

## wrangle-report

- **Data rescores:**

<https://www.stats.govt.nz/large-datasets/csv-files-for-download/>

- **Gathering:**

After download data from internet i import that data into my programming environment ( Jupyter Notebook) using (read\_csv) function.

- **Assess:**

Assess data using ( info() - describe() - sample() - value\_counts() - duplicated() - sort\_values() - isnull() ) functions.

### Issues:

1- delete rows have no values in (Victimisations\_calendar\_year\_2015) and i(Population\_mid\_point\_2015) columns

2- Erroneous datatypes (Area\_unit\_2013\_code, Urban\_area\_2013\_code, Territorial\_authority\_area\_2013\_code, Rate\_ratio\_NZ\_average\_rate, Region\_2013\_code, Region\_2013\_label, Urban\_area\_type ) columns

3-reset index of dataset . and delete "Index" column

4-delete the word "urban area" in colume "Urban\_area\_type"

5- Inaccurate "Rate\_per\_10000\_population" Calculate by ("Victimisations\_calendar\_year\_2015"\*10000)/"Population\_mid\_point\_2015")

6-delete the word "Region" in "Region\_2013\_label" column

- **Clean**

After define the issues i convert those definitions to code and run that code and i tested my dataset, visually and with code, to make sure my cleaning operations worked.

\*\*\*\*\*