# 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 Data Sources 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 that 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

**Index of Multiple Deprivation (IMD) score**: This is a comparative tool used to identity the most deprived areas nationally, regionally and locally. The general rule is that the higher an areas IMD score and the lower its rank on the scale the more deprived an area is. It consists of Crime(%), Health(%), the rest of IMD’s and % here

For reference the current ranking system ranges from 1 to 32,844, with 1 being the most deprived and 32,844 being the least deprived. National and local organisations use the Index of Multiple Deprivation, sometimes in conjunction with other data, to distribute funding or target resources to areas. From the results of the crime data analysis we would be able to see which areas would benefit from resources such as more policing or neighbourhood watch initiatives

Lower Layer Super Output Area (LSOA) are geographic hierarchy designed to improve. The reporting of small area statistics in England and Wales.

# Research Questions

1. Q1
2. Q2
3. Q3
4. Q4

## 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

1. West Midlands Crime Data 2019 by month
2. Local Income Deprivation Data
3. Lower Layer Super Output Areas December 2011 Boundaries
4. D4
5. D5

# Decisions and results

**Notes**

1. IMD – indicators
2. Scatter plot crime rate vs IMD
   1. Coding
   2. Outliers
3. Bar charts – district level overview
   1. Potentially independent mean?
4. Plot heatmap
   1. Coding
   2. Group by
   3. Further analysis where we find interesting things – e.g. bike theft Warwick
      1. Go deeper and let readers know exactly where not to leave your bike!

## Results from analysis

which crime has the highest number of occurrences and in what area?

which areas have more severity of crime and does this correlate with the IMD ?

What would make the analysis better ?

Increased scope – other IMD aspects?

Use API’s