**Data Analysis of an Indian Delivery Company**

Calculating business metrics

**PROBLEM**

For this task, we will use a dataset from a delivery company operating in India. The delivered items can include entrees, sides, drinks, desserts, or grocery items and are typically delivered in boxes or bags. The deliverer will generally drive a car, but in bigger cities where homes and restaurants are closer together, they may use bikes or motorized scooters.

We need to build a dashboard showing how vehicle type, type of order, weather conditions, and road traffic affect the average time taken for delivery.

Calculate the following measures and add them to the dashboard:

a. the median time that is taken for delivery

b. average time for delivery for sunny and stormy weather

c. difference between the average delivery time of ordinary scooters and motorcycle

Also, we need to answer the following question:

***What would you suggest to the delivery company so they can collect better data and get more precise results?***

**METHODOLOGY**

All the data needed for this project can be accessed through this [link](https://drive.google.com/file/d/1lJ8z4boUaTC2tXnFkaecUxTAbzz1Kflc/view).

**Clean Data**

The first column to clean is the *Weatherconditions* column. I’ll just transform the column and then **Extract using Text After Delimiter** option.

Next, I’ll also extract the time in the *Time\_taken(min)* column using Text after Delimiter as well. I also changed the *time\_taken(min)* column to time data type.

The *Order\_Date* and *Time\_ordered* columns are in text data type so I’ll change them to date and time respectively.

After changing the data type of *Time\_ordered* column, there are a lot of errors due to null empty values. I selected all cells and then right-click, then chose Remove Errors. For *type\_of\_vehicle* column, I replaced the values for *electric\_scooter* to *electric scooter*.

Lastly, I deleted columns that are not important: *ID, delivery\_person\_ID, vehicle\_condition, multiple\_deliveries,* and *festival*. That’s it for our data cleaning for this dataset.

**Analyze**

I created new measures for KPI to show in my dashboard.

* Longest delivery

Graphical user interface, application

Description automatically generated

* Median delivery

Graphical user interface, application, Word

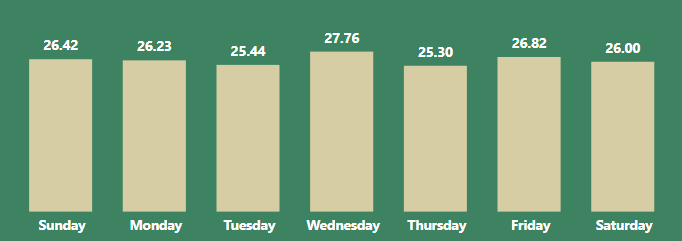
Description automatically generated

* The highest number of orders in a day

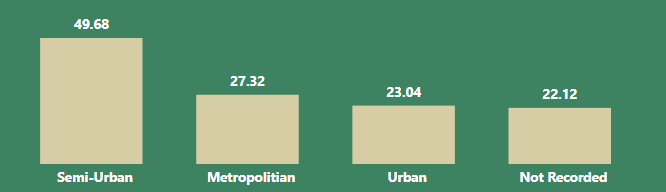
Graphical user interface, application, Word

Description automatically generated

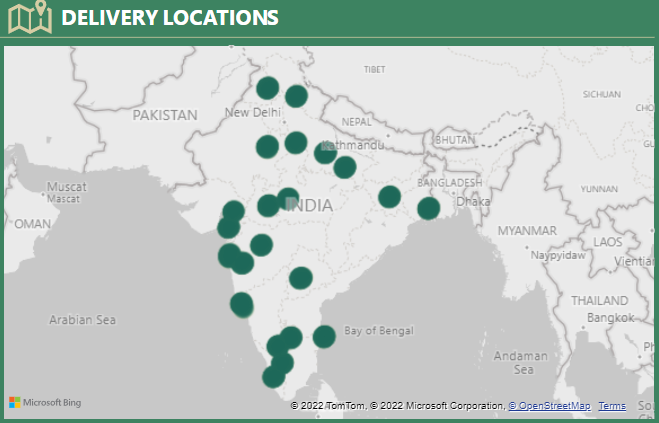
I created a column chart to show the average delivery time by day of the week.



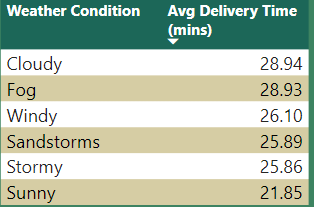
Another column chart shows the average delivery time by city.



I created a heat map showing customer locations.

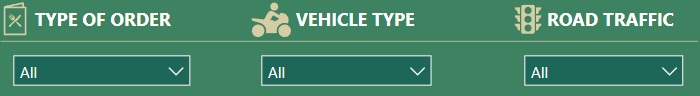


Lastly, two tables show the average delivery time by conditions and vehicle type.

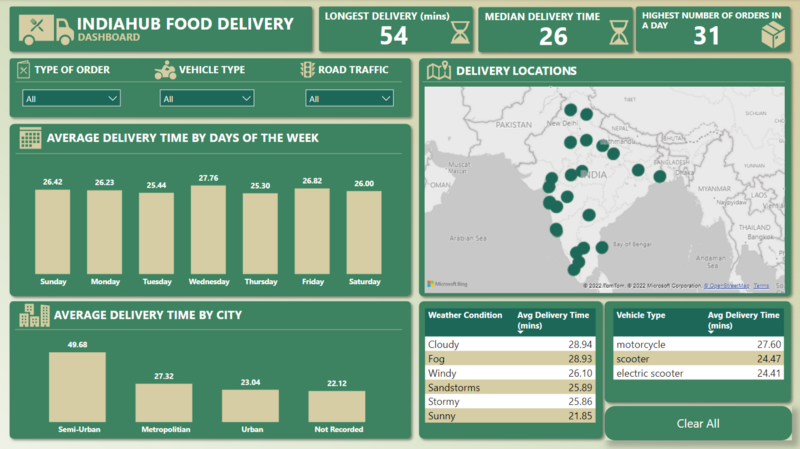
Graphical user interface

Description automatically generated with medium confidence

For filters, I created three slicers for the type of order, vehicle type, and road traffic.



The final dashboard looks like this:



**RECOMMENDATION**

*As a data analyst, what would you suggest to the delivery company so they can collect better data and get more precise results?*

As a data analyst, I would suggest the delivery company implement a system for collecting and storing data in a structured manner. This could include using a database or data warehouse to store information such as delivery times, distances, customer feedback, and any other relevant data. I would also recommend implementing processes and procedures for regularly collecting and updating this data, as well as regularly reviewing and cleaning the data to ensure its accuracy and completeness. Additionally, I would suggest implementing metrics and key performance indicators (KPIs) to track the company’s performance and identify areas for improvement. This could include metrics such as on-time delivery rate, customer satisfaction, and average delivery distance. By regularly tracking and analyzing these metrics, the company can gain insights into its operations and make data-driven decisions to improve its performance.