# Análise da amostra

**Tamanho da amostra:** 20.

**Unidade de análise:** funcionários.  
  
**Comando:** *table(scores$Graduacao)*

|  |  |  |
| --- | --- | --- |
|  | **Sim** | **Não** |
| **Possui graduação?** | 09 (**45%**) | 11 (**55%**) |

**Comando:** *table(scores$Anos\_de\_Experiencia,scores$Graduacao)*

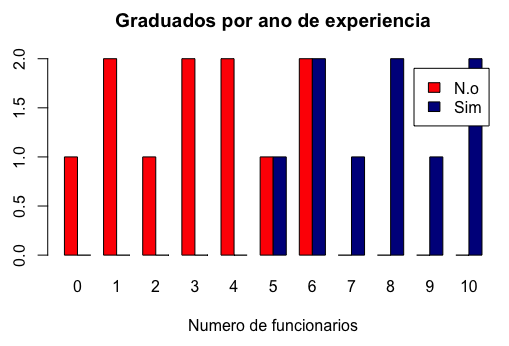
|  |  |  |
| --- | --- | --- |
|  | **Possui graduação?** | |
| **Anos de experiência** | **Sim** | **Não** |
| 00 | 00 | 01 |
| 01 | 00 | 02 |
| 02 | 00 | 01 |
| 03 | 00 | 02 |
| 04 | 00 | 02 |
| 05 | 01 | 01 |
| 06 | 02 | 02 |
| **07** | **01** | **00** |
| **08** | **02** | **00** |
| **09** | **01** | **00** |
| **10** | **02** | **00** |

**Comandos:** *anos\_graduacao<-table(scores$Graduacao,scores$Anos\_de\_Experiencia)*

*barplot(anos\_graduacao, main="Graduados por ano de experiencia",*

*xlab="Numero de funcionarios", col=c("red","darkblue"),*

*legend = rownames(anos\_graduacao),beside=TRUE)*

****

**Observações:**  
Dentro da amostra de funcionários da empresa, 55% (11) deles não possuem graduação. Ao cruzar anos de experiência e escolaridade, é possível observar que todos os funcionários com mais de 07 anos de experiência possuem graduação. Em contrapartida, todos os funcionários com 04 anos ou menos de experiência não possuem graduação.

# Análise descritiva

### Anos de experiência

**Comando:** *summary(scores$Anos\_de\_Experiencia)*

* **Mínimo:** 0 anos.
* **Percentil 25 (primeiro quartil):** 03 anos.
* **Percentil 50 (mediana):** 5,5 anos.
* **Média:** 5,2 anos.
* **Percentil 75 (terceiro quartil):** 7,25 anos.
* **Máximo:** 10 anos.

**Comando:** *boxplot(scores$Anos\_de\_Experiencia)*



**Comando:** *hist(scores$Anos\_de\_Experiencia, ylab="Frequencia",main="Histograma de anos de experiência", xlab="Anos de experiência")*



**Observações:** Ao analisar o tempo de experiência, é possível verificar que 50% dos funcionários possuem no máximo 5,5 anos de atuação profissional. Ao analisar a média aritmética, tem-se um valor próximo de 5 (5,2), o que dá uma ideia de uma distribuição de experiência próxima de algo simétrico (média, moda e mediana com valores iguais), que permitiria dizer que 5 anos é uma medida que divide a amostra de funcionários ao meio. Outro ponto a se considerar é de que apenas 25% dos funcionários (total de 5) possuem mais do que 7,25 anos experiência.

### Escore do teste

**Comando:** *summary(scores$Escore\_teste)*

* **Mínimo:** 70.
* **Percentil 25 (primeiro quartil):** 77,25.
* **Percentil 50 (mediana):** 82,5.
* **Média:** 82,75.
* **Percentil 75 (terceiro quartil):** 82,75.
* **Máximo:** 100.

**Comando:** *boxplot(scores$Escore\_teste)*



**Comando:** *hist(scores$Escore\_teste, ylab="Frequencia",main="Histograma de escore", xlab="Escore")*

**

**Observações:** Ao analisar o escore (desempenho dos funcionários na avaliação aplicada), é possível constatar que o resultado mínimo foi de 70. Se a avaliação aplicada teve como intervalo valores entre 0 à 100, todos os funcionários atingiram, no mínimo, 70% da nota máxima. Outra informação relevante diz respeito a proximidade entre a média, mediana e percentil 75 (82,75, 82,5 e 82,75 respectivamente), o que nos mostra que a diferença entre o desempenho de 50% e 75% dos funcionários foi pequena (uma diferença de 0,25, sendo que a mediana foi de 82,5 e o p75 foi de 82,75).

### Salário

**Comando:** *summary(scores$Salario)*

* **Mínimo:** R$ 22,2k.
* **Percentil 25 (primeiro quartil):** R$ 27,8k.
* **Percentil 50 (mediana):** R$ 30,85k.
* **Média:** R$ 31,24k.
* **Percentil 75 (terceiro quartil):** R$ 34,67k.
* **Máximo:** R$ 43k.

**Comando:** *boxplot(scores$Salario)*

****

**Comando:** *hist(scores$Salario, ylab="Frequencia",main="Histograma de salário", xlab="Salário")*

****

**Observações:** Assumindo que a unidade de salário está expressa em milhares de R$ anuais, é possível observar um cenário onde 50% dos funcionários recebem um salário anual de até R$ 30,85k. Outro ponto interessante para analisar é que apenas 25% (5) dos funcionários recebem um salário maior do que R$ 34,67k.

# Correlações

**Comando:** *cor(scores[-3])*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Anos\_de\_Experiencia** | **Escore\_teste** | **Salario** |
| **Anos\_de\_Experiencia** | 1.0000000 | 0.3355080 | 0.8553203 |
| **Escore\_teste** | 0.3355080 | 1.0000000 | 0.5887223 |
| **Salario** | 0.8553203 | 0.5887223 | 1.0000000 |

**Comando:** *cor(scores[-3], method="spearman")*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Anos\_de\_Experiencia** | **Escore\_teste** | **Salario** |
| **Anos\_de\_Experiencia** | 1.0000000 | 0.3872121 | 0.8883546 |
| **Escore\_teste** | 0.3872121 | 1.0000000 | 0.5384036 |
| **Salario** | 0.8883546 | 0.5384036 | 1.0000000 |

**Comando:** *pairs(scores[-3])*

**

**Observações:** Independente do método de correlação (*pearson* ou *spearman*), percebesse que há uma relação forte e positiva entre as variáveis salário e anos de experiência. Outra relação que aparenta ter uma relação moderada se dá entre as variáveis salário e pontuação no teste. Já as variáveis pontuação no teste e anos de experiência aparentam não ter uma relação relevante para análise.

# Programa executado

scores<-read.csv2("/Users/raphaelalbino/Dropbox/doutorado/EAD5917/exercicio 1/Escores.csv")

table(scores$Graduacao)

anos\_graduacao<-table(scores$Graduacao,scores$Anos\_de\_Experiencia)

barplot(anos\_graduacao, main="Graduados por ano de experiencia",

xlab="Numero de funcionarios", col=c("red","darkblue"),

legend = rownames(anos\_graduacao),beside=TRUE)

summary(scores$Anos\_de\_Experiencia)

boxplot(scores$Anos\_de\_Experiencia)

hist(scores$Anos\_de\_Experiencia, ylab="Frequencia",main="Histograma de anos de experiência", xlab="Anos de experiência")

summary(scores$Escore\_teste)

boxplot(scores$Escore\_teste)

hist(scores$Escore\_teste, ylab="Frequencia",main="Histograma de escore", xlab="Escore")

summary(scores$Salario)

boxplot(scores$Salario)

hist(scores$Salario, ylab="Frequencia",main="Histograma de salario", xlab="Salario")

cor(scores[-3])

cor(scores[-3], method="spearman")