YULU: CASE STUDY

Introduction

Yulu is India's leading micro-mobility service provider, which offers unique vehicles for the daily commute. Starting of a mission to eliminate traffic congestion in India. Yulu provides the safest commute solution through a user-friendly mobile app to enable shared, solo and sustainable commuting. Yulu zones are located at all the appropriate locations to make those first and last miles smooth, affordable and convenient!

Business Problem

Yulu has recently suffered considerable dips in the revenues. They want to know the factors on which shared electric cycles depend. Specifically, they want to understand the factors affecting the demand for these shared electric cycles in the Indian market.

Approach

- General Analysis of the data followed by Visual understanding.
- Tracking the count of consumers according to different situations.
- Finding variables which are significant in predicting the demand for shared electric cycles.
- All the recommendations and insights are backed by data and visual representation.

Non-Graphical Analysis of Sample Data:

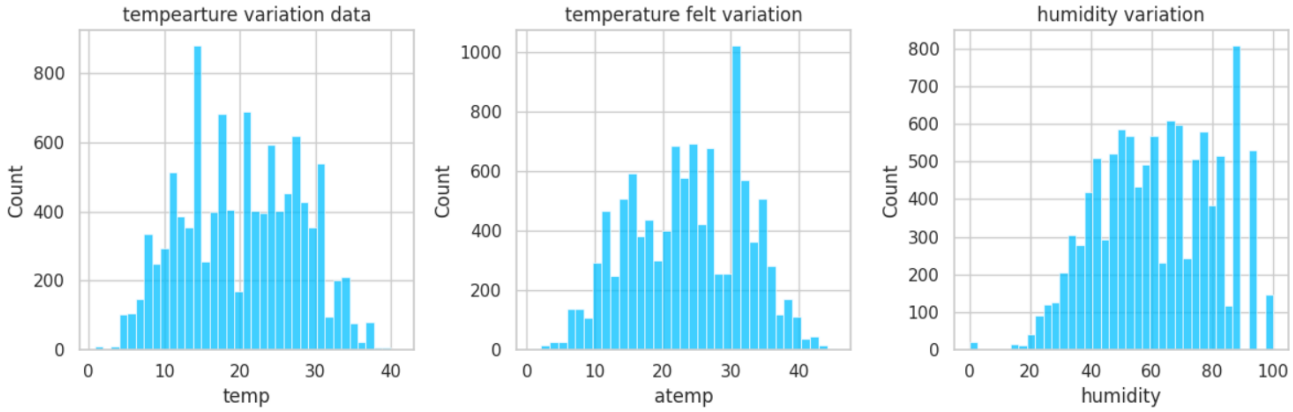
- Shape of data= (10866 rows x 12 columns)
- Size of data = 130392 in raw form
- Attributes of Data: datetime, season, holiday, workingday, weather, temp, atemp, humidity, windspeed, casual, registered, count.
- Any N/A or NaN value present: “Not found.”
- Data Types initially present: ‘Object’, ’int64’, ‘float64’.
- Number of distinct seasons: 4 (Spring coded as 1, Summer coded as 2, Fall Coded as 3, Winter coded as 4).
- There are 2686 winter season data, 2733 Summer & Fall season data each and 2734 winter data.
- Number of distinct Weather: 4
- There are 7412 working and 3474 non-working days.
- There are 10575 non-Holiday and 311 Holiday data present.
- Maximum and Minimum temperature: 41.0 and 0.82 degree Celsius respectively.
- Maximum and Minimum temperature felt: 45.455 and 0.76 degree respectively.

Insights:

- The data is clean and there are several parameters which can be related to the count parameter to get good insights and recommendations.
- The data has all weather conditions involved.
- The data clearly differentiate between working/ non-working and holiday/non-Holiday, these parameters will help to understand the demand better.

Setting Up Tune:

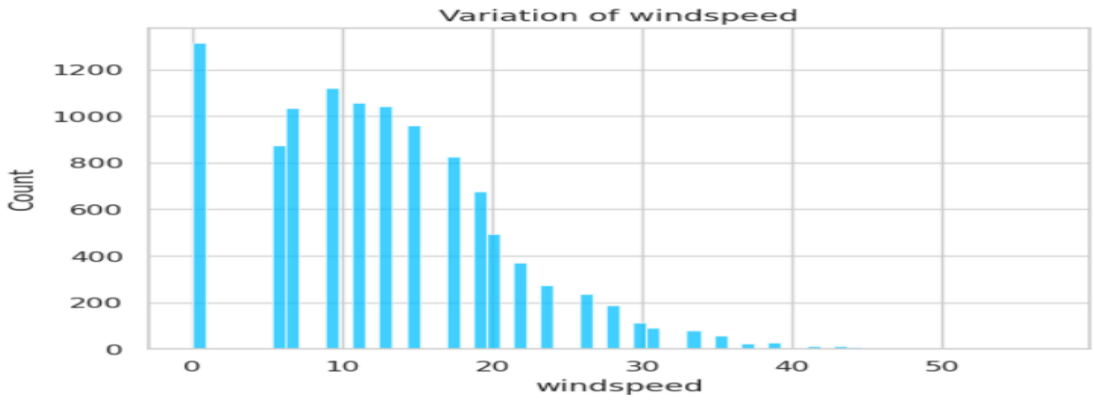
Checking how temperature, felt temperature and humidity is present in the graph.



Insights:

- The temperature variation has wide spectrum but most of it lay between 10-30.
- Temperature felt showed a bit variation but still most of it lies between 10-30, with outliers present.
- Humidity variation is also high, and it is one of the factors for the shift in temperature felt.

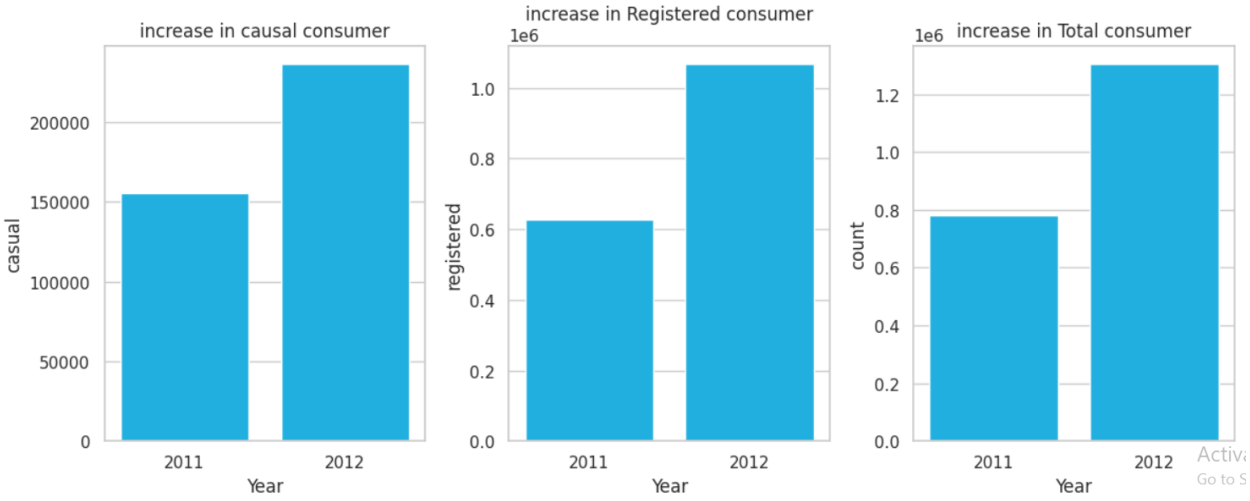
Checking how the wind speed is varying:



Insights:

- With some outliers, most of the wind speed lies between 0-20.

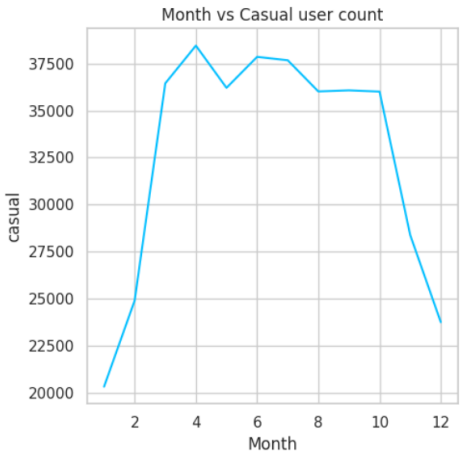
Tracking how casual, registered and overall consumer usage is increasing/decreasing:

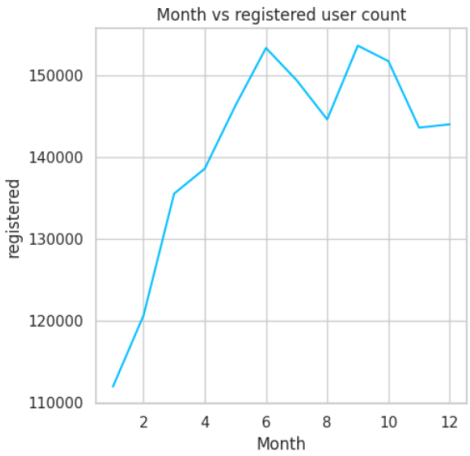


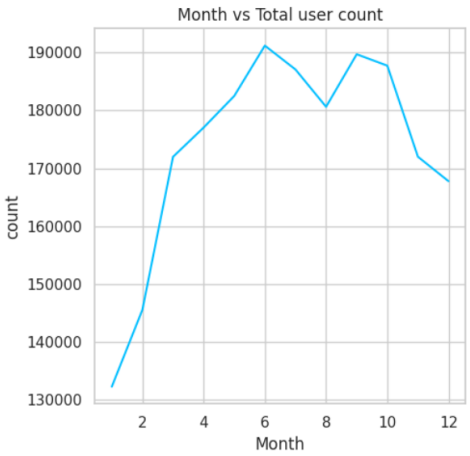
Insights:

- The number of uses by casual consumer in year 2011 was 155817 and it increased to 236318.
- The number of uses by registered consumer in year 2011 was 626162 and it increased to 1067179.
- The overall number of uses in year 2011 was 781979 and it increased to 1303497.
- So, the number of uses increased considerably in one year.

Tracking month to get any specific pattern for usage:



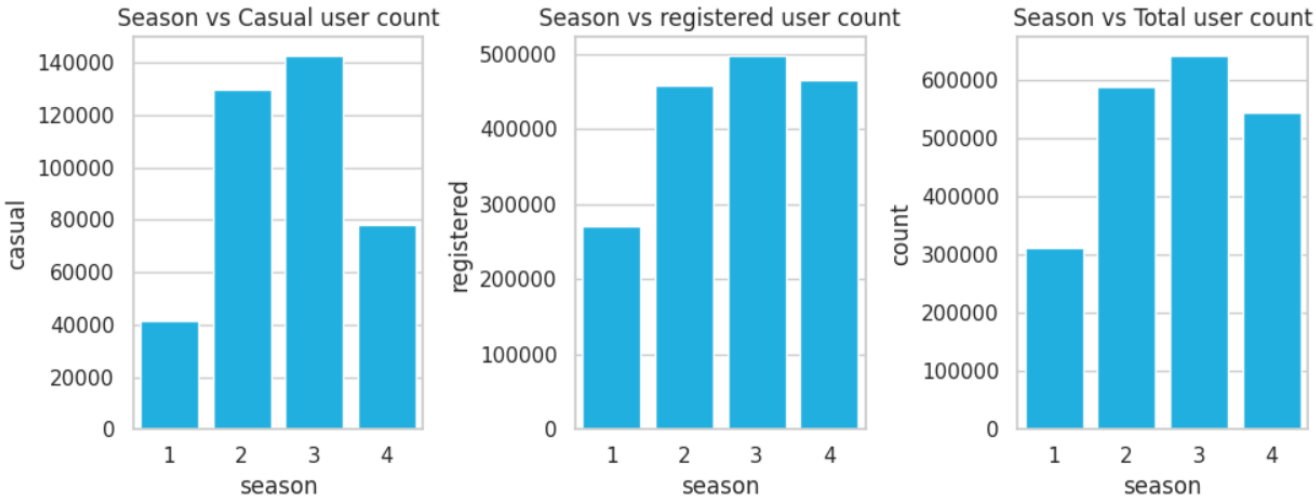




Insights and Recommendations:

- Some interesting observations are being unveiled.
- Usage by casual consumers is high in between 3rd and 10th month of the year, but then there is drop which follows for another 4 months.
- Registered consumers too have less consumption, and it only increases after 4th month of the year, and it again starts to decrease after 10th month of the year.
- Overall consumer base usage follows similar trend.
- There is a need to deeper analysis and see if other recorded attributed present in the data is contributing to steep decrease.

Tracking season to get any specific pattern:



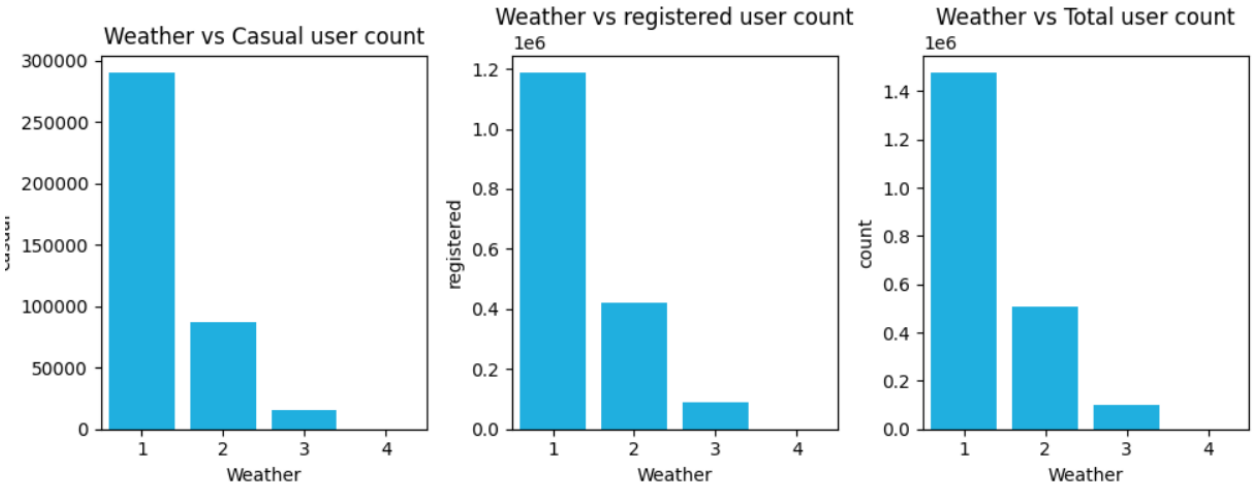
Assumption made:

- Spring coded as 1, Summer coded as 2, Fall Coded as 3, Winter coded as 4.

Insights and Recommendations:

- Customer usage is high for Summer and Fall season and it is lesser for Spring and Winter.
- Casual users seem to drop the usage of cycles by considerable number in Spring and Winter Season.
- There is similar decrease in registered users too, but winter usage does not see that much drop.
- Taking feedback from these customers about what motivates them to use cycles in winter, and then displaying that feedback through video or blog can help increase casual customer usage.
- Also discount ride and spring centric marketing campaigns can help.

Tracking weather to get any specific pattern:



Assumptions Made:

- 1 for Clear Weather, 2 for Misty and cloudy, 3 for Light snow and Light rain, 4 for heavy rain and ice pallets.

Insights and recommendations:

- Clear weather sees a good usage of the service but as the weather starts to degrade the service also decreases, reason maybe people preferring to stay away from 2-wheeler as the weather degrade as they are most exposed to harsh weather.

- Safety measures and new precaution measure/guidelines specially designed for harsh weather which ensures safety can gain customer trust and can eventually increase customer usage.

- Discount ride in those situations can help.

Final Confirmations

Visual and Numerical Observations till now have shown the impact of different attributes on the usage of cycles. There is need for some tests to give statistical confirmation ‘to the observations made and to get some understanding to increase customer satisfaction and eventually usage base.

Which test?

Ttest, Chi Square test and ANOVA test are done.

Note: Before going for ANOVA, we will use shapero test and levene test to check normality and variance.

Steps to be followed:

1. Construct Hypothesis.
2. Choose the type of test from the above three.
3. Significance level=0.05
4. Test conduction and p_value.
5. Conclude.

**Step 4 will be coded, and we will use p_value in Business case to conclude our Hypothesis. (Detailed coding is done in Google Colab).

From Observation made above, some important questions which can help in answering most of the company questions are:

1. Effect of working day on number of electric cycles used.
2. Effect of season on number of electric cycles used.
3. Effect of weather on number of electric cycles used.
4. Effect of Season on weather.

Effect of working day on number of electric cycles used.

Null Hypothesis(H0): The electric cycle rented is not dependent on season.

Alternate Hypothesis (Ha): The electric cycle rented is dependent on season.

Test used: Independent T-test.

Significance level(alpha): 0.05

Casual Usage:

p_value(calculated):1.706*10^-(-37) =0

alpha>p_value

Conclusion: Null Hypothesis is rejected, the electric cycles rented on non-working days are more than that of working days.

Registered Usage:

p_value(calculated):1.

alpha<p_value

conclusion: Null Hypothesis is not rejected so the electric cycles rented on working days are more than that of non-working days.

Overall usage:

p_value(calculated):0.7123.

alpha<p_value

conclusion: Null Hypothesis is not rejected so the electric cycles rented on working days are more than that of non-working days.

Insights and recommendations:

- Registered consumer usage is more on working day whereas the casual user shows a difference.

Effect on season on number of electric cycles used:

Null Hypothesis(H0): The electric cycle rented is not dependent on season.

Alternate Hypothesis (Ha): The electric cycle rented is dependent on season.

Test used: ANOVA.

Sub-Tests: Shapiro for gaussian and levene for variance.

Significance level(alpha): 0.05

Shapiro test for Gaussian:

H0: Data is following Gaussian.

Ha: Data is not following Gaussian.

P_value: 0(nearly zero)

alpha>p_value

Conclusion: Data is not following Gaussian.

Levene test for variance:

H0: variances are equal.

Ha: variances are not equal.

P_value:0 (nearly zero)

alpha> p_value

Conclusion: Variances are not equal.

Casual Usage:

p_value(calculated):7.93100*10^-(-214) =0

alpha>p_value

Conclusion: Null Hypothesis is rejected, the electric cycles rented depends on the season.

Registered Usage:

p_value(calculated):0(nearly zero)

alpha<p_value

conclusion: Null Hypothesis is rejected so the electric cycles rented depends on season.

Overall usage:

p_value(calculated):0(nearly zero)

alpha<p_value

conclusion: Null Hypothesis is rejected so the electric cycles rented depends on season.

Insights:

- The usage of electric cycle depends directly on the season, which we saw in visual plots too.

Effect on weather on number of electric cycles used:

Null Hypothesis(H0): The electric cycle rented is not dependent on weather.

Alternate Hypothesis (Ha): The electric cycle rented is dependent on weather.

Test used: ANOVA.

Sub-Tests: Shapiro for gaussian and levene for variance.

Significance level(alpha): 0.05

Shapiro test for Gaussian:

H0: Data is following Gaussian.

Ha: Data is not following Gaussian.

P_value: 0(nearly zero)

alpha>p_value

Conclusion: Data is not following Gaussian.

Levene test for variance:

H0: variances are equal.

Ha: variances are not equal.

P_value:0 (nearly zero)

alpha> p_value

Conclusion: Variances are not equal.

Casual Usage:

p_value(calculated):3.313100*10^-(-44) =0

alpha>p_value

conclusion: Null Hypothesis is rejected, the electric cycles rented depends on the weather.

Registered Usage:

p_value(calculated):0(nearly zero)

alpha<p_value

conclusion: Null Hypothesis is rejected so the electric cycles rented depends on weather.

Overall usage:

p_value(calculated):0(nearly zero)

alpha<p_value

conclusion: Null Hypothesis is rejected so the electric cycles rented depends on weather.

Insights:

- The usage of electric cycle depends directly on the weather, which we saw in visual plots too, where there was sharp decrease in usage as the weather worsened.

Effect of season on weather:

Null Hypothesis(H0): Weather does not depend on the season.

Alternate Hypothesis (Ha): Weather depends on the season.

Test used: Chi-square.

Significance level(alpha): 0.05

p_value(calculated): 1.545*10^ (-7)

alpha>p_value

conclusion: Null Hypothesis is rejected so the weather depends on the Season as the earlier visual plot showed.

Recommendations (Things which are helpful):

There is variation in usage of service, according to the various parameters. Good thing is that there are consumers in every condition, from harsh weather conditions to nonworking days to different season company do have customer usage. There is need for proper feedback in some condition and some experimentation in another to build trust.

- There are casual consumers who are engaging more on non-working days, proper identification of those individual in terms of job detail or their purpose of usage can help us grab more similar consumers/business.

- Harsh weather conditions need precautions, new safety measures and awareness can help to gain trust. Also, Discounted rides can help. There should also be a defined boundary so that the product itself don't get damaged in harsh conditions.

- There is need for consumer feedback in terms of their experience in different seasons, this positive message will generate trust in consumers who use the service seasonally. Discounted rides and season-based marketing campaign can help.