 2.0, 1.7, 2.1, 3.2)

plot(experiencia, renda, main = "Correlação entre experiência e
renda", xlab = "experiência", ylab = "renda", col = "blue", pch=20)
abline(lm(renda~experiencia), col = "red")

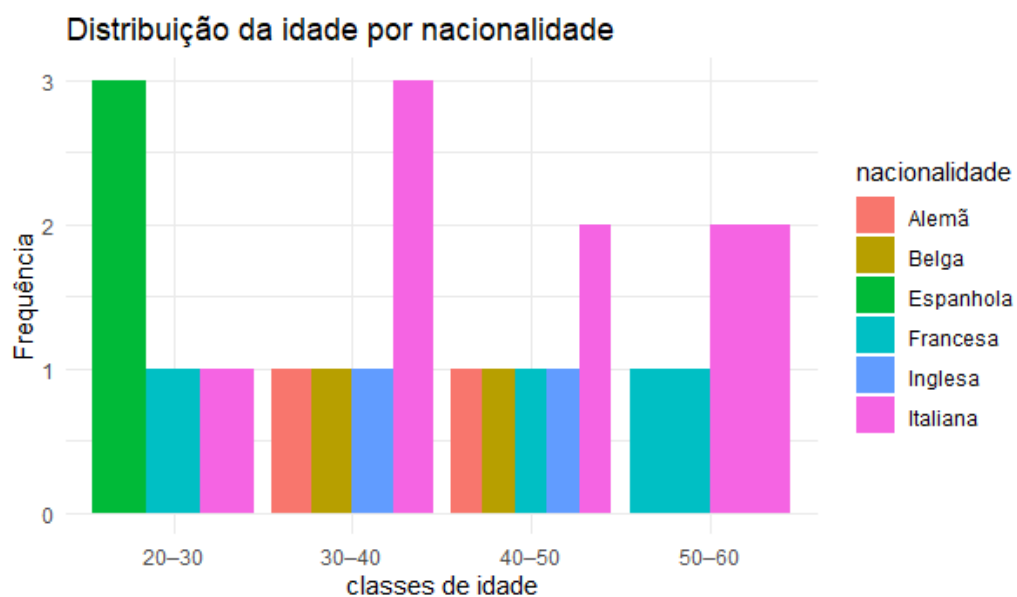
cor(experiencia, renda)
```

## Item 4.

IDADE	NACIONALIDADE	RENDA	EXPERIÊNCIA
46	Belga	1.2	21
51	Francesa	1.8	28
39	Italiana	1.2	13
52	Italiana	1.1	29

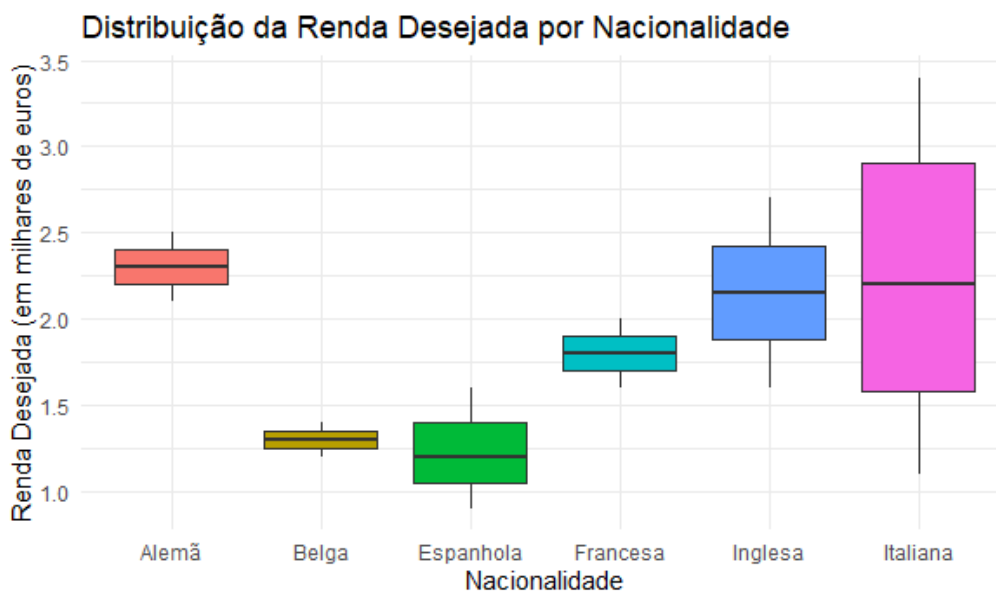
Tem-se quatro candidatos que correspondem aos critérios dados.

## Item 5.



## - Diferenças entre grupos de idade

- Todas as pessoas de nacionalidade espanhola estão presentes na classe de pessoas mais jovens.
- A nacionalidade italiana é a única que está presente em todas as classes, estando a maior parte delas na classe 30-40.
- A nacionalidade francesa está presente em todas as classes, exceto em 30-40.
- Todas as pessoas de nacionalidade alemã têm entre 30 a 50 anos de idade, tal qual pessoas de nacionalidade belga e inglesa.



## - Diferenças entre grupos de renda

- O grupo da nacionalidade italiana apresenta a maior dispersão de renda desejada, assim como o maior máximo de renda desejada. A tendência central do grupo para renda, coincide com a tendência central dos grupos das nacionalidades inglesa, e alemã.
- O grupo da nacionalidade belga apresenta a menor dispersão de renda desejada, a tendência central do grupo para renda, coincide com a tendência central do grupo da nacionalidade espanhola.
- O grupo de nacionalidade espanhola apresenta o menor mínimo de renda desejada.

## Code R

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

# ----- histograma para idade -----
dadosA <- data.frame(
  idade = c(28, 34, 46, 26, 37, 29, 51, 31, 39, 43, 58, 44, 25, 23, 52, 42,
48, 33, 38, 46),
  nacionalidade = c("Italiana", "Inglesa", "Belga", "Espanhola",
                    "Italiana", "Espanhola", "Francesa", "Belga",
                    "Italiana", "Italiana", "Italiana", "Inglesa",
                    "Francesa", "Espanhola", "Italiana", "Alemã",
                    "Francesa", "Italiana", "Alemã", "Italiana")
)

#faixas etárias
dadosA$classe_idade <- cut(
  dados$idade,
  breaks = seq(20, 60, by = 10),
  labels = c("20-30", "30-40", "40-50", "50-60"),
  include.lowest = TRUE
)

# Histograma agrupado
ggplot(dadosA, aes(x = classe_idade, fill = nacionalidade)) +
  geom_bar(position = "dodge") +
  labs(
    title = "Distribuição da idade por nacionalidade",
    x = "classes de idade",
    y = "Frequência"
  ) +
  theme_minimal()

# ----- boxplot para renda -----

dadosB <- data.frame(
  renda_desejada = c(2.3, 1.6, 1.2, 0.9, 2.1, 1.6, 1.8, 1.4, 1.2, 2.8,
3.4, 2.7, 1.6, 1.2, 1.1, 2.5, 2.0, 1.7, 2.1, 3.2),
  nacionalidade = c("Italiana", "Inglesa", "Belga", "Espanhola",
                    "Italiana", "Espanhola", "Francesa", "Belga",
                    "Italiana", "Italiana", "Italiana", "Inglesa",
                    "Francesa", "Espanhola", "Italiana", "Alemã",
                    "Francesa", "Italiana", "Alemã", "Italiana")
)

# Boxplot
ggplot(dadosB, aes(x = nacionalidade, y = renda_desejada, fill =
nacionalidade)) +
  geom_boxplot() +
  labs(
    title = "Distribuição da Renda Desejada por Nacionalidade",
```

```
x = "Nacionalidade",  
y = "Renda Desejada (em milhares de euros)"  
) +  
theme_minimal() +  
theme(legend.position = "none")
```