**ADTA 5250 Final Project**

**Dallas city Revenue Budget Analysis**

Points to say:

: I have checked for Missing data and there is no Missing data in the data set and checked for the outliers. Finally, I have even looked for duplicates and the data set doesn’t contain any duplicates.

Before beginning any analysis or model-building, data cleaning is an essential stage in the data preparation process. Its goal is to increase the dataset's quality and dependability. The procedure usually entails a number of steps, some of which may be processing missing data, finding and treating outliers, and eliminating duplicate entries.

Handling Missing Data: Missing data refers to some data points where values are missing. It may happen due to a number of factors, including incorrect data entry, missing information, or problems with data collecting. There are a number of methods you can employ to deal with missing data:

Deletion: In some circumstances, you could decide to eliminate the rows or columns that contain missing values if the amount of missing data is minimal. The loss of information could result from doing this, thus it should be done with caution.

Imputation includes substituting estimated or expected values for missing data. The use of regression models, advanced imputation techniques like K-nearest neighbours (KNN) or expectation-maximization (EM) algorithms, or mean, median, and mode imputation are examples of common methods.

Outliers are data points that dramatically differ from the rest of the data. They should be identified and dealt with accordingly. They may be the result of data gathering errors or extraordinary events. The outcomes of analysis and modelling can suffer from outliers. There are several ways to spot and deal with outliers, including:

Visualisations: To find outliers visually, use plots like box plots, scatter plots, or histograms.

b. Z-score or IQR method: To locate and eliminate outliers, get the Z-score (number of standard deviations from the mean) or use the Interquartile Range (IQR).

b. Capping: To limit extremely high or low values, set a predetermined upper and lower limit.

d. Transformation: To lessen the impact of outliers, use data transformations such log transformation.

Removing Duplicates: Duplicate items in a dataset are those that are exactly the same or very similar. Additionally to wasting computer resources, duplicate data can skew analytical conclusions. You can use the following options to eliminate duplicates: a. There are numerous built-in procedures and strategies for finding and eliminating duplicates in programming languages and tools.

b. Check several columns: Duplicates might occasionally only be taken into account if particular columns have the same values.

c. Manual examination: You can manually examine the data to find duplicates for smaller datasets.

The data cleaning procedure should be customised to the unique properties of your dataset and the goals of your research, it's vital to remember that. Additionally, reliability and openness in the data preparation process depend on the steps conducted during data cleansing being documented.

Q1) What is the Overall Revenue for each Fund type from 2022 -2023?

A screenshot of a graph

Description automatically generated

**Fig-1**

The above chart tells us about the Revenue generated by each fund type in 2022 -2023 in Dallas city. We use the dataset to find the future revenue predictions and how that can be helpful to the people in the city, and which will help in developing the city.

Business Insights:

1)**Unexpected finding:** As we are into the second half of the year of 2023 the revenue generated was quite amazing compared to 2022. It’s already about 75% revenue in the half year of 2023 which generated in the year 2022. For example, if we see the revenue of **Enterprise operating fund** in 2022 is around $1177.5M and in 2023(till now) it’s $839M which is around 72% of 2022.

2) We can clearly see that most of the revenue is through the General Fund in both 2022 and 2023 which is around $1618M and $1424M respectively.

3) We can also see that the least revenue is from another operating fund in both 2022 and 2023 which is around 22M and much lesser in 2023. It is around 1.5% of the general fund.

From the above visualization we can see that most of the revenue is through the General fundraising which means the people are coming to the city and which is a sign of development of the city. The people in the city are paying the taxes and spending more which interim raises the taxes to the authority.

Q2) What is the Overall Budget for each Fund type from 2022 -2023?

A screenshot of a computer

Description automatically generated

**Fig-2**

The above chart tells us about the Budget generated by each fund type in 2022 -2023 in Dallas city. The above visualization on the budget allocations in 2022 & 2023 will help to allocate the budget in a manner where the development of the city will be in all the aspects. This will help us to analyze the future allocations of budget in the city.

Business Insights:

1) As we are in the second half of the year of 2023 the Budget allocated was quite amazing compared to 2022. It’s already about 105% revenue in the half year of 2023 which allocated in the year 2022. For example, if we see the budget of **Enterprise operating fund** in 2022 is around $1217.5M and in 2023(till now) it’s more than that of 2022.

2) We can clearly see that most of the budget is allocated through the General Fund in both 2022 and 2023 which is around $1589M and $1727M respectively.

3) **Unexpected finding:** We can also see that the least budget is from another operating fund in both 2022 and 2023 which is around 19M and lesser in 2023. It is around 1.5% of the general fund.

From the above visualization we can see that most of the budget is allocated to the General fundraising which means the people are government is spending more on the general welfare of the citizens in the city and which is a sign of development of the city.

From Fig-1 and Fig-2 we can see that the revenue and budget pattern was highest for General Fund in both years 2022-2023 and the least was Other Operating Funds but there is Numerical Difference. The second Highest Budget was from Enterprise Operating and next to it highest budget was with General Obligation. Last but one Budget

was from Internal Fund service.

Q3) Compare Revenue, Budget, and profit (in simple terms) in 2023.

**A screenshot of a computer

Description automatically generated**

**Fig-3**

**Revenue, Budget, and Profit by department each year/overall**

Here in fig-3 I have compared the Revenue, Budget, and Profit by department each year/overallusing parameters, filters.

**Business Insights:**

1. We can see that the Budget allocated and revenue which got more is from the office of the budget department. Surprisingly the allocated budget is more into the department than the revenue which resulted in negative profit i.e., taxes/ amount of money coming into the department is less than the spendings. The difference is around $240M.
2. **Unexpected finding:** Surprisingly, the departments which have less budget allocation have gained more revenue even in smaller amounts they are having more profits than the ones which have more budget allocations. For example, from fig-3 we can see that the Aviation/ sustainable development and construction department which have fewer budget allocations gained more revenue than the office of budget department which has higher budget allocations and revenue, but net is less for them.
3. There are a few of the departments which have a budget of 0 and have got revenue, which is quite an interesting fact. The departments are Mayor & council, City Manager’s office, and civil service.

The above visualization is helpful in the future allocations of budget and revenue predictions which will play a crucial role in the development of the city. The mayor and government officials can see the analysis and make the proper decisions when allocating the budget. Also, This visualization can tell that how much revenue each department is making and able the officials to look into the more developments in the area to increase the revenue more.