Pedestrians on the Brooklyn Bridge

By Wafa Berri, Cierra Blackett, Elhadji Keita, Karina Rose Sanchez, Arturo Monroy Ortiz

Our Research Question:
How do temporal factors such as time of day and day of the week, and environmental factors such as weather conditions, temperature, and precipitation, affect hourly pedestrian counts on the Brooklyn Bridge?

Our Dataset

- Title: Brooklyn Bridge Automated Pedestrian Counts Demonstration Project
- Publicly available through NYC Open Data
- Data provided by the New York City Department of Transportation (DOT)
  - “DOT is testing automated technology to count pedestrians. The counter is located on the Manhattan approach of the Brooklyn Bridge.”
- This makes our project an **observational study**.
- Our goal with this project is to make **predictions**.

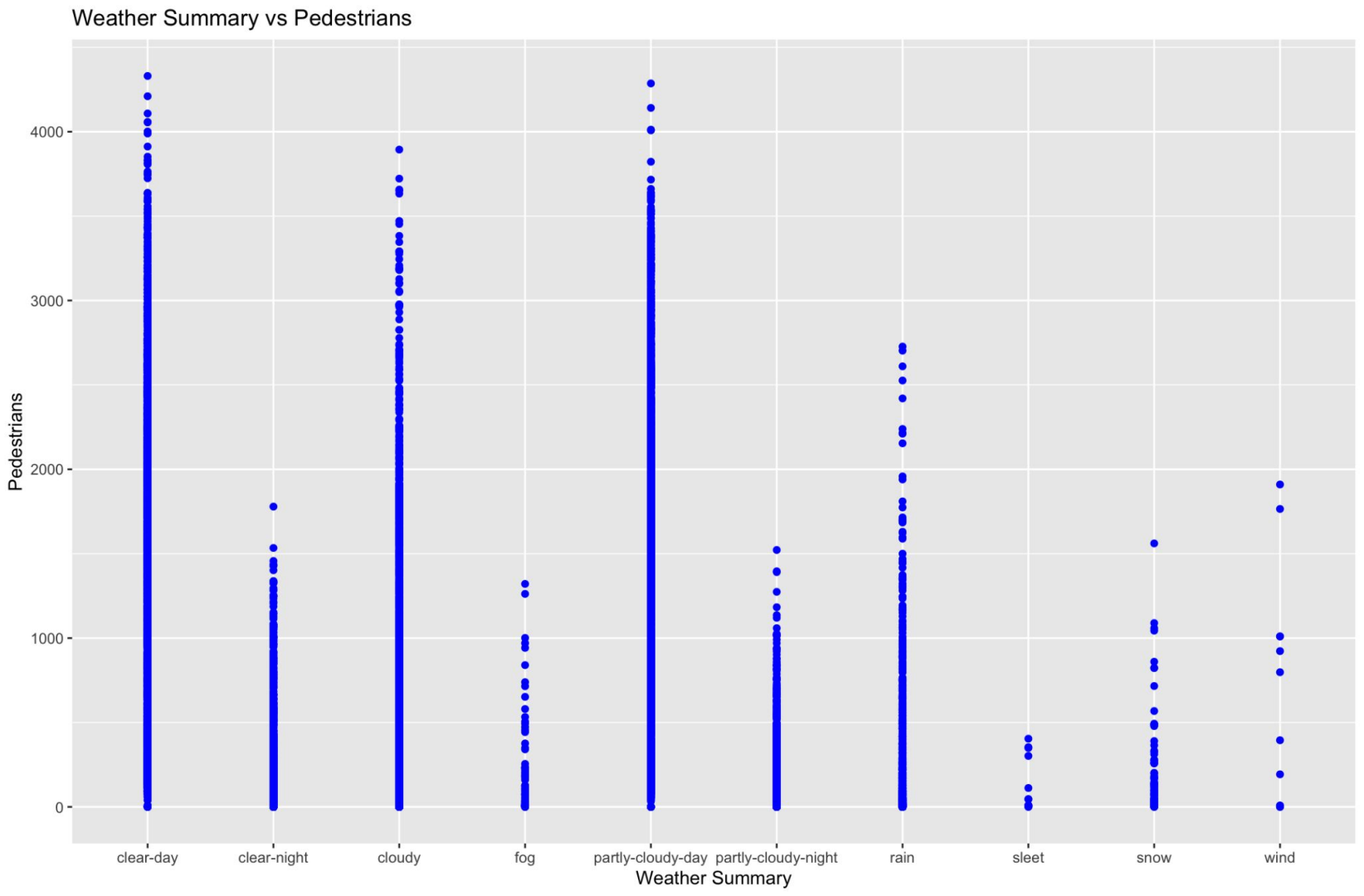
Our Features:

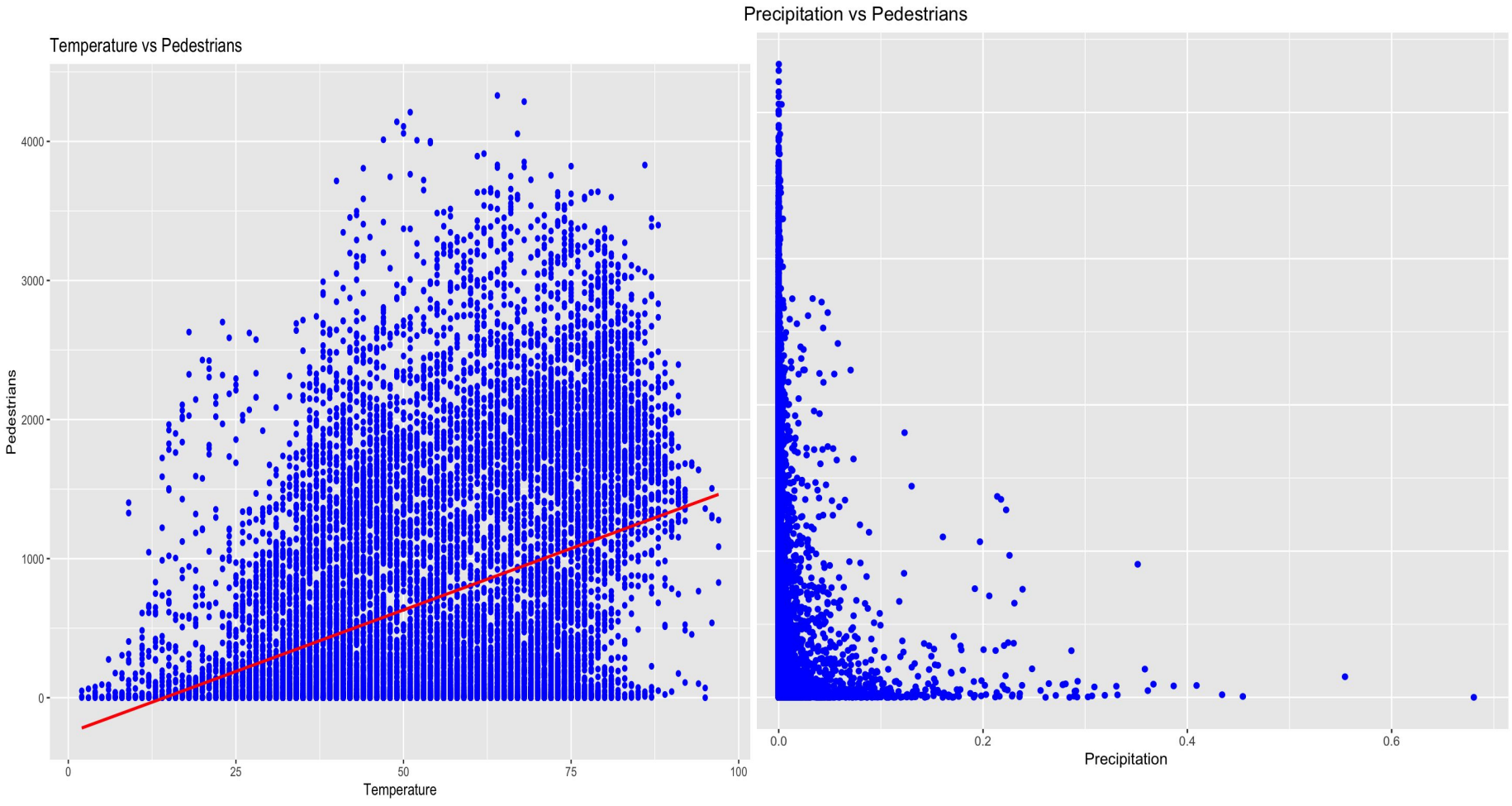
- **Hour_beginning -** Date and time of hourly count (categorical)
- **Weather_summary -** Overall daily weather (cloudy, clear, rain, etc.) (categorical)
- **Temperature -** Hourly temperature, in Fahrenheit degrees (numerical)
- **Precipitation -** Hourly precipitation, in inches (numerical)
- **Events -** holidays (categorical, 1 for event 0 for no event)

Target Variable:

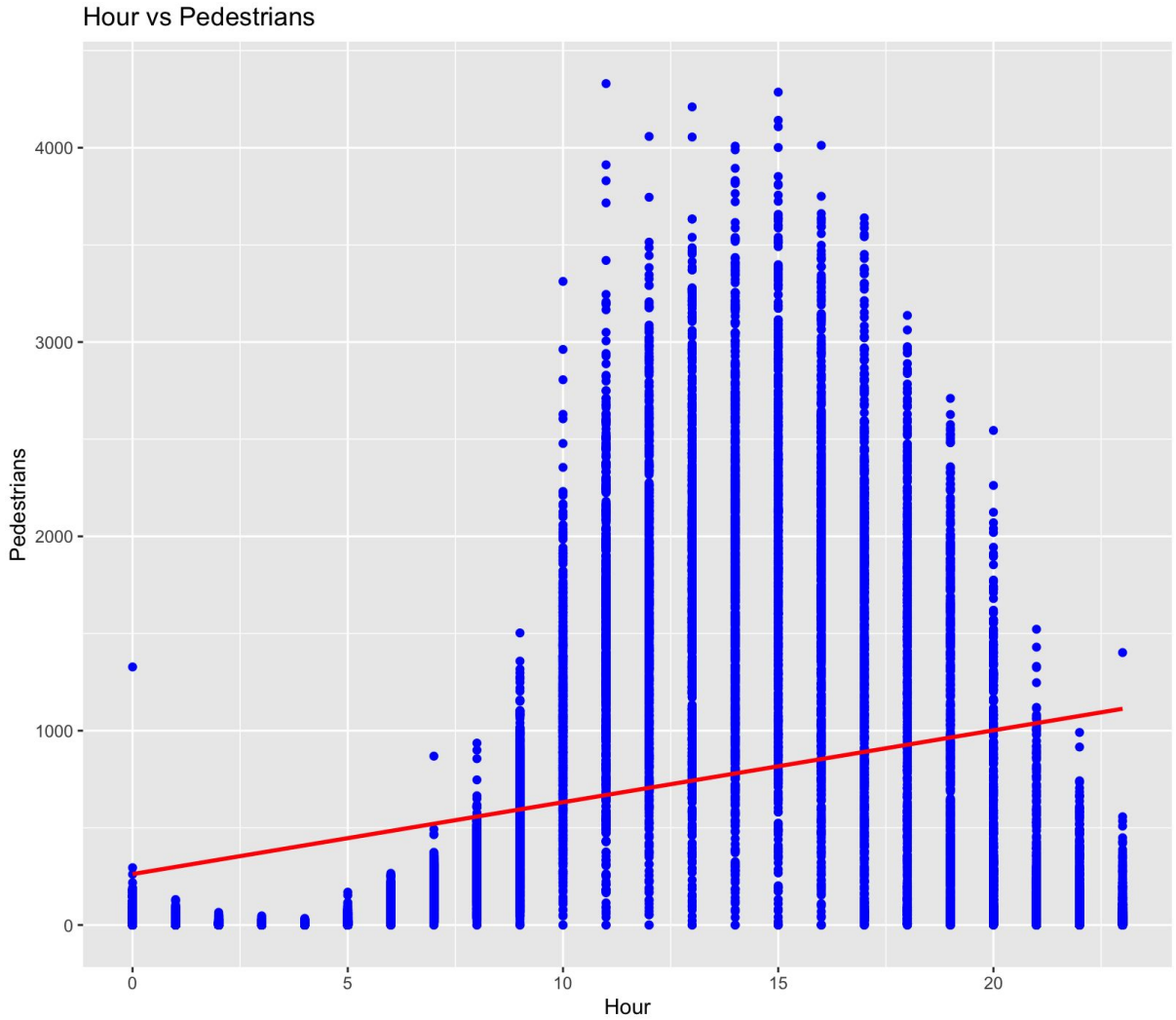
- **Pedestrians -** Total count (sum of directions to Brooklyn and to Manhattan) (numerical)

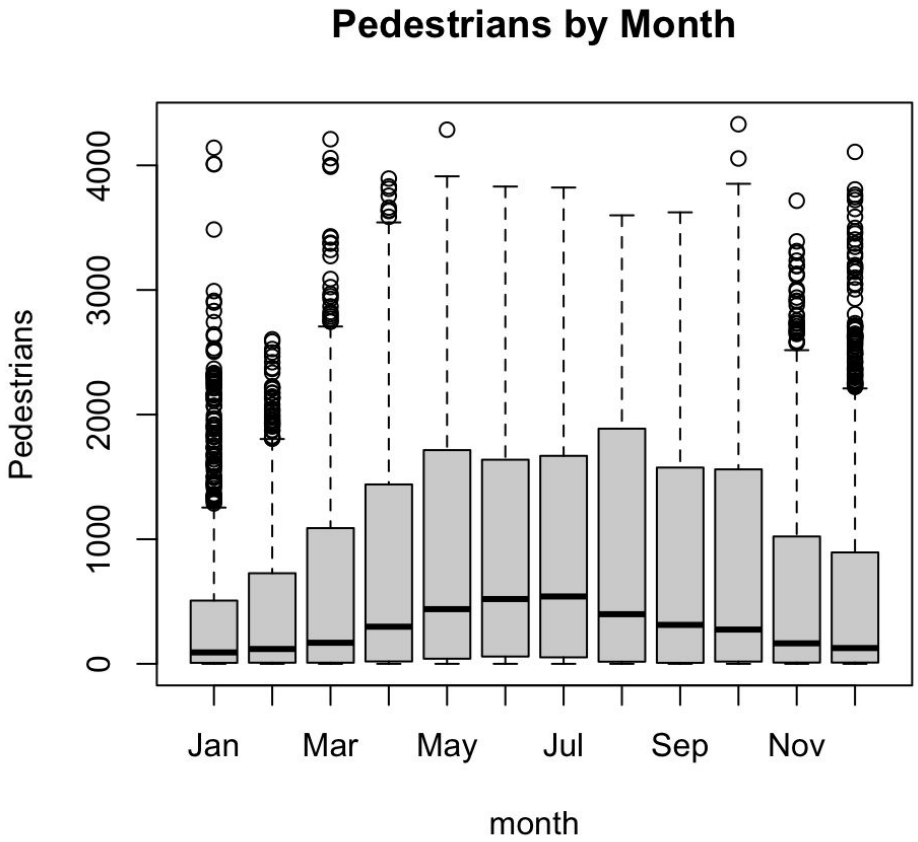

Plots

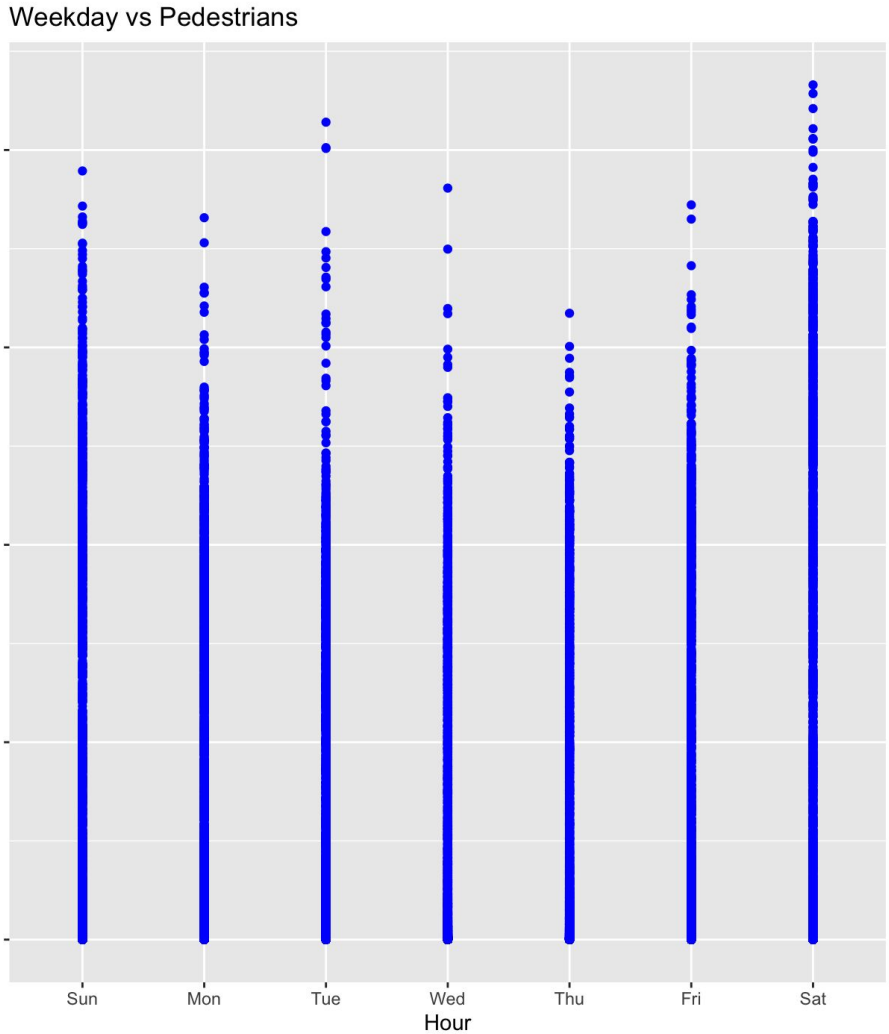

Weather Summary vs Pedestrians
Pedestrians
4000
3000
2000
1000
0
clear-day
clear-night
cloudy
fog
partly-cloudy-day
partly-cloudy-night
rain
sleet
snow
wind
Weather Summary


Temperature vs Pedestrians
Pedestrians
4000
3000
2000
1000
0
0
25
50
75
100
Temperature
Precipitation vs Pedestrians
0.0
0.2
0.4
0.6
Precipitation

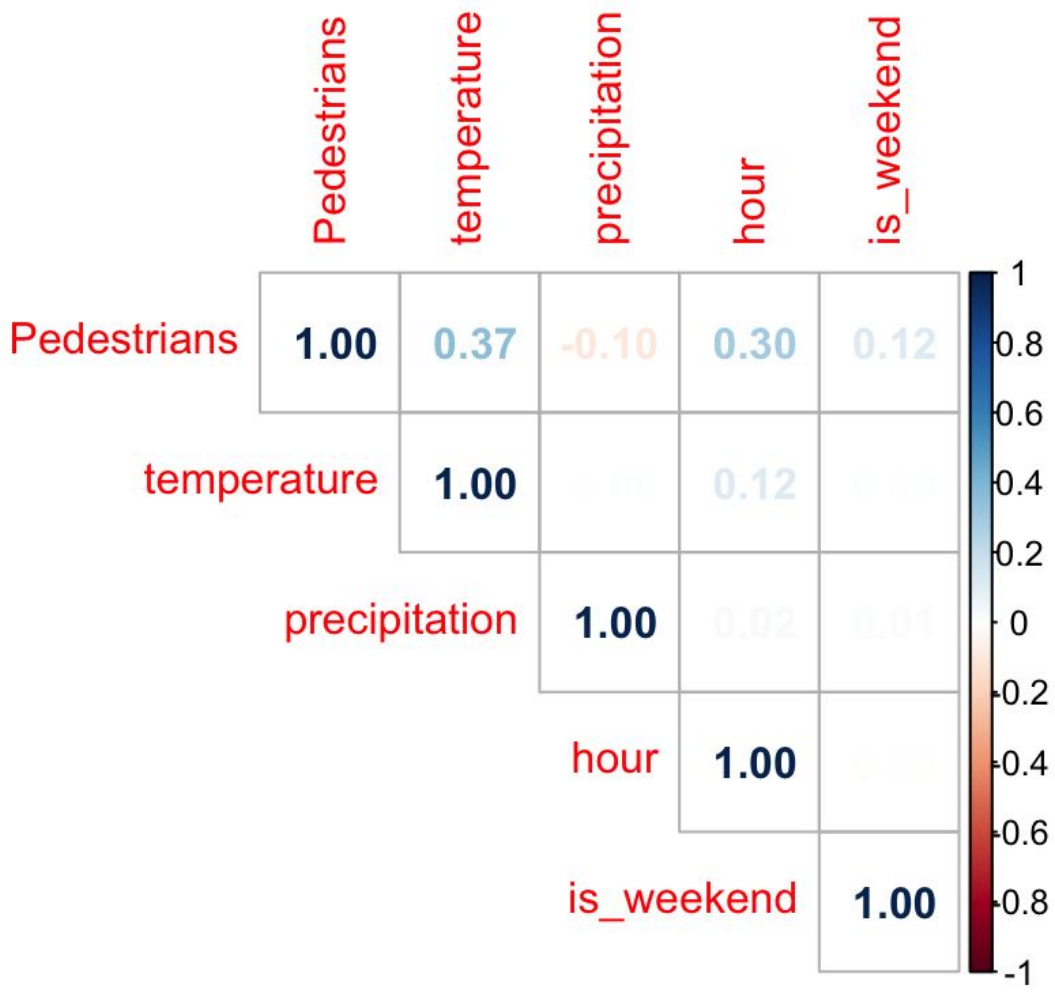
*The trend is not linear; it appears to be cyclic.




Pedestrians by Month
Pedestrians
0
1000
2000
3000
4000
Jan
Mar
May
Jul
Sep
Nov
month


Weekday vs Pedestrians
Sun
Mon
Tue
Wed
Thu
Fri
Sat
Hour

Correlation plot of our variables:



Our Base Model and Exploratory Analysis:

```
Residuals:
     Min       1Q   Median       3Q      Max
-2022.59  -345.56   -20.04   281.32  2842.57

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                       -33.5704    31.7316  -1.058 0.290094
weather_summaryclear-night      -1093.7141    13.8901 -78.740  < 2e-16 ***
weather_summarycloudy            -751.5066    15.2471 -49.288  < 2e-16 ***
weather_summaryfog               -955.5921    64.4766 -14.821  < 2e-16 ***
weather_summarypartly-cloudy-day  -17.7051    14.1781  -1.249 0.211772
weather_summarypartly-cloudy-night -1186.7555    15.0725 -78.737  < 2e-16 ***
weather_summaryrain              -1012.4223    26.4915 -38.217  < 2e-16 ***
weather_summarysleet              -975.7120   149.4861  -6.527 6.91e-11 ***
weather_summarysnow               -809.0133    59.3166 -13.639  < 2e-16 ***
weather_summarywind               -660.0741   160.3380  -4.117 3.86e-05 ***
temperature                         20.4869     0.5411  37.864  < 2e-16 ***
precipitation                    -1172.2879   248.1974  -4.723 2.34e-06 ***
has_event                           73.1385    17.6464   4.145 3.42e-05 ***
hour                                20.4506     0.6550  31.223  < 2e-16 ***
weekday.L                          109.8148    11.7584   9.339  < 2e-16 ***
weekday.Q                          254.5987    11.6476  21.858  < 2e-16 ***
weekday.C                           30.4630    11.6014   2.626 0.008653 **
weekday^4                           41.9739    11.5993   3.619 0.000297 ***
weekday^5                           -2.8478    11.5870  -0.246 0.805855
weekday^6                           -1.4143    11.5909  -0.122 0.902890
month.L                            -88.9634    17.4173  -5.108 3.30e-07 ***
month.Q                            673.1774    30.1923  22.296  < 2e-16 ***
month.C                            283.2989    17.6387  16.061  < 2e-16 ***
month^4                           -194.0804    15.7950 -12.287  < 2e-16 ***
month^5                           -124.4125    16.0759  -7.739 1.06e-14 ***
month^6                            166.1523    15.7419  10.555  < 2e-16 ***
month^7                            -39.2870    15.6803  -2.505 0.012238 *
month^8                            -19.5111    15.5614  -1.254 0.209927
month^9                             45.3662    16.3068   2.782 0.005408 **
month^10                            54.5145    17.2081   3.168 0.001538 **
month^11                            58.6434    15.8413   3.702 0.000215 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 553.7 on 16010 degrees of freedom
Multiple R-squared:  0.5886,    Adjusted R-squared:  0.5878
F-statistic: 763.5 on 30 and 16010 DF,  p-value: < 2.2e-16
```

Our Base Model:

Our multivariate linear regression model analyzed how weather, time variables, and special events affect the number of pedestrians on the Brooklyn Bridge. The **adjusted R-squared value of 58.8%** means that the model explains a little more than half of the changes in pedestrian counts using factors like weather, time of day, and seasonal patterns. The **remaining 41% of the variation is likely due to factors we did not include**, such as construction, transit disruptions, or random day-to-day fluctuations. Our base model also **does not include higher order terms and interaction terms**. In terms of accuracy, **the model's predictions are typically within about plus or minus 550 pedestrians per hour, which reflects the natural variability in real-world pedestrian traffic**.

Exploratory Analysis:

When we look at each factor **while keeping all other variables the same**, weather has a clear impact on pedestrian activity. Rain is associated with about **1,013 fewer pedestrians**, snow with about **809 fewer**, and sleet with about **976 fewer pedestrians** on the Brooklyn Bridge. Temperature has the opposite effect, with pedestrian counts increasing by about **20 people per degree**, meaning a **10-degree warmer day could bring around 200 additional pedestrians**. Pedestrian traffic also increases by roughly **20 people per hour throughout the day**, and days with special events see an average increase of about **73 pedestrians**. Overall, these results show that pedestrian activity is strongly influenced by predictable weather and time-related patterns.

Implications and Limitations

The model shows that factors such as time of day, day of the week and environmental factors are important predictors of pedestrian traffic on the Brooklyn Bridge this implies that.

1. Peak Pedestrians hours can be anticipated which allows city agencies to manage congestion than reacting after overcrowding.
2. Maintenance and bridge operations can be schedule efficiently to during times where there less pedestrians this can be during the winter particularly in the night.
3. Staffing and public safety resources can be increased during high traffic periods especially on the weekends and good weather conditions which will help reduce safety risks and improve the pedestrian flow.

The model may take insight on environmental factors, time of day and the day of the week however it has limitations that need to be considered.

1. Analysis is observational so the relationships between time weather and the pedestrian counts represents correlations and dont alway imply causations.
2. Dataset covered a limited time periods which can reduce model ability to generalize findings across different years. It also doesn't account for long terms trends like the growing effects of climate change which will show the gradual shifts in commuter behavior.