**SALES FORECASTING**

Prathamesh Bhagwan Bagul

**Tools Used: -**

1. Data Cleaning - Python + Excel
2. Data Analysis - Python
3. Sales Forecasting - Python + Power BI
4. Visualization - Power BI

**Procedure: -**

1. Data Cleaning: -
2. Entries with empty spaces on **Total Fees** were discarded.
3. Entries with empty values and numbers on **Counsellor Name** were given the entry of **No Counsellor**.
4. Entries with empty spaces only on **Fees Pending** or **Fees Received** were calculated. **Total Fees = Fees Pending** **+ Fees Received**
5. Entries with duplicate counsellor names (upper - lower case) were replaced.
6. Course names were uniformized.
7. Data Analysis: -
8. Libraries used were matplotlib, seaborn, pandas and numpy.
9. Countplot was plotted for various columns to get an idea of the count of each counsellor.
10. Boxplot was plotted to identify outliers, which were subsequently removed.
11. A pivot table was constructed containing the total sales of each counsellor per month.
12. A line plot was plotted to visualize the trend of each course in the given period.
13. A bar graph and a pie chart showing the total sales for each course were constructed.
14. The trend of each course was shown in a single line plot.
15. Sales Forecasting: -
16. The library used was statsmodels, from which adfuller, ARIMA, SARIMA and graphics for Autocorrelation plot were imported.
17. Dickey-Fuller test was used to check the stationarity of the data and a p-value of less than 0.05 as a standard value for achieving stationarity.
18. If the p-value was more than 0.05, second difference was calculated, and this process was continued till it became less than 0.05.
19. The d-value was equal to the order of the difference finally obtained.
20. Once the model was stationary, partial auto-correlation plot and auto-correlation plot were used to identify p and q values respectively.
21. These p-d-q values were used in the ARIMA and SARIMA models to forecast the sales.
22. Visualization: -
23. Microsoft Power BI was used to visualize the relationship between datapoints and gain useful insights.
24. **Total Sales** was forecasted course-wise and counsellor-wise (with the option of choosing any combination).
25. Pie charts and bar graphs of **Total Sales** were constructed course-wise and counsellor-wise.
26. Miscellaneous plots containing **Total Sales** of the institute with various other parameters were shown graphically.

**Observations: -**

1. The forecasting was done for the next four months since the dataset was not rich enough for a good forecast over longer duration.
2. Course-wise trends: -
3. CDCW sales is seasonal, and its trend is increasing over time. Its absolute contribution to the revenue is maximum.
4. GST sales has been increasing almost consistently since April and will continue to do so.
5. Analytics sales has been decreasing over time and its trends show a gradual decrease in the future. It might not have a severe effect on the revenue since its contribution to the entire sales is the least of all the courses. The reason for decrease is that it is a new course and not many people seem to be interested in it. There is a need of improvement in this sector.

Chart, scatter chart

Description automatically generated

1. Chart, line chart

   Description automatically generatedIf we consider the past six months, the trends of GST and PMP are increasing rapidly, that of SSBB, SSGB and CDCW are increasing in a slower pace and that of Analytics is decreasing. The reason for the rapid increase in trend for PMP might be that it is in its early stages, and it is rapidly generating interest in people to try out a new course.
2. Fees per enrollment for SSBB is almost double that of other courses. Average sales per enrollment for other courses is at a range of 10K – 13K whereas that of SSBB is about 25K.
3. Chart, bar chart

   Description automatically generatedBetween April and July, the total sales of each course are relatively higher, probably because it is the summer season and people are willing to improve their skillsets by enrolling in crash courses.
4. Table

   Description automatically generatedCounsellor-wise trends: -
5. Manish has contributed the most to the revenue of the company among all the counsellors and has a consistent sales pattern from the beginning.

1. Table

   Description automatically generatedIf we consider the past six months, Akansha has generated the highest revenue, which is almost double that of Priyanka, who is at the second position.

Chart, bar chart

Description automatically generated

1. There is a lot of pending fees on SSBB, CDCW and GST, which aggregates to about 3M.

Chart, bar chart

Description automatically generated

Fees Pending Forecast.

Chart, line chart

Description automatically generated

Chart, line chart

Description automatically generatedCourse-Wise Trend