✨ Overview

In Sprint #3 my goal was to analysing subscription data for streaming service going through three critical datasets — Cohort, A/B test and CLV. As a final output observations, conclusions on opportunities and recommendations for performace to the Product team for three separate tasks (Task 1: Cohort, Task 2: A/B Testing, Task 3: CLV) are delivered.

Steps were taken:

- **Data Cleaning**

All of the following were handled:

✅ No duplicates
✅ No missing values
✅ No negative values
✅ No date logic mismatches

- **KPI 1: Weekly Average Revenue Per User (ARPU) by Cohorts (USD)** (per week w_0 to w_6):

ARPU = Total Revenue / Total number of User in the Cohort. Focus -> revenue per User.
Purpose: To understand **how much revenue each user** contributes and evaluate monetization strategies, i.e. -> to get insight of how valuable each user is over time (per week, per month).
ARPU use for:

- Understand customer lifetime value (**total revenue a customer brings over their entire relationship with your company**)
- Measure the impact of upsells, pricing changes, or new features.
- Segment by user types (e.g., mobile vs. desktop ARPU).

cohort	W_0	W_1	W_2	W_3	W_4	W_5	W_6
2020-11-01	2.48	2.15	2.16	2.06	2.25	2.02	1.75
2020-11-08	2.9	2.82	2.65	2.29	2.2	2.13	
2020-11-15	3.43	3.1	2.97	2.47	2.5		
2020-11-22	3.87	3.72	3.24	3.17			
2020-11-29	4.93	4.86	4.12				
2020-12-06	9.31	6.79					
2020-12-13	18.13						

Key Observations:

1. **ARPU Increases with Newer Cohorts (Initial Weeks)**
   - Cohorts starting later (e.g., Nov 29) show a higher ARPU in W_0:
     - Nov 1 cohort: $2.48
     - Nov 29 cohort: $4.93
   - This may indicate improved monetization strategies, more engaged users, or better marketing effectiveness for later cohorts.
2. **ARPU Declines Over Time Within Each Cohort**
   - Almost every cohort shows a gradual weekly decline in ARPU:
     - Nov 1 cohort drops from $2.48 in W_0 to $1.75 in W_6
     - Nov 15 cohort drops from $3.43 in W_0 to $2.47 in W_3
   - This is typical in subscription model, where early engagement is high, followed by a drop-off.
3. **Missing Values in Later Weeks**
   - Most recent cohorts don't have data beyond W_2 or W_3:
     - Nov 29 cohort only has W_0 to W_2
   - Indicates either users haven't reached later weeks yet or churn may be affecting activity levels.
4. **Early Cohorts Show Stable but Lower Monetization**

- The Nov 1 cohort has the longest visible lifetime (W_0 to W_6) but maintains a lower ARPU overall, suggesting:
    - Less mature monetization or onboarding early in the product lifecycle
    - Possibly higher user volume with more variability in behavior

Business insight:

1. **Later cohorts are spending more early on** — this is a positive signal and worth investigating further:
    - What changed in product, pricing, or onboarding?
    - Can those changes be sustained or applied to earlier/lower-performing cohorts?
2. **Retention appears to be a challenge** — ARPU drop-off suggests declining engagement over time:
    - Follow-up analysis on **user retention per week** and **purchase frequency** would be valuable.
3. **Customer segmentation** could reveal deeper insights:
    - Break down ARPU by **category** (mobile vs desktop) or **country** to optimize targeting.
4. **CLV modeling** is the next logical step:
    - Combining ARPU trends with retention rates will help forecast **Customer Lifetime Value**, critical for budgeting, marketing ROI, and growth strategies.

- **KPI 2: Weekly Average Order Value (AOV) by Cohorts (USD)** (per week w_0 to w_6):

AOV = Total Revenue / Number of Total Orders. Focus -> revenue per order.

Purpose: To understand the **spending intensity per order**, i.e. to get insight of how much people spend per order. AOV use for:

- Optimize **checkout flow** or **product bundling**.
- Assess effectiveness of **promotions** (e.g., free shipping threshold).
- Track behavioral shifts (e.g., larger carts on weekends).

cohort	W_0	W_1	W_2	W_3	W_4	W_5	W_6
2020-11-01	15.57	13.97	14.3	13.95	16.17	15.43	14.84
2020-11-08	15.8	15.71	15.32	13.91	14.27	14.63	
2020-11-15	15.42	14.72	14.74	13.03	14.23		
2020-11-22	13.73	14.04	13.36	14.99			
2020-11-29	13.88	14.47	13.34				
2020-12-06	17.6	14.42					
2020-12-13	18.13						

Key Observations

- **AOV is stable** within each cohort, suggesting consistent pricing or product offerings. Average revenue per order 413.97 / 28 = 14.79
- Some **variation appears across cohorts**:
    - Earlier cohorts (e.g., Nov 1, Nov 8) have **higher AOV** (~$14–$16) compared to later cohorts (e.g., Nov 22, Nov 29).
- **Missing values** in later weeks are expected due to recency — users haven't yet reached those weeks.

- **KPI 3: Revenue per Order** by comparing ARPU to AOV

- **ARPU (Average Revenue Per User)** = Total Revenue / Total Users in Cohort
- **AOV (Average Order Value)** = Total Revenue / Total Orders

So:

ARPU = Average Revenue per Order × Average # of Orders Per User

This means the difference between ARPU and AOV reveals insights into **user purchasing frequency**:

- If ARPU ≈ AOV → users placed about 1 order per week
- If ARPU < AOV → some users didn't make purchases that week
- Larger gaps = more drop-off or inactive users

Average # of orders per user = Total orders / Total number of users

Average # of Orders per User per Week for each cohort:

cohort	W_0	W_1	W_2	W_3	W_4	W_5	W_6
2020-11-01	0.16	0.15	0.15	0.15	0.14	0.13	0.12

2020-11-08	0.18	0.18	0.17	0.16	0.15	0.15	
2020-11-15	0.22	0.21	0.2	0.19	0.18		
2020-11-22	0.28	0.26	0.24	0.21			
2020-11-29	0.36	0.34	0.31				
2020-12-06	0.53	0.47					
2020-12-13	1						

Key Insights:

1. **Order Frequency Improves with Newer Cohorts**:
   - Nov 1 cohort: **~0.16 orders/user in W_0**
   - Nov 29 cohort: **~0.36 orders/user in W_0**
   - Suggests newer users are making **more frequent purchases early**.
2. **Drop-off Over Time**:
   - Most cohorts show a **decline in weekly purchase frequency**, supporting what we saw in ARPU.
   - Since AOV is stable, but ARPU varies, the drop in ARPU is likely due to fewer users buying, not smaller orders – meaning retention and re-activation matter more, than changing offer size.
3. **Focus on increasing the number of buyers per week, rather than the value of purchases per user!**

- **KPI 4: (Predictive) Customer Timelife Value (CLV)**. CLV = Sum of ARPU across all weeks

6-week Customer Lifetime Value (CLV) per cohort, calculated as the sum of ARPU from **Week 0 to Week 6**.

cohort	CLV_6_Weeks
2020-11-01	14.87
2020-11-08	14.99
2020-11-15	14.47
2020-11-22	13.99
2020-11-29	13.91
2020-12-06	16.11
2020-12-13	18.13

Observations:

- **CLV is relatively stable across cohorts** (~$13.91 to $14.99), despite variations in AOV and user engagement.
- The **highest CLV** was seen in the **Nov 8** cohort: **$14.99**
- This suggests that even small improvements in weekly ARPU across mulrtiple weeks meaningfully affect Lifetime Value.

Business insight:

- CLV highlights how much revenue we can expect from each user — and helps in budgeting, forecasting, and evaluating ROI.
- Cohorts with slightly higher CLV (like 2020-11-08) might be worth studying for behavioral patterns or promotional triggers that contributed to stronger long-term value.

Strategic Recommendations:

Area	**Opportunity**	**Suggested Actions**
Early Monetization	Capitalize on strong W_0–W_2 engagement	Improve onboarding offers, discounts, upsells
Retention	Users fade quickly after 2–3 weeks	Introduce email reminders, in-app missions, loyalty points
User Segmentation	All users treated the same in current analysis	Segment by platform (mobile/desktop), country, or product preference

Predictive Customer Lifetime Value (CLV) per cohort:
The total revenue a business can expect from a customer within a specific cohort during their lifetime as a customer.

- Predictive CLV (1 cohort), formula:
  - **Average Purchase Value**: $14.89
  - **Average Purchase Frequency**: 0.1428 times per week / 1.71 per quarter (12 weeks)
  - **Average Customer Lifespan**: 2 quarters (24 weeks)
- **Formula**: CLV = (Average Purchase Value) x (Average Purchase Frequency) x (Average Customer Lifespan)

1 cohort's CLV = ($14.89) x (1.71) x (2) = $50.92
2 cohort's CLV = ($14.94) x (1.98) x (2) = $59.16
3 cohort's CLV = ($14.42) x (2.4) x (2) = $69.22
4 cohort's CLV = ($14.03) x (2.97) x (2) = $83.34

Retention over multiple weeks is a *critical driver* of Customer Lifetime Value (CLV), and prioritizing it can yield exponential gains in long-term revenue and user engagement.

Impact Area	**Description**
CLV Growth	The longer a user stays active, the more chances they have to subscribe, repurchase, or engage with monetized features.
CAC Efficiency	Higher retention improves return on customer acquisition costs — every retained user lowers your CAC.
Predictability	Strong early-week retention correlates with more stable revenue and better forecasting.
Upsell Potential	Retained users are more likely to convert to premium, make repeat purchases, or refer others.

Link to conversation with ChatGPT
https://chatgpt.com/c/6843f292-bb6c-8011-b9ac-86760f3f431c