CS 4485 - 002 – T2 – Data Ingestion

Thinh Nguyen

Nam Nguyen

Duy Vu

Minh Vu

**Report #1**

This is a quick summary of our first assignment about obtaining a list of all US zip codes and importing it into the database. Database information:

Database name: zipcode

Username: root

Password: password

For the list of zip codes, we decide to download it free from the source “http://federalgovernmentzipcodes.us/download.html”. This website provides a list of all US zip codes for free, which contains 42522 unique codes. From other resources, this is currently a correct total number of US zip codes. The CSV file was last updated on January 22, 2022. The information provided on this website indicates that this zip codes data has been used for many other research works, which means it is pretty reliable.

The problem with this source is the CSV file need to be downloaded and stored manually. However, according to our research, most US zip code numbers remain the same. Furthermore, we do not consider the Zip+4 Code, which may change monthly. Therefore, from our viewpoint, it is not necessary to automatically update the CSV file for now to save system resources. Our current routine to import this data is running the python script manually at this moment.