

Project Plan

Chih-Kai (Kyle) Chang

Project choice: Option 3

Perform an analysis of “data scientist” jobs listed on job boards and on the employment pages of major companies. * What are the most common skills that employers look for? * What are the most unique skills that employers look for? * Where are the types of companies that employ the most data scientists?

Project Plan

1. Scrape data from job postings website such as **Indeed.com**, **Glassdoor.com** with search key word *Data scientist*.
 - Use **rvest** package, and understand the pattern of URLs to scrape job description.
 - Sort search results and save them into .csv file.
2. Process data
 - Create a dataframe that have title, company, industry, city, state, and the link of that job description.
 - Scrape requirement skills information from the job link.
 - Get coordinate information for U.S. state. (Define those skills that are commonly required or essential for data scientist jobs such as programming skill) (Define what type of programming skill: Java, Python ect.)
3. Exploratory analysis
 - Find the most related skills identify common and important elements for different website by using bar chart.
 - Get a rough idea about how these data scientist jobs distribute around U.S. (ggmap)
4. Perform an proper statistical analysis and put analysis result into report.