ETL Project

By: Nishant, John and Anam

**Pre-Processing**

Data analyst is very demanding role. As part of project requirements our team has been hired to analyze data scientist data to address whether data scientists are very demanding or not with location.

The following table illustrates the observations and actions taken by the group to ensure a clean data set.

| **Pre-process Step** | **Data Need** | **Observation** | **Action** |
| --- | --- | --- | --- |
|  | 1000+ rows of data needed | More than 501 rows for each set. | None taken. |
|  | Unique identifier | Job title is the unique identifier for two of the data sets. | None taken. |