DATA ANALYSIS REPORT

Upfront price vs Metered Price

By: Numan Akhtar **akhtarnuman66@gmail.com**

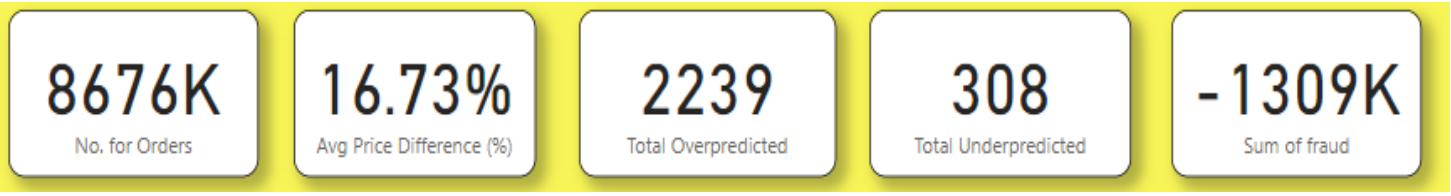
EXECUTIVE SUMMARY

This report provides an analysis of the accuracy of upfront pricing predictions for a ride-hailing service, focusing on the differences between predicted and actual distances and times of rides. The study examines the various factors affecting pricing precision, such as GPS confidence, destination changes, and data entry methods. By leveraging advanced data visualization techniques and statistical analysis, the report identifies key patterns and trends that influence the accuracy of price predictions. The objective is to offer data-driven insights that help enhance the upfront pricing model, improve customer satisfaction, and reduce pricing discrepancies for more reliable ride experiences.

KEY FINDINGS (DOWNLOAD DASHBOARD)

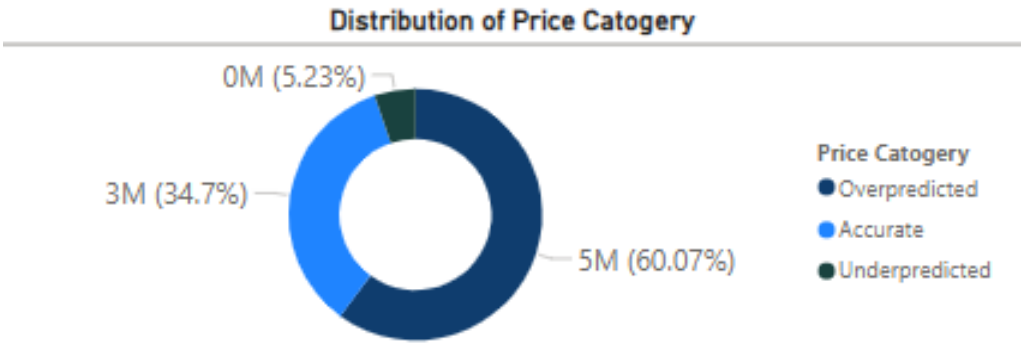
Top Metrics (Cards) Overview:

- **Number of Orders (8676K):** A total of 8.676 million orders were analysed, providing a robust dataset for identifying trends in pricing accuracy.
- **Average Price Difference (16.73%):** The average deviation between the predicted and actual prices is 16.73%, indicating room for improvement in the pricing model to better align predictions with actual ride costs.
- **Total Overpredicted (2239):** There are 2,239 instances where the actual price was significantly lower than the predicted price, leading to potential customer dissatisfaction due to overpricing.
- **Total Underpredicted (308):** There are 308 cases of underpredicted prices where customers paid less than the predicted amount, which could lead to losses for the service provider.
- **Sum of Fraud (-1309K):** The negative fraud score (-1.309 million) indicates possible fraudulent activities where predictions and actuals are manipulated or result in discrepancies.



Distribution of Price Category (Donut Chart):

- **Accurate Predictions (60.07%):** The majority of rides (60.07%) fall within the accurate prediction range (within ±20%), indicating a generally effective pricing model.
- **Overpredicted (34.7%) and Underpredicted (5.23%):** A significant portion of rides are overpredicted (34.7%), while a smaller percentage are underpredicted (5.23%). This suggests that the model tends to overestimate rather than underestimate ride prices.



Note: Price Category = IF(ABS([Price Difference (%)]) <= 20, "Accurate", IF([Price Difference (%)] > 20, "Overpredicted", "Underpredicted"))
AND
Price Difference (%) = ([metered_price] - [upfront_price]) / [upfront_price] * 100

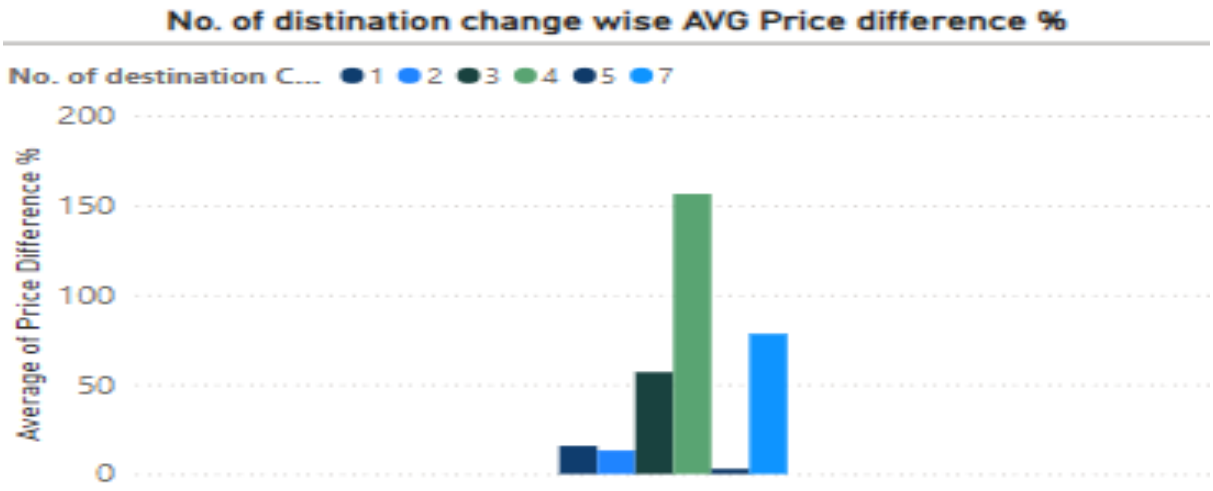
Change Reason Pricing (Table):

- **Client Destination Changed (1998 Overpredicted, 301 Underpredicted):** Most pricing discrepancies arise from clients changing their destination after the ride starts.
- **Driver Destination Changed (127 Overpredicted, 7 Underpredicted):** Instances where the driver changes the destination also lead to inaccuracies, although to a lesser extent.
- **Driver Waypoint Changed (5 Overpredicted):** Few discrepancies are caused by changes in waypoints.

change_reason_pricing	Total Overpredicted	Total Underpredicted
	1998	301
client_destination_changed	109	7
driver_destination_changed	127	
driver_waypoint_changed	5	
Total	**2239**	**308**

Number of Destination Changes vs. Average Price Difference % (Bar Chart):

- **Price Differences by Destination Changes:** There is a clear correlation between the number of destination changes and the average price difference. As the number of destination changes increases (up to 7), the average price difference also increases significantly, indicating that frequent changes during the ride lead to larger pricing discrepancies.



Conclusion:

Overall, the dashboard reveals that while most ride price predictions are accurate, there is a considerable number of overpredicted prices. Destination changes by clients are the primary driver for these discrepancies, and fraud seems more common in overpredicted pricing cases. Efforts to improve real-time data collection, minimize destination changes, and enhance fraud detection could enhance the accuracy and reliability of the upfront pricing model.