

# Context

## Business intelligence and data-driven decision making

This project provides insights on how different groups of people have accessing to green space and how the government can help increase the neighbourhood ratings

## Domain knowledge and the business context

Scottish Household servers is run by Scottish government. It is designed to help the Scottish Government and other bodies to plan services and policies for Scotland. The project helps the government to understand the differences in how people accessing green space in order to make better policies.

# Data

## Internal and external data sources

All data used were from the organisation and are open.

## Types of data

Mostly are categorical strings. Also has numerical types.

## Data formats

CSV formats

## Data quality and bias

For Ethnicity groups, all the 'Other' groups are from Edinburgh and Glasgow, The data has a lot of missing values. Data was pre-grouped and can not compare between columns.

# Ethics

## Ethical issues in data sourcing and extraction

No, The Scottish Household Survey (SHS) is a continuous survey based on a random sample of the general population in private residences in

Scotland.Data used for this project is open to public. There is no identifiable information.

## **Ethical implications of business requirements**

No, the business requirement were to make better decision for Scotland and this is ethical.

# **Analysis**

## **Stages in the data analysis process**

- 1.Data exploring and cleaning
- 2.Feature engineering
- 3.Converting 'values' to proportions
- 4.Creating and interpreting plots
5. Hypothesis test
6. Building models

## **Tools for data analysis**

I used Python for my analysis, within it, I used Pandas, Numpy, Seaborn, Matplotlib, scitick-learn

## **Descriptive, diagnostic, predictive and prescriptive analysis**

Descriptive, my model explains what features have impacts on the neighbourhood rating based on the data.