We obtained our dataset from Kaggle website. It is a CSV file with compiled data set from Indeed website. This dataset is 9 months old and is pre-cleaned. It has over 5000 rows and few columns with relevant column names for our work, which are Job Type, Location, Salary, Type of skill, Number of skills and Industry.

1. Job Type was categorized into Data Analyst, Data Scientist, and Data Engineer.

2. Location was different sates within the USA.

3. Type of Skills varied from Python and SQL to Machine learning.

4. Industry: We chose top 5 industries which were: Healthcare, Business/Consulting, Software /IT, Finance /Insurance.

From our data analysis, we found that highest number of Data Science jobs were available in California, followed by New York, Virginia, Texas and then Massachusetts. Relatively high number of Data Science jobs were available in California, where as in Texas, the availability all these job types were comparatively similar.

For a specific skill, the number of jobs is significantly different among 3 job types. Python and SQL are essential skills required by all 3 job types. R is needed for both data analyst and data scientist which also requires tableau and machine learning respectively. Whereas, Data engineer needs Hadoop, spark and java.

For a specific skill, the number of jobs is significantly different in salary categories. The need for python, spark, java, Hadoop and machine learning increases as the salary increasing; while the need for SQL, r, tableau and SAS stays stable across all salary categories.