

## Sources

<https://www.kaggle.com/datasets>

<https://github.com/awesomedata/awesome-public-datasets>

## Potential datasets:

Pavan:

<https://www.kaggle.com/saulfuh/bart-ridership> - Bay Area Rapid Transit ridership data, provides a couple of exploratory questions as well

<https://earthquake.usgs.gov/earthquakes/search/> - earthquake data, can get CSV based on magnitude/time (limited to 20000 events)

<https://github.com/emorisse/FBI-Hate-Crime-Statistics/tree/master/2013> - FBI hate crime data in counties by quarter and broken down by demographic info

Ameeshi:

<https://archive.ics.uci.edu/ml/datasets/Sentiment+Labelled+Sentences> - It contains sentences labelled with positive or negative sentiment.

<http://tomslee.net/airbnb-data-collection-get-the-data> - Airbnb Data for different cities

<http://insideairbnb.com/> How is Airbnb really being used in and affecting the neighbourhoods of your city?

<https://www.yelp.com/dataset> Yelp reviews, businesses etc.

<https://github.com/narudarurasya/yelp-analysis>

<https://github.com/scku/Yelp-Data-Analysis>

<https://nycdatascience.com/blog/student-works/project-1-exploratory-visualizations-of-yelp-academic-data-set-draft/>