
Fundamentos de Ciência de Dados, Big Data e Inteligência Artificial - Visualização de Dados

Visualizações de dados para o *dataset* Mercado:

Nele tínhamos apenas 5 colunas, porém uma era de index então pouco provável de utilizar em alguma análise, com 4 colunas a única análise que consegui pensar era saber quanto do estoque poderia gerar lucro, quanto ele já tomou prejuízo e quais produtos estão prestes a dar prejuízo e para ver quais produtos estão em cada situação analisei a data de validade a quantidade em estoque e seu valor.

Ambos os gráficos no fim servem para esta mesma análise, o gráfico pizza podemos observar o valor total que poderia ser obtido com as vendas, e a divisão de como está se ele pode lucrar, quanto já perdeu e quanto está para perder, nele podemos ver os dados principais separados por cores e os valores em porcentagem, porém ao colocar o mouse em cima podemos ver o valor que representa essa porcentagem para uma análise mais precisa. Neste ponto me perguntei e se eu quisesse saber quais são os produtos que estão em cada grupo, a primeira solução que pensei, e deixei implementada, era uma lista com os produtos ao dar *hover* em algum grupo do gráfico, porém não achei eficiente e decidi criar o segundo gráfico.

No segundo gráfico mostro apenas os produtos com quantidade em estoque, nele mostro quais os produtos pertencem a cada grupo usando as mesmas cores do gráfico 1, além da quantidade de produtos, para saber além de qual produto gerou prejuízo, qual precisa vender rapidamente, e quais ainda tem tempo.

O primeiro gráfico escolhi o de pizza pois tinha um valor total e queria mostrar em quantas partes ele se dividia, então esse tipo de gráfico seria o que melhor demonstraria isso diretamente. Já para o segundo gráfico optei por barras, pois o agrupamento ficaria claro nas cores, e tínhamos duas variáveis simples a serem analisadas, o nome do produto em um eixo e a quantidade em estoque

no outro, como um dos eixos era o nome do produto optei pela orientação horizontal das barras para facilitar a leitura.

Visualizações para o *dataset* Taxis:

Aqui escolhi fazer um *dataset* disponível na biblioteca *seaborn* chamado Taxis, e optei por fazer um dashboard contendo 4 gráficos e 2 filtros que considero útil para análise de todos os gráficos. Sobre os filtros, o primeiro é um filtro *multiseletor* que permite escolher quatro períodos do dia diferentes e o segundo permite escolher um range de datas.

O primeiro gráfico é um de linha pois temos uma variável contínua que são as horas do dia no eixo x, e no eixo y é possível ver claramente a separação entre os diferentes bairros. Neste gráfico podemos analisar o volume de corridas por hora do dia e região, então um exemplo de uso seria caso queremos saber qual bairro tem mais chamadas de corrida as três da tarde por exemplo, podemos utilizar o filtro de data para acompanhar esse valor por semana e tirar análises ainda mais precisas.

O segundo gráfico é por dispersão com linha de tendência, aqui escolhi este tipo de gráfico por ter duas variáveis quantitativas e alto volume de dados, separei por cores os períodos do dia, e nele que aplico a linha de tendência. A análise obtida é relação da tarifa por distância, com a linha de tendência podemos observar qual período do dia corridas mais distantes valem mais a pena.

O Terceiro gráfico é de barras, escolhido por se tratar de uma variável quantitativa e outra descritiva a qual optei por colocar no eixo y e deixar a orientação como vertical, neste gráfico também optei por deixar agrupamentos dos períodos do dia nas barras e elas não sobrepostas, para deixar claro a análise, aqui podemos a princípio concluir qual bairro tem maior quantidade de corridas, e juntando o período do dia saber a maior quantidade de corrida por período.

O Último gráfico é por barras pelo mesmo motivo do anterior, porém optei por deixar orientação na vertical mesmo com a variável descritiva neste eixo, pois como eram muitos valores caso essa variável ficasse no eixo y o gráfico seria longo demais e complicaria sua leitura. Neste gráfico podemos observar a quantidade de passageiros por zona da cidade.