# Introduction

This is a Git Collaboration project conducted by the participants below:

Shannon Watts

Nathaniel Mweemba

Maxwell Acha

Gary Whitney

# Purpose of the Project

To carry out an in-depth analysis of crime data obtained from various sources which are listed in the DATASETS section below. The crime data is specific to the west midlands region and is interrogated and cross examined in python, making use of various libraries to establish corelations and draw conclusions. The main aspects of the data include:

Crime ID

Crime type

Last outcome category

Month

Longitude

Latitude

Location

LSOA code

LSOA name

# Hypothesis

Within the west midlands during 2019, did the index of multiple Deprivation (IMD) score have any corelation to the crime rate in the respective area. Are certain areas pre-disposed to higher rates of crime if so, what else can the IMD score inform us about the crime rates.

# Some descriptions

Lower Layer Super Output Area (LSOA)

Are small areas designed to be of a similar population size, with an average of approximately 1,500 residents or 650 households. They are designed to improve the reporting of small area statistics in England and Wales.

The Index of Multiple Deprivation (IMD)

Is a measure of relative deprivation for small areas (Lower Super Output Areas (LSOAs). It is a combined measure of deprivation based on a total of 37 separate indicators that have been grouped into seven domains, each of which reflects a different aspect of deprivation experienced by individuals living in an area.

Every LSOA in England is given a score for each of the domains and a combined score for the overall index. This score is used to rank all the LSOAs in England from the most deprived to the least deprived, allowing users to identify how deprived areas are relative to others.

The 7 domains are Income, Employment, Education, Skills and Training, Health and Disability, Crime, Barriers to Housing Services, Living Environment. In this project we focus on Crime.

# Research Questions

## Research Question 1:

## Does the crime data source from the West Midlands Police Database correlate with the Index of Multiple Deprivation (IMD) score?

## Research Question 2:

## Which district and Lower Layer Super Output Areas (LSOAs) have higher rates of crime?

## Research Question 3:

## What type of crime might you be exposed to in certain Districts/LSOAs?

## Research Question 4:

## Where is the safest place to park your bike?

## BREAKDOWN OF TASKS

1. Identify and obtain data as csv or excel files
2. Studied/discussed
3. Exploring and cleaning of data
4. Lists questions / choose questions that will be used
5. Plan visualizations
6. Allocate tasks to group members
7. Coding
8. Create charts
9. Summarization/conclusion
10. Preparing the presentation

## DATASETS

* Lower Layer Super Output Areas December 2011 Boundaries
* West Midlands Crime Data 2019 by month
* Local Income Deprivation Data

# Decisions and results

## Results from analysis

Do some areas have severe rates of crime? does this fit with our hypothesis?

Is there a corelation between IMD score and crime rate?

Which crime has the highest number of occurrences?

What assumptions can be proved or disproven?

Are there any trends in relation to IMD score?

What would make the analysis better?