Rawad - Analysis of EDA for Draft project Report

To analyze the asylum seekers dataset, a structured data preparation and analysis workflow was implemented. The methodological approach followed a systematic process that included data collection, cleaning, and visualization in Google Colab.

Data Collection

The dataset was sourced from an official government repository, containing records on asylum seekers receiving support in the UK. The raw data included variables such as date, nationality, region, support type, accommodation type, UK region, and the number of people receiving assistance.

Data Cleaning

The collected dataset underwent preprocessing to ensure consistency, accuracy, and usability. Key cleaning steps included:

- **Handling Missing Values:** The dataset was examined for missing values in critical columns such as date, nationality, and support type. Rows with substantial missing data were removed, while minor gaps were imputed where necessary.
- **Data Type Conversion:** The 'Date' column was converted to a standardized datetime format to allow for time-based analysis.
- **Categorical Data Standardization:** Inconsistent spellings and variations in categorical fields (e.g., nationality, region, and support type) were corrected to ensure uniformity.
- **Duplicate Removal:** Duplicate records were identified and eliminated to prevent redundancy in analysis.
- **Numerical Data Cleaning:** The 'People' column was checked for anomalies, such as non-numeric values, and was converted to an appropriate numeric format for aggregation and statistical calculations.
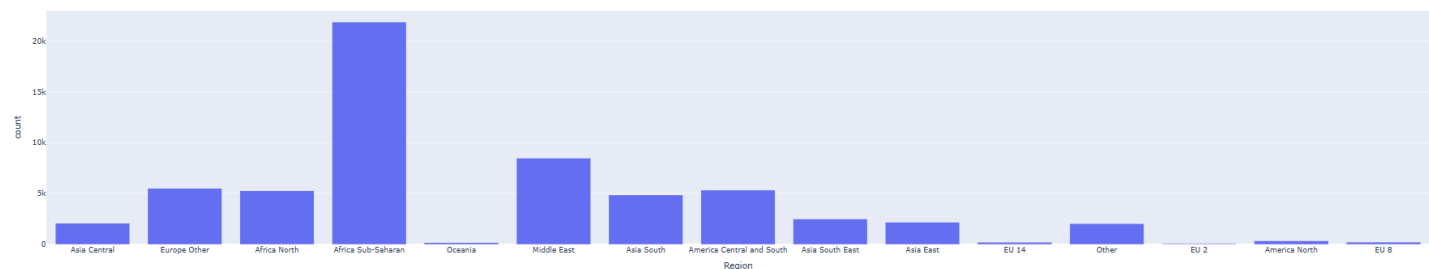
Data visualisations, Results & Key takeaways

1. Regional Distribution of Asylum Seekers

A histogram was generated to visualize the distribution of asylum seekers across different regions. The results indicate that certain regions have a significantly higher concentration of asylum seekers, suggesting a potential link between regional policies, available infrastructure, or government support programs.

Key Insight:

- Some regions have a disproportionately higher number of asylum seekers compared to others, which may be influenced by housing availability or local government policies.
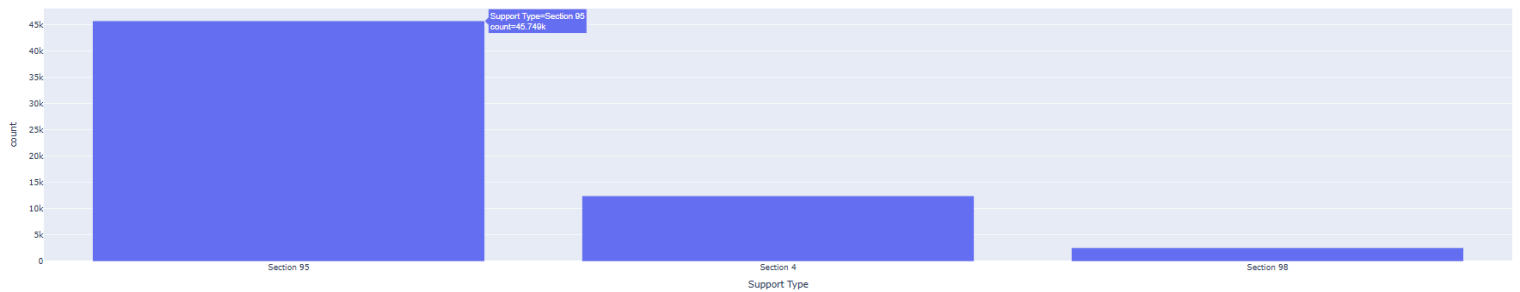

2. Distribution of Support Types

The histogram illustrating different **support types** highlights the variation in the type of assistance provided. Certain support types appear more frequently, indicating that some forms of assistance (e.g., dispersed accommodation) are more widely utilized than others.

Key Insight:

- The dominance of certain support types suggests a standard approach in providing assistance, possibly due to government funding allocation or policy preferences.
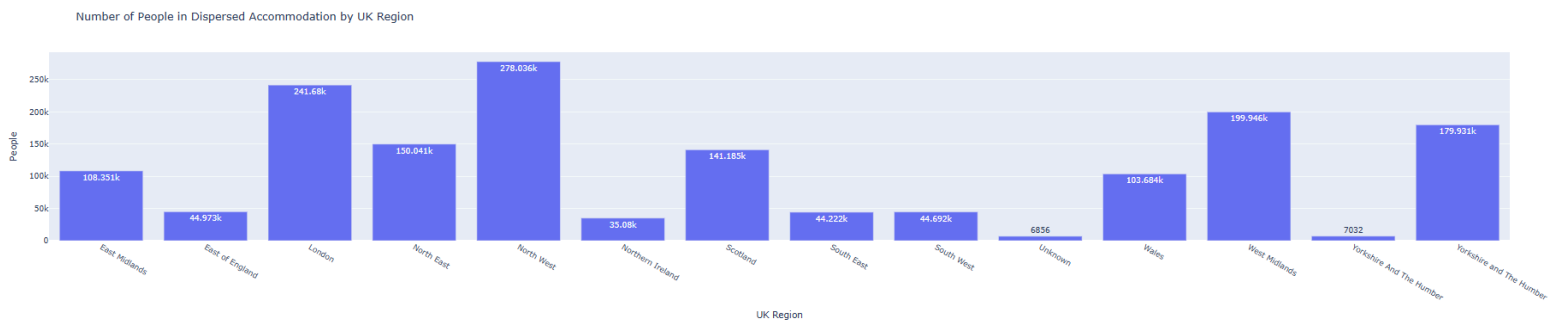
---

3. Number of People in Dispersed Accommodation by UK Region

A bar chart was created to showcase the number of individuals in dispersed accommodation across various UK regions. The data shows that certain regions accommodate significantly more asylum seekers than others.

Key Insight:

- The uneven distribution suggests that asylum dispersal policies are not uniform across the UK. Regions with higher numbers might have more available housing or specific policies favoring asylum placements.

Number of People in Dispersed Accommodation by UK Region

---

4. Trend of Top 5 Nationalities Seeking Support Over Time

A line chart was used to analyze trends in the number of asylum seekers from the top five nationalities over time. The visualization reveals fluctuations in asylum applications, with certain nationalities experiencing sharp increases or decreases in specific years.

Key Insight:

- The variations in nationality trends might be linked to geopolitical events, conflicts, or policy changes affecting migration patterns.
- A steady increase in certain nationalities may indicate a long-term shift in asylum-seeking demographics.

Trend of Top 5 Nationalities Seeking Support Over Time

Reference Dataset: This dataset was obtained from: