Fetched from:

<https://www.kaggle.com/elroyggj/indeed-dataset-data-scientistanalystengineer/downloads/indeed-dataset-data-scientistanalystengineer.zip/2>

=======================================================================================================================

=======================================================================================================================

Job\_Title: Title of Role

Link: Weblink of Job Posting

Queried\_Salary: Salary Range of the Job Posting (Estimated/Actual if available)

Job\_Type: 3 Categories of Job Types - data\_scientist, data\_analyst, data\_engineer

Skill: List of desired skills on indeed site

No of Skill: Count of the number of desired skills

Company: Company that posted the job posting

No of Reviews: Number of Reviews for the Company

No of Stars: Ratings for the Company

Date Since Posted: Number of days since the job was posted - if less than a day, will be rounded up to a full day

Description: Web scrape of part of the job description

Location: State the job opening is located in

Company\_Revenue: Annual revenue of hiring company

Company\_Employees: Employee count of hiring company

Company\_Industry: Industry of hiring company

=======================================================================================================================

=======================================================================================================================

Expansion of Skill column (Top 10 overall skills):

- python

- sql

- machine learning

- r

- hadoop

- tableau

- sas

- spark

- java

- Others

Expansion of Location column (Top 10 overall states):

- CA

- NY

- VA

- TX

- MA

- IL

- WA

- MD

- DC

- NC

- Other\_states

Expansion of Company\_Industry columns (Top 5 overall industries)

- Consulting and Business Services

- Internet and Software

- Banks and Financial Services

- Health Care

- Insurance

- Other\_industries