**Problem 2: The Hiring Statistics**

An Italian company received 20 résumés from both Italian and foreign candidates during the hiring process for the position of Foreign Relations Manager. Table 2 reports the information considered relevant in the selection process: age, nationality, minimum desired income (in thousands of euros), and years of work experience.

| | Idade | Nacionalidade | Renda | Experiência |
|---|---|---|---|---|
| 1 | 28 | Italiana | 2.3 | 2 |
| 2 | 34 | Inglesa | 1.6 | 8 |
| 3 | 46 | Belga | 1.2 | 21 |
| 4 | 26 | Espanhola | 0.9 | 1 |
| 5 | 37 | Italiana | 2.1 | 15 |
| 6 | 29 | Espanhola | 1.6 | 3 |
| 7 | 51 | Francesa | 1.8 | 28 |
| 8 | 31 | Belga | 1.4 | 5 |
| 9 | 39 | Italiana | 1.2 | 13 |
| 10 | 43 | Italiana | 2.8 | 20 |
| 11 | 58 | Italiana | 3.4 | 32 |
| 12 | 44 | Inglesa | 2.7 | 23 |
| 13 | 25 | Francesa | 1.6 | 1 |
| 14 | 23 | Espanhola | 1.2 | 0 |
| 15 | 52 | Italiana | 1.1 | 29 |
| 16 | 42 | Alemã | 2.5 | 18 |
| 17 | 48 | Francesa | 2.0 | 19 |
| 18 | 33 | Italiana | 1.7 | 7 |
| 19 | 38 | Alemã | 2.1 | 12 |
| 20 | 46 | Italiana | 3.2 | 23 |

- 1. Compute the mean, median, and standard deviation for the variables age, desired income, and years of experience. What can you infer from these values about the typical candidate profile?

In order to calculate the mean, we need to sum all values, and divide them by the number of values. Since Age,Income, and Years are quantitative data, this can be done without problems.

Let's start with **Age**:

The sum of all ages: 28+34+46+26+37+29+51+31+39+43+58+44+25+23+52+42+48+33+38+46 is 773, and the number of candidates is 20, so the mean for the age data is 773/20 = 38.65 .

Now for the income mean:

The incomes are abbreviated from thousand of euros, so 2.3 = 2300 euros. Doing the conversion on all income numbers, the sum is: 2300+1600+1200+900+2100+1600+1800+1400+1200+2800+3400+2700+1600+1
= 38400, dividing by 20, the mean for the income data is **1920** euros.

The Years of experience mean:

Years of experience is also a quantitative data, so the sum is: 2+8+21+1+15+3+28+5+13+20+32+23+1+0+29+1 = 280, dividing by 20, the result for the mean is 14 years.

Now, to calculate the median,

The median is calculated by listing all of the quantitative data in a crescent order, and because of that the value in the middle do not deform from outliers.

For the Age, the ordered list is: 23, 25, 26, 28, 29, 31, 33, 34, 37, 38, 39, 42, 43, 44, 46, 46, 48, 51, 52, 58. Since the total is 20, the median is between the 10th and 11th position, so find this, we sum both values and divide by 2. The result is: 38+39/2 = 38.5 .

The ordered income is: 0.9, 1.1, 1.2, 1.2, 1.2, 1.4, 1.6, 1.6, 1.6, 1.7, 1.8, 2.0, 2.1, 2.1, 2.3, 2.5, 2.7, 2.8, 3.2, 3.4 . The median can be calculated similar to the age, so 1.7+1.8/2 = 1.75 .

At last, the ordered years of experience is: 0, 1, 1, 2, 3, 5, 7, 8, 12, 13, 15, 18, 19, 20, 21, 23, 23, 28, 29, 32. The median is 13+15/2 = 14 .

Finally, to calculate the standard deviation, we continue from the mean. This process involves finding the variance first, which is the average of the squared differences from the mean. The standard deviation is simply the square root of the variance.

Como este é um conjunto de 20 currículos (uma **amostra**), usamos a fórmula do desvio padrão amostral ($s$), que divide por $n - 1$ (ou seja, $20 - 1 = 19$).

A fórmula da Variância Amostral ($s^2$) é:

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n - 1}$$

E o Desvio Padrão Amostral ($s$) é:

$$s = \sqrt{s^2}$$

**Desvio Padrão da Idade:**

- **Média ($\bar{x}$):** 38.65 anos
- **Soma dos Quadrados ($\sum(x_i - \bar{x})^2$):** $(28 - 38.65)^2 + ... + (46 - 38.65)^2 = 1872.55$
- **Variância ($s^2$):** $\frac{1872.55}{19} \approx 98.555$
- **Desvio Padrão ($s$):** $\sqrt{98.555} \approx$ **9.93 anos**

**Desvio Padrão da Renda (em milhares de euros):**

- **Média ($\bar{x}$):** 1.92 (ou 1920 euros)
- **Soma dos Quadrados ($\sum(x_i - \bar{x})^2$):** $(2.3 - 1.92)^2 + ... + (3.2 - 1.92)^2 = 9.672$
- **Variância ($s^2$):** $\frac{9.672}{19} \approx 0.509$
- **Desvio Padrão ($s$):** $\sqrt{0.509} \approx$ **0.713** (ou **713 euros**)

**Desvio Padrão dos Anos de Experiência:**

- **Média ($\bar{x}$):** 14.0 anos
- **Soma dos Quadrados ($\sum(x_i - \bar{x})^2$):** $(2 - 14.0)^2 + ... + (23 - 14.0)^2 = 2004.0$
- **Variância ($s^2$):** $\frac{2004.0}{19} \approx 105.474$
- **Desvio Padrão ($s$):** $\sqrt{105.474} \approx$ **10.27 anos**

**Inferência sobre o perfil do candidato:**

Com base nas medidas de tendência central (Média e Mediana), o perfil típico do candidato é:

- **Idade:** Média de 38.65 anos e Mediana de 38.5 anos.
- **Renda:** Média de 1920 euros e Mediana de 1750 euros.
- **Experiência:** Média e Mediana de 14 anos.

As medianas (38.5 anos, 1750 euros, 14 anos) são provavelmente os melhores indicadores do candidato "típico", pois não são afetadas por valores extremos (como o candidato de 58 anos ou o de 0 anos de experiência).

Os desvios padrão nos dizem sobre a diversidade do grupo:

- **Idade ($s \approx 9.93$ anos):** A dispersão é moderada. Indica que, embora a média seja de 38.65, há uma variedade considerável de candidatos, desde os mais jovens (23) até os mais velhos (58).
- **Renda ($s \approx 713$ euros):** A dispersão é relativamente baixa em comparação com a média (1920 euros). Isso sugere que a maioria dos candidatos tem expectativas salariais semelhantes, girando em torno de 1750-1920 euros.
- **Experiência ($s \approx 10.27$ anos):** Este é um desvio padrão *muito alto*, especialmente quando comparado à média de 14 anos. Isso revela que o grupo é extremamente heterogêneo em termos de experiência, contendo tanto candidatos em início de carreira (0, 1, 2 anos) quanto profissionais muito seniores (28, 29, 32 anos).

**Conclusão:** O perfil "típico" é um profissional de meia-idade (38.5) e meio de carreira (14 anos), com uma expectativa salarial de 1750 euros.

- 2. Group the candidates by nationality and compute the average desired income and average years of experience for each group. Which nationality has the highest average desired income? Which group appears most experienced?

In total we have 6 different nationalities: Italian, English, Belgic, Spanish, French and German.The 8 **Italians** candidates have an average desired income of (2.3+2.1+1.2+2.8+3.4+1.1+1.7+3.2)/8 = 17.8/8 = **2.225** or 2225 euros, and the average experience is: (2+15+13+20+32+29+7+23)/8 = 141/8 = **17.65 years**.
For the two **English** candidates, the mean desired income are (1.6+2.7)/2 = 4.3/2 = **2.15** or 2150 euros, with an average experience of (8+23)/2 = 31/2 = **15.5 years**.
The two **Belgian** candidates have an average income of (1.2+1.4)/2 = 2.6/2 = **1.3** or 1300 euros,

and an average experience of (21+5)/2 = 26/2 = **13 years**.
For the three **Spanish** candidates, the average income is (0.9+1.6+1.2)/3 = 3.7/3 ≈ **1.233** or 1233 euros, and the average experience is (1+3+0)/3 = 4/3 ≈ **1.33 years**.
The three **French** candidates have an average income of (1.8+1.6+2.0)/3 = 5.4/3 = **1.8** or 1800 euros, and an average experience of (28+1+19)/3 = 48/3 = **16 years**.
Finally, the two **German** candidates have an average income of (2.5+2.1)/2 = 4.6/2 = **2.3** or 2300 euros, with an average experience of (18+12)/2 = 30/2 = **15 years**.
According to the calculations, the **German** candidates have the highest average desired income at 2300 euros, followed very closely by the Italians (2225 euros).