

# TP3 - Análise de Grafos com Spark e GraphX

Os arquivos disponibilizados constroem um grafo que representa estações de aluguel de bicicletas e onde elas foram alugadas e devolvidas.

Você deve criar um script e submete-lo ao spark-submit.

1 - Importe os arquivos que se encontram em:

<https://github.com/leoinfnet/aulas-spark/tree/master/tp3>

station.csv contem as estações e trips.csv contém as informações sobre retirada e entrega das bicicletas.

2 - Trate os arquivos da maneira que achar necessário.

3 - Crie o grafo no spark;

4 - Imprima:

Total De Estações: X

Total de Rotas: Y

Onde X é numero de vertices do grafo e Y seu número de arestas.

5 - Imprima estação com maior numero de saídas

6 - Imprima estação com menor numero de saídas

7 - Imprima estação com maior numero de entradas

8 - Imprima estação com menor numero de entradas

9- Crie uma novo dataframe com os campos: nome da estação e DegreeRatio , onde:

DegreeRatio = *inDegree/outDegree* de cada estação

10 - Exiba a estação com maior Degree e Menor Degree

11 - Calcule a estação com maior e menor PageRank.

12 - Qual o menor caminho(Se existir entre as estações)

**San Jose Diridon Caltrain Station e Powell Street BART**

13 - Qual(is) os menores caminhos distintos entre:

**Clay at Battery e Redwood City Caltrain Station**

14 - Plote em um mapa todas as estações.

15 - Plote em um mapa o caminho encontrado na questão 12.