

Data Science Job Opportunity Analysis in the United States

- MS in Business Analytics Capstone Project

Overview

Data Cleaning

EDA

Text Analysis

Applications

Works Cited

About Us

Project Goal

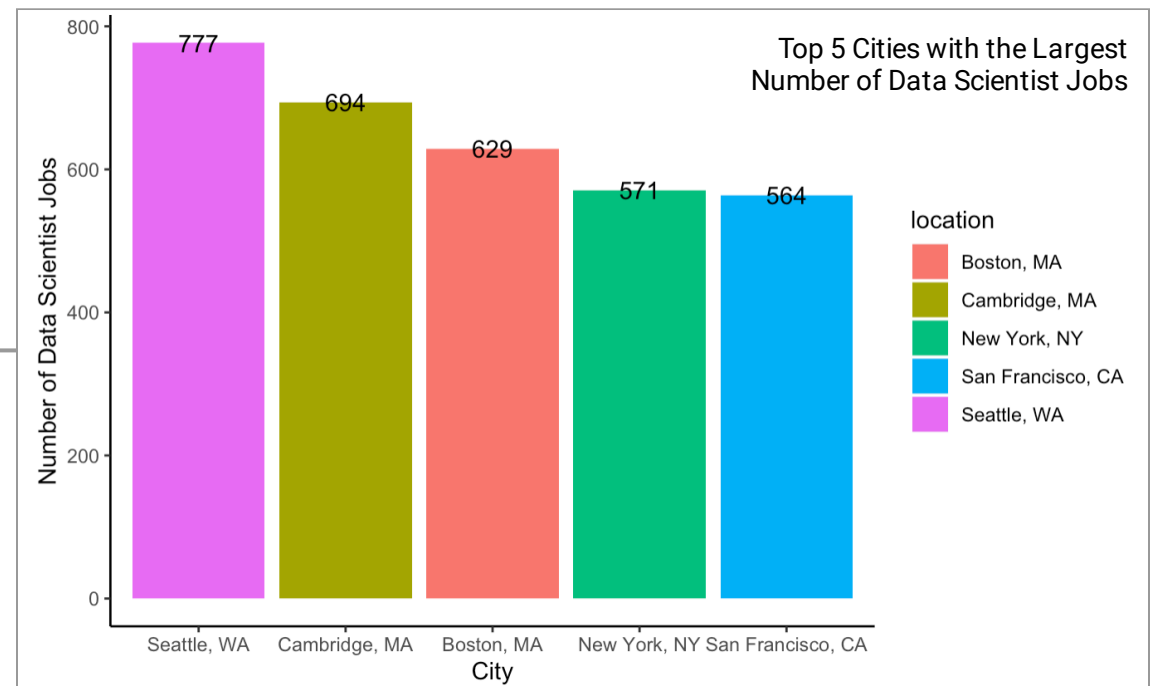
We suppose all students in MSBA program want to get a job after the graduation in order to apply what we learned at school to the real business environment.

Here are some business problems we want to solve:

1. What type of talents do employers want concerning tools, skills, degrees and majors?
2. What is the difference between different job roles(e.g. Engineer vs. Data Scientist vs. Analyst) in the data science job market?
3. How much would you make as a data scientist in Seattle on average?

Data Sources

1. Data Scientist Job Market in the U.S. (Updated in 2019)
2. City of Seattle Wage Data: Comparison by Gender (Updated in 01/2020)



Since Seattle is the top data scientist job market in the U.S., we would love to explore possibilities of working in this city.

Data Science Job Opportunity Analysis in the United States

- MS in Business Analytics Capstone Project

Overview

Data Cleaning

EDA

Text Analysis

Applications

Works Cited

About Us

Data Summary

1. Data Scientist Job Market in the U.S. (Updated in 2019)

This dataset is from the Kaggle.com website. In this dataset, we have in total of 6964 rows and 5 columns. Columns include information such as position, company, description, reviews, and location. "

2. City of Seattle Wage Data: Comparison by Gender (Updated in 01/2020)

Comparison by Gender" dataset is from the Data.gov website. We have in total of 759 rows and 10 columns. Columns include information such as job classification, female and male average hourly rate, number of female and male employed, total average hourly rate and total number of people employed.

Data Cleaning

1. Re-format the column names for wage data

```
colnames(wage)[colnames(wage)=="Job Classification"] <- "Job_Classification"
colnames(wage)[colnames(wage)=="Total Avg Hrly Rate"] <- "Total_Avg_Hrly_Rate"
wage2 = wage[grepl("Analyst", wage$Job_Classification),]
```

2. Re-format the location column and add a new column called 'state' in all data.

```
city_trim <- function(s) gsub('([A-z]+, [A-Z]{2})\\ [0-9]{5}', '\\1', s)
all$location <- sapply(all$location, city_trim)
state_extract <- function(s) gsub('^[A-z ]+, ([A-Z]{2}).*', '\\2', s)
all$state <- sapply(all$location, state_extract)
```

Since the location column in the original dataset include both the city and state in one column. We decided to split the strings and save 'states' to a new column. Now, we have 6964 rows and 4 columns.

Data Science Job Opportunity Analysis in the United States

- MS in Business Analytics Capstone Project

Overview

Data Cleaning

EDA

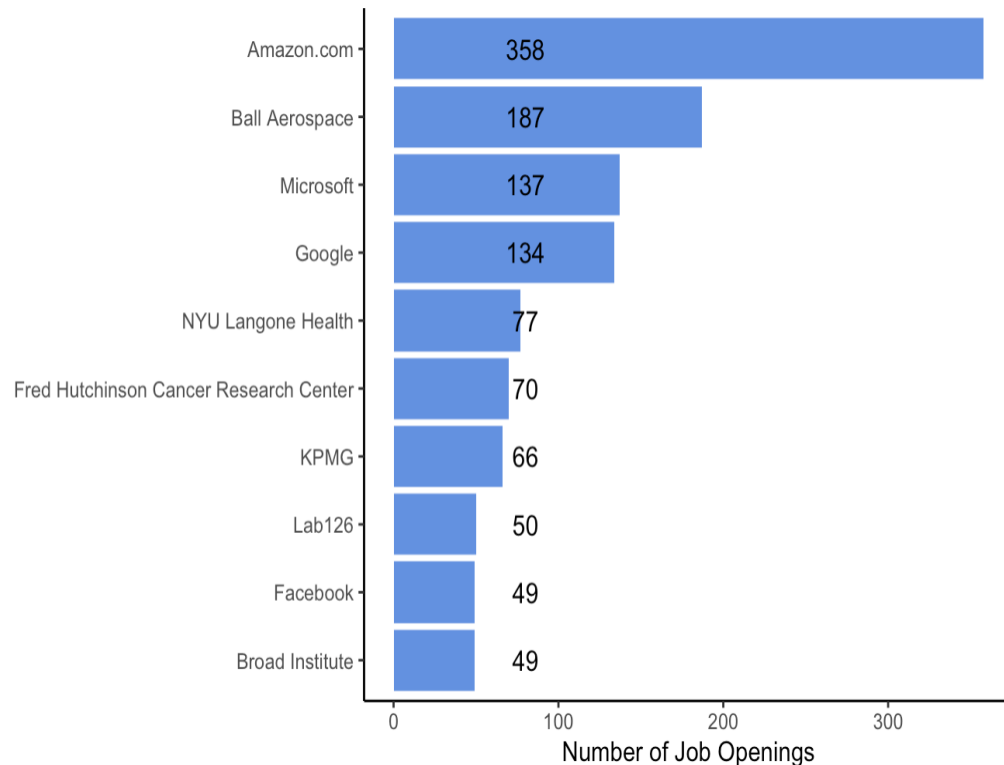
Text Analysis

Applications

Works Cited

About Us

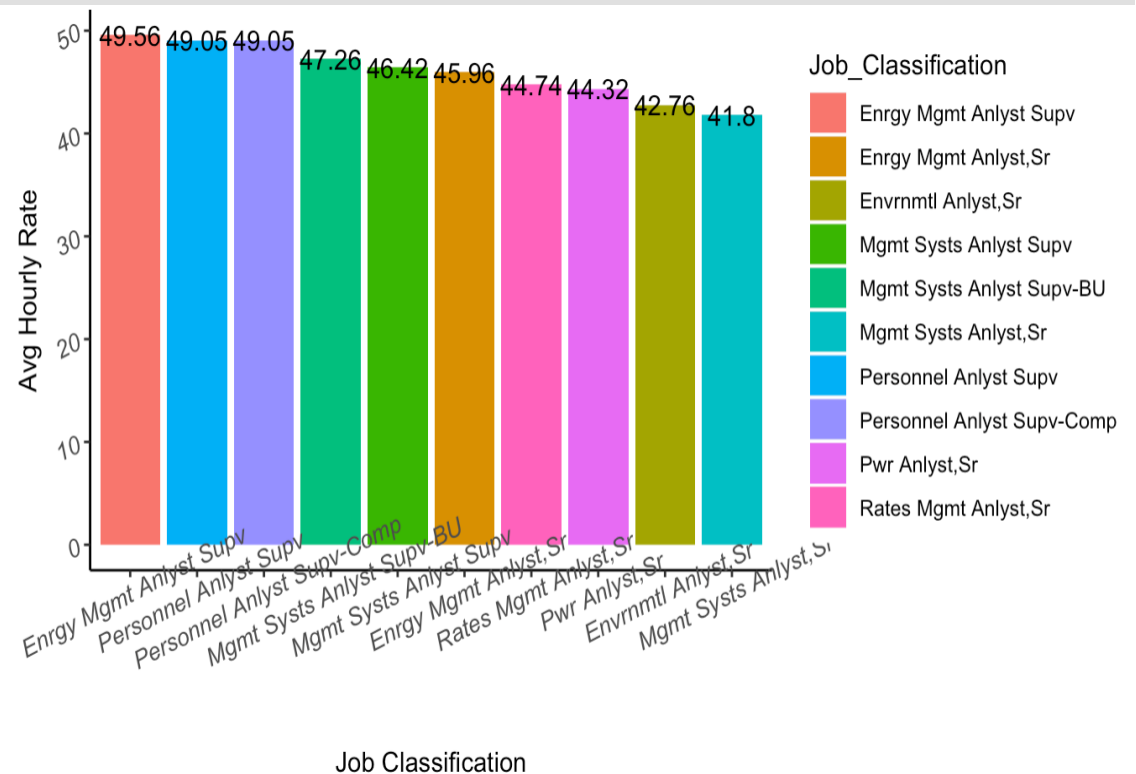
Companies with Most Data Science Job Openings



Analysis Insight

Companies who are actively hiring data scientists and/or data analysts are mostly based in Seattle, WA.

Top 10 analyst hourly wages



Analysis Insight

Energy and Management Systems are top fields for data analysts in terms of hourly wages.

Data Science Job Opportunity Analysis in the United States

- MS in Business Analytics Capstone Project

Overview

Data Cleaning

EDA

Text Analysis

Applications

Works Cited

About Us

Data Science Job Opportunity Analysis in the United States

- MS in Business Analytics Capstone Project

Overview

Data Cleaning

EDA

Text Analysis

Applications

Works Cited

About Us

Data Science Job Opportunity Analysis in the United States

- MS in Business Analytics Capstone Project

Overview

Data Cleaning

EDA

Text Analysis

Applications

Works Cited

About Us

1. <https://www.kaggle.com/sl6149/data-scientist-job-market-in-the-us#alldata.csv>
2. <https://catalog.data.gov/dataset/city-of-seattle-wages-comparison-by-gender-wage-progression-job-titles-417b2>

Data Science Job Opportunity Analysis in the United States

- MS in Business Analytics Capstone Project

Overview

Data Cleaning

EDA

Text Analysis

Applications

Works Cited

About Us

Hello Everyone!

Thanks for your interest in our project!

We are a group of Masters students who are interested in the field of Data Science. We understand the challenging process of job search and hope that our project can help aspiring data scientists and data analysts to ease their job searching experience with more targeted search plan.

If you have any question or comment regarding the project, please do not hesitate to reach out to us at questromMSBA@bu.edu.

MSBA Class of 2020
Cohort B Team 3