CAPSTONE PROJECT
Bike
Sharing Demand Prediction

NIKHIL BABURAO SHINDE

CONTENT

- ❑ BUSINESS UNDERSTANDING
- ❑ DATA SUMMARY
- ❑ FEATURE ANALYSIS
- ❑ EXPLORATORY DATA ANALYSIS
- ❑ DATA PREPROCESSING
- ❑ IMPLEMENTING ALGORITHMS
- ❑ CHALLENGES
- ❑ CONCLUSIONS

BUSINESS UNDERSTANDING

- Bike rentals have become a popular service in recent years and it seems people are using it more often. With relatively cheaper rates and ease of pick up and drop at own convenience is what making this business thrive.
- Mostly used by people having no personal vehicles and also to avoid congested public transport which that's why they prefer rental bikes.
- Therefore, the business to strive and profit more, it has to be always ready and supply no. of bikes at different locations, to fulfil the demand.
- Our project goal is a preplanned set of bike count values that can be a handy solution to meet all demands.

DATA SUMMARY

- This Dataset contains 8760 lines and 14 columns.
- Three categorical features ‘Seasons’, ‘Holiday’, & ‘Functioning Day’.
- One Datetime features ‘Date’.
- We have some numerical type variables such as temperature, humidity, wind, visibility, dew point temp, solar radiation, rainfall, snowfall which tells the environment conditions at that particular hour of the day.

	Date	Rented Bike Count	Hour	Temperature(°C)	Humidity(%)	Wind speed (m/s)	Visibility (10m)	Dew point temperature(°C)	Solar Radiation (MJ/m2)	Rainfall(mm)	Snowfall (cm)	Seasons	Holiday	Functioning Day
8755	30/11/2018	1003	19	4.2	34	2.6	1894	-10.3	0.0	0.0	0.0	Autumn	No Holiday	Yes
8756	30/11/2018	764	20	3.4	37	2.3	2000	-9.9	0.0	0.0	0.0	Autumn	No Holiday	Yes
8757	30/11/2018	694	21	2.6	39	0.3	1968	-9.9	0.0	0.0	0.0	Autumn	No Holiday	Yes
8758	30/11/2018	712	22	2.1	41	1.0	1859	-9.8	0.0	0.0	0.0	Autumn	No Holiday	Yes
8759	30/11/2018	584	23	1.9	43	1.3	1909	-9.3	0.0	0.0	0.0	Autumn	No Holiday	Yes

FEATURE SUMMARY

- Date : Year-Month-Day
- Rented Bike Count - Count of bikes rented at each hour
- Hour - Hour of the day
- Temperature - Temperature in Celsius
- Humidity - %
- Wind Speed - m/s
- Visibility - 10m
- Dew point temperature -Celsius
- Solar radiation -MJ/m2
- Rainfall –mm
- Snowfall -cm
- Seasons -Winter, Spring, Summer, Autumn
- Holiday -Holiday/No Holiday
- Functional Day - NoFunc(Non Functional Hrs),Fun(Functional Hrs)

INSIGHTS FROM OUR DATASET

➢ There are No Missing Values present

➢ There are No Duplicate values present

➢ There are No null values.

➢ And finally we have 'rented bike count' variable which we need to predict for new observations

➢ The dataset shows hourly rental data for one year (1 December 2017 to 31 November(2018)(365 days).we consider this as a single year data ➢ So we convert the "date" column into 3 different column i.e "year","month","day".

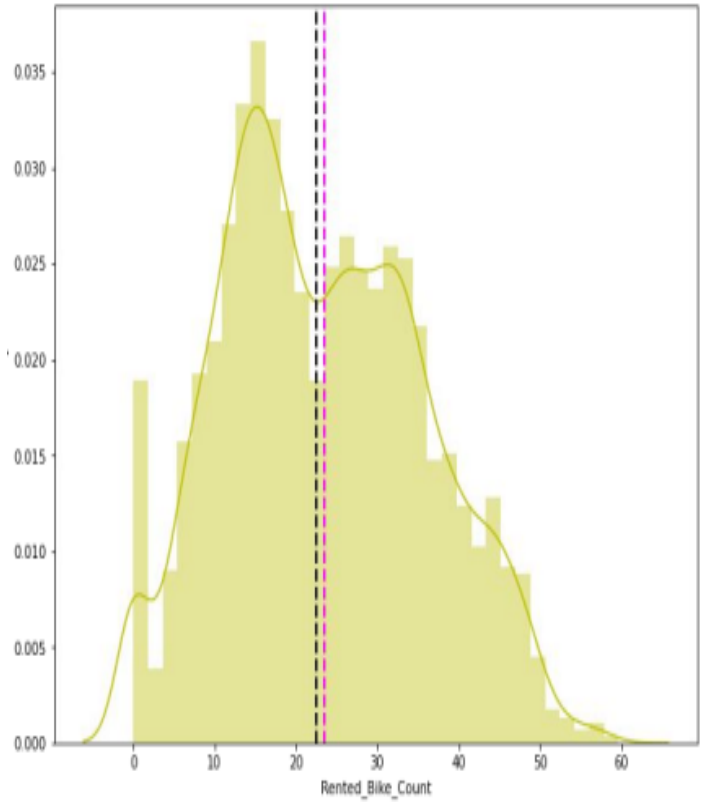
➢ We change the name of some features for our convenience , they are as below 'Rented_Bike_Count', 'Hour', 'Temperature', 'Humidity', 'Wind_speed', 'Visibility', 'Dew_point_temperature', 'Solar_Radiation', 'Rainfall', 'Snowfall', 'Seasons', 'Holiday', 'Functioning_Day', 'month','weekdays_weekend'

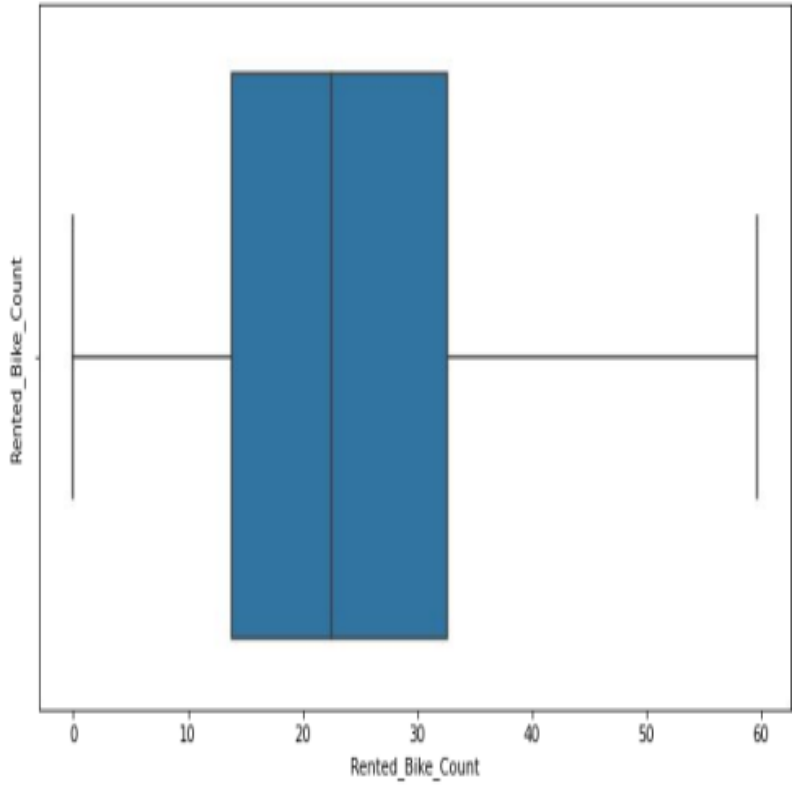
ANALYSIS OF RENTED BIKE COLUMN

➢ The above graph shows that Rented Bike Count has moderate right skewness.

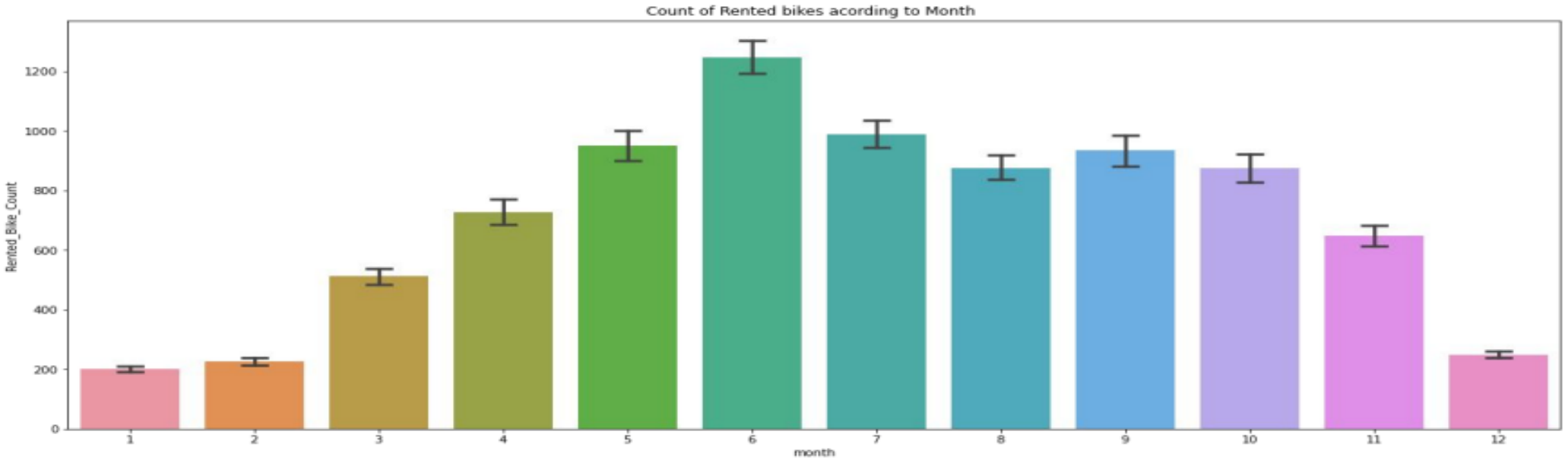
➢ The above boxplot shows that we have detect outliers in Rented Bike Count column

➢ Since the assumption of linear regression is that 'the distribution of dependent variable has to be normal', so we should perform Square root operation to make it normal



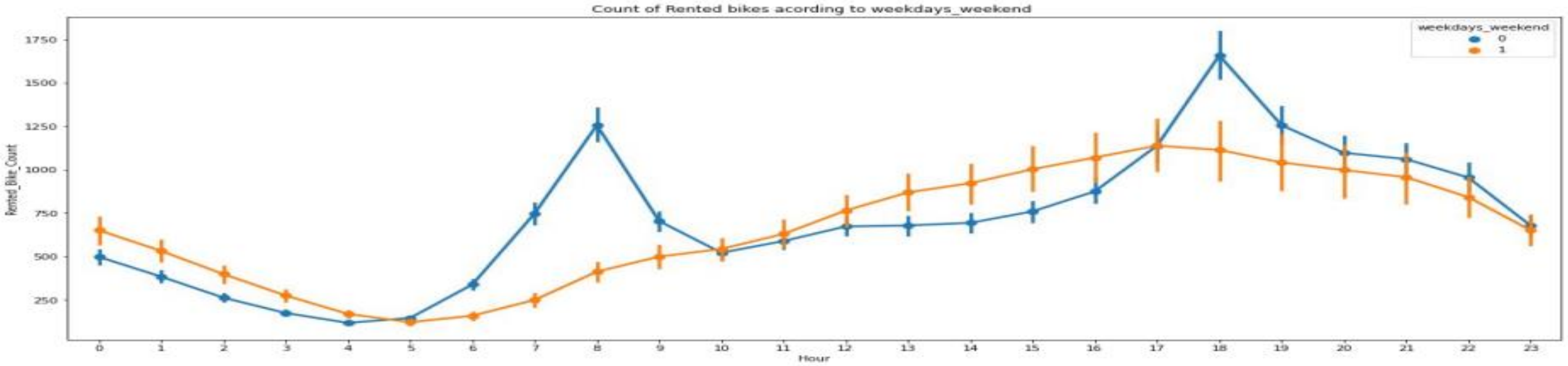


ANALYSIS OF MONTH VARIABLE



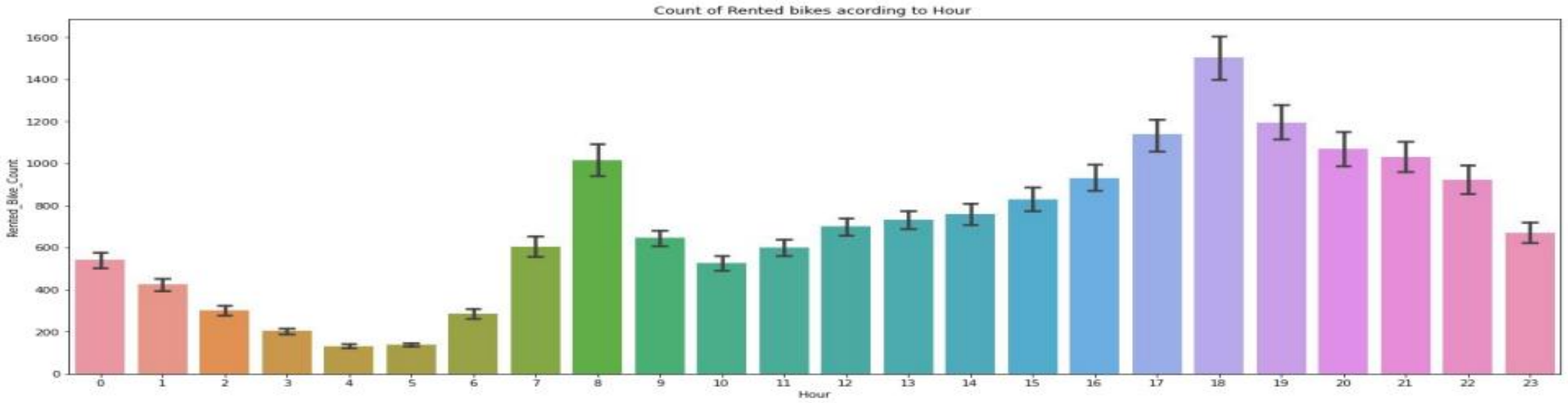
➢ From the above bar plot we can clearly say that from the month 5 to 10 the demand of the rented bike is high as compare to other months.These months are comes inside the summer season.

ANALYSIS OF WEEKDAYS_WEEKEND VARIABLE



- From the above point plot and bar plot we can say that in the weekdays which represent in blue color show that the demand of the bike higher because of the office.
- Peak Time are 7 am to 9 am and 5 pm to 7 pm
- The orange color represent the weekend days, and it show that the demand of rented bikes are very low especially in the morning hour but when the evening start from 4 pm to 8 pm the demand slightly increases.

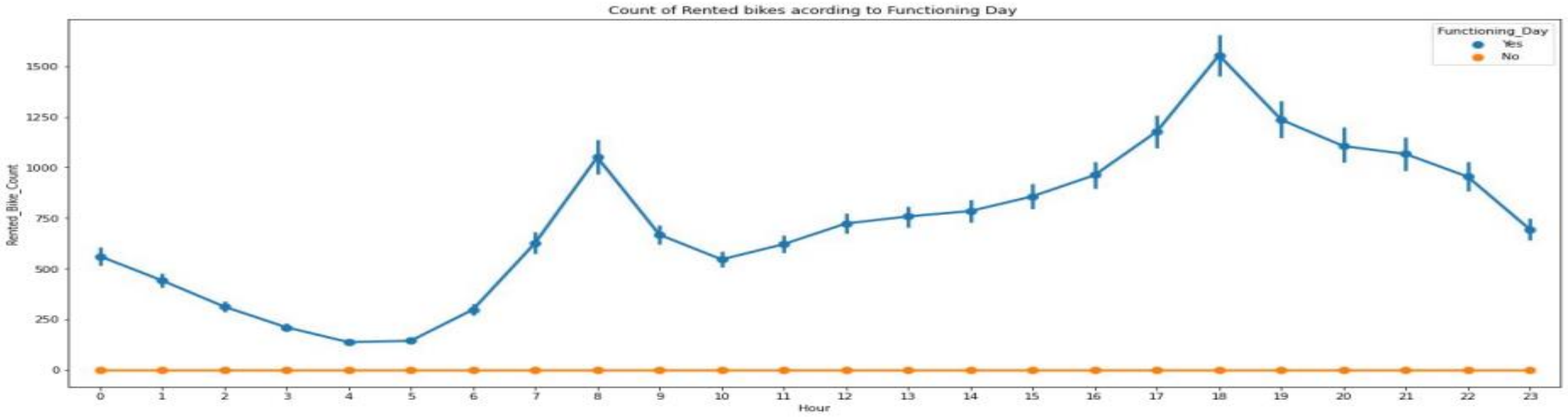
ANALYSIS OF HOUR VARIABLE



➢ In the above plot which shows the use of rented bike according the hours and the data are from all over the year.

➢ generally people use rented bikes during their working hour from 7am to 9am and 5pm to 7pm

ANALYSIS OF FUNCTIONING DAY VARIABLE



➢ In the above point plot which shows the use of rented bike in functioning daya or not, and it clearly shows that,

➢ Peoples dont use rented bikes in no functioning day.