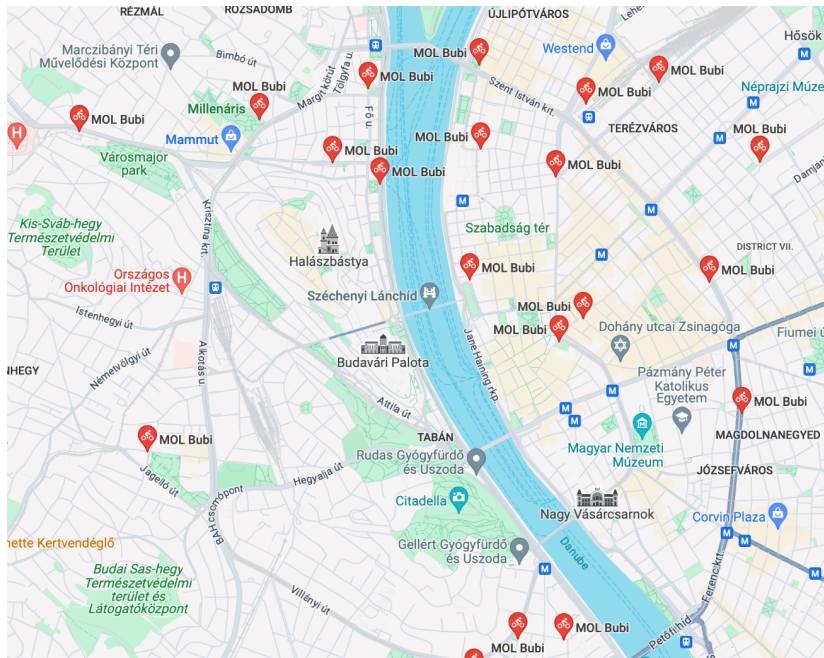# Exploring the MOL Bubi dataset

Tamás Barczikay

November 24, 2024

MOL Bubi Public Bike Stations

Objective 1: Understanding the effects of weather

▶ General effects of weather / regression analysis

▶ Time of the day / von Mises diagram

▶ Route analysis / frequency table

Objective 2: Recommendation for ad placement

▶ Extending **Objective 1's** scope.

▶ Effect of rack number / frequency table - scatterplot

Data Cleaning

- ▶ In general there were no missing entries in rides and stations
- ▶ Dates and times were character encoded (changed to POSIXct format)
- ▶ Rides over 10 hrs were eliminated considered to be outliers (10hrs can be a work shift, only 0.1% of data lost)
- ▶ Duplicated stations (reallocated 1 min away ✓, rack increase days/few weeks before end ✓, rack increase before that ✗)
- ▶ For the last reason Deák tér, Erzsébet tér, Baross utca and Keleti pályaudvar were handled separately in some applications
- ▶ *Wind direction* is a factor that cannot be interpreted easily due to station surroundings (drop).
- ▶ 57.6% of *visibility in kms* are mainly missing (-9999, drop)
- ▶ Same for *wind chill* (50% are -999, drop)
- ▶ *Hail* had no variation (all values were 0, drop)

General effects of weather - latent effects
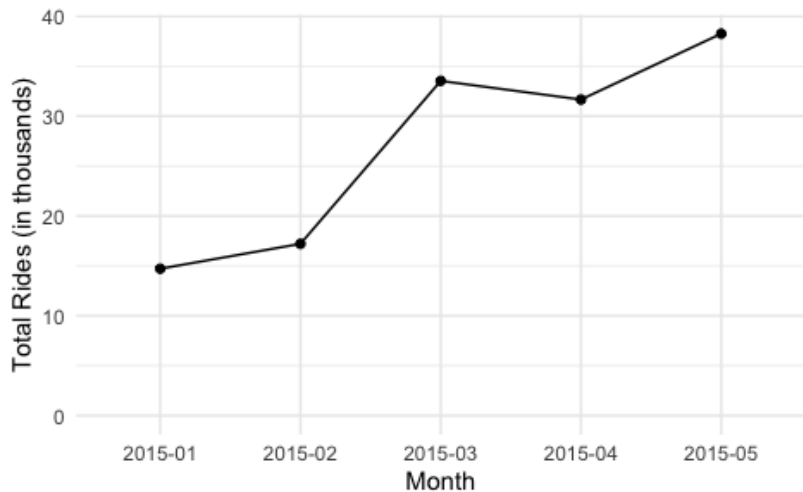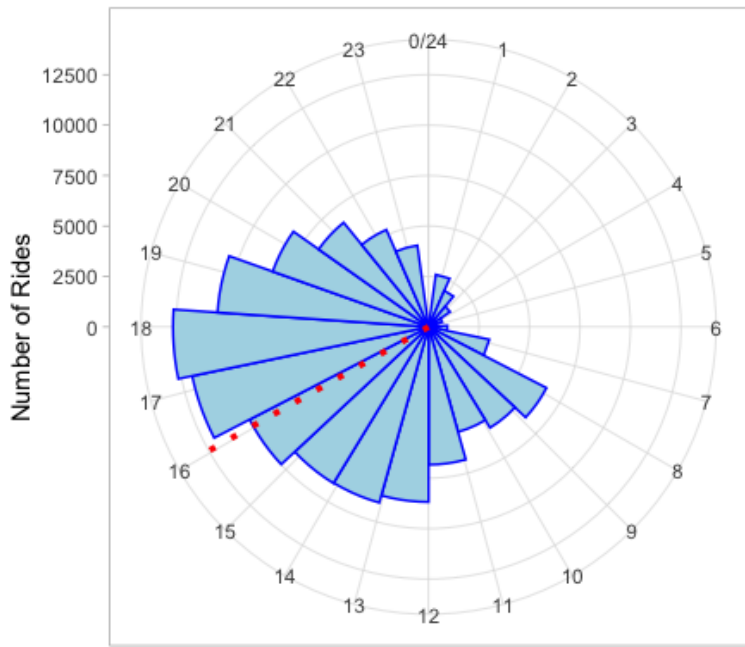
General effects of weather - regression results
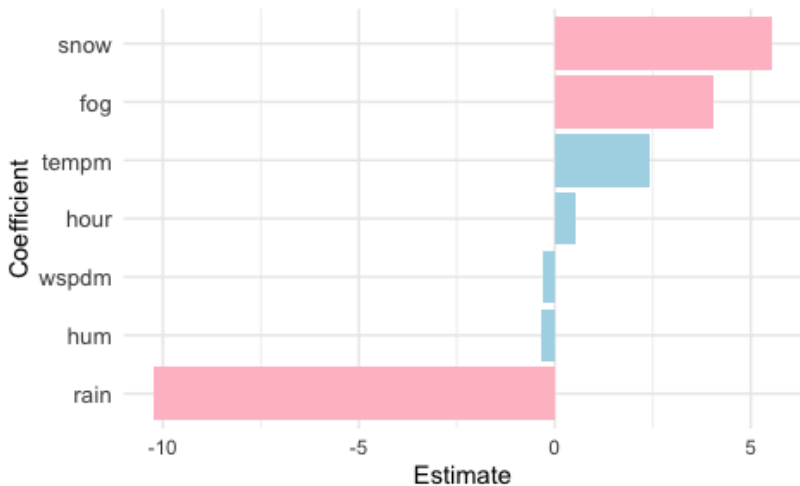
Note: Variables indicated with pink are binary. Insignificant factors are not shown (pressure and thunder). The model explains around 64% of the variation in rides.

Time of the day

Route Analysis

| Route	Total Rides	% share
Margitsziget round trip	1224	0.9
Jászai Mari tér to **Margitsziget** and return	1021	0.75
Kodály körönd to **Oktogon** and return	906	0.67
Hősök tere to **Oktogon** and return	620	0.46
Jászai Mari tér to Szent István park and return	590	0.44
Garay tér to Keleti pályaudvar and return	520	0.38
Margitsziget to Szent István park and return	498	0.37
Ferenciek tere to Kálvin tér and return	480	0.35
Dózsa György út - Dembinszky utca to Keleti	414	0.31
Városháza Park round trip	380	0.28

Table: The three most visited stations

Effect of rack number I.

Station	Start	Days	Rides	Avg.
Deák tér	01-01	130.0	1906	14.6
Deák tér	05-11	20.5	477	23.3
Erzsébet tér	01-01	127.0	2814	22.1
Erzsébet tér	05-08	23.5	837	35.7
József körút - Baross utca	01-01	131.0	1316	10.1
József körút - Baross utca	05-11	20.4	287	14.1
Keleti pályaudvar	01-01	77.9	660	8.47
Keleti pályaudvar	03-19	64.7	993	15.4
Keleti pályaudvar	05-23	8.38	88	10.5

Table: Summary of Rides by Stations with increasing racks

Effect of rack number II.
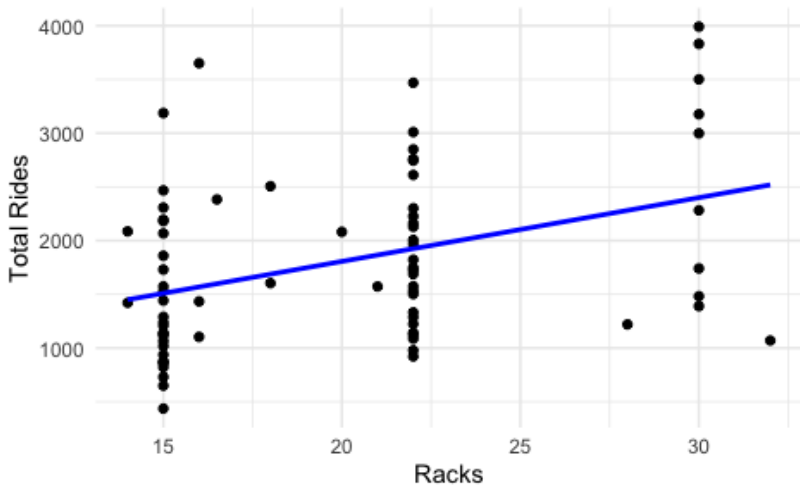
A new rack increases usage in the sample by 59.4 rides ($R^2 = 14\%$)

Recommendations

- Use advertisements in **hot months**, from the start of spring, but preferably during the summer (higher temperature adds more rides, but sample ends at the end of May)

- Snowing and fog is easily offset by hot temperature

- If rolling poster displays are available, they should show the advertisement with a higher frequency **during afternoons** (from 16.00 to 19.15)

- The ads **should be placed at 1)** Margitsziget, **2)** Oktogon and **3)** Jászai Mari tér, where the 3.6% percentage of the whole sample showed up

- It should be noted, that Kálvin tér for example had almost as many start and destination rides as Margitsziget - around 4000 - , but all routes have two endpoints and this location location went to minor ones.

- Place ads near places with **high number of racks** (Margitsziget and Oktogon 30, Jászai was just 15)