Project Report
Project: Bike Renting

Aparajita Mallick Sarkar

19.08.2020

Contents

1. <u>Introduction</u>

1.1. Problem Statement

The objective of this Case is to predict the daily bike rental count based on the environmental and seasonal settings. It is categorised under Regression Problem where regression model in to be built on the historical data.

2.1. Data

The historical data is provided, detailing the environmental and seasonal settings. There are 15 independent variables and one target variable, that is, 'cnt'. Total number of observations in 731. Below is a sample data used to predict bike counts.

Table 1.1: Sample Data (Columns: 1-9)

instant	dteday	season	yr	Mnth	holiday	weekday	workingday	weathersit
1	01-01-2011	1	0	1	0	6	0	2
2	02-01-2011	1	0	1	0	0	0	2
3	03-01-2011	1	0	1	0	1	1	1
4	04-01-2011	1	0	1	0	2	1	1
5	05-01-2011	1	0	1	0	3	1	1

Table 1.2: Sample Data (Columns: 10-16)

Temp	atemp	Hum	windspeed	casual	registered	cnt
0.344167	0.363625	0.805833	0.160446	331	654	985
0.363478	0.353739	0.696087	0.248539	131	670	801
0.196364	0.189405	0.437273	0.248309	120	1229	1349
0.2	0.212122	0.590435	0.160296	108	1454	1562
0.226957	0.22927	0.436957	0.1869	82	1518	1600

The details of data attributes in the dataset are as follows:

- instant: Record index
- dteday: Date
- season: Season (1:springer, 2:summer, 3:fall, 4:winter)
- yr: Year (0: 2011, 1:2012)
- mnth: Month (1 to 12)
- holiday: whether day is holiday or not (extracted from Holiday Schedule)
- weekday: Day of the week
- workingday: If day is neither weekend nor holiday is 1, otherwise is 0.
- weathersit: (extracted from Freemeteo)

1: Clear, Few clouds, Partly cloudy, Partly cloudy
2: Mist + Cloudy, Mist + Broken clouds, Mist + Few clouds, Mist
3: Light Snow, Light Rain + Thunderstorm + Scattered clouds, Light Rain + Scattered clouds
4: Heavy Rain + Ice Pallets + Thunderstorm + Mist, Snow + Fog

- temp: Normalized temperature in Celsius. The values are derived via (t-t_min)/(t_max-t_min), t_min=-8, t_max=+39 (only in hourly scale)
- atemp: Normalized feeling temperature in Celsius. The values are derived via (t-t_min)/(t_maxt_min), t_min=-16, t_max=+50 (only in hourly scale)
- hum: Normalized humidity. The values are divided to 100 (max)
- windspeed: Normalized wind speed. The values are divided to 67 (max)
- casual: count of casual users
- registered: count of registered users
- cnt: count of total bikes rented including both casual and registered users. This is the target variable.

2. <u>Methodology</u>

2.1. Pre-Processing

Data Pre-processing is the step in which the data gets transformed, to bring it to such a state that the machine can easily understand it. In other words, the features of the data can then be easily interpreted by the algorithm. It includes exploring the data, cleaning the data as well as visualizing the data through graphs and plots. This is often called as Exploratory Data Analysis.

The following steps are taken under the pre-processing:

a. Deleted the variable 'instant' as it acts as an index
b. Deleted the variables 'casual' and 'registered' as we are interested in predicting the total count of bikes rented that is the variable 'cnt'. Here cnt is the sum of 'casual' and 'registered'.
c. Converting the data types of 'season', 'yr', 'mnth', 'holiday', 'weekday', 'workingday' and 'weathersit 'to category.
d. As the mnth and yr already denotes the month and the year respectively, we derive the date part from the 'dteday' variable and save it as a categorical variable with levels from 1 to 31.

After the above mentioned steps there are 731 observations and 13 variables.

2.1.1. Missing Value Analysis

Missing Value analysis is done to check for any missing values in the dataset. Missing values can be dealt in the following ways:

a. Dropping the variables in which the values are missing
b. Dropping the observations where the values are missing
c. In case the number of missing values in a variable is less than 30% of the total count of observation, the values can be imputed using central statistics, distance based methods like knn or prediction methods.

It is observed that there are no missing values in the data set.

Table 2.1: No. of missing values in each variables.

	Variables	Missing_Values
0	dteday	0
1	season	0
2	yr	0
3	mnth	0
4	holiday	0
5	weekday	0
6	workingday	0
7	weathersit	0
8	temp	0
9	hum	0
10	windspeed	0
11	cnt	0

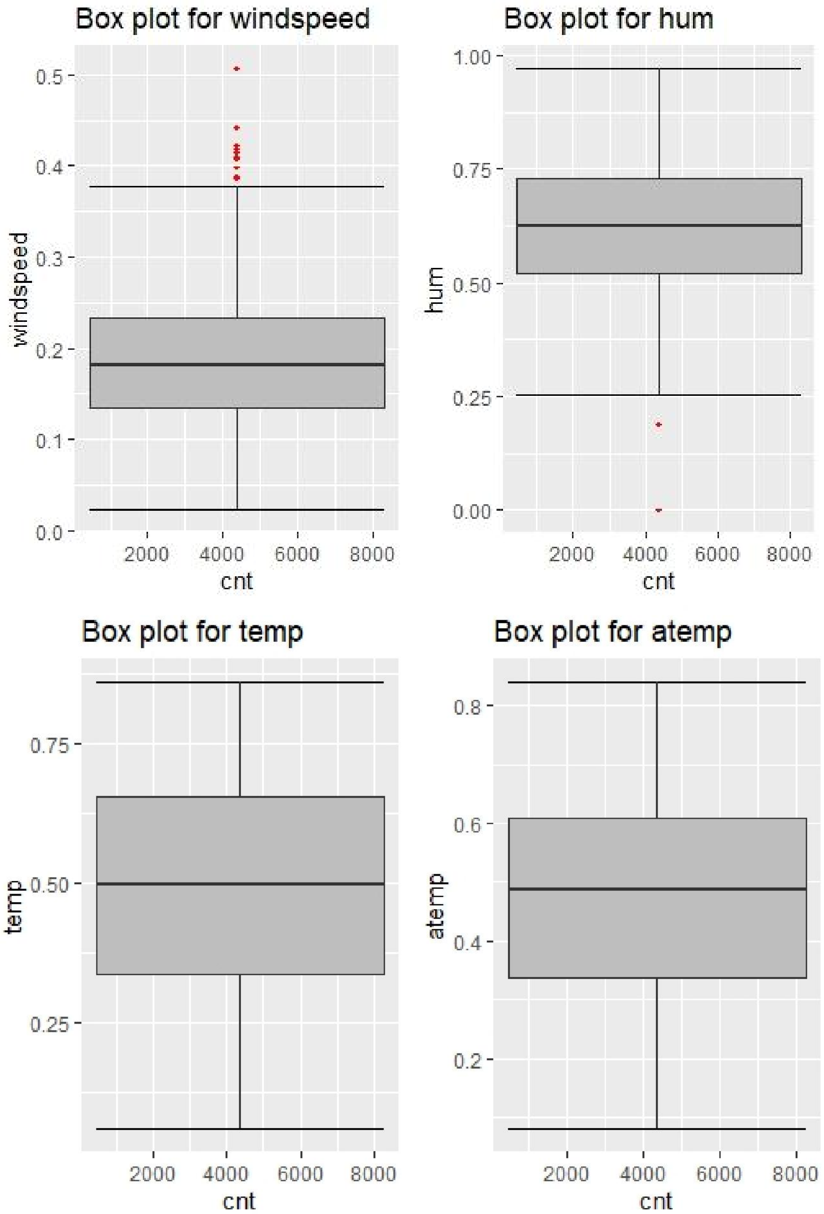
2.1.2. Outlier Analysis

Outliers are observations which are inconsistent with rest of the data set and are defined only for continuous variables. In the given data set the following are the continuous independent variables:

a. Temp
b. Atemp
c. Hum
d. Windspeed

Boxplot is used to detect outliers.

Fig 2.1: Box Plots for Outlier Detection



As it can be seen outliers are present in ‘hum’ and ‘windspeed’. As the variables define humidity and windspeed, it can be considered that the outliers are a result of some unexpected weather condition. Thus the observation with the outliers is deleted reducing the number of observations to 717.

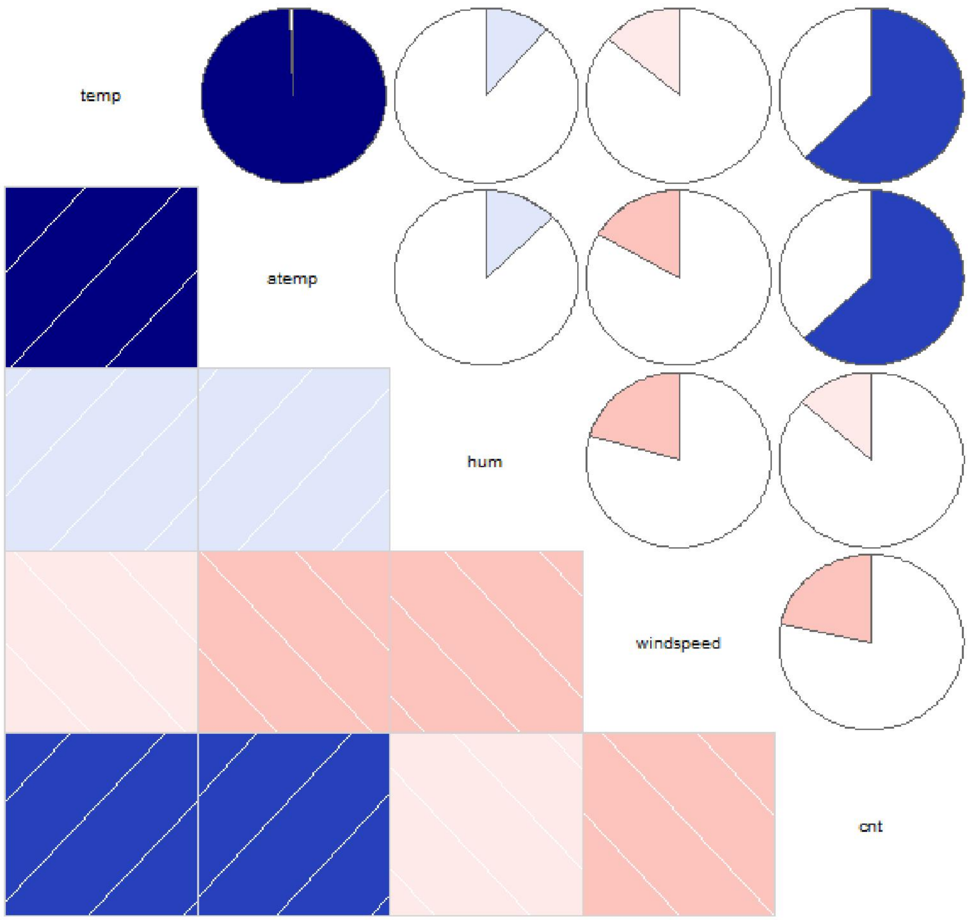
2.1.3. Feature selection

Feature selection includes extracting meaningful feature from the dataset. This includes removing the unimportant data thus reducing the complexity of the model. It is done considering that the independent variables should have minimum correlation between them and the correlation between the target variable and the independent variables should be high.

Table 2.2: Correlation values for Continuous variables, calculated in python

	temp	atemp	hum	windspeed	cnt
temp	1.000000	0.991738	0.114191	-0.140169	0.625892
atemp	0.991738	1.000000	0.126587	-0.166038	0.629204
hum	0.114191	0.126587	1.000000	-0.204496	-0.136621
windspeed	-0.140169	-0.166038	-0.204496	1.000000	-0.216193
cnt	0.625892	0.629204	-0.136621	-0.216193	1.000000

Fig. 2.2: Correlation Plot in R



Plotting the correlation plot in R and python shows high correlation between 'temp' and 'atemp', that is 0.991738. Thus it is unnecessary to have both the variables. Thus we can keep either of them.

2.1.4. Dimension reduction

On the basis of correlation analysis of the continuous variables, the variable 'atemp' is dropped.

2.1.5. Data after pre-processing

After the Pre-processing, the data is as followed:

a. 12 variables
b. 717 observations
c. 11 independent variable and 1 target variable
d. 'dteday', 'season', 'yr', 'mnth', 'holiday', 'weekday', 'workingday' and 'weathersit' are categorical variables.
e. 'temp', 'hum', 'windspeed' and 'cnt' are the continuous variables.

2.2. Modelling

2.2.1. Model Selection

The problem is categorised under Regression Problem where a continuous target variable, 'cnt', is to be predicted using the historical data. Thus various regression models are built on the historical data. The motive of building of the model is to know how well the model predicts the new data rather than how well it fits the data it was trained on. Thus the historical data is divided into 'Train Data' on which the model is built and the 'Test Data' which is used to compare the actual value to that of the predicted values by the model, to estimate the performance of the model. The difference between the actual value and the predicted value is the 'error'. Low error is desirable.

Following steps are taken for each of the model:

- 80 % of the data is selected by simple random sampling as the 'Train Data' and the rest of the 20% as the 'Test Data'.
- Model is built on the Train Data.
- Prediction is made using the model for the 'Test Data'.
- The predictions are in float, thus it is converted to integer to truncate the decimal part as the count of bikes has to be integers.
- The predictions are stored and further used to calculate MAPE.

Following regression models are built:

- Decision tree for regression target variable
- Random Forest (with 200 trees)
- Multiple Linear Regression

2.2.2. Decision Tree

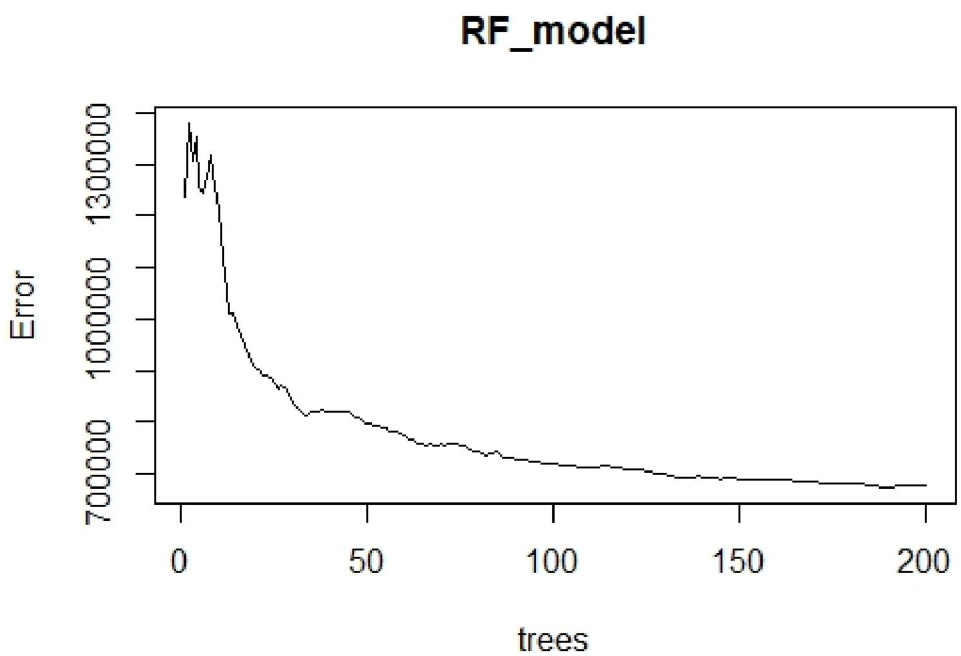
A decision tree is a predictive model based on branching series of Boolean Tests and can be used for classification and regression problems. It has a flow chart type structure. To build the model, the CART(Classification and Regression Tree) algorithm is used. Here the nodes have a binary split.
The R implementation of CART is called as RPART(Recursive Partitioning And Regression Trees) available in the package with the same name. CART implementation for regression problem in python has been done using DecisionTreeRegressor from sklearn library.

2.2.3. Random Forest

Random Forest is an ensemble technique that includes multiple decision trees. Ensemble techniques use multiple learning algorithms back to back to obtain better predictive performance than could be obtained from any one of the constituent learning algorithms alone. Random forests can be used for both regression and classification problems.

While implementing Random Forest we can specify the number of trees to be created. Generally, the more the number of trees the better is the accuracy. If the number of trees to be created is not defined, decision trees are created and checked for error. The number of trees are created till the time the error no longer reduces. In the below plot it can be seen that error reduces no further after 200 trees. Thus we consider 200 trees in the Random Forest.

Fig 2.3: Plot of Random Forest in R



In R, randomForest() function from the library of the same name is used. In Python, RandomForestRegressor from the sklearn library is used.

2.2.4. Multiple Linear Regression

Multiple Linear Regression is a statistical model used for regression problems, where the coefficients or weights are calculated for each of the independent variable. This coefficient explains the amount of information the independent variable carries to explain the target variable.

It can be represented as below:

$$y = b_0 + b_1x_1 + b_2x_2 + \ldots\ldots + b_nx_n$$
$$b_0 \rightarrow \text{Intercept}$$
$$b_1, b_2, \ldots., b_n \rightarrow \text{Coefficients of independent variable}$$
$$x_1, x_2, \ldots., x_n \rightarrow \text{Independent variables}$$
$$y \rightarrow \text{Target Variable}$$

In python, dummy variables for each category of the categorical variables is created, thus having total of 63 independent variables on which we apply the linear regression model by using the ols() function from the statsmodels library. This calculates the weight of each category of the categorical variables.

Creating dummy variables is not required in R, as the lm() function is used which itself calculates the weights of all the categories of the categorical variables separately.

Table 2.3: Parameters calculated for Linear Regression Model in R

Multiple R-squared	0.8563	Adjusted R-squared	0.8404
F-statistic	53.85	p-value	< 2.2e-16

As can be seen by the Adjusted R-squared value, the model explains 84.04% of the variance of the target variable, which is acceptable. F-statistic is 53.85, which shows a good relationship between the independent and the target variables. Also from the combined p-value it can be inferred that the null hypothesis can be rejected. (Null Hypothesis, Ho-> Target variable doesn't depend on predictor variables). Indicating it is a good model.

Table 2.4: Parameters calculated for Linear Regression Model in Python

Multiple R-squared	0.863	Adjusted R-squared	0.847
F-statistic	55.17	p-value	1.55e-180

As can be seen by the Adjusted R-squared value, the model explains 84.7% of the variance of the target variable, which is acceptable. F-statistic is 55.17, which shows a good relationship between the independent and the target variables. Also from the combined p-value it can be inferred that the null hypothesis can be rejected. (Null Hypothesis, Ho-> Target variable doesn't depend on predictor variables). Indicating it is a good model.

3. Conclusion

3.1. Model Evaluation

One of the measures of errors is Mean Absolute Percentage Error (MAPE), which is defined as follows:

$$MAPE = \left(\frac{1}{\mathrm{n}}\right) \left|\frac{\sum_{t=1}^{n}(\mathrm{A_t} - \mathrm{F_t})}{\mathrm{A_t}}\right| * 100$$

$$\mathrm{A_t} \rightarrow \text{Actual Value}$$
$$\mathrm{F_t} \rightarrow \text{Predicted Value}$$
$$n \rightarrow \text{Number of observations}$$

Thus the lower MAPE is desirable while selecting model.

Following are the MAPE calculated in R:
- MAPE of Decision Tree= 21.25035%
- MAPE of Random Forest= 21.88941%
- MAPE of Multiple Linear Regression= 17.00978%

Following are the MAPE calculated in Python
- MAPE of Decision Tree= 26.511570167563264 %
- MAPE of Random Forest= 13.926401687733211 %
- MAPE of Multiple Linear Regression= 17.917377200794842 %

3.2. Model Selection

It is seen in R, the Linear Regression Model has the least MAPE of 17.00%. This gives an accuracy of 83%. Thus in R the Linear Regression Model is selected and the output for the test data is stored in LR_op_R.csv.
In Python, the Random Forest Model has the least MAPE of 13.92%. This gives an accuracy of 86.08%. Thus in Python the Random Forest Model is selected and the output for the test data is stored in RF_op_python.csv.

Appendix I: Summary of Multiple Linear Regression Model in R

```
Call:
lm(formula = cnt ~ ., data = train)

Residuals:
    Min      1Q  Median      3Q     Max 
-3589.8  -377.5    77.0   471.6  2514.3 

Coefficients: (1 not defined because of singularities)
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  1645.337    340.557   4.831 1.79e-06 ***
dteday02      103.234    270.083   0.382 0.702449    
dteday03      234.034    246.498   0.949 0.342845    
dteday04      281.820    276.244   1.020 0.308121    
dteday05      195.889    244.160   0.802 0.422750    
dteday06      220.703    256.903   0.859 0.390691    
dteday07      -79.477    254.044  -0.313 0.754524    
dteday08      -26.535    271.241  -0.098 0.922108    
dteday09       90.511    248.842   0.364 0.716209    
dteday10      221.743    270.508   0.820 0.412751    
dteday11      349.584    250.026   1.398 0.162657    
dteday12      186.197    250.981   0.742 0.458501    
dteday13      192.176    262.532   0.732 0.464494    
dteday14      183.395    249.467   0.735 0.462584    
dteday15      407.182    244.122   1.668 0.095934 .  
dteday16      379.090    247.385   1.532 0.126040    
dteday17      548.908    250.200   2.194 0.028690 *  
dteday18      133.489    256.855   0.520 0.603493    
dteday19      282.723    282.227   1.002 0.316930    
dteday20      326.106    255.475   1.276 0.202365    
dteday21      371.519    249.019   1.492 0.136330    
dteday22      -94.602    248.888  -0.380 0.704029    
dteday23       65.358    260.711   0.251 0.802153    
dteday24     -243.444    246.115  -0.989 0.323056    
dteday25      117.037    253.256   0.462 0.644182    
dteday26      185.644    257.868   0.720 0.471901    
dteday27       -5.893    258.492  -0.023 0.981820    
dteday28     -106.472    261.585  -0.407 0.684159    
dteday29     -318.264    251.314  -1.266 0.205942    
dteday30     -209.500    250.457  -0.836 0.403278    
dteday31      236.381    283.262   0.834 0.404389    
season2       884.810    212.025   4.173 3.53e-05 ***
season3       823.410    242.447   3.396 0.000736 ***
season4      1551.310    205.601   7.545 2.06e-13 ***
yr1          2032.838     68.096  29.853  < 2e-16 ***
mnth2         194.510    170.256   1.142 0.253794    
mnth3         605.631    199.931   3.029 0.002575 ** 
mnth4         538.476    286.888   1.877 0.061089 .  
mnth5         771.716    310.462   2.486 0.013246 *  
mnth6         623.533    324.541   1.921 0.055248 .  
mnth7         103.087    356.129   0.289 0.772342    
mnth8         501.886    344.840   1.455 0.146163    
mnth9        1118.605    297.174   3.764 0.000186 ***
mnth10        624.360    274.445   2.275 0.023316 *  
mnth11         61.895    265.918   0.233 0.816041    
mnth12         10.413    208.203   0.050 0.960131    
holiday1     -706.442    207.910  -3.398 0.000732 ***
weekday1       78.902    129.852   0.608 0.543701    
weekday2      178.418    127.806   1.396 0.163315    
weekday3      331.412    124.242   2.667 0.007883 ** 
weekday4      285.664    128.290   2.227 0.026398 *  
weekday5      417.299    124.444   3.353 0.000857 ***
weekday6      380.322    126.430   3.008 0.002757 ** 
workingday1        NA         NA      NA       NA    
weathersit2  -431.999     96.408  -4.481 9.16e-06 ***
```

```
weathersit3 -1963.601    239.926  -8.184 2.16e-15 ***
temp         4245.195    487.948   8.700  < 2e-16 ***
hum         -1755.239    376.927  -4.657 4.09e-06 ***
windspeed   -2849.104    521.127  -5.467 7.14e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 783.2 on 515 degrees of freedom
Multiple R-squared:  0.8563,	Adjusted R-squared:  0.8404 
F-statistic: 53.85 on 57 and 515 DF,  p-value: < 2.2e-16
```

Appendix II: Summary of Multiple Linear Regression Model in Python

OLS Regression Results

Dep. Variable:	cnt	**R-squared:**	0.863
Model:	OLS	**Adj. R-squared:**	0.847
Method:	Least Squares	**F-statistic:**	55.17
Date:	Wed, 19 Aug 2020	**Prob (F-statistic):**	1.55e-180
Time:	03:14:27	**Log-Likelihood:**	-4472.5
No. Observations:	559	**AIC:**	9061.
Df Residuals:	501	**BIC:**	9312.
Df Model:	57		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
yr	2021.6172	67.016	30.166	0.000	1889.949	2153.285
holiday	237.3595	183.669	1.292	0.197	-123.496	598.215
workingday	670.8007	92.074	7.285	0.000	489.903	851.699
temp	4584.5749	474.136	9.669	0.000	3653.034	5516.116
hum	-1958.0551	360.849	-5.426	0.000	-2667.018	-1249.092
windspeed	-3156.8101	505.092	-6.250	0.000	-4149.170	-2164.451
season_1	-320.9896	157.511	-2.038	0.042	-630.452	-11.527
season_2	559.9716	154.681	3.620	0.000	256.069	863.874
season_3	779.0421	169.647	4.592	0.000	445.734	1112.350
season_4	1340.8380	173.575	7.725	0.000	999.813	1681.863

dteday_01	-230.0216	170.234	-1.351	0.177	-564.481	104.438
dteday_02	-64.4425	183.435	-0.351	0.726	-424.840	295.955
dteday_03	175.6987	173.968	1.010	0.313	-166.097	517.495
dteday_04	284.4375	202.171	1.407	0.160	-112.770	681.645
dteday_05	197.6954	178.743	1.106	0.269	-153.483	548.873
dteday_06	223.8940	179.865	1.245	0.214	-129.489	577.277
dteday_07	35.1255	169.711	0.207	0.836	-298.307	368.557
dteday_08	-131.8175	168.814	-0.781	0.435	-463.489	199.854
dteday_09	48.4258	169.556	0.286	0.775	-284.703	381.554
dteday_10	138.7625	169.103	0.821	0.412	-193.477	471.002
dteday_11	329.4380	164.874	1.998	0.046	5.509	653.367
dteday_12	137.8393	189.367	0.728	0.467	-234.213	509.891
dteday_13	152.2271	179.405	0.849	0.397	-200.252	504.706
dteday_14	170.3466	169.458	1.005	0.315	-162.590	503.283
dteday_15	450.8459	190.863	2.362	0.019	75.856	825.836
dteday_16	236.8906	173.993	1.361	0.174	-104.956	578.737
dteday_17	510.1627	174.882	2.917	0.004	166.571	853.754
dteday_18	180.2131	196.090	0.919	0.359	-205.047	565.473
dteday_19	164.0723	165.579	0.991	0.322	-161.242	489.386

dteday_20	287.0053	173.800	1.651	0.099	-54.462	628.473
dteday_21	274.4157	184.865	1.484	0.138	-88.791	637.623
dteday_22	-280.1661	186.546	-1.502	0.134	-646.676	86.344
dteday_23	-59.9488	174.270	-0.344	0.731	-402.339	282.442
dteday_24	-104.4899	174.556	-0.599	0.550	-447.442	238.462
dteday_25	-114.1619	169.716	-0.673	0.501	-447.606	219.282
dteday_26	-65.2660	178.630	-0.365	0.715	-416.223	285.691
dteday_27	-189.5462	180.678	-1.049	0.295	-544.527	165.434
dteday_28	-109.4532	190.326	-0.575	0.565	-483.388	264.482
dteday_29	-496.5715	180.132	-2.757	0.006	-850.480	-142.663
dteday_30	-123.9182	195.612	-0.633	0.527	-508.239	260.402
dteday_31	331.1697	220.338	1.503	0.133	-101.731	764.070
weathersit_1	1596.3116	102.868	15.518	0.000	1394.206	1798.417
weathersit_2	1149.4967	129.896	8.849	0.000	894.289	1404.705
weathersit_3	-386.9462	229.309	-1.687	0.092	-837.471	63.579
mnth_1	-74.5913	195.277	-0.382	0.703	-458.253	309.071
mnth_2	81.0127	185.417	0.437	0.662	-283.278	445.303
mnth_3	516.9163	144.150	3.586	0.000	233.704	800.128
mnth_4	483.8973	177.747	2.722	0.007	134.676	833.118

mnth_5	703.7222	187.687	3.749	0.000	334.971	1072.473
mnth_6	335.5922	180.075	1.864	0.063	-18.204	689.388
mnth_7	-314.8764	214.161	-1.470	0.142	-735.641	105.889
mnth_8	80.2696	199.694	0.402	0.688	-312.070	472.609
mnth_9	662.8288	163.688	4.049	0.000	341.230	984.428
mnth_10	321.6010	181.577	1.771	0.077	-35.145	678.347
mnth_11	-203.2090	190.619	-1.066	0.287	-577.720	171.302
mnth_12	-234.3012	167.671	-1.397	0.163	-563.727	95.124
weekday_0	489.4840	126.932	3.856	0.000	240.099	738.869
weekday_1	-18.7033	84.185	-0.222	0.824	-184.103	146.696
weekday_2	164.3126	87.394	1.880	0.061	-7.392	336.017
weekday_3	264.4015	90.688	2.916	0.004	86.226	442.577
weekday_4	283.3881	88.721	3.194	0.001	109.077	457.699
weekday_5	214.7613	84.527	2.541	0.011	48.691	380.831
weekday_6	961.2180	121.243	7.928	0.000	723.010	1199.426

Omnibus:	72.512	**Durbin-Watson:**	1.415
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	159.468
Skew:	-0.715	**Prob(JB):**	2.35e-35
Kurtosis:	5.191	**Cond. No.**	1.49e+16

	coef	std err	t	P>|t|	[0.025	0.975]
yr	1975.6288	65.047	30.372	0.000	1847.826	2103.432
holiday	96.5565	170.037	0.568	0.570	-237.526	430.639
workingday	700.9029	90.662	7.731	0.000	522.773	879.033
temp	4229.2791	483.858	8.741	0.000	3278.610	5179.948
hum	-1821.0286	349.133	-5.216	0.000	-2506.994	-1135.063
windspeed	-2454.1473	486.087	-5.049	0.000	-3409.195	-1499.100
season_1	-277.5246	157.195	-1.765	0.078	-586.377	31.328
season_2	619.0050	159.462	3.882	0.000	305.700	932.310
season_3	556.6920	172.718	3.223	0.001	217.340	896.044
season_4	1444.8710	167.333	8.635	0.000	1116.101	1773.641
dteday_01	-187.7470	159.131	-1.180	0.239	-500.403	124.909
dteday_02	80.7300	176.313	0.458	0.647	-265.684	427.144
dteday_03	183.8695	171.060	1.075	0.283	-152.225	519.964
dteday_04	322.4527	176.732	1.825	0.069	-24.785	669.690
dteday_05	137.1227	170.978	0.802	0.423	-198.809	473.055
dteday_06	227.3658	148.828	1.528	0.127	-65.047	519.778
dteday_07	-20.5889	187.327	-0.110	0.913	-388.642	347.464
dteday_08	-133.7719	170.308	-0.785	0.433	-468.388	200.844
dteday_09	70.1351	171.692	0.408	0.683	-267.199	407.470
dteday_10	170.9158	170.687	1.001	0.317	-164.445	506.277
dteday_11	171.7904	169.834	1.012	0.312	-161.895	505.476
dteday_12	252.2981	171.108	1.474	0.141	-83.889	588.485
dteday_13	118.3908	158.659	0.746	0.456	-193.338	430.120
dteday_14	20.2776	158.022	0.128	0.898	-290.198	330.754
dteday_15	303.7314	177.494	1.711	0.088	-45.004	652.466
dteday_16	248.6039	161.900	1.536	0.125	-69.492	566.700
dteday_17	346.4479	188.770	1.835	0.067	-24.442	717.338
dteday_18	-37.1311	176.425	-0.210	0.833	-383.765	309.503

dteday_19	235.3860	166.134	1.417	0.157	-91.029	561.801
dteday_20	265.5072	187.843	1.413	0.158	-103.561	634.575
dteday_21	284.1737	163.824	1.735	0.083	-37.703	606.050
dteday_22	-143.5600	161.349	-0.890	0.374	-460.573	173.453
dteday_23	79.5271	181.314	0.439	0.661	-276.713	435.767
dteday_24	-164.6041	167.066	-0.985	0.325	-492.850	163.641
dteday_25	122.1906	167.770	0.728	0.467	-207.439	451.820
dteday_26	130.7514	158.734	0.824	0.410	-181.124	442.626
dteday_27	-233.4749	184.072	-1.268	0.205	-595.135	128.185
dteday_28	-26.8576	194.927	-0.138	0.890	-409.844	356.129
dteday_29	-569.4312	182.724	-3.116	0.002	-928.442	-210.420
dteday_30	6.3500	188.459	0.034	0.973	-363.928	376.628
dteday_31	82.1925	220.320	0.373	0.709	-350.685	515.070
weathersit_1	1549.5737	98.989	15.654	0.000	1355.083	1744.064
weathersit_2	1139.7232	120.111	9.489	0.000	903.732	1375.714
weathersit_3	-346.2535	219.136	-1.580	0.115	-776.806	84.299
mnth_1	-144.5279	193.808	-0.746	0.456	-525.316	236.260
mnth_2	-20.5732	184.957	-0.111	0.911	-383.970	342.824
mnth_3	321.7423	147.654	2.179	0.030	31.637	611.848
mnth_4	258.1476	179.350	1.439	0.151	-94.233	610.528
mnth_5	627.1998	184.187	3.405	0.001	265.314	989.085
mnth_6	329.0795	184.349	1.785	0.075	-33.124	691.283
mnth_7	31.9070	212.611	0.150	0.881	-385.824	449.638
mnth_8	397.2540	201.352	1.973	0.049	1.645	792.864
mnth_9	942.1029	163.202	5.773	0.000	621.448	1262.757
mnth_10	311.2818	182.918	1.702	0.089	-48.109	670.673
mnth_11	-400.9409	185.100	-2.166	0.031	-764.620	-37.262
mnth_12	-309.6295	166.769	-1.857	0.064	-637.293	18.033
weekday_0	601.8075	121.915	4.936	0.000	362.272	841.343

weekday_1	35.0356	81.879	0.428	0.669	-125.838	195.910
weekday_2	221.0080	85.801	2.576	0.010	52.430	389.586
weekday_3	234.6507	86.791	2.704	0.007	64.127	405.174
weekday_4	126.9013	84.419	1.503	0.133	-38.962	292.764
weekday_5	179.8639	83.438	2.156	0.032	15.928	343.800
weekday_6	943.7764	117.932	8.003	0.000	712.067	1175.486

Omnibus:	48.010	**Durbin-Watson:**	1.407
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	83.166
Skew:	-0.570	**Prob(JB):**	8.72e-19
Kurtosis:	4.520	**Cond. No.**	1.47e+16