**1.0 Extract:** your original data sources and how the data was formatted (CSV, JSON, pgAdmin 4, etc).

**1.1 Data sources**

In this project, two original criminal datasets were retrieved and downloaded as CSV files.

* Criminal data from Toronto Police:

<https://data.vancouver.ca/datacatalogue/crime-data-details.htm>

* Criminal data from City of Vancouver official website: <https://data.torontopolice.on.ca/pages/catalogue>

**1.2 Data columns**

The crime rate in Vancouver dataset (2003-2018) contains the following information:  
- Crime type  
- Crime date  
- Crime time  
- Neighbourhood

The crime rate in Toronto dataset (2014-2018) contains the following information:  
- Offence type  
- Reported month  
- Reported day of the week  
- Occurrence month  
- Occurrence date of the week  
- Longitude  
- Latitude

**2.0 Transform:** what data cleaning or transformation was required.

**2.1 Questions**

Three questions are resolved in this report based on the 2 achieved datasets.

* Compare the crime rates by category in Toronto and Vancouver.
* Analyze the crime rate trends in both cities over the years.
* Analyze the crime rate trends month by month in both cities.

**2.2 Data cleaning**

**2.2.1 Crime data of City of Toronto**

* crime rates by category

list all categories recorded

group data by category and calculate the number of crimes of each category

draw a bar chart to show the count

* crime rate trends from 2014 to 2018

calculate the number of total crime records of all crime category of each year

order the result above by year

draw a line chart to show the trend

* crime rate trends by months of the year

**2.2.1 Crime data of City of Vancouver**

* crime rates by category

list all categories recorded

group data by category and calculate the number of crimes of each category

draw a bar chart to show the count

* crime rate trends from 2003 to 2018

calculate the number of total crime records of all crime category of each year

order the result above by year

draw a line chart to show the trend

* crime rate trends by months of the year

**3.0 Load:** the final database, tables/collections, and why this was chosen.