

# Exploring the Impact of Environmental Factors on Pedestrian Footfall in Dublin City

Alphons Zacharia James

*Msc in Data Analytics*

*National College of Ireland*

Dublin, Ireland

x23169702@student.ncirl.ie

Ankith Babu Joseph

*Msc in Data Analytics*

*National College of Ireland*

Dublin, Ireland

x23185813@student.ncirl.ie

Abhilash Janardhanan

*Msc in Data Analytics*

*National College of Ireland*

Dublin, Ireland

x23121424@student.ncirl.ie

**Abstract**—This paper presents an analysis of pedestrian footfall data in Dublin city while correlating it with weather conditions and air quality indices to understand and explore the relation between environmental factors and footfall. Using advanced analytics programming techniques and database management systems we process large datasets to extract actionable insights. Our method involves extracting and transforming data within a Docker container, this approach makes sure that our approach are scalable and can be consistently replicated. Our/The study uses statistical analysis and visualizations to identify trends, changes and irregularities in footfall data across different times and locations. Our findings indicates significant correlations between pedestrian traffics and both weather patterns and pollution levels. This study can help with urban planning, public health, and environmental changes. This paper helps us to present a new way to use footfall data in Dublin city to understand urban mobility and environmental effects. It shows how it can combine database technology and analytics programming for the society.

**Index terms**—Weather, Air Quality Index, Pedestrian Footfall, MongoDB, PostgreSQL, Docker, Dagster, ETL-Extract Transform Load

## I. INTRODUCTION

In an era where urban spaces are increasingly monitored for efficiency and safety, understanding the dynamics of human mobility becomes very crucial. Pedestrian footfall data serves as a significant indicator of urban vibrancy and economic activity, offering insights into how people interact with their environment under varying conditions. This research taps into the intersection of environmental factors—specifically weather conditions and air quality—and pedestrian movement within urban landscapes.

Recent advancements in data collection and analytics have enabled more detailed observations of urban environments, allowing researchers to analyze patterns that were previously

obscured. The integration of big data technologies and analytical programming facilitates a deeper understanding of complex datasets, uncovering relationships that can inform public policy, urban planning, and resource allocation.

The objective of this study is to employ a rigorous analytical approach to discern the impact of environmental variables on pedestrian footfall. By harnessing extensive datasets encompassing weather patterns, air quality indices, and hourly footfall across multiple urban locations, this paper aims to:

Identify trends in pedestrian traffic relative to environmental changes. Evaluate the influence of air quality and weather conditions on urban mobility. Provide actionable insights that can assist policymakers and urban planners in designing more livable and responsive cities. The significance of this study lies in its potential to contribute to the sustainable development of urban areas, enhancing the well-being of residents and promoting environmentally friendly urban planning practices. As cities continue to grow, the need to integrate environmental health with urban design becomes increasingly important, making studies like this one vital for future urban development strategies. To integrate the research objectives into the introduction section of your paper effectively, I will elaborate on the specific goals that the study aims to achieve within the context of the provided information. Here's the revised introduction with the addition of the research objectives:

In an era where urban spaces are increasingly monitored for efficiency and safety, understanding the dynamics of human mobility becomes crucial. Pedestrian footfall data serves as a significant indicator of urban vibrancy and economic activity, offering insights into how people interact with their environment under varying conditions. This research taps into the intersection of environmental factors—specifically weather conditions and air quality—and pedestrian movement within urban landscapes.

Recent advancements in data collection and analytics have enabled more detailed observations of urban environments, allowing researchers to analyze patterns that were previously obscured. The integration of big data technologies and analyti-

cal programming facilitates a deeper understanding of complex datasets, uncovering relationships that can inform public policy, urban planning, and resource allocation.

**Research Objectives:** The primary aim of this study is to elucidate the interconnections between environmental conditions and urban footfall patterns, thereby providing a robust analytical basis for urban planning and health-related policymaking. Specifically, the research will:

- Identify trends and correlations between various weather conditions (like temperature, humidity, and precipitation) and changes in pedestrian footfall.
- Analyze the impact of air quality levels on pedestrian density in different urban locales, highlighting how pollution may deter or alter pedestrian behaviors.
- Develop predictive models to forecast changes in footfall based on anticipated environmental conditions, supporting dynamic urban planning and public health initiatives.
- Map spatial and temporal footfall variations, providing urban planners with detailed insights into peak and off-peak pedestrian traffic patterns across different urban zones.

The significance of this study lies in its potential to contribute to the sustainable development of urban areas, enhancing the well-being of residents and promoting environmentally friendly urban planning practices. As cities continue to grow, the need to integrate environmental health with urban design becomes increasingly important, making studies like this one vital for future urban development strategies **【3】**.

## II. RELATED WORK

In this paper we explore the intersection of urban footfall and environmental factors in Dublin city using datadriven approaches to analyse pedestrian traffic patterns. Our work is informed by similar studies in the field, each contributing different insights into how urban environments influence pedestrian behaviors and wellbeing.

One such study conducted by Anita Ratnasari Rakhmatulloh, Diah Intan Kusumo Dewi, and Dinar Mutiara Kusumo Nugraheni in Semarang, Indonesia, focused on how urban design influences pedestrian activity. They employed ArcGIS for spatial mapping and SPSS for data analysis, concluding that pedestrian frequency is higher in areas with diverse attractions and well-planned spaces. These findings align with our investigation into Dublin's urban layout, reinforcing the importance of strategic urban planning in enhancing pedestrian experiences and environmental quality.[1]

Similarly a study by Babatunde Olasunkanmi Folasayo and Abimbola A. Babatunde examined the impact of environmental pollution on pedestrians in Lagos, Nigeria. They utilized the Thermo Scientific MIE pDR-1500 instrument to measure air

quality across 20 local government areas, finding that six areas exceeded acceptable pollution standards. The results, substantiated by a one-sample T-test with a t-value of 22.226, underscore the significant effect of air pollution on pedestrian health and underscore the need for urgent governmental interventions like restricted vehicle hours and enhanced urban greenery.[2]

## III. METHODOLOGY

The study employs a structured Knowledge Discovery in Databases (KDD) lifecycle ensuring a systematic approach from data gathering to knowledge extraction. The refined KDD lifecycle, as depicted in the updated Figure 1, consists of the following stages:

### A. *Data Selection:*

The first step in our study is choosing the right datasets. We use historical pedestrian footfall data which is collected from SmartDublin in CSV format along with historical weather and air quality data, which we gather through APIs from Open-Meteo. This two approach allows us to combine different types of data thereby giving us a clear picture of historical trends. This method make sure that we have a rich dataset that captures both the human activity in urban spaces and the environmental conditions they experience.

### B. *Storage and Preprocessing/Transformation:*

Upon collection, the data is stored in MongoDB, a NoSQL database which is utilized for its efficiency in handling large volumes of unstructured or semi-structured data. Python scripts are then employed for preprocessing and transformation tasks, cleaning the data and converting it into a unified format conducive to analysis.

### C. *Storage and Visualizations:*

Post-transformation, the data is loaded into a PostgreSQL database, an SQL database renowned for its robust data warehousing and complex querying capabilities. This transition marks a shift towards more intensive data analysis and the subsequent creation of visualizations.

### D. *Knowledge:*

The culmination of the process is the extraction of actionable knowledge. Data visualizations translate complex datasets into interpretable graphics, facilitating insight generation and supporting decision-making processes.

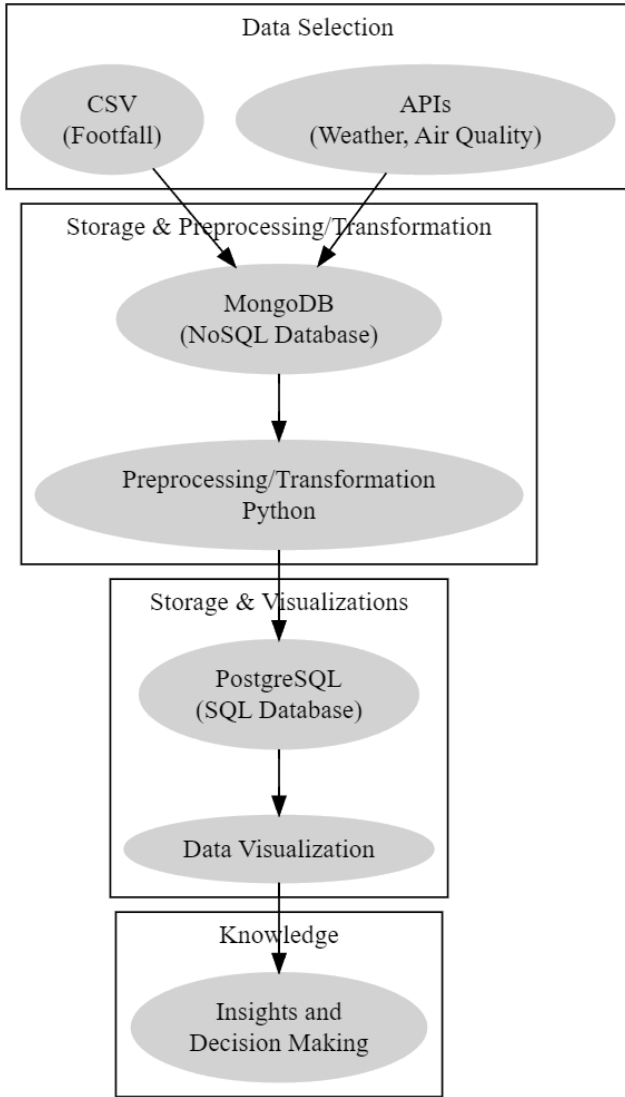


Figure 1: KDD

The iterative nature of this KDD lifecycle fosters the refinement of analysis methods, ensuring the research remains closely aligned with the defined objectives. The entire process, meticulously detailed in Figure 1, underscores the commitment to extracting meaningful and actionable knowledge from vast datasets.

#### E. Technologies:

In our database and analytical programming project a variety of technologies are used to simplify data processing and analysis. We organised our coding efforts to maximize efficiency and maintainability using python's adaptability and strength. In python we incorporated several libraries like Pandas, dagster, Pymongo, sqlalchemy, bokeh, Panel etc for scripting the analysis. The central orchestration tool for our workflow is Dagster which seamlessly facilitates the Extract, Transform, Load (ETL) process. The dagster guarantees the orchestrated flow of

data throughout the various stages. MongoDB a robust NoSQL database is used to store the raw data that are extracted from API calls and csv file read. We used MongoDB over other because of its document-oriented approach which simplifies data representation thereby reducing development complexity. PostgreSQL which is reliable relational database known for its superior querying capabilities and structured data management is to store the data after wrangling and merging processes which is used for further analysis. PostgreSQL is our first preference as it ensures robust ACID compliance, data integrity and reliability even in complex transactional scenarios. We integrated mongo-express and PG admin within the environment for visually interacting with databases. We wrapped our whole setup in Docker containers to ensure portability and uniformity across several computer settings. By merging these tools, we were able to establish a robust database management and analytical programming ecosystem capable of providing effective data driven insights.

## IV. RESULT ANALYSIS

## V. RECOMMENDATIONS AND CONCLUSIONS

### REFERENCES

- [1] A. R. Rakhmatulloh, D. I. K. Dewi, and D. M. K. Nugraheni, "The Built Environment and its Impact on Transit based Transportation Users Walking Activity in Semarang, Indonesia," *pertanika journal of science and technology*, 2021, doi: 10.47836/pjst.29.2.05.
- [2] A. Onifade, A. O. Onifade, B. Folasayo, B. Folasayo, A. Babatunde, and A. Babatunde, "ENVIRONMENTAL POLLUTION: EFFECT ON PEDESTRIANS WHILE WALKING IN URBAN STREET," *International Journal of Environmental Science*, 2022, doi: 10.47604/ijes.1436.