

Data Science Job Market

...

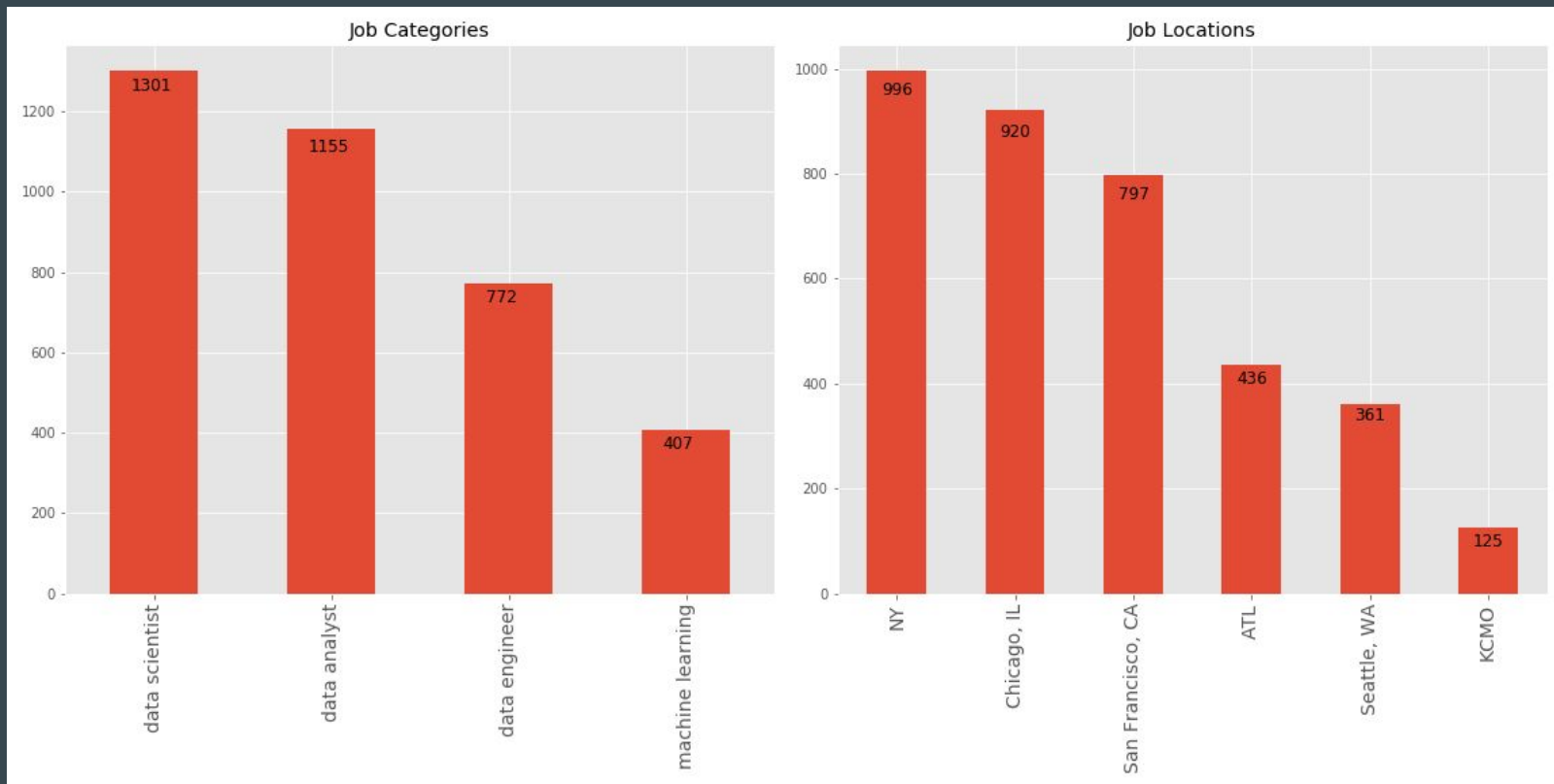
Objectives

- Discern difference between data scientist vs data analyst
 - Junior vs senior roles
- Explore any regional variations
- Make recommendations to various departments to better align Flatiron offerings and experience with job market

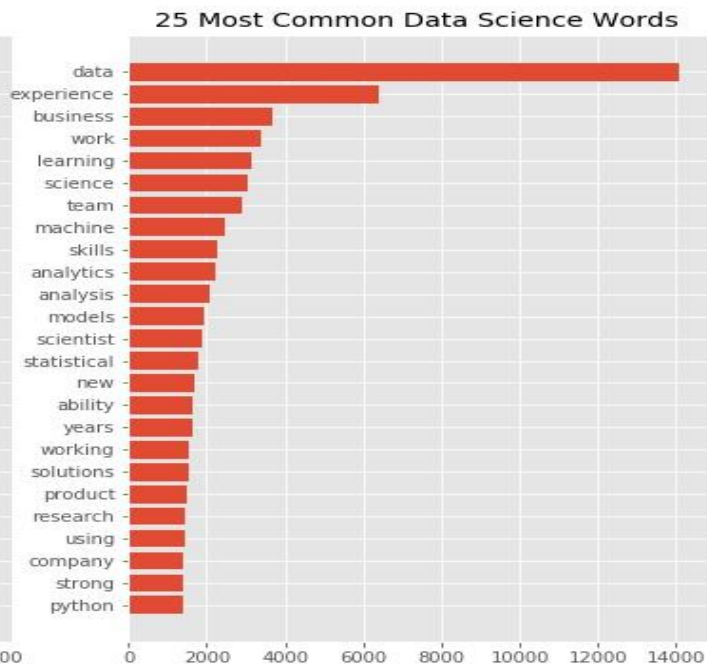
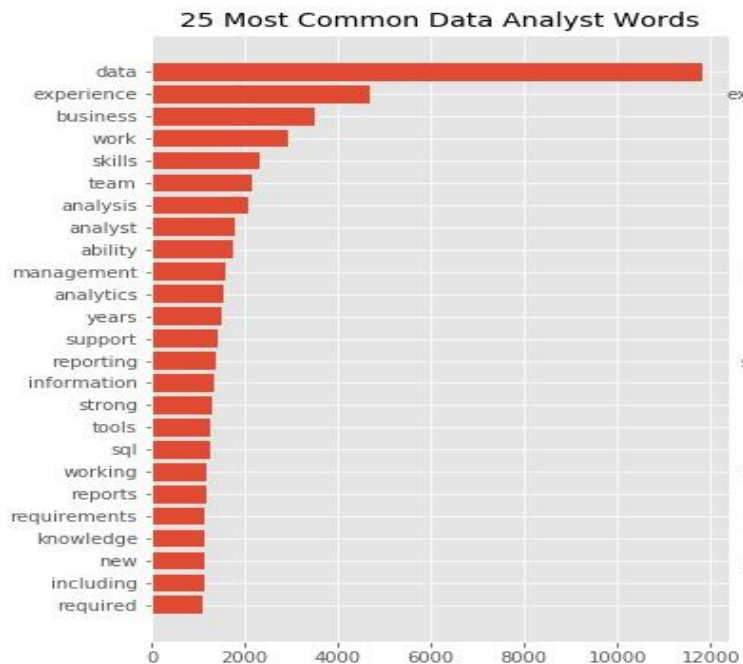
Methodology

- Data - LinkedIn scrape
- EDA - frequency counts for various jobs
- Modeling - random forest with >90% precision/recall
- Recommendations
 - Curriculum
 - Career Services
 - Marketing

Scrape Job Descriptions from LinkedIn



Most Common Words - Data Analyst vs Data Scientist



Word Clouds



Model

- Used data scraped from LinkedIn
 - TF-IDF
 - bi-grams
- Tested variety of classifiers to predict whether a job description was data analyst or data scientist
 - Logistic Regression, Naive Bayes, Random Forest
 - Gridsearch used to tune hyperparameters
- Random Forest Best Performing Model
 - F1 Score - 93%

Most Important Features - Data Scientist vs Data Analyst

Seperating

- Mathematics_Data
- Perform_Large
- Develop_Functionality

Not Seperating

- Python
- Data Scientist
- Degree requirements
- Machine learning
- Data Mining
- SQL
- NLP
- Statistical Modeling

Most Common Features - Jr vs Sr Data Scientist

Sr Level Roles

- Management
- Development
- Big_Data
- Communication Skills
- Statistical Analysis
- R_Python*
- Deep_Learning*
- Statistical_Modeling*

Jr Level Roles

- Engineering
- Modeling
- Support
- Data_Analysis
- Quantitative_Field*
- Unstructured_Data*
- Predictive_Modeling*

Most Common Features - Data Analyst vs Sr Data Scientist

Data Analyst

- Data_Analyst
- Business_Intelligence
- Data_Management
- Reports
- Bachelors_Degree
- SQL
- Data_Warehouse
- Communication_Skills

Sr Level Data Scientist Roles

- Machine_Learning
- Models
- Deep_Learning
- Big_Data
- Data_Mining
- R_Python
- Statistical_Analysis
- Advanced_Analytics

Most Common Features - Data Analyst vs Jr Data Scientist

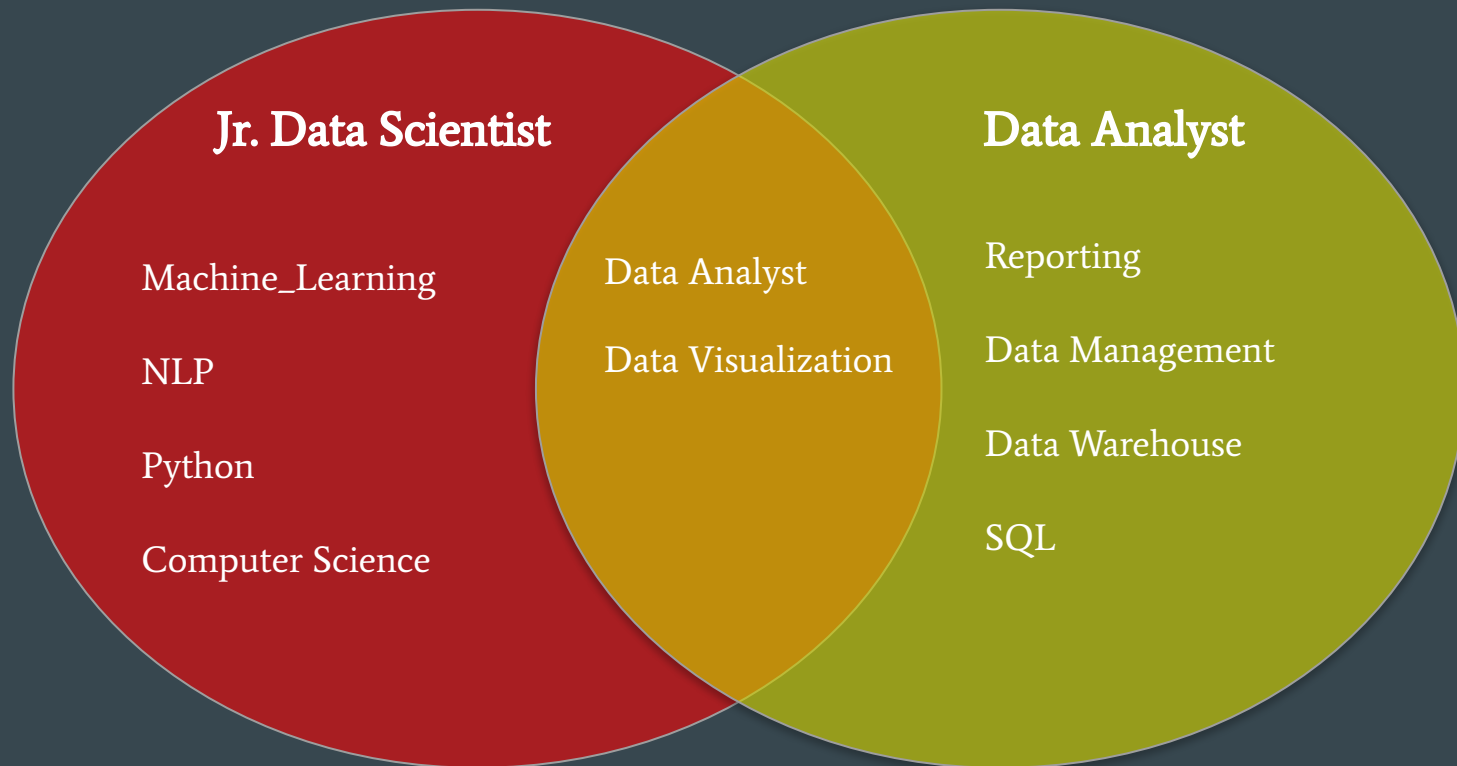
Data Analyst

- Data_Analyst
- Data_Visualization
- Reporting
- Data_Management
- Data_Warehouse
- SQL

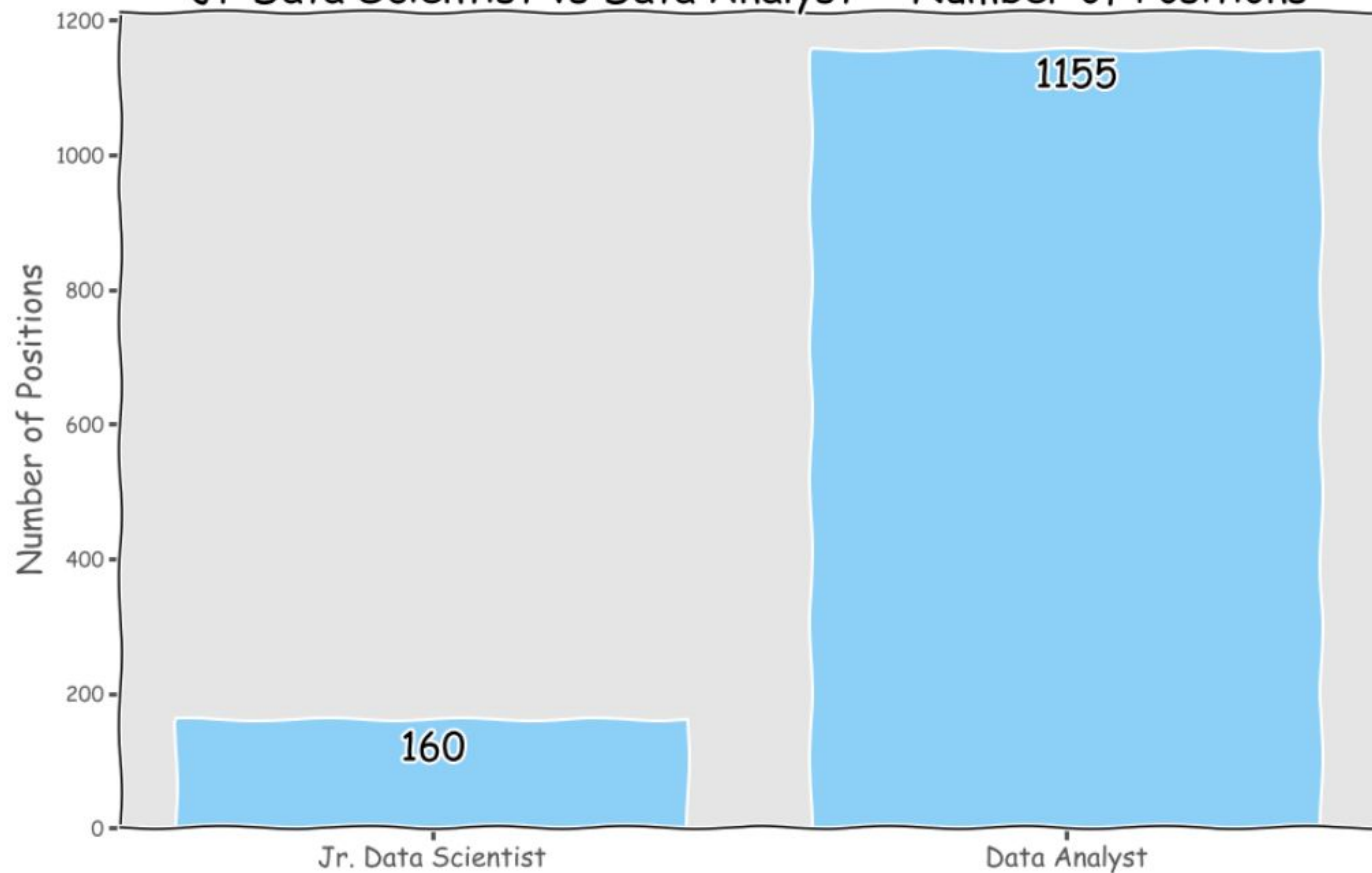
Jr Level Data Scientist Roles

- Data_Analyst
- Data_Visualization
- Machine_Learning
- NLP
- Python

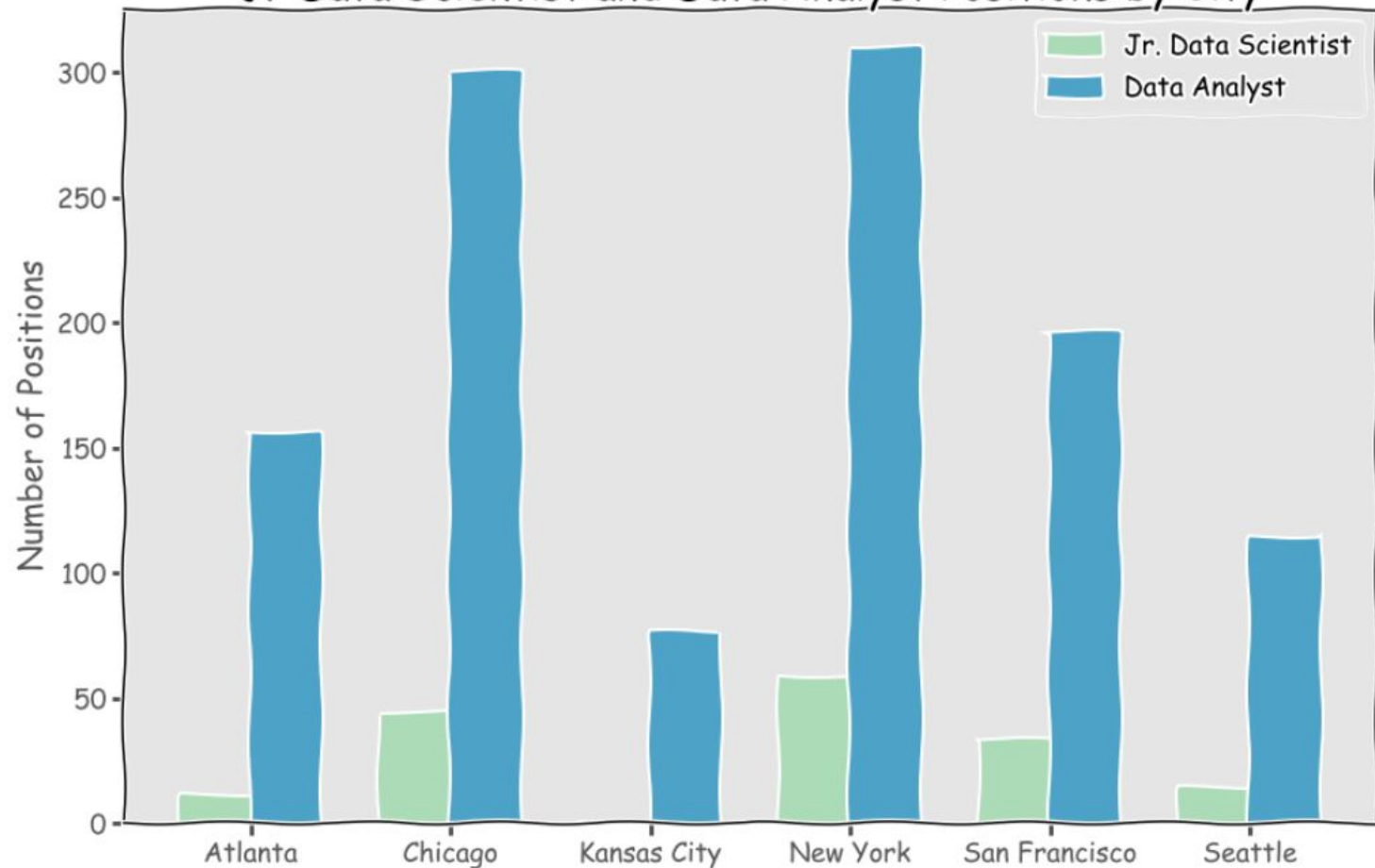
Jr. Data Scientist vs Data Analyst



Jr Data Scientist vs Data Analyst - Number of Positions



Jr Data Scientist and Data Analyst Positions by City



Recommendations - Curriculum

CURRICULUM

- Extend program from 15 to 18 weeks
- Split program into DS/DA tracks with optional specialized modules
 - Must determine cost implications and ability to consume all modules
 - DS specialization - deep learning project
 - DA specialization - deployed and hosted data web app; add additional sections on data warehousing
- Recommend applying to MS DS programs for students who weren't able to complete/couldn't get admitted to DS track during program

Recommendations - Career Services

CAREER SERVICES

- Provide DA/DS ratio for major markets/regions to help target portfolio/CV to majority class in desired job market
- Have corporate partners focus on DA jobs/employers
- Default job search recommendation for analyst positions

Recommendations - Marketing

MARKETING

- Compare ratio of DA/DS jobs to google trends ratio of DA/DS to gauge market interest and education
- Shift emphasis from DS to DA

Future Work

- Do similar analysis for data engineer and machine learning engineer
- Keyword filter for different stacks and possible subfields (NLP, Big Data, time series, etc)

Thank you!

For further inquiries, please feel free to reach out to us:

jeff.herman@flatironschool.com

brandon.lewis@flatironschool.com