Bike-Share Usage in London Network Analysis

Authors

- **Mulin Wan**
  · 🟢 [mulin-wan](#)
  CEE, University of Illinois Urbana-Champaign

- **Jingwen Yao**
  · 🟢 [jingwenyaooo](#) · 🐦 [Yaojune](#)
  CEE, University of Illinois Urbana-Champaign

- **Yunze Guo**
  · 🟢 [cyfcx2](#)
  CEE, University of Illinois Urbana-Champaign

- **Bo-Yang Wang**
  · 🟢 [byw-5](#)
  CEE, University of Illinois Urbana-Champaign

Abstract

Introduction

Description

Dataset:

For this project, we are going to focus on the usage of bike sharing in London. The data come from two datasets, "London and Taipei Bike-Share Data" and "London bike sharing dataset."

London and Taipei Bike-Share Data

The raw data is collected from the respective cities open data sites.

[London](#)

The data has been reformatted into CSV in order to be easier to use and compare. The content is unchanged.

This dataset would be the main dataset for the project as it contains every single bike rental in London in the duration. This gives the columns in the data comprise:

Object	Description
rental_id	id of people who rent the bike

Object	Description
duration	duration of rental
bike_id	id of bike
end_rental_date_time	date and time of end rental
end_station_id	id of end station
end_station_name	name of end station
start_rental_date_time	date and time of start rental
start_station_id	id of end station
start_station_name	name of start station
start_rental_date_time	date and time of start rental

London bike sharing dataset

This dataset is playing a more supporting role, as it helped providing informations on weather conditions. Although the timespan doesn't overlap too much with the first dataset (January 4th 2015 to January 3rd 2017), it gives the idea to incorporate weather conditions into consideration. In the future we may try to find weather datasets that fit the first dataset better to help analyzing.

The data has been formatted into CSV in order to be easier to use and compare. This gives The columns in the data comprise:

Object	Description
mestamp	timestamp field for grouping the data
cnt	the count of a new bike shares
t1	real temperature in Celsius
t2	temperature in Celsius "feels like"
hum	humidity in percentage
windspeed	wind speed in km/h
weathercode	different weather condition
Clear	mostly clear but have some values with *haze*, *fog*, *patches of fog*, *fog in vicinity*
scattered clouds	few clouds
Broken clouds	clouds were broken
Cloudy	cloudy weather
Rain/ light Rain shower/ Light rain	rainy weather
rain with thunderstorm	rain with thunderstorm
snowfall	snowy weather
isholiday	boolean field - 1 holiday / 0 non holiday - refers to bank holidays
isweekend	boolean field - 1 if the day is weekend / 0 if a working day
season	category (0-spring; 1-summer; 2-autumn; 3-winter)

Description of weathercode:

weathercode	Description
1	clear; mostly clear but have some values with haze/fog
2	scattered clouds / few clouds
3	broken clouds
4	clear; cloudy
7	clear; light rain shower / rain / light rain
10	clear; rain with thunderstorm
26	snow fall
90	freezing fog

In addition to season and isweekend, from the timestamp feature we can extract many separate time features - day of the week (as one scaled column or as seven columns of ismonday, istuesday etc.), month number, day of the month, week number, hour, minute. In combination with external data, we could add islight for after dawn times and is_schoolholiday to match London school holiday times.

Link of dataset:

[London and Taipei bikeshare](#)

[London bikeshare](#)

Proposal "SPECIFIC"

Recently, bike-sharing in big cities has become an important part of residents' daily life, and its role in urban transportation system has never been more significant. Looking at the spatiotemporal bike-sharing data in London, we could explore patterns, describe variations, or modeled the data in many different ways. From the two datasets, we may have a chance to take a peek at the residents' bike-renting behavior through many angles.

"PREVIOUS" work has shown that weather is a key driver for variation in usage. [1] [2] By utilizing those datasets to analyze how extreme weather events like winter rains in London will affect bike-sharing system, it is safe to draw some conclusions to guide the process of making contingency plans. The locations of start-trip and end-trip is also considered to have the potential of revealing hot spots of bike-renting usage.

The result will be able to offer some suggestions for the decision maker of bike-sharing companies about the arrangement of bike density in different blocks, distribution between urban and rural areas and methods to tackle extreme weather conditions.

References

[1] AndersOhrn (2020) Bike-share usage in London and Taipei Network, Kaggle. Available at: https://www.kaggle.com/datasets/ajohrn/bikeshare-usage-in-london-and-taipei-network (Accessed: October 24, 2022).

[2] Mavrodiev, H. (2019) London Bike Sharing Dataset, Kaggle. Available at: https://www.kaggle.com/datasets/hmavrodiev/london-bike-sharing-dataset/discussion?resource=download (Accessed: October 24, 2022).