**Data Engineering Capstone Project**

**Project Resources**

### Datasets

Gathering the right data is one of the most important task for data engineers. If you are looking for data for your own, self-directed project, or for data to enrich the data for the Udacity provided project, below are some good resources to find interesting datasets.

##### Analytics

* [Google: Dataset Search](https://toolbox.google.com/datasetsearch)
* [Kaggle Datasets](https://www.kaggle.com/datasets)
* [Github: Awesome Public Datasets](https://github.com/awesomedata/awesome-public-datasets)
* [Data.gov](https://catalog.data.gov/dataset)
* [Dataquest: 18 places to find data sets for data science projects](https://www.dataquest.io/blog/free-datasets-for-projects/" \t "_blank)
* [KDnuggets: Datasets for Data Mining and Data Science](https://www.kdnuggets.com/datasets/index.html" \t "_blank)
* [UCI Machine Learning Repository](https://archive.ics.uci.edu/ml/datasets.html" \t "_blank)
* [Reddit: r/datasets/](https://www.reddit.com/r/datasets/)

##### Software Development

* [Github: Public APIs](https://github.com/toddmotto/public-apis)
* [Last Call: Top 50 Most Popular APIs on RapidAPI (2018)](https://blog.rapidapi.com/most-popular-apis/" \t "_blank)
* [Facebook: Graph API](https://developers.facebook.com/docs/graph-api" \t "_blank)

### Example

[This Kaggle dataset](https://www.kaggle.com/hugomathien/soccer/home) is a good example of someone who found disparate datasets and combined them to provide an even more valuable dataset for others to analyze.