Crimes in the UK -Week 1

Aser Abdelrahman
Chaoran Chen
Sergiu Soima

Crimes in the UK – Dataset structure

1. Street-level crime:

   12 attributes

   35,000,000 rows

2. Outcome

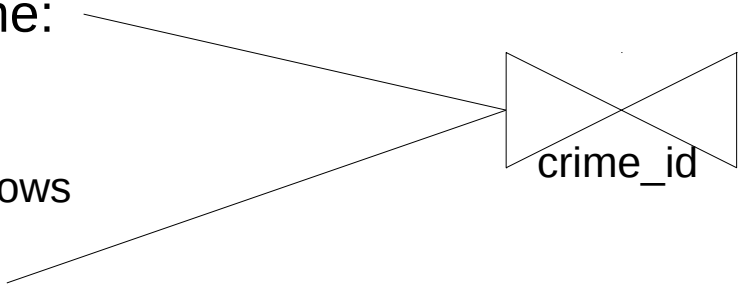
   10 attributes

   17,623,790 rows

3. Stop and search information

   14 attributes

   584,836 rows



Crimes:

15 attributes

38,550,660 rows

Quality (with respect to outcome)

1. Street-level crime:

   18,535,646 (52.96%) outcomes are missing

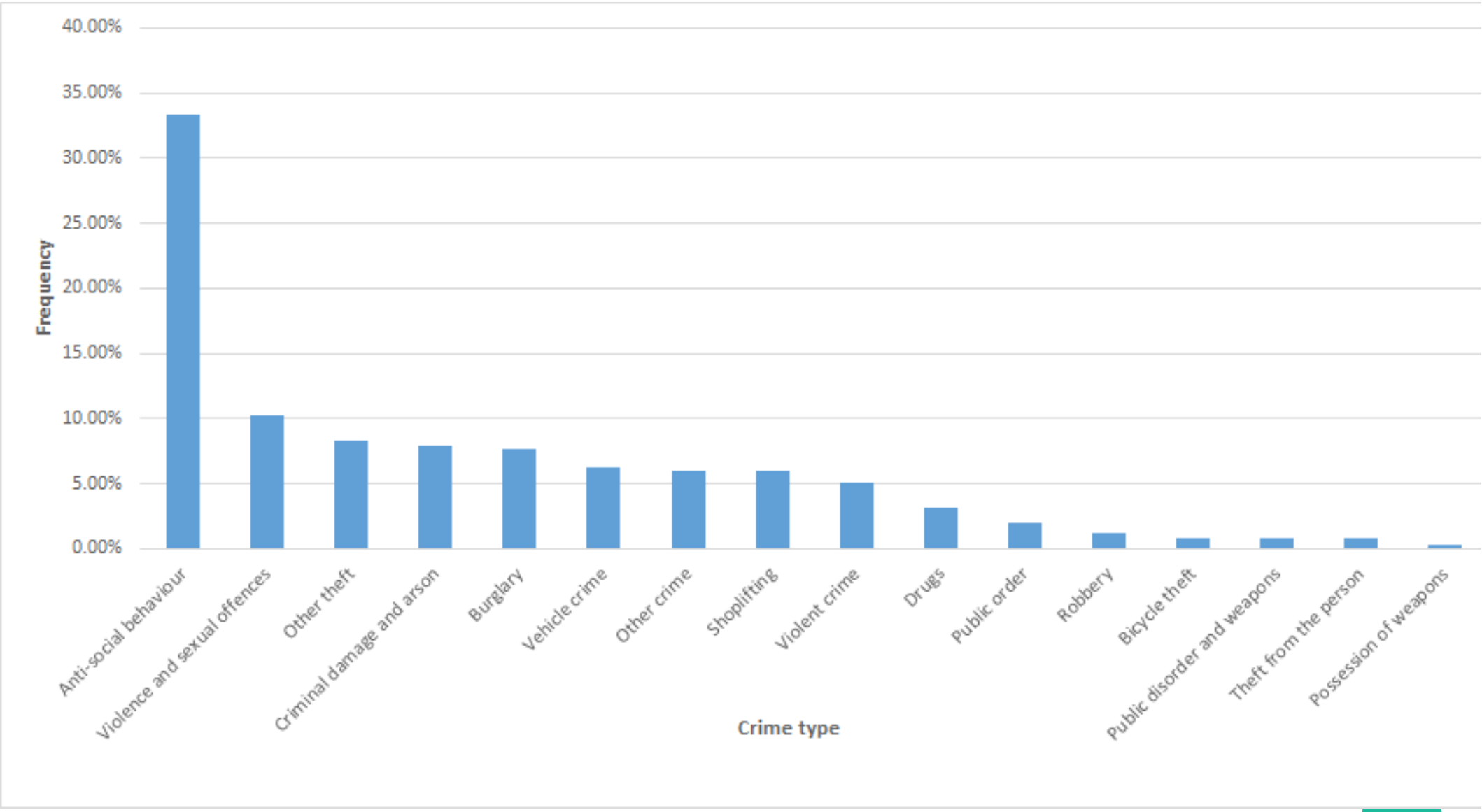
2. Crimes:

   20,765,675 (53.87%) outcomes are missing

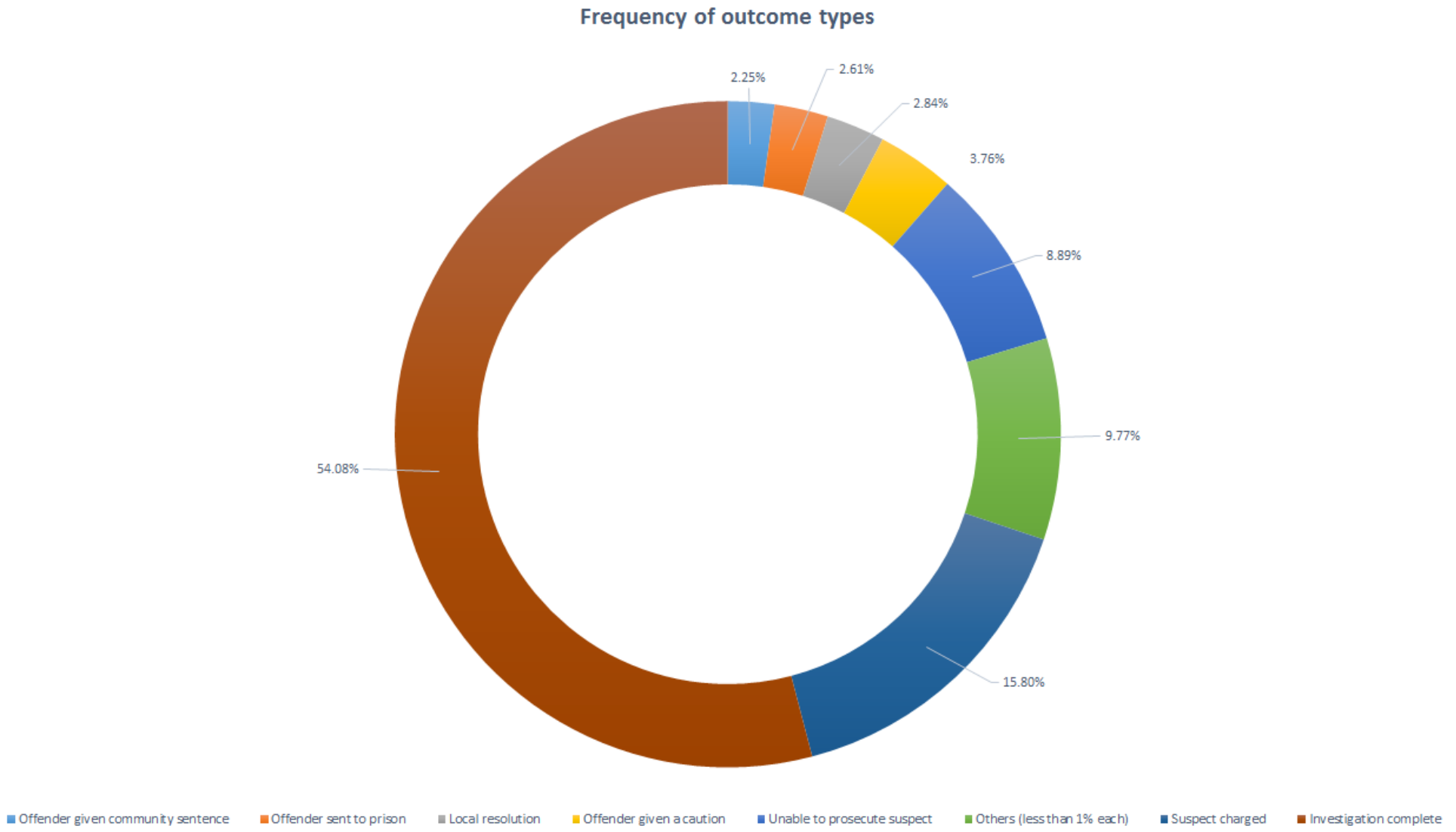
3. Stop and search information:

   17,957 (3.07%) outcomes are missing

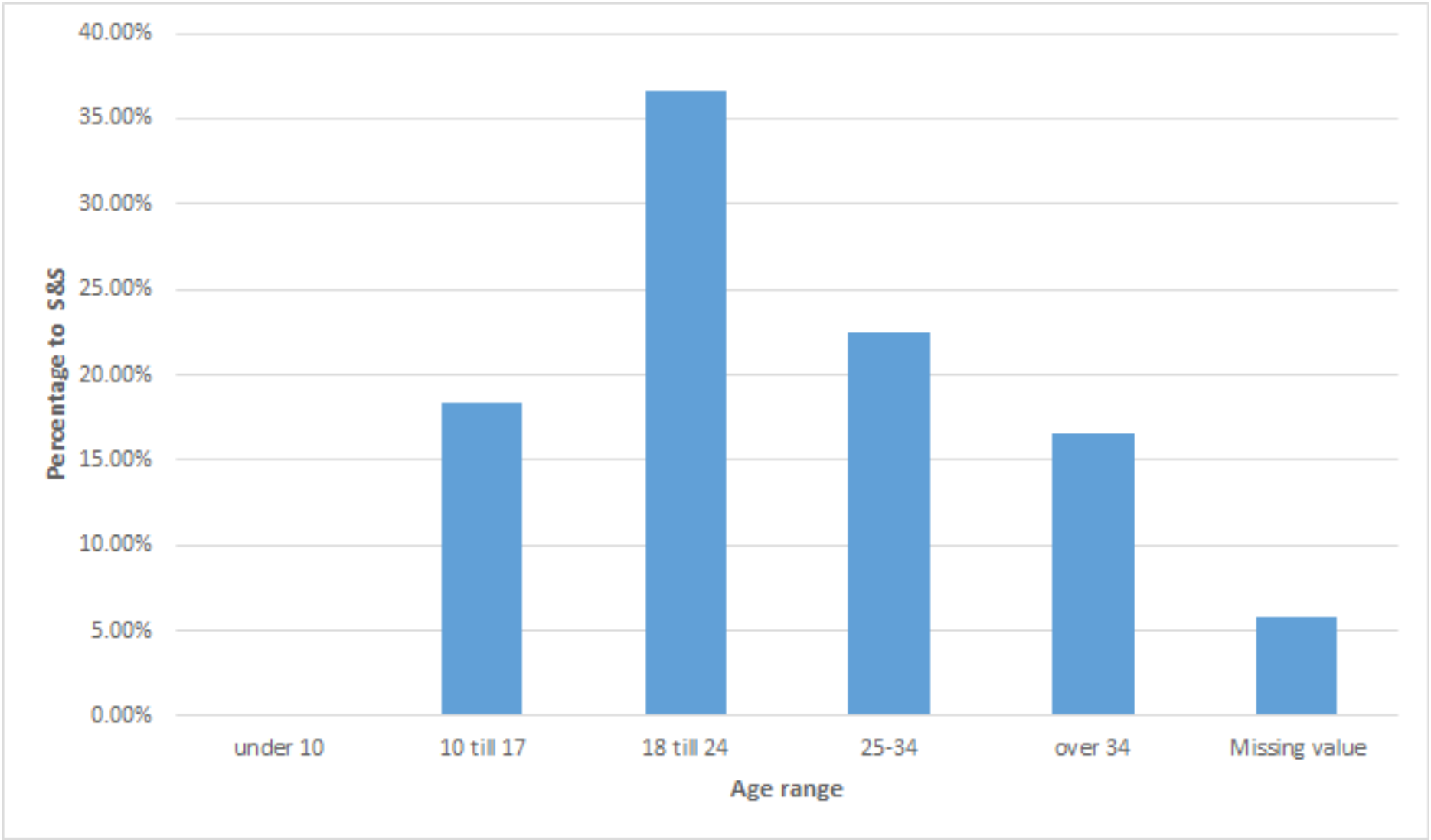
Crime Type Frequency



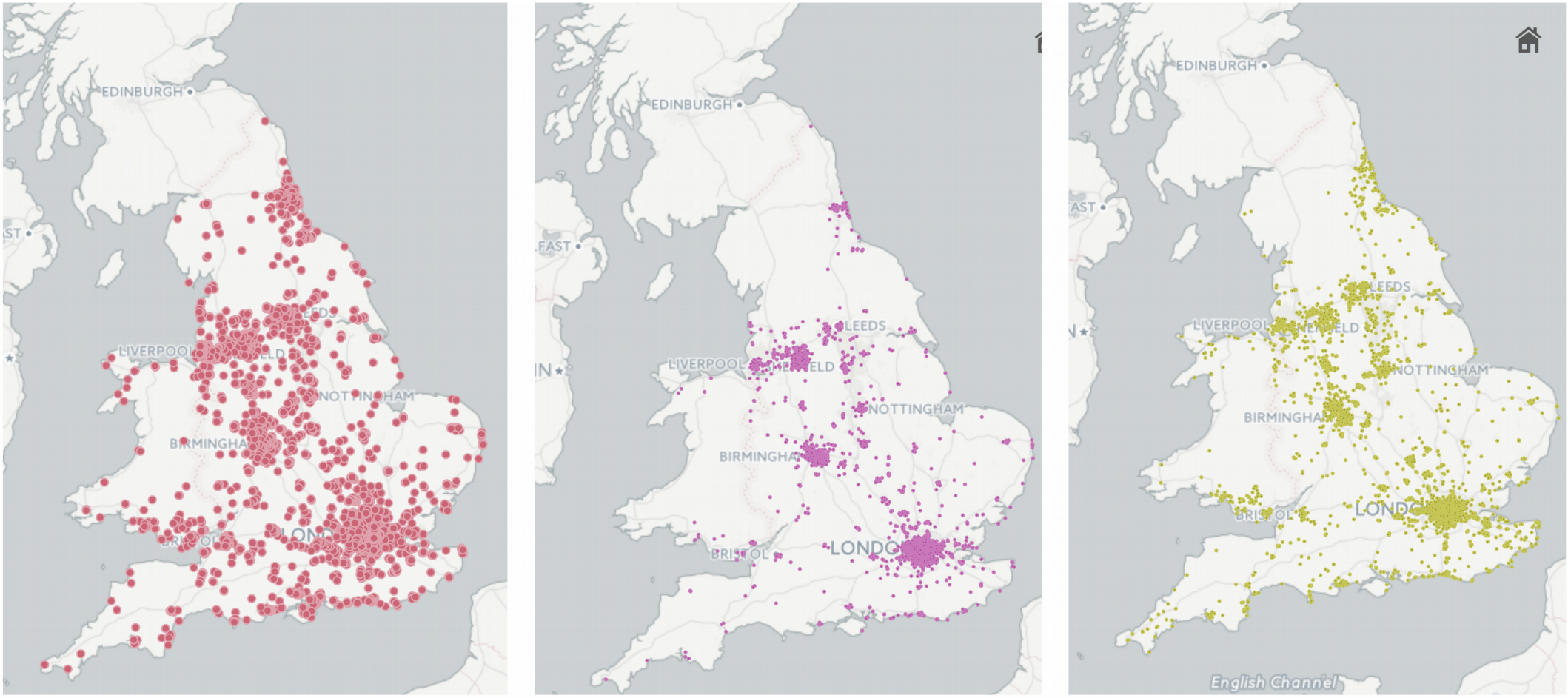
Outcomes distribution



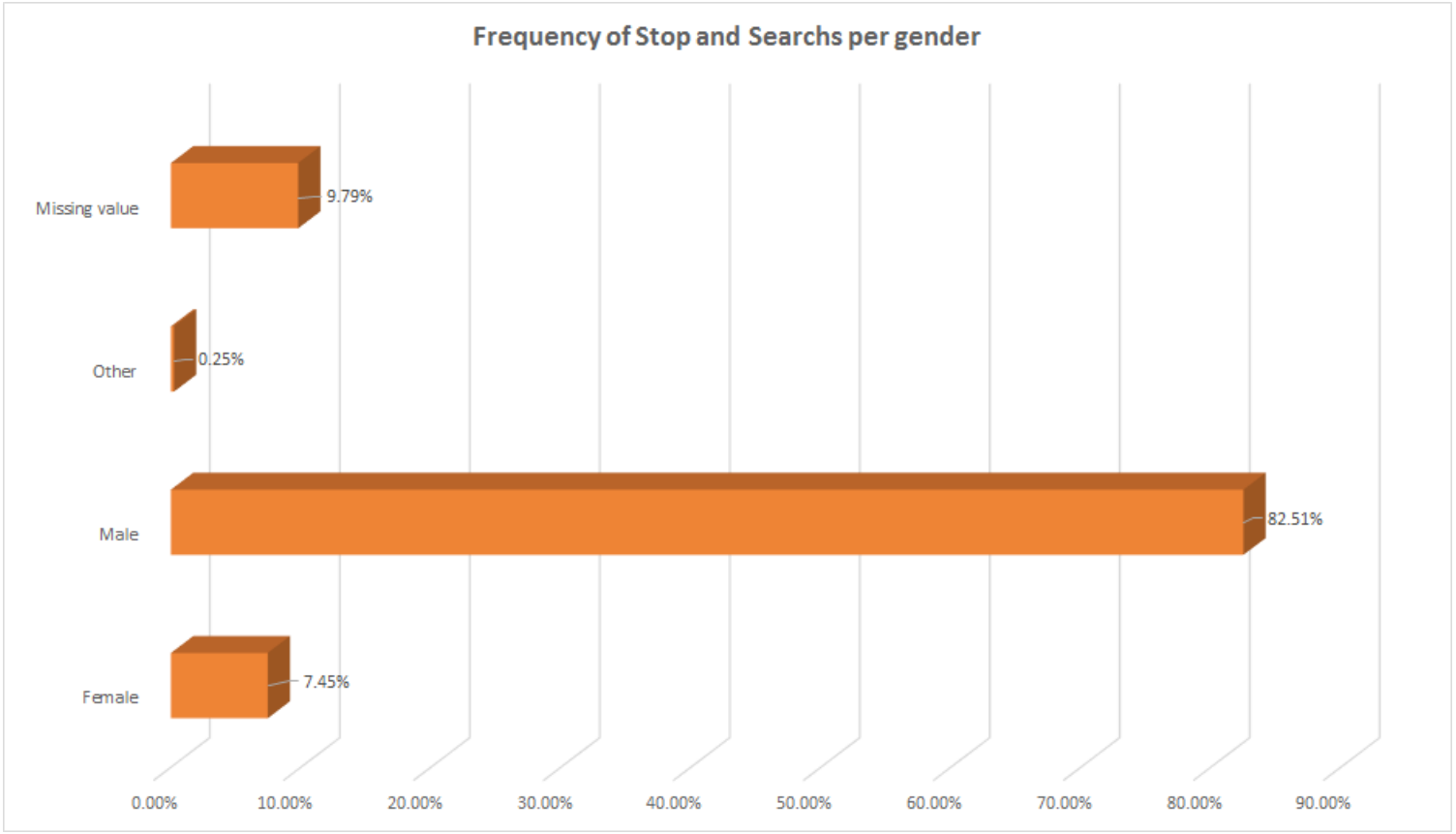
Age range frequency



Drugs, rubbery and violence locations



10 times more men investigated during stop & search



Ideas for data expansion

- Neighborhood statistics using LSOAs

  Education level - Health and care - People and society (income/lifestyle)

- Points of Interest

  ATM - Attraction - Bank - Entertainment - Food & Drink - Parking lot

- Weather data

  Are crimes influenced by weather?

Working Log:

- Reading row data into database
- Problems with MySQL
- Join of Outcome and Street
- Drawing samples and then throwing them away
- Visualization Tools

Plan for next week

- Consolidate the final table
- Cleaning and preparing the data
- Join “Crimes” and “Stop and search” datasets by aggregating information about the nearest “Stop and search” tuples located within a given range in km from the respective crime
- Write the python script for finding the nearest Stop&Search tuples
- Investigate if we can include Neighborhood statistics using LSOAs in the final table
- Investigate if we can include Points of Interest (POI) information using geolocation in the final table

Crimes in the UK -Week 1

Thank you

Aser Abdelrahman
Chaoran Chen
Sergiu Soima

Sources

https://www.police.uk/crime-prevention-advice/anti-social-behaviour/