Initial Plan: Identifying Urban Functional Regions

Author: Joshua Rimes c22002134

Supervisor: Padraig Corcoran

Project Description

With 84% of the population in the UK living in urban areas, it's important to understand the influences and impact the layouts of these areas have. Urban areas are a hive of activity that integrate functions and industries together to grow the population and economy.

The urban layout likely has a substantial effect on socio-economy and quality of living; an area that focuses on a specific function or industry might have a crucial impact on many areas such as health, wealth, and crime. These functional regions, such as nightlife or food economy areas, are the foundation of urban success and should be considered carefully in the design of the city.

When cities are developed, it's clear that there is no common process. In fact, the development of urban environments themselves could be seen as a part of the larger process of place-shaping and urban design (Carmona 2013). If we were able to find out the impacts of the layout of functional regions, we could change the approach of developing future cities to benefit urban growth.

This project aims to give an insight into this by identifying the functional regions of urban areas in an automated manner using point of interest (POI) data. A point of interest is a specific point location that is deemed useful or interesting, for example, a restaurant or a park. Using the resulting data, I will compare the layout of functional regions to data on the economy, urban growth, and quality of life, so that a conclusion can be made on the impacts of urban layout.

Considering the complexity of urban functional connotations and the characteristics of different types of data, identifying urban functional regions using data from a single source remains challenging (Yang et al. 2022). Automating data analysis and using that data along with research displays in-depth knowledge and application of computer science systems including spatial data, data mining, and data structures. In utilising my solution, it will prove that the use of computer science is imperative to the future of urban design and specifically to the problem I am posing.

Aims and Objectives

The overarching aim of this project is to discover the impact of urban layout on the economy and quality of life. To do this I will create a system that automates the identification of functional regions in a city using POI data.

Objective 1

I will implement a Python system that preprocesses geospatial POI data, automates the identification of urban functional regions using a clustering algorithm, and visualises the resulting data in an interactive manner. I am choosing to use Python for this system because it has many existing libraries for handling and processing data, this will assist me hugely in building something suited to the specific needs of my solution.

A risk of making the system myself is that the code I will need to write may require me to acquire new knowledge of libraries and algorithms that I might not understand in time to implement into the system. If this proves too challenging then I would likely have to mitigate the automated part of the system to make the programming easier, however, I believe there is a low risk of this happening, and I feel confident that I'll be able to learn and integrate these new algorithms into my system.

Objective 2

I aim to collect data about the functional regions and layout of urban areas by inputting their POI data into my system and recording the results. I've decided on analysing two cities to ensure the impact of region layout is not unique to one city. Both cities are of similar population and size so that it is fair to compare them both. Alongside those results, I will acquire data from online research and censuses about quality of life, average salaries, and economic growth in the functional regions of both cities.

The risk in this objective revolves around the access I have to data about both cities. POI data should be available for me to download without cost or ethical issues because it is public secondary data. Data obtained through research and previous censuses might be harder to find given that articles may be subjective and not truly represent the facts I want to use in my comparison. Because the main aim to seek the impacts of urban layout requires this data, I may need to focus on census and numerical data which could be harder to find.

Cardiff and Bristol are both in the southwest of the UK, meaning they might have similar economies and quality of life purely because of their location rather than because of

their layout, this poses another risk within this objective. To mitigate the risk, I would have to choose a different city in another part of the UK to compare to Cardiff, however, there are not many cities of similar population.

Objective 3

The final and most important objective of my project is to reach a conclusion on the impact of urban layout on quality of life and economy by comparing the success of different functional regions in cities such as Bristol and Cardiff, to show how the location of certain types of functional regions can play a positive or negative effect on other regions.

Finding that there is no impact on the success of industries would be just as conclusive as an extremely positive or negative impact, therefore there is no risk in reaching a conclusion.

Feasibility

The data I will be collecting in this project is all non-personal secondary data; therefore, I need no ethical approval to use it.

OS data downloaded from the Digimap portal service require agreement to an Ordinance Survey Educational User Licence. I can agree to this licence via my status as an undergraduate student at Cardiff University. I will also be using data from the 2021 UK census, this data requires no licencing to download and use.

To run the system once it has been built, there will be a few feasible requirements, these include a desktop or laptop with a processor and storage capable of dealing with small data sets, as well as the latest version of Python installed and a provided list of required libraries.

My solution is feasible because the data I need is accessible, and the system needs a low spec computer to code and run.

Work Plan

Week 1 – Writing the initial plan

Tasks:

- Write the four sections of the initial plan covering context, aims, feasibility, and work plan.

Milestone (☑): Complete Initial Plan

Deliverable (D**):** An Initial Plan document that describes the outline and plan for my solution and how I will achieve it.

Week 2 – Background reading & implementation research

Tasks:

- Research whether there are already views on the impact of urban layout.
- Write up what I've found in my research.
- Research the different datasets I can use to get POI data and which part of the 2021 census I want to analyse.
- Write up my research and decisions.

Milestone (☑): Background reading on current urban planning and layout standards

Milestone (☑): Choose and download datasets for POI data and census data

Week 3 – Implementation research

Tasks:

- Research previous projects and articles on region identification and POI data analysis to identify potential data analysis methods I could use in my system.

Milestone (☑): Review of related works on identifying functional regions

Week 4 – Implementation - Preprocessing

Tasks:

- Use geopandas to handle inputting geospatial data.
- Normalise data entries so that processing the data is consistent.

- Make a meaningful categorisation system for the POI data.
- Use a Python connection to PostgreSQL to create a storage for the data.

Milestone (☑): Data in the system will be categorised and stored in a spatial database

Milestone (☑): The system will be able to take in and clean data

Deliverable (D**):** A Python system for collecting data (used via a terminal at this stage).

Week 5 – Implementation - Clustering

Tasks:

- Program methods for calculating features such as POI density and proximity using Python's shapely library.

Milestone (☑): The system will be able to calculate relevant variables to be used in clustering

Week 6 – Implementation – Clustering & Visualisation

Tasks:

- Use methods from scikit-learn to create a DBSCAN clustering method that can define functional regions.
- Make a classifier for assigning functional labels to the clusters.

Milestone (☑): Clustering algorithms will be implemented into the system

Deliverable (D**):** A system that clusters collected POI data into functional regions, executed from a terminal.

Week 7 – Implementation - Visualisation

Tasks:

- Code a visual map using Folium that displays regions with different colours depending on categories.

Milestone (☑): The system displays the functional regions on a visual map

Week 8 – Implementation - Correlation

Tasks:

- Make overlay heatmaps to find a connection between the functional regions and socio-economic data.

Milestone (☑): Completed system with all functions

Deliverable (D**):** A system that runs from a terminal and opens an interactive map showing the regions calculated through clustering.

Week 9 – Testing & Analysis

Tasks:

- Write down testing criteria and tests.
- Test the system for ability to handle data, cluster the regions, and visualise the results.

Milestone (☑): Successful testing

Week 10 - Analysis

Tasks:

- Use the regions generated by my system to compare the success of industries concerning the regions they're in proximity to.
- Analyse multiple urban areas to remove any anomalies.
- Write up my conclusion on the impacts of urban layout.

Milestone (☑): Completed analysis of the impact of functional regions

Week 11 - Report

Tasks:

- Write the final report on my project, starting with the following sections: Introduction, Background, Methodology, Implementation.

Week 12 - Report

Tasks:

- Finish writing the report on my project with the following sections: Testing, Evaluation, Conclusion, & Learning Reflections.

Milestone (☑): Completed the final report

Deliverable (D): Final report.

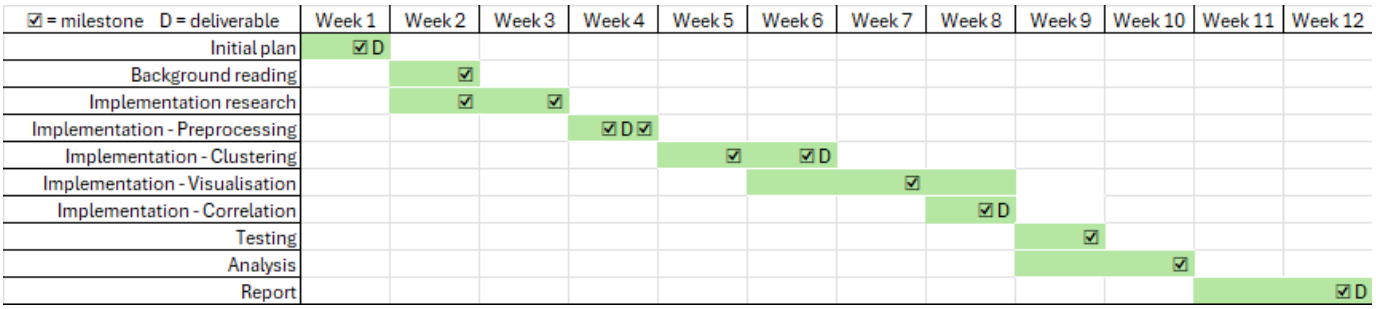

Figure1. Gantt chart visualising my 12-week work plan

References

Carmona, M. 2013. The Place-shaping Continuum: A Theory of Urban Design Process. *Journal of Urban Design* 19(1), pp. 2–36. Available at: https://www.tandfonline.com/doi/full/10.1080/13574809.2013.854695#d1e681.

Yang, M., Kong, B., Dang, R. and Yan, X. 2022. Classifying urban functional regions by integrating buildings and points-of-interest using a stacking ensemble method. *International Journal of Applied Earth Observation and Geoinformation* 108, p. 102753. Available at: https://www.sciencedirect.com/science/article/pii/S0303243422000794#ab005.