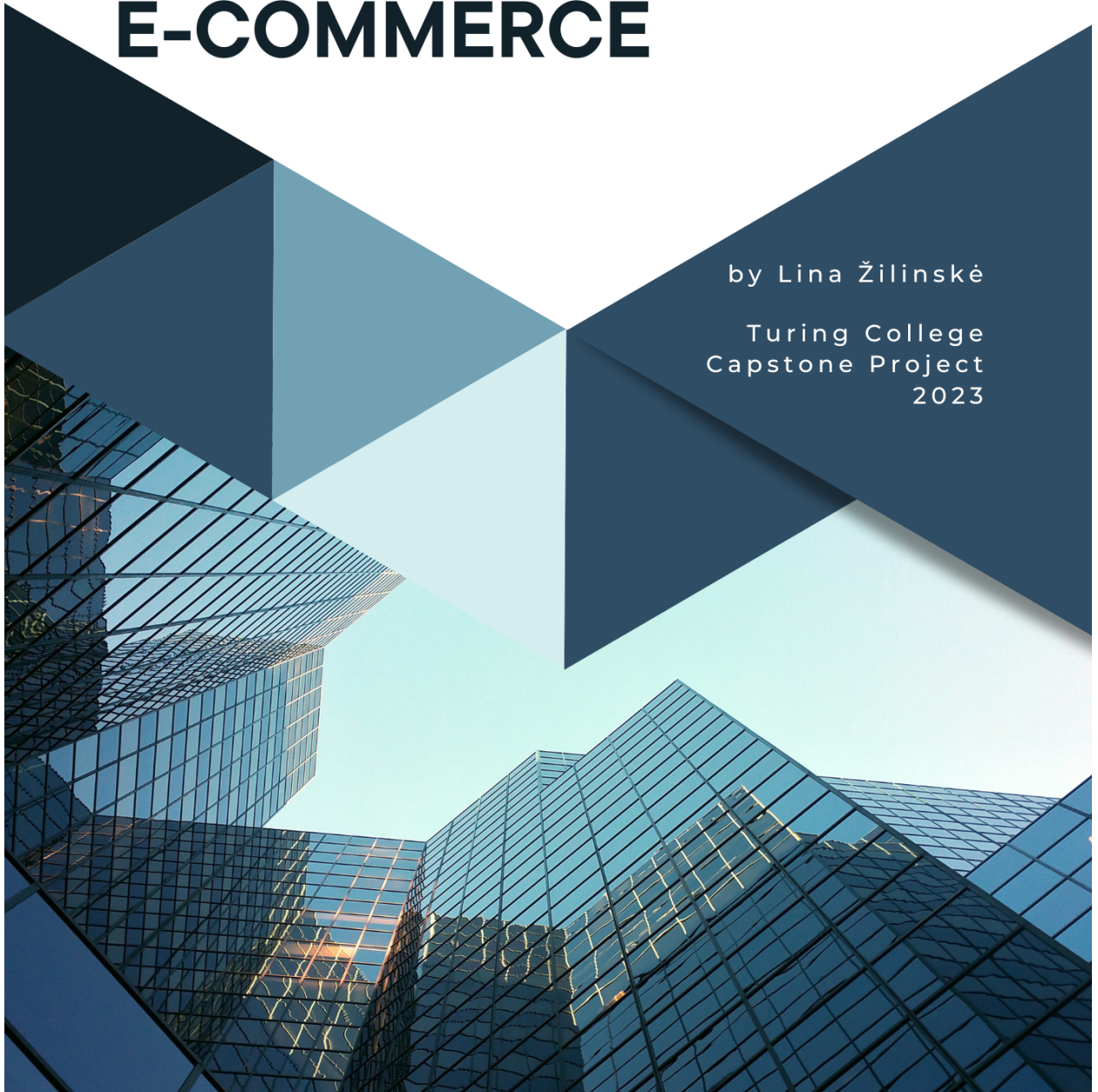




DATA ANALYSIS AND INSIGHTS OF THELOOK E-COMMERCE

by Lina Žilinskė

Turing College
Capstone Project
2023



Contents

Contents	1
Executive Summary	2
1. Project Background	3
Datasource	3
Targeted audience	3
Hypothesis and Aim	3
Context of the data	4
2. Analysis	5
RFM Analysis	9
Cohort Analysis	14
Customer Lifetime Value	16
3. Results	20
4. Conclusion	22
5. Recommendations	22

Executive Summary

TheLook is a fictitious E-Commerce clothing site developed by the Looker team and is based in the USA. The data source for this project is a public dataset found in the BigQuery platform, dataset ID: bigquery-public-data.thelook_ecommerce.

The targeted audience of this report is theLook marketing team. They want to determine if there are any traffic sources that should be discontinued or, conversely, require increased investment in terms of making the purchase. The marketing manager and team aim to obtain key insights regarding the traffic source that attracts the most clients and generates the highest revenue. Additionally, they want to identify which source is more likely to have returning customers.

The hypothesis I am going to test is that paid sources are more likely to bring more customers and revenue.

A comprehensive analysis was conducted, including general timeline analysis, customer segmentation using RFM, Cohort and Customer Lifetime Value (CLV) analysis.

Hypothesis that paid sources are more likely to bring more customers and revenue confirms partly. Most revenue and clients are brought through Search registered clients (70%), however we need to acknowledge that those clients further make their purchases depending on different marketing campaigns, like Email and Adwords and then generate better CLV.

Paid registration traffic cohorts showed better user retention rates, *Display* having the best retention rate at month 12.

A deeper analysis is needed to evaluate marketing campaigns (Email, Adwords, Facebook, YouTube and Organic) separately based on other metrics like ROI, CAC. Recommended to improve customer retention in the onboarding stage and pay attention to customer segments that need attention and are at risk. It is always best to have a variety of traffic sources. This minimizes the risk of the website being decimated if the main traffic source dries up.

1. Project Background

Datasource

For this Capstone project I will follow the steps of the data analysis process: Ask, Prepare, Process, Analyze, Share, and Act.

Data cleaning, validation, initial exploration of datasets and further analysis done in SQL and Google Spreadsheets. Data visualization done using Google Looker Studio.

Targeted audience

The marketing department wants to determine if there are any marketing traffic sources that should be discontinued or, conversely, require increased investment in terms of making the purchase. The marketing manager and team aim to obtain key insights regarding the traffic source that attracts the most clients and generates the highest revenue. Additionally, they want to identify which source is more likely to have returning customers.

Hypothesis and Aim

Paid sources are more likely to bring more customers and revenue.

Business Task:

Analyze TheLook E-Commerce and gain insights into customers' behavior after registering to the website through various user traffic sources and determine which marketing campaigns are most effective in driving purchases.

Business Objectives:

- What are the traffic sources customers can register to the website through?
- How do these sources differentiate during the timeline?
- Sources ranking in terms of revenue and number of orders.
- What is the customer segmentation according to the traffic sources?

- What do customers' cohorts look like?
- Do customers tend to make purchases after clicking on different marketing campaigns, or do they tend to stick to one particular campaign?
- What is the customer lifetime value and does it depend on the registration source?

Context of the data

TheLook is a fictitious eCommerce clothing site developed by the Looker team. The dataset contains information about customers, products, orders, logistics, web events and digital marketing campaigns.

File can be accessed:

<https://console.cloud.google.com/marketplace/product/bigquery-public-data/thelook-ecommerce?project=tc-da-1&pli=1>

Data contains 7 tables (1 figure), nearly 125K rows in the *orders* table and almost 2,418K rows in the *events* table, 100K distinct registered users, 125K orders, total revenue >10M USD.

The dataset contains events from 2019-01-02 until current date. Company is selling clothing categories for men and women, such as: Tops, Accessories, Socks, Swim, Outwear etc. Company is getting traffic from Google Adwords, Email campaigns, Youtube, Facebook and through organic searches.

2. Analysis

70 % of the users have registered through *Search* traffic source, 15 % through *Organic* search, the rest are taking 5 % each (*Facebook*, *Email* and *Display* advertising).

Now we can dig in to see how this is reflecting in marketing campaigns used for purchases and customer dynamics. Even though a customer registered through Facebook, for the purchase he might still use any other marketing campaign, like Youtube or Adwords.

Row	traffic_source	n_orders	revenue	AOV
1	Email	68175	5798325.285183...	85.05060924361...
2	Adwords	48315	3891748.503724...	80.54948781381...
3	YouTube	17677	1307751.431234...	73.98039436753...
4	Facebook	17448	1285355.521126...	73.66778548409...
5	Organic	8957	636869.5205056...	71.10299436258...

Figure 2.1. Purchase traffic source totals

Most of the orders and revenue 5,8 M eventually come from Email Marketing. Google Adwords also play an essential role and generate 3,8 M gross revenue. YouTube and Facebook both have very similar metrics and both take 22 % of the business. Organic source, being in second place by registrations, gets only 5 % of the purchase generated revenue.

It is not uncommon for paid ads to generate more business compared to organic searches. Paid ads have the advantage of being strategically placed in front of targeted audiences, allowing business to reach potential customers who may not have discovered them through organic searches. In our case we must admit that Organic registrations can provide long-term benefits, as a well-optimized website can attract consistent organic traffic over time without ongoing advertising costs.

Canceled and returned orders are taking 25% of all orders in monetary value.

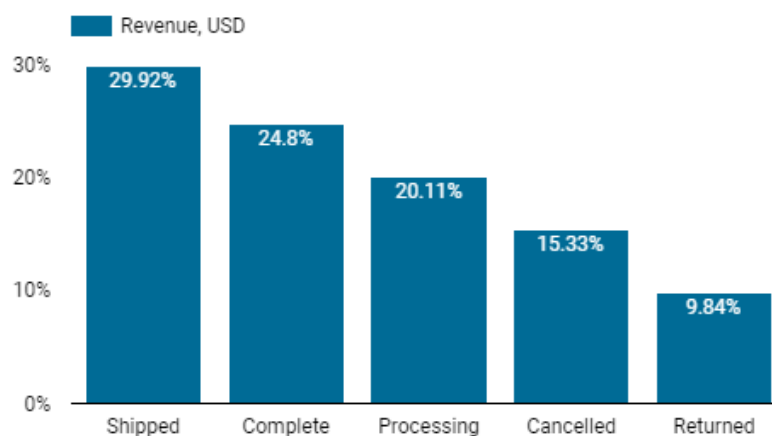


Figure 2.2. Order status

Revenue, including *Shipped, Complete, Processing, Canceled and Returned* orders is considered as Gross Revenue.

Revenue, excluding *Canceled and Returned* orders is considered as Net Revenue.

Returned and Canceled orders excluded from Looker revenue tables, now showing that most of the Net_revenue is generated through Email and Adwords campaigns.

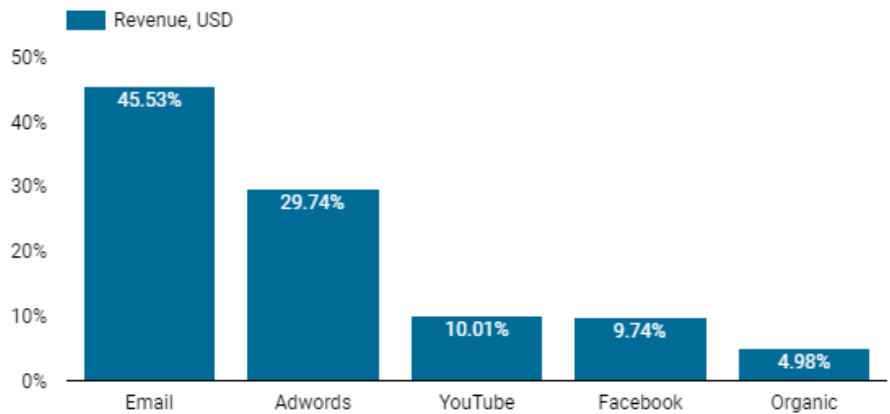


Figure 2.3. Purchase traffic source Net revenue

There is an increase in Net revenue month by month and high increase since February 2023 due to increasing daily revenue. Also dramatical increase in May 2023 from 369K in February to 605K Net Revenue, 64% more orders.

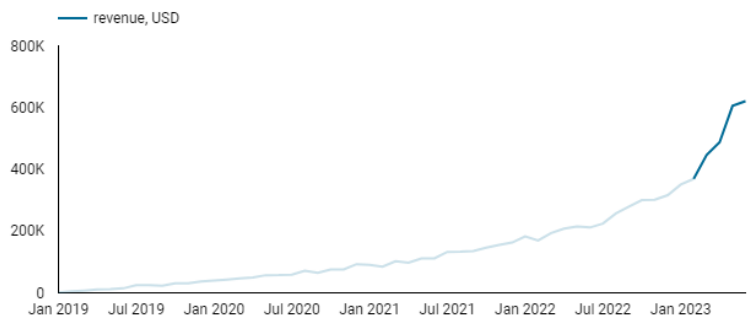


Figure 2.4. Revenue increase

Comparing year to date results to the previous year, there is also significant revenue increase in 2023 both by revenue and by number of orders. There is only half month data for June, but so far we can see that sales are growing.

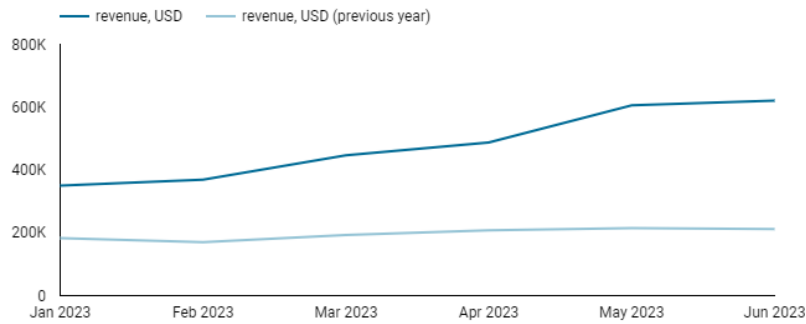


Figure 2.5. Revenue comparison

Search registration traffic source stands out, taking 70 % of the revenue, leaving all the rest of traffic way behind.

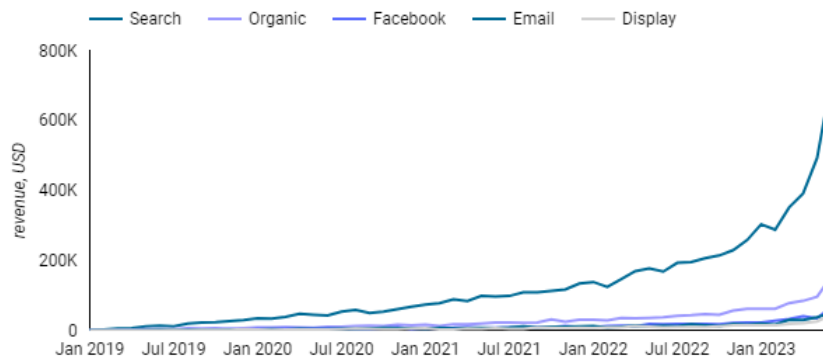


Figure 2.6. Sales by registration traffic source

There is an increase in all purchase traffic sources, however *Email* and *Adwords* are taking 75% of all Net revenue.

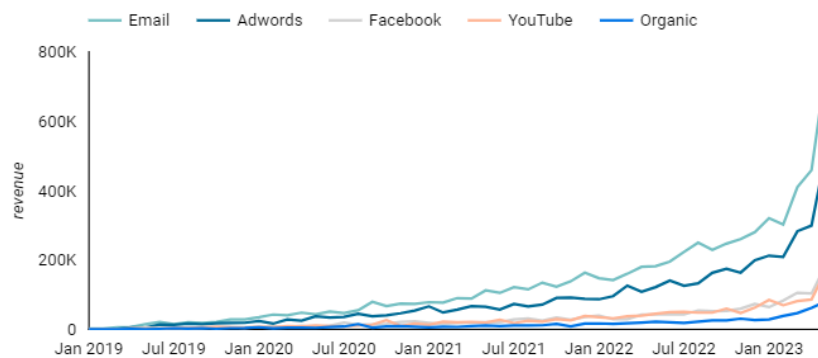


Figure 2.7. Purchase traffic source revenue

To compare with other channels (*Facebook*, *Youtube*, and *Organic*) both *Email* and *Adwords* have similar *Shipped*, *Completed* and *Processing* orders ratio vs *Cancelled* and *Returned*, approximately 3.00. Therefore no excessive cancellations or returns are expected in either source, that might be considered a bad result.

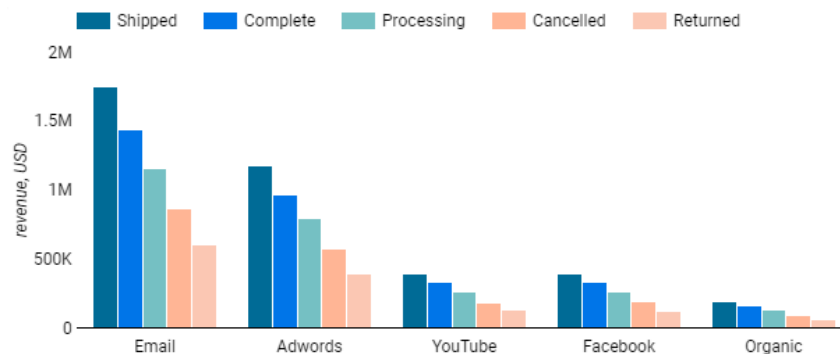


Figure 2.8. Purchase traffic source by status

RFM Analysis

I wanted to have a look at how customers are behaving in different traffic sources in terms of purchase frequency, revenue and days since last order, this is why RFM analysis was chosen to be processed.

Customers segmented into following groups:

- Champions,
- Loyal Customers,
- Potential Loyalists,
- Recent Customers,
- Promising,
- Need Attention,
- About to Sleep,
- At Risk,
- Lost.

Customers not active for 533-1579 days are considered as *Lost*.

Row	rfm_segment	customers	min_recency	max_recency	min_frequency	max_frequency	min_monetary	max_monetary	sum_monetary
1	Champions	7856	1	130	2	4	318.75	165968.1	51629024.8
2	Need Attention	11180	131	532	1	4	27.3	74942.6	13353909.7
3	Lost	14559	533	1579	1	4	14.0	130890.2	28526289.7
4	Loyal Customers	13145	1	275	1	4	59.8	278203.0	59280589.5
5	About to Sleep	2457	276	532	1	1	7.5	315.2	426095.9
6	Promising	3974	45	275	1	1	10.5	315.0	706429.6
7	Recent Customers	1632	1	44	1	1	7.5	315.4	284496.1
8	At Risk	4171	276	532	2	4	316.0	104579.1	25675124.9
9	Potential Loyalists	13826	1	275	1	4	0.6	108699.8	18195968.6

Figure 2.9. Customer segmentation

Lost customers take 20% of total customers. Best customers - *Champions* take only 12,6 %, however generate 35% of total revenue.

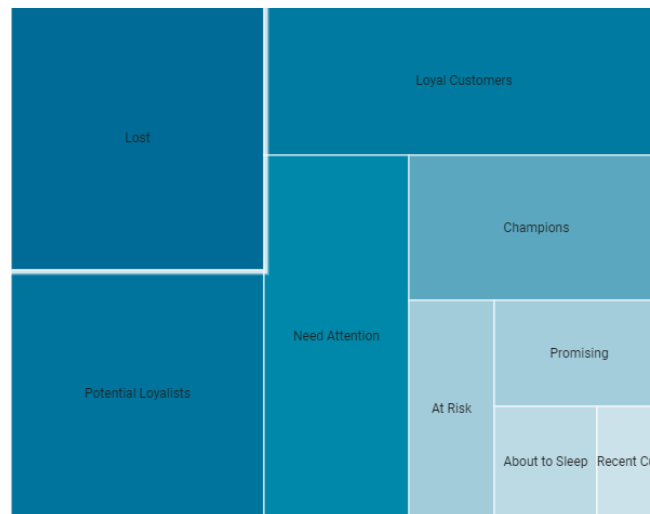


Figure 2.10. Customer segmentation treemap

Highest average monetary value belongs to *Champions*. These are our best customers having the biggest item basket, buying most frequently (3,1 times) and placing orders not long ago. Second and third positions go to segments *At Risk* and *Loyal Customers*. Their baskets are also expensive and they tend to buy frequently to compare to other segments.

rfm_segment	monetary ▾	frequency	recency
Champions	703.1	3.1	44.9
At Risk	551.5	2.6	419.2
Loyal Customers	424.1	2.3	134.1
Lost	150.9	1.3	878.6
Potential Loyalists	111.3	1.3	102.3
Need Attention	98.7	1.1	355
Promising	22.5	1	162
About to Sleep	22.4	1	432.5
Recent Customers	22	1	20.7

Figure 2.11. Customer segmentation values

In the following chart we can see that *Lost* customers are separated from the rest of the group. They were lost some 2,5 years ago. Company should pay attention to customers *Needing Attention* and *At Risk*, as it has been nearly a year since the last order.

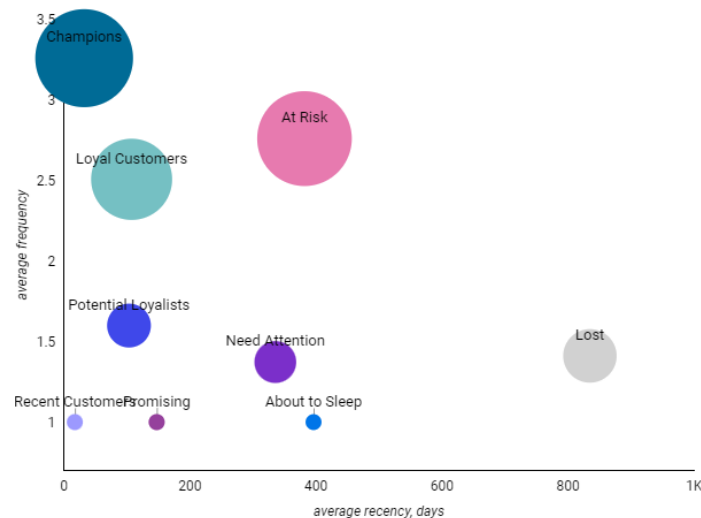


Figure 2.12. Customer segmentation bubbles.
Bubble size shows average monetary value

The company's most valuable customers are *Champions* and *Loyal Customers*. To attract *Champions*, introduce new products and implement loyalty programs. For *Loyal Customers* focus on upselling higher-value products, ask for reviews and engage them in different ways. There are 2 segments that need most of the attention at the moment. Segment *At Risk* needs to get personalized emails to reconnect, offer them renewals and provide helpful resources. Segment *Need Attention* need to get limited time offers. Send them reminder emails with offers what was trending and similar items to what was bought in the past.

There is a main difference between *Search* and *Facebook* registration traffic sources, where *Search* Champions on average spend \$565 and Facebook Champions - \$278. In general there is more revenue generated from Loyal Customers in Facebook source, rather than best customers Champions.

Order frequency by purchase traffic source is shown in the table below. Despite the fact that our best customers *Champions* have an average frequency of 2.95 and we have 5280 of them, numbers below show that customers do not stay 4 times in a row with the same marketing source a lot.

All five sources are mostly used for singular purchase, then clients tend to buy using different channels. However *Email* and *Adwords* got few clients attention 4 times in a row to make a purchase. *Facebook* and *YouTube* both have very similar results.

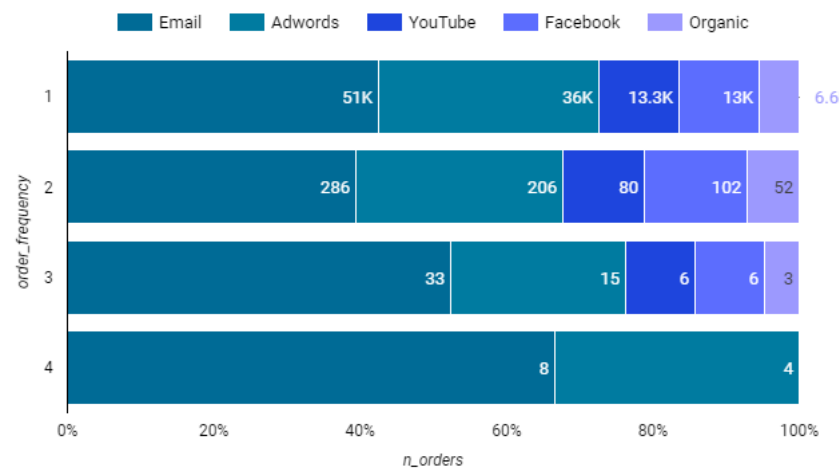


Figure 2.13. Purchase traffic order frequency

Among the customers from registration traffic sources, the results are very consistent. Regardless of the traffic source, the second purchase is made by 22 %, the third by 3%, and the fourth by 0.3%. Email cohort stands out, as its customers have never made more than three purchases.

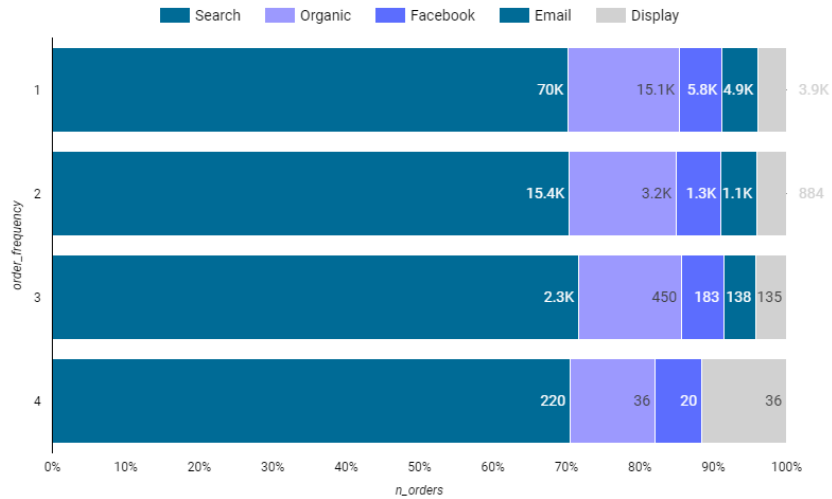


Figure 2.14. Registration traffic order frequency

Among all purchase traffic sources, customers tend to most frequently buy items with values ranging between \$10 and \$100. This trend is consistent for both Facebook and YouTube, as they show similar results. Additionally, items priced up to \$10 exhibit comparable popularity across the entire customer population.

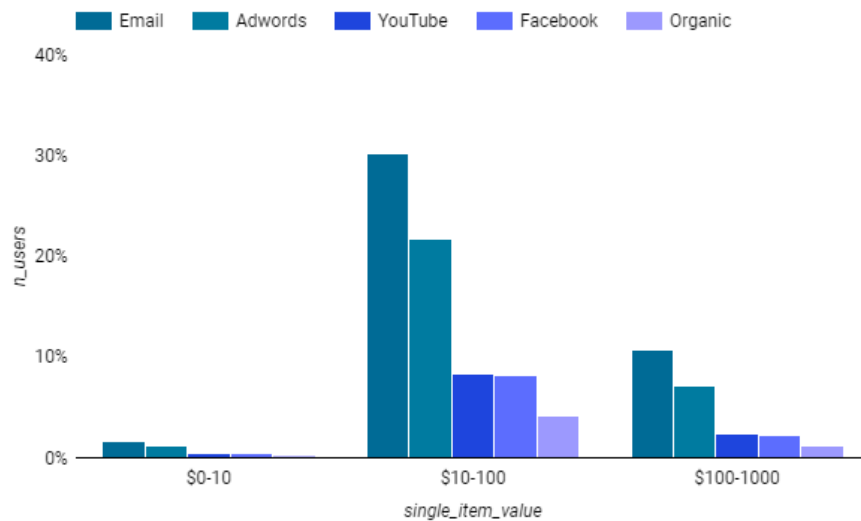


Figure 2.15. Single item value frequency

Cohort Analysis

It is important to understand our users, when and what they do, but also why they do it. Cohort analysis comes handy here to understand which cohorts were most successful and which struggled.

The table below indicates that after the registration, 5% of customers never made a purchase. Cohorts from 2022-04 until 2022-07 are having relatively better retention rates. Overall cohorts of 2022-04 and 2022-05 have best retention results for the first 4 months, cohorts of 2022-06 to 2022-08 have good results for the first 3 months, 2022-09 and 2022-10 for the first 2 months. 2022-08 cohort starts having really bad retention in month 8. All cohorts are struggling when 2023 March and April comes. Could be the offseason months. Need to check what marketing campaigns were used for cohorts of 2022-04 and 2022-05 and what were the most popular items sold between April and August to check if the summer season may have influenced customer activity.

COHORT	n_users	month_0	month_1	month_2	month_3	month_4	month_5	month_6	month_7	month_8	month_9	month_10	month_11	month_12
2022-04-01	100.00%	97.95%	93.91%	89.25%	84.46%	79.47%	73.10%	66.19%	59.48%	51.40%	43.33%	36.34%	27.79%	18.21%
2022-05-01	100.00%	97.75%	91.80%	87.70%	82.54%	76.85%	70.04%	62.30%	54.56%	45.97%	38.43%	28.24%	18.65%	
2022-06-01	100.00%	97.26%	91.71%	86.16%	79.88%	72.96%	65.90%	59.41%	50.83%	41.82%	30.50%	19.90%		
2022-07-01	100.00%	97.22%	90.71%	83.80%	76.68%	68.88%	60.95%	53.15%	43.12%	31.25%	20.41%			
2022-08-01	100.00%	97.33%	91.09%	84.06%	76.77%	65.71%	57.25%	47.23%	35.46%	22.90%				
2022-09-01	100.00%	97.59%	90.01%	80.77%	71.12%	62.44%	51.34%	38.80%	24.60%					
2022-10-01	100.00%	96.70%	87.92%	79.34%	69.44%	57.56%	44.03%	28.51%						
2022-11-01	100.00%	96.71%	85.38%	72.96%	61.70%	47.29%	31.02%							
2022-12-01	100.00%	94.65%	83.46%	71.58%	55.52%	37.11%								
2023-01-01	100.00%	93.59%	81.03%	64.14%	40.97%									
2023-02-01	100.00%	93.35%	72.65%	50.48%										
2023-03-01	100.00%	90.63%	62.57%											
2023-04-01	100.00%	85.46%												
2023-05-01	100.00%													
AVERAGE	100.00%	95.09%	85.19%	77.29%	69.91%	63.14%	56.70%	50.80%	44.68%	38.67%	33.16%	28.16%	23.22%	18.21%
WEEKLY CHURN		4.91%	9.90%	7.89%	7.38%	6.77%	6.44%	5.90%	6.12%	6.01%	5.50%	5.00%	4.94%	5.01%
Onboarding Stage					Nurture Stage									

Figure 2.16. Retention rate for the previous year

Looking at the average churn rates, we can observe that churn rates are relatively high in the early months (month_0, month_1, month_2). This suggests that the onboarding phase, where customers are new and still familiarizing themselves with the products or service, occurs within the first few months.

As the months progress, the churn rates gradually decrease and stabilize to lower percentages. This indicates that the nurture stage, where customers have become more engaged and accustomed to the product or service, starts around the 4th or 5th month and extends beyond the first year.

Therefore we can say that the onboarding phase lasts approximately for the first 2-3 months, and the nurture stage begins around the 4th or 5th month onwards.

It is important to note that the onboarding stage retention rate is getting worse. Improving onboarding stage retention is crucial for setting a strong foundation for long-term customer engagement.

It is necessary to have a look at how customers behave if we compare them registering from different user traffic sources (*Search, Facebook, Organic, Email* and *Display*). To compare user retention rate month by month and by user traffic sources we can see that there is no significant difference among paid and non paid traffic sources, however paid traffic sources tend to have slightly better retention rates, Display showing the best result at month 12.

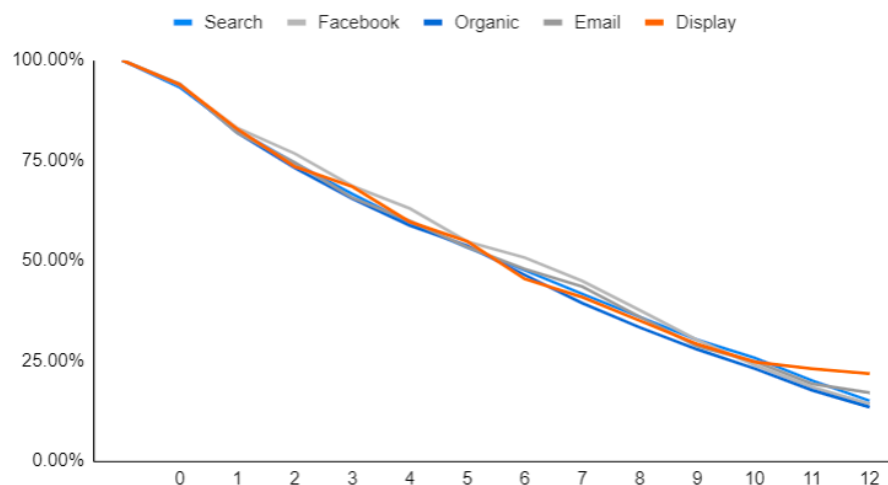


Figure 2.17. Retention rate for the previous year

Customer Lifetime Value

This approach calculates the average value of a cohort of customers who registered through the same traffic source during a specific time period. It tracks the collective behavior and value generated by that group of customers over time.

This approach will help us understand how much time needs to pass for the customer to start buying. It will also help us measure the effectiveness of marketing campaigns and customer acquisition strategies. It still helps identify the most valuable customers and tailor retention strategies accordingly.

Cohort lifetime value is used to evaluate the performance of specific cohorts or assess the impact of marketing campaigns. And with customer lifetime value we focus on understanding the long-term value and behavior of individual customers. Both approaches in conjunction help gain a comprehensive understanding of our customers' value and behavior.

In the table below we can see that since 2022-11 registration customers started to buy more on month 0. Sales increase is very well reflected in the last cohorts of 2023, where we can see the increasing revenue per customer. Customers tend to buy most in the first 4 months after the registration, and the average customer lifespan is 15 months, where the start date counted from customer registration date.

start_date	n_users	month_0	month_1	month_2	month_3	month_4	month_5	month_6	month_7	month_8	month_9	month_10	month_11	month_12	month_13	month_14	month_15	month_16	month_17
2022-01-01	1477	3.83	6.69	9.93	9.24	7.64	8.68	9.85	7.58	7.60	9.21	11.67	8.31	6.41	6.76	9.61	9.57	9.78	4.44
2022-02-01	1270	4.78	12.24	10.94	8.34	11.02	10.63	7.73	9.00	7.42	10.49	8.82	10.30	7.58	10.45	9.21	8.54	8.13	
2022-03-01	1521	5.53	10.16	9.70	8.91	10.76	13.13	11.10	9.30	9.17	12.85	11.76	8.38	10.77	10.19	9.20	6.72		
2022-04-01	1469	4.01	9.28	11.35	12.31	10.79	8.58	10.67	13.20	9.85	11.61	10.73	12.48	12.47	10.50	8.04			
2022-05-01	1455	4.26	9.98	12.78	10.28	11.38	12.10	11.23	13.65	13.55	11.01	13.26	10.93	12.82	7.65				
2022-06-01	1426	6.25	15.38	12.30	13.56	15.33	13.74	11.76	12.05	11.04	15.02	11.17	11.74	7.98					
2022-07-01	1512	5.91	13.49	15.28	13.79	13.26	14.67	17.41	15.06	16.04	13.23	14.42	8.40						
2022-08-01	1479	7.64	13.33	16.74	15.70	16.24	17.27	15.13	18.34	15.36	14.86	8.29							
2022-09-01	1436	7.26	17.60	16.32	17.85	18.29	16.51	18.65	17.82	18.14	10.17								
2022-10-01	1492	8.76	21.77	19.34	21.38	17.65	22.30	17.55	18.61	12.14									
2022-11-01	1445	11.82	21.54	23.93	20.99	23.97	21.44	21.88	13.84										
2022-12-01	1502	14.71	24.17	25.05	26.92	30.05	28.39	15.98											
2023-01-01	1511	12.60	27.81	29.73	32.25	32.54	19.23												
2023-02-01	1383	19.82	43.92	37.81	41.10	22.05													
2023-03-01	1462	23.27	45.64	53.95	31.46														
2023-04-01	1338	28.76	84.02	47.93															
2023-05-01	1436	64.23	109.57																
2023-06-01	2837	211.38																	
Grand Total		\$24.71	\$28.62	\$22.07	\$18.94	\$17.21	\$15.90	\$14.08	\$13.50	\$12.03	\$12.05	\$11.27	\$10.08	\$9.67	\$9.11	\$9.01	\$8.28	\$8.95	\$4.44

Figure 2.20. Customer Lifetime Value Cohorts

Paid versus non paid traffic sources have lower customer value (CV) at month 1, however later numbers show that non paid sources are more consistent and eventually bring higher CLV.

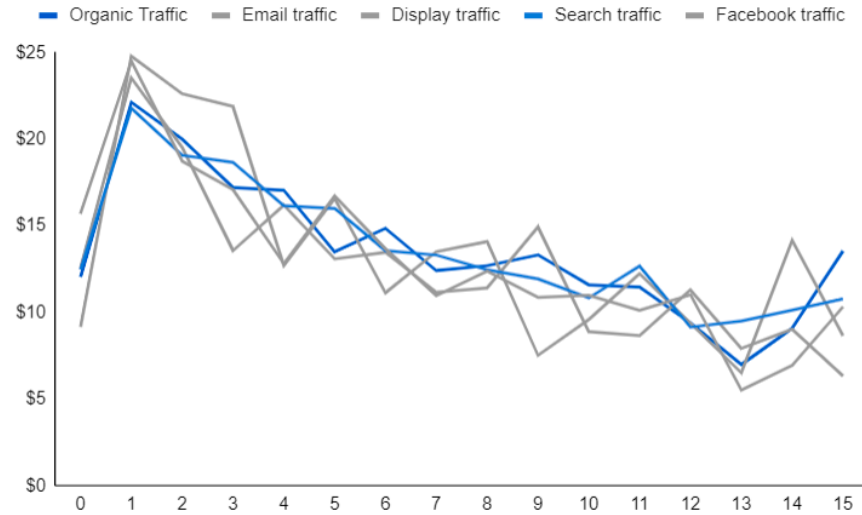


Figure 2.21. Customer Value in different user traffic sources

Search CLV prediction table shows that the 2023-05 cohort will generate 4 times more revenue than the year before. All monthly cohorts average prediction for the 15 months is \$246. Facebook predicted CLV is \$211.

start_date	month_1	month_2	month_3	month_4	month_5	month_6	month_7	month_8	month_9	month_10	month_11	month_12	month_13	month_14	month_15	
2022-03-01															149.45	149.45
2022-04-01														155.98	164.07	160.03
2022-05-01													166.16	174.68	183.74	174.86
2022-06-01												158.50	166.50	175.04	184.12	171.04
2022-07-01											167.63	176.20	185.09	194.59	204.68	185.64
2022-08-01										156.34	168.26	176.86	185.79	195.32	205.45	181.34
2022-09-01									150.99	161.52	173.83	182.72	191.95	201.79	212.26	182.15
2022-10-01								149.95	162.41	173.73	186.98	196.54	206.46	217.05	228.31	190.18
2022-11-01							154.55	169.25	183.32	196.10	211.05	221.84	233.04	244.99	257.70	207.98
2022-12-01						154.79	172.29	188.67	204.36	218.60	235.26	247.29	259.78	273.10	287.27	224.14
2023-01-01					166.07	187.69	208.91	228.77	247.79	265.06	285.27	299.85	315.00	331.15	348.33	262.17
2023-02-01				147.80	174.63	197.37	219.68	240.57	260.57	278.73	299.98	315.31	331.24	348.23	366.29	265.03
2023-03-01			151.55	185.56	219.25	247.80	275.81	302.04	327.14	349.95	376.62	395.88	415.87	437.20	459.88	318.81
2023-04-01		167.45	226.05	276.79	327.03	369.62	411.40	450.51	487.96	521.98	561.77	590.49	620.31	652.12	685.96	453.53
2023-05-01	143	222.12	299.87	367.16	433.82	490.31	545.73	597.62	647.29	692.41	745.20	783.30	822.86	865.06	909.94	571.03

Figure 2.22. Predicted Customer Lifetime Value

Overall, more customers wait a bit for their first order rather than buying products just after the registration. Email registered clients buy least straight after the registration but they show best results for the next 3 months.

For the selected period from 2022-01-01 to 2023-05-31, the number of users and CLV generated by each traffic source is as follows:

Traffic source	CLV	Predicted CLV	Growth
Search	\$218	\$246	12.84%
Facebook	\$208	\$211	1.44%
Organic	\$217	\$236	8.76%
Email	\$214	\$217	1.40%
Display	\$201	\$220	9.45%

Non paid traffic sources generated better CLV than paid, and predicted growth is on average 10%.

The *Display* traffic source has a relatively lower CLV compared to other user traffic sources. However, the predicted CLV is higher, indicating potential growth and improved customer value in the future. This suggests that the *Display* campaign is performing well and has the potential to yield higher returns over the next 15 months.

Focusing on optimizing *Email* and *Facebook* campaigns and nurturing customer relationships could lead to higher CLV in the future, however predicted growth is quite low and should be investigated further.

While the number of users for each traffic source varies significantly, it is important to consider the quality of those users and their CLV. In this case, the *Search* traffic source has the highest predicted CLV at \$246, followed closely by *Organic* traffic at \$236.

However, making a conclusive judgment solely based on these numbers would be premature. To fully assess which traffic source is better, we should consider additional metrics such as conversion rates, customer acquisition costs, and overall return on investment (ROI) for each traffic source.

Additionally, it is essential to conduct statistical analysis to determine if the observed differences in CLV are statistically significant. Factors like seasonality, market trends, and other external factors should also be taken into account to make a well-informed decision.

Therefore, it is not possible to definitely state which source is better in terms of CLV. Further analysis and consideration of multiple factors are necessary to draw meaningful conclusions.

3. Results

Business task to analyze TheLook E-Commerce and gain insights into customers' behavior after registering to the website through various user traffic sources and to determine which marketing campaigns are most effective in driving purchases was analyzed.

The following questions were raised and analyzed:

1. What are the traffic sources customers can register to the website through?

- Customers can register through various traffic sources: Search, Organic, Facebook, Email and Display.
- 70% of the customers used Search traffic source, 15% Organic, Facebook, Email and Display - 5%.
- Purchase though is processed through Email, Adwords, *Facebook, Youtube, and Organic* traffic, Email getting 45% of the revenue and Adwords - 30.

2. How do these sources differentiate during the timeline?

- As the number of customers increases, so does the revenue. Most new customers originate from **Search** registration traffic source, and there is a significant increase in customer acquisition starting from February 2022. Another notable peak occurs in February 2023, indicating further growth in customer augmentation.
- Purchase traffic sources exhibit a similar trend of growth, however there are two marketing categories that predominantly attract customers to make a purchase: **Email** and **Adwords**.

3. Sources ranking in terms of revenue and number of orders.

- Search registration traffic source generates 70% of the revenue and number of orders. Organic 15%, Facebook 6%, Email 4.8 % and Display 4.2%.
- Following the number of customers distribution during the registration (70/15/5/5/5) we can see that Facebook customers generated a bit more revenue, whereas Email and Display customers - less.

- Purchase traffic sources sequence is as follows: Email 45% revenue, Adwords 30%, Facebook 10%, YouTube 10%, Organic 5%.
- Highest **ARPU** value belongs to Email purchase campaign - \$98 and lowest to Organic purchase campaign - \$73.

4. What is the customer segmentation according to the traffic sources?

- Only general customer segmentation was counted, revealing that the company has 20% Lost customers. The best customers are named Champions and Loyal Customers, we have 26% of them. The third category to be mentioned is customers who need attention at the moment, and they take 21% out of the total.
- All five purchase traffic sources are mostly used for singular purchase, then clients tend to buy using different channels. However *Email* and *Adwords* got *few clients attention* 4 times in a row to make a purchase. *Facebook* and *YouTube* both have very similar results.

5. What do customers' cohorts look like?

- Although revenue is increasing, cohort analysis revealed that customer retention has been dropping down in the past cohorts. Retention experiences significant drop during the months of March and April throughout all cohorts.
- The onboarding stage takes 3 months, when 30% of the customers churn.
- Paid registered traffic source cohorts have slightly better retention rates.
- Acquisition cohorts used to identify when new users are churning. Most new users churn during the first 3 onboarding months.

6. Customer Lifetime Value

- Analysis revealed that CLV increases significantly after February 2023.
- Organic and Email registered cohorts have better CLV for the first 5 months, however Organic continues to have the best results throughout the whole period.

4. Conclusion

Hypothesis that paid sources are more likely to bring more customers and revenue confirms partly. Most revenue is brought through Search registered clients, and this source brings the

highest CLV. However, we need to acknowledge that those acknowledged clients further make their purchases depending on different marketing campaigns, like Email and Adwords; Email having the highest ARPU - 98\$, while Organic purchases lowest - \$73.

A deeper analysis is needed to include conversion rates, bounce rates, and additional marketing metrics, such as ROI, CAC, CPC and others.

5. Recommendations

1. Further analysis is needed to analyze good and bad performing marketing campaigns. Compare the content of what was published, including impressions and clicks, conversion rates, acquisition costs, ROI and other relevant metrics to make informed decisions about the effectiveness of each campaign. Additionally, conducting A/B testing and monitoring the performance of different campaigns over time can provide a more accurate understanding of their relative impact on customer lifetime value.
2. Improve customer retention in the **onboarding stage**. Despite increasing sales, the retention of customers might be declining. If customers are not staying engaged with the business for as long as expected, it can lower the CLV eventually. Enhance the customer experience, implement loyalty programs, offer personalized recommendations, and provide exceptional customer service to increase customer loyalty and lifetime value. Gather feedback from customers during and after the onboarding process to understand their pain points, challenges, or suggestions for improvement.
3. Pay attention to customer segments that **Need Attention** and **are At Risk**. Engage Loyal and Potential To Loyal to become champions.
4. It is always best to have a variety of traffic sources. This minimizes the risk of website being decimated if the main traffic source dries up.