ETL Project

Brandon Su

Andrea Rosenberger

**Sources of Data**

Kaggle: Significant Earthquake data. <https://www.kaggle.com/usgs/earthquake-database/data>

We chose this dataset because of a general interest in the trends and frequency of recent significant earthquakes, and their locations. There have been rumors for quite some time that we are due for a very big earthquake, and this dataset was a way of just taking a glimpse of a possible trend.

**Types of Transformation Needed**

For the earthquake dataset from Kaggle, we did not need much transformation aside from cleaning. There were extra columns dedicated to depth error, magnitude types and stations that were not needed for what we wanted to see, which was a basic measure of the earthquake, location, and time.

**Moving Data into PgAdmin4 Database**

We chose PgAdmin4 due to general preference and usability of the tool. We found that it was much easier to work with and to visualize the data we import into the database.