# Assignment-based Subjective Questions

1. **From your analysis of the categorical variables from the dataset, what could you infer about their effect on the dependent variable?**
   Ans)

   1. *Demand for the year 2019 has grown*
   2. *Season 3 - Fall has the highest demand*
   3. *Demand is continuously growing each month till June. September month has the highest demand. After September, demand is decreasing. During the year-end and beginning, it is less, which could be due to extreme weather conditions.*
   4. *On Holiday there is less demand as compared to Non-Holiday days.*
   5. *Weekday sales are almost closed on all days.*
   6. *The good weathershit has the highest demand*

2. **Why is it important to use drop_first=True during dummy variable creation?**
   *Ans) drop_first=True is important to use, as it helps in reducing the extra column created during dummy variable creation. Hence it reduces the correlations created among dummy variables.*

3. **Looking at the pair-plot among the numerical variables, which one has the highest correlation with the target variable?**
   *Ans) 'temp' & 'atemp' has the high correlation with the 'cnt' column and are highly correlated with each other*

4. **How did you validate the assumptions of Linear Regression after building the model on the training set?**
   *Ans) We basically did the Model evaluation, by performing the following operations*

   1. *Residual Analysis*
      a. *Errors are normally distributed here with a mean 0. So everything seems to be fine*
      b. *Actual and Predicted results follow almost the same pattern so this model seems ok*
      c. *R2 Same as we obtained for our final model*
   2. *Linearity Check*

3. *Predict values for test data set*
4. *R-Squared value for test predictions*
5. *Plot Error Terms for test data*

5. **Based on the final model, which are the top 3 features contributing significantly towards explaining the demand of the shared bikes?**

   *Ans)*

   1. *Company should focus on expanding business during Spring.*
   2. *Company should focus on expanding business during September.*
   3. *Based on previous data it is expected to have a boom in number of users once situation comes back to normal, compared to 2019.*