AUTOCORRELATION

The dataset shows the historical data on avocado prices, sales and sales volume in multiple US markets.  It represents weekly 2015 retail scan data for national retail volume units and their Average price, and Avocado sold volume. We will try to find out whether this timeseries data has any Autocorrelation problem.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| year | Date | Average Price | Total Volume | Total Bags | Sales |
| 2015 | 27-12-2015 | 1.33 | 64236.62 | 8696.87 | 85434.7046 |
| 2015 | 20-12-2015 | 1.35 | 54876.98 | 9505.56 | 74083.923 |
| 2015 | 13-12-2015 | 0.93 | 118220.22 | 8145.35 | 109944.8046 |
| 2015 | 06-12-2015 | 1.08 | 78992.15 | 5811.16 | 85311.522 |
| 2015 | 29-11-2015 | 1.28 | 51039.6 | 6183.95 | 65330.688 |
| 2015 | 22-11-2015 | 1.26 | 55979.78 | 6683.91 | 70534.5228 |
| 2015 | 15-11-2015 | 0.99 | 83453.76 | 8318.86 | 82619.2224 |
| 2015 | 08-11-2015 | 0.98 | 109428.33 | 6829.22 | 107239.7634 |
| 2015 | 01-11-2015 | 1.02 | 99811.42 | 11388.36 | 101807.6484 |
| 2015 | 25-10-2015 | 1.07 | 74338.76 | 8625.92 | 79542.4732 |
| 2015 | 18-10-2015 | 1.12 | 84843.44 | 8205.66 | 95024.6528 |
| 2015 | 11-10-2015 | 1.28 | 64489.17 | 10123.9 | 82546.1376 |
| 2015 | 04-10-2015 | 1.31 | 61007.1 | 8756.75 | 79919.301 |
| 2015 | 27-09-2015 | 0.99 | 106803.39 | 6034.46 | 105735.3561 |
| 2015 | 20-09-2015 | 1.33 | 69759.01 | 9267.36 | 92779.4833 |
| 2015 | 13-09-2015 | 1.28 | 76111.27 | 9286.68 | 97422.4256 |
| 2015 | 06-09-2015 | 1.11 | 99172.96 | 7990.1 | 110081.9856 |
| 2015 | 30-08-2015 | 1.07 | 105693.84 | 10306.73 | 113092.4088 |
| 2015 | 23-08-2015 | 1.34 | 79992.09 | 10880.36 | 107189.4006 |
| 2015 | 16-08-2015 | 1.33 | 80043.78 | 10443.22 | 106458.2274 |
| 2015 | 09-08-2015 | 1.12 | 111140.93 | 9225.89 | 124477.8416 |
| 2015 | 02-08-2015 | 1.45 | 75133.1 | 11847.02 | 108942.995 |
| 2015 | 26-07-2015 | 1.11 | 106757.1 | 13192.69 | 118500.381 |
| 2015 | 19-07-2015 | 1.26 | 96617 | 11287.48 | 121737.42 |
| 2015 | 12-07-2015 | 1.05 | 124055.31 | 24431.9 | 130258.0755 |
| 2015 | 05-07-2015 | 1.35 | 109252.12 | 29898.96 | 147490.362 |
| 2015 | 28-06-2015 | 1.37 | 89534.81 | 26662.08 | 122662.6897 |
| 2015 | 21-06-2015 | 1.27 | 104849.39 | 21875.65 | 133158.7253 |
| 2015 | 14-06-2015 | 1.32 | 89631.3 | 29002.59 | 118313.316 |
| 2015 | 07-06-2015 | 1.07 | 122743.06 | 22775.21 | 131335.0742 |
| 2015 | 31-05-2015 | 1.23 | 95123.62 | 23681.01 | 117002.0526 |
| 2015 | 24-05-2015 | 1.19 | 101470.91 | 29355.13 | 120750.3829 |
| 2015 | 17-05-2015 | 1.43 | 109857.47 | 26657.44 | 157096.1821 |
| 2015 | 10-05-2015 | 1.26 | 120427.91 | 16821.3 | 151739.1666 |
| 2015 | 03-05-2015 | 1.2 | 59197.67 | 12570.51 | 71037.204 |
| 2015 | 26-04-2015 | 1.22 | 49585.46 | 12778.44 | 60494.2612 |
| 2015 | 19-04-2015 | 1.19 | 49064.73 | 14301.92 | 58387.0287 |
| 2015 | 12-04-2015 | 1.13 | 48364.29 | 17104.37 | 54651.6477 |
| 2015 | 05-04-2015 | 1.16 | 47362.13 | 10728.94 | 54940.0708 |
| 2015 | 29-03-2015 | 1.02 | 67799.08 | 7684.08 | 69155.0616 |

**MODEL BUILDING: -**

Here we are taking the sales as dependent variable and the independent variables as Total Volume and Total Bags sold since in general the sales will be dependent on these variables.

**DESCRIPTIVE STATISTICS: -**



**CORRELATION: -**



As we can see there is high degree of correlation between sales of avocado and the total volume of sales, while there is only possibility of correlation between the sales Vs number of bags sold and total volume Vs total bags.

**SCATTER PLOT: -**



**LINE & SYMBOL PLOT: -**



**REGRESSION OUTPUT: -**



**Estimated model:** sales = (10280.16) +(1.015) \*(Total Bags) +(0.901) \*(Total Volume)

From the above output we can draw the value of Durbin Watson statistic value which is 1.53 which does not lie in the range of 1.75 to 2.3 hence we conclude that this data has Autocorrelation problem. Now for the higher order autocorrelation we test for

Breusch Godfrey Serial Correlation LM Test.



Here we look at the probability values of chi-square. If the probability of chi-square is less than 0.05 then we can say that there is presence of autocorrelation. But the probability of chi-square in our model is 25.28% which is more than 5%. With this we conclude that in higher order autocorrelation test of our data there is no autocorrelation.

**REMEDIAL MEASURE: -**

To reduce the effect of autocorrelation in our model we do the regression analysis taking the lagged values of the sales as one of the independent variables.



**Estimated model:** sales = (19051.14) +(1.21) \*(Total Bags) +(0.732) \*(Total Volume) +(0.050) \*(lsales)

**CONCLUSION: -**

We can see that probability values of only one independent variable are statistically insignificant rest all are having values below 0.05 which is good sign. The vales of R-squared and Adjusted R-squared are also 0.848 and 0.834 which implies our model has high explanatory power. The Durbin Watson statistic value which is 1.74 which means the problem of autocorrelation is eliminated.

NAME: S.KRISHNAN

REG.NO: 200968068

SEC: DSE-A

ROLL.NO: 19