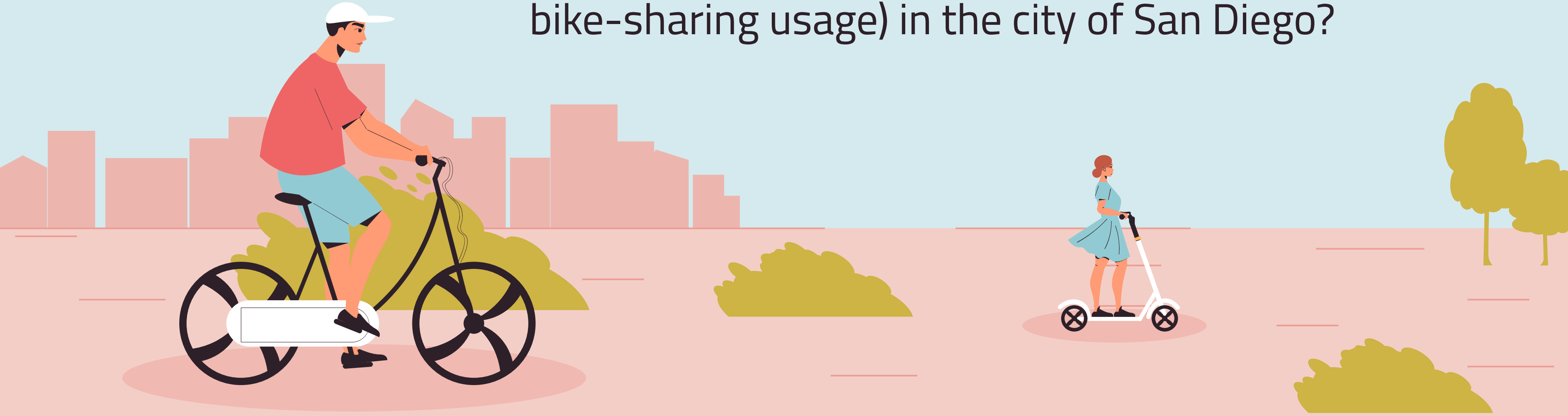
Weather & Wheels

Evaluating San Diego Bike-Sharing Usage in Varying Weather Conditions

By: Keanu Ventura & Nathan Park

Problem

How do weather conditions (such as temperature, wind, and precipitation influence bike-sharing usage) in the city of San Diego?



Why is it Important?

Bike-sharing systems rely on accurate demand forecasting for deployment and maintenance decisions.

Weather strongly influences sudden spikes or drops in bike-sharing usage.

Reliable weather data helps operators allocate bikes efficiently across the city.

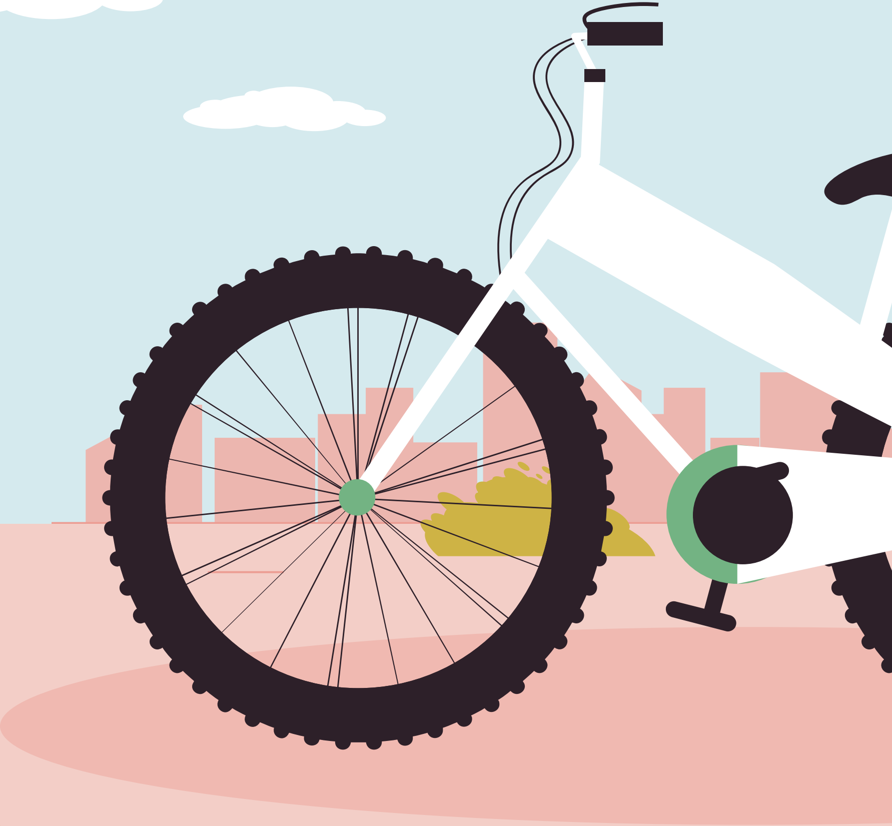
Better forecasting reduces empty or overcrowded stations and improves availability.

Optimization by understanding how weather affects local usage.

Improved reliability encourages more people to use bike-sharing.

Hypothesis

Bike-sharing usage in San Diego decreases on days with extreme weather conditions, and we anticipate that bike-sharing services will peak on milder, more moderate weather conditions. We also expect rainfall to be the strongest predictor, causing the largest sudden drops.



Datasets Required

Bike-Sharing Usage Data for San Diego

https://bikeshare.metro.net/about/data/

San Diego Temperature and Rainfall Data

https://www.weather.gov/wrh/climate?wfo=sgx

Wind Speed in San Diego

https://oceaninformatics.ucsd.edu/datazoo/catalogs/ccelter/datasets/13

2025 San Diego Variable Means

63.0

Temperature (°F)

On average, San Diego saw temps of 63.0° F.

11.5

Wind speed (m/s)

Described as a “Fresh Breeze” by the National Weather Service.

0.02

Precipitation (mm)

Throughout the entire year (up to now), San Diego has received an average of 0.02 mm precipitation (per day).

Partially Cloudy

Condition (clear, rain, cloudy, etc.)

During 2025, San Diego was, on average, usually partially cloudy.

Variables With Most Bike-Share Trips (Biased Raw Count)

68

Temperature rounded to nearest 1°F

37537 bike-share trips → 11% of total usage

10.5

Wind Speed rounded to nearest 0.5 m/s

57389 bike-share trips → 17% of total usage

0.0

Precipitation rounded to the nearest .01 mm

308109 bike-share trips → 91% of total usage

Partially cloudy

Weather Condition

191960 bike-share trips → 57% of total usage

We Normalized → Highest # Rides on a Typical Day by Variable

63

Temperature rounded to nearest 1°F

1562 bike-share trips
Lowest 52 → 877 bike-share trips

13.0

Wind Speed rounded to nearest 0.5 m/s

1338 bike-share trips
Lowest 18.5 → 754 bike-share trips

No Rain

Precipitation binned into rain or no rain

No Rain: 1270 bike-share trips
Rain: 1028 bike-share trips

Rain

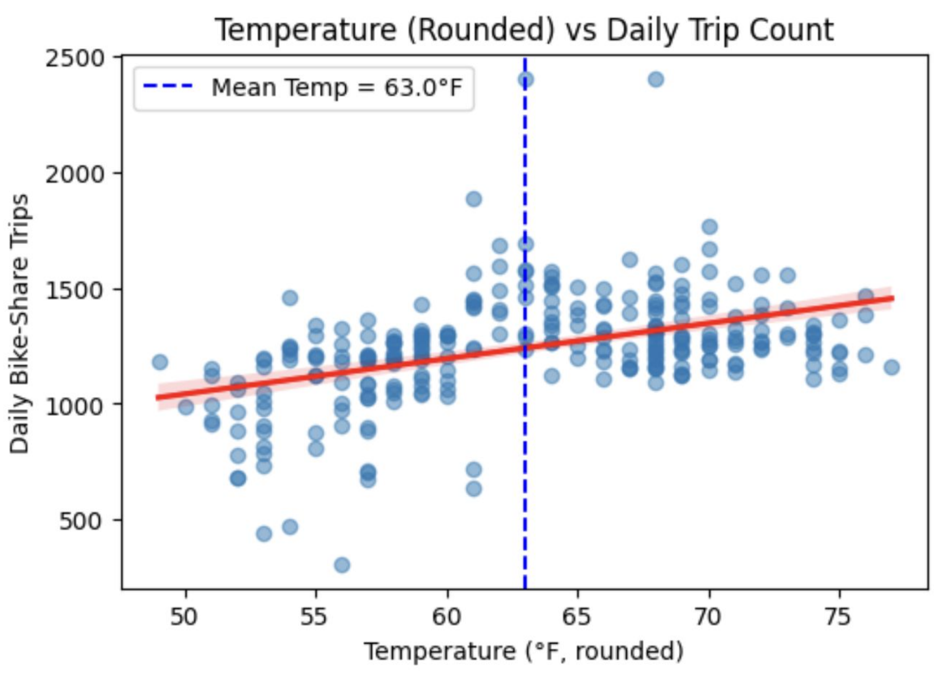
Weather Condition

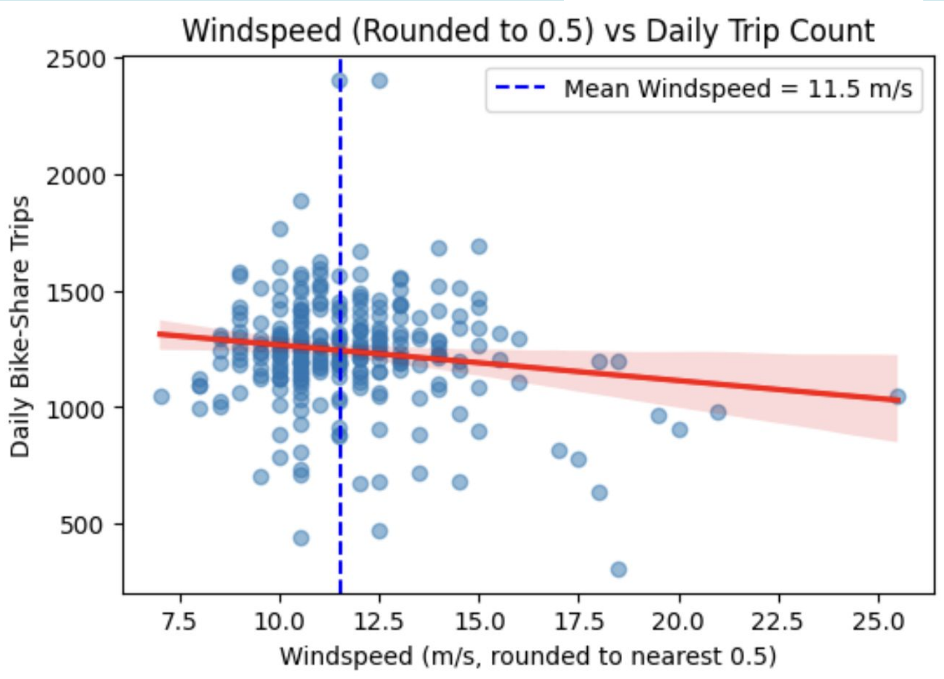
1344 bike-share trips
Lowest Rain, Overcast → 929 bike-share trips

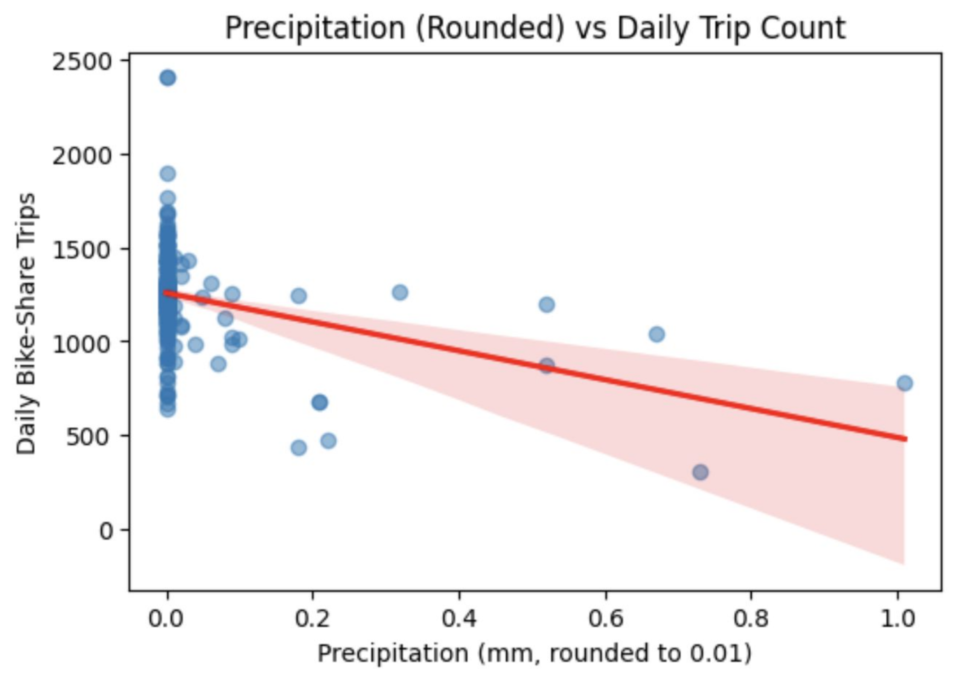
Rain Has the Highest Average Daily Trips?

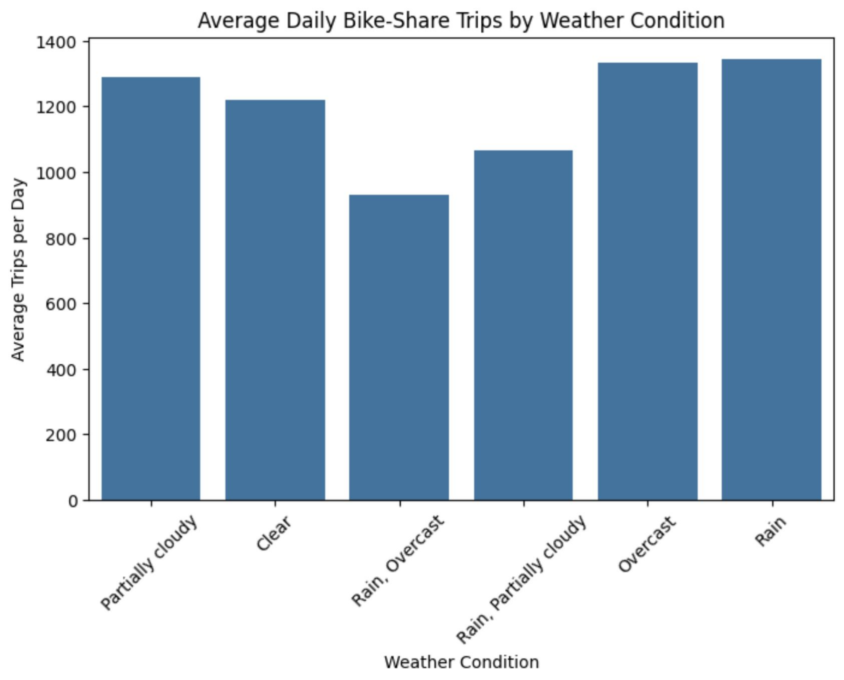
"Rain" days are extremely rare in our dataset (appears once), making the estimate unreliable.

conditions	
Clear	95
Overcast	15
Partially cloudy	181
Rain	1
Rain, Overcast	11
Rain, Partially cloudy	31


Temperature (Rounded) vs Daily Trip Count
Mean Temp = 63.0°F
Daily Bike-Share Trips
2500
2000
1500
1000
500
50
55
60
65
70
75
Temperature (°F, rounded)


Windspeed (Rounded to 0.5) vs Daily Trip Count
Mean Windspeed = 11.5 m/s
Daily Bike-Share Trips
2500
2000
1500
1000
500
7.5
10.0
12.5
15.0
17.5
20.0
22.5
25.0
Windspeed (m/s, rounded to nearest 0.5)


Precipitation (Rounded) vs Daily Trip Count
Daily Bike-Share Trips
Precipitation (mm, rounded to 0.01)
2500
2000
1500
1000
500
0
0.0
0.2
0.4
0.6
0.8
1.0


Average Daily Bike-Share Trips by Weather Condition
Average Trips per Day
Weather Condition
1400
1200
1000
800
600
400
200
0
Partially cloudy
Clear
Rain, Overcast
Rain, Partially cloudy
Overcast
Rain

Multiple Linear Regression Model

Lastly, we built a multiple linear regression model using the variables as predictors, with daily bike-share ridership as the response variable.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                trip_id   R-squared:                       0.275
Model:                            OLS   Adj. R-squared:                  0.253
Method:                 Least Squares   F-statistic:                     12.51
Date:                Tue, 02 Dec 2025   Prob (F-statistic):           3.36e-15
Time:                        14:48:16   Log-Likelihood:                -1837.2
No. Observations:                 273   AIC:                             3692.
Df Residuals:                     264   BIC:                             3725.
Df Model:                           8
Covariance Type:            nonrobust
=============================================================================================
                                coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------------------
const                       392.7003    150.343      2.612      0.010      96.677     688.724
temp_round                   12.5220      1.974      6.345      0.000       8.636      16.408
windspeed_round               3.8042      6.169      0.617      0.538      -8.342      15.950
precip_round               -400.9966    156.791     -2.558      0.011    -709.716     -92.277
cond_Overcast               121.5094     58.154      2.089      0.038       7.006     236.013
cond_Partially cloudy        45.1381     29.151      1.548      0.123     -12.260     102.536
cond_Rain                   -24.3738    209.761     -0.116      0.908    -437.390     388.643
cond_Rain, Overcast        -126.8221     75.877     -1.671      0.096    -276.223      22.579
cond_Rain, Partially cloudy -44.3684     58.171     -0.763      0.446    -158.906      70.169
==============================================================================
Omnibus:                       86.909   Durbin-Watson:                   1.167
```

Weather Explains Only 27.5% of the Variation

It makes sense that weather covers a moderate portion of variation in ridership because variation also comes from:

- Day of Week
- Special Events
- Holidays
- Seasonality
- Tourism
- Commute Cycles
- Bike Availability
- Random Day to Day Factors

Variable Predictor Results

Temperature is strongest predictor

Each 1 °F increase → ~12.5 more rides per day. P-value < 0.0001 → temperature statistically significant

Wind speed is not meaningful

Coef = +3.80 → effect is very small and likely noise. P-value = 0.54 → not statistically significant

Precipitation has a significant (-) effect

Light rain reduces ridership by ~401 trips per 1 mm of rain.More precipitation → less bike-share usage

Overcast has a significant (+) effect

Overcast days ~121 more trips than the baseline condition. Mild conditions → increased ridership

Challenges & Limitations Encountered

Data

Original datasets we found didn’t have the data we needed (not from 2025, no precipitation data, no wind data, bikeshare data up to 2025 q3). Final weather dataset: https://www.visualcrossing.com/weather-api/

Moderately Weak Predictive Model

Again, weather alone cannot explain most ridership variation, so the model’s predictive power is limited. There are other important factors previously mentioned that can be included.

Future Ideas

Bike Type

The bike-share usage data includes a bike type column (standard or electric) which may be interesting to explore later.

Hourly-level data

If we examine hourly-level bike ridership instead of daily we may discover more detailed patterns and improve prediction.

More Variables

If we incorporate additional variables that drive ridership variation we may learn more.