Dataset Documentation:

Contribution: Md Aminul Islam Rana and Md Abu Sufian

In this section, we used a dataset from Edmonton city of Canada to get a glimpse of real-life analysis using machine learning tools by python code.

Dataset Source: https://www.edmonton.ca/city_government/citizen-dashboard-city-by-the-numbers

Here, we have completely mentioned in detail the dataset that I have used in my project as dataset documentation for Edmonton city which following below:

Dataset Name: Edmonton City by the Numbers

Source: https://www.edmonton.ca/city_government/citizen-dashboard-city-by-the-numbers

Dataset Overview:

The dataset Edmonton city by the numbers offers focus on the development and success of the city of Edmonton across a few metrics pertaining to population, economics, environment, transit, public safety, and more. The dataset contains more than 100 different indicators, all of which are updated often to give the most recent statistics on the condition of the city.

Variables:

UNEMPLOYMENT_RATE: The unemployment rate measures the percentage of the labor force that is unemployed but actively seeking employment within the past four weeks.

National_Unemployment_Rate: The National Unemployment Rate is the unemployment rate for Canada as a whole.

Impaired Driving Incidents: This variable represents the number of incidents of impaired driving that have been reported to the Edmonton Police Service.

90_RIGHT_ENERGY: This variable represents the percentage of customers whose power is restored within 90 minutes after an outage has occurred.

Edmonton CMA - Working Age Population Growth: This variable represents the percentage change in the working-age population (aged 15-64) within the Edmonton Census Metropolitan Area (CMA) over a specified period.

Edmonton CMA - Labour Force Growth: This variable represents the percentage change in the labor force within the Edmonton Census Metropolitan Area (CMA) over a specified time.

Edmonton CMA - Employment Growth: This variable represents the percentage change in employment within the Edmonton Census Metropolitan Area (CMA) over a specified time.

Governance: This variable represents the city of Edmonton's score on a governance index, which measures the quality of government and public services.

YEAR: This variable represents the year in which the data was collected.

Data Collection:

The City of Edmonton has accumulated information from several sources, including government departments, non-profit groups, and commercial businesses. The city of Edmonton's open data team gathers and arranges the data, working to make sure it is accurate, current, and open to the public.

Data Usage:

Subject to the restrictions and conditions specified by the City of Edmonton, the public may freely use the Edmonton City by the number's dataset for non-commercial purposes. Users of the data must credit the data's creator in any publications or software programmes that use it. Before utilising the data for any analysis or application, users are advised to review the dataset documentation and use guidelines.

Limitations:

There are some restrictions on the data, even though Edmonton city by the Numbers offers a thorough overview of the development and advancement of the city of Edmonton. To safeguard people's privacy, some data points might be restricted or anonymised, and some data might not be accessible for all years or time periods. To make sure they are utilising the data properly and within the permissible scope, users are advised to study the dataset description and use guidelines. Moreover, the dataset was in small size that why we face problems with dataset tuning time.

Recommendation:

The large dataset gives more accurate accuracy and results. So, further study and analyse of the big data set as size would be better for taking leadership decision-making.