

Homework 2

Group Members:

1. Lina Ziaka
2. Cherry O'Connell
3. Pallavi Atkarne
4. Iwona Izdebsk
5. Diana Mutheu

Topic: Healthcare and Criminology, original idea: England street crime analysis before and during the COVID-19 pandemic.

We will analyse street crime before and during the Covid-19 Pandemic with focus on London and compare it with other major cities in England. We will cover the years 2018 until 2021. We will analyse if lockdowns during the pandemic had any significant effect on the number of street crimes reported. With this, we aim to answer the following questions:

- Did pandemic affect street crime levels?
- What are the top street crimes before and during the pandemic and what are the outcomes of these crimes?
- What are the locations with the most and least number of street crimes and were they affected by the pandemic?
- Was there a significant change in street crime outcomes?
- What other factors influenced by the pandemic coincided with changes? For example, major social events, mobility of the population.

Our datasets will be coming from different agencies and organisations. The main source of our datasets will be the UK Police database where reports from different constabularies and police services across England are stored. Other relevant data will be sourced from APIs, London Public Database, UK government website for Covid-19 Pandemic information.

To manage this project, we are using agile methodology. We held a sprint planning meeting and decided on the sprint backlog, which will be managed with the help of github. We are working in short iterations, to be able to quickly react to any new ideas and challenges as we dig deeper into the data. At the moment we are in the data collection, cleaning and exploratory analysis phase, to which the whole team contributes based on every person's strength and weakness. The next sprint will be devoted to data analysis.

During the sprint planning meeting everyone will pick a backlog item to work on, but at the same time any team member can help and share any ideas and resources they have come across and can contribute to any backlog item if they have time or the other team member struggles. Moreover, to keep up with all the tasks we have a project roadmap displayed on a Gantt Chart to keep track of completed and upcoming tasks (see image below).

For code management we will be using Git and GitHub for version control and also for reviewing each other's code. For managing file storage, we are using Google Drive for example for the final report, powerpoint presentation slides and csv data files. For the data, since it is too large (spanning over millions of rows) we will be using SQL to create a database. To make the database accessible to everyone we are hosting it using RDS which is offered by AWS which stores the database in the cloud.

Gantt Chart:

Epic		NOV							DEC							DEC							DEC						
		T	F	S	S	M	T	W	T	F	S	S	M	T	W	T	F	S	S	M	T	W	T	F	S	S	M		
CFG-1 Project Understanding	DONE																												
CFG-2 Data Mining & Cleaning																													
CFG-9 London street crime	IN PROGRESS																												
CFG-10 London population	IN PROGRESS																												
CFG-11 London tourist data	IN PROGRESS																												
CFG-12 Covid-19 Lockdown timeline	IN PROGRESS																												
CFG-13 Other - news	TO DO																												
CFG-3 Data Analysis																													
CFG-14 London street crime analysis	TO DO																												
CFG-4 Predictive Modelling & Visualisation																													
CFG-5 Complete Homework 2																													
CFG-6 Daily Scrum																													
CFG-7 Code Review																													
CFG-8 Project Report & Review																													
CFG-16 Report Writing	TO DO																												
CFG-17 Review	TO DO																												
CFG-15 Project Submission																													
CFG-18 Presentation																													