

Gerando gráficos

Transcrição

Anteriormente, aprendemos a validar nossos clusters por meio de diferentes métricas e fórmulas matemáticas. É chegado o momento de interpretarmos o resultado dos clusters, nosso objetivo principal.

Precisamos primeiramente encontrar uma maneira de visualizar esses clusters. Temos 16 atributos, isto é, 16 dimensões. Não podemos inserir tudo isso em um gráfico, portanto começaremos por analisar dois atributos: purchase payments.

Utilizaremos uma biblioteca chamada `matplotlib.pyplot`. Em seguida utilizaremos o `scatter`, que nos fornece o gráfico de pontos, com o mesmo dataframe que utilizamos desde o início. Teremos, ainda, as variáveis `s` que definem o tamanho dos pontos e a `cmap`, que define o esquema de cores. Por fim, atribuiremos novos nomes aos atributos, de forma que eles se tornem mais facilmente identificáveis.

```
import matplotlib.pyplot as plt
plt.scatter(dataframe["PUERCHASES"], dataframe['PAYMENTS'], c=labels, s=5, cmap='rainbow' )
plt.xlabel("valor total pago")
plt.ylabel("valor total gasto")
plt.show
```

Como resultado da compilação, teremos o seguinte gráfico com duas dimensões:

