Number of Bike Rented Prediction

Results

**With Outliers**

* **Without Removing outliers with all 12 variables**

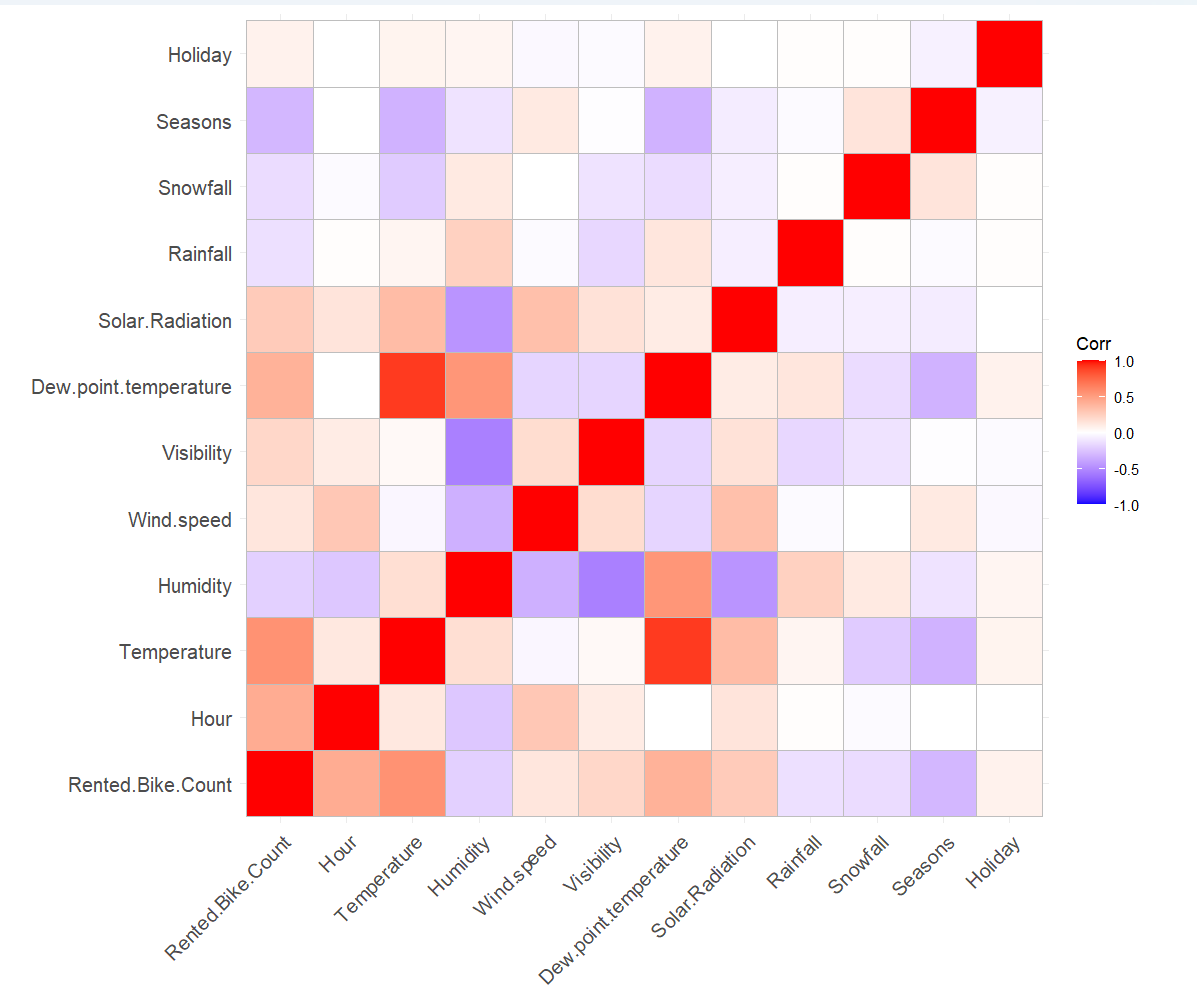
|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5562101 | 0.1625676 | 0.1159608 |
| **Random Forest** | 0.8360311 | 0.1019917 | 0.06444064 |
| **Decision Tree** | 0.6900172 | 0.1298076 | 0.09031582 |
| **Support Vector Regression** | 0.75701 | 0.1278503 | 0.07971049 |
| **Xgboost** | 0.8339927 | 0.09352312 | 0.0652922 |

* **PCA method**

**Without Rainfall, Holiday**

|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.525037 | 0.1253098 | 0.09430913 |
| **Random Forest** | 0.8632942 | 0.06756251 | 0.04333183 |
| **Decision Tree** | 0.6907592 | 0.1011512 | 0.07288018 |
| **Support Vector Regression** | 0.76581 | 0.08902849 | 0.05473097 |
| **Xgboost** | 0.8522732 | 0.06995845 | 0.04634666 |

* **Correlation based feature selection - 11 variables**

****

* **Removed Dew point temperature**

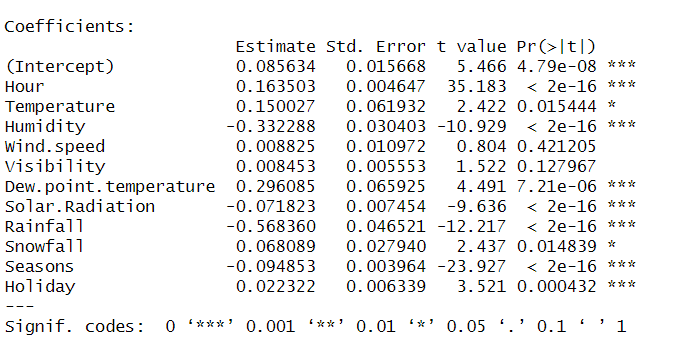
|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5370758 | 0.1213251 | 0.09062209 |
| **Random Forest** | 0.8624813 | 0.06635607 | 0.04261616 |
| **Decision Tree** | 0.6575196 | 0.1043817 | 0.07430606 |
| **Support Vector Regression** | 0.7542247 | 0.08887181 | 0.05605392 |
| **Xgboost** | 0.8546704 | 0.0679566 | 0.04554153 |

* **Removed Dew point temperature and humidity**

|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5163104 | 0.1264635 | 0.09493526 |
| **Random Forest** | 0.8555107 | 0.06944341 | 0.04443756 |
| **Decision Tree** | 0.681659 | 0.1025901 | 0.0734392 |
| **Support Vector Regression** | 0.7414528 | 0.09351784 | 0.05789927 |
| **Xgboost** | 0.8457627 | 0.07149336 | 0.04658759 |

* **10 variables - based on p value from mlr**

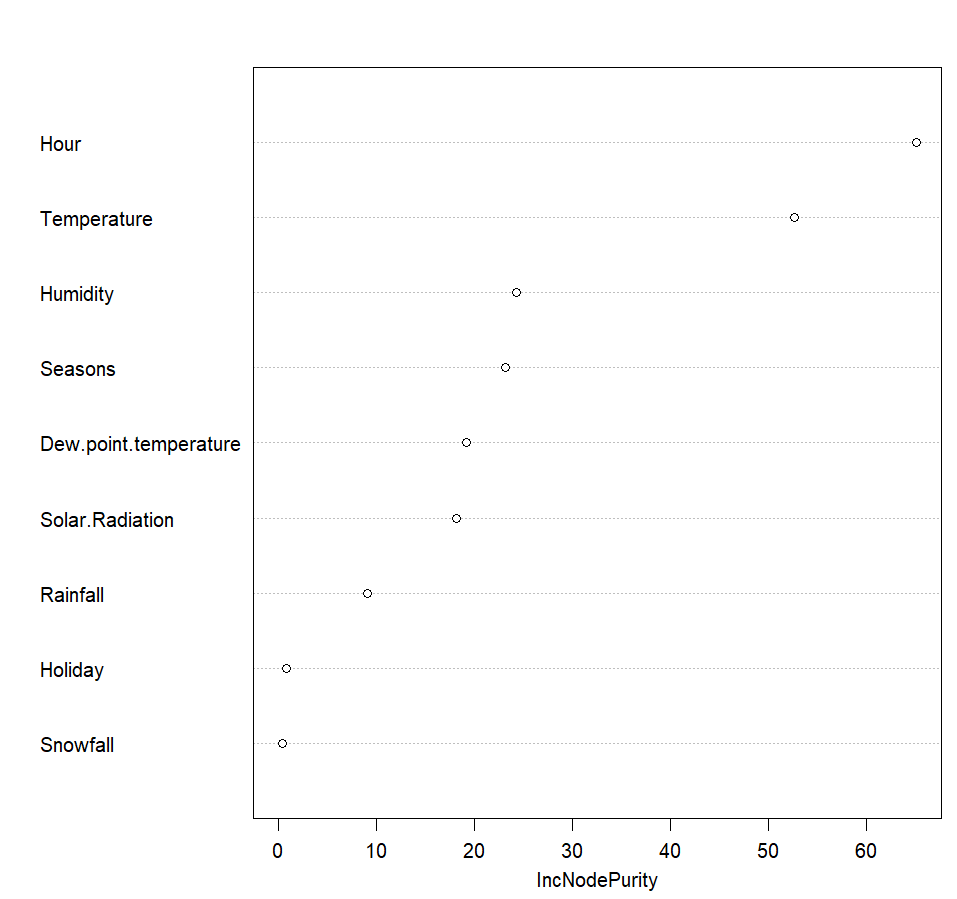
**Removed Visibility and Wind speed**

****

|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5398373 | 0.1233447 | 0.09292138 |
| **Random Forest** | 0.8694684 | 0.06586419 | 0.04214869 |
| **Decision Tree** | 0.6907592 | 0.1011512 | 0.07288018 |
| **Support Vector Regression** | 0.7757286 | 0.08731926 | 0.05376273 |
| **Xgboost** | 0.8648636 | 0.06688322 | 0.04450271 |

* **9 variables - based on varImp (IncNodePurity) from rf that gave max r square**

**Removed Snowfall**

****

|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5504721 | 0.1220056 | 0.09070551 |
| **Random Forest** | 0.8642006 | 0.06768718 | 0.04459111 |
| **Decision Tree** | 0.6570556 | 0.1068039 | 0.07604703 |
| **Support Vector Regression** | 0.7769461 | 0.08706164 | 0.05305995 |
| **Xgboost** | 0.8515197 | 0.07008074 | 0.04579884 |

* **8 variables - based on varImp (IncNodePurity) from rf that gave max r square**

**Removed Snowfall, Holiday, Visibility, Wind Speed**

|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.6281587 | 0.1109072 | 0.08858721 |
| **Random Forest** | 0.8808706 | 0.06310622 | 0.04386928 |
| **Decision Tree** | 0.6976142 | 0.09593143 | 0.07322589 |
| **Support Vector Regression** | 0.8598966 | 0.06353324 | 0.04423193 |
| **Xgboost** | 0.860088 | 0.06403106 | 0.04491506 |

**Without Outliers**

* **Without outliers**

|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5473801 | 0.1194195 | 0.09101804 |
| **Random Forest** | 0.8649464 | 0.06533112 | 0.04215836 |
| **Decision Tree** | 0.6599218 | 0.1034622 | 0.07462639 |
| **Support Vector Regression** | 0.7674741 | 0.08592157 | 0.053666867 |
| **Xgboost** | 0.8519141 | 0.06847069 | 0.04597584 |

* **Removing outliers & Snowfall, Holiday**

**varImp(Decision Tree)**

Overall

Dew\_point\_temperature 0.4081039

Hour 1.5270646

Humidity 0.9034749

Rainfall 0.6673689

Seasons 0.4266246

Solar\_Radiation 1.0006181

Temperature 0.8119720

Visibility 0.1504542

Wind\_speed 0.1632385

Snowfall 0.0000000

Holiday 0.0000000

|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5456778 | 0.1196279 | 0.09110341 |
| **Random Forest** | 0.8639102 | 0.06553144 | 0.04196065 |
| **Decision Tree** | 0.6599218 | 0.1034622 | 0.07462639 |
| **Support Vector Regression** | 0.7734043 | 0.08479327 | 0.05261507 |
| **Xgboost** | 0.848573 | 0.06927428 | 0.04662244 |

* **Removing outliers & Snowfall**

**varImp(Random Forest)**

Overall

Hour 199.31903

Temperature 70.62899

Humidity 75.97422

Wind\_speed 35.70376

Visibility 57.83856

Dew\_point\_temperature 43.42105

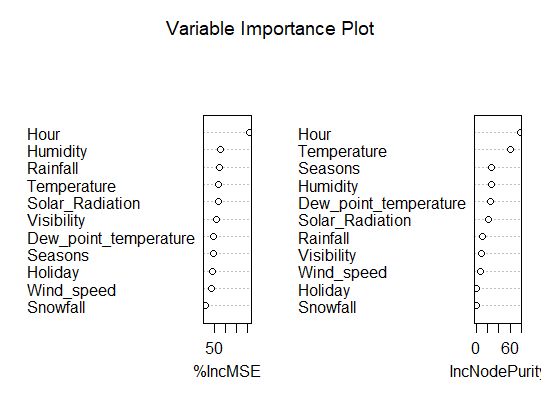
Solar\_Radiation 64.25378

Rainfall 63.80079

Snowfall 12.57202

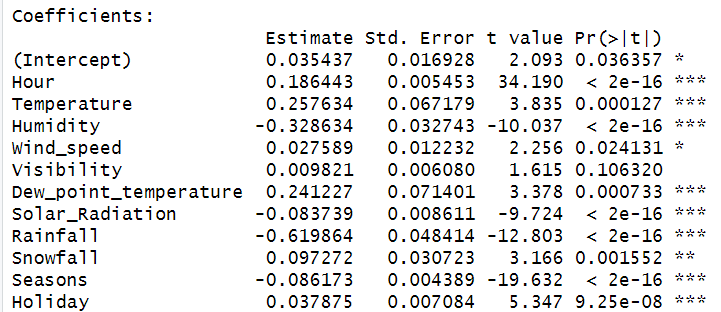
Seasons 46.79967

Holiday 38.50003



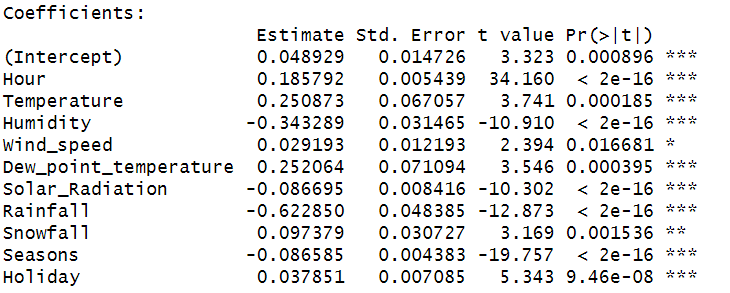
|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5469032 | 0.1194784 | 0.09104717 |
| **Random Forest** | 0.8652598 | 0.06522884 | 0.0419889 |
| **Decision Tree** | 0.6599218 | 0.1034622 | 0.07462639 |
| **Support Vector Regression** | 0.7712073 | 0.0852177 | 0.05302658 |
| **Xgboost** | 0.8485897 | 0.06923033 | 0.04627449 |

* **Without outliers & Visibility- based on p value from MLR feature selection**

****

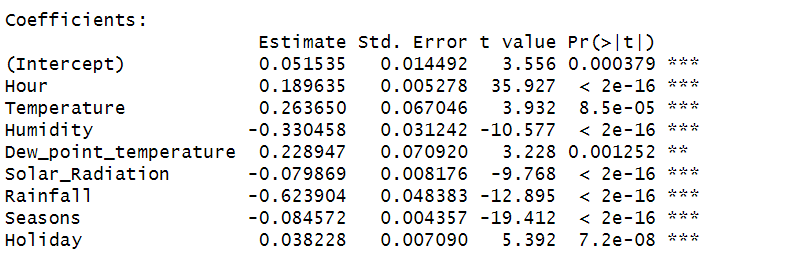
|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5469524 | 0.1194744 | 0.09108662 |
| **Random Forest** | 0.8649428 | 0.06529069 | 0.041873 |
| **Decision Tree** | 0.6599218 | 0.1034622 | 0.07462639 |
| **Support Vector Regression** | 0.7691268 | 0.08554957 | 0.05362547 |
| **Xgboost** | 0.8528036 | 0.06826666 | 0.04577217 |

* **Without outliers & Visibility, Wind Speed- based on p value from MLR feature selection**

****

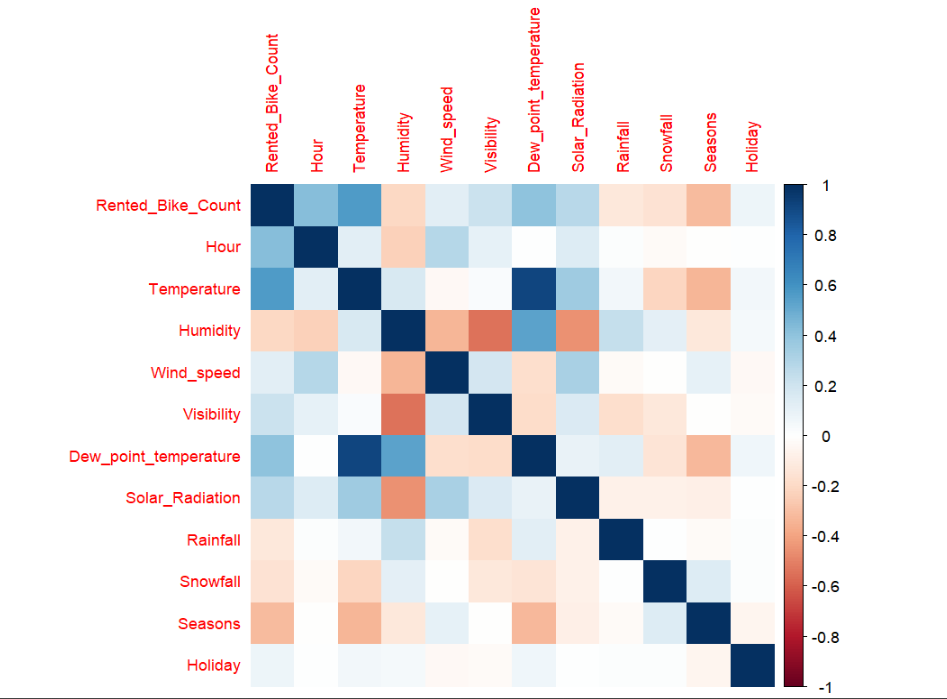
|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5458229 | 0.1196308 | 0.09130943 |
| **Random Forest** | 0.8632628 | 0.06566137 | 0.04163264 |
| **Decision Tree** | 0.6599218 | 0.1034622 | 0.7462639 |
| **Support Vector Regression** | 0.7706958 | 0.08534717 | 0.05331999 |
| **Xgboost** | 0.8516487 | 0.0685164 | 0.04564682 |

* **Without outliers & Visibility, Wind Speed, Snowfall- based on p value from MLR feature selection**

****

|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5453539 | 0.1196888 | 0.09134141 |
| **Random Forest** | 0.8602185 | 0.06651306 | 0.04381451 |
| **Decision Tree** | 0.6599218 | 0.1034622 | 0.074622639 |
| **Support Vector Regression** | 0.7754553 | 0.08443863 | 0.05253969 |
| **Xgboost** | 0.8508893 | 0.06871091 | 0.04566968 |

* **Without outliers & Temperature- Correlation graph**

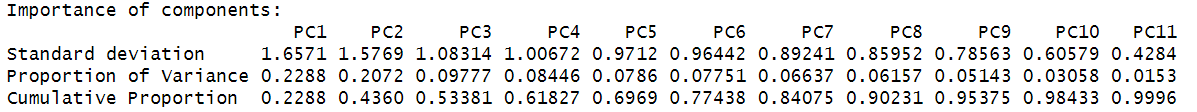


|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5454764 | 0.1196828 | 0.0909999 |
| **Random Forest** | 0.8575832 | 0.06712838 | 0.0439132 |
| **Decision Tree** | 0.6854849 | 0.09959865 | 0.6988999 |
| **Support Vector Regression** | 0.7589608 | 0.08739153 | 0.05524969 |
| **Xgboost** | 0.8280768 | 0.07385864 | 0.05112498 |

* **Without outliers & Dew Point Temperature- Correlation graph**

|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5472807 | 0.1194219 | 0.0911495 |
| **Random Forest** | 0.8679278 | 0.0645929 | 0.04180181 |
| **Decision Tree** | 0.6743593 | 0.1013429 | 0.07246293 |
| **Support Vector Regression** | 0.7672732 | 0.08587945 | 0.05381912 |
| **Xgboost** | 0.8359511 | 0.07202127 | 0.04883039 |

* **PCA (Principal Component Analysis)- Drop PC9,PC10,PC11**

****

|  | **R2** | **RMSE** | **MAE** |
| --- | --- | --- | --- |
| **MLR** | 0.5197979 | 0.1229673 | 0.09398925 |
| **Random Forest** | 0.8479512 | 0.06934761 | 0.04630832 |
| **Decision Tree** | 0.6576969 | 0.1037768 | 0.07483919 |
| **Support Vector Regression** | 0.7645784 | 0.08651047 | 0.05455491 |
| **Xgboost** | 0.8339409 | 0.07254684 | 0.04925406 |

CUBIST LightBGM