

# Bike demand summary

March 10, 2023

Univariate analysis- 1. Temperature: The peak is around 0.68 or 0.70, accounting for temperature around 27 deg Celsius. The distribution is normal. 2. No outliers present in atemp field. 3. No visible abnormally high values in humidity field. 4. Cnt: Plot is positively skewed, which indicates that most of the outliers are present after 600 and to the right side.Boxplot also indicates the presence of outliers after 600.Both plots provide similar insights that there are high values present in cnt field

Bivariate Analysis- 1. The Count vs Hour boxplot shows that bike rides are more in office going hours (between 7am to 9am) and coming hours(4pm to 7pm). 2. The Count vs Weekday boxplot shows outliers for monday,tuesday, wednesday and thursday, with maximum count on weekends without any outliers. 3. Count vs Month boxplot shows less bike rides during Nov, Dec, Jan,Feb,March(less IQR and more outliers) which increased from April till October January stands out the most with very less rides in average and most outliers 4. Spring season has the least count of rides(months=Jan,Feb,March), with fall being the highest(Aug,Sep,Oct)

Reporting R2 set on train set= 0.670263437848587 Reporting R2 on test set= 0.66449761515553