

BIKE SHARE TREND ANALYSIS



Bhakti Raichura: 016020628

Khushee Thakker: 015271529

Shreya Srirama: 016029845

Varsha Srinivasan: 016001544

1. Abstract

It is estimated that there are more than 2 billion bikes used in the world and this number is constantly growing. It is also known that over 50 percent of the world's population knows how to use a bike. Owing to social and health benefits, people are now moving from car-sharing to the bike-sharing model. In such a world where there are several bike riders, our focus is to analyze the users' usage of bikes, their dependability on the weather conditions, and correlations with other nodes such as stations, trips, users' age groups. We also aim at analyzing the distribution of first mile/last mile research along with the weather factors, to detect the measures which help in the increase of bike-sharing. The data analysis will be done in MySQL and Python and data visualization will be done in Tableau. Through our trend analysis, we will be able to gain insights into bike ride patterns and deduce suggestions to increase profitability for a bike-sharing company.

2. Motivation for the project

In the world of increasing global temperature, bike riding not only provides health benefits but shifting trips from cars to bikes also helps reduce congestion, air pollution, and CO2 emissions. Owing to this, we aim to provide an insightful analysis of bike riding for an average user. The project will help us find patterns in bike-sharing data which can be utilized by companies like Lime, Lyft, etc. to boost their bike-share business. Knowledge of the busiest stations, high-demand areas and days of the week, impact of the weather forecast on bike trips, etc. will help these companies source more bikes to cater to the demand and thus increase sales.

3. Literature Survey

[1]. **The effect of weather and climate on bicycle commuting**

Weather effects (long and short terms) on bicycle commute for tertiary students in Australia and Melbourne. The dataset focuses on students who are young and healthy, who commute using bicycles to their universities. The dataset includes various facets like places the students commuted, distance covered, riding time, % of riders based on months and semester weeks, etc.

[2]. **Weather, transport mode choices and emotional travel experience**

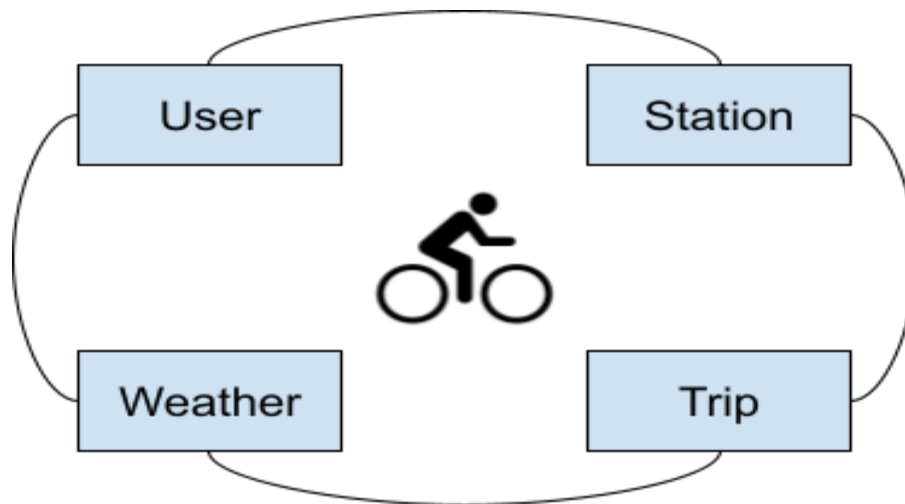
Effect of weather conditions affecting choices of transport mode, travel experiences along outdoor thermal perceptions. Insights of users' emotional experiences influence travel behavior in terms of public and active mode transports.

[3]. Factors influencing bike share membership: An analysis of Melbourne and Brisbane

The factors affecting membership in Australia's bike-sharing plan. The findings assist in the prospective increase of potential users in the future and the reliability of the transport system.

4. Methodology

We intend to use multiple datasets in our bike sharing database. This will have multiple tables to include user details, trip details, data on weather conditions, etc. The first step will be to create a data model in the form of an ER diagram with all the involved entities. We will then define relationships between these entities and convert them to a relational database. We will perform Normalization up to 3NF to remove all data inconsistencies and dependencies. The aim is to Join related tables using their primary keys and perform analysis to gain insightful information. Through the use of DDL and DML commands in MySQL, we will be able to answer the questions and identify trend patterns within the dataset. A part of the analysis will also be done in Python and Tableau will be used for final data visualization. A simple block diagram is represented below.



Keywords from syllabus - Enhanced ER Model, modeling entities and attributes, normalization, SQL commands/queries, joins, python

5. Deliverables and milestones

Topic	Details	Tentative Week
Data selection and proposal	Topic and data set to be used finalized	March 2 nd week
Scope finalization	Defining abstract and problem statement	March 4 th week
Data cleaning	Data cleaning and Normalization	April 1 st week
Analyzing data set	DDL, DML queries in MySQL, Processing across multiple tables using Joins	April 2 nd week
Documentation	Document the approach and publications	April 3 th week
Data Visualization	Visualization using Python and Tableau	April 4 th week

6. Team members and their roles

Topic	Team Member
Data selection and proposal	Bhakti, Khushee, Shreya, Varsha
Scope finalization	Bhakti, Khushee, Shreya, Varsha
Data cleaning	Khushee, Shreya
Analyzing data set	Bhakti, Khushee, Shreya, Varsha
Data Visualization	Bhakti
Documentation	Varsha

7. References

- [1].MaxNankervis, The effect of weather and climate on bicycle commuting, Transportation Research Part A Policy and Practice., Volume 33, Issue 6, August 1999, Pages 417-431 [https://doi.org/10.1016/S0965-8564\(98\)00022-6](https://doi.org/10.1016/S0965-8564(98)00022-6)
- [2].Lars Böcker, Martin Dijst, Jan Faber, Transportation Research Part A: Policy and Practice Volume 94, December 2016, Pages 360-373, <https://doi.org/10.1016/j.tra.2016.09.021>
- [3].Elliot Fishmana, Simon Washington, Narelle Haworth, Angela Watsonc, Transportation Research Part A: Policy and Practice Volume 71, January 2015, Pages 17-30 <https://doi.org/10.1016/j.tra.2014.10.021>