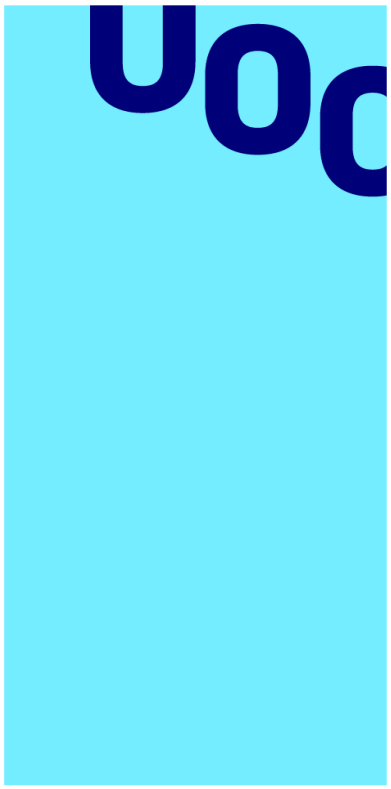
Are U Query-ous? A Web-Based Platform for Democratizing Open Geospatial Data Access

From Queries to Maps, A New Way to See the World!



Universitat Oberta de Catalunya

TFG - Localization Based Systems and Intelligent Spaces

Bachelor's degree in Techniques for Software Application Development

Author:

Nicolas D'Alessandro Calderon

Project supervisor:

Joaquín Torres Sospedra

Coordinating professor:

Antoni Perez-Navarro

Abstract

This study explores the challenge of making open data more accessible to the general public, addressing the gap between the availability of geospatial information and its practical use. *Are U Query-ous?* is a web-based application designed to enable individuals, regardless of their technical background, to explore and interpret geographic and demographic data intuitively. By integrating interactive maps and user-friendly visualization tools, the platform allows users to analyze regions based on economic activity, population distribution, and local trends.

Through a simplified interface, users can explore and filter publicly available data, identifying patterns relevant to their interests. The application is intended for individuals who are curious about urban development, seeking optimal locations for personal or professional activities, or analyzing demographic trends for research or decision-making. Additionally, the project examines the potential integration of artificial intelligence to facilitate data retrieval through natural language queries, further enhancing accessibility.

The development and results of the analyses in this data scientific report are intended to address all the concepts targeted in the foregoing. The author also expects to depict some fundamental principles underlying web apps development and data science.

Keywords

Geospatial Data, Open Data Visualization, Interactive Maps, Urban Analytics, Location Intelligence

Index

List of figures

01 INTRODUCTION

02 PLANNING

03 METHODS AND RESOURCES

List of Tables

02 PLANNING

01 INTRODUCTION

1.1 Summary of the Proposal

This final project focuses on developing an intuitive **geo-analytics platform** that simplifies the exploration of open data through **interactive maps**. Many open data sources provide valuable insights into **urban planning, demographics, and economic activity**, but non-experts often struggle to extract meaningful information from them.

Are U Query-ous? aims to solve this problem by creating a **user-friendly interface** that enables individuals to explore and analyze spatial data without requiring technical expertise.

Additionally, if time permits, the project will explore the integration of a natural language processing model, such as **those available from Hugging Face**, to enable users to interact with the data using natural language queries. **This feature is considered an enhancement rather than a core requirement and will be evaluated based on project timelines** and feasibility.

At the end of the project, the system will provide a **fully functional prototype** that allows users to:

- **Filter** and **visualize** location-based open data.
- Identify **regional patterns and trends** based on economic and demographic factors.
- Utilize **intelligent search capabilities** to access relevant insights more intuitively.

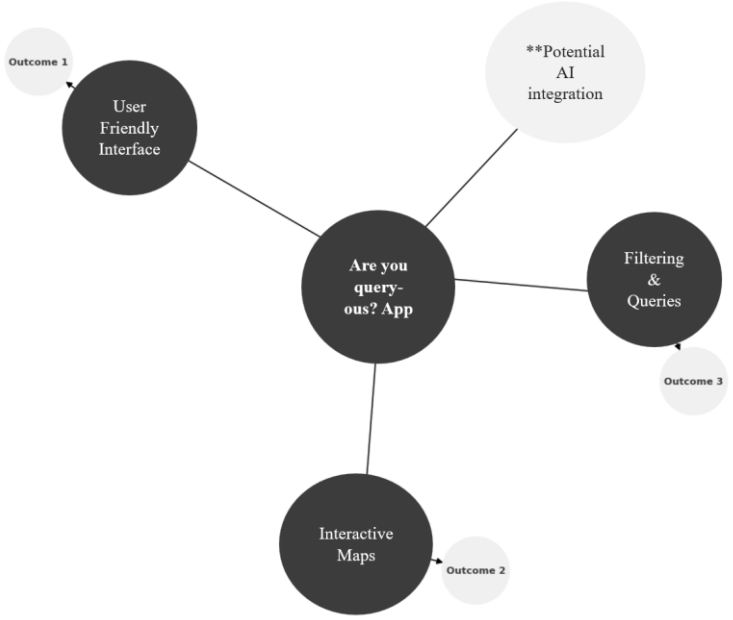


Figure 1 1: Conceptual Representation of Project Summary

1.2 Scope of the Project

This project focuses on developing a web-based platform that allows users to explore and analyze open geospatial data in an intuitive way. The platform will integrate publicly available datasets from Barcelona and Madrid, two cities with well-structured open data portals that provide reliable and detailed information.

The scope includes the following key aspects:

- **Cities Covered:** The platform will use **open data from Barcelona and Madrid**, ensuring access to urban mobility, economic activity, and demographic datasets for meaningful analysis.
- **Data Integration**: The system will collect and process geospatial, economic, and mobility-related data, allowing users to filter and visualize insights interactively.
- **User Interaction**: The web application will feature an interactive map where users can explore regional trends, compare different areas, and extract useful insights without needing technical expertise.
- **Core Functionalities:**
    - *Mapping & Visualization*: Users will view geospatial data overlaid on maps, with filtering options.
    - *Urban & Economic Insights*: The system will present mobility trends, population density, and economic indicators based on selected areas.
    - *Accessibility & Usability*: The interface will be simple and user-friendly, ensuring that both professionals and non-experts can use it effectively.

By **limiting the project scope to Barcelona and Madrid**, the system will leverage well-organized open datasets while **maintaining a manageable level of complexity within the project timeline**. The structured data availability from these cities will support the development, testing, and validation of the platform, ensuring that it meets its intended objectives.

1.2 Justification

Access to open data has grown exponentially, yet many users struggle to transform this data into actionable insights. While businesses and government entities benefit from sophisticated **geo-analytics tools**, individuals and small organizations often lack the resources or expertise to use these datasets effectively. This project is relevant because it seeks to **bridge this gap**, making open data truly accessible and usable for the **general public, students, researchers, and small businesses**.

Moreover, the relevance of geospatial analysis has expanded in fields such as **smart cities, sustainable urban development, and socio-economic research**. By providing an easy-to-use tool, this project supports the broader goal of promoting **data-driven decision-making at all levels of society**.

1.3 Motivation

In the past eight years I have been working and developing my professional career in the field of **data analysis**, so this project aligns with both academic and professional aspirations. The motivation for this project is coming from:

- A personal interest in **making complex data more understandable** for non-experts.
- A desire to provide **access to urban and economic and geomarketing insights** through intuitive visualization for the general public, students, researchers, and small businesses.
- The opportunity **to apply geospatial analytics in a real-world application**.

Additionally, the skills developed through this project, including data processing, backend development, frontend visualization, and user interface design, will be valuable in both academic research and professional settings.

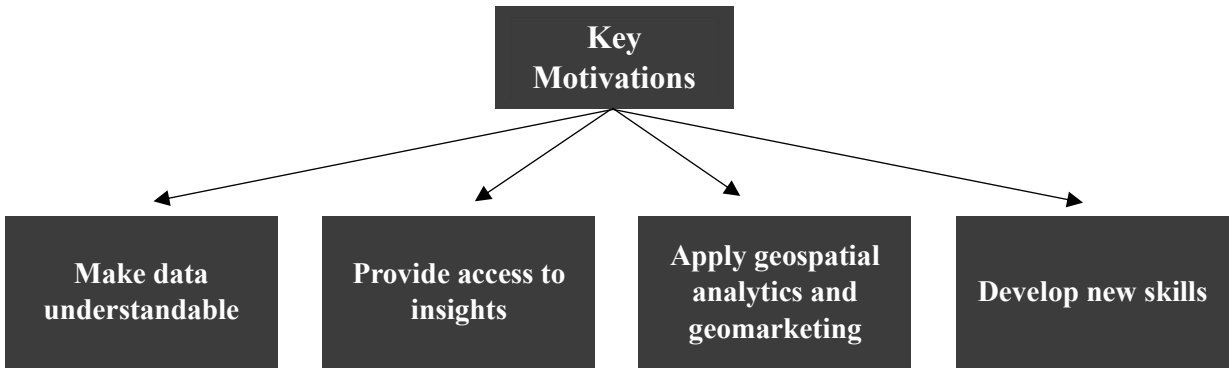


Figure 1 2: Key Motivations for the Project

1.4 Objectives

1.4.1 Main Objective

- To **develop a web-based app that enables users to intuitively explore, filter, and analyze open geospatial data**, making location-based intelligence more accessible to a non-technical audience.

1.4.2 Sub-Objectives

- **Develop an interactive mapping system** that allows users to visualize open data in an intuitive and engaging way.
- **Implement filtering and querying functionalities** to help users refine their search and extract relevant insights.
- **Ensure usability and accessibility** by designing a simple and intuitive user interface.
- **If time permits, experiment with integrating an NLP model** to allow natural language queries for filtering and searching data. This is considered a potential enhancement and not a primary project requirement.
- **Validate the effectiveness of the platform** through user feedback and iterative improvements.

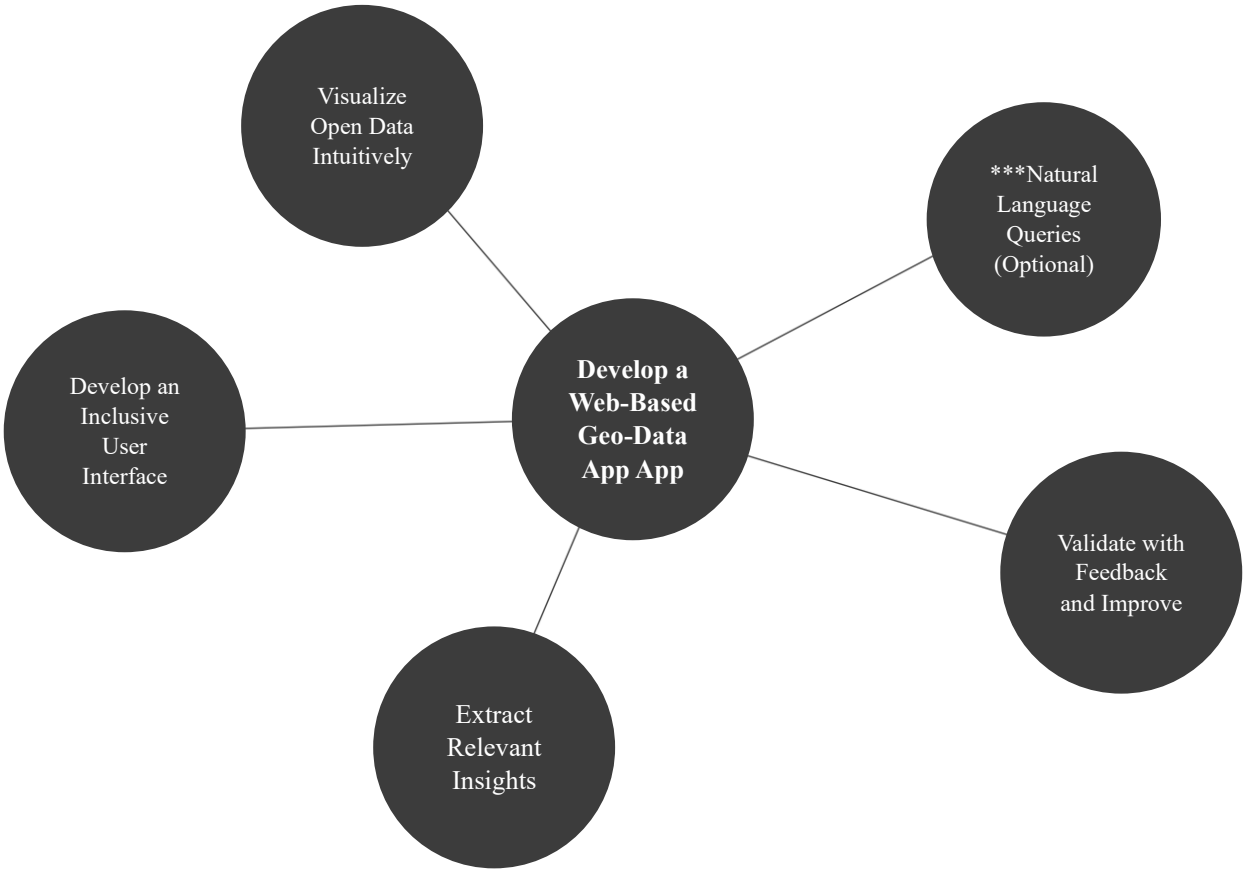


Figure 1 3: Objectives Breakdown

1.5 Hypothesis

Developing an **interactive web-based platform for querying and visualizing open geospatial data** will improve accessibility for non-technical users, allowing them to explore urban and economic data more intuitively. This will be evaluated through a structured user testing phase with a selected group of participants (e.g., students, professors, family, friends, etc.) who will provide qualitative feedback and complete pre-defined tasks.

1.6 Sustainability, diversity, and ethical/social challenges

1.6.1 Sustainability

This project promotes sustainability by supporting the efficient **use of open data to improve urban and social planning**. By making geospatial information more accessible, individuals and organizations can make informed decisions about resource allocation, mobility, and land use, reducing unnecessary waste and inefficiencies. Additionally, the project aligns with the United Nations Sustainable Development Goals (SDGs), particularly Goal 11 (Sustainable Cities and Communities), by providing insights that encourage responsible urban development. The project has a minimal ecological footprint as it relies on existing digital infrastructure, avoiding additional resource consumption.

1.6.2 Ethical behaviour and social responsibility

The project considers ethical principles by **ensuring data privacy and security**. Since the system processes **publicly available open data**, it does not involve personal or sensitive information. However, the potential risks of misinterpreting data or using insights unethically are acknowledged. To mitigate this, the platform will provide **transparent data sources and disclaimers to ensure users understand the limitations of the information**.

1.6.3 Diversity, gender and human rights

The project is designed to be inclusive and accessible, allowing all individuals, regardless of background, gender, or technical expertise, to explore geospatial data. By offering a simple and user-friendly interface, it ensures that users with limited data experience can engage with geographic insights. Furthermore, **accessibility features will be considered**. The project aligns with the principle of equal access to information, promoting diversity and reducing barriers to data-driven knowledge.

02
PLANNING

2.1 Project Planning

A detailed **schedule with milestones** has been defined, ensuring that the development process is structured, manageable, and aligned with the semester timeline.

CAT1: Project Planning Phase (Feb 19 - Mar 04)

Task ID	Task Description	Start Date	End Date	Dependencies	Status
1.1	**Define project scope and objectives**	Feb 19	Feb 20	-	Done
1.2	**Research available open geospatial datasets for Barcelona and Madrid**	Feb 20	Feb 22	1.1	Done
1.3	**Identify key visualization requirements for non-technical users**	Feb 21	Feb 22	1.1	Done
1.4	**Define technology stack and architecture**	Feb 22	Feb 24	1.1, 1.2	Done
1.5	**Assess potential integration challenges with open data sources**	Feb 23	Feb 24	1.2	Done
1.6	**Create detailed project timeline and milestones**	Feb 24	Feb 26	1.1, 1.4	Done
1.7	**Document sustainability, ethical, and diversity considerations**	Feb 26	Feb 27	1.1	Done
1.8	**Set up development environment and repository structure**	Feb 27	Feb 28	1.4	Done
1.9	**Compile comprehensive CAT1 documentation**	Feb 28	Mar 02	1.1-1.8	Done
1.10	**Review and finalize CAT1 documentation**	Mar 02	Mar 04	1.9	Done

Table 2. 1: Project Planning Phase

CAT2: Design Phase (Mar 05 - Apr 01)

Task ID	Task Description	Start Date	End Date	Dependencies	Status
2.1	**Define comprehensive user stories and use cases**	Mar 05	Mar 08	1.10	To Do
2.2	**Create database schema for geospatial and demographic data**	Mar 08	Mar 11	2.1	To Do
2.3	**Design UI wireframes for interactive mapping interface**	Mar 09	Mar 12	2.1	To Do
2.4	**Design system architecture diagram**	Mar 12	Mar 14	2.1, 2.2	To Do
2.5	**Research and document filtering algorithms for geospatial data**	Mar 14	Mar 17	2.1	To Do
2.6	**Create class diagrams for frontend components**	Mar 17	Mar 19	2.3, 2.4	To Do
2.7	**Design API endpoints for backend services**	Mar 19	Mar 21	2.4	To Do
2.8	**Implement login screen prototype**	Mar 21	Mar 24	2.3, 2.6	To Do
2.9	**Implement basic map visualization prototype**	Mar 24	Mar 28	2.8	To Do
2.10	**Document state of the art in geospatial data visualization**	Mar 16	Mar 23	-	To Do
2.11	**Compile and review CAT2 documentation**	Mar 28	Apr 01	2.1-2.10	To Do

Table 2. 2: Design Phase

CAT3: Implementation Phase (Apr 02 - May 06)

Task ID	Task Description	Start Date	End Date	Dependencies	Status
3.1	**Set up backend database with PostGIS**	Apr 02	Apr 04	2.11	To Do
3.2	**Implement data ingestion pipelines for Barcelona open data**	Apr 04	Apr 08	3.1	To Do
3.3	**Implement data ingestion pipelines for Madrid open data**	Apr 08	Apr 12	3.1	To Do
3.4	**Develop FastAPI backend services**	Apr 05	Apr 14	3.1	To Do
3.5	**Implement authentication and user management**	Apr 14	Apr 17	3.4	To Do
3.6	**Develop React frontend components for map visualization**	Apr 10	Apr 18	2.9	To Do
3.7	**Implement filtering and querying functionalities**	Apr 18	Apr 22	3.4, 3.6	To Do
3.8	**Integrate Leaflet.js for interactive mapping**	Apr 22	Apr 25	3.6	To Do
3.9	**Implement data visualization components for demographic insights**	Apr 25	Apr 29	3.6, 3.8	To Do
3.10	**Conduct unit testing for backend components**	Apr 29	May 01	3.4	To Do
3.11	**Perform integration testing of frontend and backend**	May 01	May 04	3.6, 3.9	To Do
3.12	**Deploy prototype to development environment**	May 04	May 05	3.10, 3.11	To Do
3.13	**Compile implementation documentation for CAT3**	May 01	May 05	3.1-3.12	To Do
3.14	**Review and finalize CAT3 documentation**	May 05	May 06	3.13	To Do

Table 2. 3: Implementation Phase

CAT4: Final Product & Report (May 07 - Jun 03)

Task ID	Task Description	Start Date	End Date	Dependencies	Status
4.1	**Refine UI based on testing feedback**	May 07	May 10	3.14	To Do
4.2	**Implement advanced filtering capabilities**	May 10	May 14	4.1	To Do
4.3	**Add economic activity visualization features**	May 14	May 17	4.1	To Do
4.4	**Optimize database queries for performance**	May 17	May 20	4.1	To Do
4.5	**Implement (optional) NLP capabilities for natural language queries**	May 20	May 24	4.1	To Do
4.6	**Conduct comprehensive system testing**	May 24	May 27	4.2-4.5	To Do
4.7	**Fix bugs and performance issues**	May 27	May 29	4.6	To Do
4.8	**Prepare final deployment package**	May 29	May 30	4.7	To Do
4.9	**Create user documentation and installation guide**	May 25	May 29	4.2-4.5	To Do
4.10	**Write results and analysis sections for final report**	May 15	May 25	-	To Do
4.11	**Draft conclusions and future work sections**	May 25	May 28	4.10	To Do
4.12	**Compile comprehensive bibliography**	May 28	May 30	4.10, 4.11	To Do
4.13	**Finalize and format complete project report**	May 30	Jun 02	4.9-4.12	To Do
4.14	**Final review and submission of CAT4**	Jun 02	Jun 03	4.8, 4.13	To Do

Table 2. 4: final Product and Report

CAT5: Presentation Preparation (Jun 04 - Jun 10)

Task ID	Task Description	Start Date	End Date	Dependencies	Status
5.1	**Create presentation outline and storyboard**	Jun 04	Jun 05	4.14	To Do
5.2	**Design presentation slides**	Jun 05	Jun 07	5.1	To Do
5.3	**Prepare demonstration script and data**	Jun 06	Jun 07	5.1	To Do
5.4	**Record demonstration of key platform features**	Jun 07	Jun 08	5.2, 5.3	To Do
5.5	**Create narrated presentation video**	Jun 08	Jun 09	5.2, 5.4	To Do
5.6	**Review and finalize presentation**	Jun 09	Jun 10	5.5	To Do

Table 2. 5: Presentation Preparation Phase

Defence Preparation (Jun 11 - Jun 17)

Task ID	Task Description	Start Date	End Date	Dependencies	Status
6.1	**Prepare defence presentation based on feedback**	Jun 11	Jun 13	5.6	To Do
6.2	**Anticipate potential questions and prepare responses**	Jun 13	Jun 15	6.1	To Do
6.3	**Practice presentation delivery**	Jun 15	Jun 16	6.1	To Do
6.4	**Final defence presentation**	Jun 17	Jun 17	6.1-6.3	To Do

Table 2. 6: Defence Preparation Phase

Ongoing Tasks Throughout the Project

Task ID	Task Description	Start Date	End Date	Dependencies	Status
7.1	**Trello board updates**	Feb 19	Jun 17	-	To Do
7.2	**Weekly supervisor check-ins**	Feb 19	Jun 17	-	To Do
7.3	**Documentation updates**	Feb 19	Jun 03	-	To Do

Table 2. 7: Ongoing Tasks Throughout the Project

For a detailed breakdown of the project plan, tasks, and timeline, please refer to the attached **are-you-queryous-planning.xlsx file.**

2.2 Expected Outcome

By the end of the semester, the project will deliver:

- A functional web application where users can **explore and visualize geospatial data** interactively.
- An intuitive filtering system allowing users **to refine results based on key indicators**.
- A structured API serving **open datasets** with a focus on usability and efficiency.
- A research report detailing **the impact of accessible open data visualization**.

This project contributes to the broader goal of **making open data actionable and meaningful** for a **diverse audience**, reinforcing the importance of **geospatial intelligence in everyday decision-making**.

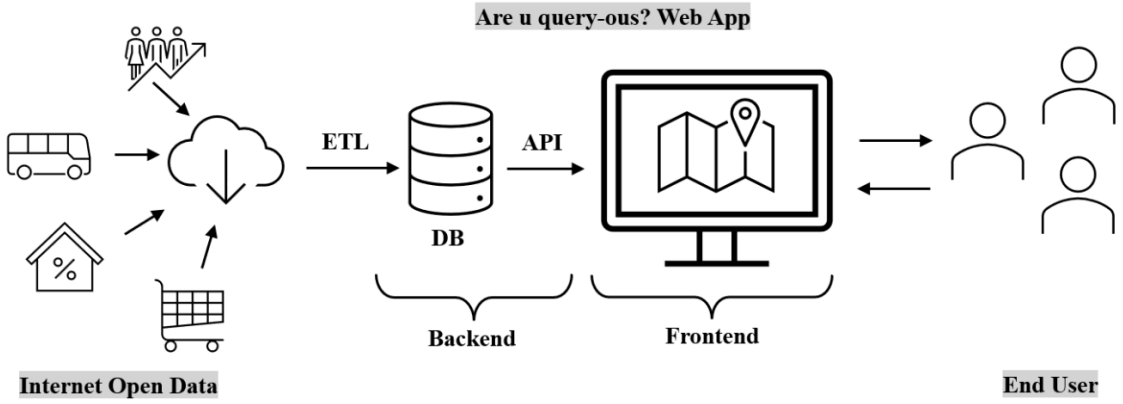


Figure 2 1: Expected Outcome Visual Representation

03
METHODS AND RESOURCES

3.1 Development Approach and Methodology

This project follows a structured development process **based on agile principles**, allowing for **iterative testing and feedback** throughout the semester. The key steps include:

3.1.1 Data Collection and Processing

- **Identify and integrate open datasets** (demographic, economic, urban mobility, etc.).
- **Clean and preprocess data** to ensure usability in the application.

3.1.2 Backend Development

- Build a **RESTful API** to serve geospatial data.
- **Store data in a database** optimized for efficient queries.

3.1.3 Frontend Development

- **Design a responsive user interface** with map-based interaction.
- **Implement data visualization tools** for filtering and exploration.

3.1.4 Testing and Refinement

- **Gather feedback** to enhance usability and features.

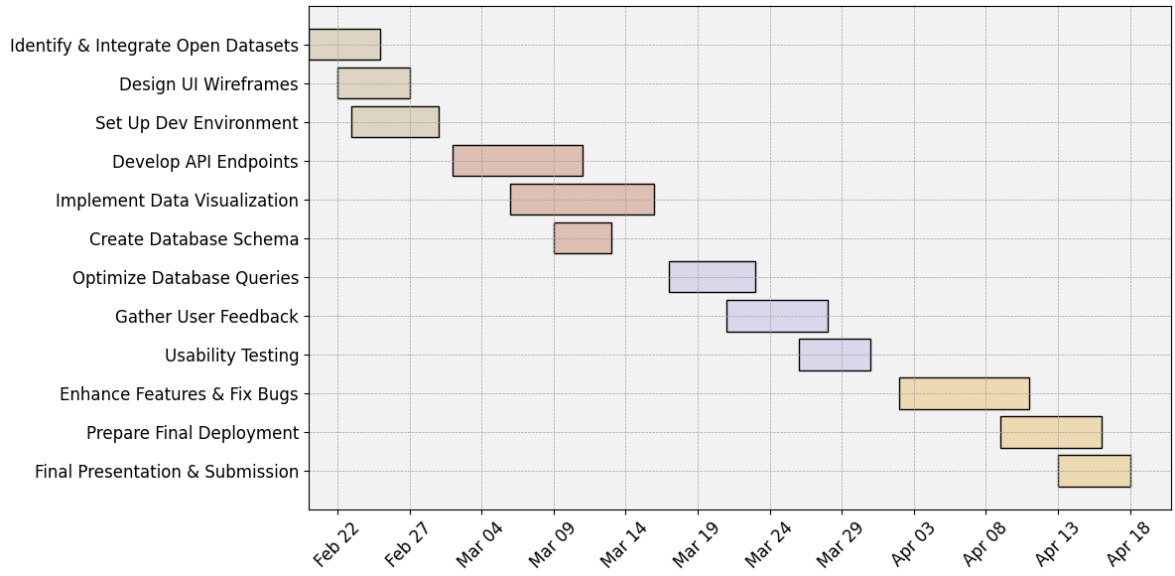


Figure 3 1: Process Diagram

3.1.5 Project Management and Development Workflow

To manage the development process efficiently, we will use a **Kanban dashboard in Trello**. The board will include the following columns:

- **To Do:** List of all planned tasks and ideas.
- **Doing:** Tasks that are ready to be worked on.
- **Deferred:** Tasks that are postponed for later stages.
- **Done**: Completed tasks.

This dashboard will include all the detailed tasks from the previous planning (Task 02) and will be updated regularly to track progress and keep the project on schedule.

3.2 Technology Stack and Implementation Details

The project will leverage a modern web technology stack to ensure efficiency, scalability, and a seamless user experience. The core technologies include:

3.2.1 Frontend:

- **React.js** for building a dynamic and interactive user interface.
- **Leaflet.js** for mapping and geospatial visualization.

3.2.2 Backend:

- **FastAPI** to handle API requests and serve processed geospatial data.
- **PostgreSQL/PostGIS** for storing and querying spatial data efficiently.

3.2.3 Data Processing & Integration:

- **Python** and **Pandas** for data preprocessing and transformation.
- **GeoJSON** format to represent geospatial data and serve it dynamically.

3.2.4 Deployment & Hosting:

- **Docker** for containerized development and deployment.
- **GitHub Codespaces** for cloud-based development and collaboration.
- **Vercel** for frontend deployment and **Fly.io** or **Heroku** for backend deployment.

3.2.5 Potential AI Integration (Time-Permitting Feature):

- The project may experiment **with integrating a pre-trained NLP model from Hugging Face** to process natural language queries.
- If implemented, a small widget will allow users to enter queries in plain text (*e.g., "Show me the most populated districts in Madrid").*
- This AI-based functionality is considered an exploratory addition, meaning it will only be developed if time and resources allow.

These technologies will enable the development of a robust and scalable application that can effectively serve users with varying levels of expertise in data analysis and geospatial exploration.

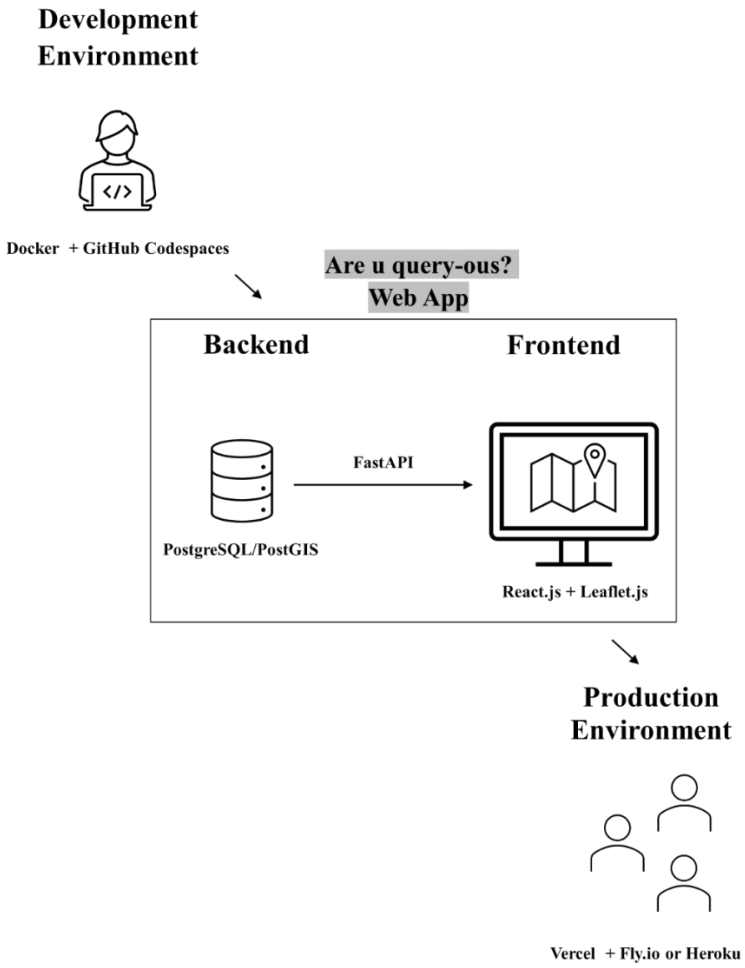


Figure 3 2: Technology Stack Used

04
RESULTS

This section will be further developed in future Continuous Assessment Tests (CATs) and project iterations. Additional details, refinements, and validations will be incorporated as the project progresses.

05
CONCLUSIONS AND FUTURE WORK

This section will be further developed in future Continuous Assessment Tests (CATs) and project iterations. Additional details, refinements, and validations will be incorporated as the project progresses.

06
GLOSSARY

This section will be further developed in future Continuous Assessment Tests (CATs) and project iterations. Additional details, refinements, and validations will be incorporated as the project progresses.

07
BIBLIOGRAPHY

This section will be further developed in future Continuous Assessment Tests (CATs) and project iterations. Additional details, refinements, and validations will be incorporated as the project progresses.

08
APPENDICES

This section will be further developed in future Continuous Assessment Tests (CATs) and project iterations. Additional details, refinements, and validations will be incorporated as the project progresses.