# Installation and Usage Instructions

## Installing the relevant libraries

The project is built in python. The folder includes a file called “requirements.txt”. To install the relevant dependencies, simply run in terminal:

pip install -r requirements.txt

## Database Setup

1. Install MySQL from <https://dev.mysql.com/downloads/mysql/>
2. Once installation is complete, open the mysql shell using mysql -u “username” -p and then enter the password.
3. Create a database with the following command :

“CREATE DATABASE sentiment\_store;”

## Data Management and Preprocessing Pipeline

The entry point of the project is the “main.py” file. The program assumes that a mysql instance with a database called sentiment\_store exists in the system.

To connect with the mysql instance, go inside the DatabaseConnection.py file and change the host, username and password to the relevant setting.

Then, execute the main.py through python main.py to populate the database with the review data and then preprocess the same.

This step performs the following tasks:

1. Combine the 2 data sources into a single table called data\_dump;
2. Read data from data dump
3. Cleans and Pre-Processes the entire dataset
4. Stores the preprocessed data into a new table called preprocesstable for persistence.
5. Generates multiple machine learning models and clustering models
6. Generates insights using models generated above
7. Visualizes wordclouds for each cluster