



Analysis of Google Analytics Data

by Gaelim Holland
Tools: PowerBI and Python

Key Metrics and K.P.Is

8,386

Total Rows of data

133K

New Users

190K

Sessions

\$24.44

Avg. Order Value

0.92

Avg. QTY

3.07

Conversion Probability

Initial Impressions of the Data

Quick Analysis

At a glance, there is a low average order quantity when aggregating the data as a whole. The overall traffic to the website may be over-leveraged in organic traffic. This makes traffic and revenue streams vulnerable to Google algorithm changes.

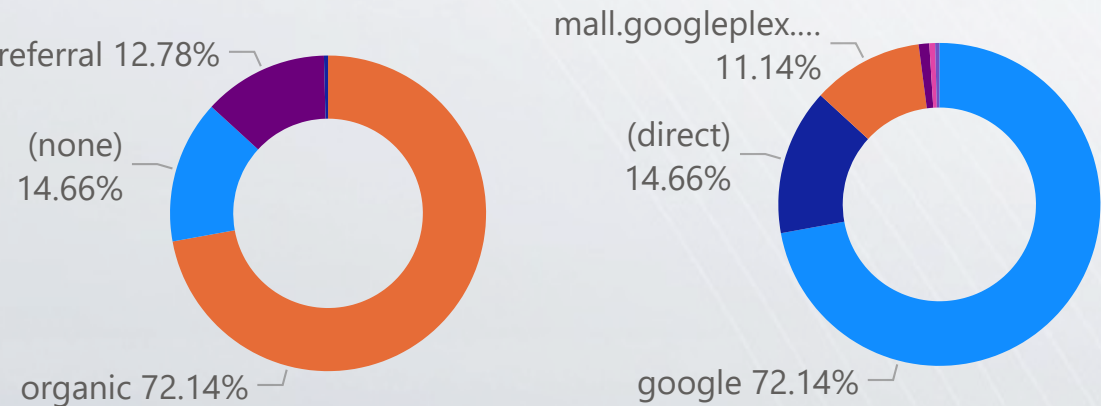
Preliminary Questions

Is organic traffic the most valuable source of traffic?
Why is order quantity low? Which segments show promise? Are there correlating factors?
Are users converting on a dominant device?

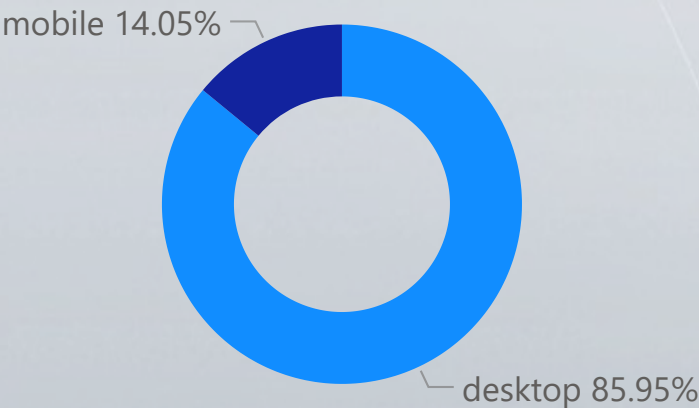
Stipulations

I ignored the day of week column because the dates given do not align with an actual day of week on the dates given. Due to the limited amount of data and context, exploratory analysis will be focused on conversion metrics.

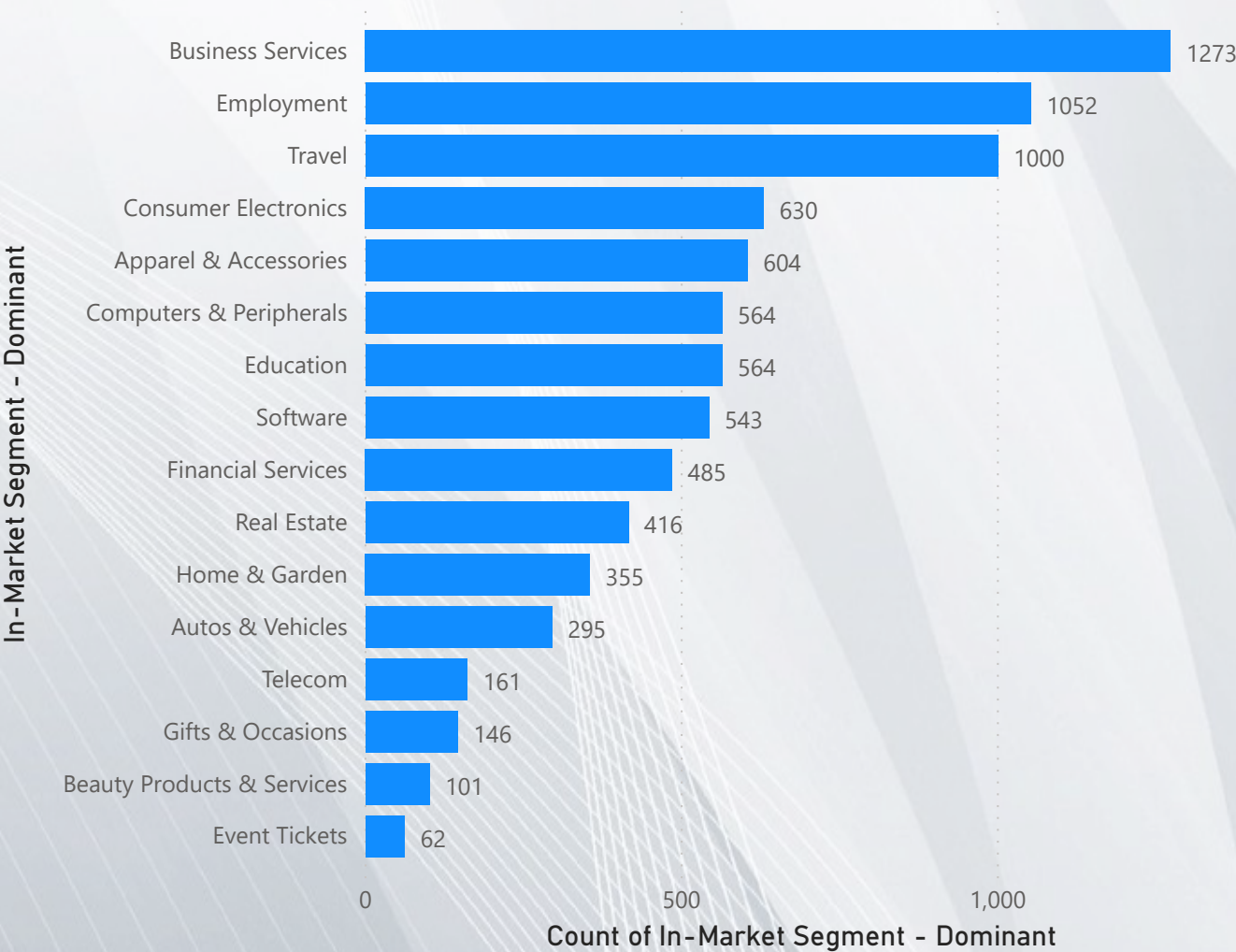
Where are the Dominant Channels?



What Devices are Preferred ?



What are Our Users Key Interest?



Exploring the Initial Data Questions

The initial hypothesis is traffic is over-leveraged for organic channels and underrepresented on mobile devices. The reason to explore this is due to the importance of diversified traffic that is not vulnerable to algorithm changes and ensuring that we have a optimized mobile experience due to the growing trend in mobile transactions.

Highly Dependent on One Channel

In the chart below, we can see that **76.83%** of all traffic is from the organic channels with **6.93%** of this traffic converting compared to the **13% of overall traffic** originating from referral traffic converting at up to **90%**.

Challenging Mobile Experience

Also we are under-represented on mobile with **9.8%** of all sessions coming from mobile with a conversion probability of **<1%**.

Key Value Metrics by by Channel & Device

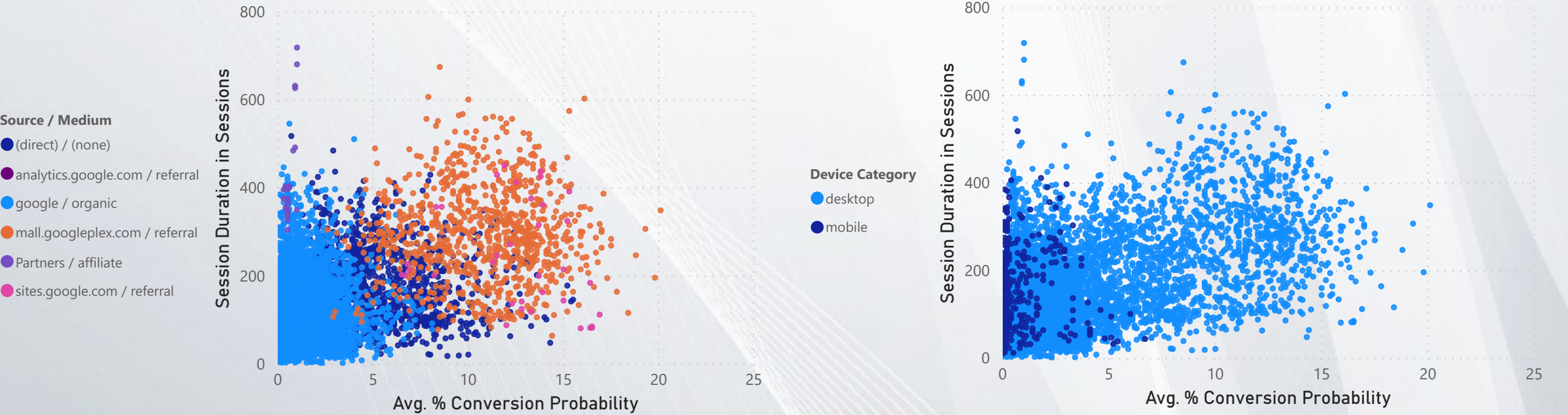
Calculated value metrics help us determine what percentage of the traffic converts (% of Converted Traffic) and the the traffic value. We can quickly determine which channels drive value. Some channels may provide opportunity for optimization and remarketing.

Source / Medium	Sessions	Data	Session Duration	New Users	Oder Value	Conversion Probability	% of Converted Traffic	Traffic Value	Avg Order
mall.googleplex.com / referral	9.95%	11.14%	298.89	5.96%	41.61%	10.66	90.79%	\$4.51	\$91.31
(direct) / (none)	11.97%	14.66%	177.88	11.90%	17.14%	5.34	38.36%	\$1.54	\$28.58
google / organic	76.83%	72.14%	120.74	80.96%	41.26%	1.42	6.93%	\$0.58	\$13.98
analytics.google.com / referral	0.70%	1.06%	152.53	0.76%	0.00%	0.29	0.00%	\$0.00	\$0.00
Partners / affiliate	0.22%	0.42%	420.83	0.21%	0.00%	0.60	0.00%	\$0.00	\$0.00
sites.google.com / referral	0.34%	0.58%	230.04	0.21%	0.00%	12.63	0.00%	\$0.00	\$0.00

Device Category	Sessions	Order Value	Data	Session Duration	New Users	Conversion Probability	% of Converted Traffic	Traffic Value	Avg Order
desktop	90.18%	99.50%	85.95%	155.96	88.90%	3.45	20.67%	\$1.19	\$28.29
mobile	9.82%	0.50%	14.05%	121.96	11.10%	0.74	3.07%	\$0.06	\$0.88

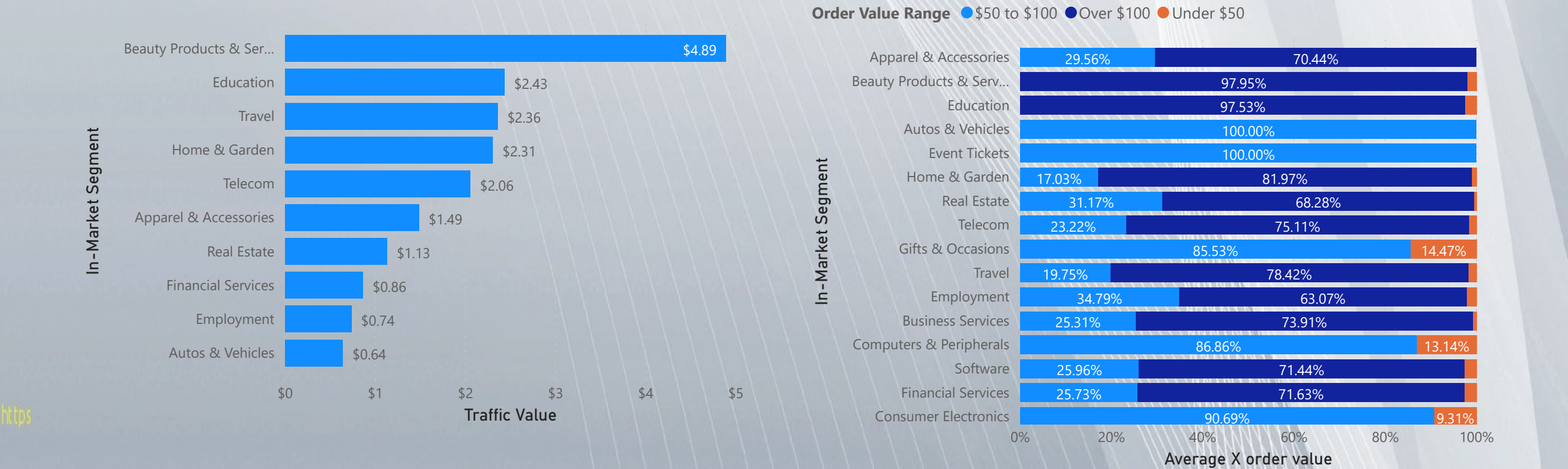
Leveraging Conversion Probability

We leverage Google probability algorithm which use features selection and machine learning to predict probability of conversion. We have 8000 plus rows of data . However this algorithm learns over time which can mitigate the small sample size. This metric has a weak correlation with session duration. However categories of device and channel indicate the significance of conversion intent for the shopping and direct channels. Also showing the lack of mobile conversion optimization.



Opportunities to Grow

Due to being over-leveraged on one channel, we can evaluate the traffic value of various In-Market segments that will provide insight into further research for targeting keywords that can be used in Paid Media, interest categories for Display Targeting and Remarketing. Using both traffic value (order value/sessions) and categorical bins of average order value we can identify the most effective segments for ROI.



Next Steps & Recommendations

Present to Stake Holders

These findings are the result of exploratory data analysis with little business context. First these findings should be presented to the business stake holders to provide context and elicit new critical hypothesis and objectives for more data utilization.

Optimization of Channels

The initial hypothesis was proven that we are over-leveraged for organic channels which provides a serious business vulnerability.

- Expand into paid search with targeted traffic based on keyword research
- Expand high quality referral traffic using content marketing strategies and outreach
- Expand traffic to the Google Merchandise Store due to its high conversion
- Optimize UX to improve mobile traffic and conversion

What is the Overall Omni-Channel Experience?

Analysis should be conducted to understand the user omin-channel journey to conversion.

Further Analysis

The sample provided was not substantial enough to provide time series insights. It would be better to have a full month to a year of data to discover seasonality trends and forecast.

[10]: <AxesSubplot:>

