IMPERIAL

FROM DATA TO PRODUCT

DEPARTMENT OF DESIGN ENGINEERING

BoroughMatch

Author:
Suhani Gujral (02122070), Nora Luo(02109089)

Module Head
Pierre Pinson

Date: 10 December, 2024

Contents

	Nora Luo CID: 02109089	Suhani Gujral CID: 02122070
Ideation	50%	50%
Data sourcing	50%	50%
Data curation	50%	50%
Data product canvas	50%	50%
Conceptual and system design	50%	50%
Programming	50%	50%
Testing & validation	50%	50%
Archiving & documentation	50%	50%
Report writing	50%	50%

Table 1: Task Allocation Table

1 Abstract

This report aims to discuss the motivations and developments behind a data product created to help first-time renters in London find an area to rent in. The primary goal of this product is to empower users who are unfamiliar with the city, to make well-informed decisions by addressing the challenges posed by fragmented and inaccessible information sources.

The first part of this report outlines the ideation process, which was heavily influenced by our own experiences as first-time renters in London. From this, we developed a user persona that helped identify the key performance indicators and data types users would find most valuable, such as crime statistics, access to amenities, grocery stores, and area-specific rental pricing. Next, datasets with relevant data points were curated pertaining to the different boroughs in London along with their geographical location coordinates and geometry to overlay on the map. Finally, the report discusses potential avenues for future development, including innovative features and enhancements that could further distinguish this product as a prominent tool for first time renters navigating the rental market in London.

2 Introduction

2.1 The Problem

Currently, finding a rental property in London can be overwhelming, particularly for newcomers who lack local knowledge. Important factors such as crime rates, the availability of amenities like grocery stores and restaurants, and rental affordability are dispersed across different platforms. Moreover, existing platforms fail to offer tailored area recommendations that align with a user's specific needs. The novelty in our approach lies in creating an integrated platform that bridges this gap, offering an intuitive, centralised solution for area selection.

2.2 Ideation and Motivation

Starting off with the ideation process, we reflected on the challenges we face as international students in the UK. It quickly became apparent that securing rental housing in a city like London, where housing availability is limited, is a significant challenge. Through initial research, it became evident that private renters in London often face considerable difficulties in identifying suitable properties. As first-time renters with limited knowledge about essential factors such as available amenities, neighborhood safety dynamics, and pricing trends, it is crucial to have a clear understanding of what to look for in the rental market.

Safety remains one of the most pressing concerns for renters. Analysing crime statistics and avoiding areas with higher crime rates can significantly reduce stress and mitigate the need for additional expenditures on security measures. For university students, who often return home late at night after extended study sessions, residing in neighborhoods that support safe and secure commuting is not just desirable but essential for peace of mind and overall well-being.

For university students who spend the majority of their day on campus, proximity to amenities that cater to their needs is crucial. Access to social hubs such as restaurants, cafes, and takeaway shops not only enhances convenience but also ensures a less hectic lifestyle. Moreover, many students rely on part-time jobs in nearby establishments for supplemental income, making such environments near their residence both practical and advantageous.

Finally, aspects like transport availability have been a significant focus of previous data products, particularly those providing real-time information on tube and bus schedules, such as Citymapper, which offers strong competition to Google Maps. It is reasonable to conclude that transport accessibility is generally very good to excellent across most areas of Greater London hence this won't take on a role in our data-product.

2.3 Why London

London is a vibrant tapestry of neighborhoods, each with distinct charm and appeal. From the historic streets of Notting Hill to the lively markets of Camden, the city offers something for

everyone[1]. Young professionals can enjoy a rich social life with bustling marketplaces, live music, and theaters, while those seeking tranquility can find peaceful havens like Dulwich or Kew, offering greenery and quiet escapes from urban noise. London has something in store for every type of user. Moreover, renting in London is a mutli-dimentional process as compared to the rest of the UK. Most areas outside London have more availability of rentals and little diversity in living styles.

This project aims to help first-time renters, particularly international students and those early in their careers to navigate London's rental market by addressing challenges related to safety, amenities, and pricing before they start looking for their ideal house to rent.

2.4 Target Audience

2.4.1 User Persona

Our target users are primarily international students and young professionals planning to move to London and rent a flat. There top needs are to live in a safe locality, obtain the cheapest rents and have amenities or external sources of income at a stones throw from their new residence. An example user persona is detailed in **Figure 1**.

2.4.2 User Journey

Once our users plan their move to London, they can navitage to our initial app to explore area-specific amenities and get an initial recommendation on what area suits their needs the best. Post **BoroughMatch**, they would continue to carry on detailed property searches with a better overview of London's landscape. This is highlighted in **Figure 2**

3 Product Requirements

3.1 Solution Description

Our data product enables users to explore various boroughs in London through a dashboard comprising four tabs. The first three tabs present descriptive insights on crime data, rental prices, and amenities across boroughs, providing users with an over- view of London. Users can delve deeper into specific areas using an interactive map. The final tab uses a linear programming optimisation algorithm to recommend the most suitable boroughs based on user preferences.

The crime data tab features a color-coded map where green indicates lower crime rates and red signifies higher crime rates. This map offers a comprehensive overview of total crimes in each borough, with an interactive feature allowing users to click on a borough for a detailed breakdown of various crime categories. This tab provides users with an insightful perspective on the safety levels across boroughs without having to look at additional resources and get haunted by regular crime updates.

The rental price map tab enables users to explore average rental prices across different boroughs, offering valuable insights for first-time renters seeking to understand rental costs in London. Users can begin by selecting the desired flat type, which dynamically updates the map in the main panel. Boroughs are color-coded, with green representing lower average prices and red indicating higher prices. Due to data-set limitations only borough-level average pricing could be incorporated in the present minimum viable product (MVP). Additionally, a bar chart displayed in the side panel illustrates the average monthly rental prices for the selected number of bedrooms. This would help the user navigate house rentals for not just living alone but also within house-shares, which is common within our target user-base.

The Amenity Map tab incorporates six types of amenity data: restaurants, cafes, takeaways, clubs, bars, and grocery stores. To enhance usability, users can adjust sliders in the side panel to set thresholds for each amenity of interest, highlighting regions that meet their preferences in light blue. Exact slider values are displayed to improve clarity and credibility. A layer control panel allows users to select specific amenities to avoid overwhelming the map visualisation with all layers. When multiple layers are displayed boroughs are color-coded with gradients, where darker colors indicate areas meeting more user-defined criteria. Once users narrow down their areas of interest, they can click on a borough to view
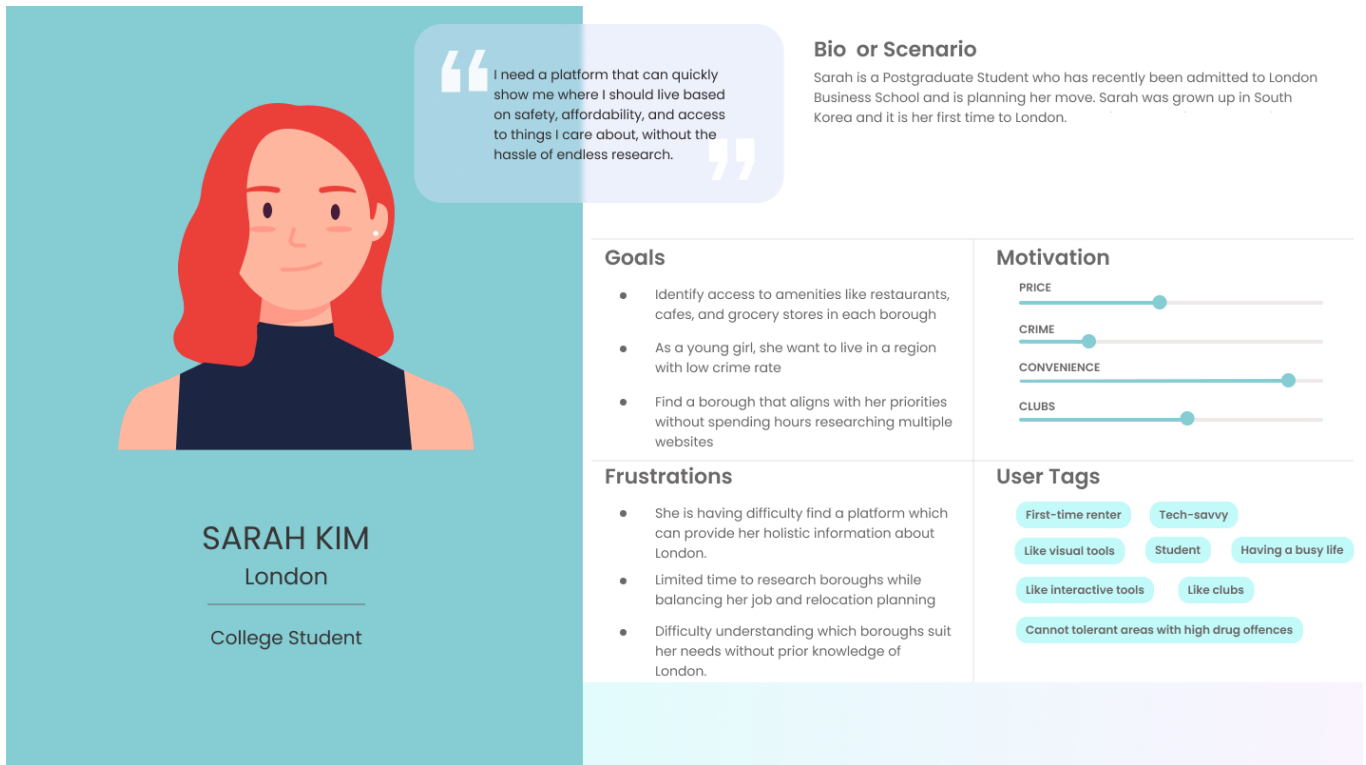
Figure 1: User Persona

Figure 2: User Journey

give weight to their most wanted amenity type and their least wanted crime type. Users can choose the price range they can accept using another slider at the bottom. The final optimal borough is dynamically updated as the user navigates their needs on our platform. Through this visualisation, the user gets a brief overview of the trade-offs that exist within the provided three metrics.

3.2 Expected Impacts

The expected impact of our project would be the successful compilation of currently scattered resources of information available to decide on a place to rent. Combining the functionality of individually searching for amenities on Google Maps, referring to crime maps and finding cheapest places to live according to your needs.

3.3 KPIs - Our Mission

BoroughMatch's early stage KPIs focus on three key areas:

- **Technical:** Ensure data accuracy using the latest official datasets, with a fully functional product delivering both descriptive and prescriptive features. Responsive front-end design

detailed amenity data through a popup window. The retailers layer provides the exact location and names of grocery stores, offering a clear view of retailer distribution across different areas.

Our final tab helps users to make better decisions by giving recommendations based on three perspectives (rental price, crime, and amenity). Through this final interpretation the user can be guided to the borough that most suits their needs for further research.

The sliders on the side panel allows users to

and accurate mapping are validated against source data.

- **User-Friendliness:** Provide an intuitive, interactive interface with clear layouts, responsive controls, and minimal confusion. Effectiveness is measured through user feedback on task completion and visualisation clarity.

- **Project Delivery:** Follow structured milestones, starting with defining objectives, creating mockups, and developing a Minimum Viable Product (MVP). The final stage integrates optimization algorithms for borough recommendations. Progress is tracked weekly over a six-week timeline.

3.4 Competitor Analysis

BoroughMatch addresses key gaps in the competitive landscape. While platforms like Crystal Roof provide crime and demographic data, they lack rental insights or personalised recommendations. Similarly, Rightmove and SpareRoom offer extensive rental listings but fail to integrate crime and amenity data that are core to the moving in process, requiring users to consult multiple sources.

Key differentiators of BoroughMatch include:

- **Integrated Data:** Combines rental, crime, and amenity data for holistic insights

- **Personalised Recommendations:** Utilises optimisation algorithms to suggest boroughs tailored to user preferences.

- **User-Friendly Interface:** Simplifies complex data through intuitive maps and plots making the rental process more streamlined.

4 Data Sourcing and Curation

In order to render the map, we first source the London borough data including the name, code, and geojson data. Individually, three different sets of data analysis and cleaning was done on separate files for the descriptive analysis - **"CrimeTest.R", "AmenityMap.R", "PricingMap.R"**. The maps for these regions was developed by mapping borough geojson data to the borough descriptive data. The output of this data analytics was curated in the **"main.R"** folder.

4.1 Amenity Data

The amenity data sourced included five types of amenities within each borough along with the corresponding borough area codes. In the original raw dataset borough names and codes were contained within a single column separated by a colon. This information was split to enable the use of the name for linking the borough GeoJSON data file to map amenity information accurately. For retailer data, the dataset includes the names of the retailers as well as their longitude and latitude coordinates.

4.2 Crime Data

The original crime dataset sourced from the London DataStore comprises borough-level monthly crime statistics categorised under various "major categories" of offenses. The data was aggregated to reflect the total crime count over the preceding year (October 2023–2024) as a subset of the last 1.5 years of data. The focus was restricted to borough-level data, excluding lower district levels (LSOAs), to provide an overview of crime rates and ensure consistency with data available for other descriptive metrics for better user insights. Borough-level data for the "City of London" was excluded due to the absence of collated crime data for the area and its separate police force distinct from the rest of London. For further development of this segment of our dashboard, we could incorporate LSOA or MSOA level data to have an indepth overview of each boroughs.

4.3 Pricing Data

The pricing data was initially obtained at both borough and individual ward levels categorised by the number of rooms. However, due to inconsistencies in the definition of London wards and limited user familiarity with these definitions only borough-level information was used. For further analysis, ward-level data was averaged within each borough for each house type (e.g., one-bedroom, two-bedroom, etc.). An initial bar plot visualisation showed a skewed towards higher prices for houses with more rooms. Consequently, the data for each column was individually standardised before being displayed on the map. For future development, this data

could be presented as a range of prices derived from a more comprehensive dataset.

5 Development Process

5.1 Conceptual Design - Front-End

The product is designed to be a one-stop-shop for all early stage rental look-outs hence the UI needed to be carefully designed to not look overwhelming. The main focus was the map in all four tabs.

The Crime tab uses a green-to-red color scheme to represent the range of "total crime types" across boroughs. Total crime numbers are not displayed to avoid overwhelming users, as understanding severity requires examining specific crime types, such as in high-crime areas like Westminster. Clicking on individual boroughs gives the user a holistic understanding of the crime types to make personal judgement.

The Amenities tab avoids overwhelming users with total amenity numbers, which are revealed only upon clicking a borough. Designed for user interaction, the tab has toggle bars with thresholds labeled "Min," "Less," "Average," "More," and "Max," offering a clearer representation compared to raw numerical values.

The Pricing Map features an overview bar chart ranking boroughs by price, from highest to lowest, to help users identify affordable areas. To ensure clarity, especially for houses with more rooms and skewed pricing data, axis "breaks" provide a detailed view of variations. The color scheme was created by standardising the data and applying quantile-based groupings, where boroughs are color-coded based on price ranges. This method ensures evenly distributed colors based on frequency of observations, even with skewed data, by adjusting group sizes according to the frequency of observations within each range.

The prescriptive analytics allows users to select their highest-priority amenity and exclude a specific crime type, simplifying the interface by avoiding over 20 permutations. A two-sided slider lets users set a price range, accommodating predefined budgets and flexibility for additional expenses. On the back end, the code works by formualting the following optimisation problem solved using linear programming:

$$\text{Maximize } S = W_a \cdot A - W_c \cdot C$$

Subject to:

$$P_{\min} \leq P \leq P_{\max}$$

where:

- S is the score to be maximized,

- A represents the amenity value,

- C represents the crime rate,

- W_a and W_c are the weights assigned to amenity and crime, respectively,

- P is the price, bounded by $P_{\min}$ and $P_{\max}$

5.2 Interaction

Our interaction design is accessible and engaging. Maps dynamically update based on users inputs from sliders and selection. We also leverage the hovering, clicking and popup window features for users to see more detailed data.

6 Testing and Validation

The integration between R and R Shiny was thoroughly tested to ensure proper functionality with calling functions and variables from files. Color-coded mapping leveraged quantile-based bins as opposed to consistent increments to evenly distribute data across the color scale, effectively highlighting differences between boroughs and helping user interpretation. All development errors and warnings were resolved.

User testing through interviews provided positive feedback, confirming the platform's utility, intuitiveness, and effectiveness in assisting first-time renters in London. Since our users were university students like us it was good to recieve lots of feedback during the development of this MVP. The product simplifies borough exploration with clear, interactive visualisations of safety, affordability, and amenities. The personalised recommendation feature, powered by linear programming optimisation, was particularly praised for helping users narrow down borough choices based on their preferences. General feedback

highlighted the need for more detailed area specifications, which require richer datasets and to some extent, enhanced R Shiny functionality. These limitations are discussed further in the next section.

Overall, the project successfully met all key performance metrics, including technical database cleaning, developing a user-centric dashboard, and achieving mini-milestones with consistent feedback guiding the final output.

7 Reflections and Future Steps

From the onset of the product development process, our team has acquired a user-centric design approach. The product ideas stem from user insights and user pain points, ensuring its relevance and effectiveness. The UI prioritises simplicity and ease of use while visualisations such as color-coded maps, pie charts, and bar charts, communicate the data in an effective and visually engaging manner, enhancing the user experience.

7.1 Limitations

While the current product provides a strong foundation, it still exhibits some limitations: The current product only offers insights at the borough level, making it suitable only as a preliminary search tool. In some cases, the situation within a borough can be varied. The product focuses on three key data sources: crime rates, rental prices, and amenities. While these are highly relevant to our target users, other important factors, such as transportation convenience and proximity to workplaces or universities, are not yet included. At the current stage, the optimisation algorithm considers only the most desired amenity, the least desired crime, and the rental price range as constraints. A more sophisticated approach with additional user preferences could yield more precise recommendations.

7.2 Future Steps

Looking ahead, we aim to further enhance the product value-proposition.

- The current product targets international students and young professionals, but we aim to expand our user base by incorporating a broader range of metrics that could benefit

more user groups while ensuring the interface remains intuitive and user-friendly.

- We also aim to source data at the Lower Super Output Area (LSOA) level to add more granularity and precision.

- To tailor recommendations more precisely, we will enable users to specify detailed preferences and assign the relative importance of various factors, such as prioritizing specific types of amenities (e.g., grocery stores over cafes) or placing higher weight on certain crime categories.

- Potentially using API form relocation services to provide users with seamless transitions from borough exploration to property searches.

We will also continuously gather user feedback through surveys and analytics to identify areas for improvement. By implementing these enhancements, our goal is to transform **BoroughMatch** into a comprehensive yet user-centric tool for first-time renters in London.

7.3 Conclusion

In conclusion, **BoroughMatch** successfully met all defined KPIs, delivering a user-friendly platform for first-time renters in London. With access to more sophisticated datasets and continuous user feedback, **BoroughMatch** has the potential to evolve into a premier tool for providing early-stage insights on choosing where to live. Looking ahead the platform could expand to other cities, integrating tailored parameters to meet the unique needs of diverse across the country. This ambitious vision would require extensive research and behavioral insights to adapt to the nuances of different countries and regions. **BoroughMatch** represents a significant step forward in simplifying the rental decision-making process through data-driven, user-centric solutions.

8 References

[1], https://mytravelaffairs.com/where-to-rent-a-house-and-which-neighbourhoods-to-avoid-in-london/Where to Rent a House and Which Neighbourhoods to Avoid in London.