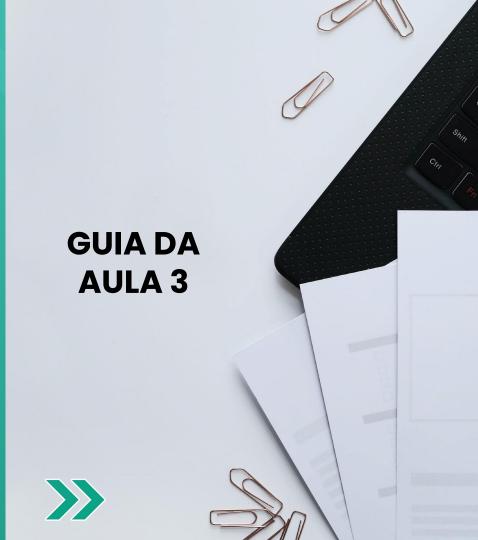


Profissão: Analista de dados





FUNDAMENTOS DE ESTATÍSTICA







Faça a correlação

- Introdução
- Coeficiente de Correlação de Pearson



Acompanhe aqui os temas que serão tratados na videoaula







Introdução

Correlação são métricas que medem a dependência estatística entre conjuntos de dados, correlações estas que podem ser causais ou não.

Exemplo:

Sobe o preço do dólar, sobe o preço de insumos importados, sobe o preço dos produtos, logo, sobe a inflação.





Coeficiente de Correlação de Pearson

O coeficiente de correlação de Pearson (\$\textbf{r}_{xy}\$) resume a correlação linear entre dois conjuntos de dados em um único número entre -1 e 1,, sendo que:

- \$\textbf{r}_{xy} < 0 \$, enquanto x cresce, y decresce;
- \$\textbf{r}_{xy} = 0 \$, n\u00e4o h\u00e4 rela\u00e4\u00f4o entre x e y;
- \$\textbf{r}_{xy} > 0 \$, enquanto x cresce, y cresce.

```
\textbf{r}_{xy} = \frac{\sum_{i=1}^{n} (x_i-x_m)(y_i-y_m)}{\sqrt{\sum_{i=1}^{n} (x_i-x_m)^{2}} \sqrt{n} (y_i-y_m)^{2}} (y_i-y_m)^{2}}
```





Exemplo:

```
In [ ]:
         %%writefile
         nba.csv
         height; weight; wage
         2.01;86.2;17150000
         1.93;106.1;898310
         2.11;120.2;9881598
         1.88;85.7;15643750
         1.88;84.8;2875000
         2.11;106.1;2376840
         1.98;86.6;2625717
         2.08;104.3;37199000
         2.03;117.9;28942830
         1.83;81.6;522738
In [ ]:
         df = pd.read csv('nba.csv', sep=';')
```









• Peso e altura





Peso e salário





Altura e salário

