Summary

I opted to take the data files and load them into Databricks. I created a small Bronze Layer and moved all usable data into a silver layer. I used a combination of both python and sql scripts to cleanse the data. I build out a small “warehouse” on the data and created a semantic model in Power BI. While I did some data analysis and exploring in Power BI, I also used SQL and Python within a Databricks Notebook called Analysis. All work is in there.

A screenshot of a computer

AI-generated content may be incorrect.

Analysis

Overall, the distribution of the data is fairly. Starting with some basic querying, I first wanted to understand the number of searches for a specific DMA Yielded a conversion. This was looking at the entire dataset, not a specific date/time

A screenshot of a data

AI-generated content may be incorrect.

So going a bit deeper, I then looked at the Conversion Rate at the Source Level. A slight difference between the lowest (direct) and the highest (organic).

A screenshot of a computer

AI-generated content may be incorrect.

Then looking at channel, we can start to see some larger differences in Conversion Rate. At an aggregate, reddit as the worst conversion rate compared to google maps, which has the best.

A screenshot of a graph

AI-generated content may be incorrect.

So the next step was to look at the Conversion Rate by Channel by DMA. We then got some really interesting results!

A screenshot of a graph

AI-generated content may be incorrect.

While at the aggregate, reddit may have the lowest Conversion Rate, within the difference DMAs, it ranks higher than other channels, particularly in Atlanta and Chicago.

**Recommendations**

1. Marketing campaigns should be targeted based on channel and DMA. For DFW, move some of your marketing dollars from Facebook and by more Reddit advertising.
2. For New York, Google Maps has over a 15% Conversion Rate, whereas Reddit has a little over 10. Shift the campaign dollars away from Reddit and improve google maps search and increase Google Ads.
3. Similarly, the Conversion Rate for Reddit in LA is the lowest of all DMAs. However, Facebook is the second highest conversion within the DMA. Move more marketing dollars to Facebook from Reddit.

**Data Needs**

I looked at the various tables and was really trying to understand conversion rate at the property level and try to understand how appropriately 40% of all listings went without a reservation, while at the same time a smaller portion of the listings were rented multiple times. Approximately 23% of all listings had 2 or more reservations. Data to help understand why a large portion never got rented, while others got rented multiple times could include:

* **Pricing**: Were the properties not rented too expensive?
* **Reviews**: Did the properties with multiple reservations have favorable reviews?
* **Location**: Was socio-economic, traffic and size of space a factor

So data to support these hypothesis would be helpful. Such as pricing, terms, length of stay requested, traffic density, etc.

Looking at overall Conversion Rates, understanding what type of searches, property size needed, what areas are being searched, property availability, how long did a potential customer stay on the site searching.