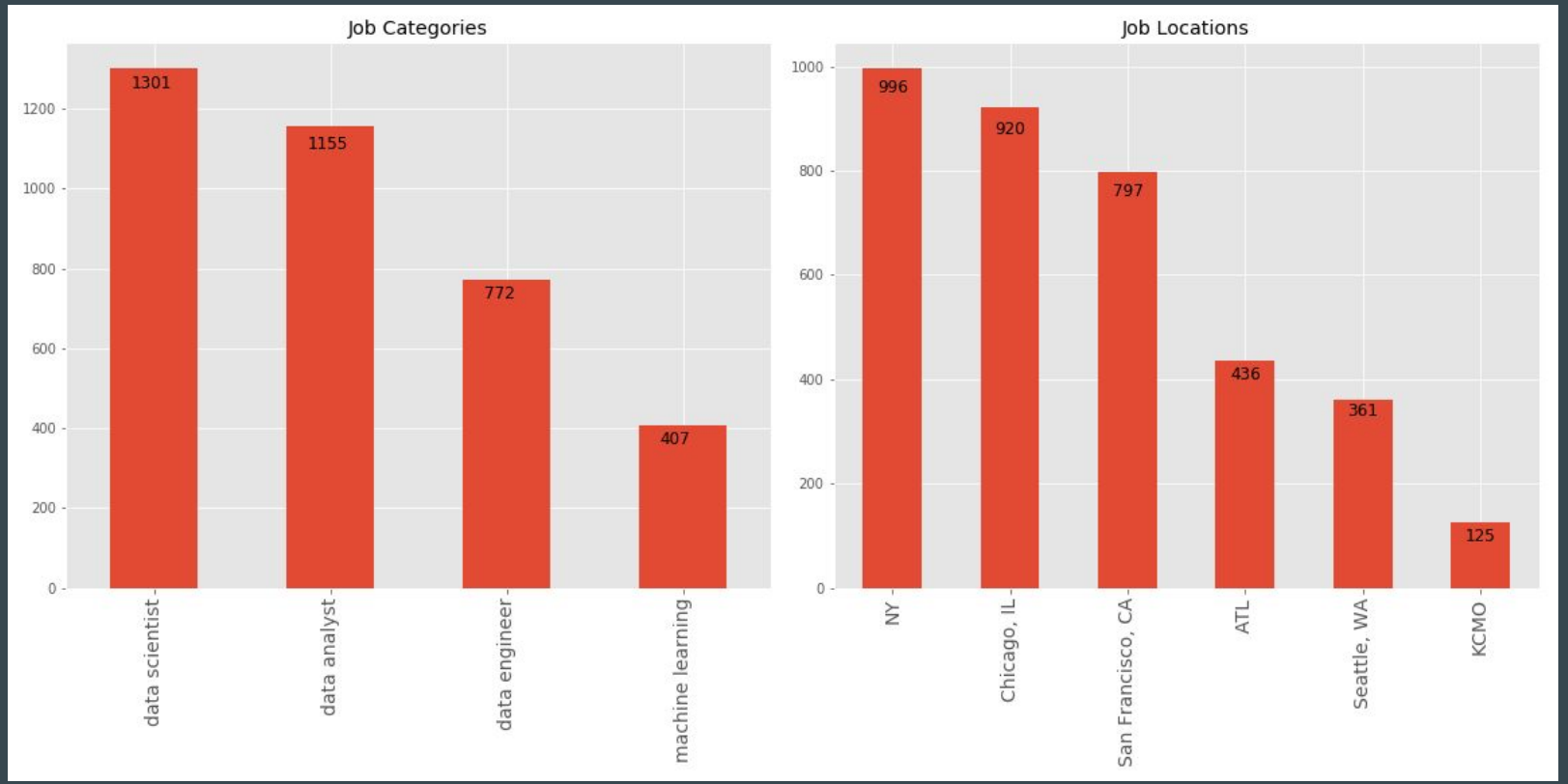
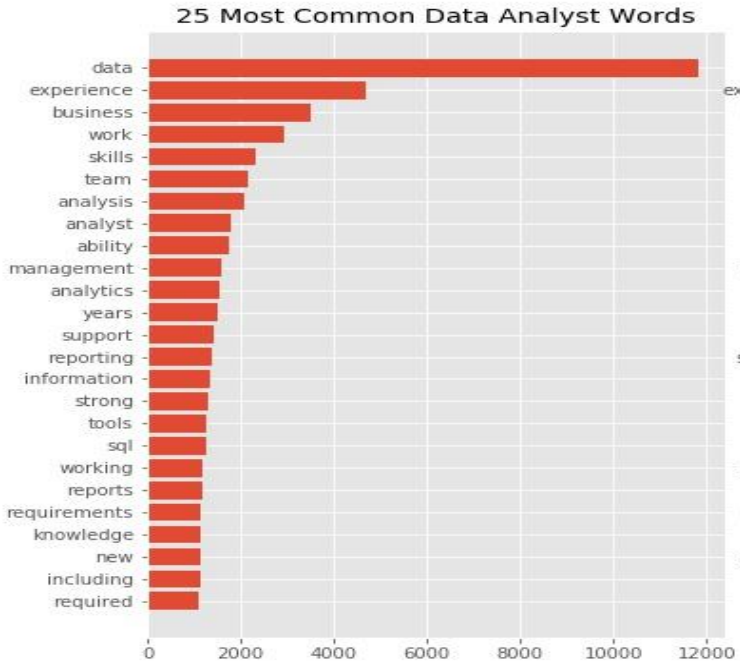
Data Science Job Market

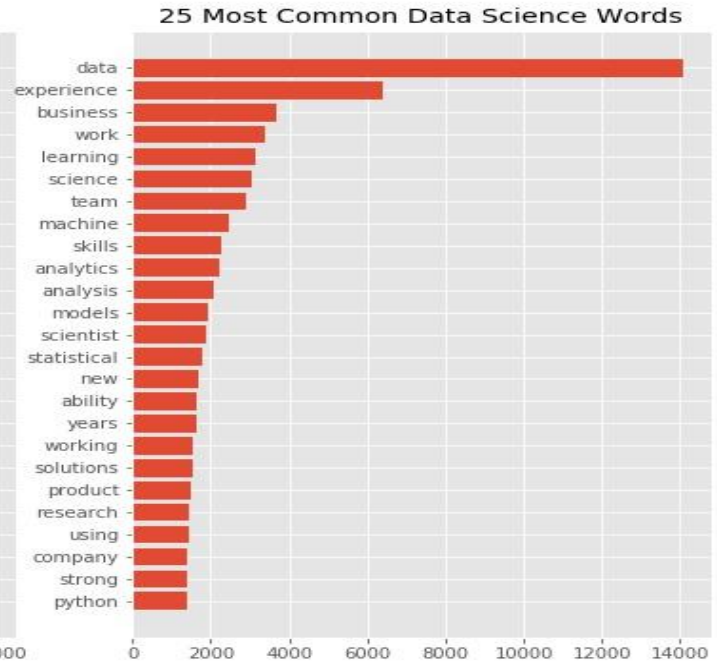
...

Scrape Job Descriptions from LinkedIn



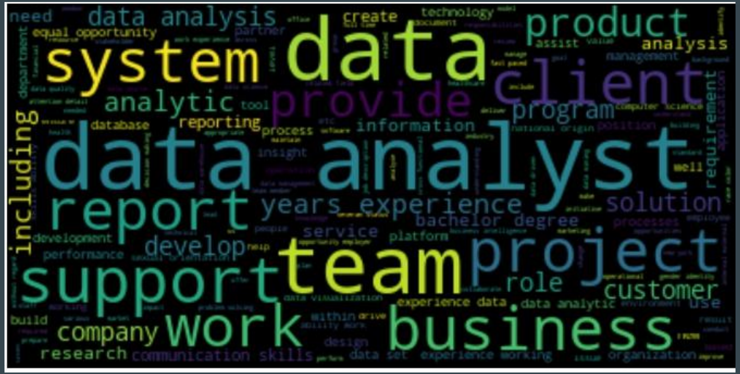
Most Common Words - Data Analyst vs Data Scientist





Word Clouds


data scientist
machine learning
data science
product
analytic
research
solution
model
including
company
customer
client
computer
science
work
team
support
platform
years experience
technology
big data
role
system
build
need
help
insight
data analysis
project
well
develop
data set
national origin
experience working
equal opportunity
people
service
use


data analyst
data
system
client
provide
product
data analysis
analytic
report
support
team
project
business
work
company
including
years experience
bachelor degree
solution
customer
role
develop
information
program
reporting
database
process
insight
service
platform
analysis
create
technology
need
equal opportunity
research
communication skills
development
performance
build
use
organization
requirement

Model

- Utilized Data Scraped from Linkedin
- Built Classifier to Predict Whether a Job Description Was Data Analyst/Data Scientist

Highlights

- TF-IDF (also bi-grams)
- Logistic Regression, Naive Bayes, Random Forest
- Gridsearch used to tune hyperparameters
- Random Forest Best Performing Model
- Model Performance: F1 Score - 93%

Most Important Features

Seperating

- Mathematics_Data
- Perform_Large
- Develop_Functionality

Not Seperating

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

Next Steps

- Break data analyst/data scientist jobs by years experience
- Do similar analysis for data engineer and machine learning engineer
- Keyword filter for different stacks and possible subfields (NLP, Big Data, time series, etc)