Data Analytics Bootcamp  
etl projet Proposal

Alexis Summey, Andrew Davis, Ellen Colborn, Justin Butler, Timisha Martin, Scott Caldwell

## Project Background and Description

|  |  |
| --- | --- |
|  | The city of Toronto has a solid plan to increase public and transportation safety to effectively continue to decrease the amount of collisions in the area. The introductions of “The Vision Zero Road Safety Plan,” the “Congestion Management Plan,” and the “Transit Network Expansion” plan have all led to the curiosity surrounding the city’s success in these areas over time.  This project will be an analysis of the recent implementation of the various traffic congestion and public transit safety plans over the past 10 years in Toronto. We will analyze the city’s traffic congestion, pedestrian and cyclist traffic, and vehicular collision data. We will also analyze the relationship between city weather data and traffic congestion/vehicular collision data. |

## Predictions

|  |  |
| --- | --- |
|  | We predict that because the city of Toronto has introduced solid plans to effectively increase public and transportation safety, the vehicular, pedestrian and cyclist collisions will decrease over time. We will only focus on data collected between 2008 and 2018 to account for a 10-year data analysis. |

## High-Level Requirements

|  |  |
| --- | --- |
|  | The high-level requirement of this project will be to Extract, Transform and Load data by querying database. |

ETL must include the following:

* Placing CSV or other datafiles into a Pandas DataFrame
* Transforming data to be included and queried within a database
* Establishing a successful connection to a database to query

## Deliverables

|  |  |
| --- | --- |
|  | The project will be submitted in our “etl\_toronto” GitHub repository. We will include our proposal, data files, data cleaning files, SQL files, Pandas/Python files and report. |

## Implementation Plan

|  |  |
| --- | --- |
|  | Include recommendations that lead to your proposed solution. Summarize what you’re proposing to do and how you’re going to meet the goals. You’ll be able to expand on the details within the ‘Our Proposal’ section. |

## High-Level Timeline/Schedule

|  |  |  |
| --- | --- | --- |
|  | **September 7th, 2019:**    **September 9th, 2019:**  **September 11th, 2019:** | Explore various topics and find data  Create the project proposal  Clean the Data  Place data files into Pandas DataFrame  Load data into database  Create report? Presentation? |

## Specific Exclusions from Scope

|  |  |
| --- | --- |
|  | Describe how you plan to implement the project. For example, will all parts of the project be rolled out at once or will it be incremental? What will be included in each release? |