

Analysing and predicting bike rental demand based on weather, date and time



Sten-Egert Märtson
Kaur Kivilaan
Project E8

Bike rental

Bike rental systems, such as Capital Bikeshare, provide an **accessible** and **eco-friendly** transportation option for **urban areas**. These programs allow users to rent bicycles **via an app** or **kiosk** for short-term use, enabling convenient travel for **commuting, errands, or leisure**.

Our goal

The goal of this project was to analyze the **public data offered by Capital Bikeshare** to understand how various factors, such as **weather** (temperature, humidity, wind speed) and **date-related features** (holidays, weekends, seasons, and time of day), influence bike rentals. By exploring these correlations, including the impact of weather, we aimed to uncover **trends in hourly and seasonal usage patterns**.

Since the idea for this project originally came from a **Kaggle** competition , we aimed to create a sufficiently good **machine learning** model to predict **hourly bike rentals based on the features mentioned above**.

The insights gained can help **optimize bike availability during peak times** and **plan maintenance during periods of low demand**, improving overall **system efficiency** and **user satisfaction**.

The Data

The core data of our project originates from the **Washington D.C.** based bike sharing company **Capital Bikeshare**. We used the data given in their **collaborative competition with Kaggle** [1] and the **public data offered on their website** [2]. Kaggle data encompasses amounts of bikes rented during each hour from **2011 to 2013**. There are **17379** rows which equate to **3 million** bike rides. Capital Bikeshare data however covers each ride from **September of 2010 to November 2024** with new data being added each month. There are a combined number of **44.1 million bike rides** in said dataset.

Data Preparation & ML Models

We **converted** the Capital Bikeshare **per-ride data to the per-hour Kaggle format**. Most of the analysis was done on said new data using libraries like **matplotlib**, **Seaborn** and **pandas**.

Models were evaluated on **root mean squared logarithmic error (RMSLE)**. While making the models, our sights were set on achieving an RMSLE score of **at least 0.4** or lower in the Kaggle competition.

As a way to test the quality of our data preparation, we also tried training some models on the much larger self-made dataset. We also cleaned up the data as there was some outlier situations (e.g. COVID).

	Kaggle	Capital Bikeshare System Data
Random forest	0.484	0.400
XGBoost	0.399	0.398
TensorFlow	0.508	0.456

Yearly trends:

- Bike rentals have **grown** steadily from 2012 to 2019, reflecting **increasing popularity and availability**
- A **sharp decline** occurred in 2020, likely due to the **COVID-19 pandemic**
- Since 2020, rentals have rebounded rapidly, reaching **record highs** starting in 2023

Monthly trends:

- Rentals are at their **lowest** during **winter** months
- Summer** months maintain **consistently high** with September being the peak month
- Rentals **rise** sharply in **spring** and **decline** gradually in **autumn**

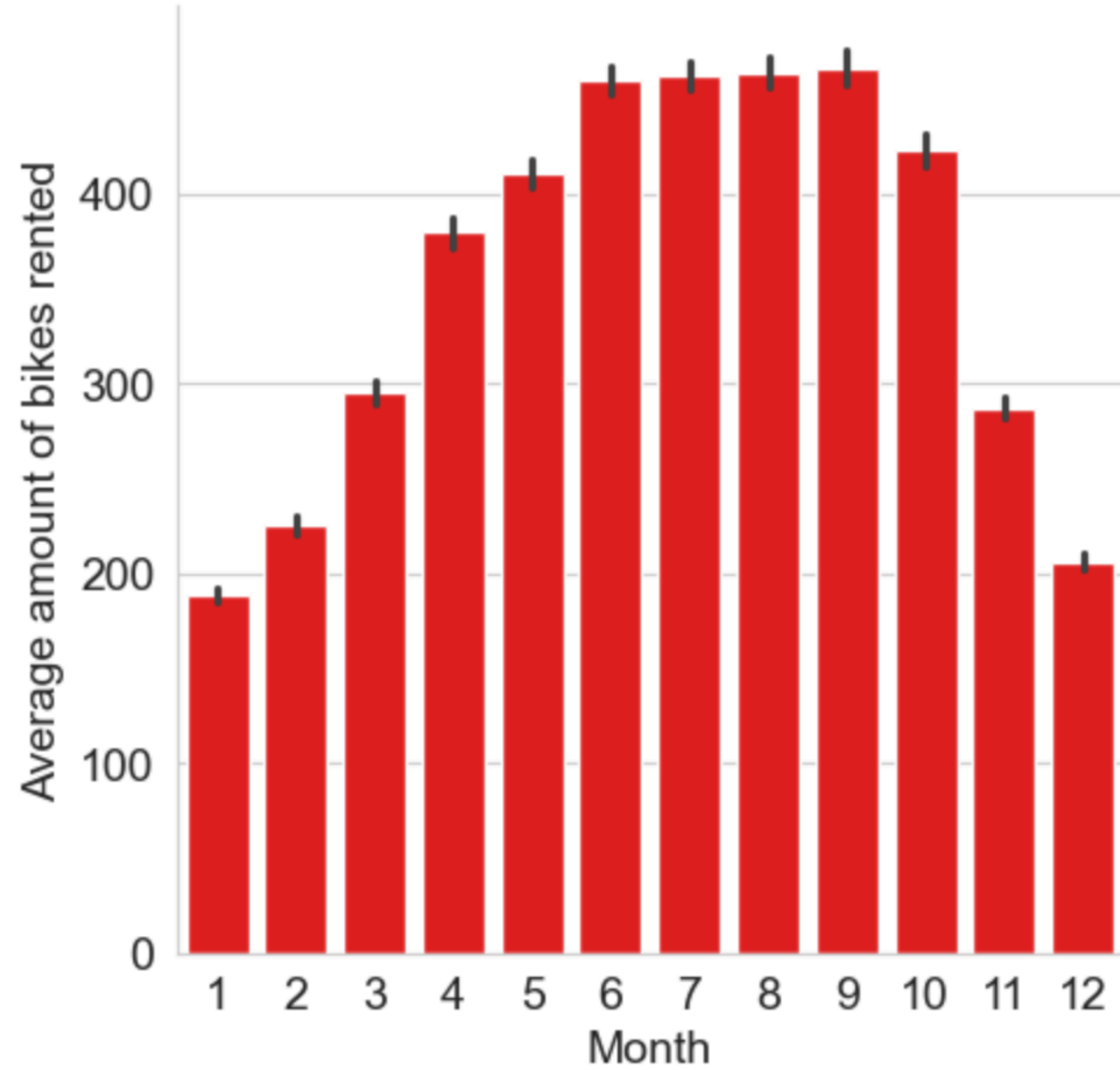
Hourly trends:

- Morning Rush:** Bike rentals **gradually increase** from early morning, **peaking sharply** around 8 AM as commuters begin their day, then taper off slightly by late morning
- Afternoon Peak:** Rentals **remain steady** through midday and early afternoon, with the **highest point of activity** occurring at 5 PM as people finish work and evening activities begin
- Evening Decline:** In the evening, rental activity **remains high** for a while but starts to **decline** after 7 PM, with usage **dropping steadily** into the late night hours

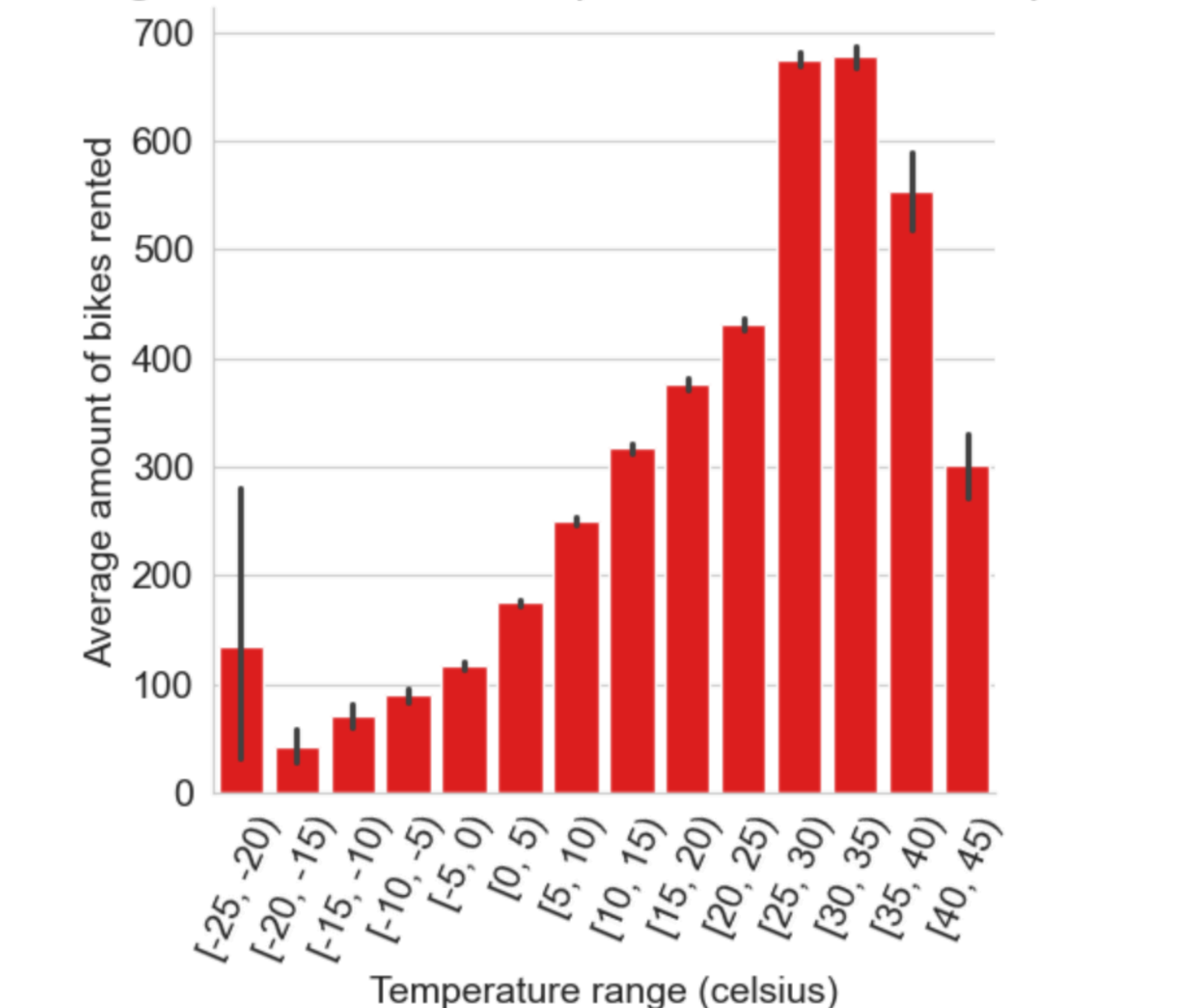
Weather-based trends:

- Higher temperatures** means **more bikes rented** up until about **35°C** after which demand sharply declines
- Demand is highest when wind speed is low** - the amount of bikes rented declines with higher speed values with sudden drops at **30+ km/h**
- The **ideal humidity for bike rental is 15-20%**, with a consistent linear decrease in demand if humidity is higher or lower

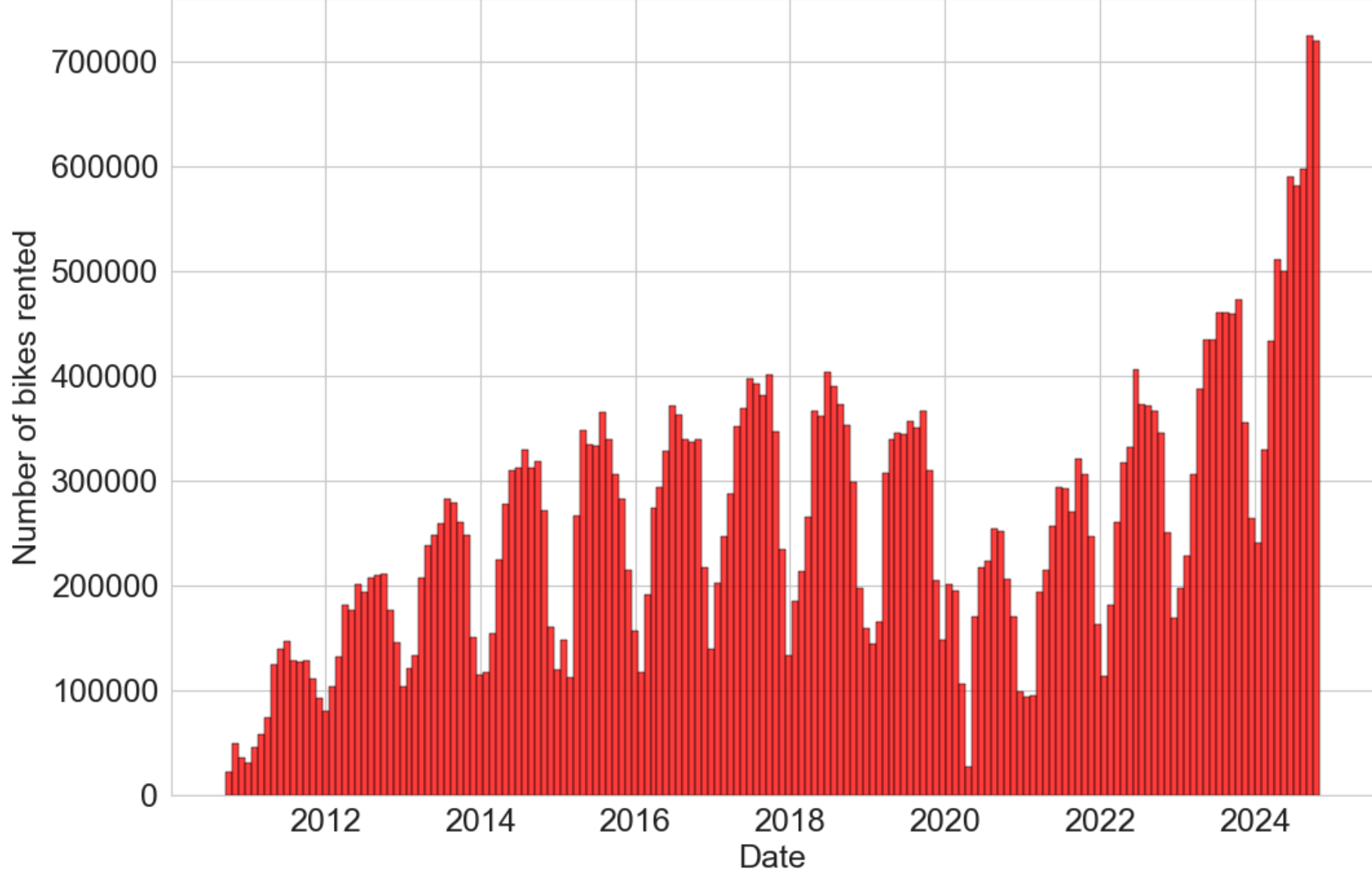
Average bike rental amount per hour based on month



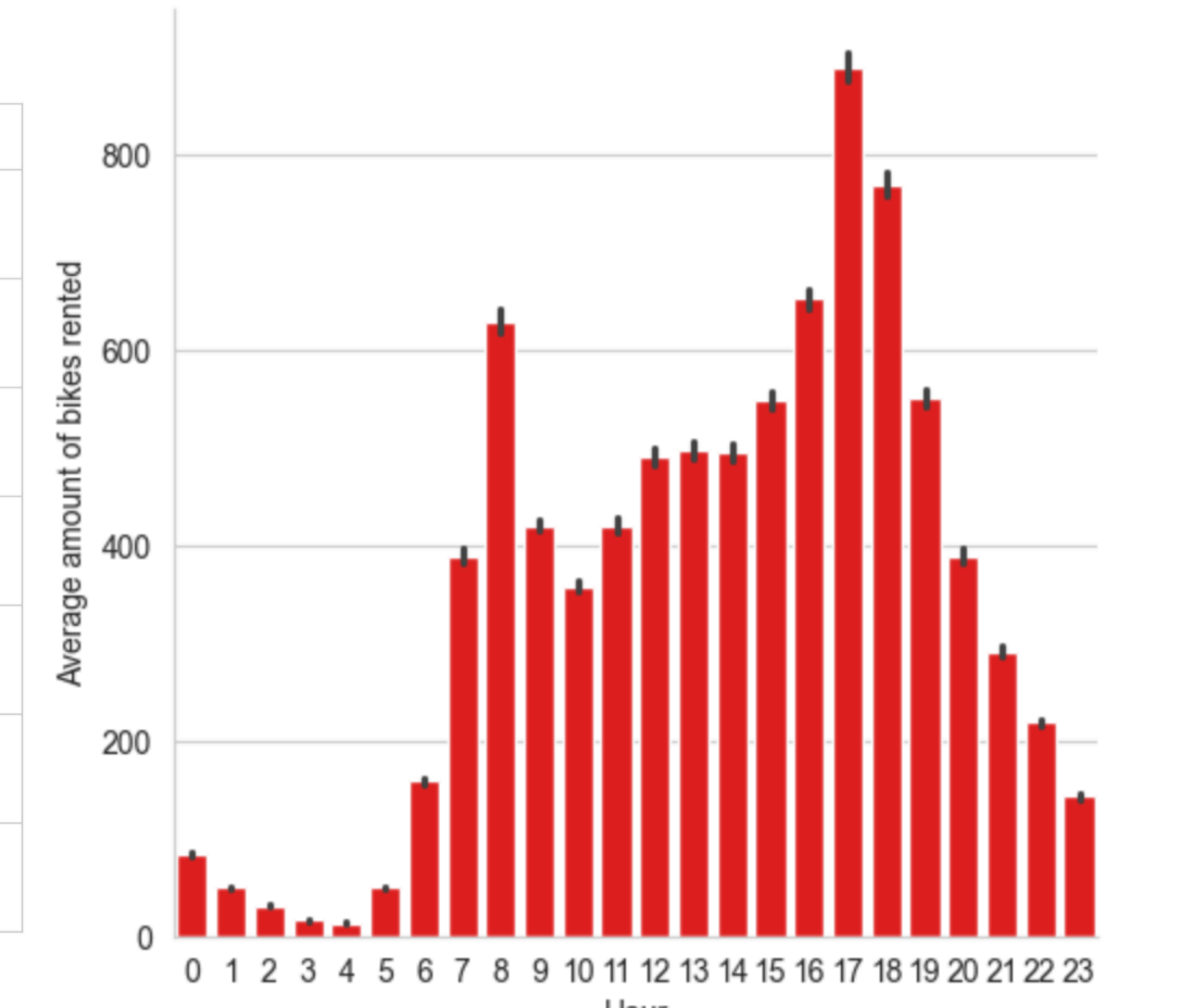
Average bike rental amount per hour based on temperature



Number of bikes rented from 2010 to 2024



Bike sharing averages by hour



References

[1] *Bike Sharing demand*. 2015. URL: <https://www.kaggle.com/competitions/bike-sharing-demand/overview>
[2] *System Data*. 2024. URL: <https://capitalbikeshare.com/system-data>

