

# Business Analytics Project for e-commerce company

Carlos Gallegos

# Table of Contents

Sheet Name	Description
<a href="#">Executive Summary</a>	Synopsis of the results from the Conversion Funnel and Retention Rates. A description of the Analysis on them as well as on the Data used.
<a href="#">Conversion Funnel</a>	Pivot Table representing the path experienced by our customers as they interact with the company. With the total conversion rates and conversion rates to the next step of views, shopping cart and purchase events
<a href="#">Retention Rates</a>	Table with the percentage of customers who remained active compared to the initial number of users who purchased on the cohorts between 2020-09 to 2021-02 divided by months.
<a href="#">Cohort Analysis</a>	Our popularity with the new customers by month. Represented with purchases from unique customers for the last six months between 2020-09 and 2021-02.
<a href="#">Purchase Activity</a>	Clean and filtered data from purchase events with the addition of 4 columns (first_purchase_date, event_month, first_purchase_month and cohort_age) to make a cohort analysis and retention rates by month.
<a href="#">First Purchase</a>	Pivot Table representing the minimum purchase event date of our customers
<a href="#">Raw Data</a>	Data from an online electronics, accessories, and supply store. Where each event is divided by user_id, type of event (views, shopping cart and purchase, category of product, brand of the product, price and date of the event.

**Legend**

- Summaries
- Analysis
- Calculations
- Raw Data

# Executive Summary

Executive Summary				
Results	Synopsis			
Conversion Funnel	Since the Data from the raw_user_activity on event_date column is not timestamp format is not suitable for a conversion funnel as a sequence of events. We used a simple conversion funnel. Out of the 10,453 unique users who viewed the product, 3,036 (or 29%) proceeded to add it to their shopping cart. From there, 1,081 users (or 36% of those who added the product to their cart) completed the purchase			
Retention Rates	The retention rate initially was high but declined significantly in subsequent months, despite an increase in the total number of purchases. January 2021 showed signs of improvement in the retention rate. Implementing a loyalty program based on the opening month's success could help sustain and improve retention rates.			
Analysis	Description			
Raw Data	Data originates from an online electronics, accessories, and supply store. It contains 43,660 recorded events captured from webpage interactions, including views, shopping cart additions, and purchases. The data consists of six columns:  user_id: Unique identifier for each user. event_type: Specifies the type of activity performed by the user (views, shopping_cart, or purchase). category_code: The category of the product involved in the event. brand: The brand of the product being sold. price: The price of the product in USD. event_date: The date of the event, ranging from September 24, 2020, to February 28, 2021. For analysis, specific sheets were created using the raw data: Conversion Funnel sheet: Utilizes the columns event_type and user_id. First Purchase sheet: Uses the columns user_id and event_date to track the user's first purchase. Cohort Analysis sheet: Uses the columns user_id, event_type (filtered for purchases), and event_date to categorize and analyze user cohorts. Retention Rate sheet: Employs the same columns as the Cohort Analysis sheet (user_id, event_type filtered for purchases, and event_date) to calculate user retention rates over time.			
	The data used to calculate the conversion rates in our simple funnel consisted of event types in rows (views, shopping cart, and purchase), with the count of unique user IDs as values. 10453 for View, 3036 for shopping cart and 1081 for purchase. We added two columns:  Total Conversion Rates: This column represents the ratio of the count of unique user IDs in the shopping cart and purchase stages to the count of unique user IDs			

# Analysis

docs.google.com/spreadsheets/d/1\_MnH4fISoSIR9M8HlpWC5gCmg8nW\_FIXkcyieqBtME/edit?gid=868644233#gid=868644233

Carlos Andres Gallegos Sprint 3 Project 04-23-24

File Edit View Insert Format Data Tools Extensions Help

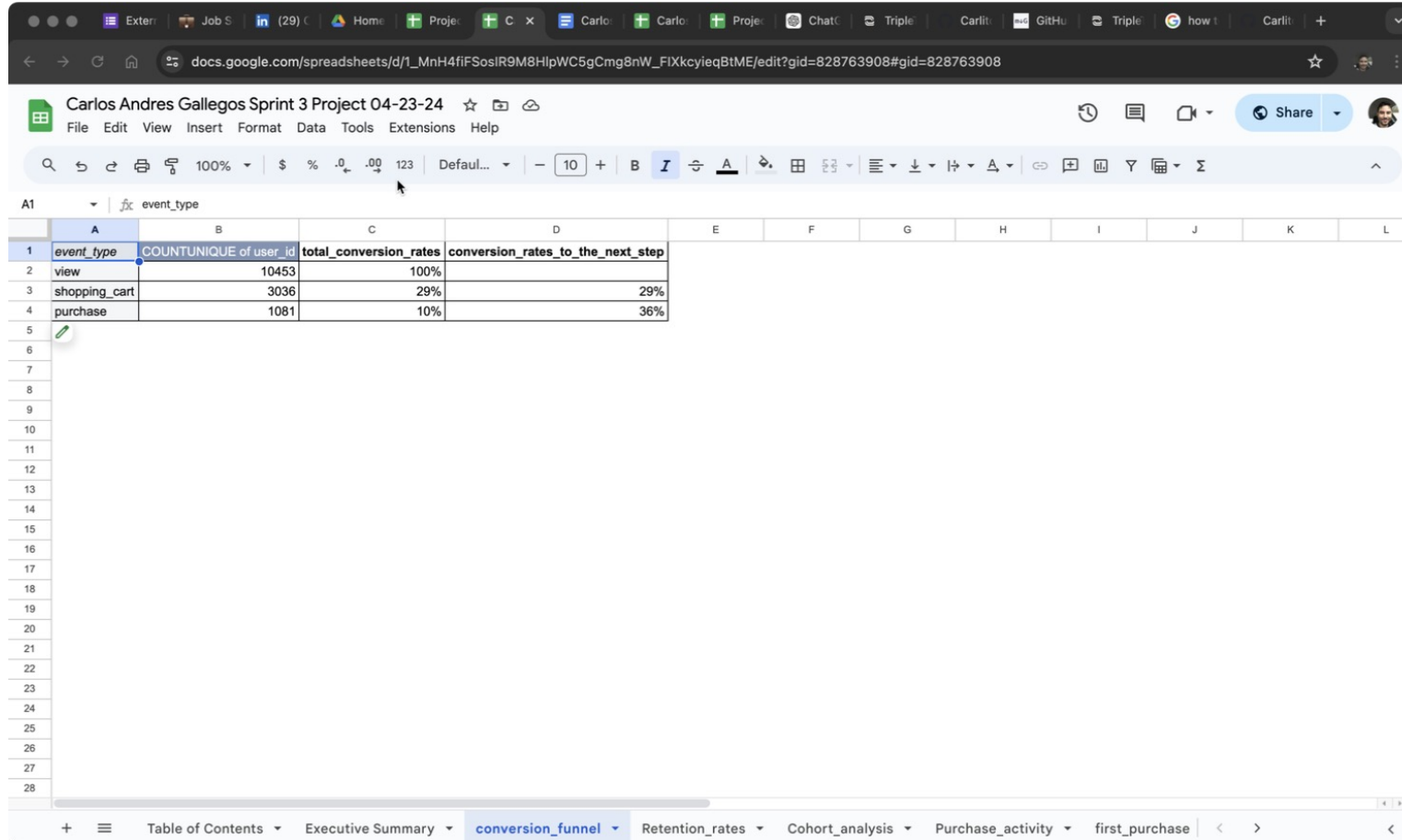
70% 123 Arial 10 B I U

B2

	A	B	C	D	E	F	G	H	I	J	K
9	<b>Analysis</b>	<b>Description</b> Data originates from an online electronics, accessories, and supply store. It contains 43,660 recorded events captured from webpage interactions, including views, shopping cart additions, and purchases. The data consists of six columns:  <b>user_id:</b> Unique identifier for each user. <b>event_type:</b> Specifies the type of activity performed by the user (views, shopping_cart, or purchase). <b>category_code:</b> The category of the product involved in the event. <b>brand:</b> The brand of the product being sold. <b>price:</b> The price of the product in USD. <b>event_date:</b> The date of the event, ranging from September 24, 2020, to February 28, 2021. For analysis, specific sheets were created using the raw data: <b>Conversion Funnel sheet:</b> Utilizes the columns event_type and user_id. <b>First Purchase sheet:</b> Uses the columns user_id and event_date to track the user's first purchase. <b>Cohort Analysis sheet:</b> Uses the columns user_id, event_type (filtered for purchases), and event_date to categorize and analyze user cohorts. <b>Retention Rate sheet:</b> Employs the same columns as the Cohort Analysis sheet (user_id, event_type filtered for purchases, and event_date) to calculate user retention rates over time.									
10	<b>Raw Data</b>	The data used to calculate the conversion rates in our simple funnel consisted of event types in rows (views, shopping_cart, and purchase), with the count of unique user IDs as values. 10453 for View, 3036 for shopping cart and 1081 for purchase. We added two columns:  <b>Total Conversion Rates:</b> This column represents the ratio of the count of unique user IDs in the shopping cart and purchase stages to the count of unique user IDs in the views stage in percentage format. 29% for shopping cart and 10% for purchase. <b>Conversion Rates to the Next Stage:</b> This column calculates the conversion rate from one stage to the next. It includes the ratio of the count of unique user IDs in the shopping cart stage to the count of unique user IDs in the views stage, and the ratio of the count of unique user IDs in the purchase stage to the count of unique user IDs in the shopping cart stage in percentage format. 29% for shopping cart and 36% for purchase.									
11	<b>Conversion Funnel</b>	We conducted an acquisition cohort analysis based on the month of a user's first purchase. By filtering event types, we analyzed 4,845 purchases in the "Purchase_activity sheet." We created four additional columns to aid the analysis:  <b>First_purchase_date:</b> A pivot table named "First purchase" was created to find the minimum date of purchase for each user ID using the VLOOKUP function. <b>Event_month:</b> The TEXT function was used to extract the month from each event_date. <b>First_purchase_month:</b> The TEXT function was applied to each first_purchase_date to derive the first purchase month. <b>Cohort_age:</b> The DATEDIF function was used to calculate the number of months between first_purchase_month and event_month. We created a pivot table in the "Cohort analysis sheet" to count unique user IDs for each first_purchase_month in the cohort_age. This allowed us to generate a "Retention_rate sheet" to analyze the retention rate of customers across the subsequent four months after their first purchase. <b>Initial Retention Rate:</b> The highest retention rate was observed in September 2020, with a rate of 12.5%. This covered only the initial six days of data. <b>Decline in Retention:</b> In the second month, the retention rate fell to 6.25%. In the fourth month, it dropped further to 3.13%. No data was available for the third month. Improvement in January 2021: In January 2021, the retention rate improved slightly to 6.87%, suggesting a positive trend. <b>Purchase Trends:</b> Despite the declining retention rates, the total number of purchases increased over the months: 187 purchases in October 2020. 238 purchases in November 2020. 203 purchases in December 2020. 233 purchases in January 2021. 188 purchases in February 2021. No retention rate data is available for February.									
12	<b>Retention Rates</b>	<b>Best Months for Returning Customers:</b> The opening month of September 2020 saw the highest retention rate, although it was based on a short period (6 days)									
13											

Table of Contents Executive Summary conversion\_funnel Retention\_rates Cohort\_analysis Purchase\_activity first\_purchase

# conversion\_funnel



Carlos Andres Gallegos Sprint 3 Project 04-23-24

File Edit View Insert Format Data Tools Extensions Help

100% 123 Default... 10 B I A

A1 event\_type

	A	B	C	D	E	F	G	H	I	J	K	L
1	event_type	COUNTUNIQUE of user_id	total_conversion_rates	conversion_rates_to_the_next_step								
2	view	10453	100%									
3	shopping_cart	3036	29%									
4	purchase	1081	10%	36%								
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												
24												
25												
26												
27												
28												

Table of Contents Executive Summary conversion\_funnel Retention\_rates Cohort\_analysis Purchase\_activity first\_purchase

# Retention\_rates

Carlos Andres Gallegos Sprint 3 Project 04-23-24

File Edit View Insert Format Data Tools Extensions Help

100% 123 Default... 10 B I A

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	COUNTUNIQUE of user_id	cohort_age													
2	first_purchase_month	0	1	2	3	4	Grand Total								
3	2020-09	100.00%	12.50%	6.25%		3.13%	100.00%								
4	2020-10	100.00%	7.49%	3.74%	0.53%	0.53%	100.00%								
5	2020-11	100.00%	5.46%	2.94%	0.42%		100.00%								
6	2020-12	100.00%	4.43%	2.96%			100.00%								
7	2021-01	100.00%	6.87%				100.00%								
8	2021-02	100.00%					100.00%								
9	Grand Total	100.00%	5.18%	2.04%	0.19%	0.19%	100.00%								
11			1	2	3	4									
12	2020-09		12.5	6.3	0.0	3.1									
13	2020-10		7.5	3.7	0.5	0.5									
14	2020-11		5.5	2.9	0.4	0.0									
15	2020-12		4.4	3.0	0.0	0.0									
16	2021-01		6.9	0.0	0.0	0.0									
17	2021-02														

Table of Contents Executive Summary conversion\_funnel Retention\_rates Cohort\_analysis Purchase\_activity first\_purchase

# Cohort\_analysis

Carlos Andres Gallegos Sprint 3 Project 04-23-24

File Edit View Insert Format Data Tools Extensions Help

100% 123 Default... 10 B I A

A1 COUNTUNIQUE of user\_id

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	COUNTUNIQUE of user_id	cohort_age											
2	first_purchase_month	0	1	2	3	4	Grand Total						
3	2020-09	32	4	2		1	32						
4	2020-10	187	14	7	1	1	187						
5	2020-11	238	13	7	1		238						
6	2020-12	203	9	6			203						
7	2021-01	233	16				233						
8	2021-02	188					188						
9	Grand Total	1081	56	22	2	2	1081						
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													

+ Table of Contents Executive Summary conversion\_funnel Retention\_rates Cohort\_analysis Purchase\_activity first\_purchase



# Purchase\_activity

docs.google.com/spreadsheets/d/1\_MnH4fiFSosIR9M8HipWC5gCmg8nW\_FIXkcyieqBtME/edit?gid=1367554732#gid=1367554732

Carlos Andres Gallegos Sprint 3 Project 04-23-24

File Edit View Insert Format Data Tools Extensions Help

100% 123 Default... 10 B I A

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	user_id	event_type	category_code	brand	price	event_date	first_purchase_date	event_month	first_purchase_month	cohort_age			
2	a1515915625519397646	purchase	computers.peripherals.camera	logitech	123.35	2020-09-24	2020-09-24	2020-09	2020-09	0			
3	a1515915625519397646	purchase	computers.peripherals.camera	logitech	123.35	2020-09-24	2020-09-24	2020-09	2020-09	0			
4	a1515915625519397646	purchase	computers.peripherals.camera	logitech	123.35	2020-09-24	2020-09-24	2020-09	2020-09	0			
5	a1515915625358482476	purchase	computers.components.tv_tuner	perfeo	19.49	2020-09-24	2020-09-24	2020-09	2020-09	0			
6	a1515915625519463641	purchase	stationery.cartridge	hp	71.02	2020-09-24	2020-09-24	2020-09	2020-09	0			
7	a1515915625519463641	purchase	stationery.cartridge	hp	71.02	2020-09-24	2020-09-24	2020-09	2020-09	0			
8	a1515915625519463641	purchase	stationery.cartridge	hp	71.02	2020-09-24	2020-09-24	2020-09	2020-09	0			
9	a1515915625503751892	purchase	appliances.environment.vacuum	topon	35.87	2020-09-24	2020-09-24	2020-09	2020-09	0			
10	a1515915625519463641	purchase	stationery.cartridge	hp	71.02	2020-09-24	2020-09-24	2020-09	2020-09	0			
11	a1515915625385888878	purchase	stationery.cartridge	cactus	24.73	2020-09-25	2020-09-25	2020-09	2020-09	0			
12	a1515915625385888878	purchase	stationery.cartridge	cactus	24.73	2020-09-25	2020-09-25	2020-09	2020-09	0			
13	a1515915625519762021	purchase		kitfort	140.63	2020-09-25	2020-09-25	2020-09	2020-09	0			
14	a1515915625519762021	purchase		kitfort	140.63	2020-09-25	2020-09-25	2020-09	2020-09	0			
15	a1515915625520029989	purchase	stationery.cartridge	hp	29.57	2020-09-26	2020-09-26	2020-09	2020-09	0			
16	a1515915625520049345	purchase	auto.accessories.player	kenwood	104.86	2020-09-26	2020-09-26	2020-09	2020-09	0			
17	a1515915625520070184	purchase	computers.peripherals.printer	epson	647.29	2020-09-26	2020-09-26	2020-09	2020-09	0			
18	a1515915625520070184	purchase	computers.peripherals.printer	epson	647.29	2020-09-26	2020-09-26	2020-09	2020-09	0			
19	a1515915625520070184	purchase	computers.peripherals.printer	epson	647.29	2020-09-26	2020-09-26	2020-09	2020-09	0			
20	a1515915625520227072	purchase	computers.peripherals.wifi	ubiquiti	155.43	2020-09-27	2020-09-27	2020-09	2020-09	0			
21	a1515915625520227072	purchase	computers.peripherals.wifi	ubiquiti	130.73	2020-09-27	2020-09-27	2020-09	2020-09	0			
22	a1515915625520227072	purchase	computers.peripherals.wifi	ubiquiti	155.43	2020-09-27	2020-09-27	2020-09	2020-09	0			
23	a1515915625371158307	purchase	stationery.cartridge	cactus	20.14	2020-09-28	2020-09-28	2020-09	2020-09	0			
24	a1515915625371158307	purchase	stationery.cartridge	cactus	20.14	2020-09-28	2020-09-28	2020-09	2020-09	0			
25	a1515915625371158307	purchase	stationery.cartridge	cactus	20.14	2020-09-28	2020-09-28	2020-09	2020-09	0			
26	a1515915625371158307	purchase	computers.peripherals.printer	canon	192.86	2020-09-28	2020-09-28	2020-09	2020-09	0			
27	a1515915625371158307	purchase	stationery.cartridge	cactus	20.14	2020-09-28	2020-09-28	2020-09	2020-09	0			
28	a1515915625520413119	purchase	stationery.cartridge	nv-print	17.62	2020-09-28	2020-09-28	2020-09	2020-09	0			

+ Table of Contents Executive Summary conversion\_funnel Retention\_rates Cohort\_analysis Purchase\_activity first\_purchase



# first\_purchase

The screenshot shows a Google Sheet with the following data:

	A	B
1	user_id	_MIN of event_date
2	a1515915625355179497	2020-11-04
3	a1515915625355805313	2020-09-29
4	a1515915625358482476	2020-09-24
5	a1515915625359833886	2020-11-20
6	a1515915625368479327	2020-11-01
7	a1515915625371158307	2020-09-28
8	a1515915625385482819	2020-10-01
9	a1515915625385888878	2020-09-25
10	a1515915625393382317	2020-11-03
11	a1515915625395096356	2020-10-17
12	a1515915625403460003	2020-10-11
13	a1515915625413257995	2020-10-09
14	a1515915625415474736	2020-10-03
15	a1515915625466195424	2020-10-10
16	a1515915625469806973	2020-10-04
17	a1515915625476168209	2020-10-10
18	a1515915625483402054	2020-11-09
19	a1515915625486521695	2020-10-05
20	a1515915625503751892	2020-09-24
21	a1515915625505008297	2020-09-29
22	a1515915625507596879	2020-11-08
23	a1515915625507754425	2020-10-13
24	a1515915625509677596	2020-12-20
25	a1515915625509808717	2020-10-03
26	a1515915625510154735	2020-10-21
27	a1515915625510668770	2020-11-17
28	j15915625511110840	2020-10-08

# Copy\_of\_raw\_user\_activity

Google Docs interface showing a spreadsheet titled "Carlos Andres Gallegos Sprint 3 Project 04-23-24". The spreadsheet is named "Copy\_of\_raw\_user\_activity" and contains user activity data.

	A	B	C	D	E	F	G
1	user_id	event_type	category_code	brand	price	event_date	
2	a1515915625519014356	view	computers.peripherals.printer	pantum	113.81	2020-09-24	
3	a1515915625519392599	view	stationery.cartridge	lomond	8.35	2020-09-24	
4	a1515915625369253995	view	construction.tools.welding		250.63	2020-09-24	
5	a1515915625369253995	view	construction.tools.welding		250.63	2020-09-24	
6	a1515915625519410002	view	computers.desktop	steelseries	58.49	2020-09-24	
7	a1515915625519413944	view	appliances.kitchen.microwave	samsung	130.11	2020-09-24	
8	a1515915625519397646	view	computers.peripherals.camera	logitech	123.35	2020-09-24	
9	a1515915625519397646	shopping_cart	computers.peripherals.camera	logitech	123.35	2020-09-24	
10	a1515915625519397646	purchase	computers.peripherals.camera	logitech	123.35	2020-09-24	
11	a1515915625519397646	view	computers.peripherals.camera	logitech	123.35	2020-09-24	
12	a1515915625519417750	view	computers.network.router	d-link	53.14	2020-09-24	
13	a1515915625519397646	shopping_cart	computers.peripherals.camera	logitech	123.35	2020-09-24	
14	a1515915625519397646	purchase	computers.peripherals.camera	logitech	123.35	2020-09-24	
15	a1515915625519397646	purchase	computers.peripherals.camera	logitech	123.35	2020-09-24	
16	a1515915625358482476	view	computers.components.tv_tuner	perfeo	19.49	2020-09-24	
17	a1515915625358482476	shopping_cart	computers.components.tv_tuner	perfeo	19.49	2020-09-24	
18	a1515915625358482476	purchase	computers.components.tv_tuner	perfeo	19.49	2020-09-24	
19	a1515915625519443407	view	computers.components.motherboard	asrock	77.73	2020-09-24	
20	a1515915625519443407	shopping_cart	computers.components.motherboard	asrock	77.73	2020-09-24	
21	a1515915625519443407	view	computers.components.motherboard	asrock	77.73	2020-09-24	
22	a1515915625519446240	view		formula	35.54	2020-09-24	
23	a1515915625519452732	view	computers.notebook	hp	27.62	2020-09-24	
24	a1515915625519460662	view	computers.peripherals.joystick	logitech	162.51	2020-09-24	
25	a1515915625519460662	shopping_cart	computers.peripherals.joystick	logitech	162.51	2020-09-24	
26	a1515915625519460662	view	computers.peripherals.joystick	logitech	162.51	2020-09-24	
27	a1515915625519460662	view	computers.peripherals.joystick	logitech	193.98	2020-09-24	
28	a1515915625519463641	view	stationery.cartridge	ho	71.02	2020-09-24	

# conversion\_funnel\_sequence\_of\_events

Google Sheet titled "Carlos Andres Gallegos Sprint 3 Project 04-23-24". The sheet displays a table with columns A through N. The first row (row 1) contains summary statistics: "MIN of event\_date", "event\_type", "10453", "304", "18", and "user\_count". The second row (row 2) contains headers for the data rows: "user\_id", "purchase", "shopping\_cart", "view", "view", "shopping\_cart", "purchase", and "user\_count". The subsequent rows (rows 3-28) contain data for individual users, including their user ID, event type, and date.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	MIN of event_date	event_type			10453	304	18	user_count						
2	user_id	purchase	shopping_cart	view	view	shopping_cart	purchase							
3	a1515915625354412996			2020-10-02		FALSE	FALSE							
4	a1515915625355179497	2020-11-04	2020-11-04	2020-09-26		TRUE	FALSE							
5	a1515915625355805313	2020-09-29	2020-09-29	2020-09-24		TRUE	FALSE							
6	a1515915625356687705		2020-10-05	2020-10-05		FALSE	FALSE							
7	a1515915625357946675			2020-09-28		FALSE	FALSE							
8	a1515915625358008174			2020-09-28		FALSE	FALSE							
9	a1515915625358482476	2020-09-24	2020-09-24	2020-09-24		FALSE	FALSE							
10	a1515915625359264087			2020-10-07		FALSE	FALSE							
11	a1515915625359268086			2020-09-30		FALSE	FALSE							
12	a1515915625359833886	2020-11-20	2020-11-20	2020-11-20		FALSE	FALSE							
13	a1515915625360350810			2020-10-04		FALSE	FALSE							
14	a1515915625360479862		2020-10-24	2020-10-04		TRUE	FALSE							
15	a1515915625362668216			2020-09-27		FALSE	FALSE							
16	a1515915625363906872			2020-09-27		FALSE	FALSE							
17	a1515915625364013669			2020-12-09		FALSE	FALSE							
18	a1515915625365126625			2020-09-26		FALSE	FALSE							
19	a1515915625367995831			2020-10-07		FALSE	FALSE							
20	a1515915625368479327	2020-11-01	2020-11-01	2020-11-01		FALSE	FALSE							
21	a1515915625368720412		2020-12-17	2020-11-08		TRUE	FALSE							
22	a1515915625369253995			2020-09-24		FALSE	FALSE							
23	a1515915625369734431			2020-10-06		FALSE	FALSE							
24	a1515915625371158307	2020-09-28	2020-09-28	2020-09-28		FALSE	FALSE							
25	a1515915625371372990			2020-10-03		FALSE	FALSE							
26	a1515915625375042249		2020-09-26	2020-09-26		FALSE	FALSE							
27	a1515915625375947733			2020-09-29		FALSE	FALSE							
28	a1515915625379735195		2020-10-12	2020-09-28		TRUE	FALSE							