



Project 3 Proposal

The Job Seeking Word Scraper

Team Members

- Emily Coxon
- Sandra Reyes
- Jonathon Wood

https://github.com/jonowood/Project_3

Project objective

The project objective is to assist job seekers in finding a job. Scraped data will be used to analyse the most popular requirements for certain fields to help job seekers identify which jobs they are qualified for, or necessary skills they may need to obtain.

Project overview

Data Sourcing - Webscraping using BeautifulSoup

Data Warehousing - MongoDB Database

Plotting and User Endpoint - We will plot the scraped data to various visuals using Leaflet & Plotly. We will then display them to the end user via a Interactive Website (Incorporating Flask, HTML/CSS, JavaScript.

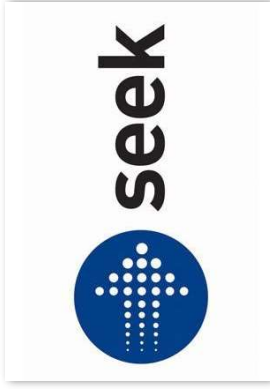
Styling - We will use the SemanticUI JavaScript Library to help create a beautiful, responsive layout for our site.

Data Sources

Using BeautifulSoup we will webscrape job listings for Seek, Jora & Indeed.

The scraped data will be stored in a MongoDB database for manipulation and plotting.

1



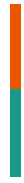
2



3



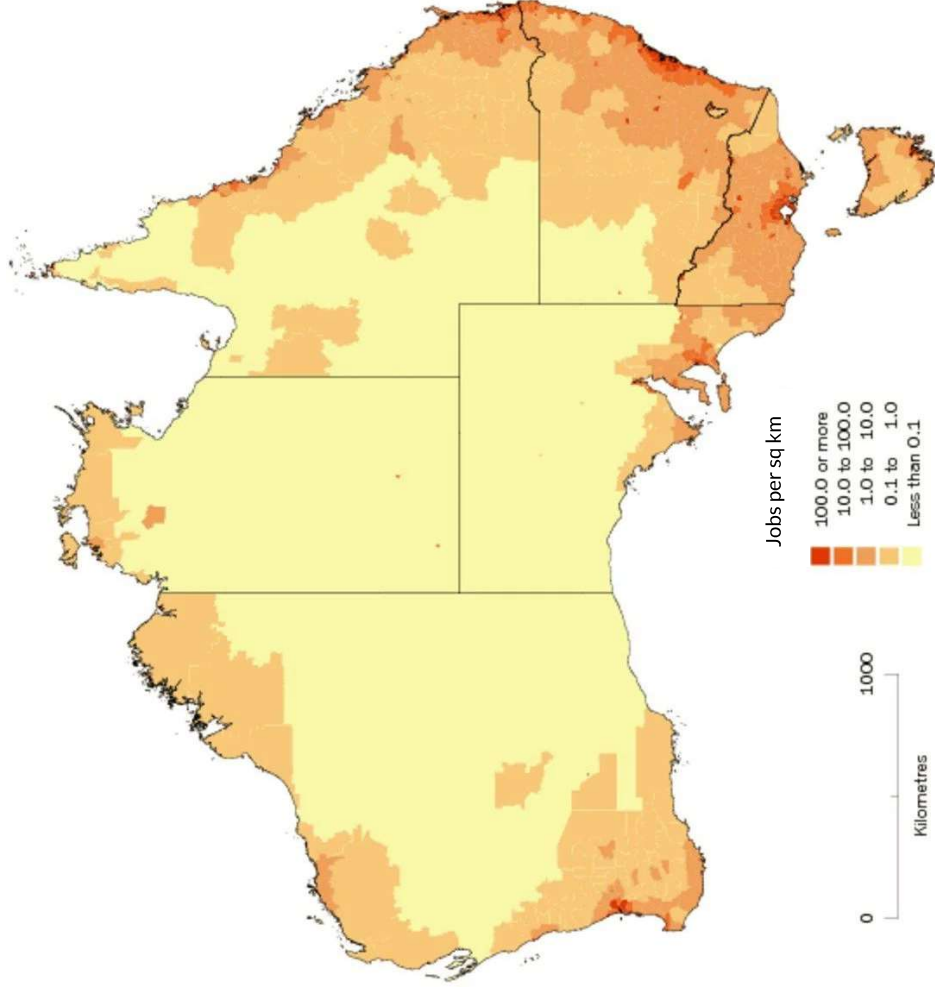
BeautifulSoup⁴



Dashboard visualisations

Choropleth Map

A choropleth map identifying the highest concentration of job postings will be produced based on the field in which the job seeker chooses.





A word cloud identifying the most common phrases will be produced based on the field in which the job seeker chooses.

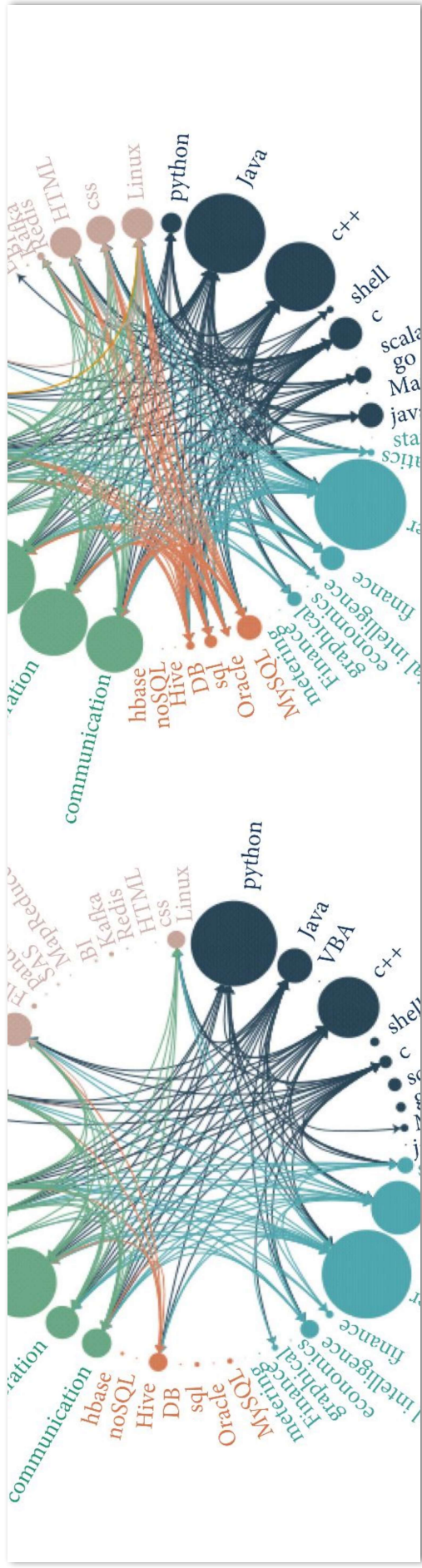




Dashboard visualisations

Dendrogram

A dendrogram plot identifying the most common phrases will be produced based on the field in which the job seeker chooses. This field is selected interactively on the plot itself, filtering the common phrases associated with that field.





Thank you.

