Mike Cancell - Branch Data Office Technical Skills Assessment

Part I

Assumptions:
The goal of the exercise is to extract business insights on the performance of taxi companies. As such, anomalous and outlier data has been scrubbed from the analysis. The following techniques were applied to the dataset:

- Exclude trips with no company information (`company IS NOT NULL`).
- Exclude trips with zero or negative miles for fare per mile calculations (`trip_miles > 0`).
- Exclude months with very low prior metrics to avoid noise:
    - < 100 for trip increases
    - < 0.1 for fare per mile decreases.
- Cap extreme percentage changes:
    - >= 1000% for trip increases.
    - <= -1000% for fare per mile decreases.

Additionally, the wording of the request, “largest month-over-month”, is unclear as to whether nominal or percentage increase is desired. The analysis was conducted according to nominal differences in metrics. Had the percent deviation been used, a different result set would be returned.

Results:
Please refer to the SQL code for details on the approach.

a. *Which three distinct taxi companies had the largest month-over-month **increase** in trips, and what were those months and **trip amounts**?*

Taxi Company	Trip Month	Insights
Taxi Affiliation Services	2014-03	In Mar 2014, trips increased by 115,543 (15.5%) compared to the previous month.
Chicago Carriage Cab Corp	2016-07	In Jul 2016, trips increased by 131,272 (370.9%) compared to the previous month.
Flash Cab	2016-01	In Jan 2016, trips increased by 276,654 (388.9%) compared to the previous month.

What are *general* reasons why **trip volume** might **increase** for a company in any *given* month?

Seasonal changes, marketing campaigns, operational adjustments, increased demand due to events, improved service offerings, increased fleet size, acquisitions and mergers (with a former competitor) and so on, are possible explanations. However, even after applying some rudimentary edge handling in the dataset, wide month-over-month fluctuations are still present in the results. This may imply issues in data quality or data collection rather than true business rationale.

b. *Which three distinct taxi companies had the largest month-over-month **decrease** in **fare-per-mile**, and what were those months and fare-per-mile values?*

Taxi Company	Trip Month	Insights
Metro Jet Taxi A.	2021-08	In Aug 2021, fare per mile decreased by $50.08 (95.1%) compared to the previous month.

Mike Cancell - Branch Data Office Technical Skills Assessment

Taxi Company	Trip Month	Insights
4732 - Maude Lamy	2016-04	In Apr 2016, fare per mile decreased by $29.30 (90.2%) compared to the previous month.
Blue Ribbon Taxi Association Inc.	2020-08	In Aug 2020, fare per mile decreased by $25.25 (49.1%) compared to the previous month.

What are *general* reasons why ***fare-per-mile*** might ***decrease*** for a company in any *given* month?

Promotional discounts or fare reductions, introduction of new pricing models, increased use of flat-rate fares, temporary subsidies or incentives, errors in fare calculation or reporting, increased competition driving down prices, operational adjustments to attract more riders. However, even after applying some rudimentary edge handling in the dataset, wide month-over-month fluctuations are still present in the results. This may imply issues in data quality or data collection rather than true business rationale.

Part II

For this request, I simply swapped the directionality of the measurements (e.g., increase to decrease).

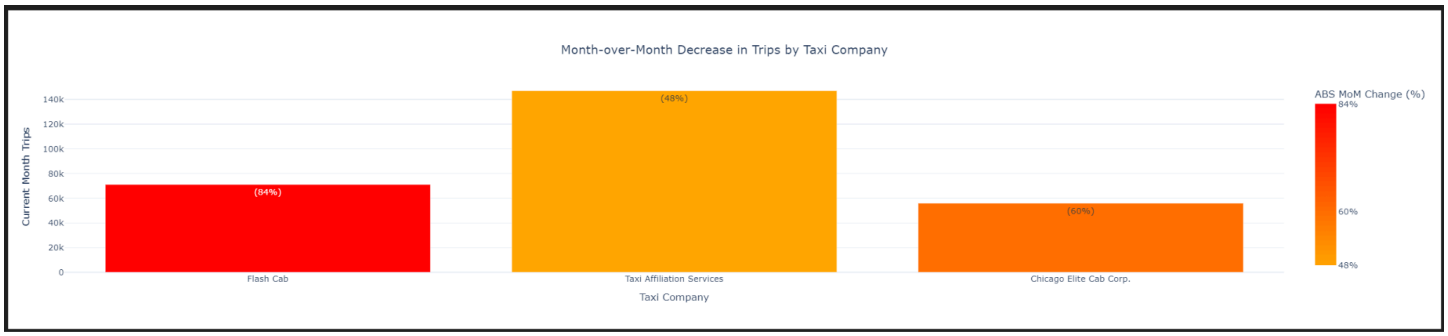
a. *Which three distinct taxi companies had the largest month-over-month **decrease** in trips, and what were those months and **trip amounts**?*

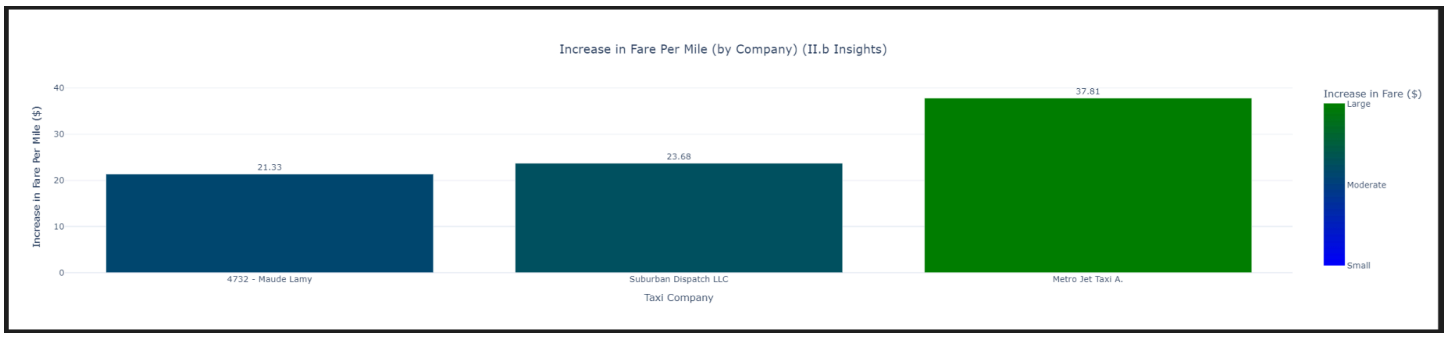
Taxi Company	Trip Month	Insights
Flash Cab	2014-11	In Nov 2014, trips decreased by 365,892 (83.7%) compared to the previous month.
Taxi Affiliation Services	2020-03	In Mar 2020, trips decreased by 134,967 (47.8%) compared to the previous month.
Chicago Elite Cab Corp.	2013-08	In Aug 2013, trips decreased by 84,206 (60.1%) compared to the previous month.

b. Which three distinct taxi companies had the largest month-over-month **increase** in fare-per-mile, and what were those months and **fare-per-mile** values?

Taxi Company	Trip Month	Insights
4732 - Maude Lamy	2015-10	In Oct 2015, fare per mile increased by $21.33 (191.0%) compared to the previous month.
Suburban Dispatch LLC	2015-08	In Aug 2015, fare per mile increased by $23.69 (313.7%) compared to the previous month.
Metro Jet Taxi A.	2022-06	In Jun 2022, fare per mile increased by $37.81 (120.5%) compared to the previous month.

Mike Cancell - Branch Data Office Technical Skills Assessment





We can see that two of the companies in the analysis are common across the top increases and decreases of trip counts (Flash Cab and Taxi Affiliation Services). This highlights a deficiency with comparing nominal metrics rather than percentages. The approach biases large taxi fleets because changes in trip count may be proportionally insignificant for them, whereas the same change amount would be relatively large to a company with fewer average taxi trips.

With respect to increases in fare-per-mile, this finding is more surprising because we would expect that this metric is fairly uniform across all taxi companies, especially in such a regulated industry and all within the same geographic market. Again, we find two companies amongst both the increase/decrease results (4732 - Maude Lamy and Metro Jet Taxi A.). For both companies we see wild fluctuations to a degree that there is no rational explanation, potentially implying quality issues in their data.

Final Note:

The referred dataset yields limited insights to the performance of Chicago taxi companies. The data is isolated to taxi company metrics but lacks contextualization of market demographics. The framework for a more robust analysis, which resolves the area of the city via lat/lon and introduces Chicago census information, is referenced below in the artifacts. This additional data would be useful in showing areas of Chicago that may be more or less lucrative to taxi service.

Artifact Links:

Artifact	**Code**	**Description**
Part I Dataset	SQL	Answers the questions posed in Part I of the assessment.
Part II Dataset	SQL	Answers the questions posed in Part II of the assessment.
Part II Final Note Reference	SQL	Trip data joined with geographic &

Mike Cancell - Branch Data Office Technical Skills Assessment

Artifact	Code	Description
		demographic information.
Combined Notebook	Python	This was used to access and analyze the dataset and generate visuals.