E-Commerce Cohort Analysis

Table of Contents

	A	B
	Table of Contents	
	Sheet Name	**Description**
	Executive Summary	A brief summary of findings and suggestions.
	conversion_funnel	Tracks user progression from product views through shopping cart additions to completed purchases, with conversion rates at each stage. Includes unique user counts and cumulative conversion rate.
	cohort_analysis	Displays monthly user retention patterns across cohorts, tracking unique user counts from acquisition (month 0) through subsequent months (1-4)
	retention_rates	Displays monthly user retention patterns across cohorts as percentages.
	purchase_activity	Lightly filtered version of raw_user_activity showing only purchase activity.
	first_purchase	Further filtered version of purchase_activity showing a unique users ID first purchase date in "yyyy-mm" format.
	raw_user_activity	Basic unfiltered data

Legend

Summaries	
Calculations	
Analysis	
Raw Data	

Executive Summary

This project shows how raw e-commerce event logs can be turned into clear, actionable business metrics using advanced spreadsheet analytics. As a junior analyst, I was asked to dig into user activity data to uncover insights about how people use the site, how well the site converts visitors into buyers, and how long customers stick around. The analysis is organized step-by-step, from cleaning the data to running a cohort analysis, and everything is documented for transparency. The findings offer a practical roadmap for improving user engagement and growing the business.

Project Overview

When I joined the e-commerce team, my first assignment was to analyze user activity and help the company better understand its customers. This project uses detailed event logs to answer two main questions:

1. How well does the website turn product views into purchases?
2. How well does the company keep customers coming back over time?

Raw Data

user_id	event_type	category_code	brand	price	event_date	first_purchase_date	event_month	first_purchase_month	cohort_age
a151591562551	purchase	computers.peripherals.camera	logitech	123.35	2020-09-24	2020-09-24	2020-09	2020-09	0
a151591562551	purchase	computers.peripherals.camera	logitech	123.35	2020-09-24	2020-09-24	2020-09	2020-09	0
a151591562551	purchase	computers.peripherals.camera	logitech	123.35	2020-09-24	2020-09-24	2020-09	2020-09	0
a151591562535	purchase	computers.components.tv_tuner	perfeo	19.49	2020-09-24	2020-09-24	2020-09	2020-09	0
a151591562551	purchase	stationery.cartrige	hp	71.02	2020-09-24	2020-09-24	2020-09	2020-09	0
a151591562551	purchase	stationery.cartrige	hp	71.02	2020-09-24	2020-09-24	2020-09	2020-09	0
a151591562551	purchase	stationery.cartrige	hp	71.02	2020-09-24	2020-09-24	2020-09	2020-09	0
a151591562550	purchase	appliances.environment.vacuum	topon	35.87	2020-09-24	2020-09-24	2020-09	2020-09	0
a151591562551	purchase	stationery.cartrige	hp	71.02	2020-09-24	2020-09-24	2020-09	2020-09	0
a151591562538	purchase	stationery.cartrige	cactus	24.73	2020-09-25	2020-09-25	2020-09	2020-09	0
a151591562538	purchase	stationery.cartrige	cactus	24.73	2020-09-25	2020-09-25	2020-09	2020-09	0
a151591562551	purchase		kitfort	140.63	2020-09-25	2020-09-25	2020-09	2020-09	0
a151591562551	purchase		kitfort	140.63	2020-09-25	2020-09-25	2020-09	2020-09	0
a151591562552	purchase	stationery.cartrige	hp	29.57	2020-09-26	2020-09-26	2020-09	2020-09	0
a151591562552	purchase	auto.accessories.player	kenwood	104.86	2020-09-26	2020-09-26	2020-09	2020-09	0
a151591562552	purchase	computers.peripherals.printer	epson	647.29	2020-09-26	2020-09-26	2020-09	2020-09	0
a151591562552	purchase	computers.peripherals.printer	epson	647.29	2020-09-26	2020-09-26	2020-09	2020-09	0
a151591562552	purchase	computers.peripherals.printer	epson	647.29	2020-09-26	2020-09-26	2020-09	2020-09	0
a151591562552	purchase	computers.peripherals.wifi	ubiquiti	155.43	2020-09-27	2020-09-27	2020-09	2020-09	0
a151591562552	purchase	computers.peripherals.wifi	ubiquiti	130.73	2020-09-27	2020-09-27	2020-09	2020-09	0
a151591562552	purchase	computers.peripherals.wifi	ubiquiti	155.43	2020-09-28	2020-09-27	2020-09	2020-09	0
a151591562537	purchase	stationery.cartrige	cactus	20.14	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562537	purchase	stationery.cartrige	cactus	20.14	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562537	purchase	stationery.cartrige	cactus	20.14	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562537	purchase	computers.peripherals.printer	canon	192.86	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562537	purchase	stationery.cartrige	cactus	20.14	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562552	purchase	stationery.cartrige	nv-print	17.62	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562552	purchase	stationery.cartrige	nv-print	17.62	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562552	purchase	stationery.cartrige	nv-print	17.62	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562552	purchase	stationery.cartrige	nv-print	17.62	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562552	purchase	stationery.cartrige	nv-print	17.62	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562551	purchase	computers.peripherals.printer		10.78	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562551	purchase	computers.peripherals.printer		10.78	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562551	purchase	computers.peripherals.printer		10.84	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562551	purchase	computers.peripherals.printer		10.78	2020-09-28	2020-09-28	2020-09	2020-09	0
a151591562552	purchase		a-data	17.89	2020-09-28	2020-09-28	2020-09	2020-09	0

Data Cleaning:

- Started with the full, cleaned event log containing all user activities and details.
- Filtered the data to include only rows where the event type was a purchase.
- Removed any duplicate purchase events for the same user on the same date.
- Sorted the filtered purchase data by user ID and event date.
- Identified the earliest purchase date for each user.
- Created a new table with just user_id and their corresponding first purchase date.
- Ensured each user appeared only once in the new table.
- Checked for and removed any users with missing purchase dates.
- Standardized the date format for consistency.
- Saved this two-column table as the foundation for cohort and retention analysis.

user_id	event_type	category_code	brand	price	event_date
a15159156255519014356	view	computers.peripherals.printer	pantum	113.81	2020-09-24
a15159156255519392599	view	stationery.cartrige	lomond	8.35	2020-09-24
a15159156253369253995	view	construction.tools.welding		250.63	2020-09-24
a15159156253369253995	view	construction.tools.welding		250.63	2020-09-24
a15159156255519410002	view	computers.desktop	steelseries	58.49	2020-09-24
a15159156255519413944	view	appliances.kitchen.microwave	samsung	130.11	2020-09-24
a15159156255519397646	view	computers.peripherals.camera	logitech	123.35	2020-09-24
a15159156255519397646	shopping_cart	computers.peripherals.camera	logitech	123.35	2020-09-24
a15159156255519397646	purchase	computers.peripherals.camera	logitech	123.35	2020-09-24
a15159156255519397646	view	computers.peripherals.camera	logitech	123.35	2020-09-24
a15159156255519417750	view	computers.network.router	d-link	53.14	2020-09-24
a15159156255519397646	shopping_cart	computers.peripherals.camera	logitech	123.35	2020-09-24
a15159156255519397646	purchase	computers.peripherals.camera	logitech	123.35	2020-09-24
a15159156255519397646	purchase	computers.peripherals.camera	logitech	123.35	2020-09-24
a15159156253358482476	view	computers.components.tv_tuner	perfeo	19.49	2020-09-24
a15159156253358482476	shopping_cart	computers.components.tv_tuner	perfeo	19.49	2020-09-24
a15159156253358482476	purchase	computers.components.tv_tuner	perfeo	19.49	2020-09-24
a15159156255519443407	view	computers.components.motherboard	asrock	77.73	2020-09-24
a15159156255519443407	shopping_cart	computers.components.motherboard	asrock	77.73	2020-09-24
a15159156255519443407	view	computers.components.motherboard	asrock	77.73	2020-09-24
a15159156255519446240	view		formula	35.54	2020-09-24
a15159156255519452732	view	computers.notebook	hp	27.62	2020-09-24
a15159156255519460662	view	computers.peripherals.joystick	logitech	162.51	2020-09-24
a15159156255519460662	shopping_cart	computers.peripherals.joystick	logitech	162.51	2020-09-24
a15159156255519460662	view	computers.peripherals.joystick	logitech	162.51	2020-09-24
a15159156255519460662	view	computers.peripherals.joystick	logitech	193.98	2020-09-24
a15159156255519463641	view	stationery.cartrige	hp	71.02	2020-09-24
a15159156255519463641	view	stationery.cartrige	hp	71.02	2020-09-24
a15159156255517481778	view			29.24	2020-09-24
a15159156255519463641	shopping_cart	stationery.cartrige	hp	71.02	2020-09-24
a15159156255517481778	view			29.24	2020-09-24
a15159156253355805313	view	appliances.sewing_machine	janome	470.38	2020-09-24
a15159156255517481778	view			29.24	2020-09-24
a15159156253355805313	view	appliances.sewing_machine	janome	314.97	2020-09-24
a15159156255517481778	view			29.24	2020-09-24
a15159156255519463641	purchase	stationery.cartrige	hp	71.02	2020-09-24

user_id	MIN of event_da
a1515915625355179497	2020-11-04
a1515915625355805313	2020-09-29
a1515915625358482476	2020-09-24
a1515915625359833886	2020-11-20
a1515915625368479327	2020-11-01
a1515915625371158307	2020-09-28
a1515915625385482819	2020-10-01
a1515915625385888878	2020-09-25
a1515915625393382317	2020-11-03
a1515915625395096356	2020-10-17
a1515915625403460003	2020-10-11
a1515915625413257995	2020-10-09
a1515915625415474736	2020-10-03
a1515915625466195424	2020-10-10
a1515915625469806973	2020-10-04
a1515915625476168209	2020-10-10
a1515915625483402054	2020-11-09
a1515915625486521695	2020-10-05
a1515915625503751892	2020-09-24
a1515915625505008297	2020-09-29
a1515915625507596879	2020-11-08
a1515915625507754425	2020-10-13
a1515915625509677596	2020-12-20
a1515915625509808717	2020-10-03
a1515915625510154735	2020-10-21
a1515915625510668770	2020-11-17
a1515915625511110840	2020-10-08
a1515915625511258041	2020-09-29
a1515915625513373329	2020-10-15
a1515915625515426823	2020-10-18
a1515915625515626372	2021-01-07
a1515915625517871720	2020-11-11
a1515915625518472854	2020-10-14
a1515915625518836854	2020-11-09
a1515915625518894671	2020-10-21
515915625519397646	2020-09-24

Conversion Funnel:

- Built a three-stage funnel (product view → cart open → purchase) using pivot tables.
- Counted unique users at each stage to measure how well the site converts visitors.
- Calculated both total and step-by-step conversion rates to spot where users drop off.

A	B	C	D
event_type	user_id_CNTUNQ	CR	Total_CR
view	10453	100.0%	100.0%
shopping_cart	3036	29.0%	29.0%
purchase	1081	35.6%	10.3%

Cohort Analysis Preparation:

- Isolated purchase events and found each user's first purchase date.
- Grouped users into cohorts based on the month of their first purchase.
- Added columns for event month, first purchase month, and cohort age to track users over time.

A	B	C	D	E	F
COUNTUNIQUE of user_id	cohort_age				
first_purchase_month	0	1	2	3	4
2020-09	32	4	2		1
2020-10	187	14	7	1	1
2020-11	238	13	7	1	
2020-12	203	9	6		
2021-01	233	16			
2021-02	188				

Retention Rate Calculation:

- Grouped purchase data by cohort and cohort age.
- Calculated month-by-month retention rates to see how customer engagement changes.
- Summarized the findings in a dedicated retention rates sheet for easy review.

cohort_age	1	2	3	4
first_activity_month				
2020-10	12.50%	6.25%	0.00%	3.13%
2020-11	7.49%	3.74%	0.53%	0.53%
2020-12	5.46%	2.94%	0.42%	0.00%
2021-01	4.43%	2.96%	0.00%	0.00%
2021-02	6.87%	0.00%	0.00%	0.00%

Results

The conversion funnel showed that many users drop off between viewing products and making a purchase. Step-by-step conversion rates highlighted where the biggest opportunities are to improve the user journey. Cohort analysis revealed that customer retention drops over time, but some groups of users stay engaged longer than others. These insights give the company a solid foundation for improving both conversion and retention.

Conclusions

User behavior on the site follows clear patterns: most people look at products, but fewer actually buy. Retention rates go down as time passes, which means it's important to engage new customers early. By focusing on the user journey and understanding how different groups behave, the company can boost both sales and long-term loyalty. Making decisions based on data is key to staying competitive in e-commerce.

Recommendations

Make the website experience smoother to reduce drop-offs between product views and purchases. Launch targeted marketing campaigns for new users to increase early retention. Regularly analyze how different customer groups perform to find high-value segments. Keep refining analytics and tracking to spot new opportunities for growth and engagement.

Executive Summary

Results	**Synopsis**
Conversion Funnel	On average, 5-10% of website visitors add items to their cart, and 20-40% of those shoppers complete their purchase. With a pageview-to-cart conversion rate of 29%, your website is significantly outperforming industry benchmarks. Your cart-to-purchase conversion rate of 36%, while solid, presents an opportunity for improvement when compared to your exceptional front-end performance. Consider implementing a "buy now" option to streamline the checkout process and capitalize on your strong initial engagement.
Retention Rates	Your retention metrics reveal a concerning pattern: user activity drops dramatically across all cohorts after the first month. While initial advertising campaigns successfully drive pageviews, customer loyalty appears to be your primary challenge. To address this, consider implementing a tiered loyalty program that rewards continued engagement and developing community-focused campaigns that foster belonging and ongoing connection with your brand. These targeted approaches should help convert one-time purchasers into long-term, valuable customers.

Analysis	**Description**
Raw Data	This dataset captures customer interactions on our e-commerce platform, tracking each user's journey from browsing to purchase. It includes unique customer identifiers, user actions (viewing, cart additions, purchases), product details (categories and brands), pricing information, and precise timestamps. This comprehensive data allows us to analyze conversion funnels, identify purchasing patterns, evaluate product performance, and create targeted customer segments to drive business growth.
Conversion Funnel	Our analysis of user interaction with the website reveals a clear three-stage conversion path from product views to completed purchases. Of 10,453 visitors who viewed products, 29% (3,036 users) progressed to adding items to their shopping carts—a strong initial engagement rate that exceeds typical e-commerce benchmarks. However, only 35.6% of users with items in their cart completed purchases, resulting in a total conversion rate of 10.3% from view to purchase. This data highlights two key insights: (1) our product pages effectively generate interest and consideration, and (2) significant revenue opportunity exists in optimizing the checkout experience to reduce the 64.4% cart abandonment rate. Targeted improvements to the purchase stage could substantially increase overall conversion performance and revenue.
Retention Rates	Our review of customer data from Sept 2020-Feb 2021 shows promising initial engagement as well as opportunities for improvement in long-term retention. Please consider: 1. A tiered loyalty program to reward continued engagement. 2. Using customer segmentation techniques to develop a better understanding of the majority demographic of the user base. 3. Community integrated advertising based on the results of the market segmentation. A loyalty-focused approach could enhance your customer lifetime value and complement your already effective acquisition efforts.