

Nome: Alexandre Porto do Nascimento

Matrícula: 220083130

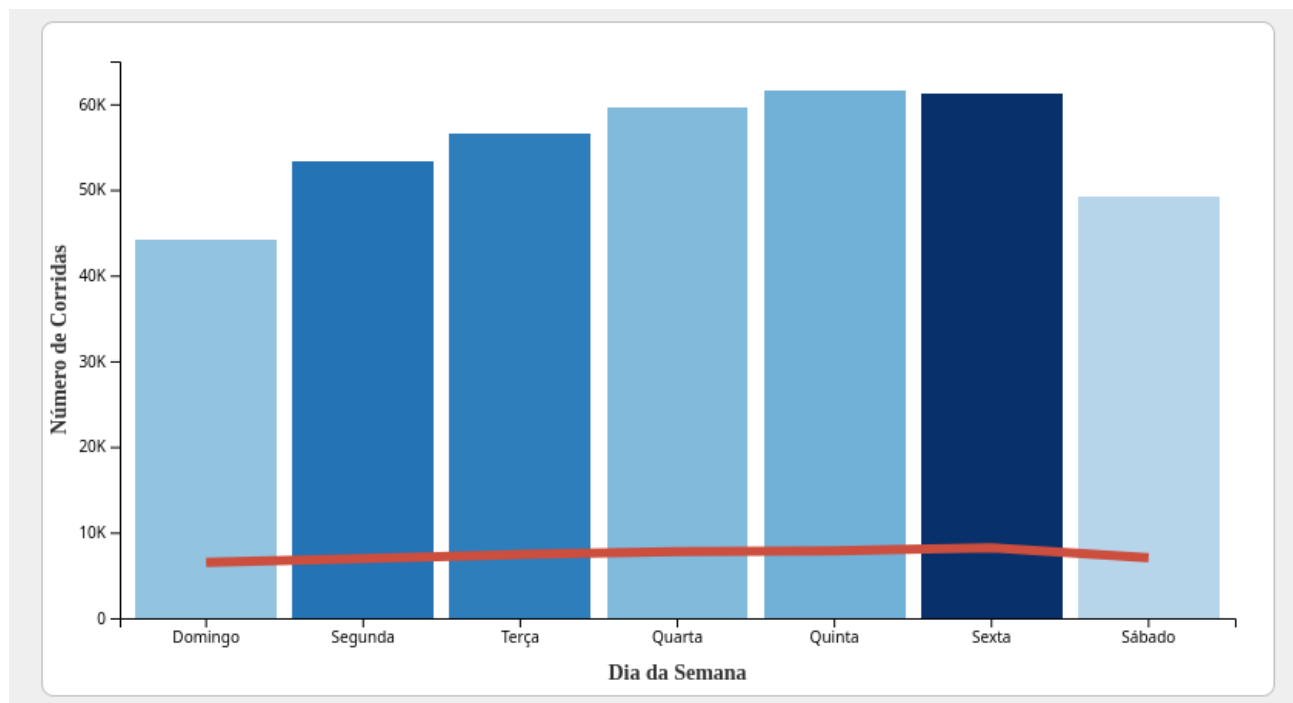
Trabalho 01:

Questão 1:

Existe alguma diferença entre o padrão de corridas de táxi durante os dias de semana e os dias de fim de semana?

Para responder a essa pergunta, foi desenvolvido um gráfico de barras, em que o eixo X estão os dados categóricos (dias da semana) e o eixo Y estão os dados ordinais (número de corridas totais).

A partir desse aspecto, é possível perceber claramente que o número de corridas nos fim de semana caem drasticamente em relação aos dias de semana, que possuem um pico na quinta-feira e sexta-feira. Os números chegam a variar em torno de 10.000 corridas.



Menos engarrafamento

Mais engarrafamento

Também há no gráfico uma linha laranja que representa a quantidade de viagens com mais de 1 passageiro. Ela se mantém bem estável durante toda a semana com uma leve queda no sábado e domingo, mas mantendo a proporção em relação a queda do número de viagens totais. Não sendo um fator que varia entre dias úteis e não úteis.

O ultimo aspecto que foi trabalhado nessa visualização foi o quão engarrafado costuma ser o dia da semana. Para isso foi feito um cálculo relacionando o momento de entrada do passageiro no táxi, o horário de saída e a distância que o veículo percorreu. Com isso, recebemos a velocidade média do carro durante toda a viagem. A partir disso, é somado todas as médias e divide-se pela quantidade de viagens no dia. Quanto mais rápido foi a média de velocidade das corridas, menos engarrafado foi o dia.

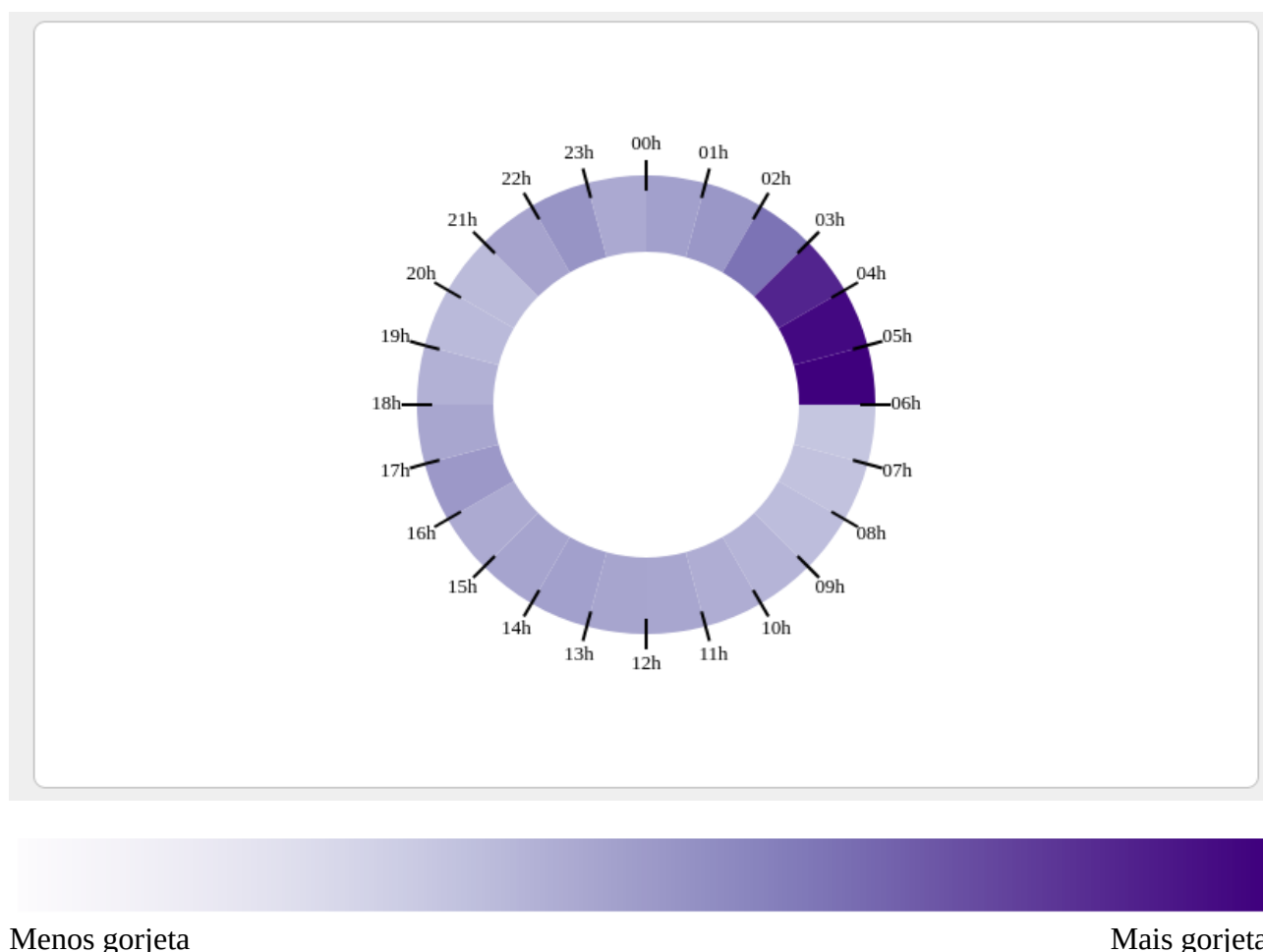
O gradiente de cores deixa bem claro que os dias de semana são bem mais engarrafados que os fins de semana. Com picos na segunda, terça e principalmente sexta-feira. Mas em um geral, pode-se esperar ruas mais vazias nos fins de semana.

Essa visualização foi escolhida pois os marcadores mais efetivos (as barras verticais) retratam o aspecto mais importante a ser tratado, que são a quantidade total de corridas. Pelo atributo “quantidade de engarrafamento” não ser facilmente medido com números objetivos e ser necessário muitos cálculos para chegar a ele, foi escolhido um gradiente de cores, que conseguem expressar bem a variabilidade dos dados, sem se prender a características ordinais fixas. O atributo menos expressivo da visualização (quantidade de corridas com mais de 1 passageiro), que não apresenta variação significativa entre os dias da semana, ficou com menor destaque, sendo uma linha e os valores são definidor por sua posição no gráfico e a variação entre os dias através da inclinação da linha.

Com todos esses aspectos trabalhados, temos uma visualização bem expressiva, que mostra todas as informações do conjunto de dados desejado e com canais visuais precisos, com boa separabilidade e com o efeito pop-out para os elementos mais importantes.

Questão 2:

Existe alguma relação entre o valor da gorjeta (tip_amout) e o horário das corridas?



Para responder essa pergunta, foi desenvolvido um gráfico em anel, dividido em 24 arcos, cada uma representando uma hora do dia. E também uma escala de cores sequencial, onde a cor roxa mais escura representa uma média maior de gorjeta para o horário definido.

Essa representação foi escolhida por dois principais motivos. O primeiro deles está relacionado ao fato de se tratar de uma representação usando horas do dia. A semelhança com um relógio faz da leitura dela mais intuitiva e o fato dos horários estarem colados uns aos outros torna bem perceptível se há uma mudança brusca na média de gorjetas da hora em relação as adjacentes.

Por se tratar de uma média para o horário, não há necessidade de representar valores absolutos ao se tratar da gorjeta, com isso foi escolhido a escala sequencial de cores, que mostra bem a tendência de gorjeta do horário.

Os marcadores são classificados pela área, representando os itens do horário. Já os canais visuais são as cores, que apesar de não ser o atributo mais efetivo para representar valores específicos, funciona bem para ter uma noção geral de que em determinado horário as gorjetas são maiores.

O gráfico deixa bem claro que as gorjetas não costumam variar muito durante todo o dia. Porém na alta madrugada, a entre 2h e 6h da manhã, as gorjetas começam a ser significativamente mais altas em relação aos outros horários, com um pico entre 5h e 6h da manhã.