**CCT College Dublin**

**Assessment Cover Page**

*To be provided separately as a word doc for students to include with every submission*

|  |  |
| --- | --- |
| **Module Title:** | *Programming for DA*  *Statistics for Data Analytics*  *Machine Learning for Data Analysis*  *Data Preparation & Visualisation* |
| **Assessment Title:** | *MSC\_DA\_CA1* |
| **Lecturer Name:** | *Sam Weiss*  *Taufique Ahmed/* *Marina Iantorno*  *Muhammad Iqbal*  *David McQuaid* |
| **Student Full Name:** | Ben Wilson |
| **Student Number:** | sbs24004 |
| **Assessment Due Date:** | 7/4/2024 |
| **Date of Submission:** | 7/4/2024 |

**Declaration**

|  |
| --- |
| By submitting this assessment, I confirm that I have read the CCT policy on Academic Misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source. I declare it to be my own work and that all material from third parties has been appropriately referenced. I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution. |

# ****Statistical and Clustering Analysis of St. Patrick's Day Footfall Counts in Dublin City Centre****

### Prepared by Ben Wilson, sbs24004, CCT

## Introduction[¶](http://localhost:8890/lab/tree/ca1-sbs24004/TourismFootfallStPatricksDay.ipynb#Introduction)

This project aims to uncover patterns and trends in pedestrian footfall within Dublin City Centre, focusing particularly on the dynamics surrounding St. Patrick's Day, and identifying similar behaviour of other events throughout the year. As Ireland's national holiday, St. Patrick's Day attracts thousands of both local and international tourists to the capital's streets. The project will mainly utilise footfall count data from 2023 to highlight trends, peak activity, and heavily used areas. By comparing everyday footfall with that on St. Patrick's Day, to reveal changes in footfall dynamics and tourist visitor volume, including a year-on-year comparison. Through the use of unsupervised k-means clustering, identify signature footfall patterns associated with large-scale events, and group them with other periods of similar activity. This should serve to provide future insights valuable for large-scale event organisation, tourism boards, festival coordinators, and local businesses.

"Fáilte Ireland is pleased to welcome thousands of visitors to locations across the country this weekend for St. Patrick’s Day celebrations. Festivals and events play a key role in delivering brilliant experiences, providing a unique reason for visitors to choose a destination and increase footfall for local businesses, supporting jobs and revenue generation. Fáilte Ireland estimates that last year 570,000 people attended the Saint Patrick’s Festival Dublin over the bank holiday weekend generating €113million in revenue for Dublin."  
  
Source: [Fáilte Ireland 2024](https://www.failteireland.ie/Utility/News-Library/failte-ireland-welcomes-visitors-st-patricks-day.aspx)

## Objectives:

* To perform statistical analysis of footfall count patterns in Dublin City Centre
* To quantify the effect of St. Patrick's Day Festival on pedestrian footfall and compare counts with previous years
* To create a k-means clustering machine learning model for identifying similar periods of activity and their characteristics

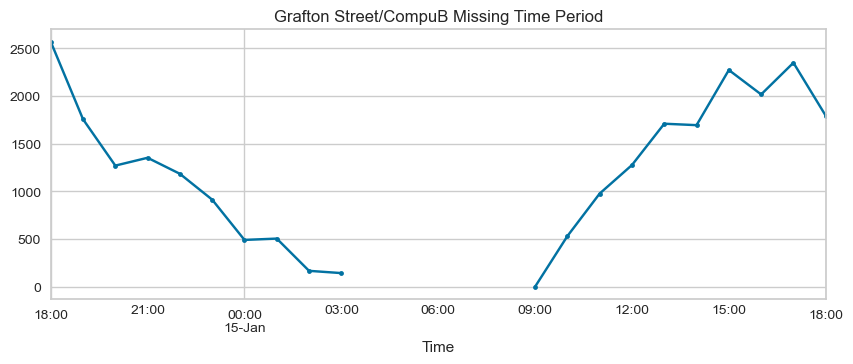
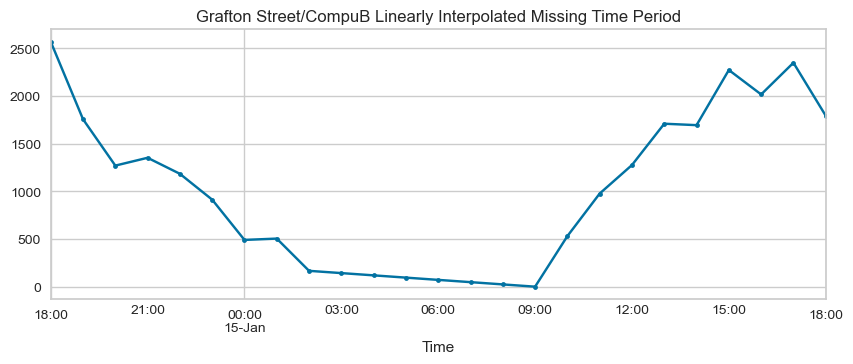
## Data Source

The data used in this project was downloaded from Smart Dublin; a data set entitled [Pedestrian Footfall DCC](https://data.smartdublin.ie/dataset/dublin-city-centre-footfall-counters). According to the "Dublinked Open Data Store", the counts are recorded using "a network of PYRO-Box people counters located throughout central Dublin". The data is supplied by Dublin City Council and the NTA (National Transport Authority).

## EDA

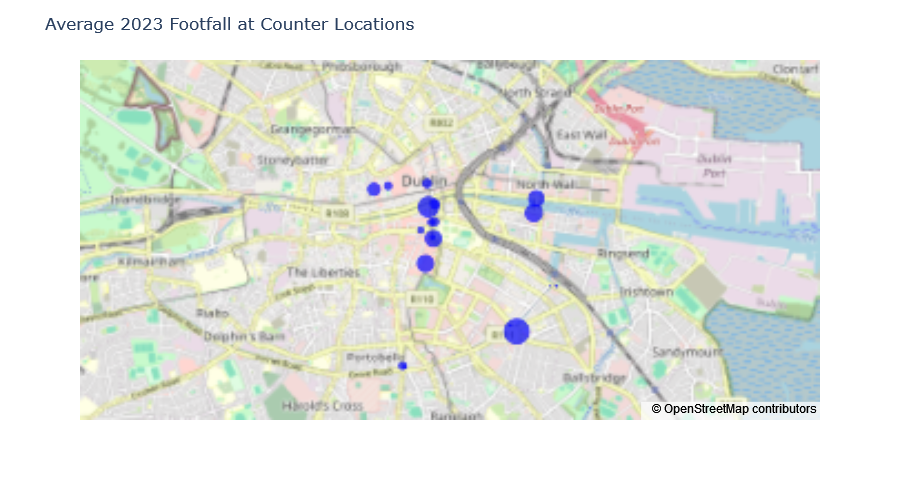
## Data Cleaning and Preparation

### Short Period NaNs

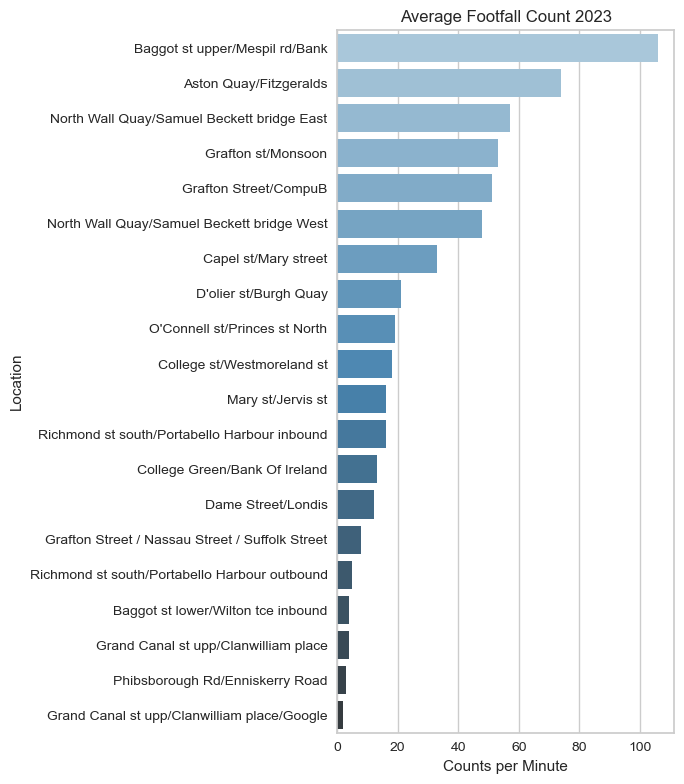
 

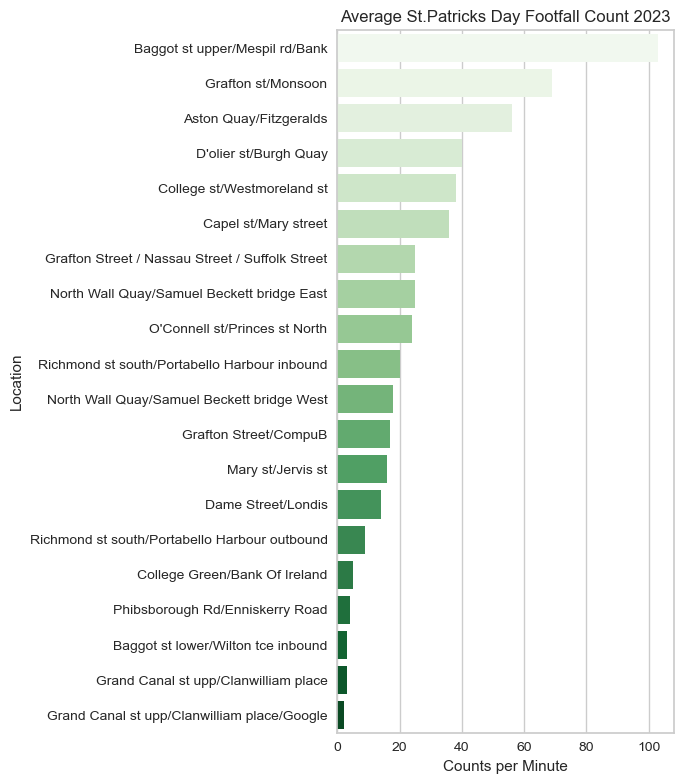
### Long Period NaNs

### 









Histograms

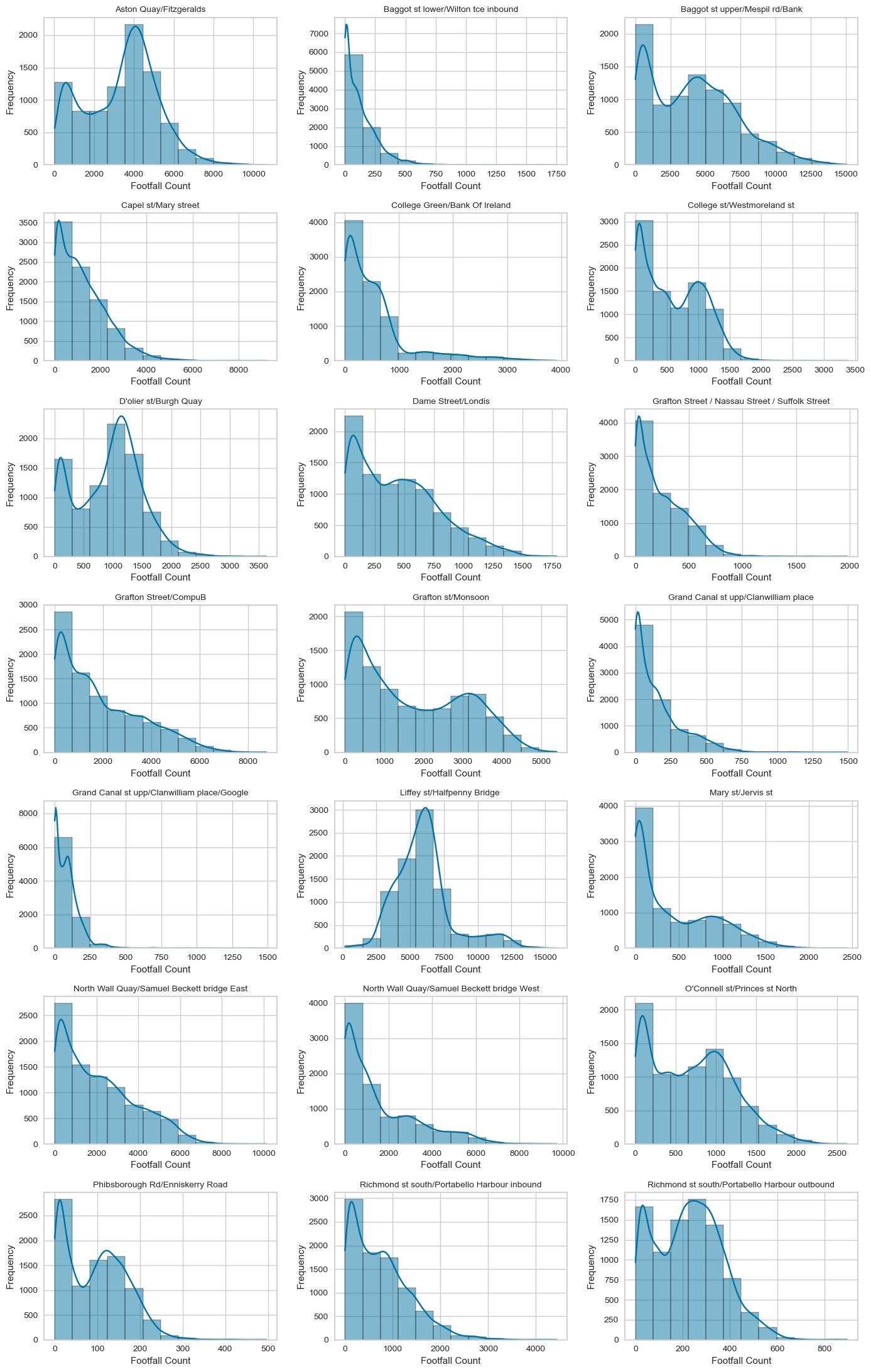


Figure 1 Histograms of Footfall Counts recorded over the year 2023



Figure 2 Histograms of Footfall Counts recorded on St. Patrick's Day 2023

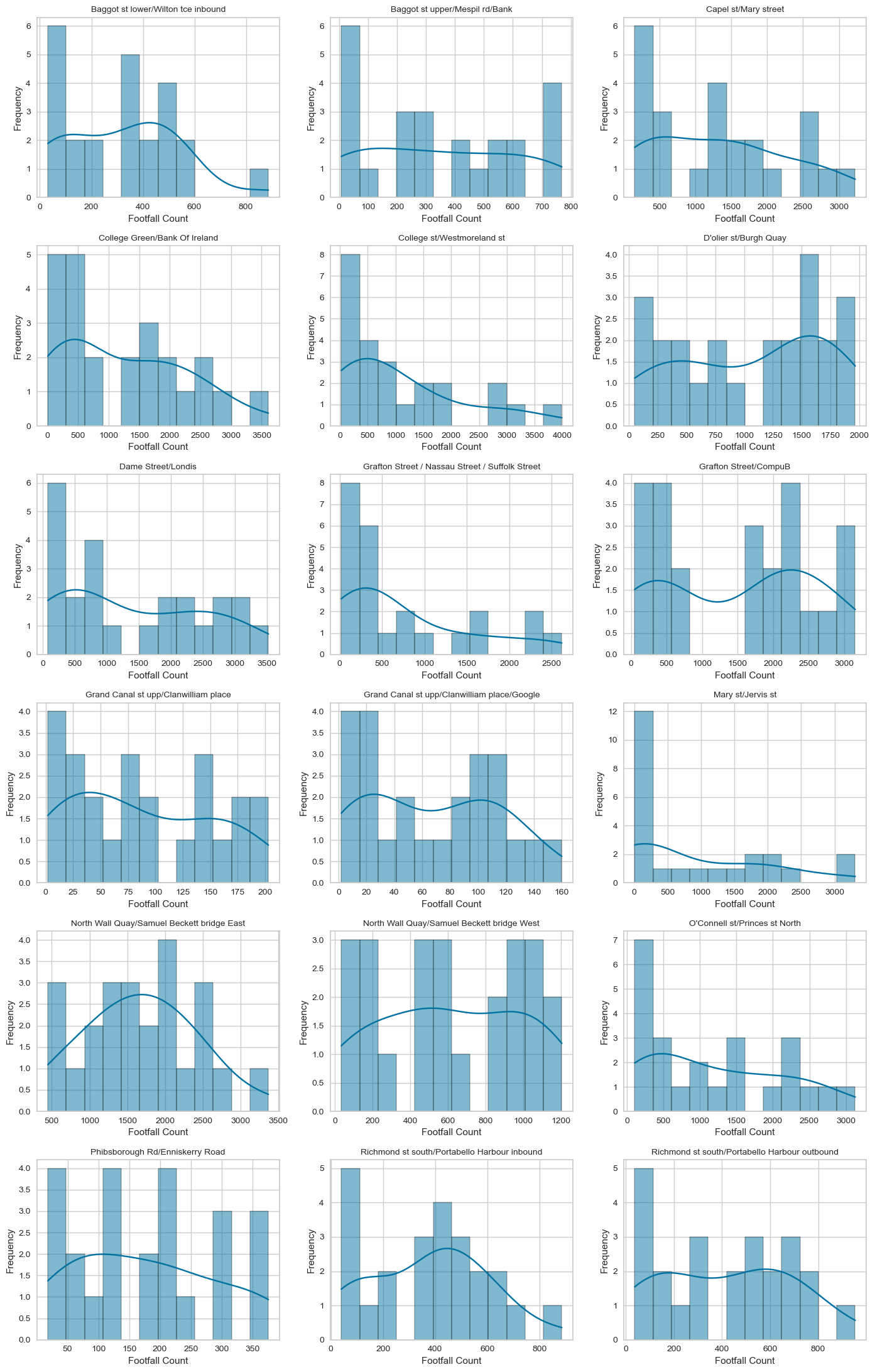
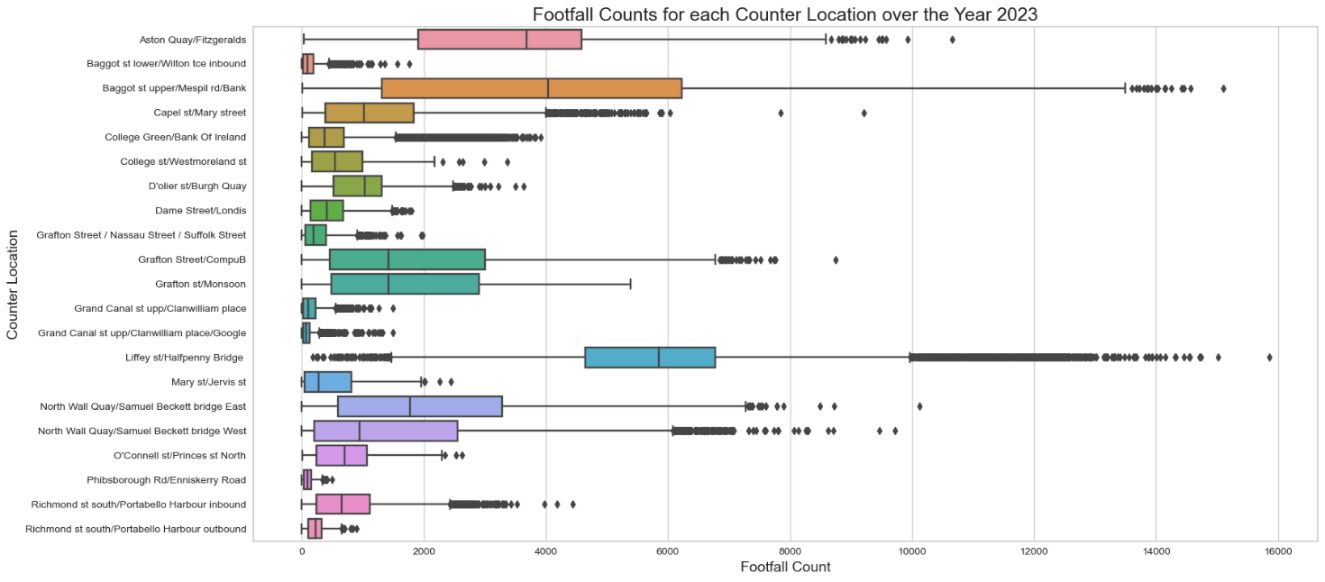


Figure 3Histograms of Footfall Counts recorded on St. Patrick's Day 2022



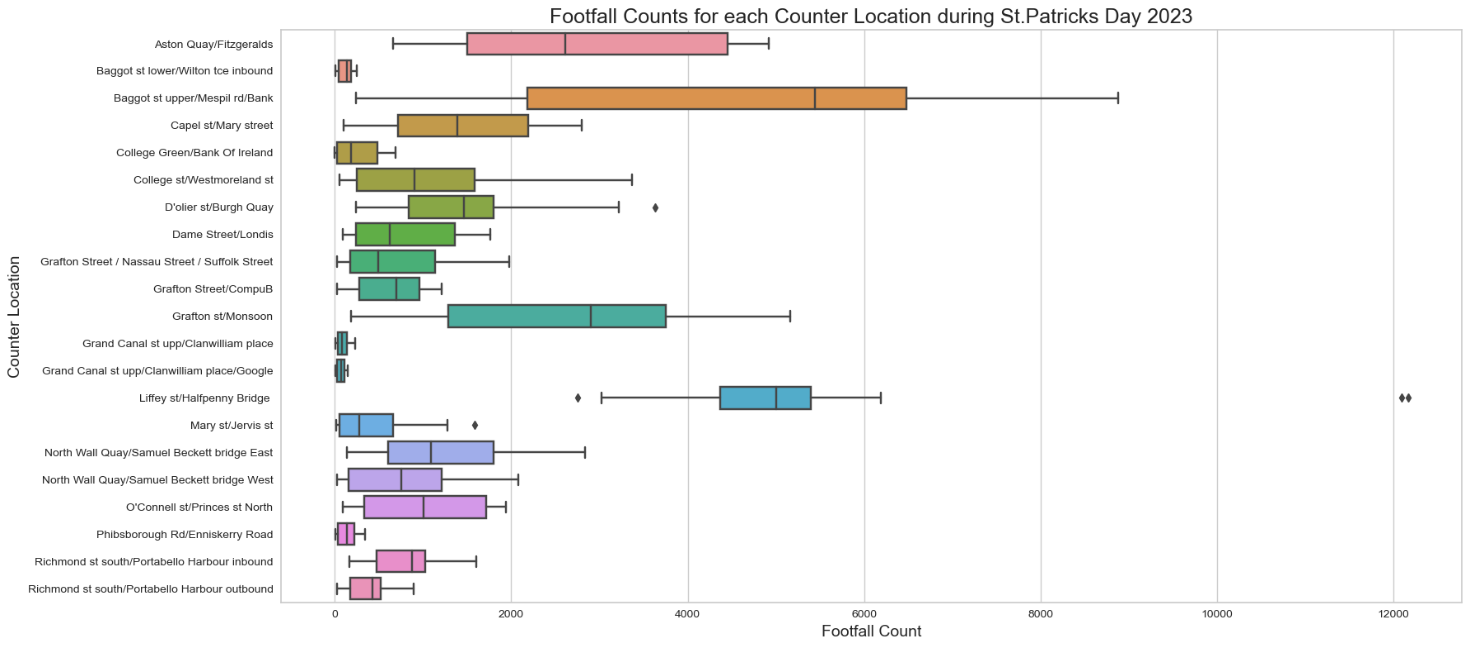
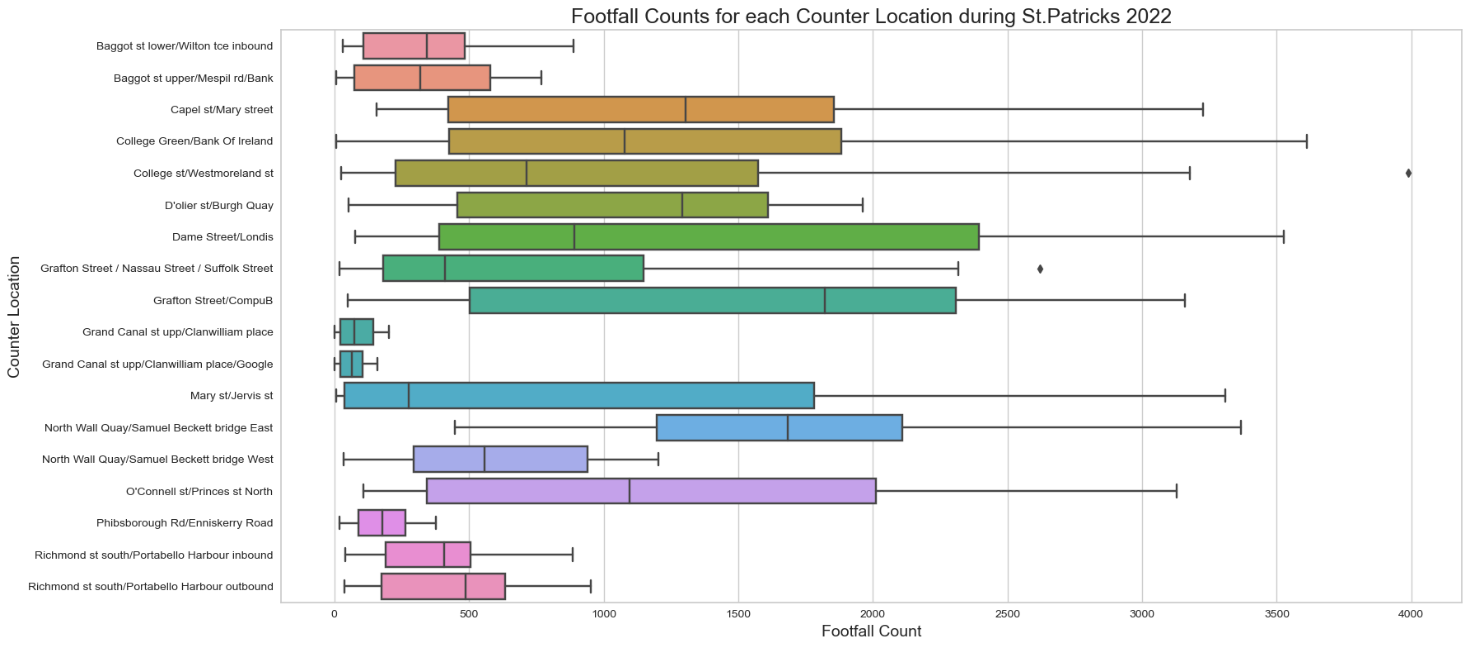


Figure 4 Boxplots 2022

Figure 5Boxplots of St.Patricks Day

Figure 6Boxplot of Counts 2023



