Table of Contents

Genral Information

- On the dataset, multiple linear regression is carried out.
- The project is completed as part of the Machine Learning module's coursework.
- We are attempting to determine how many leases the firm has issued based on a variety of independent factors, like temperature, weather, humidity, holidays, etc.
- It makes use of the Boombikes bike rental dataset.

Conclusion

Our model successfully predicts the variance in both the test and train datasets since its mean squared error is practically zero.
Following data interpretation, visualisation, data preparation, model building and training, residual analysis, and test model evaluation, the model's summary is as follows:
The test set's R-squared value is 81.13%, whereas the train set's value is 82.71%, suggesting that our model accurately and comprehensively describes the variance on the test set. As a result, we can say that the model is good.
The mean squared error of our generated model is almost zero on both the training and testing datasets, indicating that the variance is correctly predicted on the test set. Amounts of p and The relevant factors were chosen using and VIF. RFE was also carried out to pick variables automatically.
We may infer that the temperature and whether it is a working day or not affect the bike needs for the BoomBikes firm. In addition, it appears that demand for rentals

is higher in the winter than it is in the summer and spring. We had seen that rental activity was increased in September and October. The focus was mostly on holidays rather than days like Wednesday, Thursday, and Saturday.
We can get valuable insights into both the bike rental business and human behaviour thanks to these interpretations. Based on this paradigm, one suggestion is that there should be
active spring and summer marketing to increase rental rates. A solid marketing plan for the first six months of the year can help increase rental rates because the summertime also exhibits low rental levels. On days when the weather is less clear, there must be a strategy to introduce additional users, possibly with rewards or clever bargains. Rentals increased in 2019 compared to 2018, which shows that more people will become aware of this concept over time. To keep repeat clients, a thorough study must be conducted.

Technologies Used:
Python

- Numpy,
- Pandas,
- Seaborn,
- Matplotlib
- Statsmodels
- Sci-kit learn

Acknowledgement-

- https://www.youtube.com/watch?v=6HVCuXrsQBs
- https://www.youtube.com/watch?v=CSUP-o4mE_c
- https://towardsdatascience.com/end-to-end-case-study-bike-sharing-demand-dataset-53201926c8db

Contact me at @Therealnikunj on github