Text

Description automatically generated with medium confidence

Time Series Forecasting and Trend Analysis

Subject Report

on

Aurobindo Pharma's Balance Sheet Extract

Prepared by

|  |  |
| --- | --- |
| Enrollment Number | Name |
| 20231012  20231022  20231023  20231040  20231043  20231057 | Chetanya Dabi  Harshal Dhruve  Himanshu Singh  Neeral Girnara  Prathmesh Talhande  Vedant Thakkar |

In partial fulfilment of the requirements for the

**PGDM**

**in**

**Big Data Analytics**

**Batch 2023-2025**

at

**Adani Institute of Digital Technology Management**

Gandhinagar, Gujarat.

**Table 1:**

Y= f (x1, x2)

Total Liabilities = f (Non-Current Liabilities, Current-Liabilities)

Data and Source

|  |  |  |
| --- | --- | --- |
| **Variable** | **Variable Code** | **Description & Measurement** |
| Total Liabilities | TL | AMOUNT(CR) |
| Non-Current Liabilities | NCL | AMOUNT(CR) |
| Current-Liabilities | CL | AMOUNT(CR) |

**Interpretation:**

The provided information displays Aurobindo Pharma's financial structure, including total liabilities (TL), non-current liabilities (NCL), and current liabilities (CL) together with their respective currency amounts (CR). The following is a description of the data:

Total Liabilities (TL): All of a company's or organization's financial obligations, including both current and non-current liabilities, are referred to as total liabilities (TL).

Non-Current Liabilities (NCL): These are debts that are not anticipated to be settled during the company's or organization's current operational cycle or fiscal term. Long-term debts, such bonds or long-term loans, are usually included in non-current liabilities.

Current Liabilities (CL): These are the debts that the business or organization expects to pay off during its current fiscal year or operational cycle. Accounts payable, wages, and taxes are examples of short-term debts that are typically included in current liabilities.

**Table 2:**

**Descriptive statistics**



**Hypothesis testing:**

H0: Data is normally distributed (p-value >0.05)

H1: Data is not normally distributed (p-value< 0.05)

**Interpretation for Skewness for Current Liabilities , Non-Current Liabilities and Total Liabilities:-**

Positive skewness shows “Higher Mean” than “Median”

Valued of my dataset clustered at “Left” which shows “Positive Skewness” with tail of my dataset toward right.

**S=0 (No Skewness)**

**S> 0**

**(Positively Skewed)**

**Long tail on right**

**Interpretation for Kurtosis:-**

K>3 (Leptokurtic curve){High peak}

K = ~ 3 (Mesokurtic curve)

K<3 (Platykurtic curve {Broader peak}

**Interpretation for Current Liability:-**

**Kurtosis is 1.986602. It is lesser than 3. Thus it is platykurtic. Thus it shows broader peak and thinner tail. It shows value spread more from mean. Thus fewer extreme values in tail’s. Thus it shows lower potential risk or lower probability of extreme events.**

**Interpretation for Non Current Liability:-**

**Kurtosis is 5.995215. It is greater than 3. Thus it is leptokurtic. Thus it shows high peak and heavy tail. It shows values are clustered around mean. It also shows more extreme values in the tail. We may called it outlier in our dataset. It shows higher potential risk. Or higher probability of extreme events.**

**Interpretation:**

From the above table, we can observe that the variable CL has skewness of 0.74 with kurtosis of 1.99 which means that it is positively skewed data.

The variable NCL has skewness of 1.71 and kurtosis of 5.99 which means the data is positively skewed.

The variable of TL has skewness of 1.15 and kurtosis of 3.11 which means the data is positively skewed.

**Jarque Bera :-**

JB = n{ S2 - (K-3)2 }

6 24

**Assumptions from probability of Jarque-Bera probability values:-**

**For Current Liability:- Probability value is 0.132827. It is greater than 0.05. Thus statistically insignificant. Thus we will accept H0 (Null Hypothesis). It shows that current liability is “Normal”.**

**For Non Current Liability:- Probability value is 0.000002. It is lesser than 0.05. Thus statistically significant. Thus we will reject “H0 (Null Hypothesis)” , Thus we have to accept “H1  Hypothesis”. It shows that non current liability is “not Normal”.**

**For Total Liability:- Probability value is 0.035616. It is lesser than 0.05. Thus statistically significant. Thus we will reject “H0 (Null Hypothesis)” , Thus we have to accept “H1  Hypothesis”. It shows that non current liability is “not Normal”.**

**Correlation Matrix:**



**Interpretation:**

* Ideal Correlation value exists between –1 to 1.
* From the above data correlation between CL and TL is 0.9292 which is positive correlation. Thus current liability increases will result in increase in total liabilities.
* The correlation between the NCL and TL is 0.4853 which is positively correlated. Thus increase in non current liabilities will result in increase in total liabilities.

**Table 3:**

**Unit root test**

**Total Liabilities:**

* Level:



* **1st Difference:**



**Non-Current Liabilities:**

* **Level**:

**1st Difference**:



**Current Liabilities:**

* **Level**:



* **1st Difference**:

|  |  |  |
| --- | --- | --- |
| **Unit Root Test** | **Level** | **1st Difference** |
| Total Liabilities | -1.691617 (0.4220) | -1.757669 (0.3902) |
| Non-Current Liabilities | -2.666011 (0.0921) | -5.658074 (0.0001) |
| Current-Liabilities | -1.787382 (0.3790) | -9.708631 (0.0000) |

**Interpretation:**

**Total Liabilities: The test statistics, assuming a significance threshold of 0.05, are not statistically significant in either levels or first differences because values are greater than 0.05. Thus we will accept H0 “Null Hypothesis”. This implies that there is insufficient data to rule out the null hypothesis that a unit root exists. The sequence might not be stationary. Thus not normally distributed.**

**Non-Current Liabilities: At a 0.10 significance level (p-value = 0.0921), the test statistic is marginally significant in the level but not statistically significant at a 0.05 significance level. The test statistic for the first difference is very significant (p-value = 0.0001), providing evidence against the unit root null hypothesis. Following differentiation, the series might be stationary. Because it not acceptable at 1% or 5% but at 10%. Thus we will reject “Null Hypothesis” and we will accept H1 hypothesis which tells us that data do not have unit root . thus it is stationary and Thus “Normally distributed”.**

**Current Liabilities: At a significance level of 0.05, the test statistic in the level is not statistically significant. The test statistic for the first difference is very significant (p-value = 0.0000), providing evidence against the unit root null hypothesis. Following differentiation, the series might be stationary. Thus we will reject H0 “Null Hypothesis” and will accept H1 Hypothesis. Thus our data do not have unit root. Thus stationary. Thus we can say that normally distributed.**

**Table 4:**

**Regression model:**



Coefficient (C)

-1259.782 (Negative Liability = Positive Assets)

|  |  |  |
| --- | --- | --- |
|  | **Co-efficient Value** | **P. Value** |
| Non-Current Liabilities |  |  |
| Current-Liabilities |  |  |
| **C** |  |  |
| **R^2** |  | - |
| **Adjusted R^2** |  | - |
| **f-stats** |  |  |
| **DWT** |  | - |

**Interpretation:**

**Here in the above given table we can see that the Probability value if t-Statistics of NCL and CL both have values under 0.05. Thus we will reject H0 “Null Hypothesis”. Thus we will accept H1 Hypothesis. Thus Data not normally distributed.**

**The R-squared value of 0.960059 shows us that the variance in dependent variable is 96% explained by independent variable the given regression model. It shows 96% fitness of the model.**

**The Adjusted R-squared value of 0.957100 shows that the dependent variable is explained 95.71% when we take predictors into consideration.**

**The probability value of F-statistic is 0.0000 shows that the model is of high significance.**

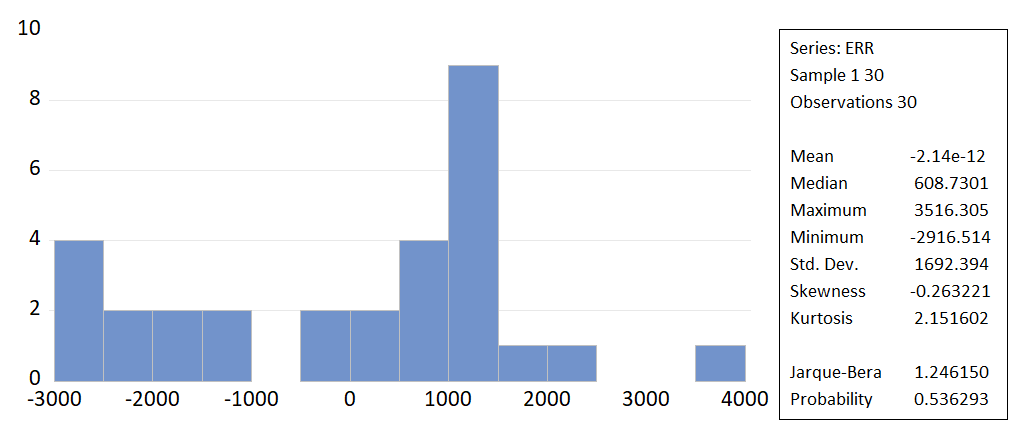
**The DWT which is Darwin Watson test which is 0.380024 which is nearer to 0 thus we can say that there is positive autocorrelation in errors. This might imply that there's some pattern or systematic variation in the errors that the model hasn't captured.**

**The F-statistics measures regression analysis to determine whether the overall regression model is significant or not. In our model F-statistics is very high thus we will reject Null Hypothesis . It suggest that at least one of our independent variable significantly influence dependent variable. Thus our model is statistically significant. Overall, we can conclude that the model is robust.**

**Table 5:**

**Error testing:**

1. **Jarque-Bera testing:**



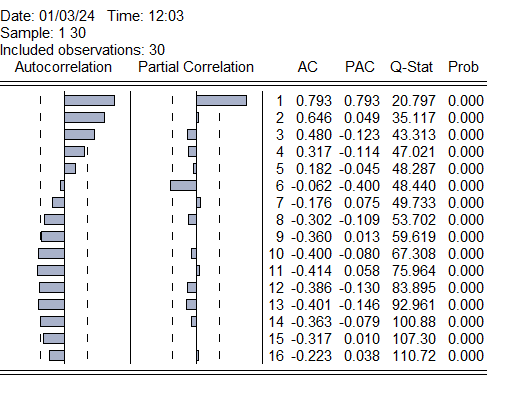
**Interpretation:**

H0: Error is normal distributed.

H1: Error is not normal distributed​.

From the above graph we can observe that the probability value of Jarque-Bera is 0.536293 which is > 0.05 so we accept the H0 and reject H1 that the error is normally distributed.

1. **Correlogram testing:**



**Interpretation:**

H0: There is no Auto correlation (AC or PAC)​

H1: There is Auto correlation (AC or PAC)

All p-value are less than 0.05 (threshold), which reject null hypothesis (H0) and accept Alternate Hypothesis (H1) which says Error is Auto correlated. (Not Normally Distributed)

1. **ADF testing:**



**Interpretation:**

Hypothesis testing:​

H0: Error has unit roots​

H1: Error does not have Unit roots​

Here, P. value of t-statistic is 0.3676 which is greater than 0.05, which leads to accept of Null Hypothesis (H0) which says Error have unit roots or trend or Auto-correlation or Not Normally distributed.

1. **BG-LM testing:**



**Interpretation:**

H0: P value is 0(Error is not correlated

H1: p value is not zero (Error is correlated)

Since my P. value of F-statistics is 0.0000 which is less than 0.05, In which we will reject Null hypothesis (H0) and accept Alternate hypothesis (H1) which says Error is not correlated or not normally distributed.

1. **Durbin-Watson Stats**



K (Independent Variable): 2

N (Observations): 30

**+ve AC**

**-ve AC**

**No AC**

**0**

**4**

**2**

**DU**

**4-DU**

**4-DL**

**DL**

Durbin-Watson Stat. value is .380024 which lies in above diagram between 0 and DL, which says Error are positively Auto correlated. Thus Not Normally distributed.

If these tests collectively suggest that errors are not normally distributed and show patterns of autocorrelation or non-stationarity, it indicates that the assumptions of the linear regression model might not hold. Violations of assumptions like normality, independence, and homoscedasticity could lead to biased parameter estimates, inefficient predictions, and unreliable hypothesis testing.

**Overall:**

|  |  |  |
| --- | --- | --- |
|  | **Stat. Value** | **P. Value** |
| **Jarque-Bera** | 1.246150 | 0.536293 |
| **Correlogram** |  |  |
| **ADF** | -1.810490 | 0.3676 |
| **BG-LM** |  |  |
| **Durbin-Watson** |  | - |

**Model Limitations:** Non-normality in errors, along with other issues, suggests that the assumptions of the regression model might not hold. It doesn't mean your model is fundamentally flawed, but rather that it may have limitations in explaining the variation in the data.

**Potential Issues:** Violations of assumptions like normality, independence, and homoscedasticity can lead to biased estimates, unreliable inference, or inefficiency in predicting outcomes. Autocorrelation, for instance, might indicate that the model doesn't account for some underlying patterns in the data.

**Room for Improvement:** Addressing these issues can improve your model. It might involve reassessing the variables, using different techniques, considering transformations, or exploring alternative models that are more suitable for your data's characteristics (e.g., time series models for time-dependent data).

**Context Matters:** Whether the model is "good" or "bad" also depends on the context, the purpose of the model, and the extent to which its predictions are accurate and reliable for your specific needs.

Remember, no model is perfect, and identifying limitations through diagnostic tests is a crucial step in refining and improving your modeling approach. It's an opportunity to explore further, adjust the model, and potentially uncover additional insights about the data generating process.

**Data source:**

