ETL Project



Monica Jones, Alicia Bricoe, Teya Robinson

TAble of  
Contents

Extraction ……………………………………………………………………………………………….1

Transformation………………………………………………………………………………………2

Load…………………………………………………………………………………………………………5

Extraction

Sources

|  |  |
| --- | --- |
| Site URL | File Type |
| http://data-wake.opendata.arcgis.com/ | CSV, JSON |

This website provided a variety of data formats for our team to manipulate and analyze.

Transformation

1. We utilized SQLAlchemy, Pandas, JSON, PPRINT and other modules pictured below to prepare and manipulate the data.

A screenshot of a cell phone

Description automatically generated

1. After importing the data from the CSV file we created a data rame using Pandas to view the restaurant data.
   1. Data collected: Restaurant name, address, city, state, postal code
   2. Renamed columns

A screenshot of a cell phone

Description automatically generated

1. The Inspection data was in a JSON format which we also created a dataframe using Pandas to view the data.
   1. Data collected: Inspection type, object ID, Score, Inspection description

A screenshot of a social media post

Description automatically generated

LOAD

We chose to load the transformed data into a mySQL database. We also decided to join the two tables for a more wholistic view of the data.

