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 datbase 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 veru 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 analytical 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

In our project we use a method called Knowledge Discovery in Databases (KDD) first described by Fayyad et al. in 1996 [3]. This approach outlines a clear set of steps for turning raw data into useful information. Our study follows this method closely by using a well organized process that starts with collecting data and ends with extracting valuable insights. The updated figure Figure 1 in our paper shows these steps thus ensuring our research is useful for improving and understanding environment's effect on pedestrian footfall.

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 therby 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:

Once we collect the data we store it in MongoDB, which is a type of NoSQL database known for its ability to efficiently manage large amounts of unstructured or semi-structured data. We then use Python scripts to prepare the data for analysis. This preparation involves cleaning the data to remove any irrelevant information by dropping columns where the precentage of null values or zero values in more than eighty, for others null values are replaced by imputing 0 to the respective columns. These null values or zero values maybe due to the improper working of the sensors. This step is crucial because it makes sure that the data is accurate and organized therby setting the stage for reliable analysis and insights.

C. Storage and Visualizations:

After preprocessing we transfer the cleaned and structured data into a PostgreSQL database which is an SQL database known for its strong capabilities in data warehousing and handling complex queries. This is a critical step towards deeper data analysis allowing us to perform analysis and create detailed visualizations. This transition is not just a technical step it's about moving closer to our goal of understanding and visualizing complex patterns within the data. By using PostgreSQL we can dig deeper into the data and uncover insights that can inform decisions. This enables us to turn raw data into meaningful visual stories that can illustrate trends, challenges, and opportunities in urban environments.

Parameters	**Description (Unit)**
id	Unique identifier for each data entry
date	Date and time of the recorded data.
temperature_2m(°C)	Temperature at 2 meters above ground level in Celsius.
relative_humidity_2m(%)	Relative humidity at 2 meters above ground level, expressed as a percentage
dew_point_2m(°C)	Dew point temperature at 2 meters above ground level in Celsius.
apparent_temperature(°C)	Perceived temperature, taking into account factors like humidity and wind, in Celsius
precipitation(mm)	Total precipitation in millimeters
rain(mm)	Amount of rainfall in millimeters
snowfall(cm)	Amount of snowfall in centimeter
cloud_cover(%)	Percentage of sky covered by clouds
wind_speed_10m(km/h)	Wind speed at 10 meters above ground level in kilometers per hour
sunshine_duration(Seconds):	Duration of sunshine in seconds
pm10($\mu g/m^3$)	Particulate Matter (PM10) concentration in micrograms per cubic meter.
pm2_5($\mu g/m^3$)	Particulate Matter (PM2.5) concentration in micrograms per cubic meter
carbon_monoxide($\mu g/m^3$)	Carbon Monoxide (CO) concentration in micrograms per cubic meter
nitrogen_dioxide($\mu g/m^3$)	Nitrogen Dioxide (NO2) concentration in micrograms per cubic meter
sulphur_dioxide($\mu g/m^3$)	Sulphur Dioxide (SO2) concentration in micrograms per cubic meter
dust($\mu g/m^3$)	Dust concentration in micrograms per cubic meter.
european_aqi	European Air Quality Index (AQI) calculated based on various pollutant concentrations.
european_aqi_pm2_5	European AQI specifically calculated for PM2.5.
european_aqi_pm10	European AQI specifically calculated for PM10
european_aqi_nitrogen_dioxide	European AQI specifically calculated for nitrogen dioxide (NO2)
european_aqi_ozone	European AQI specifically calculated for ozone (O3).
european_aqi_sulphur_dioxide	European AQI specifically calculated for sulphur dioxide (SO2)

Num	**Locations**
1	Aston Quay/Fitzgeralds
2	Baggot st lower/Wilton tce inbound
3	Baggot st upper/Mespil rd/Bank
4	Capel st/Mary street
5	College Green/Bank Of Ireland
6	College st/Westmoreland st
7	D'olier st/Burgh Quay
8	Dame Street/Londis
9	Grafton st/Monsoon
10	Grafton Street / Nassau Street / Suffolk Street
11	Grafton Street/CompuB
12	Grand Canal st upp/Clanwilliam place
13	Grand Canal st upp/Clanwilliam place/Google
14	Mary st/Jervis st
15	Phibsborough Rd/Enniskerry Road
16	North Wall Quay/Samuel Beckett bridge West
17	O'Connell st/Princes st North
18	North Wall Quay/Samuel Beckett bridge East
19	Richmond st south/Portabello Harbour inbound
20	Richmond st south/Portabello Harbour outbound

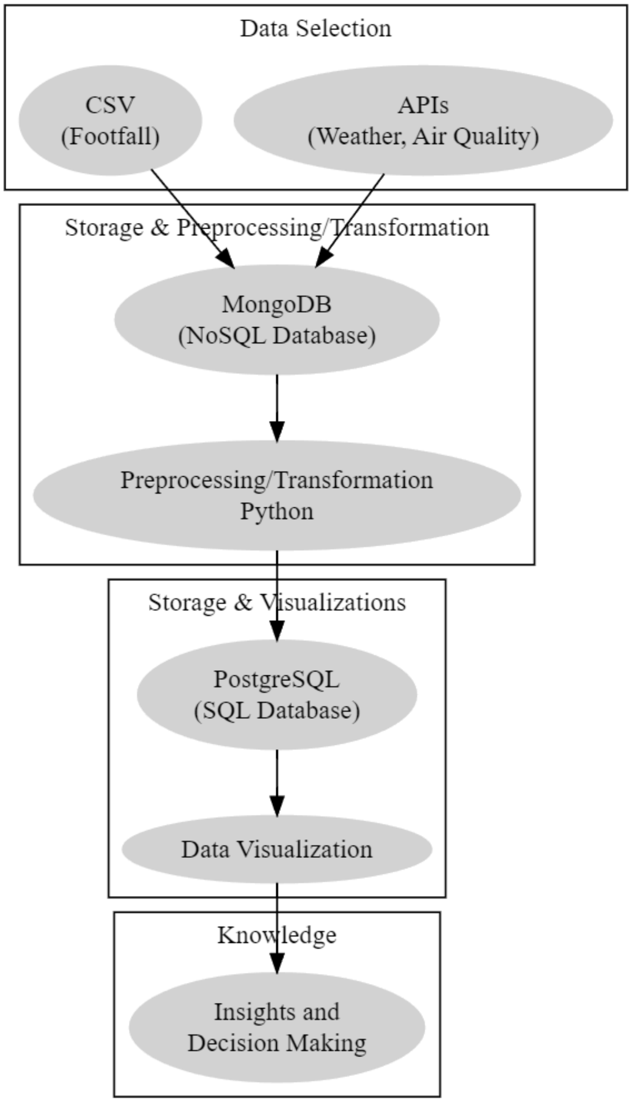

Figure 1: KDD Lifecycle

D. *Knowledge:*

The final goal of our process is to extract knowledge that can be used in decision making. We get this by converting complex data sets into clear, easyto understand visualizations. These graphs not only make it easier to spot patterns and insights therby making informed decisions. These visualizations are explained in the next section "Data Visualization"

The way we use the Knowledge Discovery in Databases (KDD) lifecycle is iterative which means we continuously refine our methods of analysis. Each step of this process is carefully detailed in Figure 1. This approach is about more than just processing data infact every step in our process builds upon the previous one.

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 incooperated 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. Data Visualization

After loading the data successfully into PostgreSQL we have a created a dashboard using panel library in python for visualizing and bokeh for plotting graphs. We have also incooperated a jupyter notebook for our analysis and vizualisation. This visualization strategy simplifies our analysis and therby making it more interactive by allowing users to engage directly with the data through dynamic visual tools. This interactive method of data visualization clearly shows how urban environments affect people's movement.

A. *Dashboard Features:*

The introduction page shows the environmental variables and final counter locations. Dataset page which shows the combined sample dataset of weather, air quality and footfall. Relationship between variable shows a scatter plot with all the environmental varaibles and counter locations. The distribution of variables pages shows a line graph with all the environmental varaibles and counter locations. The project report page has our project report displayed.

B. *Pedestrian Traffic Distribution Using Bar Chart:*

From the footfall dataset analysis we created a bar chart to visualise the average footfall at various locations across Dublin city thus revealing key pedestrian hotspots. This bar chart titled "Average Footfall by Location" Figure 2 clearly shows which areas experience the highest pedestrian traffic with locations like “Aston Quay/Fitzgeralds” and “Baggot st lower/Wilton tce inbound” leading in footfall. These insights can be crucial for urban planners and government authorities as they can design strategies to enhance urban mobility and optimize city spaces for better pedestrian flow and can also is a key metric to optimise your business performance in that locations. The visualization help to identify and address areas of high pedestrian activity thus helping in decision making processes related to urban development and infrastructure planning.

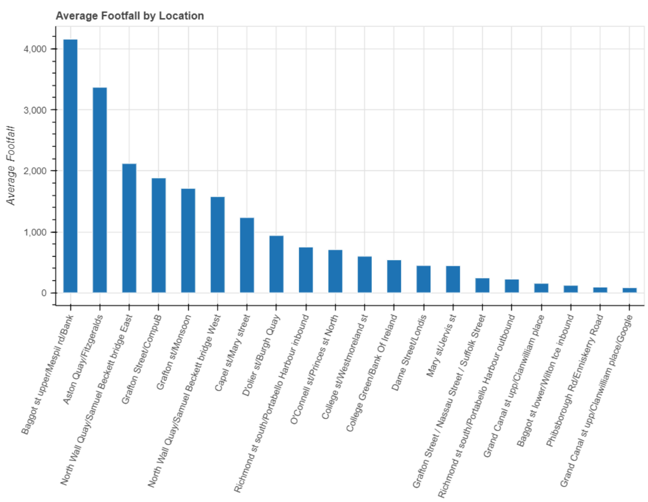

Figure 2: Average Footfall by Location

C. *Environmental Impact On Footfall Using Scatter Plot:*

Temperature vs Pedestrian Traffic at College Green/Bank Of Ireland This scatter plot shows a relatively dense clustering of points in the mid-range of temperatures, suggesting a potential correlation where footfall increases with moderate temperatures. The distribution is wide, indicating variability in footfall which could be attributed to other factors like time of day or specific events. Temperature vs Pedestrian Traffic at Baggot St Upper/Mespil Rd/Bank The plot presents a very dense cloud in the mid to higher temperature ranges, which could suggest higher pedestrian activity in warmer conditions. This plot can serve as an excellent example to discuss the influence of pleasant weather on urban mobility patterns. Inference for Selected Images: These plots are excellent representations of how temperature variations influence pedestrian movements in urban areas. The plot from College Green/Bank of Ireland, for example, could indicate that more people are likely to walk in moderate temperatures, which aligns with comfortable walking conditions. On the other hand, the plot from Baggot St Upper/Mespil Rd/Bank shows higher footfall during warmer temperatures, possibly indicating a preference for outdoor activities or commuting on foot during warmer days.

In addition to the apparent influence of weather on pedestrian traffic as shown in the scatter plot for Baggot Street Upper and Mespil Road near the bank, the area’s unique blend of commercial activities, historical significance, and aesthetic appeal contributes to its footfall patterns. This location, known for its historic Georgian architecture and proximity to the serene Grand Canal, offers a vibrant urban space that attracts both locals and visitors, especially in favorable weather conditions. Such environments underscore the importance of urban planning that considers both the functional and aesthetic components to enhance pedestrian friendliness and overall urban livability.

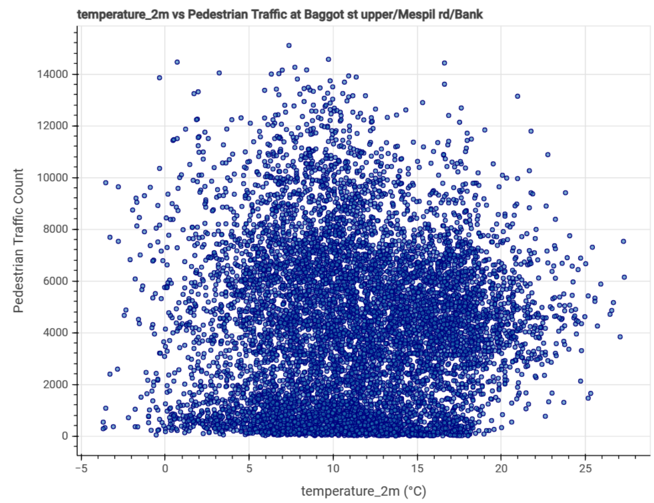

Figure 3: baggot

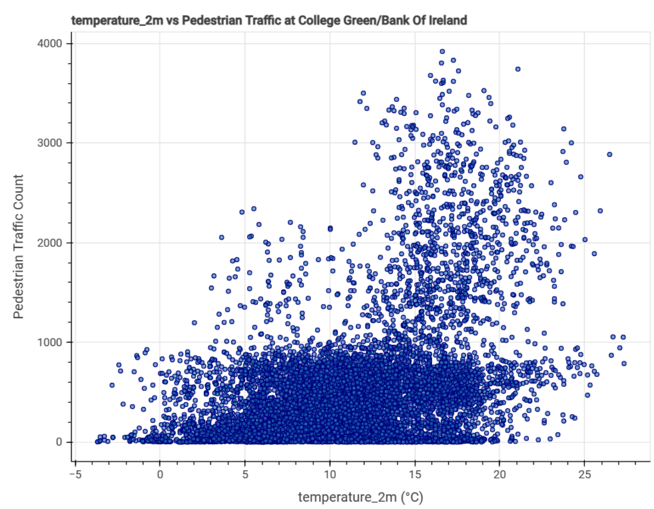

Figure 4: college green

The scatter plot for “European AQI PM2.5 vs Pedestrian Traffic at O’Connell St/Princes St North” shows a visible pattern that there is a dense clustering of data points at lower AQI levels and then it begins to spread out as AQI values increase. This suggests that more people like to walk in the area when the air quality is good and there are fewer people on road as pollution worsens. This pattern shows the how air quality impact peoples decision to participate in outdoor activities.

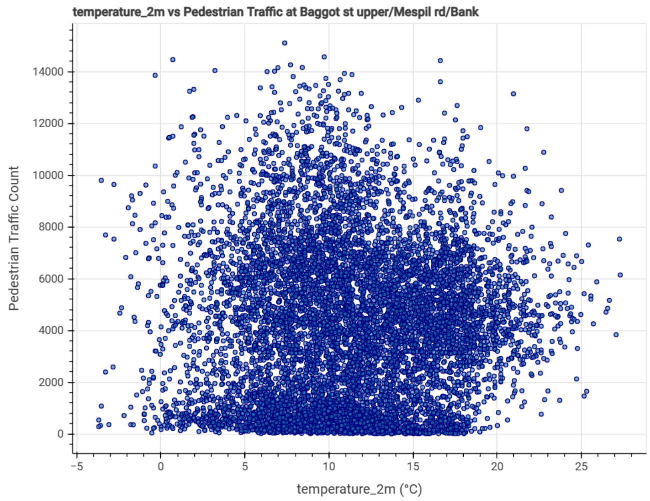

Figure 5: baggot

Similarly, the plot for “European AQI PM2.5 vs Pedestrian Traffic at College Green/Bank of Ireland” presents a high density of data points at moderate levels of air pollution, which decreases as the AQI increases beyond a certain point. This indicates that pedestrian traffic remains stable up to a certain level of air pollution but drops off as the air quality continues to decline. These trends are important for city planning and public health discussions, as they highlight the need for good air quality to maintain active and lively urban centers

*D. **Line Graph:***

V. Conclusions:

Customizable Data Views: With features allowing users to select different parameters and locations, the dashboard facilitates tailored analysis specific to particular areas or conditions. This capability is crucial for urban planners and policy-makers who need to make informed decisions based on specific data points. By converting raw data into easily interpretable visual formats, the dashboard aids in quicker and more informed decision-making processes. It provides stakeholders with a clear view of how environmental conditions affect urban mobility, which can be crucial for city planning and improving public spaces. Educational Tool: The dashboard serves as an educational tool for the community, raising awareness about how environmental quality can affect daily life in urban areas. It can also encourage public engagement with environmental issues, promoting a more informed citizenry.

VI. 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.

[3] U. M. Fayyad, U. M. Fayyad, G. Piatetsky-Shapiro, G. Piatetsky-Shapiro, P. Smyth, and P. Smyth, “From Data Mining to Knowledge Discovery in Databases,” *Ai Magazine*, 1996, doi: 10.1609/aimag.v17i3.1230.