JiaotongUniv

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

Title of the project

Bike-Share Usage in London Network Analysis

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:

1.rental_id - id of people who rent the bike

2.duration - duration of rental

3.bike_id - id of bike

4.end_rental_date_time - date and time of end rental

5.end_station_id - id of end station

6.end_station_name - name of end station

7.start_rental_date_time - date and time of start rental

8.start_station_id - id of start station

9.start_station_name - name of start station

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:

1.mestamp - timestamp field for grouping the data

2.cnt- the count of a new bike shares

3.t1 real temperature in C

4.t2 - temperature in C "feels like"

5.hum - humidity in percentage

6.windspeed - wind speed in km/h

7.weathercode - category of the weather

1 = Clear ; mostly clear but have some values with haze/fog/patches of fog/ fog in vicinity

2 = scattered clouds / few clouds

3 = Broken clouds

4 = Cloudy

7 = Rain/ light Rain shower/ Light rain

10 = rain with thunderstorm

26 = snowfall

8.isholiday - boolean field - 1 holiday / 0 non holiday - refers to bank holidays

9.isweekend - boolean field - 1 if the day is weekend / 0 if a working day

10.season - category (0-spring; 1-summer; 2-autumn; 3-winter)

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

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. 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