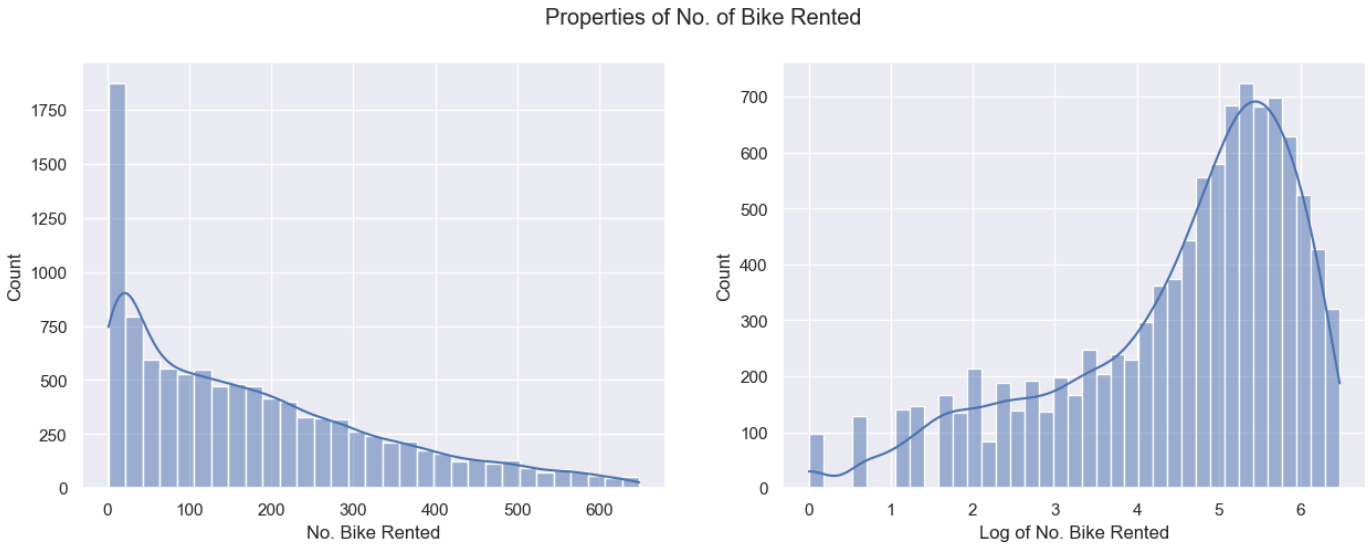
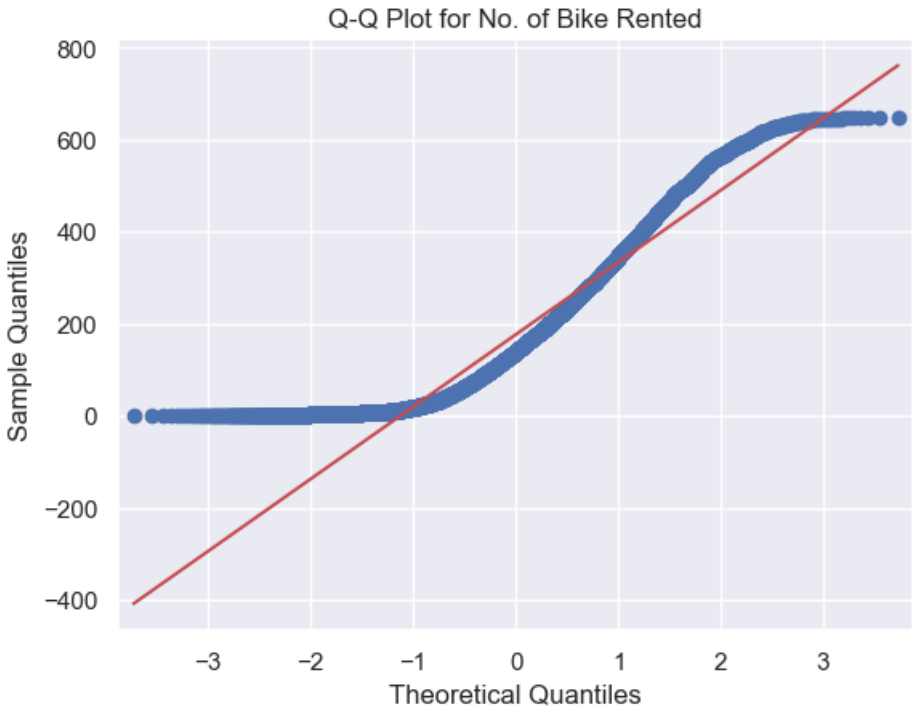
Validate Target variable's Data is `Gaussian or not`

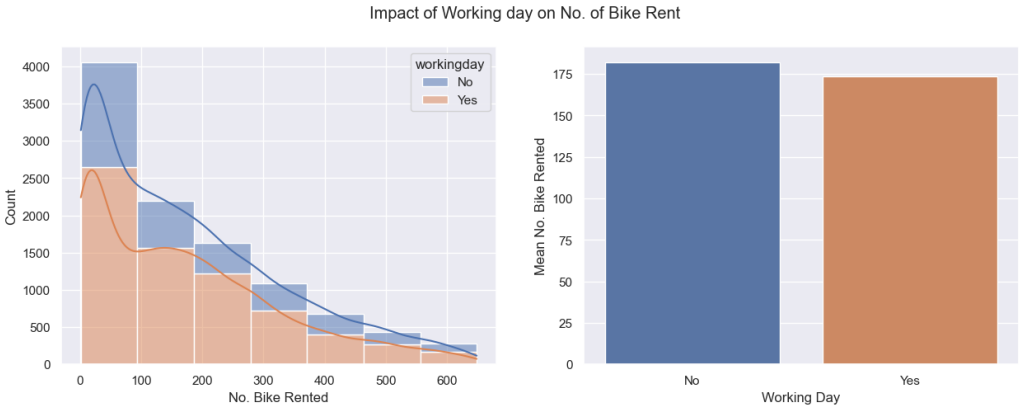
- Using **Shapiro-will's** Hypothesis Test on 'count' column, p_value found as **0%**.
- Thus **Data is not Gaussian**.





Q1. Working Day has effect on number of electric cycles rented?

Working day	Mean rent	No of rent
No	182.189881	3360
Yes	173.568507	6992



Ho -> Working Day has no effect on number of electric cycles rented

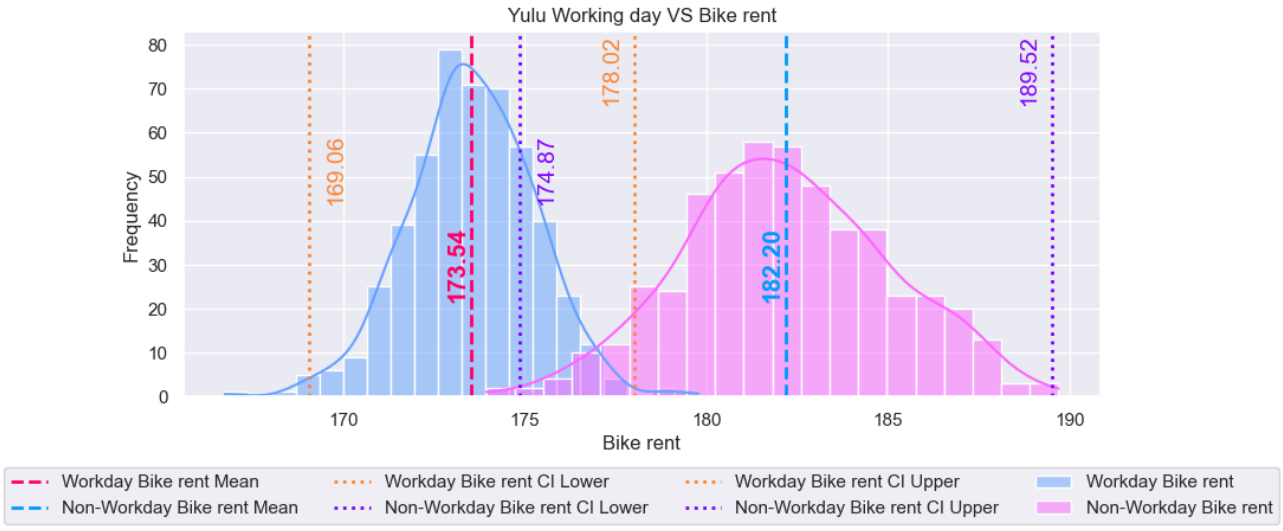
Ha -> Working Day has effect on number of electric cycles rented

Samples size = 500

Significance_level = 0.1 = 1%

```
ttest_ind(rent_on_wd, rent_on_non_wd)
```

p_value = `0.887%`



Insights

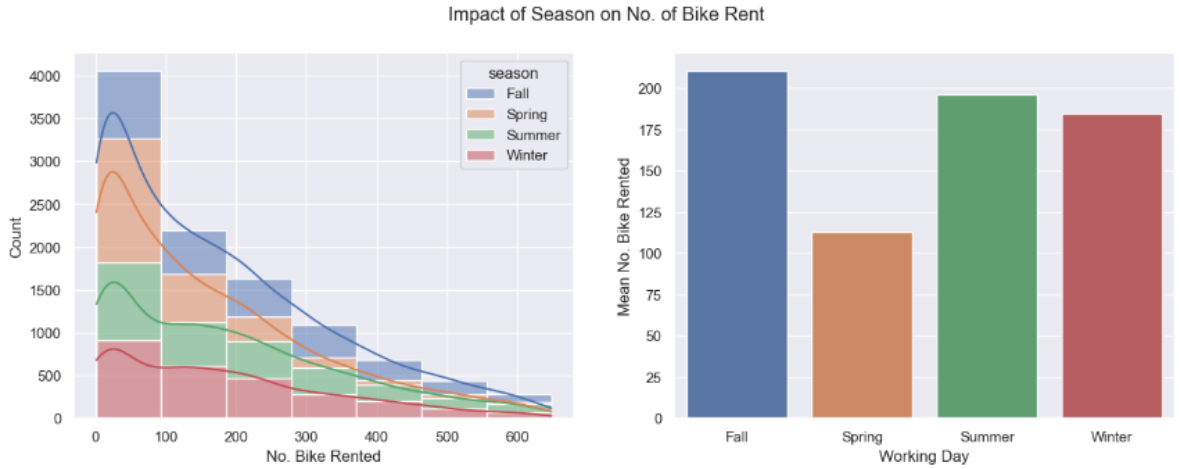
Using T-Test p_value found that **0.887%.**

- With Confidence interval of **99%** & Sample Size of **500**
  - Mean Bike rent on Working day is 173.48 with a Intervals of (168.92 - 178.04).
  - Mean Bike rent on Non-Working day is 182.21 with a Intervals of (175.20 - 189.22).
  - As 0.887% < 1% Thus Rejecting Ho & Accept Ha.

As per T-Test we can Conclude "**Working Day has effect on number of electric cycles rented**".

Q2. No. of cycles rented similar or different in different seasons?

Season	Mean rent	No of rent
Fall	210.633564	2598
Spring	112.774308	2530
Summer	195.945328	2579
Winter	184.446503	2645

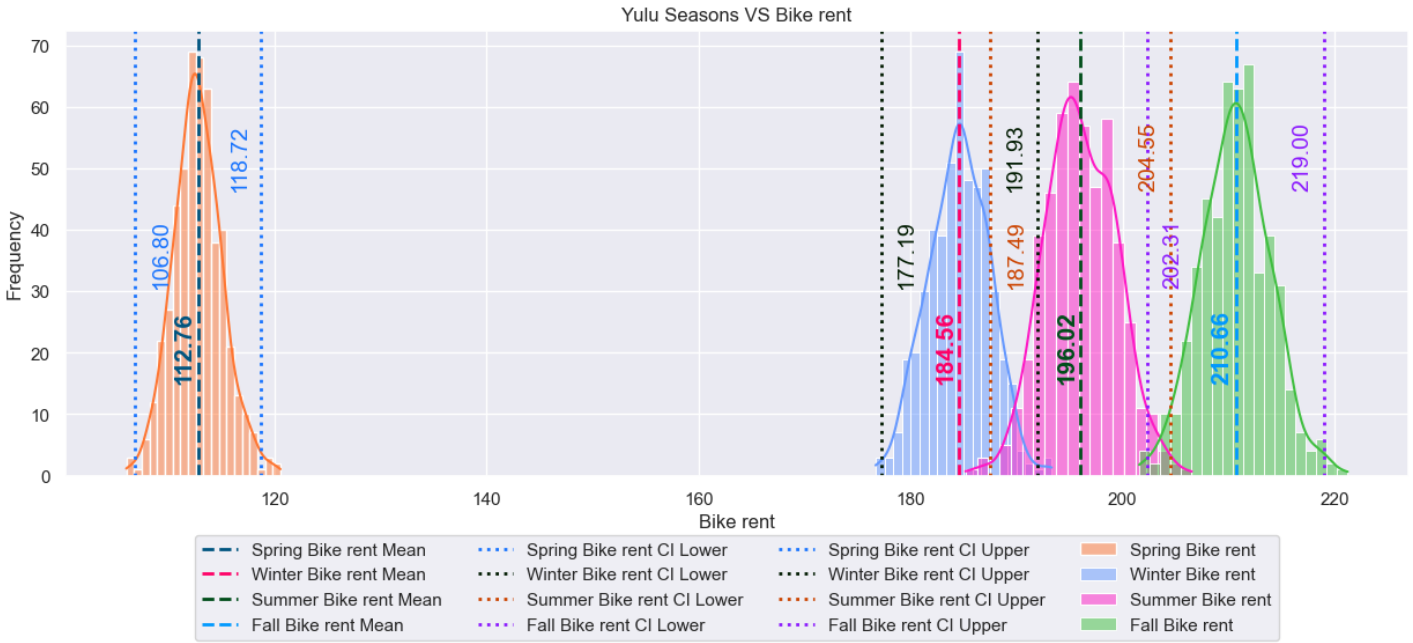


Ho -> No. of Bike rented are Similar in different seasons

Ha -> No. of Bike rented are different in different seasons

Samples size = 500

Significance_level = 0.1 = 1%



Using ANOVA Test

```
f_oneway(on_summer, rent_on_fall, rent_on_winter, rent_on_spring)
```

p_value = 0%

Using Kruskal Test

```
f_ kruskal(on_summer, rent_on_fall, rent_on_winter, rent_on_spring)
```

p_value = 0%

Using Levene Test

```
f_ kruskal(on_summer, rent_on_fall, rent_on_winter, rent_on_spring)
```

p_value = 0%

Insights

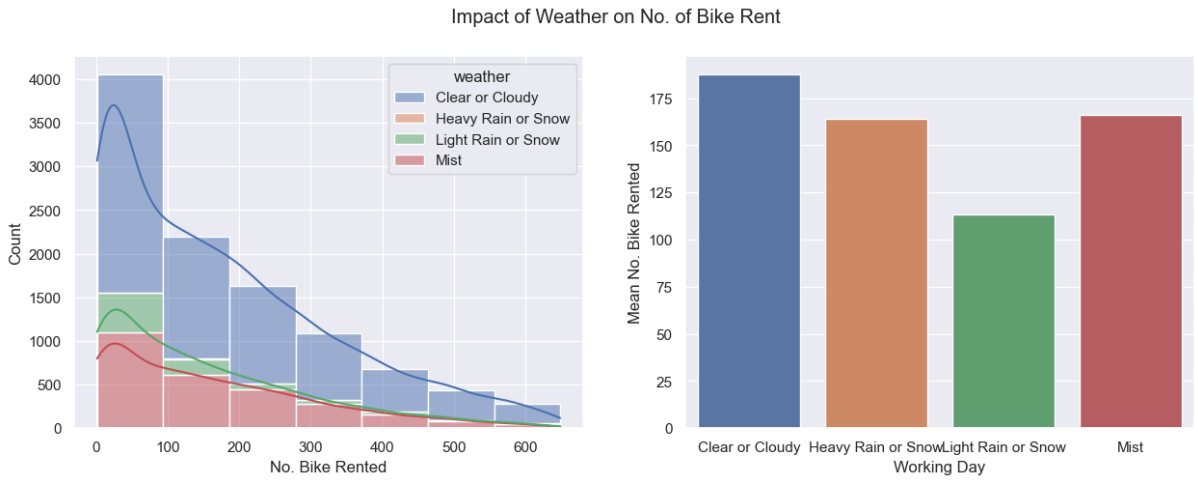
Using ANOVA, Kruskal & Levene Test p_value found that **0.0%.**

- With Confidence interval of **99%** & Sample Size of **500**
  - Mean Bike rent on Spring Season is 112.87 with a Intervals of (106.51 - 119.23).
  - Mean Bike rent on Winter Season is 184.68 with a Intervals of (176.95 - 192.40).
  - Mean Bike rent on Summer Season is 195.84 with a Intervals of (186.98 - 204.69).
  - Mean Bike rent on Fall Season is 210.84 with a Intervals of (202.32 - 219.37).
  - As 0.0% < 1% Thus Rejecting Ho & Accept Ha.

As per Anova Test we can Conclude "**No. of Bike rented are different in different seasons**".

Q3. No. of cycles rented similar or different in different weather?

Weather	Mean rent	No of rent
Clear or Cloudy	187.822607	6821
Heavy Rain or Snow	164.000000	1
Light Rain or Snow	113.562108	797
Mist	166.095134	2733



Ho -> No. of Bike rented are Similar in different weather

Ha -> No. of Bike rented are different in different weather

Samples size = 500

Significance_level = 0.1 = 1%

Using ANOVA Test

```
f_oneway(rent_on_clear_cloudy, rent_on_heavy_rain_snow,
rent_on_light_rain_snow, rent_on_mist)
```

p_value = 0%

Using Kruskal Test

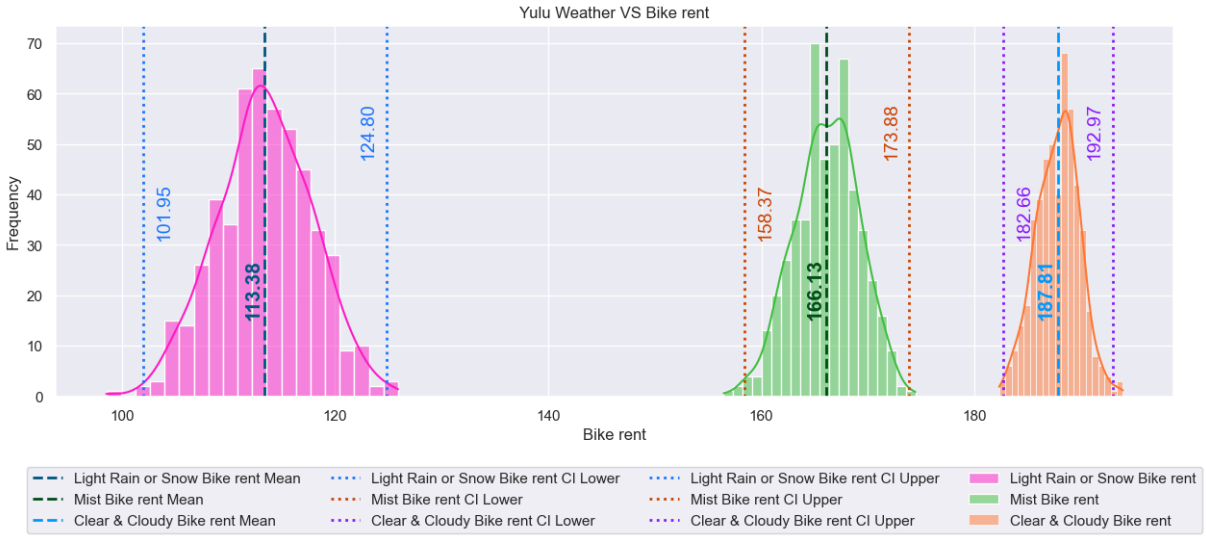
```
f_ kruskal(rent_on_clear_cloudy, rent_on_heavy_rain_snow,
rent_on_light_rain_snow, rent_on_mist)
```

p_value = 0%

Using Levene Test

```
f_ kruskal(rent_on_clear_cloudy, rent_on_heavy_rain_snow,
rent_on_light_rain_snow, rent_on_mist)
```

p_value = 0%



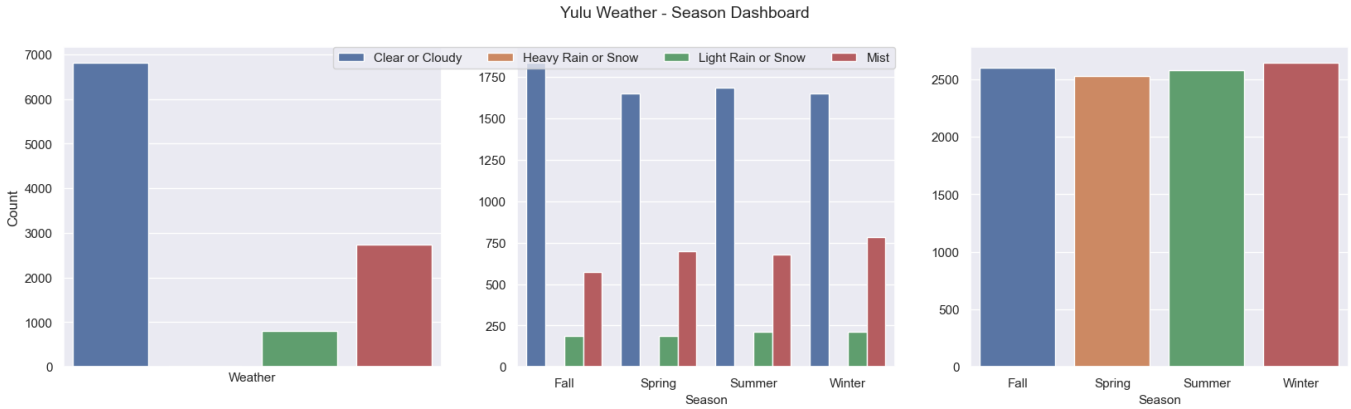
Insights

Using ANOVA, Kruskal & Levene Test p_value found that **0.0%.**

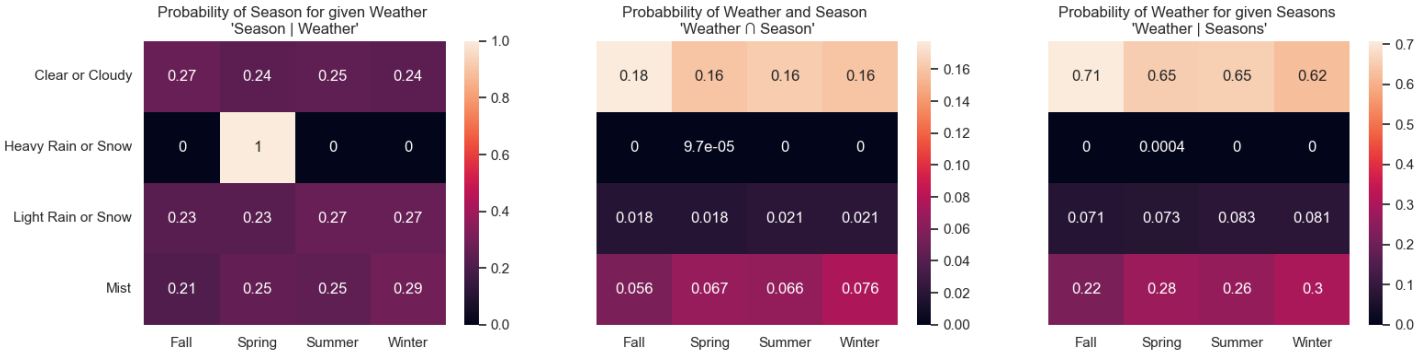
- With Confidence interval of **99%** & Sample Size of **500**
  - Mean Bike rent on 'Light Snow, Light Rain + Thunderstorm + Scattered clouds, Light Rain + Scattered clouds' Weather is 113.18 with a Intervals of (102.31 - 124.04).
  - Mean Bike rent on 'Mist + Cloudy, Mist + Broken clouds, Mist + Few clouds, Mist' Weather is 166.20 with a Intervals of (158.78 - 173.61).
  - Mean Bike rent on 'Clear, Few clouds, partly cloudy, partly cloudy' Weather is 187.83 with a Intervals of (182.51 - 193.14).
  - As only 1 record available for 'Heavy Rain + Ice Pallets + Thunderstorm + Mist, Snow + Fog' Weather, we are not going to consider this for Confidence Interval.
  - As 0.0% < 1% Thus Rejecting Ho & Accept Ha.

As per Anova Test we can Conclude "**No. of Bike rented are different in different weather**".

Q4. Weather is dependent on season?



Weather	Clear or Cloudy	Heavy Rain or Snow	Light Rain or Snow	Mist	All
Season					
Fall	1838	0	185	575	2598
Spring	1648	1	185	696	2530
Summer	1687	0	213	679	2579
Winter	1648	0	214	783	2645
All	6821	1	797	2733	**10352**



Ho -> Weather is Independent of Season.

Ha -> Weather is Dependent on Season.

Samples size = 500

Significance_level = 0.1 = 1%

Using ChiSqured Test

```
chi2_contingency(pd.crosstab(df_yulu['season'], df_yulu['weather']))
```

p_value = 0%

Insights

Using Chi-Square Test p_value found that **0.0%.**

- With Confidence interval of **99%** & Sample Size of **500**
  - As 0% < 1% Thus Rejecting Ho & Accept Ha.

As per Chi-Square Test we can Conclude **"Weather is Dependent on Season".**

- This is also absorbed from Heat Map & Probability plot
- As, Co-Relation Co-efficient for each Weather are mostly equal, except "Heavy Rain or Snow".