# ISSS602_AY2025-26Aug_Assign1_ES

Creation Date: Thursday, September 18, 2025, 08:01:11 PM

Author: edward.lim.2025

**Executive Summary_v.1**
**Creation Date: Sunday, 14 Sep, 2025**
**Author: edward.lim.2025**

# ISSS602 Data Analytics Lab Assignment 1: Show Me the Numbers

---

## Objective:

To conduct a study to discover and determine factors affecting the demand of HappyRides' bikesharing business.

## Data:

The source of the dataset is unknown and is assumed to be in a single country with 4 seasons.

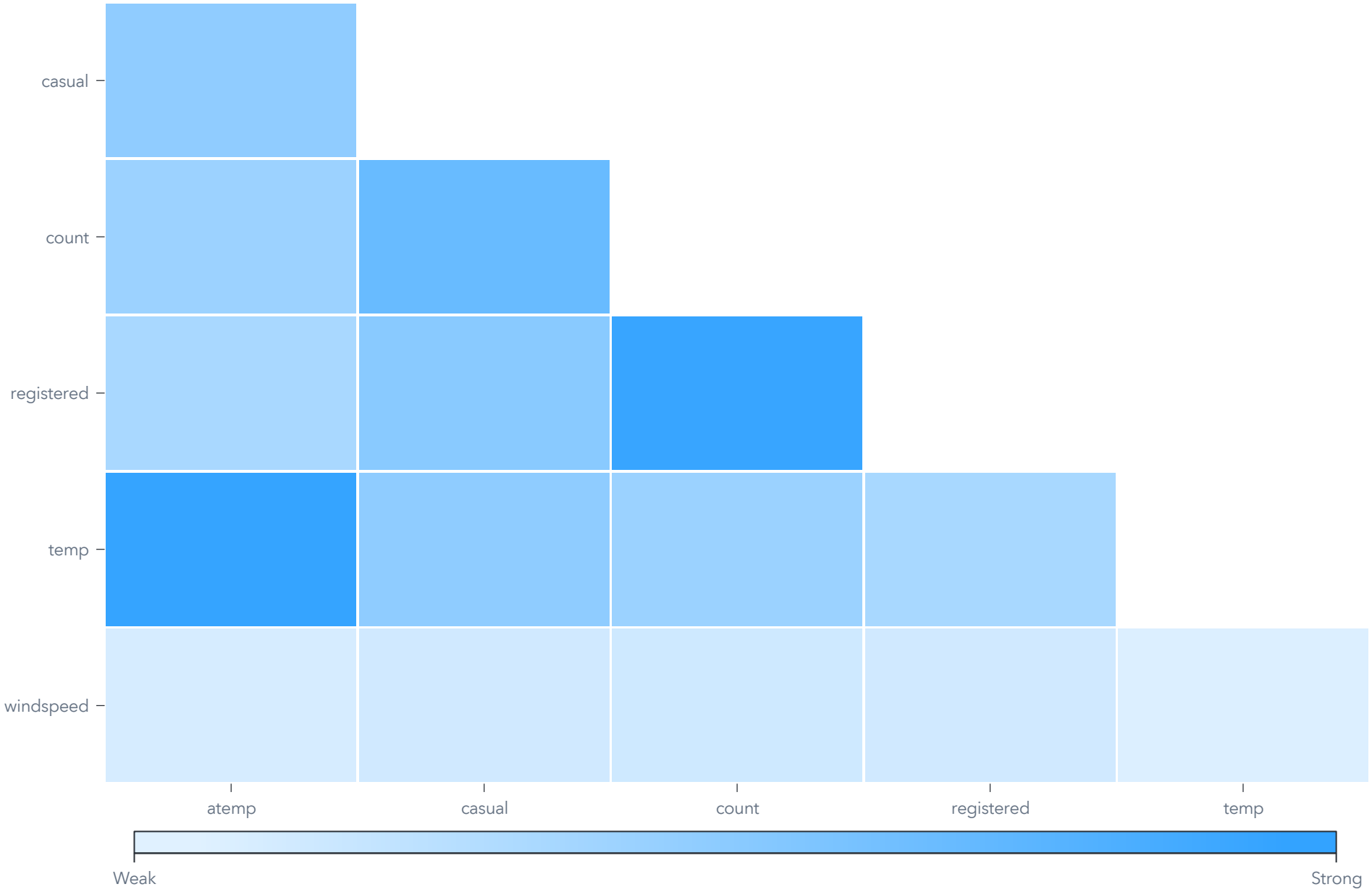
## Units of Measurements:

Windspeed is taking the unit of mph.
Humidity is taking the unit of RH. Temperate and Atemperature is taking the unit of Degree Celsius

## The Methodology:

Visualization techniques such as bar charts and box plots are employed.
Hypothesis testing is then conducted using the Chi-Square Test and Analysis of Variance (ANOVA)

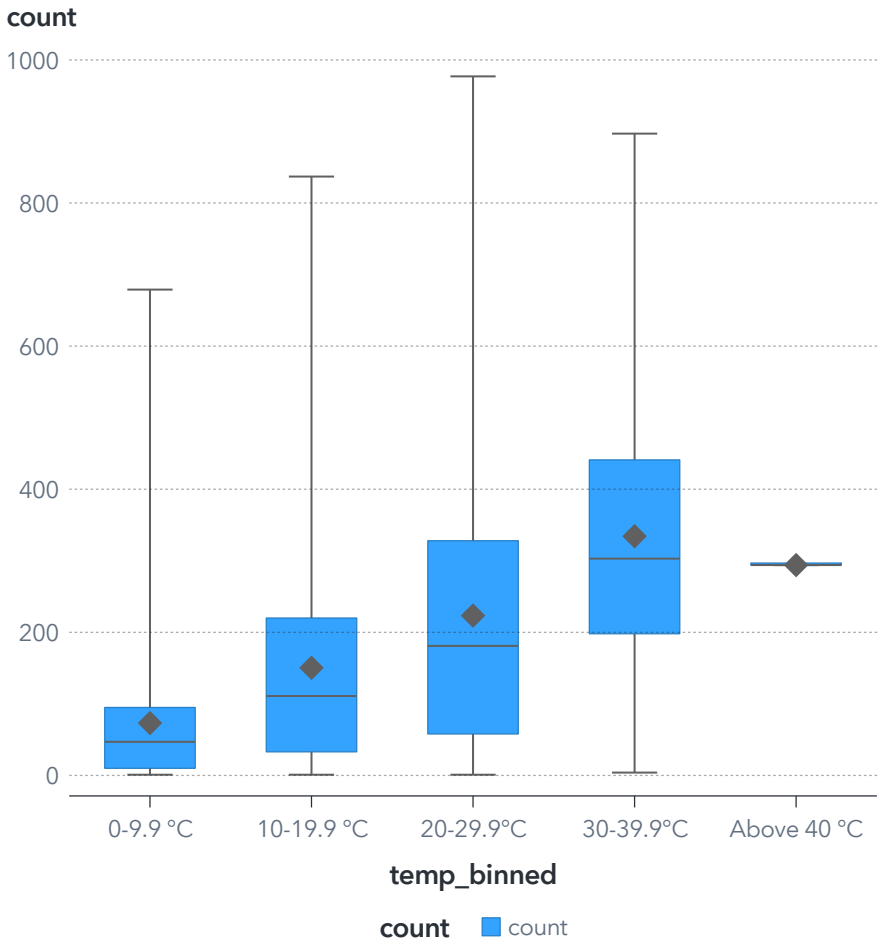
## Correlation of Selected Measures



## Insight 1: Temperature has an effect on Bike Usage

-As the temperature increases from 0 to 40 degrees Celsius, the mean of users using bikes consistently rises.
-It was observed that as the temperature increases the total user increase in the range from 0 to 39.9 degree celsius.

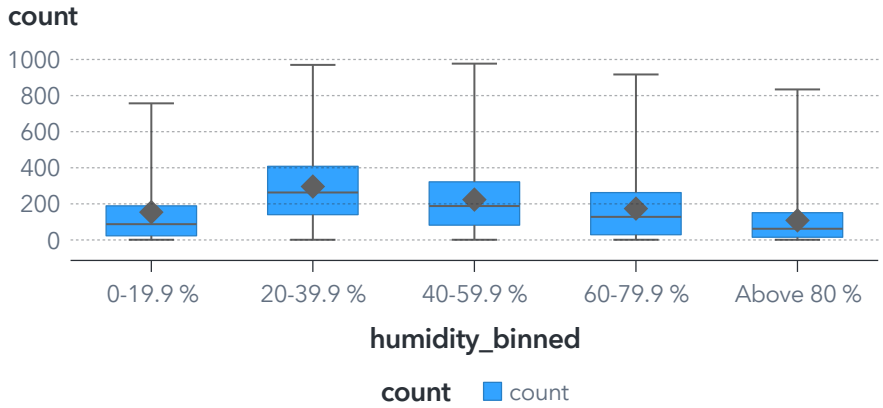
### Total Users by Temp



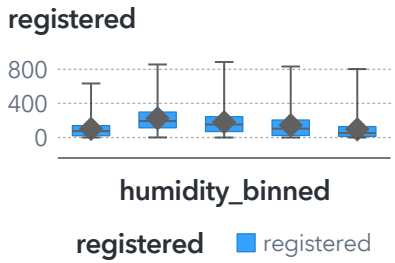

## Insight 2: Temperature has an effect on Bike Usage

-There is a relationship between humidity and bike usage, influencing both casual and registered riders.
-As humidity increases from 0 % to 40%, the mean of users also rises.
-The mean of users then drops when humidity levels climb above 60%.
-The is a high level of skewness for casual users under the 0-19.9% humidity category and median would be a better measure for this instance.

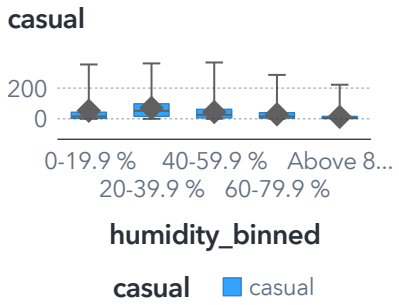
### Total Users by Humidity




### Registered Users by Humidity



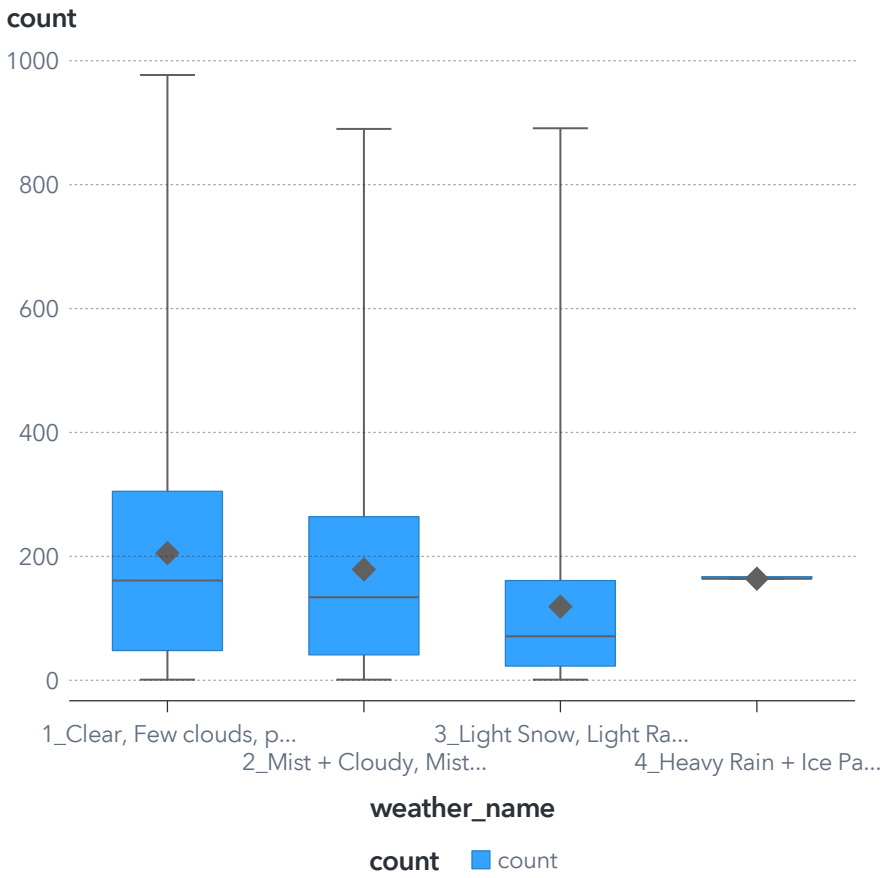

### Casual Users by Humidity



## Insight 3: Weather Type has an effect on Bike Usage

-Weather Type 1: Clear, Few clouds, partly cloudy has the highest median number of riders compared to other weather types.
-This trend should be viewed with caution, as the data for Weather Type 4 is limited (single row).
-Further data collection is needed to confirm the relationship between weather and bike usage, especially for less common weather types.

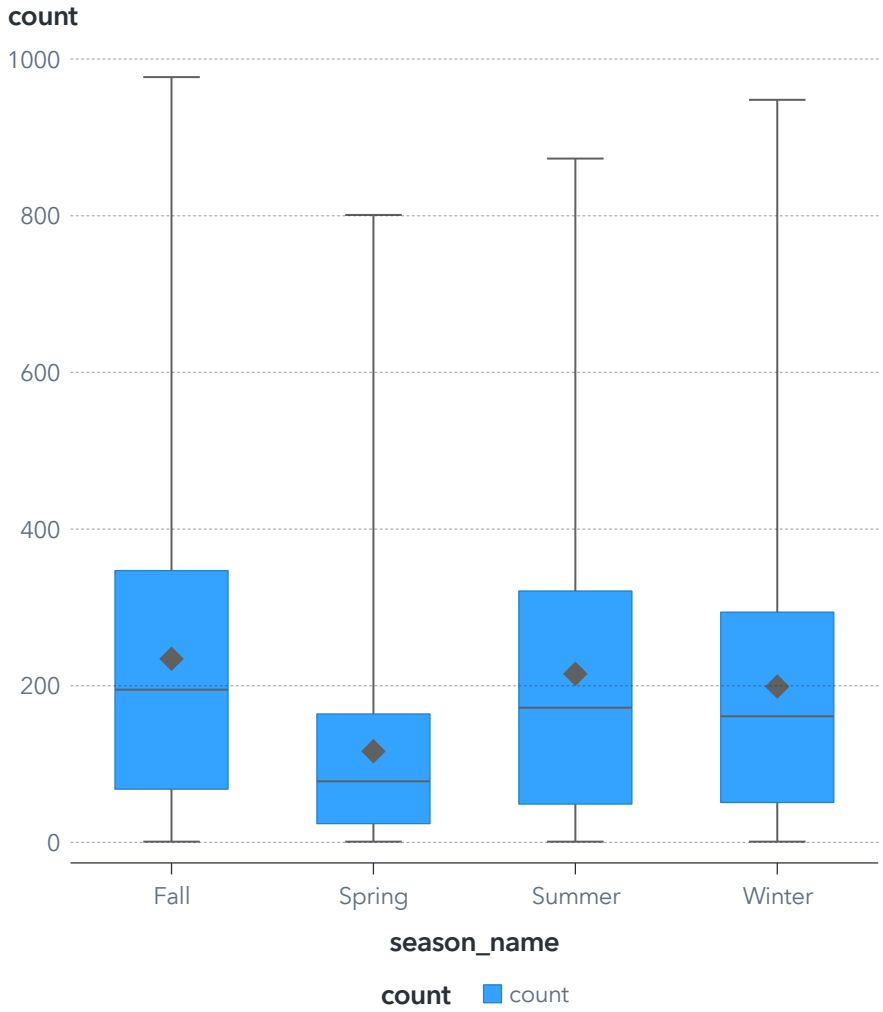
### Total Users by Weather Type




## Insight 4: Season has an effect on Bike Usage

-The mean are observed to vary across the different seasons.
-Spring is observed to have the lowest mean users as is lower than winter.

### Total Users by Season Type

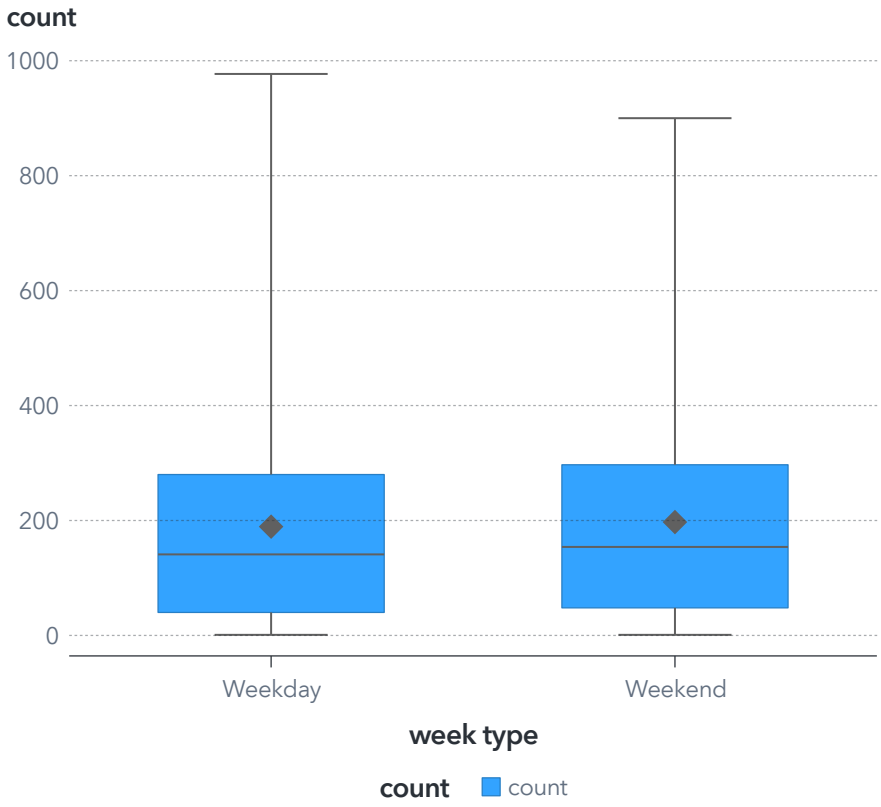


## Relationship between Casual user and Registered users

### Insight 5: Weekday/Weekday has NO significant effect on Bike Usage

-No statistically significant difference between bike user and Weekday/Weekday.
-The median and mean are slightly higher in weekend.

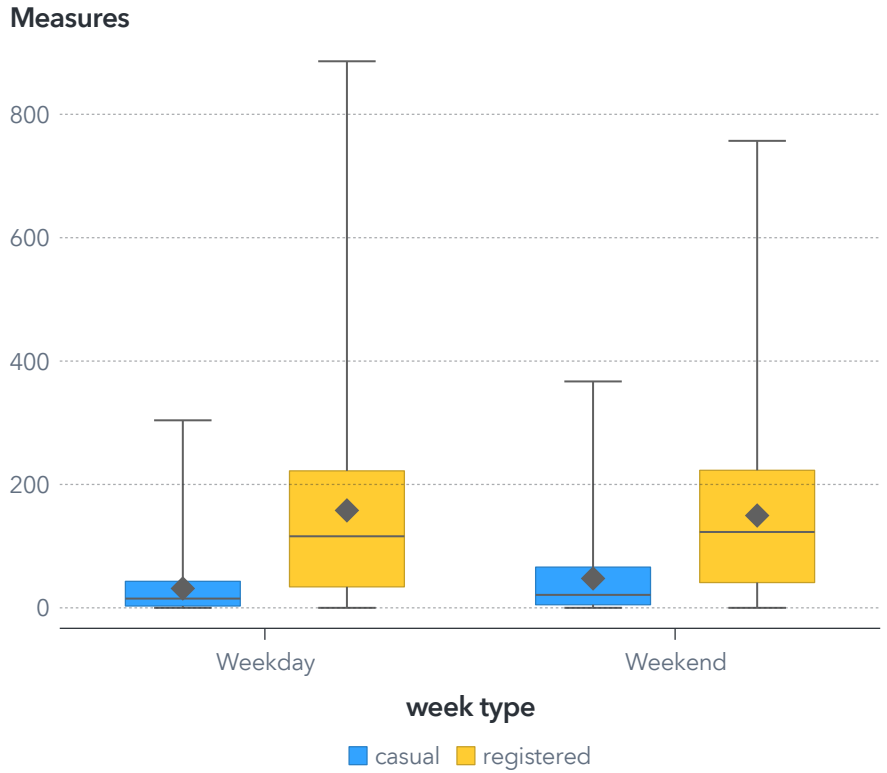
Total users by week type




### Insight 6: Weekday/Weekend have more effect on Casual Users than Registered Users

-Segregating the users by Registered and Casual, we observed a higher mean difference in the Casual Users.
-There are more Casual users on weekend then weekday where Registered users remain rather similar.

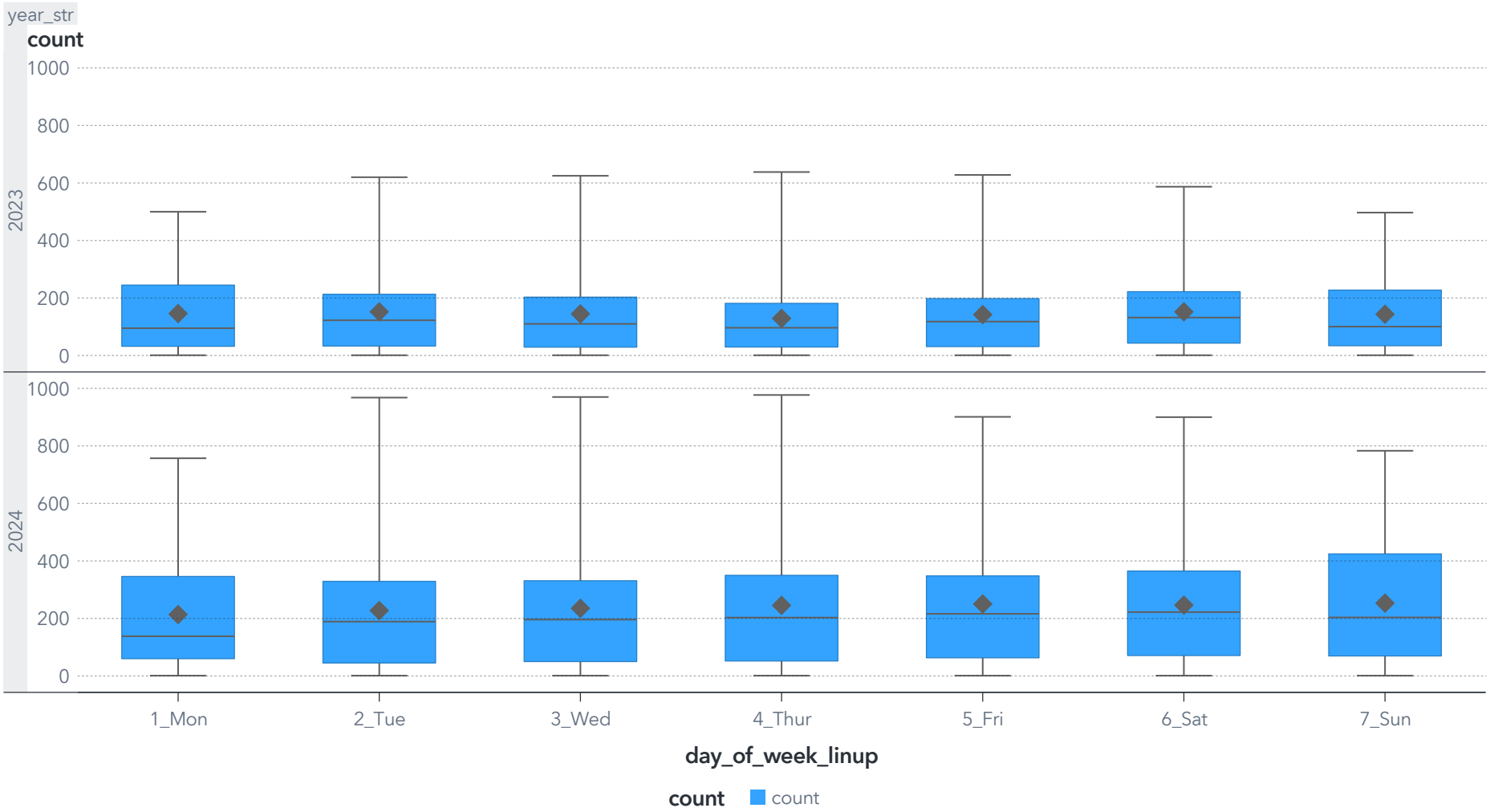
Casual, Registered Users by week type



## Insight 7: Day of week have NO significant effect on Bike Usage

-There is not much difference noted for user in the day of week.
-The mean remains rather consistent across the days in 2023 and 2024.

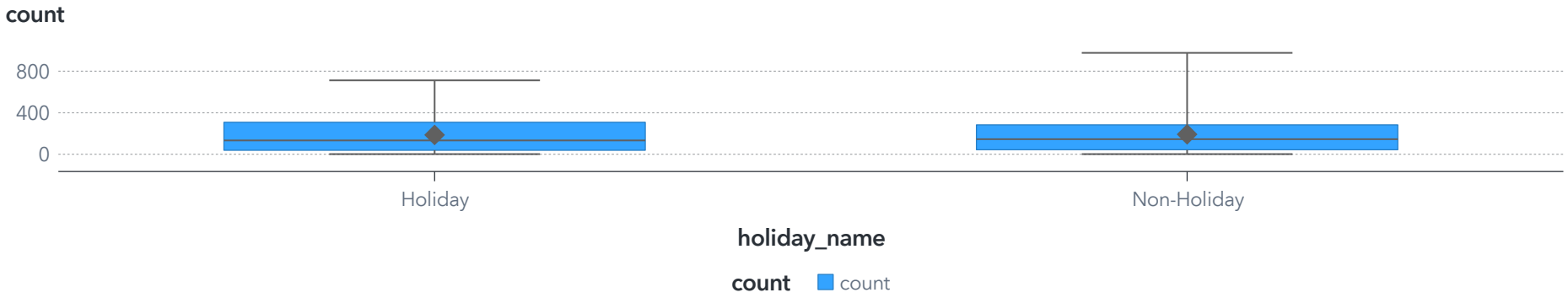
Total Users by day_of_week_name



## Insight 8: Holiday/ Non-Holiday has NO significant effect on Bike Usage

-The bike usage mean is rather similar between holiday and non - holiday period.

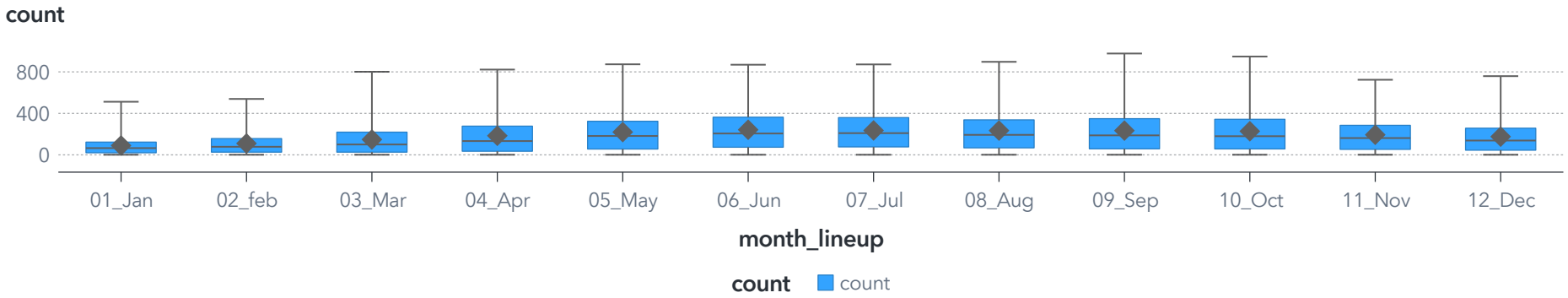
### Total Users by Holiday




## Insight 9: Month of the Year has effect on Bike Usage

-It is noted that the mean of users varies across the month of the year.
-The higher range of bike usage resides along Q2 and Q3.

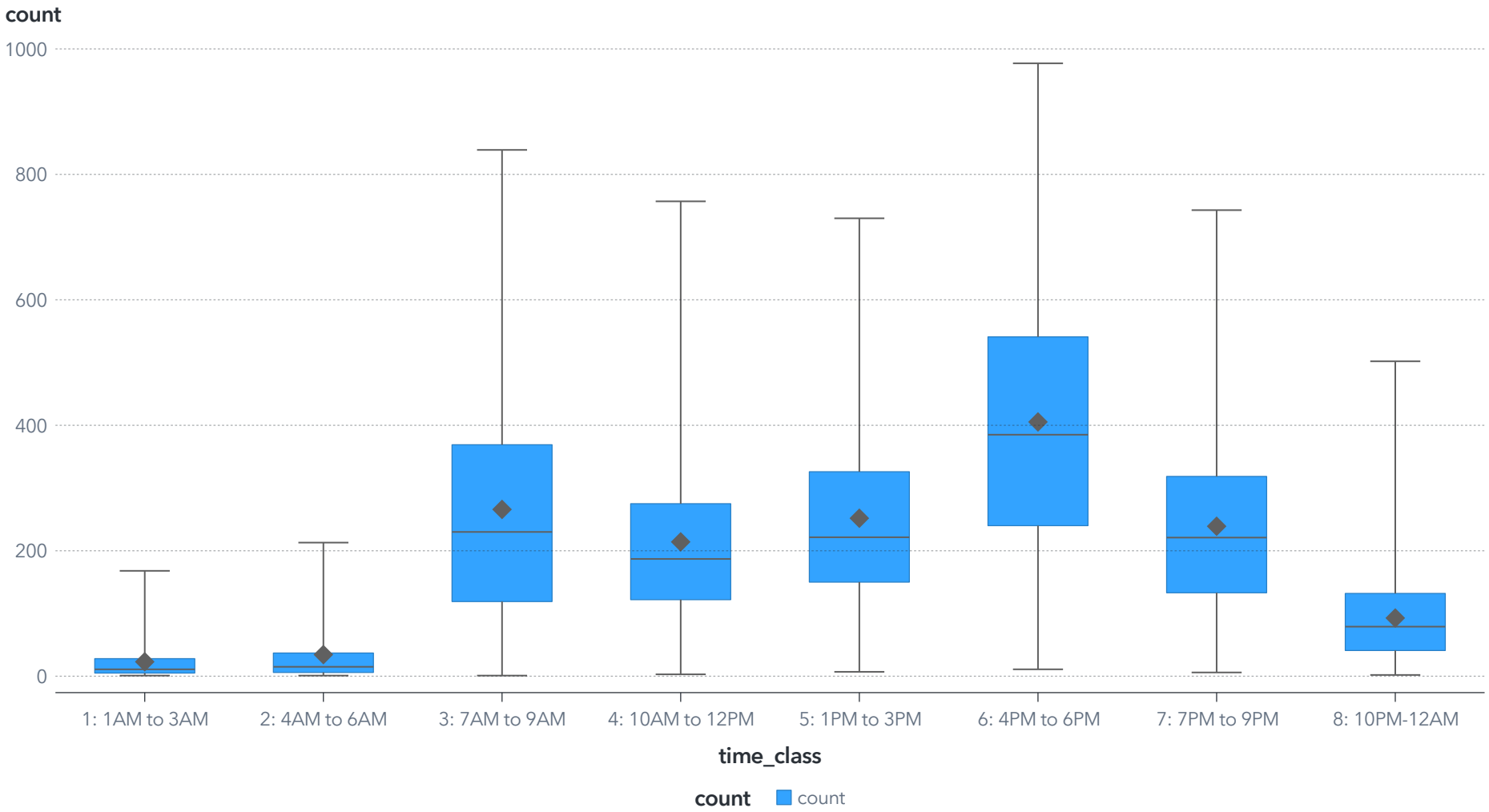
### Total Users by Month



## Insight 10: Time of the Day have effect on Bike Usage

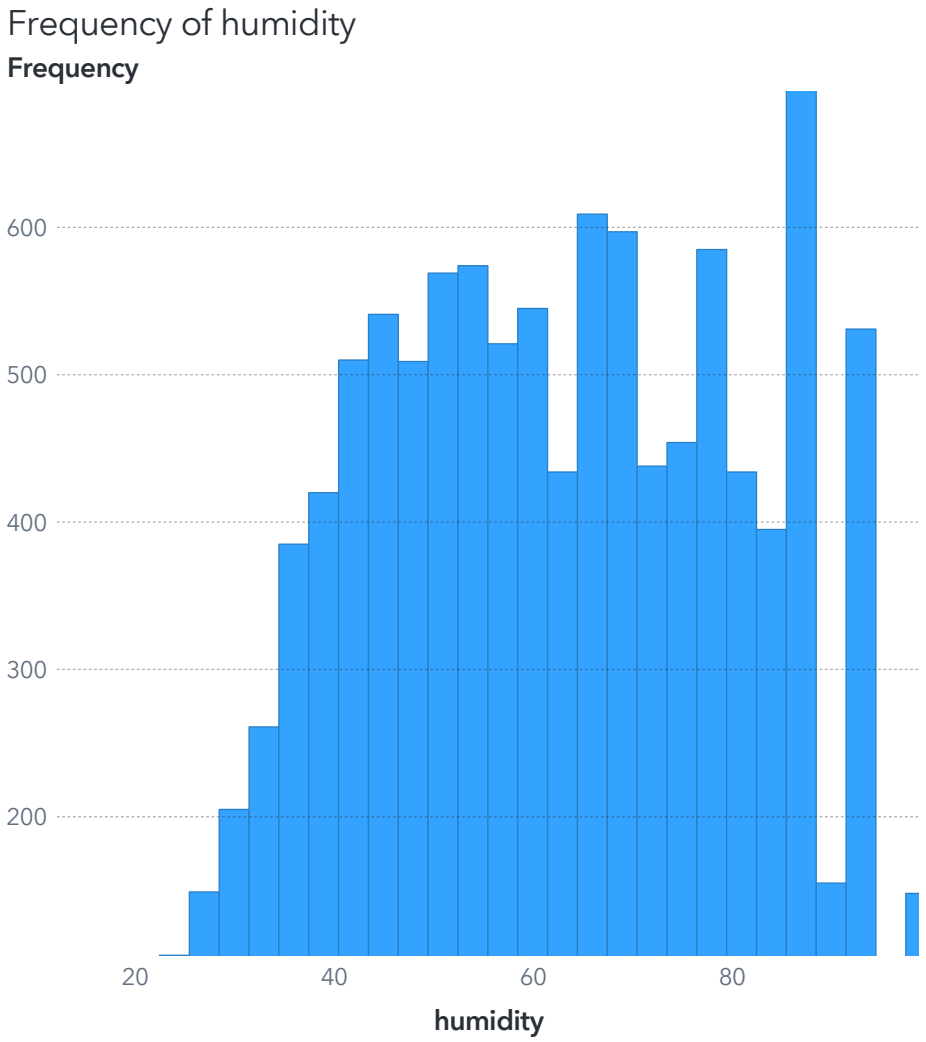
-High variance in median is noted across the different timing of the day.
-The 2 main peaks are observed from **7-9AM** and **4-6PM**.

Total Users by time_class



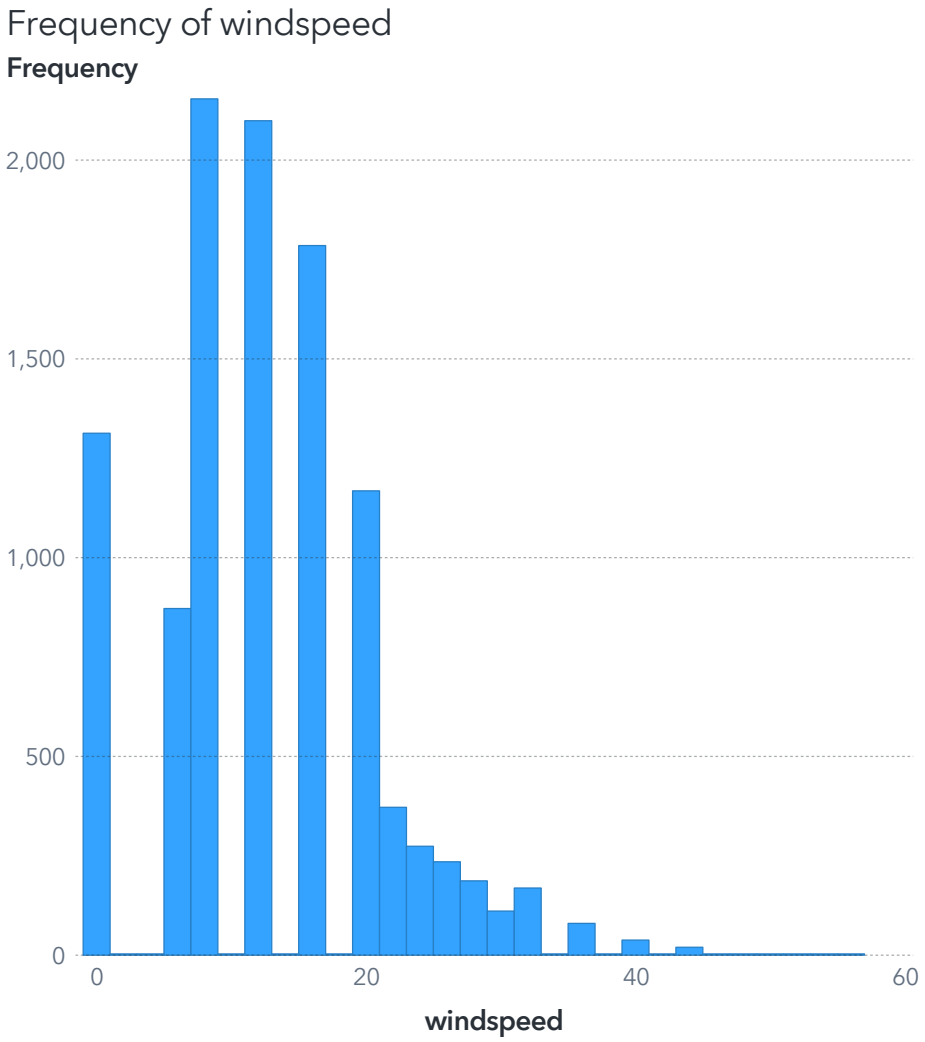
## Insight 11: Missing/Insufficient Data to Cover Range of Humidity

-Insufficient/ no data noted in the range **1-14 %** & **95-98 %**.
-There are also a small number of data that are with 0 % to be noted as it's quite unlikely for this to happen.




## Insight 12: Missing/ Insufficient Data to Cover Range of Windspeed

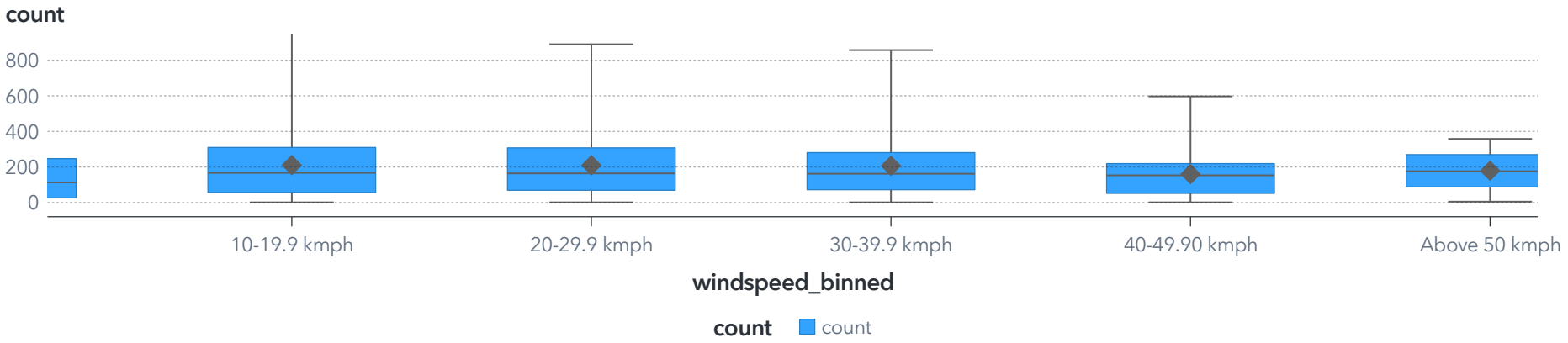
-Missing data noted in certain ranges for windspeed.
-More data might be required to make the analysis more accurate



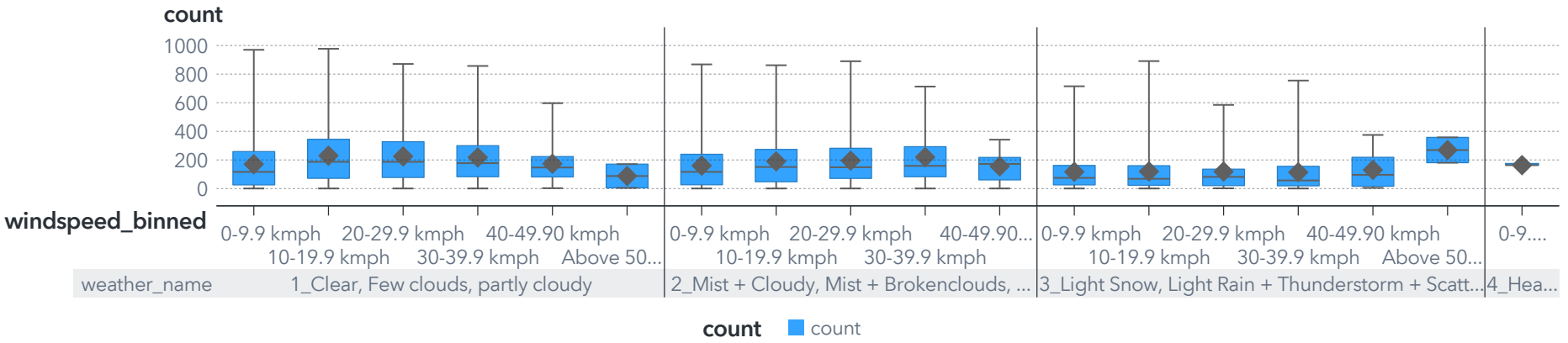
## Insight 13: Windspeed has effect on Total Users below 40 kmph

-Minimal variance noted within the category means for the category 29 to 40 kmph.
-This is also quite consistent across the different weather condition.
-As per insight 2, there are insufficient data from 0-20 kmph range.

Total Users by windspeed_binned




Total Users by windspeed_binned & weather_name



## Managerial Recommendations

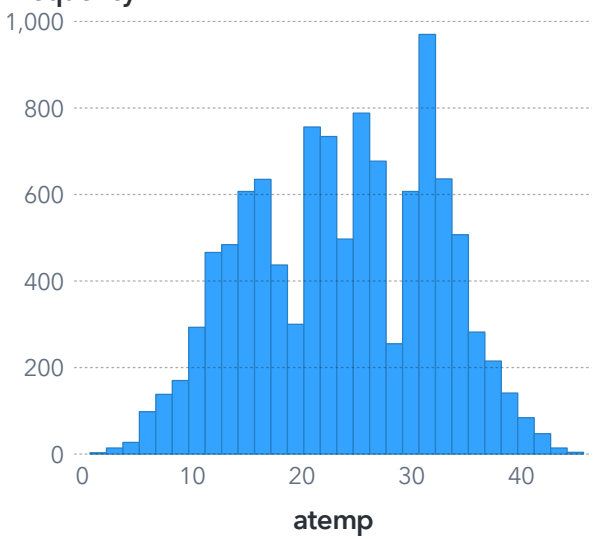
The findings of this analysis concludes that there are relation between various factors as below: Season Humidity Weather Condition xxxx xxxx

It it good to note that although there are relationship observed, it may not necessary be the cause for the demand of bike to go up or down.

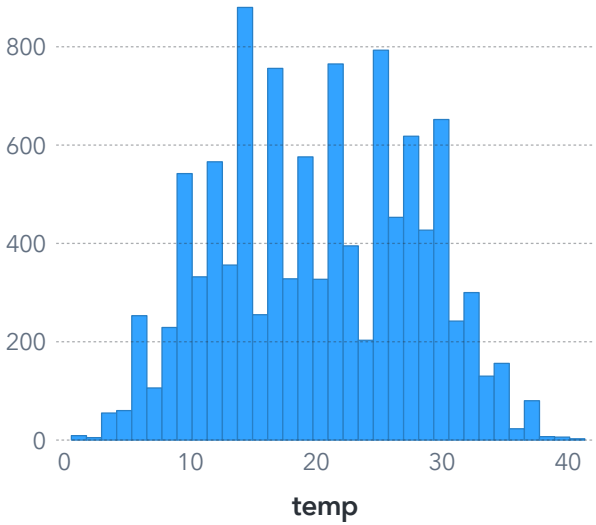
Missing/ insufficient data points for: 1. Humidity level 2. Type 4 weather

Additional data points to include for analysis:
Demographics: Gender, Age
Rent and Return location: To understand how the usage varies with location.

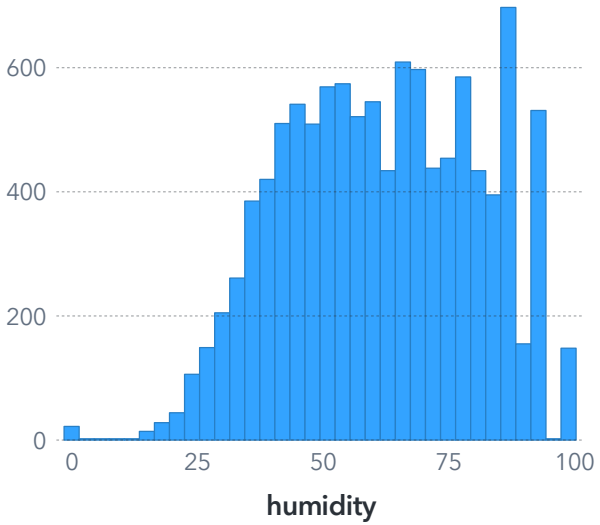
Frequency of atemp



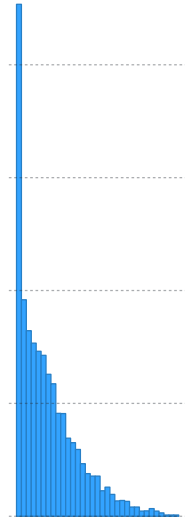

Frequency of temp



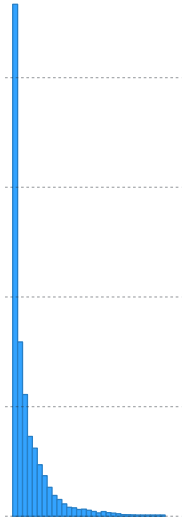

Frequency of humidity



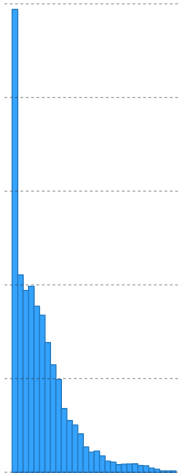

Frequency of count



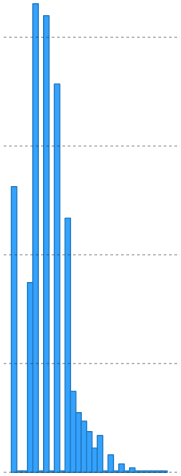
Frequency of casual



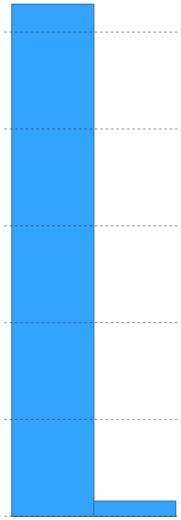
Frequency of registered



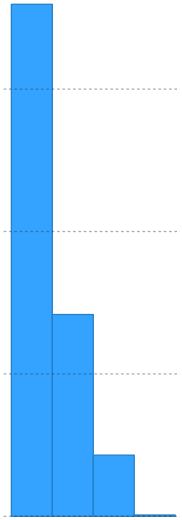
Frequency of windspeed



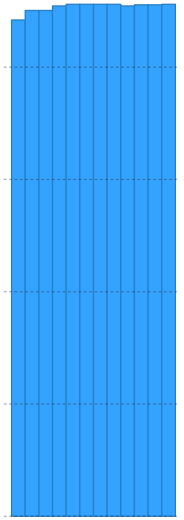
Frequency of holiday



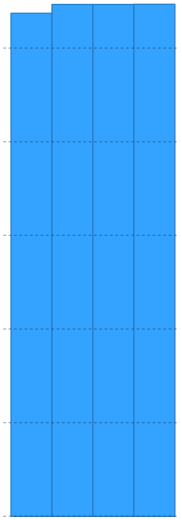
Frequency of weather



Frequency of month

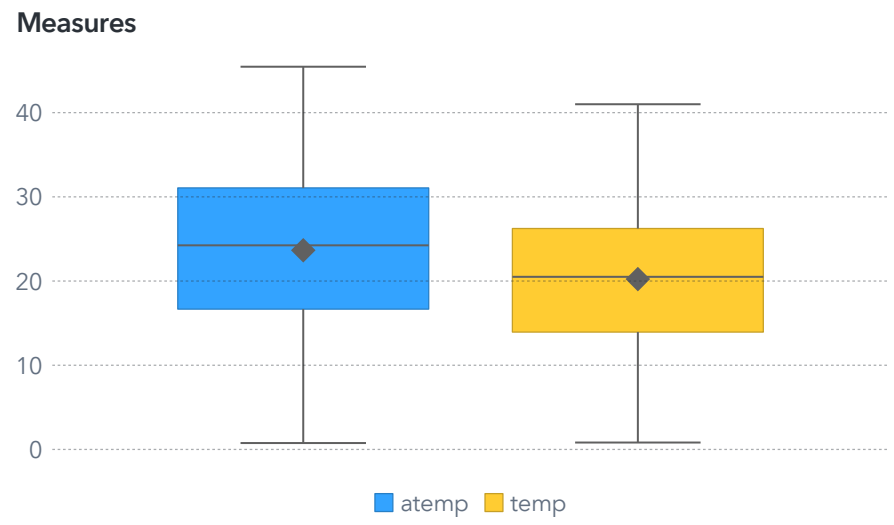


Frequency of season

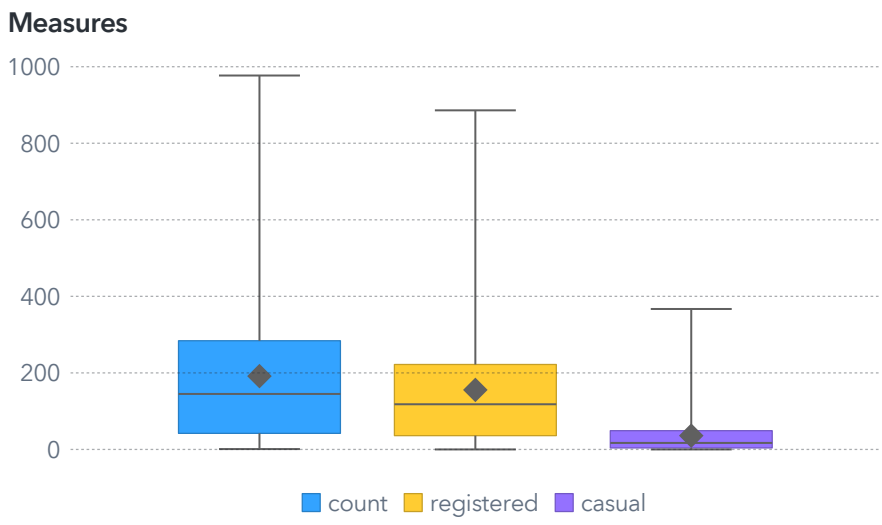


# Boxplot

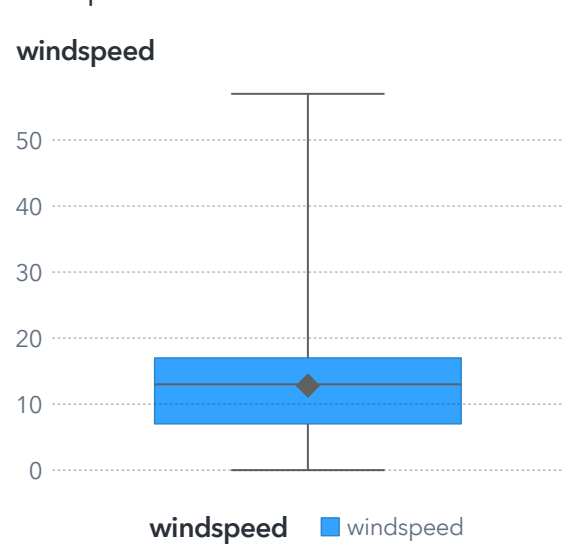
## atemp, temp



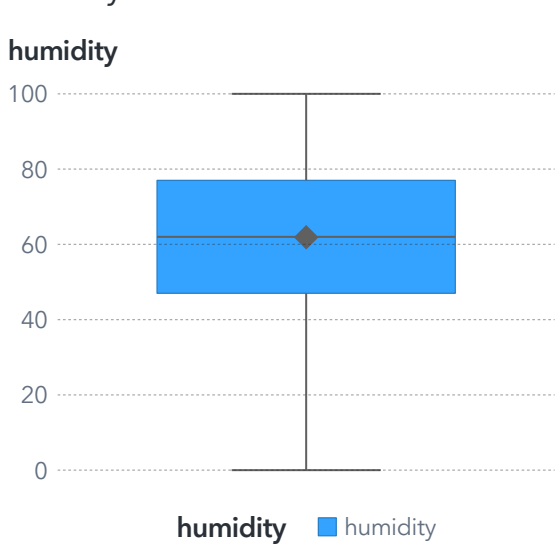

## count, registered, casual




## windspeed




## humidity




## month