

# CAPITAL BIKE SHARE

## Predictions for its Philadelphia Branch



A study by Esra Saydam

# CONTEXT

**Capital Bike Share**, is an established bike share company, which was originally founded in Washington DC in 2010.

In the year of 2023, the company holds 5600 bikes, and 700+ stations across 7 jurisdictions: Washington, DC.; Arlington, VA; Alexandria, VA; Montgomery, MD; Prince George's County, MD; Fairfax County, VA; and the City of Falls Church, VA.

It has over 3.4 million rides per year and over 9322 rides per day.

## Problem

Capital Bike Share decided to launch its 8th jurisdiction in the metropolitan area of Philadelphia next. The company is looking for an estimate on the number of days and hours they would expect in their first year in Philadelphia to assure that they can have their 10% percent of profit like in D.C area.

# DATASETS

- We were provided with hourly and daily bike ride data from 2011 and 2022.

In [7]: `day_data.head()`

Out[7]:

	instant	dteday	season	yr	mnth	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	registered	cnt
0	1	2011-01-01	1	0	1	0	6	0	2	0.344167	0.363625	0.805833	0.160446	331	654	985
1	2	2011-01-02	1	0	1	0	0	0	0	0.363478	0.353739	0.696087	0.248539	131	670	801
2	3	2011-01-03	1	0	1	0	1	1	1	0.196364	0.189405	0.437273	0.248309	120	1229	1349
3	4	2011-01-04	1	0	1	0	2	1	1	0.200000	0.212122	0.590435	0.160296	108	1454	1562
4	5	2011-01-05	1	0	1	0	3	1	1	0.226957	0.229270	0.436957	0.186900	82	1518	1600

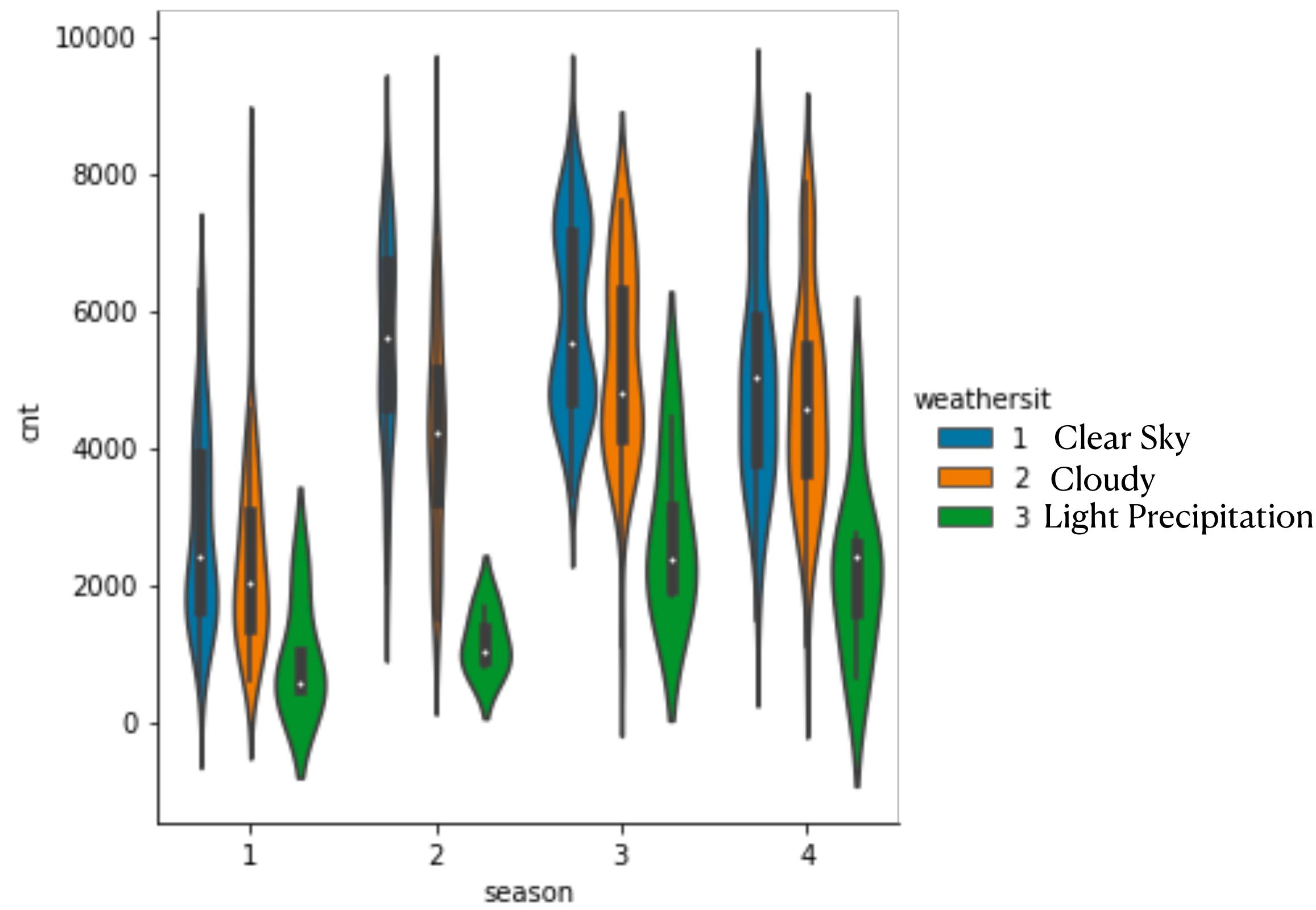
In [9]: `hour_data.head()`

Out[9]:

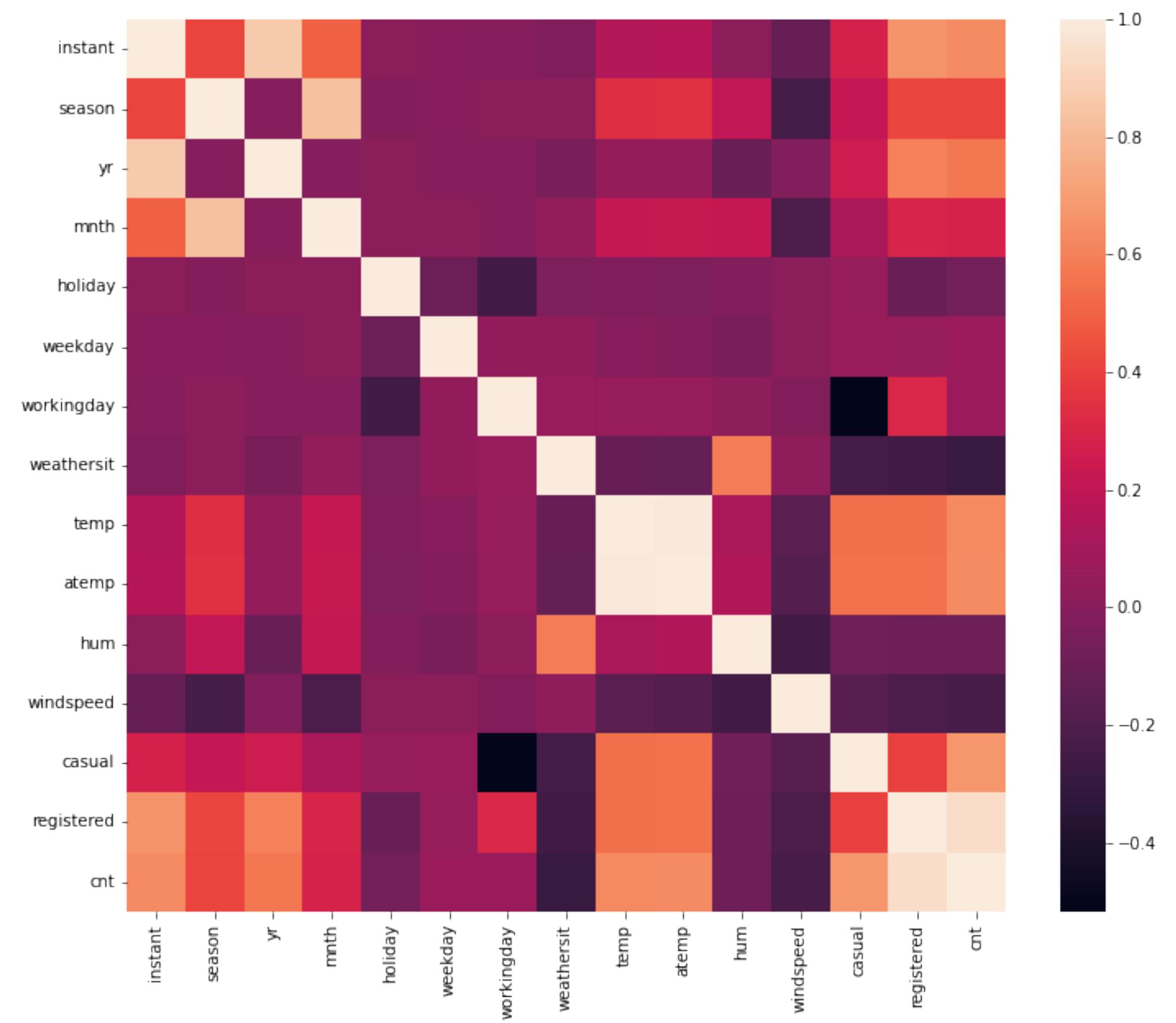
	instant	dteday	season	yr	mnth	hr	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	registered	cnt
0	1	2011-01-01	1	0	1	0	0	6	0	1	0.24	0.2879	0.81	0.0	3	13	16
1	2	2011-01-01	1	0	1	1	0	6	0	1	0.22	0.2727	0.80	0.0	8	32	40
2	3	2011-01-01	1	0	1	2	0	6	0	1	0.22	0.2727	0.80	0.0	5	27	32
3	4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

# VISUALIZATION (of day data)

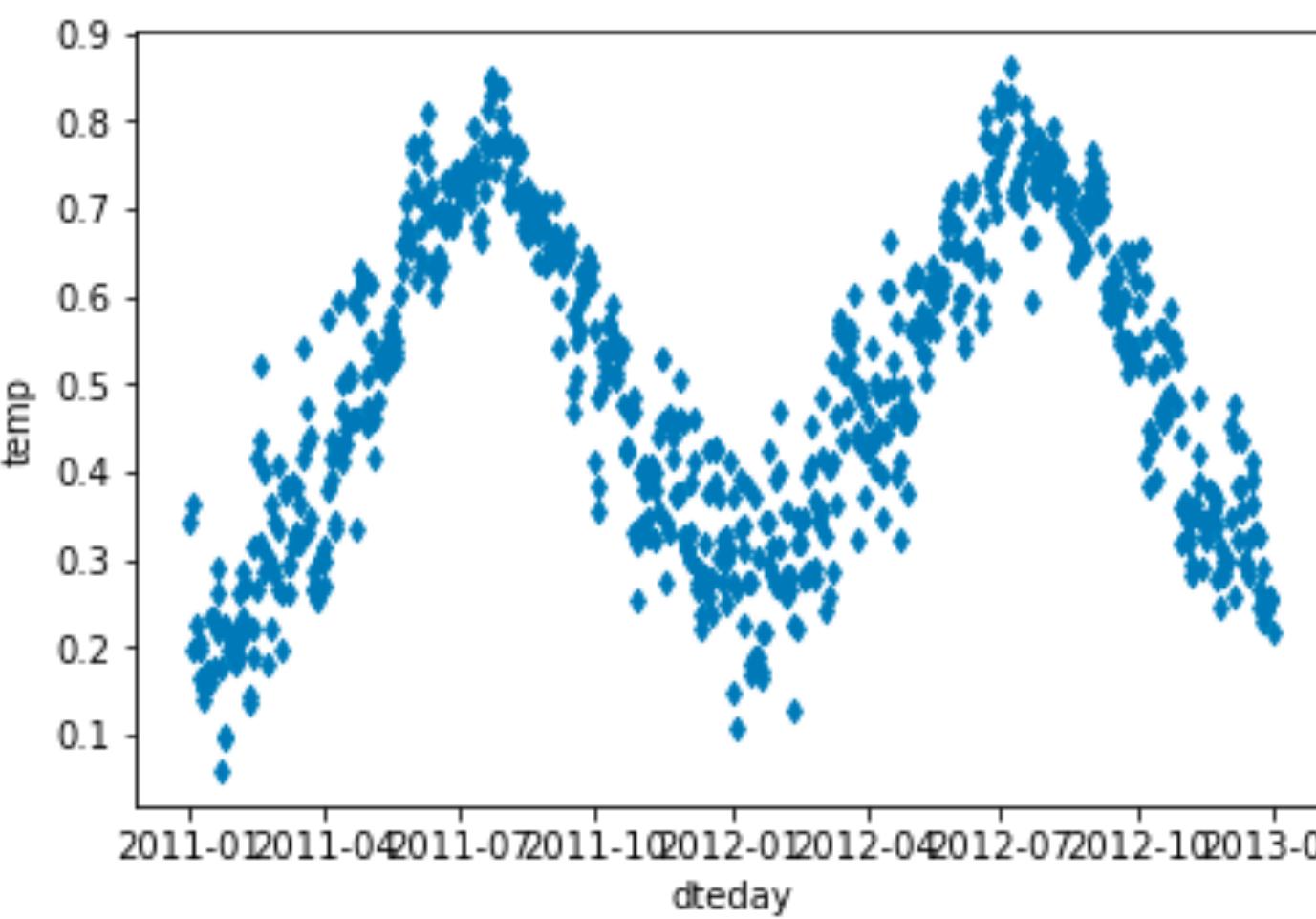
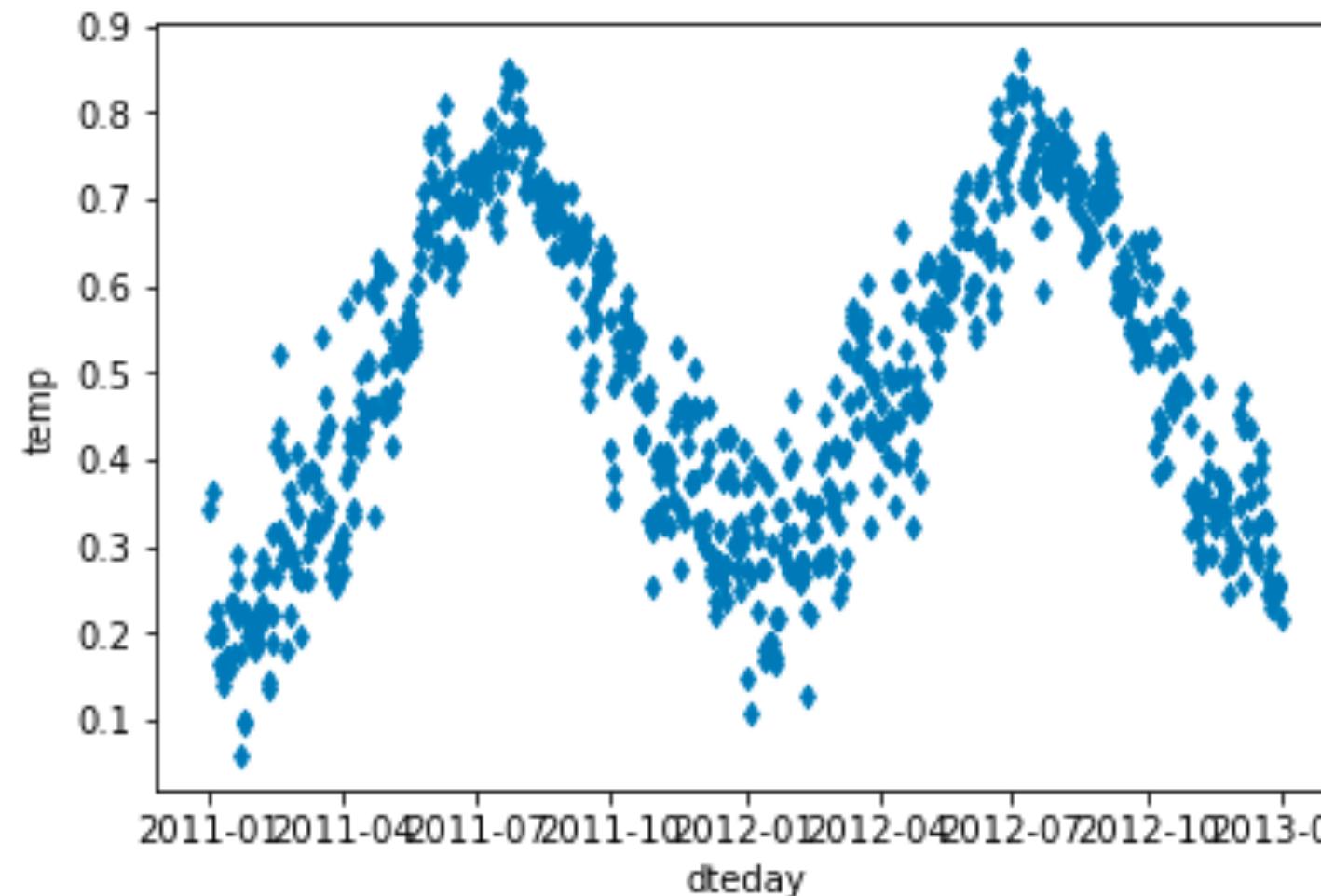
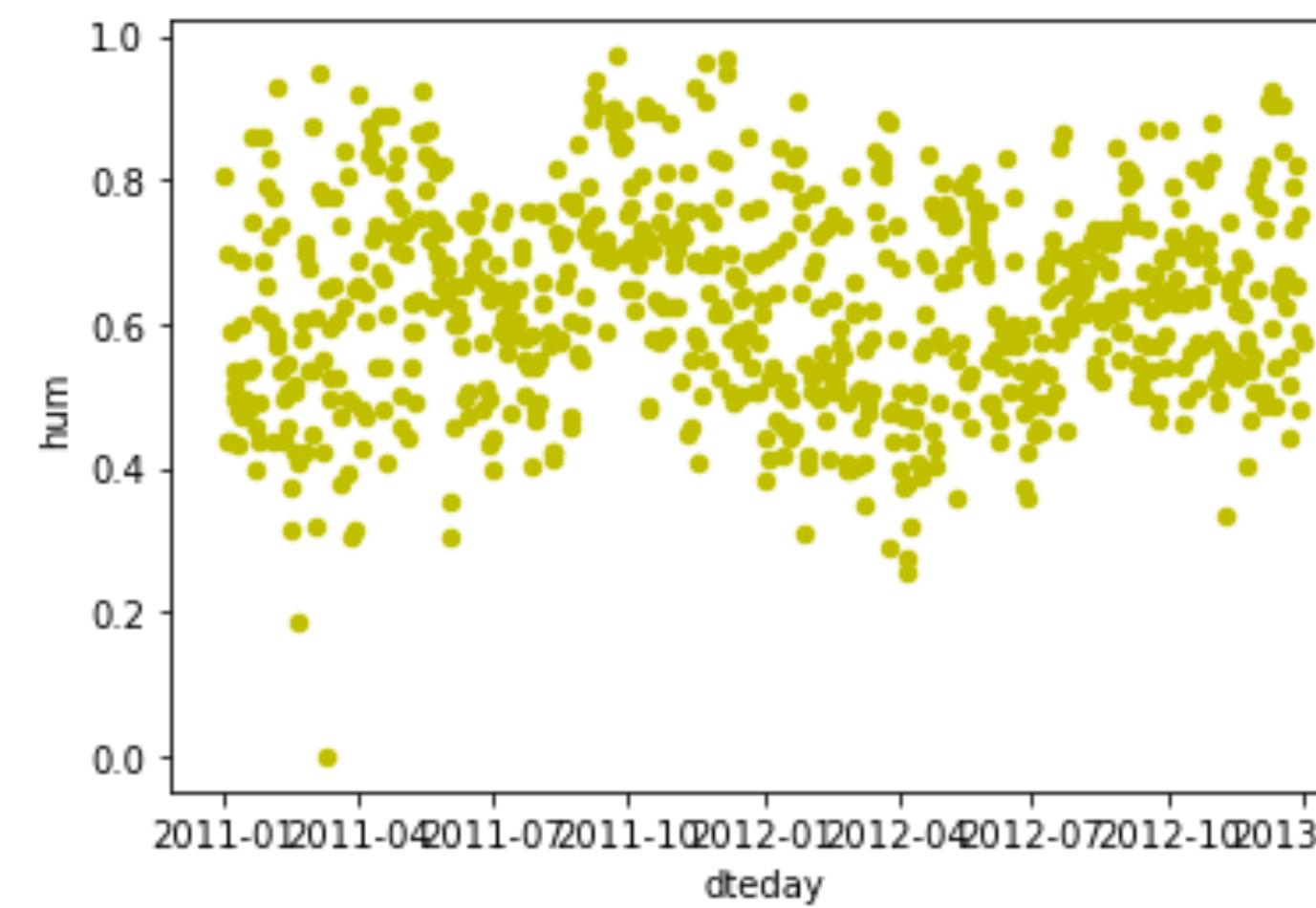
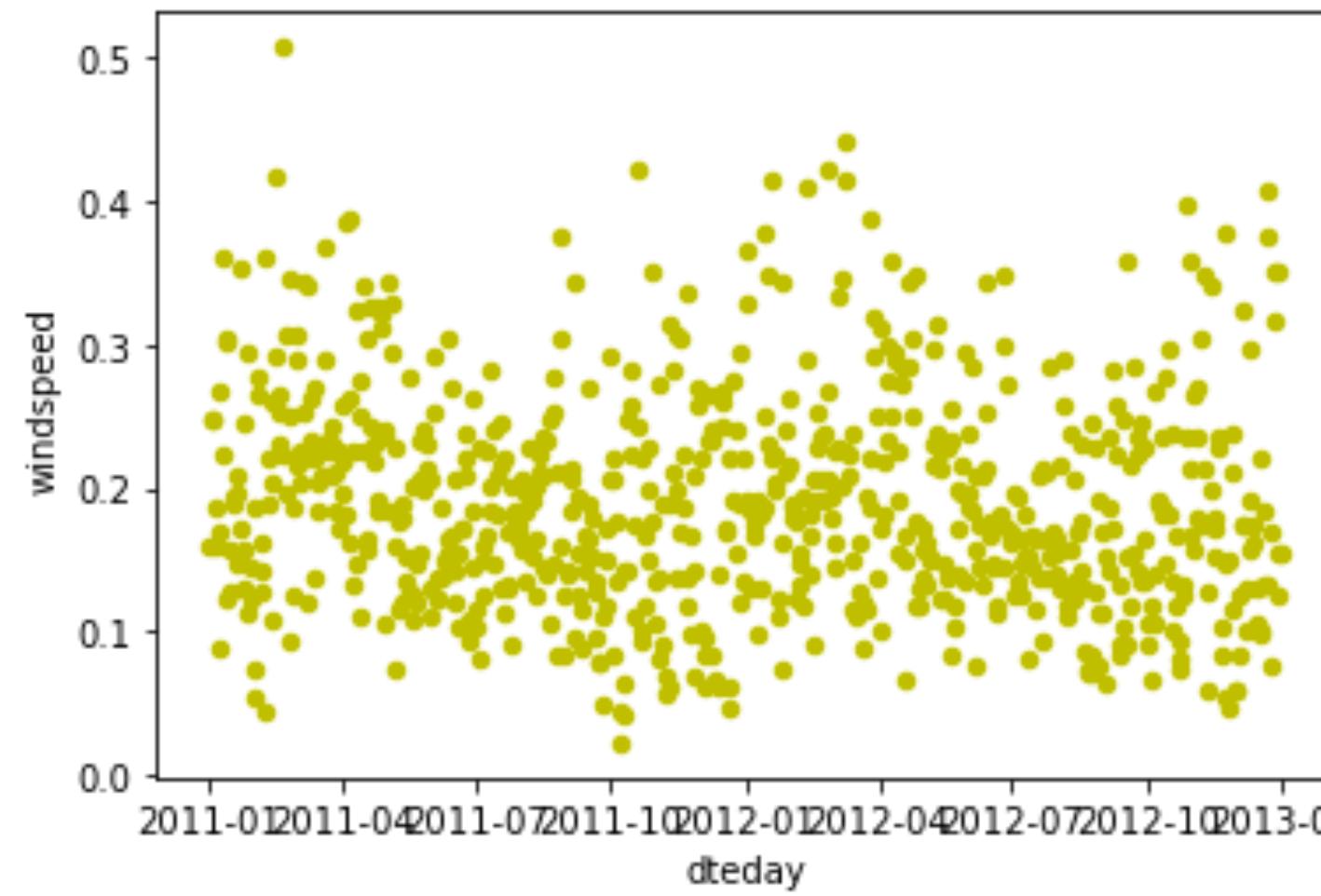
Spring and Fall are the busiest seasons with precipitation negatively impacting the bike ride count



Temperature, month, season, year are related to count

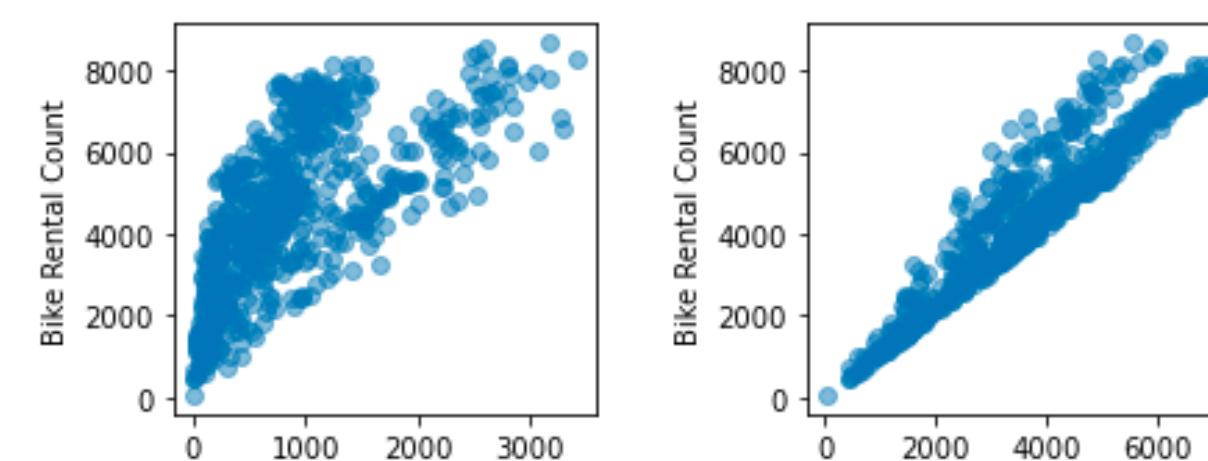
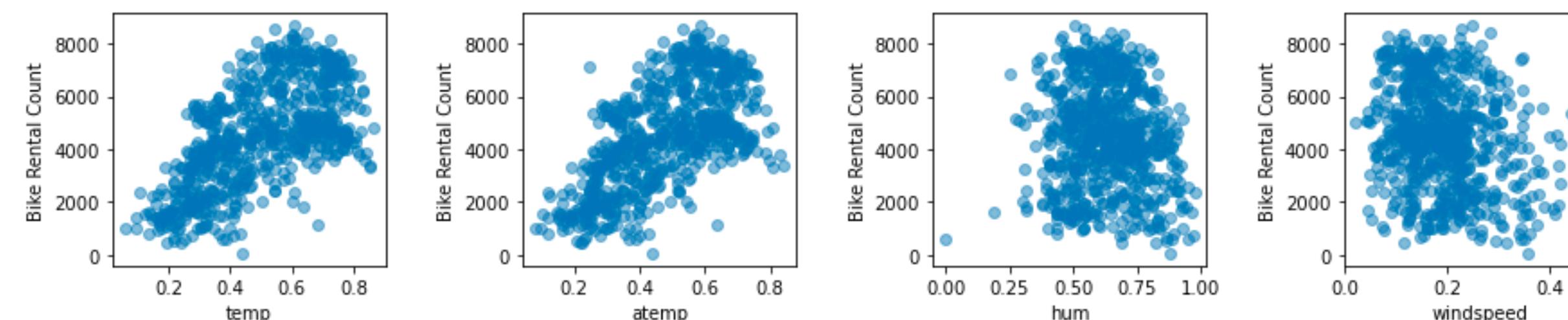
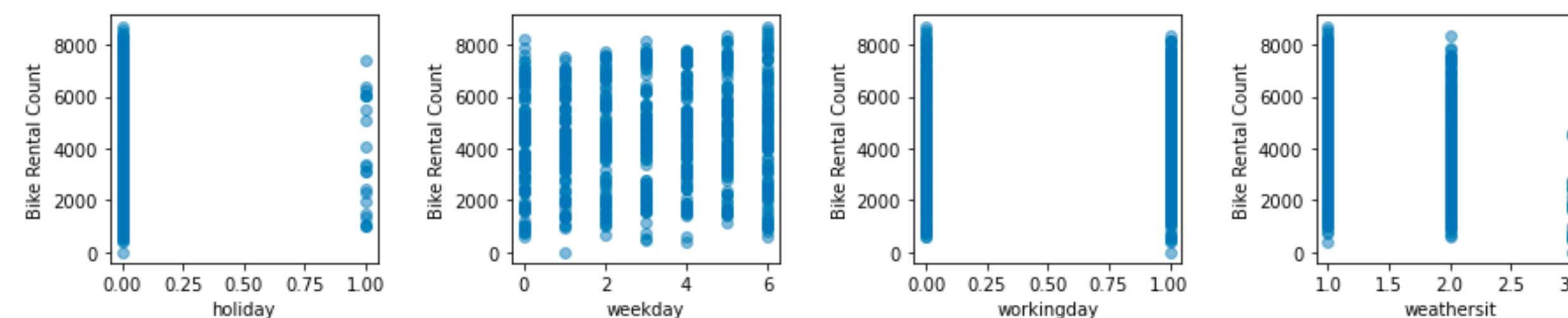
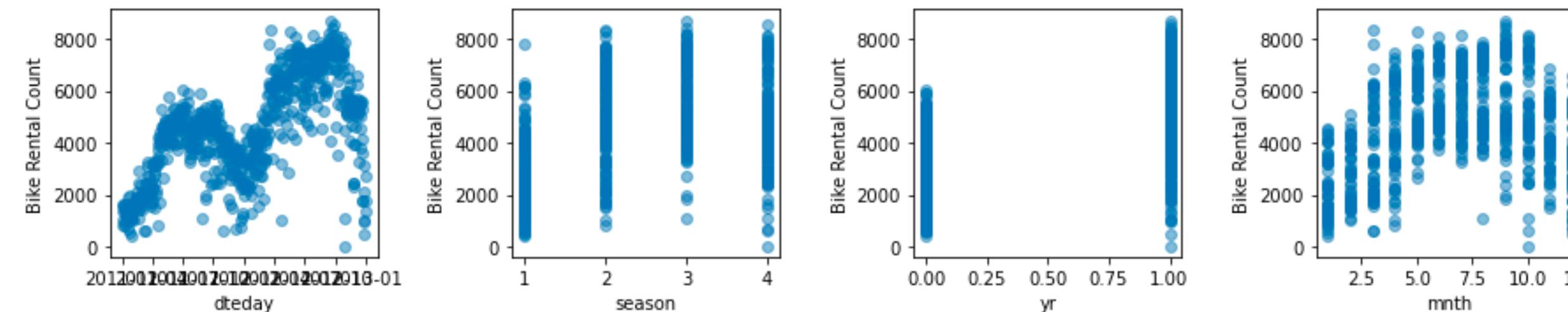


# VISUALIZATION



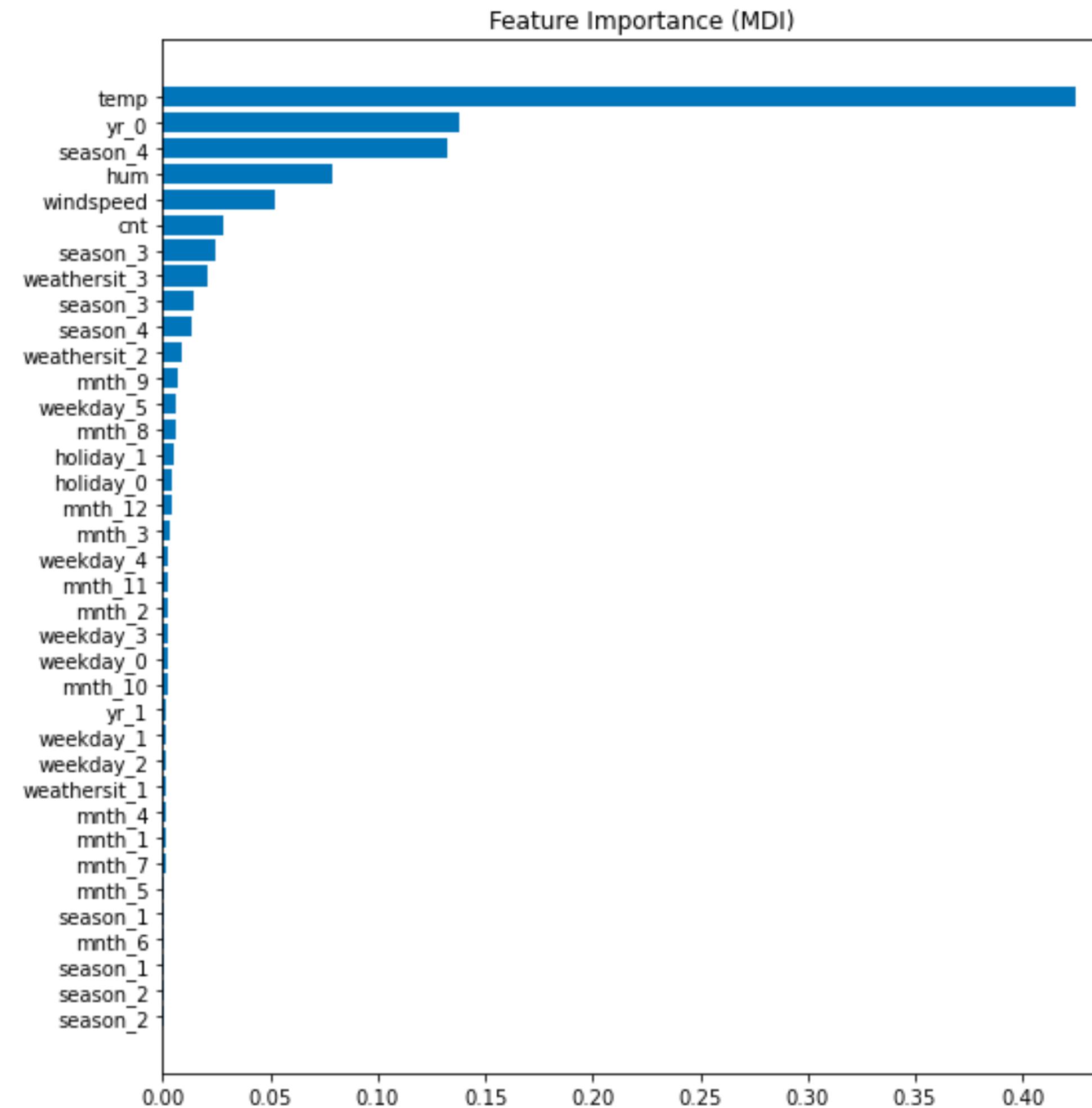
Low humidity and wind are equally correlated to low number of bike ride counts just like high humidity and wind. That's because low humidity and wind also exists in very cold weathers which are strongly affecting the bike ride counts.

# VISUALIZATION



**Correlations of features  
to Bike Ride Count**

# MODELING



## STRONGEST FEATURES AFTER MODELING

- 1) Temperature
- 2) Weather Situation
- 3) Holiday
- 4) Month
- 5) Year (Whether its the company's first year or not)
- 6) Windspeed
- 7) Humidity

\* I'll use only count values over 'registered' and 'casual' values since their sums equal to count values. I'll also focus on season rather than months

# MODELING

Model: Gradient Boosting Regressor

Measurer: Grid Search CV

Hyperparameters to use:

```
learning_rate': 0.01,  
'max_depth': 4,  
'n_estimators': 1000,  
'random_state': 1,  
'subsample': 0.5
```

# MODELING

## PREDICTING THE PHILADELPHIA VAULES

Philadelphia Population: 1,576,000

DC Population: 712,816

The population ratio: 2.21

- We created a function called ‘predict\_philly’ which would take a weather-forecasted day or list of days of Philadelphia, apply the model prediction and then multiply it with the population ratio.

# CONCLUSION

We made our predictive function predict the bike ride count of January 1st 2023 in Philadelphia. And the number was 2968. Three times more than the first day in DC. According to this result, we can predict that 10 percent profit seems likely since DC is already a profitable region for Capital Bike

**Improvements:**

**We could have been provided these materials and answers to have a more accurate prediction:**

- Philadelphia weather forecast dataset or a weather dataset that we can use to predict the future weather with ARIMA model
- Tourist numbers for each city 
- Calendar specific to Philadelphia and its local events, holidays if any.
- The effect of each city's hills on the bike rent counts. (DC has no hills but how would the hills on northwest Philadelphia prevent bikers from biking.)

To guarantee 10 percent profit, I would also suggest building bike stations next to college campuses and business headquarters since college students and employed people might have more commitment to bikes as part of daily routines.



**PHILADELPHIA WILL BE A SUCCESS!**