**Module 21 First Segment Project Deliverable**

Start Assignment

Congratulations! You’ve made it through Segment 1 of the project. You’ve already put in lots of work on this project—so now is the time to submit that work to your instructor and get the credit that you’ve earned. You’ll also get valuable feedback that will help you improve as you track to the final deliverable.

**Requirements for Segment 1**

Segment 1 is worth 33% of your overall grade for the Final Project.

Before you submit the Segment 1 deliverable, make sure that you have all the pieces in place by reviewing the requirements.

**Content (60 points)**

In this segment, the project deliverable includes the following:

* A detailed README.md file (20 points)
  + Creating and Working with a New Repository
    - <https://github.com/LouFoster/Class-Project_GroupC.git>
  + Readme File is created
* At least four commits per team member (20 points)
  + Once each member can access each project member can add 4 “four” commits – meeting the 2nd requirement
* A database that stores at least two tables (or collections) for the project (20 points)

Get input from

* **Pollution Data**

**Database Source:** https://openweathermap.org/api/air-pollution

**Table /Database #1: Air Pollution Table Screen Shot**

Graphical user interface, application, table

Description automatically generated

* **Screen Shot of 100 Cities**

**Data Year Collected in 2022 -**

**Database Source:** Kaggle

**Table /Database #2: Top Cities Screen Shot**

Graphical user interface, application, table

Description automatically generated

data source

* weather data <https://openweathermap.org/api/air-pollution>
* World Cites Data database (ask Adam for the details)
  + - * Check out Kaggle

Database Structure

Graphical user interface, application

Description automatically generated

The team should also submit the ERD,

Complete before Class Thursday (Lou)

Answer: Lou (Need to be updated with Christian data … need to see the file to create the ERD)

Graphical user interface, application

Description automatically generated

**Presentation (40 points)**

In this segment, the presentation deliverable includes the following:

* The selected topic and the reasoning for that selection (20 points)

**Overview: Welcome the “C- Group**

We are a data analytical staffing team that can help clients by supplying enhanced data and analysis to help clients successfully meet their business needs. With the rise of data in today’s economy, our staff is focused on the practical and technical skills needed to analyze and solve complex data problems.

C-Group is working with “Jason” from [Trafalgar tours](https://www.trafalgar.com/en-us/deals/september-on-sale?utm_campaign=septemberonsale&gclid=Cj0KCQjwla-hBhD7ARIsAM9tQKscaA9-Jj1xPUN3nW8evJKNtv6PoJHf4yfqpj6G6CDxtFxMtKVYcesaAudkEALw_wcB&gclsrc=aw.ds)’ Marketing Department. [Trafalgar](https://www.trafalgar.com/en-us/deals/september-on-sale?utm_campaign=septemberonsale&gclid=Cj0KCQjwla-hBhD7ARIsAM9tQKscaA9-Jj1xPUN3nW8evJKNtv6PoJHf4yfqpj6G6CDxtFxMtKVYcesaAudkEALw_wcB&gclsrc=aw.ds) is a top-rated travel company that specializes in internet-related services. To that end, Jason would like to offer prospective customers an interactive web page where they can research the weather and various weather aspects of the top 100 cities traveled.

To support Jason in this effort C-Group will offer (the client) help collecting and presenting data for customers via the search page, which they will then filter based on their preferred “weather” travel criteria in order to find their ideal city, anywhere in the world.

As part of our analysis,

* C-Group will export the data, clean it, and use the weather data to choose the best cities for vacation based on certain weather criteria.
* C-Group will also need to perform statistical calculations on the data and the weather parameters in the Northern and Southern hemispheres.
  + Suggestion We can also have a filter to search by Northern or Southern Hemisphere
  + Module 12: can help us design our new interactive web page for the top cities and read through the data about each city via a search and filter process.

**##A description of the data (20 points)**

* Weather Tracking to investigate Air Quality in Multiple Cities
  + **Table1: For our Pollution Data: air quality standards as per https://openweathermap.org/api/air-pollution**
    - **See Screen Shot for Attributes / Column headers explained.**

A screenshot of a computer

Description automatically generated

* Table 2: **Top 100 Cities**

**Data Year Collected in 2022 -**

**Database Source:** Kaggle

**Table /Database #2: Top Cities Screen Shot**

* The questions that the team plans to answer with the project (20 points)

**Project Hypothesis**

* cities with larger populations have lower air quality countries
* track air quality at certain times of the year
  + Summertime is the worst, no matter where you are
  + Reports on Years vs Air Quality in both Southern and Northern Hemispheres
* Does the country have air quality standards laws, rules, or regulations?

Game Plan

* Ping API to Obtain Air Quality
* Manually look up lat and lng for cities
* Save to dictionary
* Convert to dataframe
* Import into a database
* Visualization
* Machine Learning
  + Using the published standards for Air Quality as per <https://openweathermap.org/api/air-pollution> as the basics of our Machine Learning processes which will be focused on predicting Air Quality .

**ERD v1 (2023.03.31)**

Graphical user interface, application

Description automatically generated

[Trafalgar tours](https://www.trafalgar.com/en-us/deals/september-on-sale?utm_campaign=septemberonsale&gclid=Cj0KCQjwla-hBhD7ARIsAM9tQKscaA9-Jj1xPUN3nW8evJKNtv6PoJHf4yfqpj6G6CDxtFxMtKVYcesaAudkEALw_wcB&gclsrc=aw.ds) **https://www.trafalgar.com/en-us/deals/september-on-sale?utm\_campaign=septemberonsale&gclid=Cj0KCQjwla-hBhD7ARIsAM9tQKscaA9-Jj1xPUN3nW8evJKNtv6PoJHf4yfqpj6G6CDxtFxMtKVYcesaAudkEALw\_wcB&gclsrc=aw.ds**

**Our Group Status**