Climate Impact on Urban Mobility: Analyzing Bike-Sharing Demand

Apu Das - 23186407

Urban bike-sharing systems have emerged as a popular and eco-friendly mode of transportation in many cities, significantly mitigating climate change by reducing dependency on fossil fuels and lowering carbon emissions. The Capital Bikeshare and Seoul Bike Sharing systems are perfect examples of this trend. Understanding the factors that impact bike share rentals is crucial for optimizing efficiency and ensuring rider demand. This study analyzes the factors influencing bike share rentals using data from two different datasets from two countries. By exploring the impact of weather (temperature, humidity, and wind speed), we will try to find trends and patterns in bike rental demand. Thus, this analysis will provide valuable and practical insights into the question, "How does climate change (temperature, humidity, and wind speed) affect bike rentals?". It will also ensure sufficient bike availability and contribute to the fight against climate change by promoting sustainable transportation alternatives.

I. Data Sources

For the analysis, we utilized an SQLite dataset containing two tables: Capital Bikeshare [1] and Seoul Bikeshare [2]. The dataset is generated from an automated ETL pipeline. The Capital Bikeshare table provides hourly and daily rental bike counts from 2011 to 2012, along with weather and seasonal information. Similarly, the Seoul Bikeshare table offers hourly rental bike counts from 2017 to 2018, including weather data and holiday information. Thus, the dataset captures various features related to bike rentals and weather conditions, providing a comprehensive view of how climate factors affect bike-sharing usage.

The data pipeline ensures that the dataset is cleaned and transformed as required for the analysis. The final dataset (pipeline output) is complete with no significant missing values and no irrelevant features. Both the tables are structured with temporal, categorical, and continuous features, ensuring maximum coverage of relevant factors. The temperature, humidity, and windspeed features in Seoul Bikeshare are not min-max normalized (as the Capital Bikeshare). So, the pipeline made sure that these features were normalized. Scaling is needed for consistent comparisons between the two datasets. It allows meaningful comparisons between features. Without scaling, features with larger numerical ranges could skew the results and lead to incorrect interpretations.

A. Data Structure

The Capital bikeshare table is structured with temporal features (date, season, year, month, hour), categorical features (holiday, weekday, working day, weather situation), and continuous features (temperature, feeling temperature, humidity, wind speed). Notably, the dataset doesn't contain any missing values, ensuring a high quality of data for analysis. Samples from the dataset (pipeline output) are shown below (Fig 1 and Fig 2):

dteday	season	yr	mnth	hr	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	registered	cnt
2011-01-01	1	0	1	0	0	6	0	1	0.24	0.2879	0.81	0.0	3	13	16
2011-01-01	1	0	1	1	0	6	0	1	0.22	0.2727	0.80	0.0	8	32	40
2011-01-01	1	0	1	2	0	6	0	1	0.22	0.2727	0.80	0.0	5	27	32
2011-01-01	1	0	1	3	0	6	0	1	0.24	0.2879	0.75	0.0	3	10	13
2011-01-01	1	0	1	4	0	6	0	1	0.24	0.2879	0.75	0.0	0	1	1

Fig. 1. Samples from Capital Bikeshare table.

While Capital Bikeshare contains information for registered and casual rentals, Seoul Bikeshare contains information like Visibility, Solar Radiation, Rainfall, etc.

Date	Rented Bike Count	Hour	Temperature(°C)	Humidity(%)	Wind speed (m/s)	Visibility (10m)	Dew point temperature(°C)	Solar Radiation (MJ/m2)	Rainfall(mm)	Snowfall (cm)	Seasons	Holiday	Functioning Day
2017-12-01	254	0	-5.2	37	2.2	2000	-17.6	0.0	0.0	0.0	Winter	No Holiday	Yes
2017-12-01	204	1	-5.5	38	0.8	2000	-17.6	0.0	0.0	0.0	Winter	No Holiday	Yes
2017-12-01	173	2	-6.0	39	1.0	2000	-17.7	0.0	0.0	0.0	Winter	No Holiday	Yes
2017-12-01	107	3	-6.2	40	0.9	2000	-17.6	0.0	0.0	0.0	Winter	No Holiday	Yes
2017-12-01	78	4	-6.0	36	2.3	2000	-18.6	0.0	0.0	0.0	Winter	No Holiday	Yes

Fig. 2. Samples from Seoul Bikeshare table.

The Seoul Bike Sharing table is structured with temporal, categorical, and continuous features as well. It includes temporal features (Date and Hour), categorical features (Seasons, Holiday, Functioning Day), and continuous features (Rented Bike Count, Temperature, Humidity, Wind Speed, Visibility, Dew Point Temperature, Solar Radiation, Rainfall, Snowfall). It doesn't contain any missing values also.

B. *License*

Both datasets are licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license [3]. This license permits sharing and adapting the dataset for any purpose provided that appropriate credit is given. To comply with the licensing terms, I will ensure proper attribution is given to the UCI Machine Learning Repository and the creators of the datasets in all uses and publications. This will include citing the source in any reports, presentations, or publications resulting from this analysis. Additionally, any shared or adapted versions of the datasets will include the same CC BY 4.0 license, maintaining transparency and allowing further use by other researchers.

II. Analysis

To understand how climate change affects bike rentals, particularly temperature, humidity, and wind speed, we performed several data visualization techniques (Exploratory Data Analysis) and Correlation Analysis on the dataset. By examining these analyses on the bike-sharing systems, we found significant relationships between weather conditions (Temperature, Humidity, Wind Speed) and rental patterns, providing useful insights for enhancing the sustainability and efficiency of bike-sharing systems.

A. *Exploratory Data Analysis (EDA)*

Exploratory Data Analysis (EDA) is crucial for uncovering patterns, detecting anomalies, and gaining insights from the dataset. It enables a better understanding of data, which helps to take better and efficient decisions. By visualizing these relationships, we can better understand the impact of weather conditions on bike rental behavior, providing valuable information for bike rental efficiency. EDA also helps to identify necessary data preprocessing steps to enhance analysis accuracy.

Methodology: In the Exploratory Data Analysis (EDA) step, we explored the relationship between climate features (temperature, humidity, and wind speed) and bike rentals. To ensure comparability, we scaled the columns of the Seoul Bikeshare dataset to match those of the Capital Bikeshare dataset. We performed line plots to understand the linear relationships between the features and the number of bike rentals.

Specifically, we plotted temperature vs. bike rental counts, humidity vs. bike rental counts, and wind speed vs. bike rental counts for both tables. This approach allowed us to observe trends and patterns in the data visually. As we have data from two different cities with two different timeframes, we plotted two plots for every feature. This allowed us to see if the relationships are comparable between the two cities and timeframes.

Results: The line plots revealed varied relationships between climate features and bike rentals but similar patterns in both cities and timeframes. First, we plotted a line graph to analyze the impact of temperature on bike rental counts over time. Here, the temperature is on the x-axis, and bike rental counts are on the y-axis.

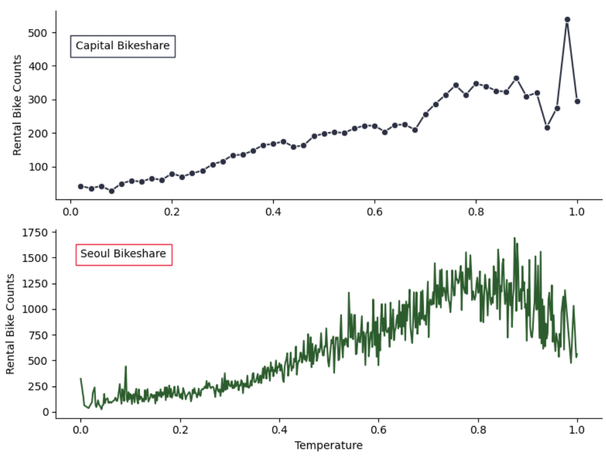

Fig. 3. Effects on Temperature on Bike Rentals: Capital vs. Seoul Bikeshare.

Then, we plotted another line graph to analyze the impact of humidity on bike rental counts over time. Here, humidity is plotted on the x-axis, while bike rental counts are on the y-axis.

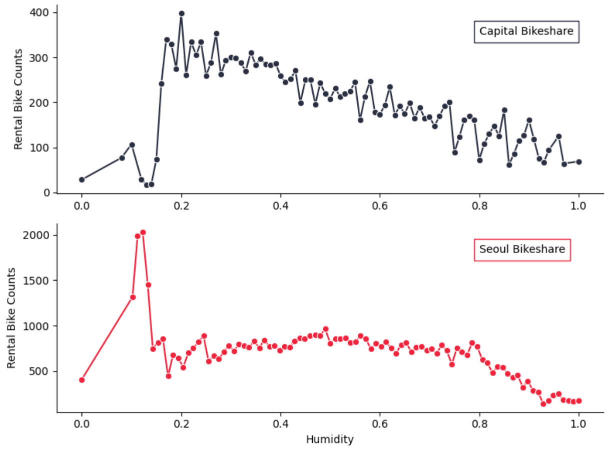

Fig. 4. Effects on Humidity on Bike Rentals: Capital vs. Seoul Bikeshare.

Finally, a line graph to analyze the impact of wind speed on bike rental counts over time. Here, humidity is plotted on the x-axis and wind speed is on the y-axis.

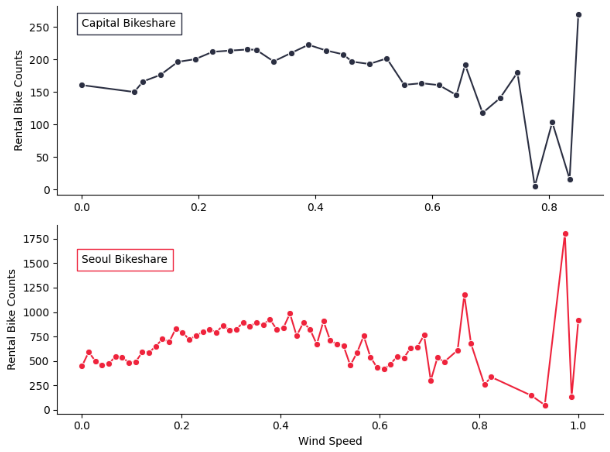

Fig. 5. Effects on Wind Speed on Bike Rentals: Capital vs. Seoul Bikeshare.

Interpretation: The exploratory data analysis (EDA) reveals significant insights into how weather conditions affect bike rental patterns. For temperature, there was a clear positive relationship with bike rental counts in both cities, with a more noticeable effect in Seoul. As temperatures increased, bike rentals also increased, likely due to increased comfort and the appeal of outdoor activities. Conversely, for humidity, there was a generally negative relationship with bike rentals, indicating that higher humidity levels corresponded to fewer bike rentals, possibly because of the discomfort associated with humid conditions.

However, wind speed showed a less clear relationship, with little variations in rental counts as wind speed changed, suggesting it is a less significant factor compared to temperature and humidity. Though wind speed does not significantly impact bike rentals in most cases, it displayed a notable spike in bike rentals when wind speeds were exceptionally high.

B. Correlation Analysis

Correlation analysis helps identify relationships between numerical features, providing insights into how changes in one feature may relate to changes in another. As temperature, humidity, and wind speed are all numerical values, it is crucial to conduct correlation analysis to understand how these weather features may be related to each other.

Methodology: As we have data from two different cities with two different timeframes, we computed two correlation matrices for each city. This allowed us to see how the weather features relate to the bike count and if the relationships are comparable between the two cities. Therefore, this analysis also helps us assess whether these relationships remain consistent globally and across different timeframes.

Specifically, we calculated correlations (Pearson correlation coefficients) between temperature, humidity, wind speed, and bike rental counts. To improve the clarity of our analysis, we adjusted the correlation matrices by setting the diagonal elements to zero. This adjustment eliminates self-correlation, enhancing the visual interpretation of relationships between weather features and bike rental counts. Using heatmaps, we effectively illustrated these correlations, showing a clear and insightful picture of how variations in weather conditions may impact bike rentals. For the heatmap, we focused on visualizing only the lower triangle since the correlation matrices are symmetric.

Results: The resulting heatmap visually displays the correlations between temperature, humidity, wind speed, and bike rentals, providing clear insights into how weather conditions influence bike rental demand between the cities and time periods.

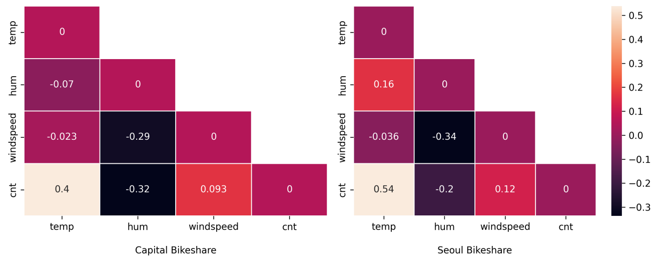

Fig. 6. Heatmap of Weather Effects on Bike Rentals: Capital vs. Seoul Bikeshare.

Interpretation: The correlation analysis shows significant insights into the relationship between weather features and bike rentals. It highlights that temperature has a significant impact on bike rentals in both datasets, with a moderate positive correlation of 0.405 in the Capital Bikeshare data and a stronger positive correlation of 0.539 in the Seoul Bikeshare data. This suggests that warmer temperatures lead to increased bike rentals. Additionally, humidity shows a negative correlation with bike rentals, particularly in the Capital Bikeshare data, with a correlation of -0.323, indicating that higher humidity may deter bike usage. These findings suggest that temperature plays an important role in bike rental activity, while humidity has a considerable but lesser role.

III. Conclusion

Our analysis shows a strong and clear relationship between climate change (temperature, humidity, and wind speed) and bike rentals. Through correlation analysis and visualization techniques, we found that temperature has a significant positive importance on bike rental counts, indicating that warmer weather encourages more people to rent bikes. Conversely, humidity negatively impacts bike rentals, suggesting high humidity discourages bike renting. However, wind speed showed weaker correlations compared to temperature and humidity, indicating it has a minor effect on rental patterns.

While the study provides clear evidence that weather conditions significantly influence bike rentals, it also has several limitations. The datasets analyzed were limited to specific regions, which may not capture all regional variations in weather impacts. Additionally, the analysis did not include other factors such as rainfall, snowfall, solar radiation, seasonal changes, and socio-economic variables.

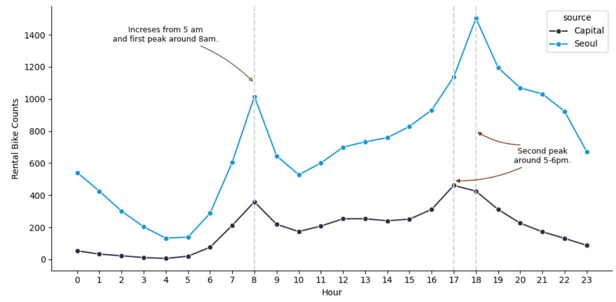

Fig. 7. Average Bike Rentals by hours: Capital vs. Seoul Bikeshare.

Notably, as we further analyzed with other features, we found that the hour of the day plays a significant role in rental behavior. Bike rentals begin to rise sharply from around 5 am, peaking first at 8 am and again around 5 pm and 6 pm, corresponding with commuting hours. This indicates that many users rely on bike rentals for their daily commute. Furthermore, the influence of weather conditions varies by hour, with higher temperatures boosting rentals in the afternoon and humidity having a greater impact in the morning. Therefore, future research can incorporate these additional factors to provide a more comprehensive understanding of how climate change affects bike rentals.

References

[1] Fanaee-T,Hadi. (2013). Bike Sharing. UCI Machine Learning Repository. https://doi.org/10.24432/C5W894.

[2] Seoul Bike Sharing Demand. (2020). UCI Machine Learning Repository. https://doi.org/10.24432/C5F62R.

[3] Creative Commons Attribution 4.0 International (CC BY 4.0): https://creativecommons.org/licenses/by/4.0/deed.en