



AceLeraDevXP

Data Science

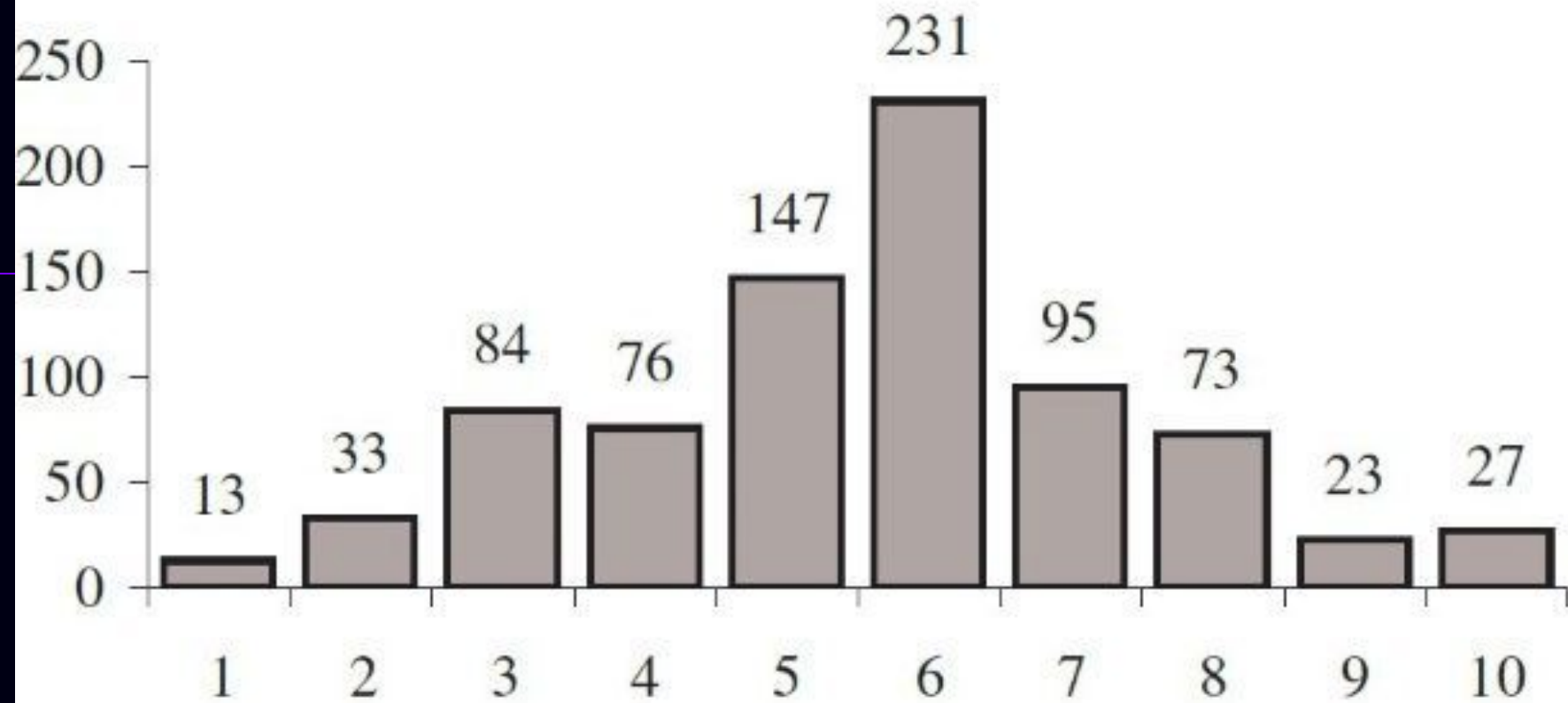
Análise de dados exploratória



Visualizações

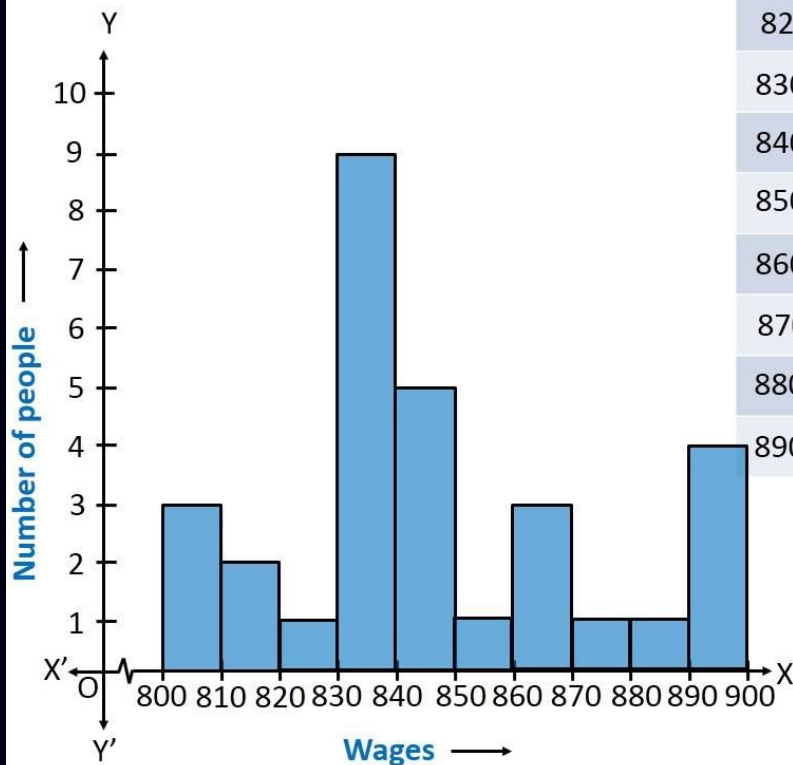
- Histograma

- Distribuição de frequências
- Representação gráfica em colunas ou em barras
- Dividido em classes uniformes ou não uniformes
- A base de cada retângulo representa uma classe
- A altura de cada retângulo
 - Quantidade
 - Frequência absoluta com que o valor da classe ocorre no conjunto de dados para classes uniformes
 - Densidade de frequência para classes não uniformes.



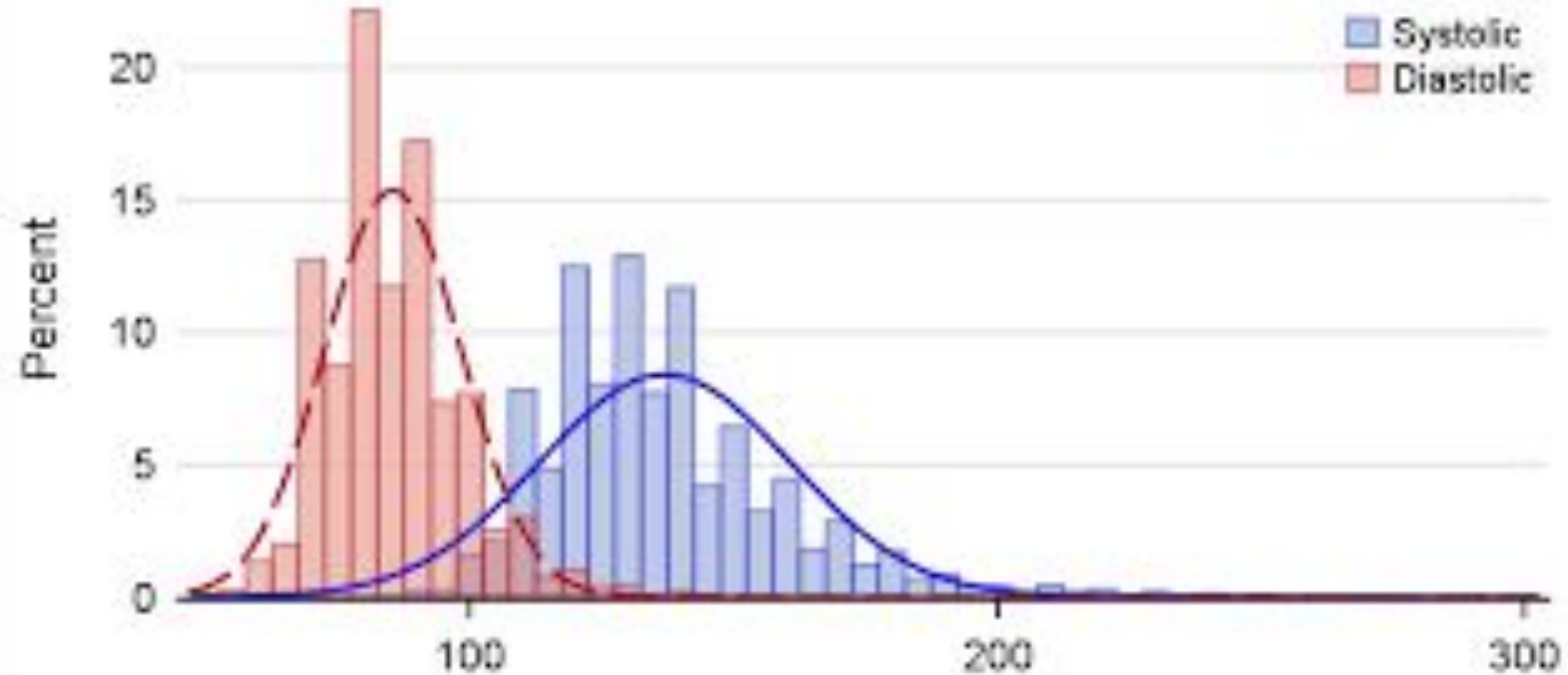
Ex 5.1, 4

Draw a histogram for the frequency table made for the data in question 3, and answer the following questions.



Wages	Frequency
800 – 810	3
810 – 820	2
820 – 830	1
830 – 840	9
840 – 850	5
850 – 860	1
860 – 870	3
870 – 880	1
880 – 890	1
890 – 900	4

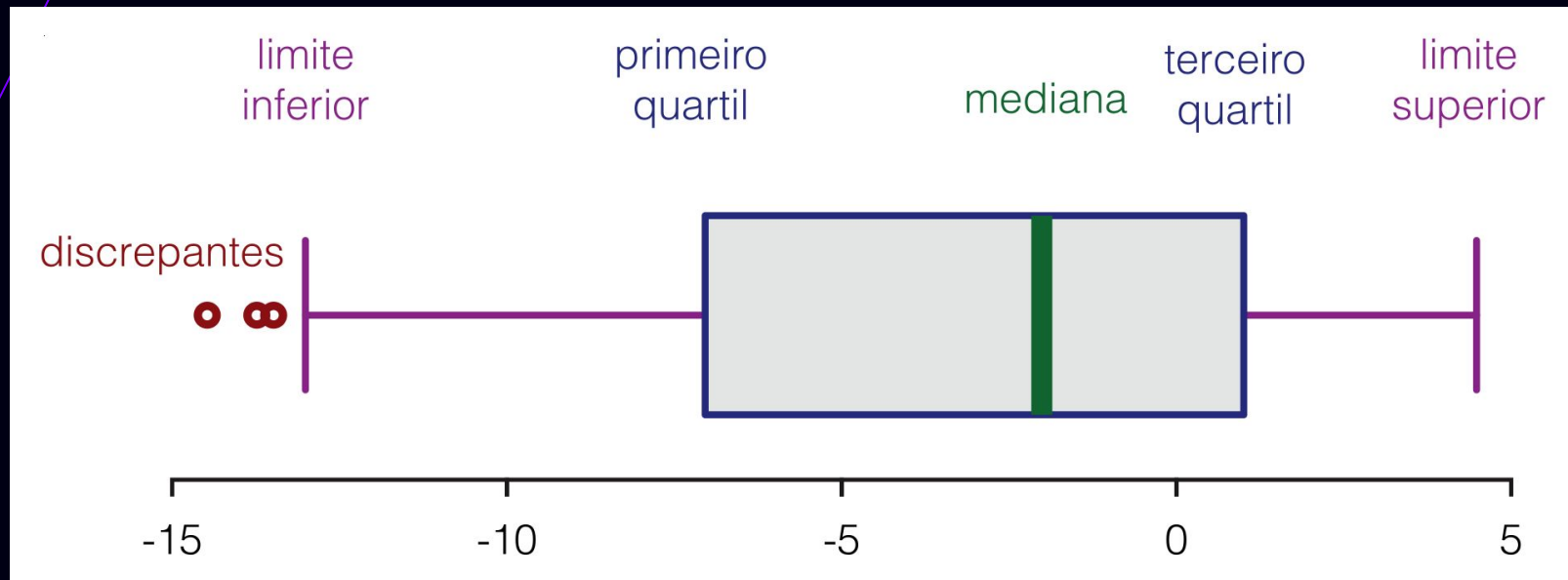
Distribution of Blood Pressure

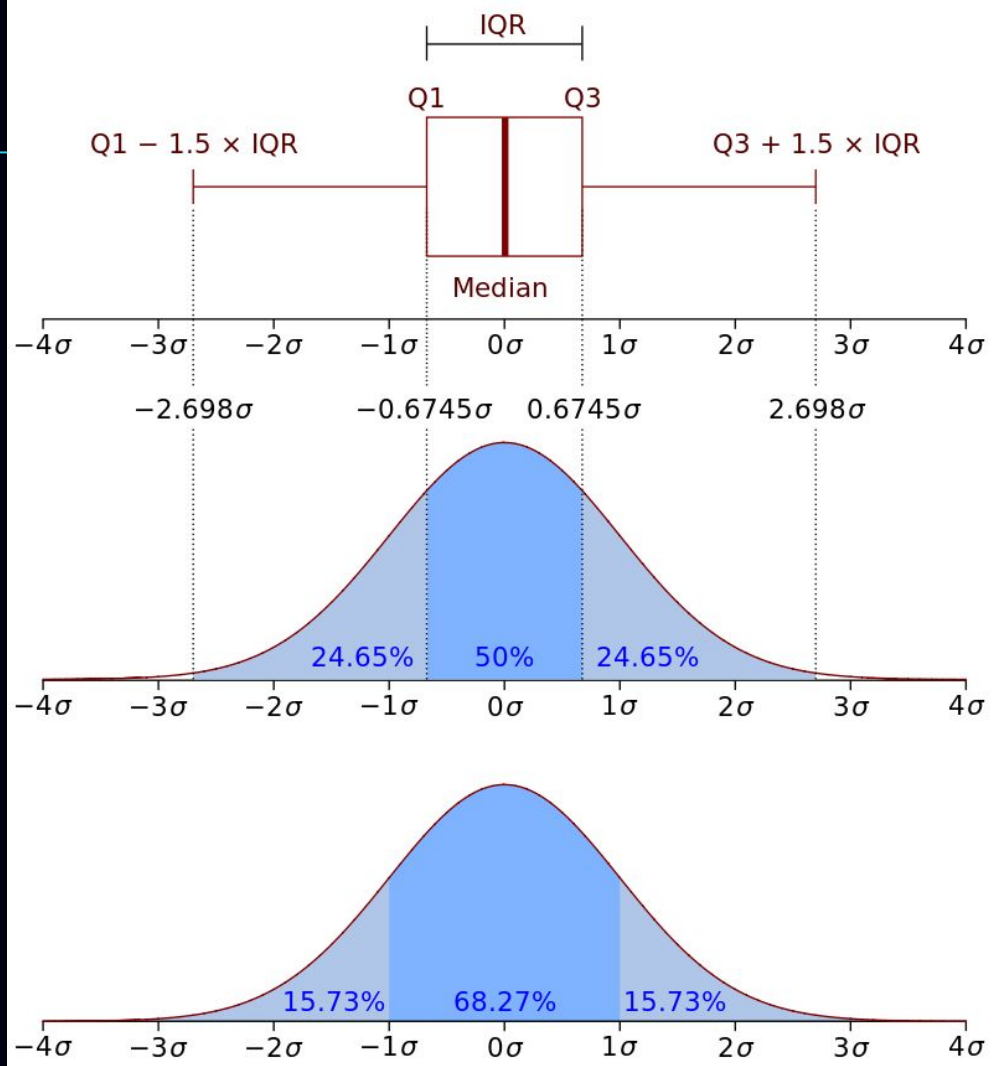


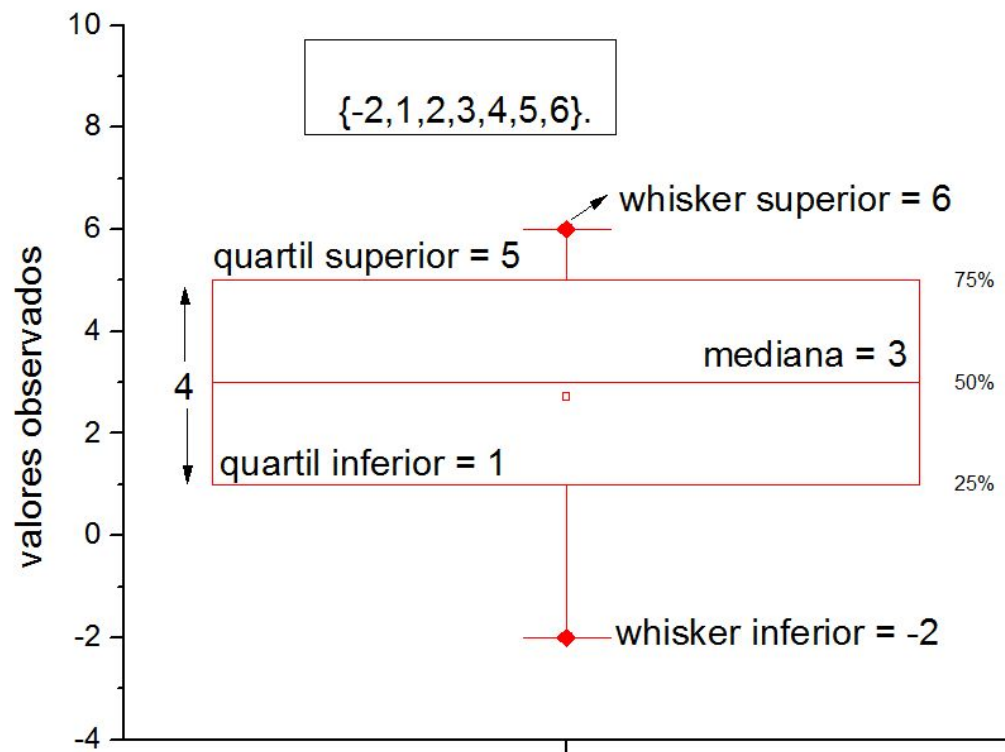
Visualizações

/

- Diagrama de Caixa (Box plot)
 - Ferramenta gráfica para representar a variação de dados observados de uma variável numérica por meio de quartis;
 - Reta (whisker ou fio de bigode) que estende-se verticalmente ou horizontalmente a partir da caixa, indicando a variabilidade fora do quartil superior e do quartil inferior.
 - Outliers ficam como pontos individuais





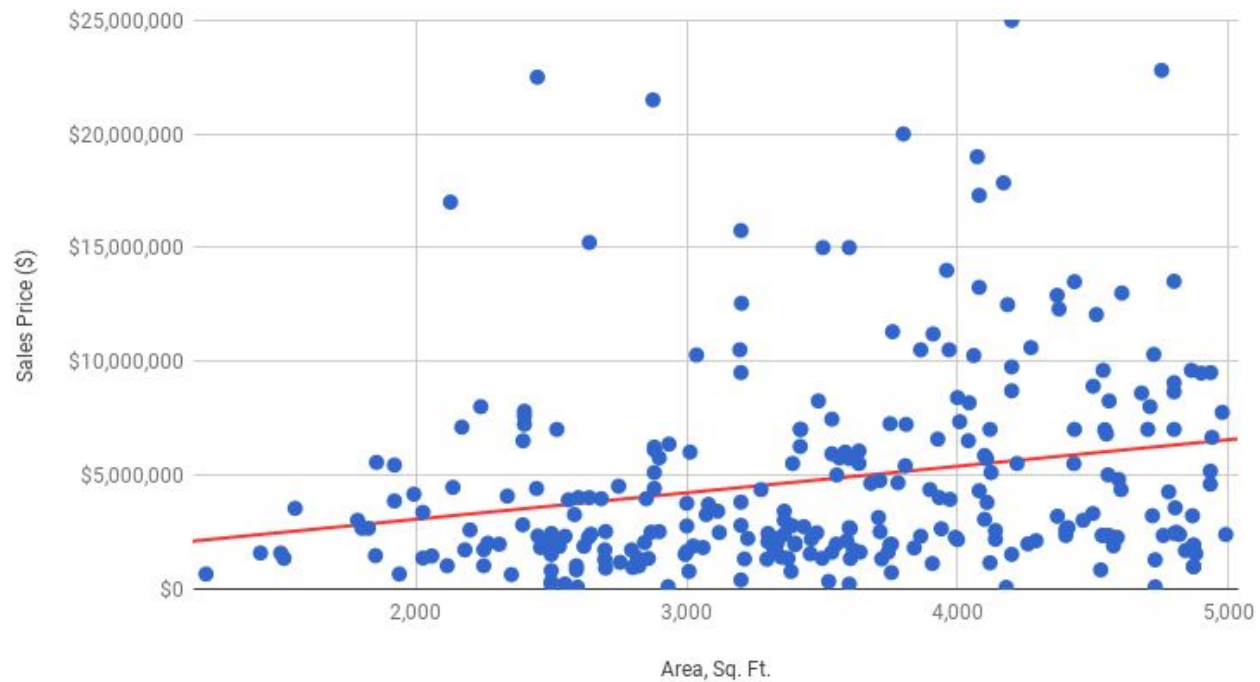


Visualizações

/

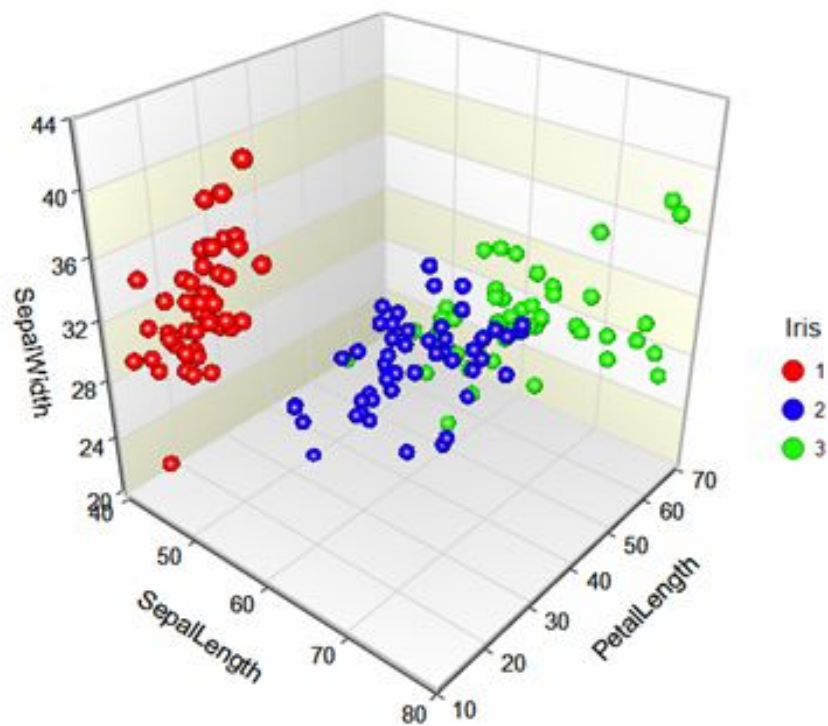
- Dispersão (Scatter plot)
 - Representa duas (tipicamente) ou mais variáveis
 - Utiliza coordenadas cartesianas para exibir valores de um conjunto de dados
 - São exibidos como uma coleção de pontos, cada um com o valor de uma variável

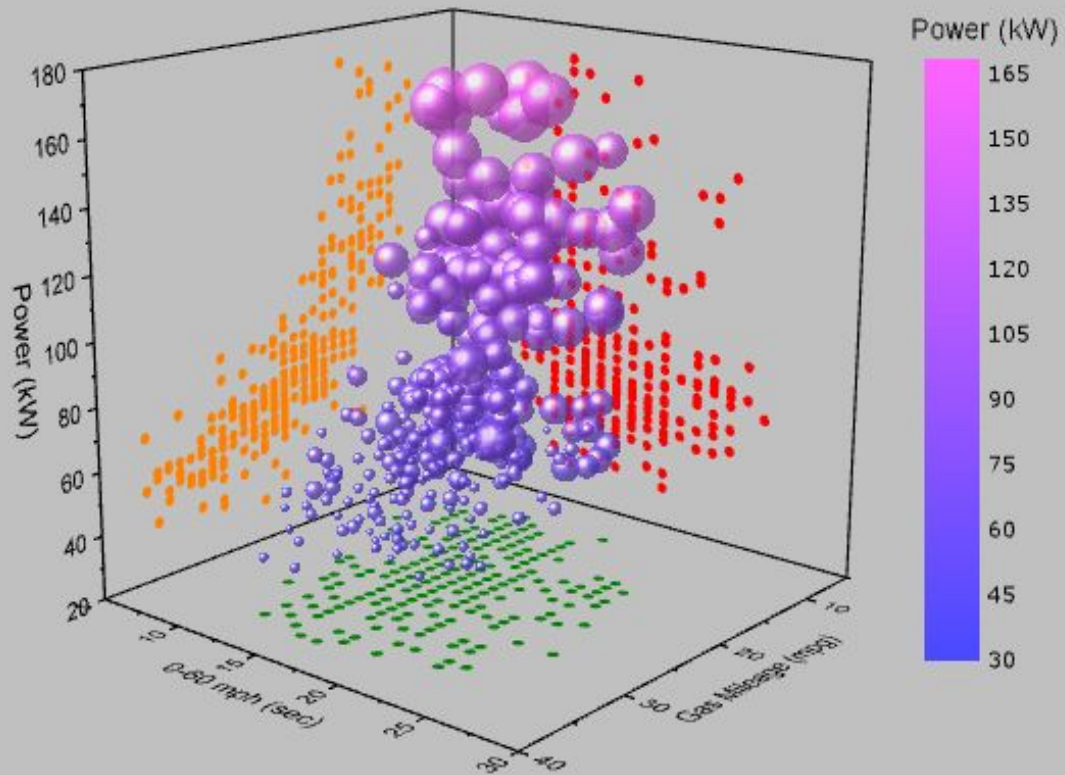
Scatterplot Example



SepalWidth vs. SepalLength vs. PetalLength

By Iris





Symbol Size proportional to Engine Displacement

Visualizações

/

- Matriz de Correlação
 - Tabela
 - Correlação entre valores de um conjunto de dados
 - Cada linha é cruzada com outra para obter o valores de correlação

