BI PROJECT
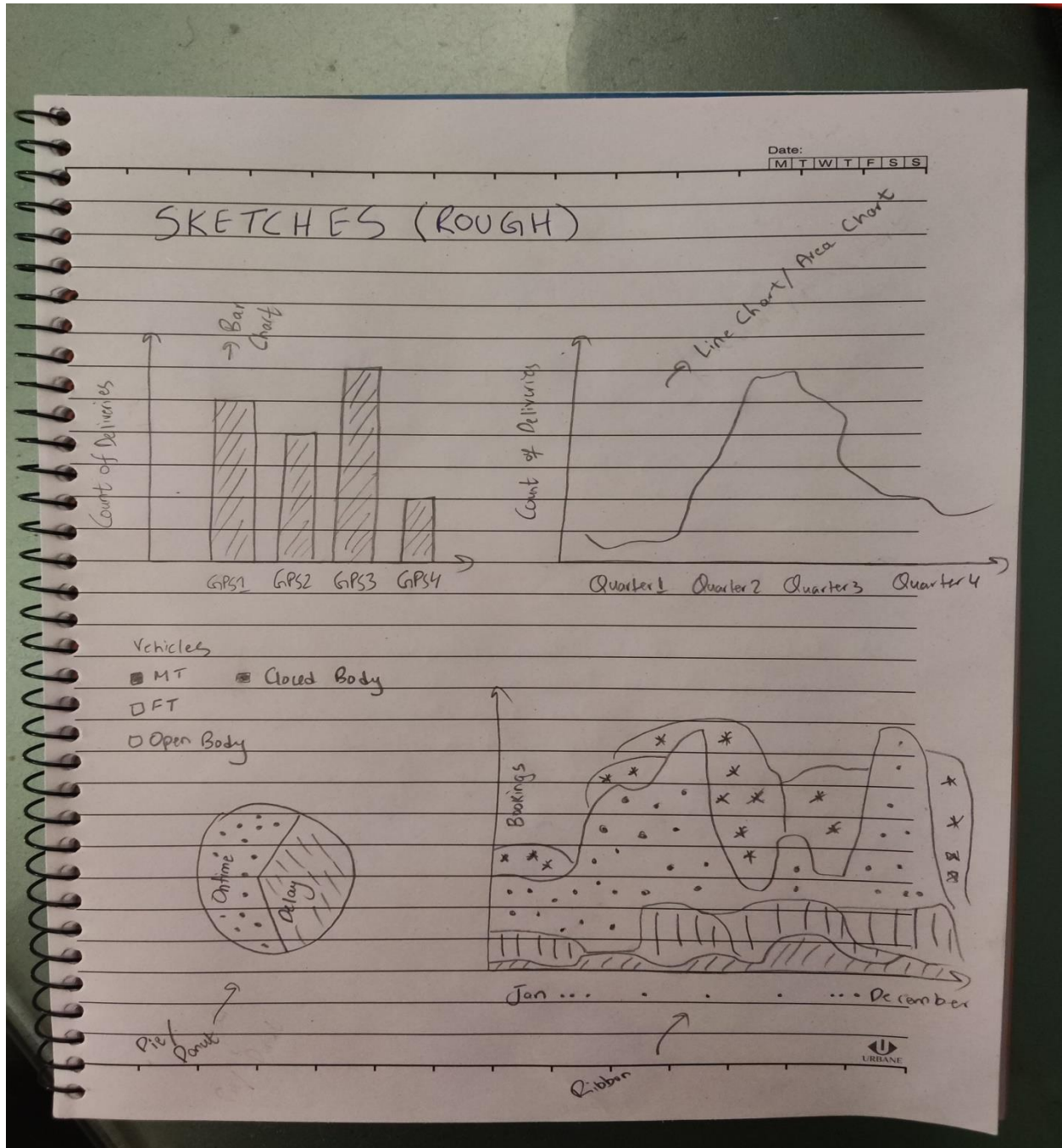
NAME: ABDUR RAFAY TANWEER

ERP: 19734

STEP 2 and STEP 3

DASHBOARDS AND STORY

→ LAYER 1

- → Individual charts showing information regarding GPS Providers, Vendor Preference as well as No. of Deliveries over some measure of time.
- → Explaination of these individual charts.
- → Show integrated view Through dashboard along with analysis.

→ LAYER 2

- → Individual charts showing/displaying information regarding delivery businesses on the basis of Starting Location and/or Destination Location.
- → Explaination of these individual charts.
- → Show integrated view through dashboard along with analysis.

→ LAYER 3

→ Individual charts displaying information regarding the delivery business/businesses by examining delayed/ on time deliveries.

→ Explaination of these individual charts.

→ Show integrated view through dashboard along with analysis.

STORY

↳ LAYER 1
↳ LAYER 2
↳ LAYER 3

DATA WRANGLING:

- Trimmed all the columns of string/text type to remove leading and trailing whitespaces from each cell in those columns
- Also removed the errors (if any) from all the columns
- Removed all the blank rows
- Changed data types where appropriate
- Dealt with the NULL values of each column on an individual basis (details for most of the columns given below

  ➢ GpsProvider, Data_Ping_Time, Current_Location, Curr_lat, Curr_lon are all related to each other hence all of these columns have their NULL values in the same rows. So the rows where these columns had NULL values were removed because it would not have been possible to fill these values and preserve their correlation at the same time.

  ➢ The NULL values in trip_end_date column were replaced by their corresponding values in the actual_eta column using the following query

  ```
  X  √  fx   = Table.ReplaceValue(#"Changed Type1",null,each [actual_eta],Replacer.ReplaceValue,{"trip_end_date"})                    ∨
  ```

  ➢ In the vehicleType column the NULL values were replaced by "Not Known" since there was no point of using a fill function in this specific column.

  ➢ In the ontime column the NULL values meant that the delivery wasn't on time hence the NULL values were replaced by "No" as in no the delivery was not on time.

  ➢ In the delay column the NULL values meant that there was no delay in the delivery hence the NULL values in the delay column were replaced by "No" as in no there was not any delay.

  ➢ In the TRANSPORTATION_DISTANCE_IN_KM column the NULL values were replaced by the most repeated value i.e. 25 since that seemed most feasible in my point of view.

  ➢ In the actual_eta column the NULL values were dealt with using the fill down function. The fill down function fills down the cell values to neighboring empty cells in the column.

- The Origin_Location and the Destination_Location column roughly followed the following format in order to display the location: Address, City, and State. So in order to perform better analysis, these columns were split by the "," delimiter to get separate columns for Address, City and State. In each of the City and State columns for the Origin_Location and Destination_Location, all of the cell values were transformed to uppercase such that there are no repeated groups for example "Tamil Nadu" and "TAMIL NADU" should be considered the same group, which they will be after this step.
- Some columns were dropped since they were not relevant to my analysis

Step 4, 5, 6

- All the charts and dashboards created are explained below with order being of relevance since this also acts as the story.
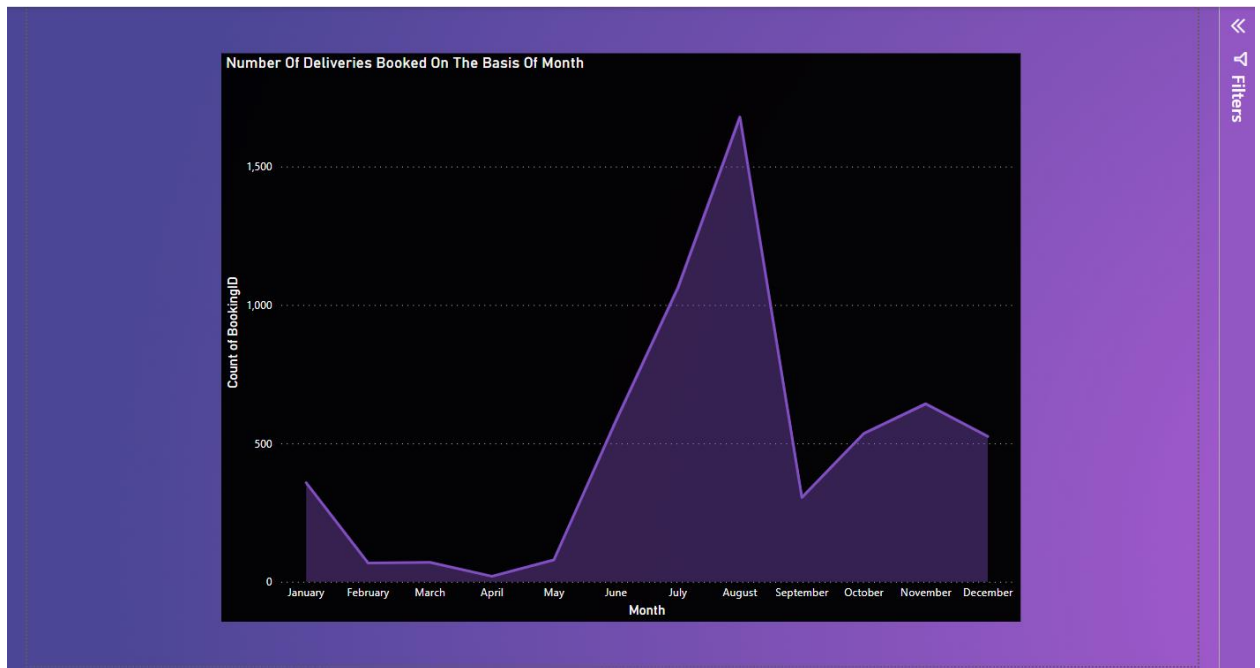
- Chart 1
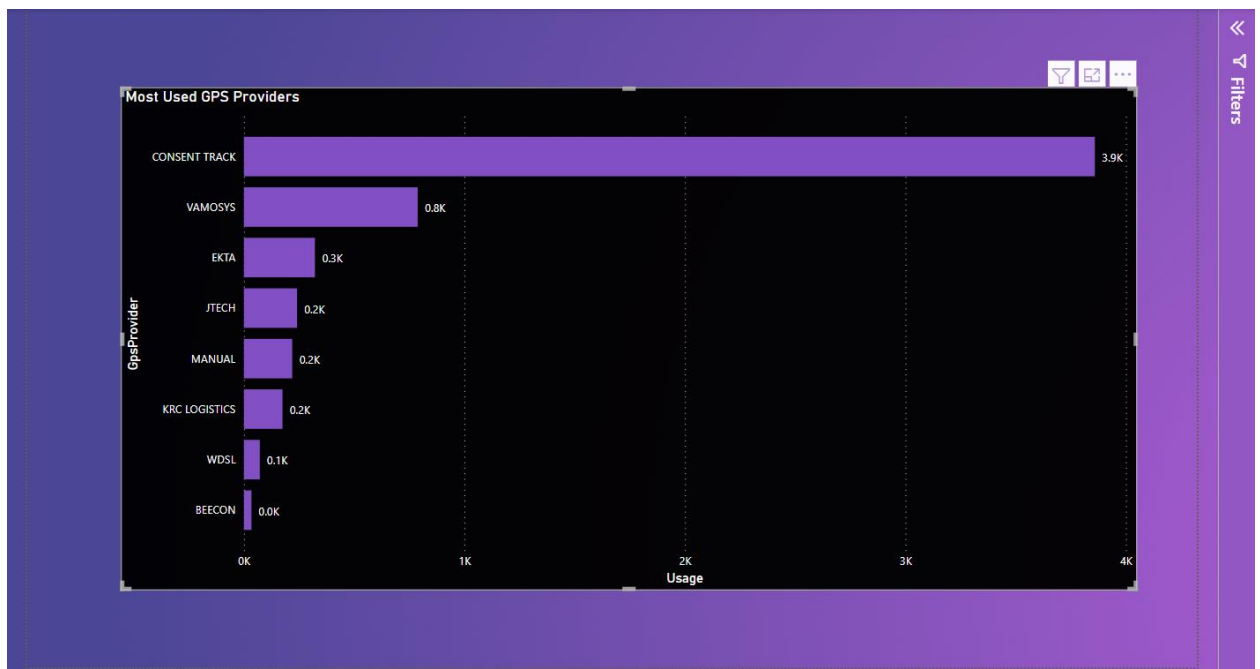
This chart tells us about the usage of the different types of GPS providers on different deliveries. In this chart we can see the top 8 GPS providers that were used to keep track of the deliveries. Through this chart we can extract the most popular/accurate/trusted GPS providers whose services are used by delivery companies. This type of information although very simple can prove to be fruit full for delivery companies that are just starting up.
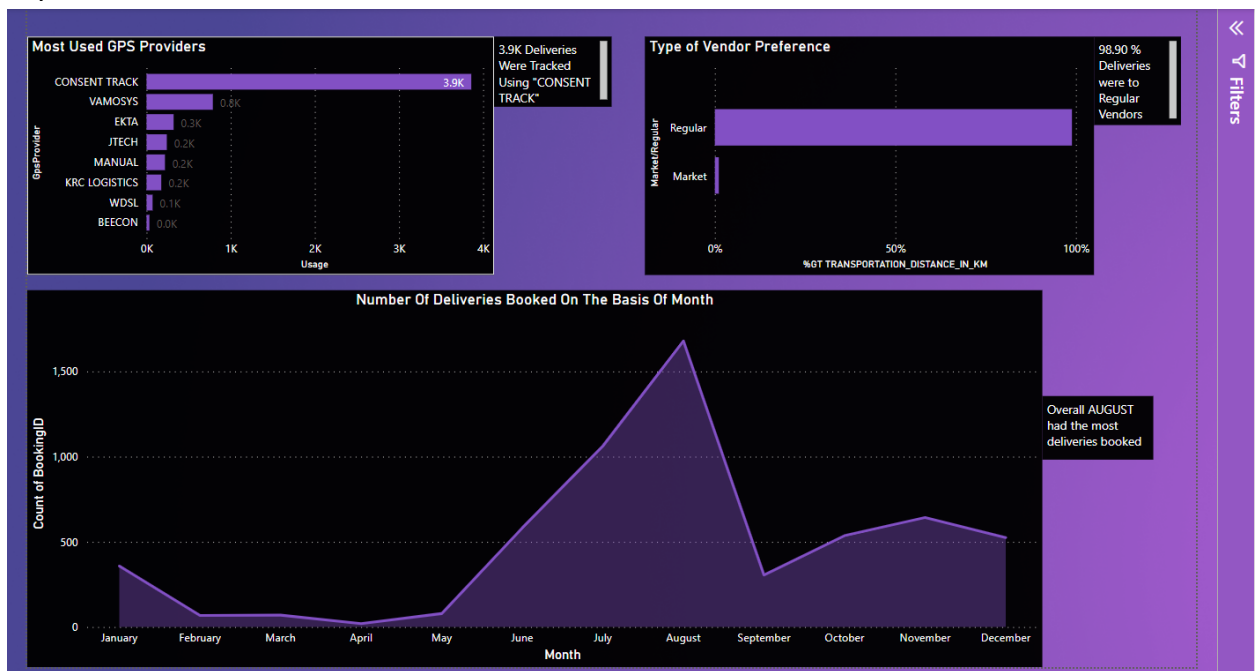
- Chart 2
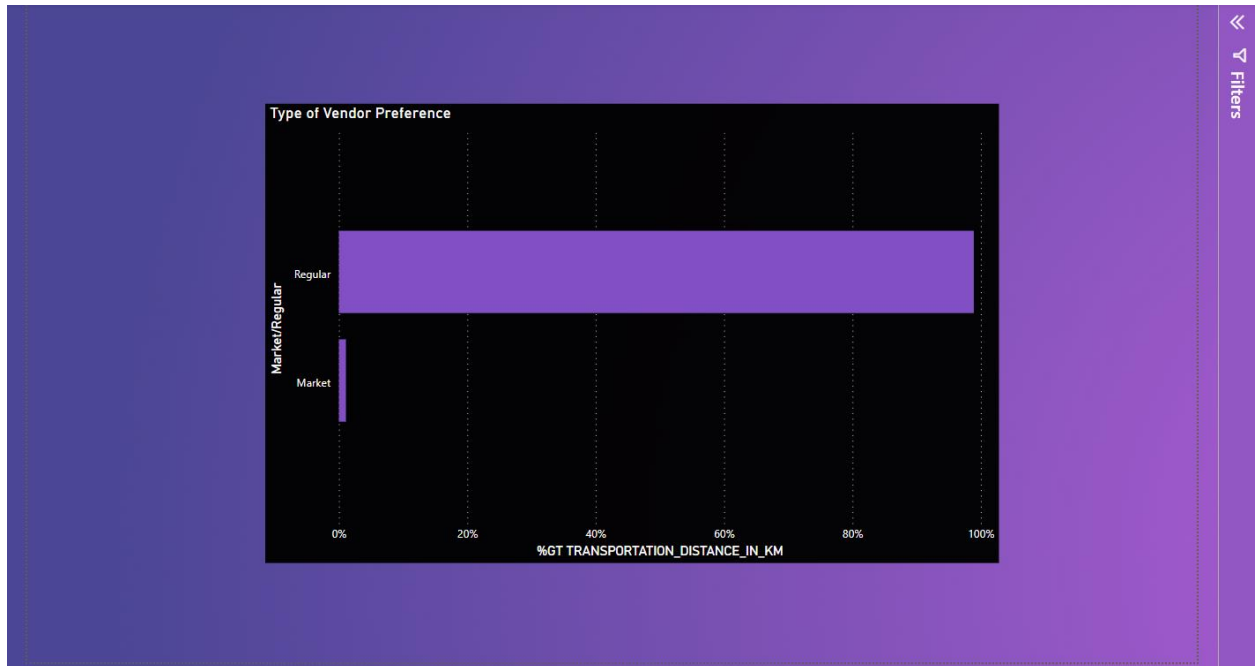
This chart shows us the percentage of the total distance travelled while delivering for two different types of vendors. A Regular vendor is that who is under contract and a Market vendor is that who will not be under contract. Through this chart we can easily see that a delivery company would much rather prefer to go through with the deliveries which are under contract because in the end a contracted vendor would be of much more prospect to them rather than a one off deal with a Market vendor.
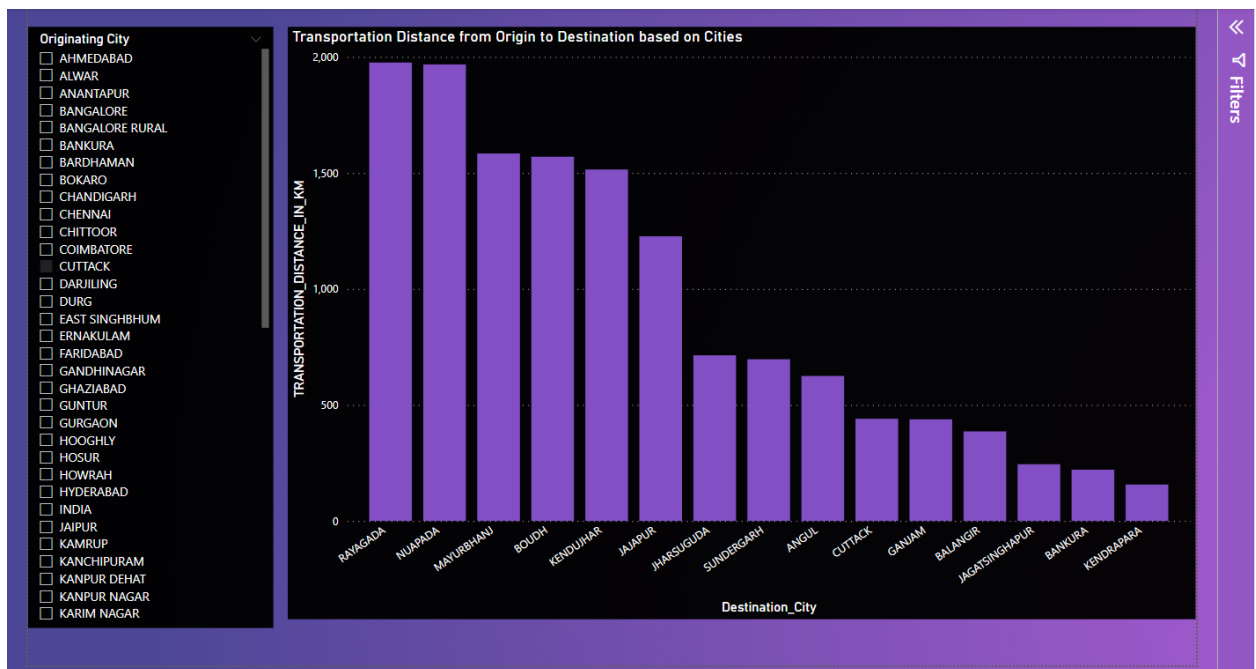
- Chart 3

This chart shows the number of deliveries that were booked over the two years in each month. Through this chart we can see those months in which the number of deliveries were maximum or minimum as well as other months that we may need to keep track of. This information may look very one dimensional but it can serve very important purposes for example upon looking at the trend of the bookings made a delivery business might declare certain months as booming months and prepare accordingly while certain months can be labeled as cool down months. This information can help build these types of dynamic working structures.

- Layer 1 Dashboard

This dashboard shows an integrated view of the previous three charts. All 3 of the charts provided us with basic/simple but important information in the view of the delivery businesses. Looking at them from a bird's eye view we can get a few takeaways from this dashboard relevant to this dataset. Firstly from chart 1 we can conclude that 3.9K deliveries of some 6k deliveries were tracked using the "CONSENT TRACK" GPS. Secondly 98.90 percent of the deliveries were to Regular vendors which were under contract with the businesses. Thirdly we can see in the Area chart that the month of "August" had the most deliveries booked while March to May can be considered as cool down months as explained earlier.
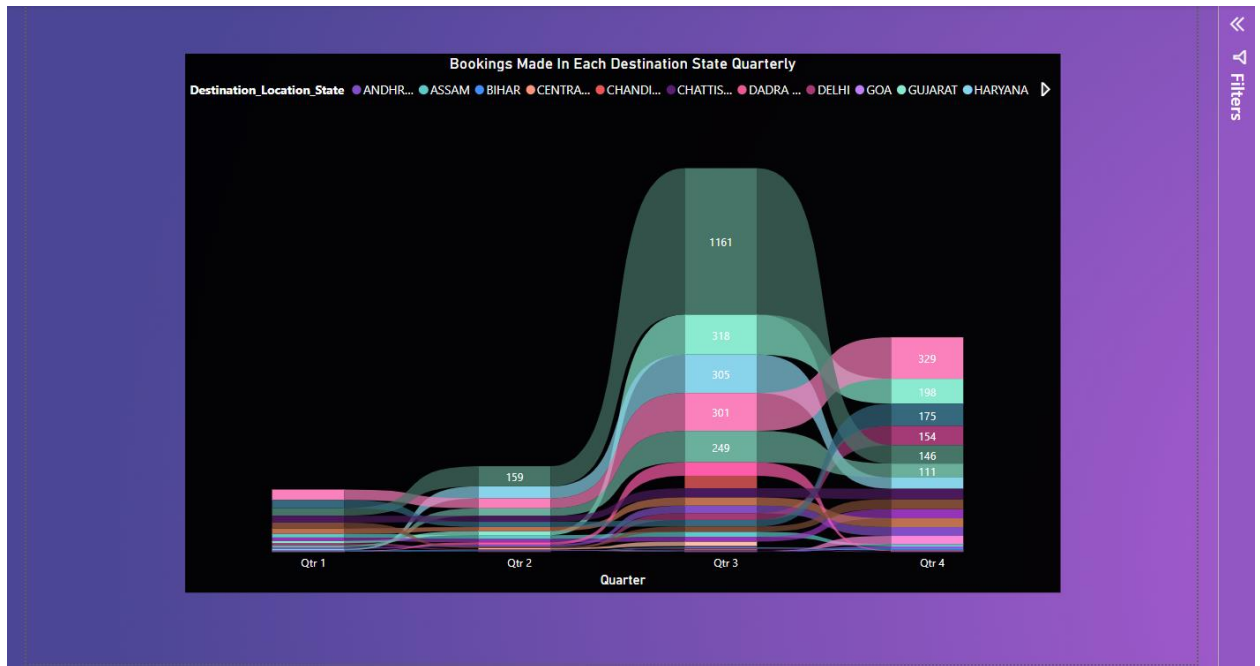
- Chart 4

This chart shows us the distance travelled in kilometers from a selected city which is part of the origin location to all the possible destination cities from that particular city. Through this chart we can see the distances recorded by the GPS providers from all the originating cities to their corresponding destinations. Note that one originating city can have more than one destinations since there can be multiple deliveries starting out from the same city and destined to end at different cities. This information can be useful when we need to filter out the travelling distances from a certain city to its destinations or when we need to know the maximum/lowest distance destination for each starting city depending on the situation.
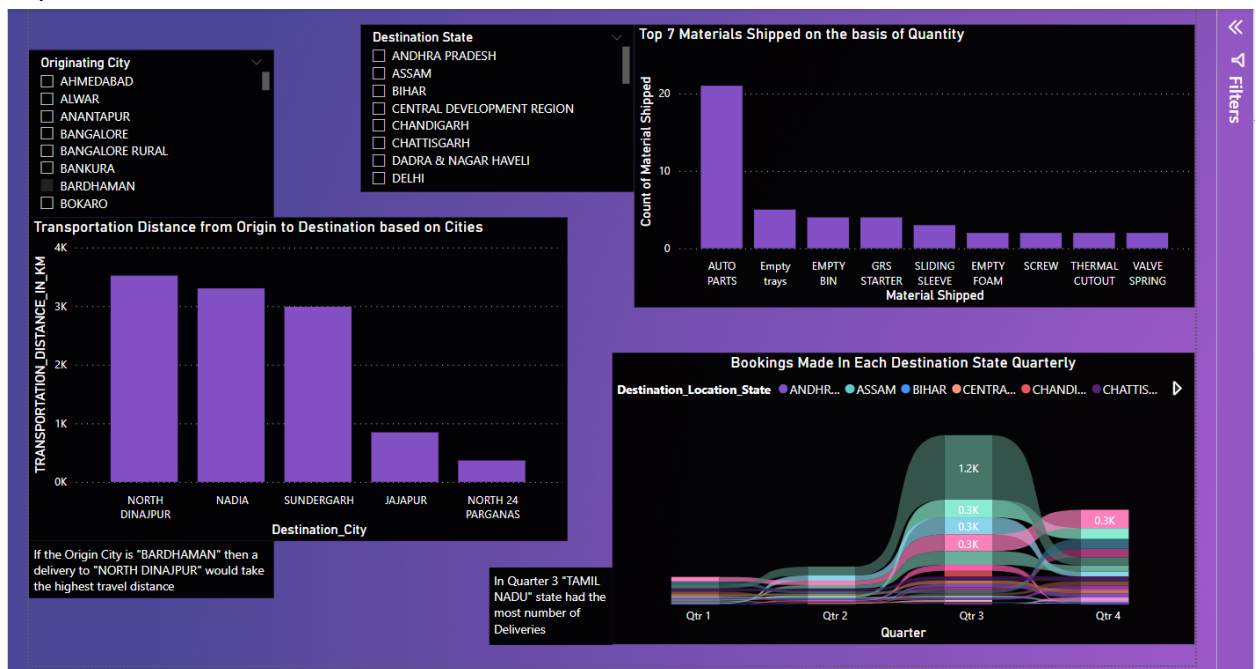
- Chart 5
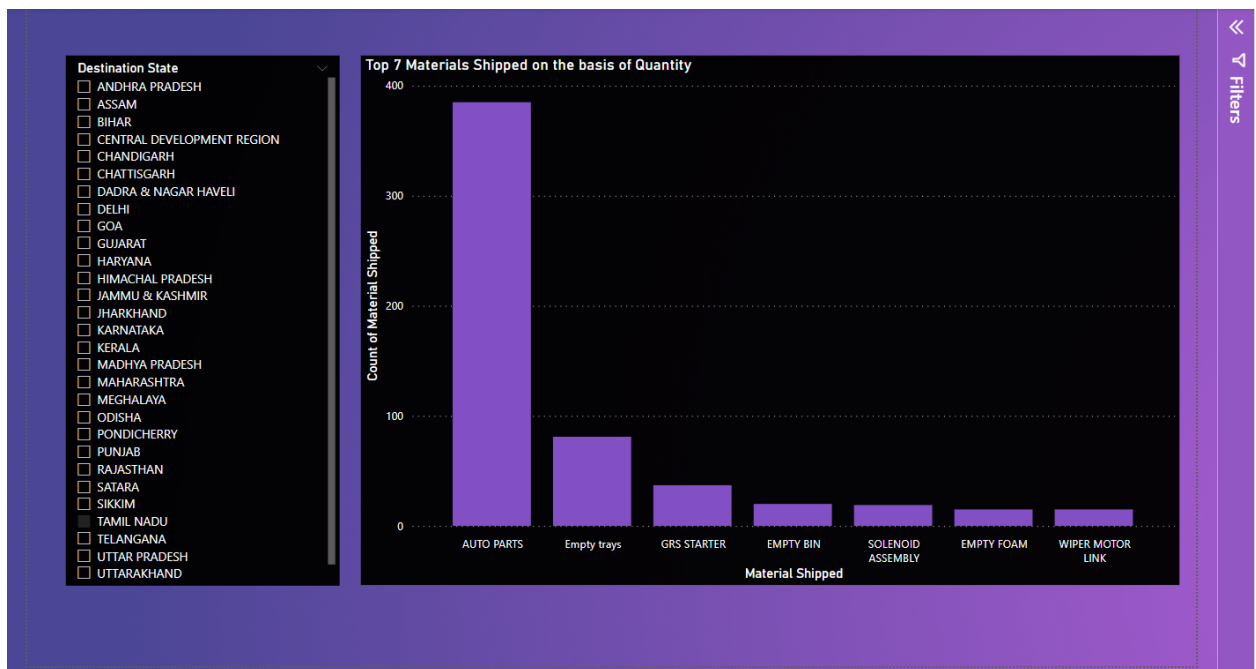
This chart shows us the top 7 materials that were shipped on the basis of quantity in each state. Through this chart we can get to know the shipment materials that vendors prefer in each state. This information can be helpful to the delivery business in a number of ways for example the materials that are on demand the most can be stored in bulk for mass delivery, in this way the delivery business can also get an edge over its competitors and on the other hand we can also look at products that are not in demand by the vendors that much such that we can free our repository of these materials for other more important products. Hence we can use this information skillfully for inventory management.
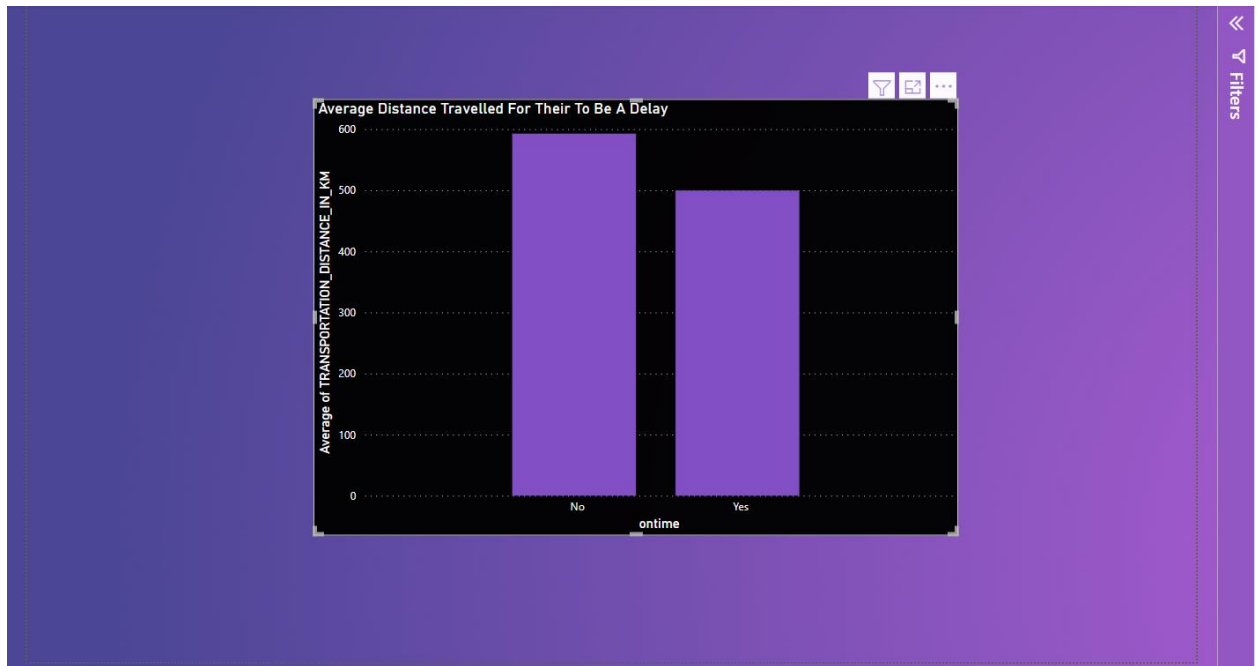
- Chart 6

This Ribbon chart shows us the number of bookings that were made over the 4 Quarters of the recorded years on the basis of the different destination states. Through this chart we can extract a lot of use full information for example those quarters in which the number of bookings was peaking as well as those states in which the most deliveries were made and these types of information is always important when dealing the demand and supply part of the delivery business making sure that their relation does not get out of balance.

- Layer 2 Dashboard

This dashboard shows an integrated view of the previous three charts i.e. Chart 4, 5 and 6. All 3 of the charts provided us with important information in the view of the delivery businesses on the basis of Origin Locations and Destination Locations. Looking at them from a bird's eye view we can get a few takeaways from this dashboard relevant to this dataset. Firstly we can read chart 4 in the following way for example: if the Origin City is "BARDHAMA" then a delivery to "NORTH DINAJPUR" would take the highest travel distance. Secondly we can read chart 5 in the following way for example: for the Destination State "PONDICHERRY" the highest material shipped is "Auto Parts". Thirdly we can conclude from chart 6 that in Quarter 3 the "TAMIL NADU" state had the most number of deliveries.

- Chart 7

This chart shows us the average distance travelled in kilometers till there isn't a delay and more importantly the average distance travelled till a delay occurs. This information can be used by the delivery business to figure out the possible reasons for delays and to counter those using creative solutions. An example of a scenario is that we know that after some average distance travelled the likelihood of their being a delay will increase so the business/company can setup small camps throughout the delivery routes after that certain distance which can provide refreshments to drivers or means to overcome any technical difficulties. This would decrease the overall probability of the delivery being delayed and the business can see improvements in terms of time management.
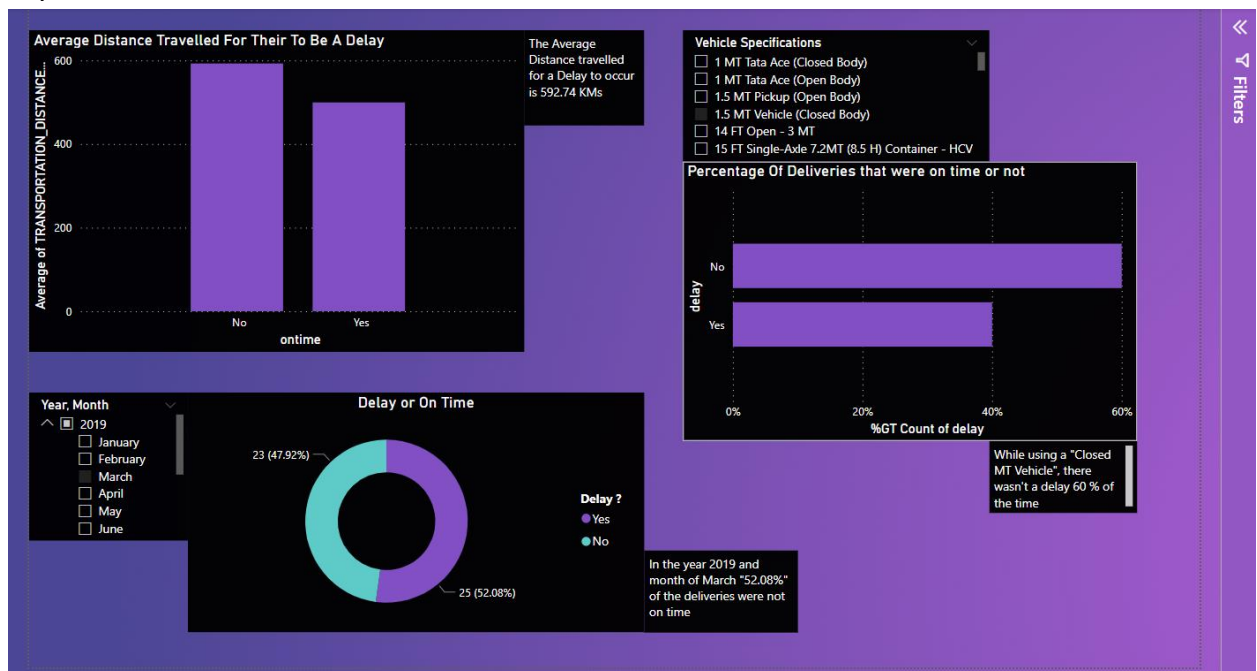
- Chart 8
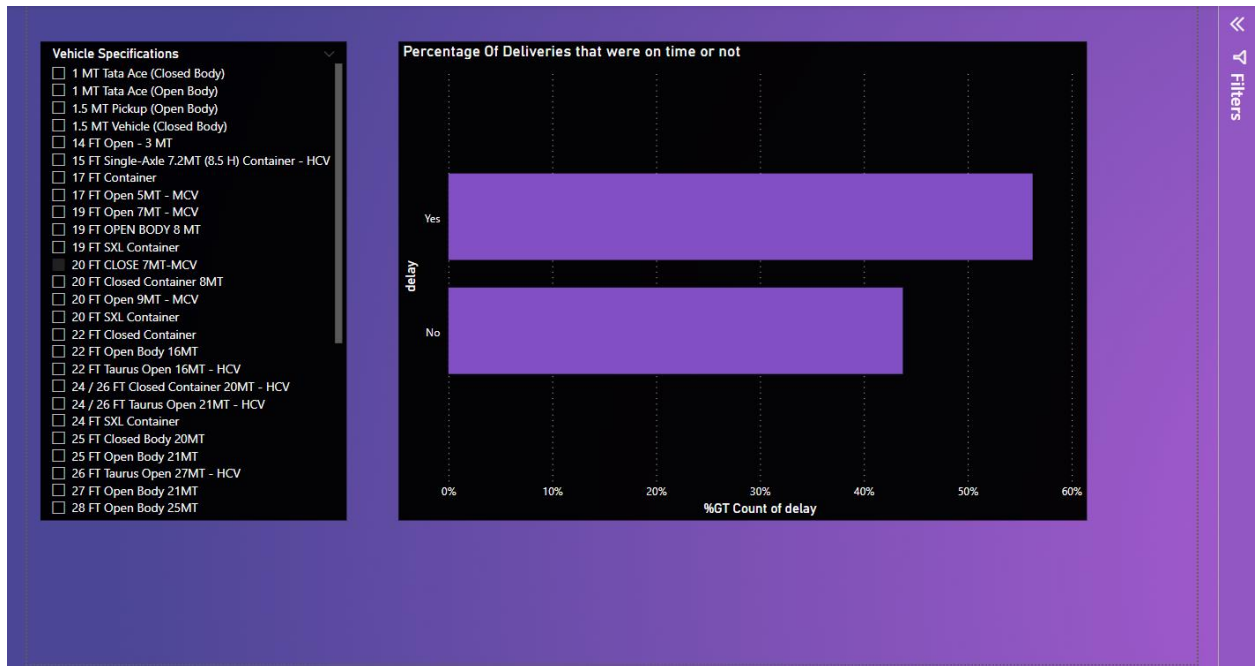
This chart shows us the percentage of deliveries that were on time as well as those deliveries which were not on time on the basis of the different Vehicle specifications. Through this chart we can again get very important information i.e. we can pinpoint the high performing and low performing vehicle models and act accordingly. A delivery business is obviously very dependent on the efficiency and quality of its vehicles and any underperforming vehicle which causes a lot of delays should be replaced or sold such that the reputation of the delivery service is not diminished. Hence this graph can help to keep track of the performances of the different vehicle models.

- Chart 9

This chart shows us the percentage/number of deliveries that were on time as well as those which were delayed on the basis of each month of the different recorded years i.e. 2019 and 2020. This chart can be used efficiently to link the past as well as present deliveries that have been delayed with the months in which most delays took place. This type of information can help delivery businesses to prepare for those "difficult" months before hand and thus this graph helps the business to come up with measures to reduce delays by studying the monthly trend of delays in the past years.

- Layer 3 Dashboard

This dashboard shows an integrated view of the previous three charts i.e. Chart 7, 8 and 9. All 3 of the charts provided us with important information in the view of the delivery businesses on the delayed and on time deliveries. Looking at them from a bird's eye view we can get a few takeaways from this dashboard relevant to this dataset. Firstly we can conclude from chart 7 that the average distance travelled for a delay to occur is 592.74 Kilo meters. Secondly we can read chart 8 in the following way for example: while using a "Closed MT Vehicle" there wasn't a delay 60 percent of the time. Lastly we can read chart 9 in the following way for example: in the year "2019" and month of "March" 52.08 percent of the deliveries were not on time.

Step 7

- Query numbers are referenced from the Excel file
- Queries number 1 and 2 are solved using the Charts 1, 2 and 3 and can also be looked at via Layer 1 Dashboard
- Queries number 3 and 4 are solved using Charts 4, 5 and 6 and can also be looked at via Layer 2 Dashboard
- Queries number 5 and 6 are solved using Charts 7, 8 and 9 and can also be looked at via Layer 3 Dashboard