

# Bubi Data Analysis – Research Plan

Virág Besenyei, Dóra Boglárka Draskóczi, Zsófia Csenge Lévai

Supervisor: Zombor Berezvai

## 1 Introduction

In recent years, public transportation in Budapest has been continuously developing, and one of the most visible elements of this progress is the public bike-sharing system (BSS), which provides an important travel mode for sustainable urban mobility. Since the launch of the first 76 stations in 2014, the system has been gradually expanded. The introduction of new docking stations not only increases the convenience of existing users but also has the potential to attract new groups of users.

At the same time, it remains a question to what extent a new station generates genuinely new demand, as opposed to merely redistributing existing traffic among the already operating stations. This issue is particularly important for evaluating efficiency, as the ultimate goal of such developments is to increase the overall usage of the network rather than simply shifting trips from one point to another. Furthermore, it is also interesting to investigate how the geographical pattern of usage alters after the inclusion of a new station.

## 2 Problem Definition and Research Objectives

The aim of our research is to analyze the impact of introducing selected new stations in Budapest on the traffic data of the Bubi system. We seek to examine whether the new stations contributed to the overall growth of the network's usage or primarily redistributed traffic from existing stations. Through this, we hope to gain a better understanding of the factors that most strongly influence the success of BSS stations.

In addition, we will attempt to develop a simple predictive model that, based on historical data, can forecast the expected traffic generated by a newly established station.

## 3 Research Plan

1. Choose several stations for comparison based on the data we have received.
2. Analyze how the previous stations' traffic has changed after the new stations were added.
3. Create Python code for generating tables, diagrams and maps about our assumptions and results.

4. Review results and data considering other circumstances that could affect the outcome, such as the geographic parameters of the chosen part of the city or seasonal variation (e.g., cyclists are less likely to use the BSS during winter).

## Ideas for the data analysis

- After the introduction of a given station, observe an increase in the number of trips in the surrounding area (e.g., within a 2 km radius).
- Identify whether new users appeared at the new station, or if existing users simply migrated from other stations.
- Construct an OD (origin–destination) matrix (e.g., the element in the 5th row and 7th column represents the number of trips from station 5 to station 7 during a given period).
- Conduct time series analysis: how does the number of departures/arrivals change at the new station and nearby older stations over time (weekly/monthly breakdown)? This can be visualized with graphs, e.g., growth of departures at the new station vs. change in traffic at neighboring stations.
- Perform map visualization: examine which stations the trips to the new station originate from, and how these patterns change over time.
- Generate heatmaps: show how traffic density in the area increased after the station’s introduction.

## 4 Literature Review

- Zombor Berezvai, *Short- and long-term effects of COVID-19 on bicycle sharing usage*, <https://www.sciencedirect.com/science/article/pii/S2590198222001348>
- Péter Bucsky, *Does Uber affect bicycle-sharing usage? Evidence from a natural experiment in Budapest: A comment*, <https://www.sciencedirect.com/science/article/pii/S0965856420306315>
- Cyrille Médard de Chardon, Geoffrey Caruso, Isabelle Thomas, *Bicycle sharing system ‘success’ determinants*, <https://www.sciencedirect.com/science/article/pii/S0965856416307674>
- *Infrastruktúra-bővítés világversenyek idején*, Közgazdasági Szemle, LXVI. évf., 2019. January (pp. 4–21).

We also have two Excel sheets containing data about the opening dates of the stations and their geographic locations within the city.