

Final-Project-Statistical-Modelling-with-Python

CityBike Lisbon: Location Data Modelling with Python

This project explores relationships between Bike Availability at Lisbon's CityBike stations and characteristics of surrounding locations & organisations. Characteristic such as; category type, average rating & average popularity.

Primary Goal

Identify predictor variables that influence Free Bikes and/or Total Bike Slots at Lisbon CityBike stations.

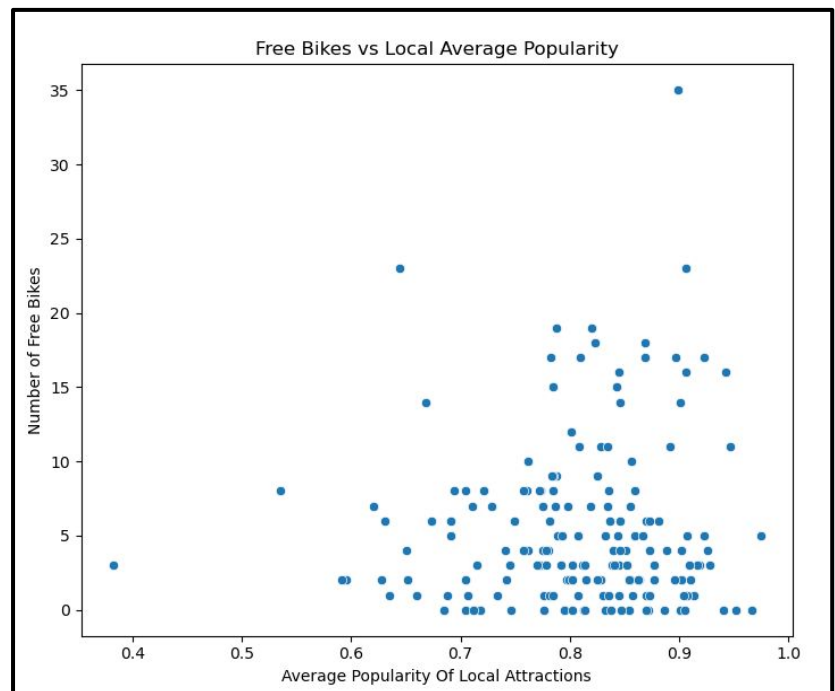
Data Timestamp:

12:39 AM Sunday, July 27 2025

Details

Key Insights

- **Independent Variables:**
 - Category Type
 - Restaurant
 - Bar
 - Cafe
 - Coffee
 - Rating
 - Popularity
- **Dependent Variables:**
 - Free Bikes
 - Total Bike Slots
- Variation in Location/Organisation density per BikeStation
 - Limited distance from BikeStation 298m address biases
- Location Category Type
 - **Poor** predictor of dependent variables
- Location Ratings & Popularity
 - **Poor** predictor of dependent variables
- No strong predictive relationships or correlations were observed
- However, **'Bar'** and **'Cafe'** offered the strongest available predictive value for **'Free Bikes'** accounting for **approximately 17.9% of the variance**.



Average Location Popularity For Each BikeStation against Free Bikes Available

Next Steps

Additional Independent Variables

- `Time Of Day`
- `Season`
- Observe trends across daily, weekly & monthly timestamps

Further cleaning and grouping of broader categories

- `Public / Government`
- `Retail`
- `Hospitality`