IntroBA Team Project
‘Seoul Bike Sharing’ Data

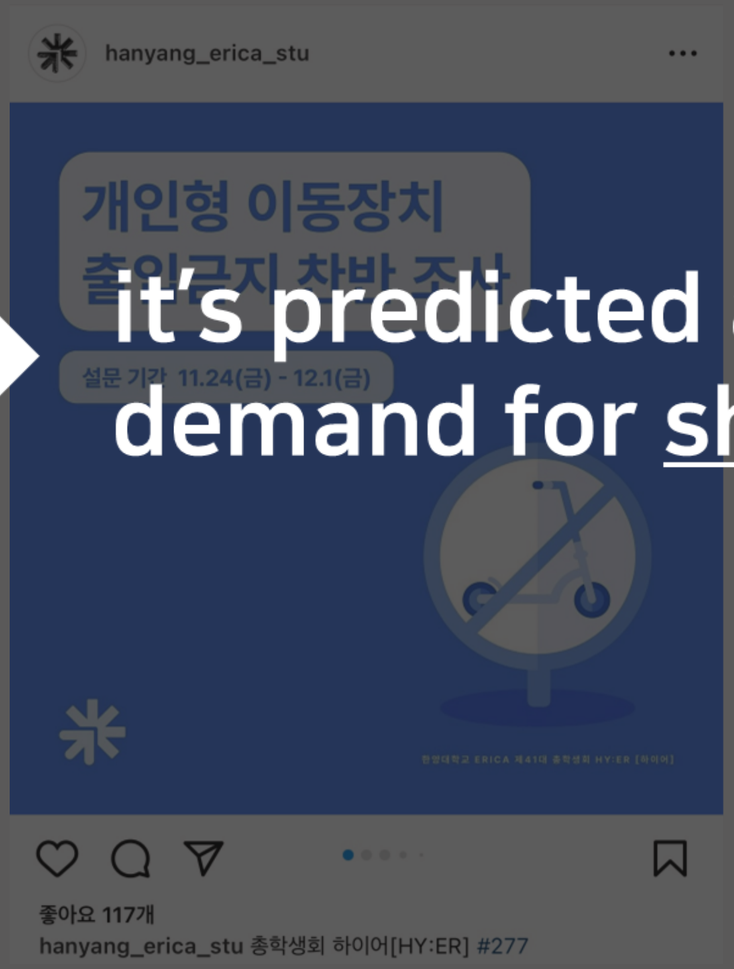
류승권 장해영 김요섭 남윤서 이주은

CONTENTS

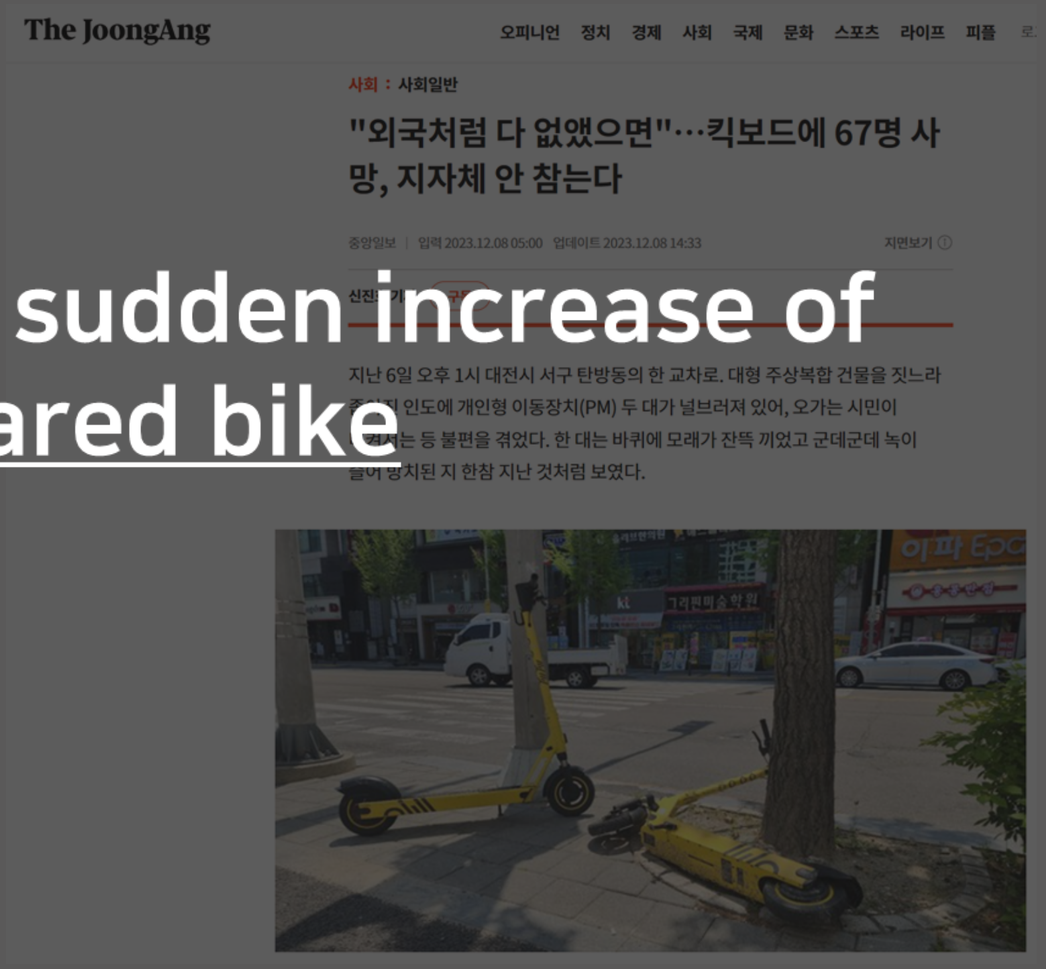
Why this Dataset is Important?

Recently, the shared kickboard problem has spread

Recently, the shared kickboard problem has spread



it’s predicted a sudden increase of demand for shared bike



Data Introduction & Preprocessing

• Data Introduction - 14 Columns

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8760 entries, 0 to 8759
Data columns (total 14 columns):
 #   Column                     Non-Null Count  Dtype
---  ------                     --------------  -----
 0   Date                       8760 non-null   object
 1   Rented Bike Count          8760 non-null   int64
 2   Hour                       8760 non-null   int64
 3   Temperature(° C)            8760 non-null   float64
 4   Humidity(%)                8760 non-null   int64
 5   Wind speed (m/s)           8760 non-null   float64
 6   Visibility (10m)           8760 non-null   int64
 7   Dew point temperature(° C)  8760 non-null   float64
 8   Solar Radiation (MJ/m2)    8760 non-null   float64
 9   Rainfall(mm)               8760 non-null   float64
 10  Snowfall (cm)              8760 non-null   float64
 11  Seasons                    8760 non-null   object
 12  Holiday                    8760 non-null   object
 13  Functioning Day            8760 non-null   object
dtypes: float64(6), int64(4), object(4)
memory usage: 958.3+ KB
```

• Data Preprocessing - Check for missing values

```
Date                          0
Rented Bike Count             0
Hour                          0
Temperature(° C)              0
Humidity(%)                   0
Wind speed (m/s)              0
Visibility (10m)              0
Dew point temperature(° C)    0
Solar Radiation (MJ/m2)       0
Rainfall(mm)                  0
Snowfall (cm)                 0
Seasons                       0
Holiday                       0
Functioning Day               0
dtype: int64
```

• Data Preprocessing

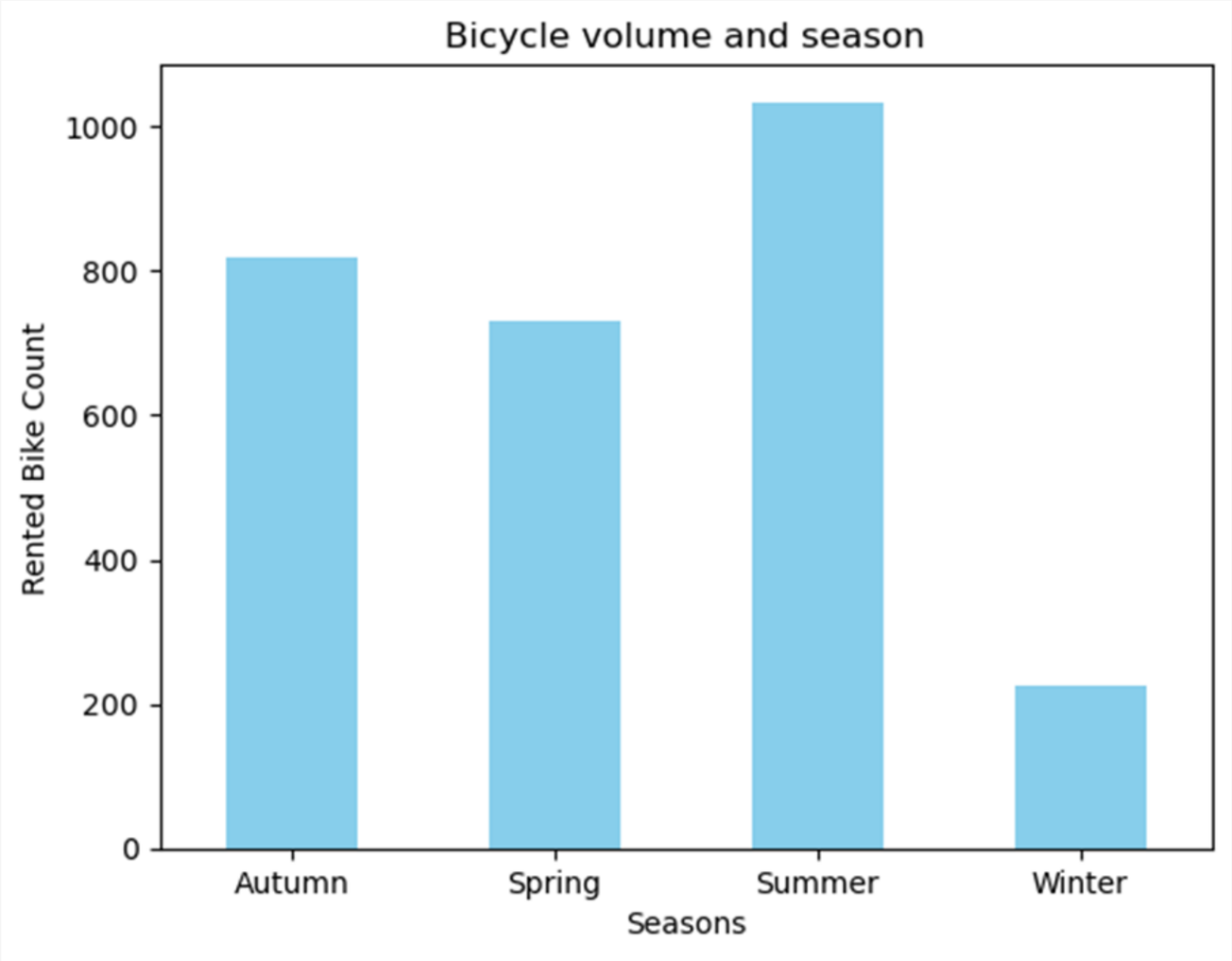
- Binary Encoding for Holiday Column

```
df['Holiday'] = df['Holiday'].apply(lambda x: 0 if x == 'No Holiday' else 1)
```

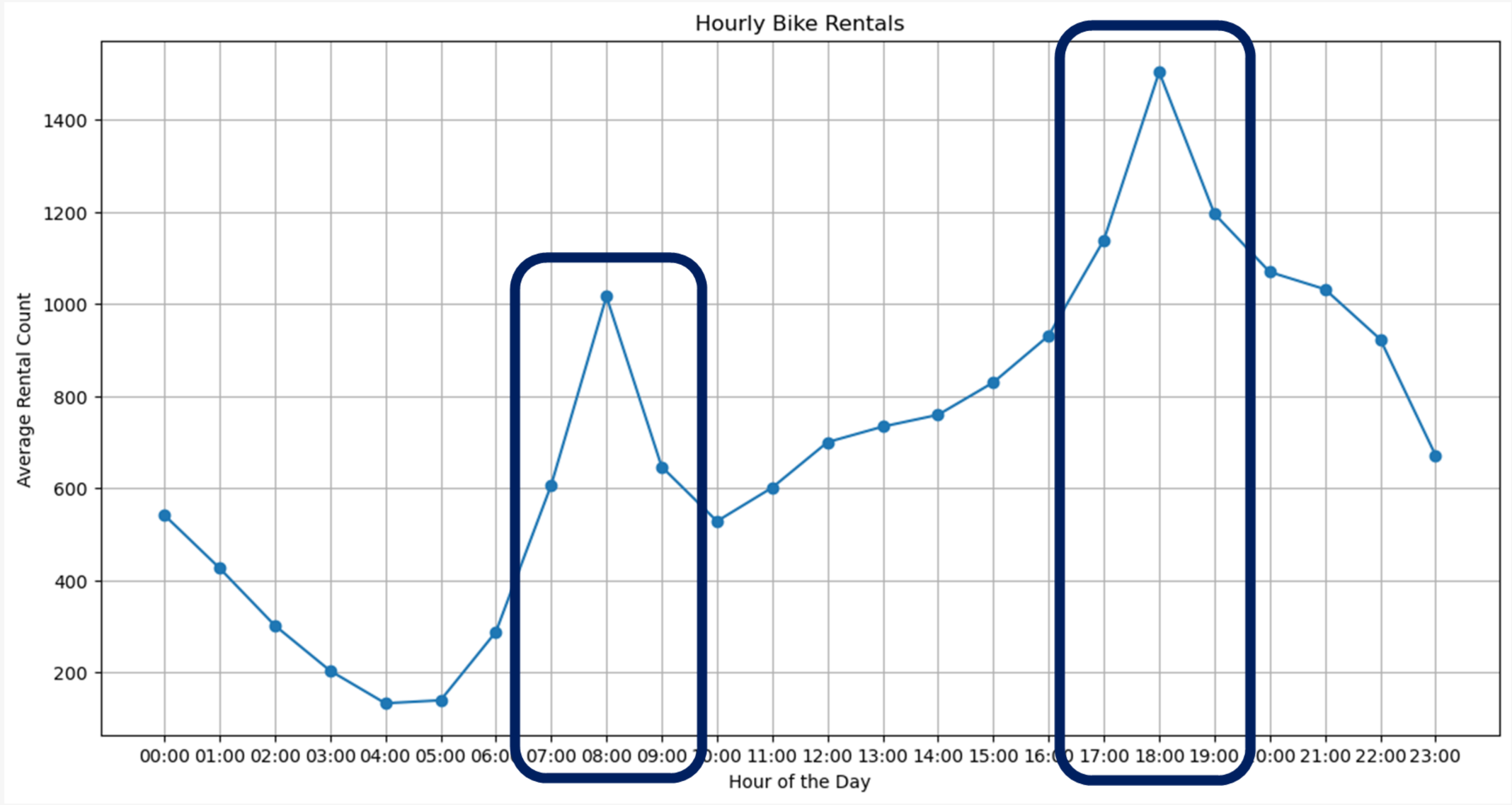
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8760 entries, 0 to 8759
Data columns (total 14 columns):
 #   Column                     Non-Null Count  Dtype  
---  ------                     --------------  -----  
 0   Date                       8760 non-null   object 
 1   Rented Bike Count          8760 non-null   int64  
 2   Hour                       8760 non-null   int64  
 3   Temperature(° C)           8760 non-null   float64
 4   Humidity(%)                8760 non-null   int64  
 5   Wind speed (m/s)           8760 non-null   float64
 6   Visibility (10m)           8760 non-null   int64  
 7   Dew point temperature(° C)  8760 non-null  float64
 8   Solar Radiation (MJ/m2)    8760 non-null   float64
 9   Rainfall(mm)               8760 non-null   float64
 10  Snowfall (cm)              8760 non-null   float64
 11  Seasons                    8760 non-null   object 
 12  Holiday                    8760 non-null   int64  
 13  Functioning Day            8760 non-null   object 
dtypes: float64(6), int64(5), object(3)
memory usage: 958.3+ KB
```

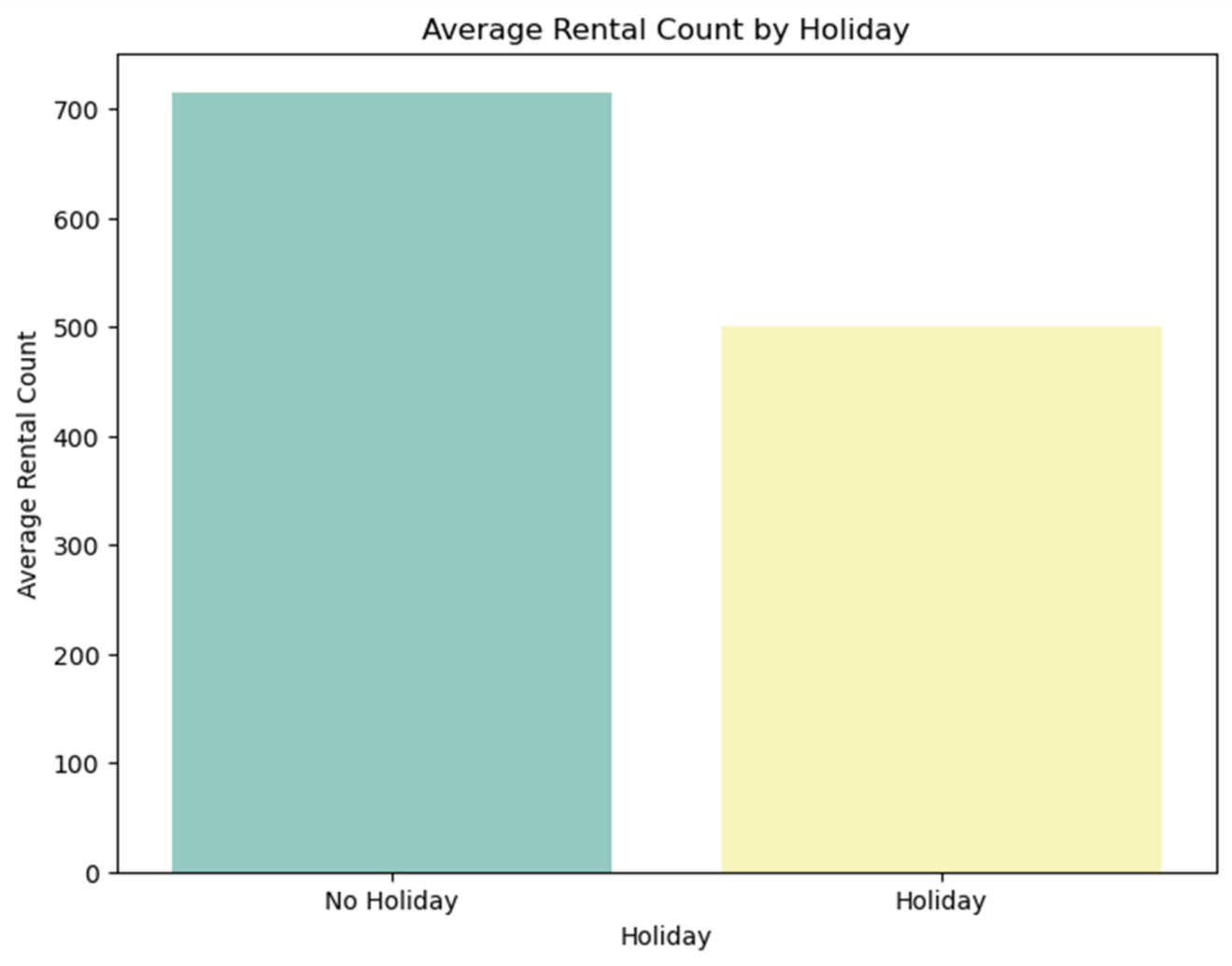
Data Visualization



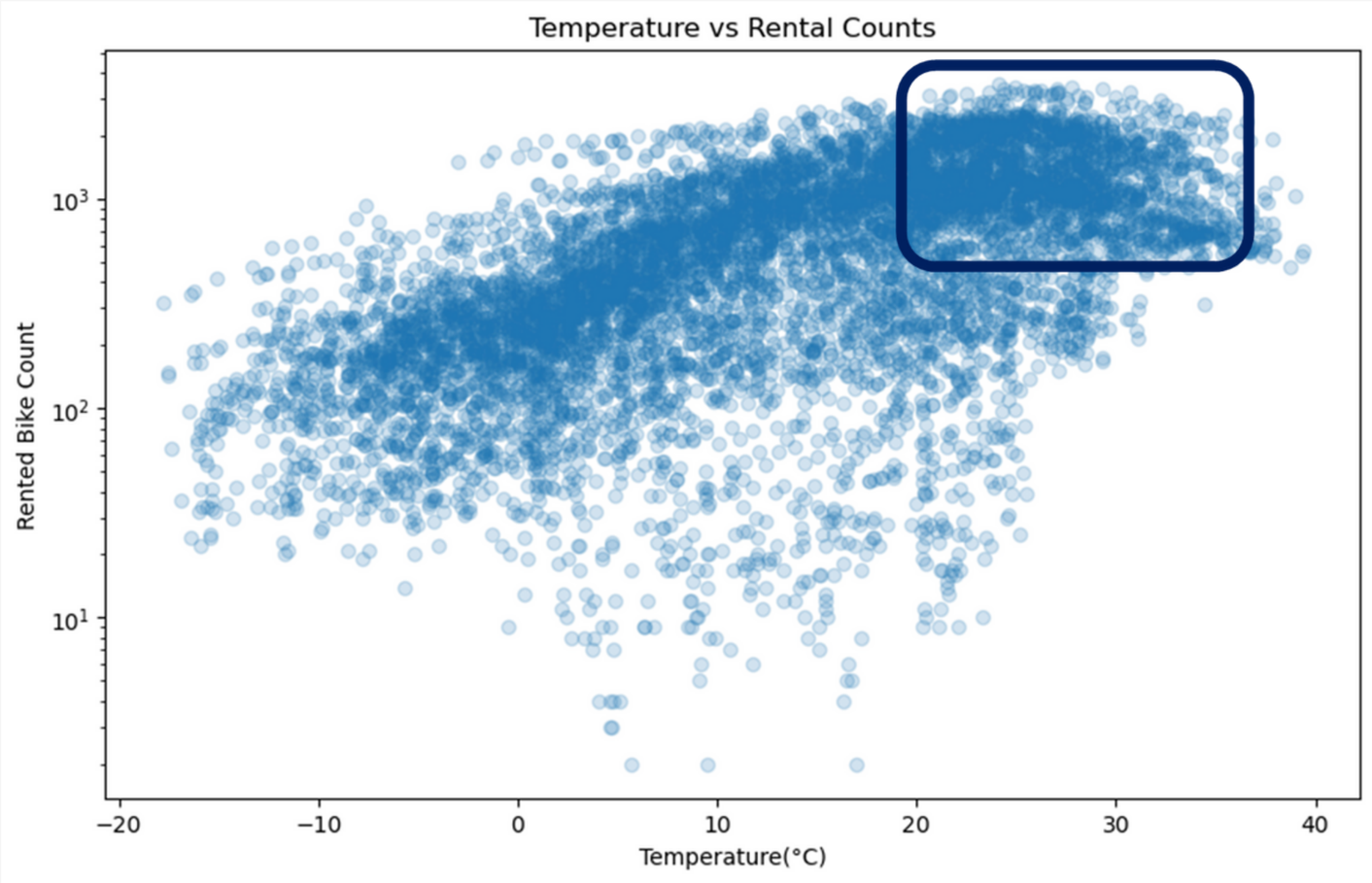
Checked for reduced volume in winter



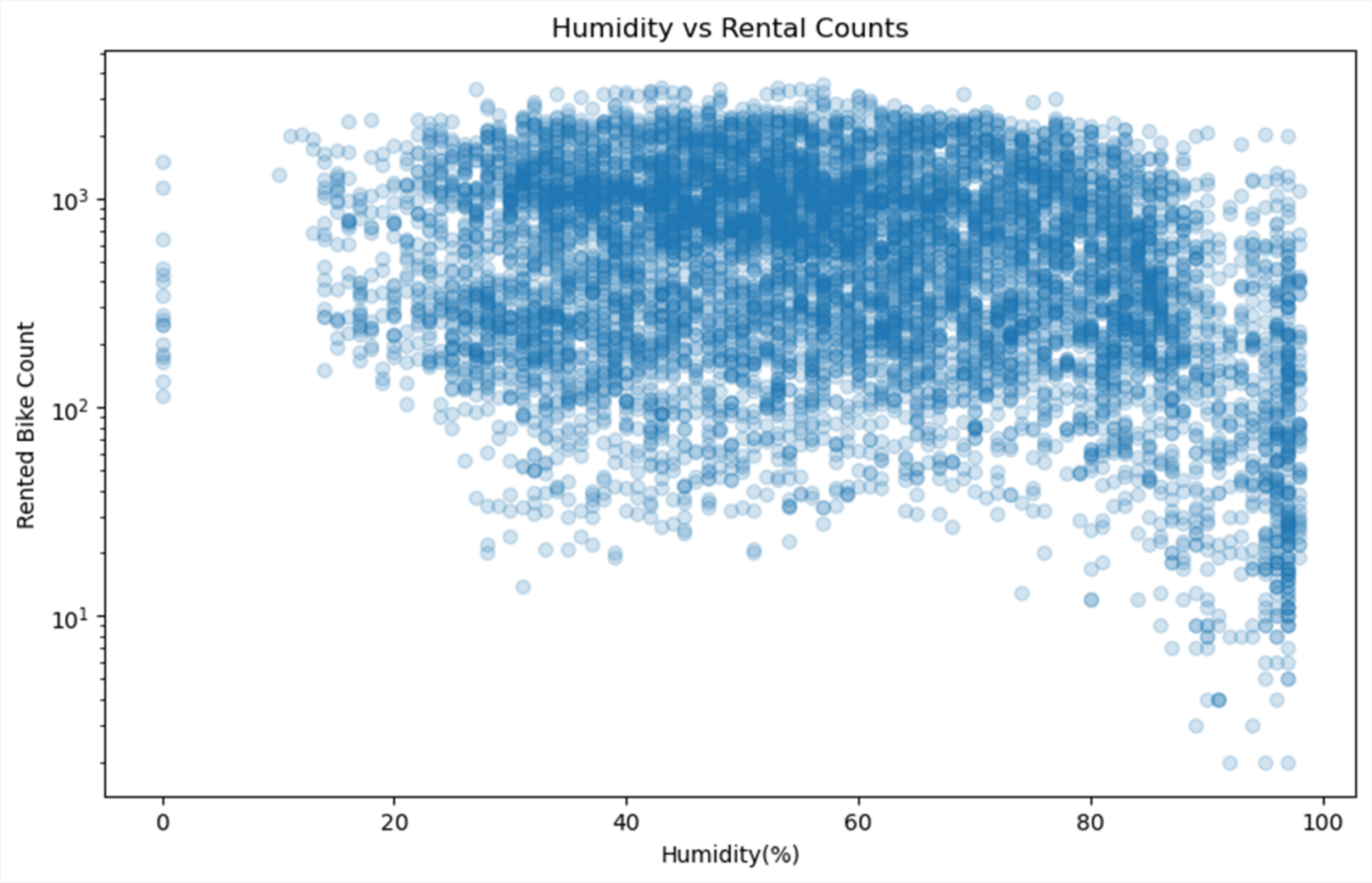
Checked for increased volume during rush hour



Checked for reduced volume on holidays

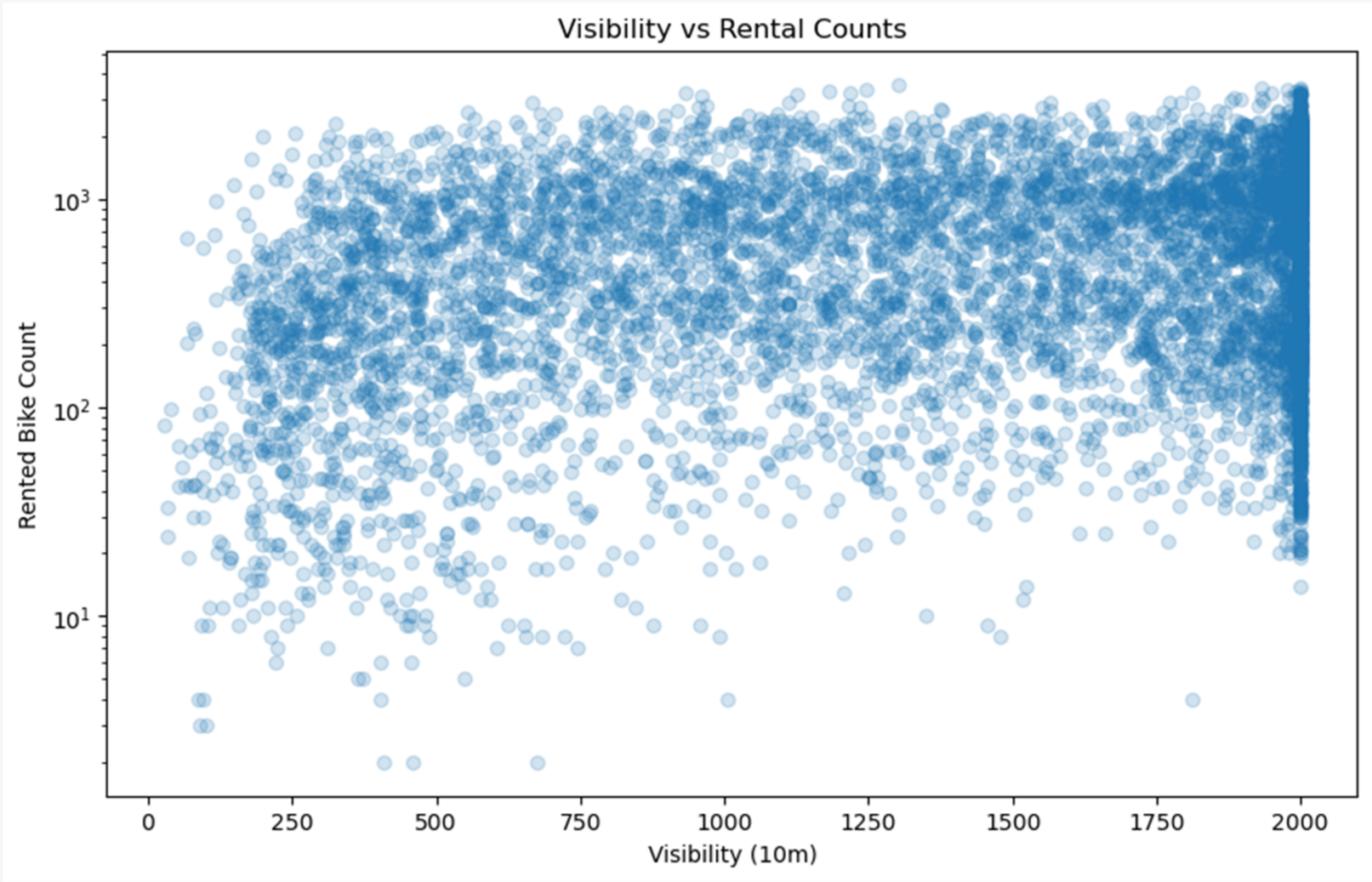


Checked for increased volume in mild weather



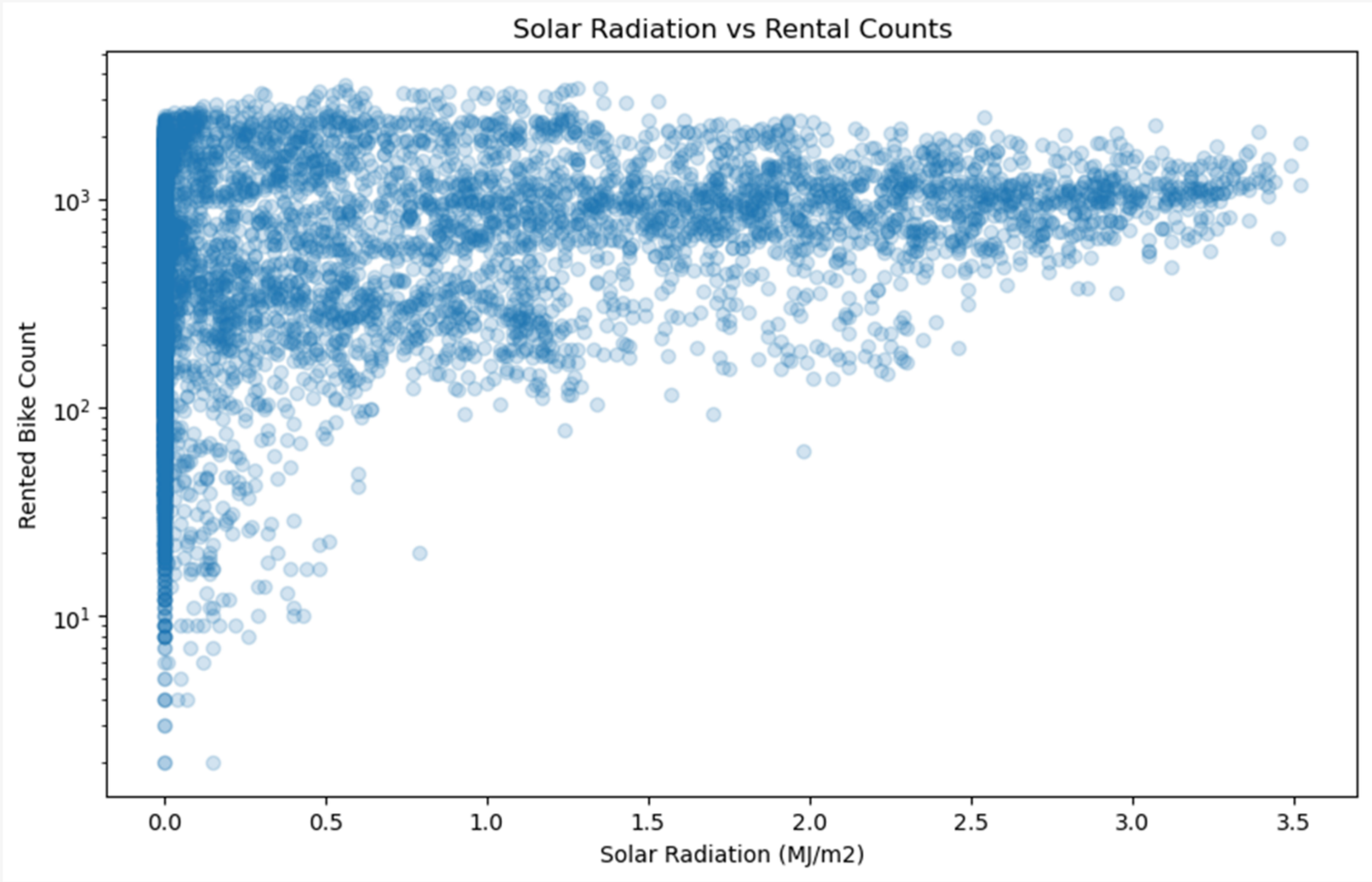
No significant results

-> decide to use as a control variable



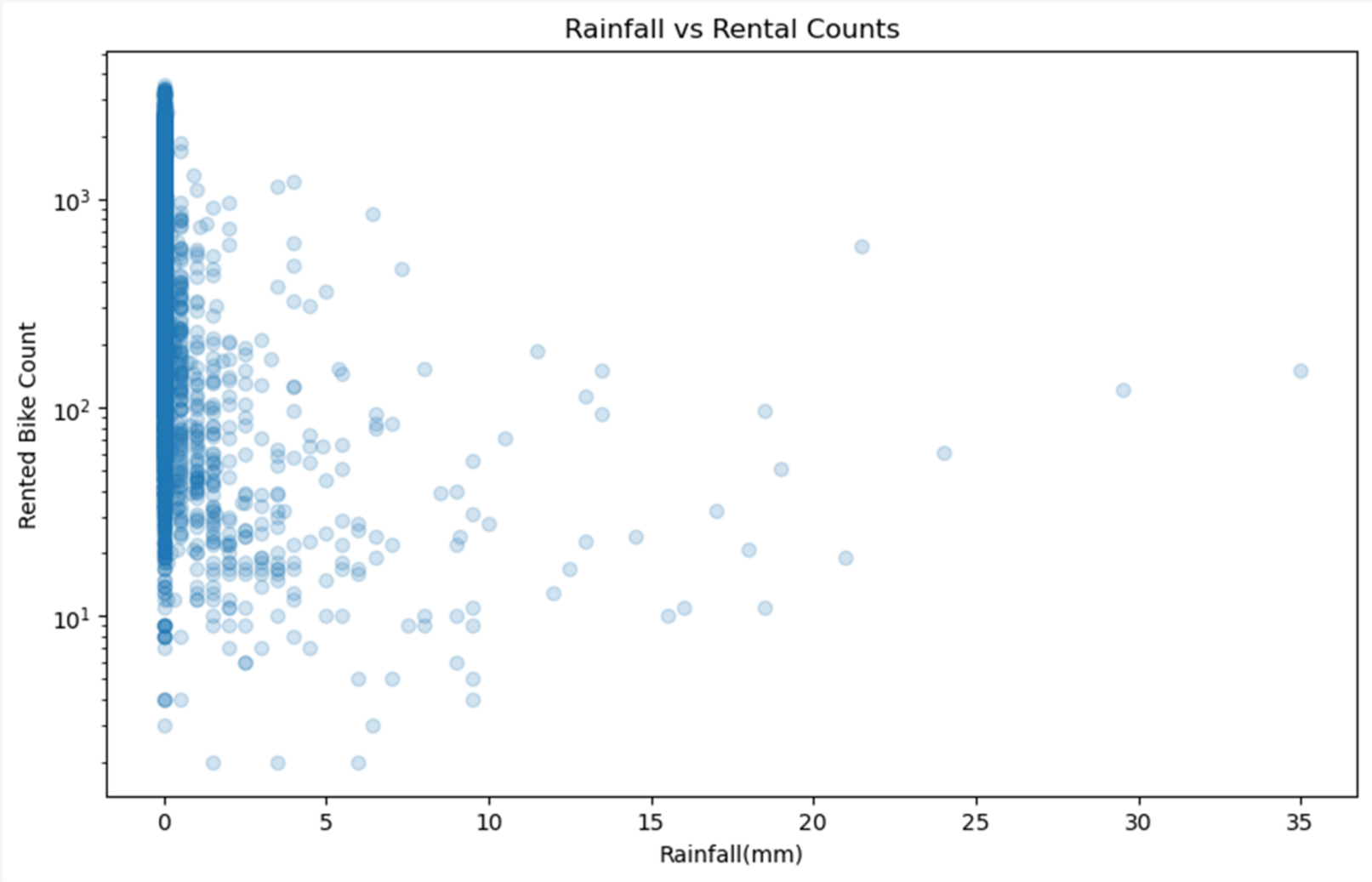
No significant results

-> decide to use as a control variable



No significant results

-> decide to use as a control variable



No significant results

-> decide to use as a control variable

Multiple Regression Analysis Model

• Multiple Regression Model

- Target Variable: : Bike Rental Counts
- Independent Variables: Temperature(°C), Holiday
- Control Variables: Other variables

  => solve omitted variable bias

• Interpretation of Multiple Regression Model

Coefficient

1. Temperature(°C): 31.24922

For every 1°C increase, the number of rental bicycles increased by about 31.

2. Holiday: -228.1334

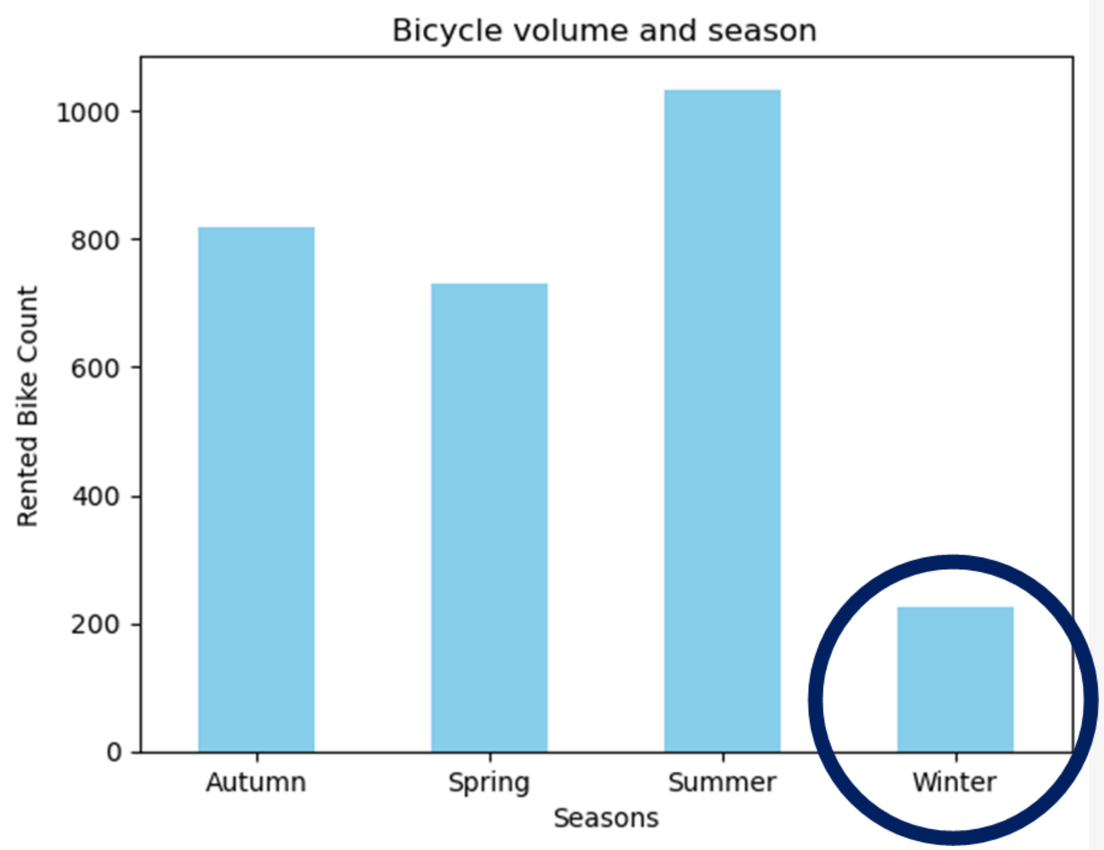
The number of bicycles rented on holidays decreased by about 228.

3. Temp_Holiday_Interact: 6.5325

= Interaction of Temperature(°C) and Holiday

An additional 6.5 bicycles increased as the temperature increased during holidays.

Setting Strategies based on Data Analysis



1. Sharp decline in rental volume in winter

It was confirmed that the bicycle capacity decreased significantly in winter.
Based on this, we thought about way to maintain and increase bicycle capacity even in winter.

Strategy to Complement Winter Bike Rental Count Reduction : Using Bicycle Handle Gloves

Using Bicycle Handle Gloves

- **Necessity**

- In fact, most sharing bikes are not equipped with gloves

- The capacity will increase if you can ride your bike warmly

- **Expectation Effectiveness**

- Providing customers with a warm and comfortable rental experience

- Maintaining or increasing volume in winter

- Improving customer loyalty and satisfaction

2. Interaction of Temperature(°C) and Holiday

It was confirmed that there were additional 6.5 bicycles increased as the temperature increased during holidays.
How can we use this for marketing strategies?

OLS Regression Results

Dep. Variable:	Rented Bike Count	R-squared:	0.474
Model:	OLS	Adj. R-squared:	0.473
Method:	Least Squares	F-statistic:	788.4
Date:	Thu, 07 Dec 2023	Prob (F-statistic):	0.00
Time:	18:40:34	Log-Likelihood:	-66286.
No. Observations:	8760	AIC:	1.326e+05
Df Residuals:	8749	BIC:	1.327e+05
Df Model:	10		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	441.5769	36.681	12.038	0.000	369.674	513.480
Hour	27.2082	0.785	34.645	0.000	25.669	28.748
Temperature(° C)	31.2492	0.541	57.767	0.000	30.189	32.310
Humidity(%)	-7.5058	0.391	-19.193	0.000	-8.272	-6.739
Wind speed (m/s)	8.0783	5.457	1.480	0.139	-2.620	18.776
Visibility (10m)	0.0230	0.010	2.246	0.025	0.003	0.043
Solar Radiation (MJ/m2)	-82.4329	7.988	-10.320	0.000	-98.091	-66.775
Rainfall(mm)	-59.3831	4.587	-12.946	0.000	-68.374	-50.392
Snowfall (cm)	17.6737	11.962	1.477	0.140	-5.775	41.122
Holiday	-228.1334	31.309	-7.286	0.000	-289.507	-166.760
Temp_Holiday_Interact	6.5325	2.087	3.130	0.002	2.441	10.624

Omnibus:	955.919	Durbin-Watson:	0.439
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1842.139
Skew:	0.715	Prob(JB):	0.00
Kurtosis:	4.733	Cond. No.	1.15e+04

Strategy for holidays

: Offering promotions during holidays

1) Placing additional bikes

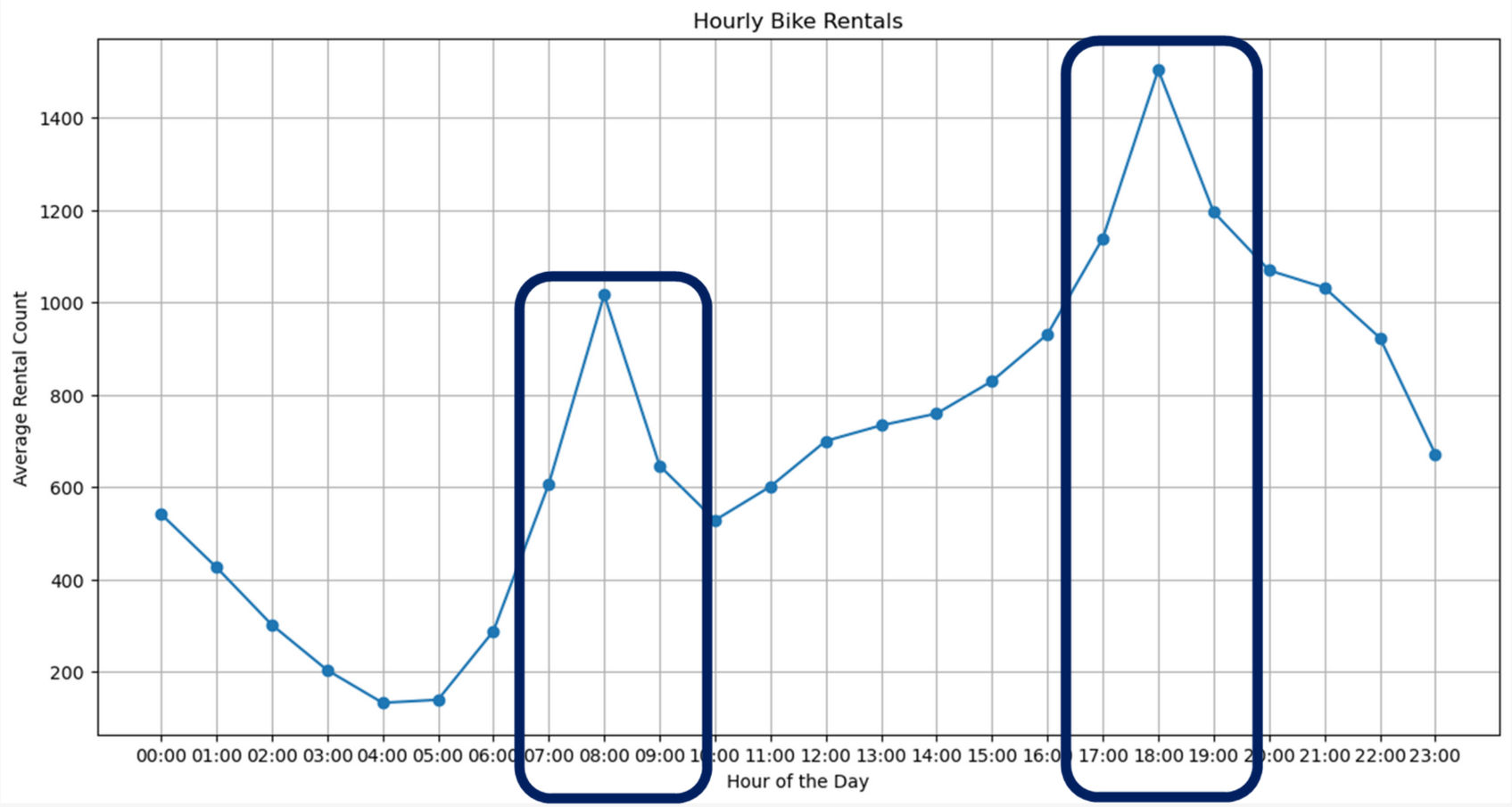
- preparing for the high temperature weather that is expected to increase demand for rentals on public holidays

2) Special events

- Conduct a challenge to post images of using a bicycle with family, friends, and lovers on social media

3.Increased usage during rush hour

It was confirmed that the bicycle rental counts increased significantly between 7 a.m. and 9 p.m. and 5 p.m. during rush hour.



- **Strategy for increased usage during rush hour**

1)Marketing and promotion targeting

- Offering special promotions (ex. Discount)

2) Improving and expanding services

- Increase the number of bicycles during rush hour

3) Promoting cooperation with local enterprises or cities

- Building partnerships through a variety of partnership businesses

Q&A

THANK YOU.