Uniswap V3 Traders Behaviour

1. Introduction 🏁

Main Dashboard:
https://public.tableau.com/app/profile/rodolfo/viz/Uniswapv3TraderBehavious/Dashboard?

publish=yes

GitHub Repo: https://github.com/LimaRods/Uniswap-User-Behaviour/tree/main

1.1 Protocol Introduction

Uniswap was launched in November 2018 and has since become one of the most popular decentralized exchanges in the cryptocurrency space. It is known for its simple and user-friendly interface, low fees, and ability to support a wide range of tokens. As of July 2023, Uniswap is the largest decentralized exchange by trading volume, with over **$1 billion** in daily trading volume on average, and has already processed over **$1.5 trillion** since its launch.

1.2 Objective

As a data consultant for **Uniswap**, I aim to deliver a comprehensive business growth report. This report will offer insights into Uniswap v3's customer and user demographics, focusing on enhancing profitability, revenue, retention, and usage metrics. The project will involve:

1. Crafting a dataset for analytical purposes.
2. Detailing the dataset creation methodology.
3. Examining pivotal customer metrics across various cohorts over time.
4. Proposing actionable strategies to augment key business performance indicators.

2. Methodology 🔬

It's important to mention that the scope of the project is restricted to v3 pools in Uniswap once is easier to find reliable data sources for structured data that provide the fee percent for hundreds of different pools.

2.1 Data Sources & Tools 🛠️

Flipside is the data provider of this project. We retrieved data from different tables such as `ethereum. core.ez_dex_swap`, `ethereum.uniswapv3.ez_pools` , `ethereum.core.fact_transactions`. We used the new Flipside feature called LiveQuery, then we were able to query external APIs to combine off-chain sources with Flipside data. We utilized the series `FEDFUNDS` from the **FRED AP**

The main language used in this project was **SQL**, following **Python**, which helped us build data pipelines and run some advanced algorithms. We also decided to go with **Tableau** in order to provide comprehensive and clear cohort charts. The dataset that feeds the dashboard is stored in **Google Sheets** and can be updated by running the Python Notebook `main.ipynb`

All codes are in the GitHub Repo

2.2 Data Transformation & Decisions 🧐

3 Datasets were crafted to run the analysis.

- **univ3_time_series.csv**: This dataset presents the metrics in monthly time-series
- **univ3_cohort_analysis.csv**: The data are prepared to perform cohort analysis
- **univ3_metrics_by_user.csv**: Here the metrics are aggregated by users. They're ordered by the top 1M traders in Uniswap v3 in terms of trading volume, due to Flipside's limitation for free resources. This dataset is dedicated to running the machine learning algorithm for clustering.

You can find these tables here.

2.2.1 Explanation

The following explanation covers the step by steps for **univ3_time_series.csv** and **univ3_cohort_analysis.csv**, the logic behind them is almost the same for **univ3_metrics_by_user.csv**, you can find the SQL queries here.

We use the following scripts to have the final datasets:

- `TIME_SERIES` table:

```
WITH new_users AS  (
  SELECT
    cohort_month,
    COUNT(ORIGIN_FROM_ADDRESS) AS new_traders
  FROM
    (SELECT
      MIN(DATE_TRUNC('month',block_timestamp)) AS cohort_month,
      ORIGIN_FROM_ADDRESS
    FROM
      ethereum.core.ez_dex_swaps
    WHERE
      platform = 'uniswap-v3'
    GROUP BY
      ORIGIN_FROM_ADDRESS
    )
  GROUP BY
    cohort_month
),

gas_price_per_tx AS (
  SELECT
    DATE_TRUNC('month', block_timestamp) as month,
    AVG(gas_price) avg_gas_price
```

```
  FROM
    (
      SELECT
            swap.block_timestamp,
            swap.tx_hash,
            origin_from_address,
            tx.gas_price,
            ROW_NUMBER() OVER(PARTITION BY swap.tx_hash order by swap.block_timestamp) as row_id

          FROM
            ethereum.core.ez_dex_swaps swap
            LEFT JOIN ethereum.core.fact_transactions tx
              ON (swap.tx_hash = tx.tx_hash AND swap.block_timestamp = tx.block_timestamp)

        WHERE
              platform = 'uniswap-v3'
        ORDER BY
          swap.block_timestamp
    )
  WHERE
    row_id = 1
  GROUP BY
     month
),

mau_tab AS (
  SELECT
    DATE_TRUNC('month',swap.block_timestamp) as month,
    COUNT(DISTINCT swap.ORIGIN_FROM_ADDRESS) AS total_traders,
    SUM(AMOUNT_IN_USD) as vol_usd,
    SUM(fee_percent * amount_in_usd) AS fee_usd,
    COUNT(*) as swap_count

  FROM
    ethereum.core.ez_dex_swaps swap
    INNER JOIN ethereum.uniswapv3.ez_pools pool ON swap.contract_address = pool.pool_address
  WHERE
    platform = 'uniswap-v3'

  GROUP BY
    month
),

dau_tab AS (
  SELECT
    DATE_TRUNC('month',day) AS month,
    AVG(total_traders) as dau
  FROM
    (
      SELECT
        DATE(block_timestamp) as day,
        COUNT(DISTINCT ORIGIN_FROM_ADDRESS) AS total_traders
      FROM
        ethereum.core.ez_dex_swaps
      WHERE
        platform = 'uniswap-v3'
      GROUP BY
        day
```

```
    )
  GROUP BY
    month
),
-- Federal Funds Rate Series - ⚡LiveQuery
interest_rate AS (
  WITH fred_data as (
  SELECT
    fred.get_series({
      'series_id': 'FEDFUNDS',
      'file_type': 'json',
      'observation_start': '2021-05-01',
      'frequency': 'm'
    }) as result
  )

  SELECT
    array.value:date::STRING AS month,
    array.value:value::FLOAT AS interest_rate_percent
  FROM
    fred_data,
    LATERAL FLATTEN(input => TO_VARIANT(result:data:observations)) as array
)

SELECT
  DATE(total.month) as month,
  total_traders,
  new_traders,
  ROUND((new_traders/total_traders),3) as new_traders_percent,
  1 - ROUND((new_traders/total_traders),3) as old_traders_percent,
  ROUND(dau/total_traders,3) AS stickiness, -- Stickiness Ratio
  swap_count,
  vol_usd,
  (vol_usd/swap_count) AS avg_tx_value,
  fee_usd as revenue_usd,
  gas.avg_gas_price,
  interest_rate_percent
FROM
  mau_tab total
  LEFT JOIN new_users new ON total.month = new.cohort_month
  LEFT JOIN dau_tab dau ON total.month = dau.month
  INNER JOIN gas_price_per_tx gas on total.month = gas.month
  INNER JOIN interest_rate rate ON total.month = rate.month
ORDER BY
  total.month ASC
```

The steps for this queries can be divided by the number of CTEs in the query:

1. CTE `new_users`
    - **Transformation**:
        - From the raw data of `ethereum.core.ez_dex_swaps`, this CTE calculates the month of the first interaction (cohort month) for each trader (`ORIGIN_FROM_ADDRESS`).

    - It then aggregates the number of new traders for each cohort month.
  - **Purpose**:
    - To identify when each user first interacted with the Uniswap v3 platform and determine the number of new traders in each month.

2. CTE `gas_price_per_tx`
  - **Transformation**:
    - From the `ethereum.core.ez_dex_swaps` table, this CTE finds the average gas price for transactions in each month.
    - It ensures uniqueness by considering only one row for each `tx_hash` (transaction hash) using the `ROW_NUMBER()` function.
  - **Purpose**:
    - To determine the average gas price users paid for transactions on a monthly basis. This helps in understanding the Ethereum network's congestion and its impact on user activity.

3. CTE `mau_tab`
  - **Transformation**:
    - From the `ethereum.core.ez_dex_swaps` table, it aggregates monthly data to calculate the total number of traders, total trading volume in USD, total fees in USD, and total swap count.
  - **Purpose**:
    - To calculate key monthly metrics like active users, volume, revenue, and activity.

4. CTE `dau_tab`
  - **Transformation**:
    - It aggregates daily data from `ethereum.core.ez_dex_swaps` to determine the daily active users (DAU).
    - It then averages the DAU values to get a monthly DAU metric.
  - **Purpose**:
    - To calculate the daily engagement of users on the platform and then average it over the month. This field will be used in the final query with `total_traders` (MAU) to calculate the retention metric stickiness ratio.

5. CTE `interest_rate`

    - **Transformation**:

        - It fetches the federal funds rate series from the FRED (Federal Reserve Economic Data) service.

        - The data is then flattened and transformed to obtain a monthly interest rate metric.

    - **Purpose**:

        - To incorporate external financial data (federal interest rates) which might influence user activity on the Uniswap v3 platform.

6. **Final Query:**

- **Transformation**:

    - It merges all the above CTEs based on the month and calculates metrics like percentage of new traders, percentage of old traders, stickiness ratio, average transaction value, and revenue in USD.

    - The data is then sorted by month in ascending order.

- **Purpose**:

    - To have a comprehensive monthly dataset which includes key metrics from both the Uniswap v3 platform and external financial data for holistic analysis.

- `COHORT_ANALYSIS`:

```
-- Setting the first cohort and date for each trader
WITH user_cohort AS (
    SELECT
      MIN(DATE_TRUNC('month',block_timestamp)) AS cohort_month,
      origin_from_address
    FROM
      ethereum.core.ez_dex_swaps
    WHERE
      platform = 'uniswap-v3'
    GROUP BY
      ORIGIN_FROM_ADDRESS
),
-- Retrieving the metrics for each user giving a cohort
following_month AS (
  SELECT
    DATEDIFF('month', uc.cohort_month, DATE_TRUNC('month', s.block_timestamp)) as cohort_ID,
    s.origin_from_address,
    SUM(AMOUNT_IN_USD) as vol_usd,
```

```
    SUM(fee_percent * amount_in_usd) AS fee_usd,
    COUNT(*) as swap_count

  FROM
    ethereum.core.ez_dex_swaps s
    LEFT JOIN user_cohort uc ON s.origin_from_address = uc.origin_from_address
    INNER JOIN ethereum.uniswapv3.ez_pools p ON s.contract_address = p.pool_address
  WHERE
      platform = 'uniswap-v3'
  GROUP BY
    s.origin_from_address,
    cohort_ID
),
-- Cohort Size
cohort_size AS (
  SELECT
    uc.cohort_month,
    COUNT(DISTINCT uc.origin_from_address) as total_users,
    SUM(AMOUNT_IN_USD) as vol_usd,
    SUM(fee_percent * amount_in_usd) AS fee_usd,
    COUNT(*) as swap_count
  FROM
    user_cohort uc
    LEFT JOIN ethereum.core.ez_dex_swaps s ON (s.origin_from_address = uc.origin_from_address
          AND DATE_TRUNC('month',block_timestamp) = uc.cohort_month)
    INNER JOIN ethereum.uniswapv3.ez_pools p ON s.contract_address = p.pool_address
  WHERE
      platform = 'uniswap-v3'
  GROUP BY
    uc.cohort_month
  ORDER BY
    uc.cohort_month
),

retention_table AS (
  SELECT
    c.cohort_month,
    f.cohort_ID,
    COUNT(*) as user_cohort,
    SUM(vol_usd) AS vol_usd_cohort,
    SUM(fee_usd) AS fee_usd_cohort,
    SUM(swap_count) AS swap_count_cohort
    --AVG(gas_price) AS avg_gas_price_cohort

  FROM
    following_month f
    LEFT JOIN user_cohort c ON f.origin_from_address = c.origin_from_address
  GROUP BY
    c.cohort_month,
    f.cohort_ID
)
-- Final view
SELECT
  DATE(r.cohort_month) as month ,
  'month_' || r.cohort_ID AS cohort_ID,
  s.total_users,
  r.user_cohort,
  s.vol_usd,
```

```
  vol_usd_cohort,
  s.fee_usd,
  fee_usd_cohort,
  s.swap_count
  swap_count_cohort,
  (s.vol_usd/s.swap_count) AS avg_tx_value,
  (vol_usd_cohort/swap_count_cohort) AS avg_tx_value_cohort

FROM
  retention_table r
  LEFT JOIN cohort_size s ON r.cohort_month = s.cohort_month
ORDER BY
  r.cohort_month,
  r.cohort_ID
```

1. CTE: `user_cohort`
   - **Transformation**:
     - From the raw data of `ethereum.core.ez_dex_swaps`, this CTE calculates the month of the first interaction (cohort month) for each trader (`origin_from_address`).
   - **Purpose**:
     - To identify when each user first interacted with the Uniswap v3 platform. This initial month will serve as the cohort for each user.
2. CTE: `following_month`
   - **Transformation**:
     - For each user and their associated cohort month, this CTE aggregates metrics such as trading volume in USD, fees in USD, and swap count for each subsequent month after their cohort month.
     - The `cohort_ID` is calculated as the difference in months between the user's activity month and their cohort month.
   - **Purpose**:
     - To track the behavior and activity of users in the months following their initial interaction with the platform.
3. CTE: `cohort_size`
   - **Transformation**:
     - For each cohort month, this CTE aggregates the total number of users, trading volume in USD, fees in USD, and swap count.

- **Purpose**:
  - To determine the size and activity of each monthly cohort during their first month of interaction.

CTE: `retention_table`

- **Transformation**:
  - This CTE aggregates user metrics for each cohort month and subsequent months (`cohort_ID`). The metrics include the number of users, trading volume in USD, fees in USD, and swap count.
- **Purpose**:
  - To build a retention table that shows how each cohort behaves in the subsequent months after their first interaction.

Final Query:

- **Transformation**:
  - Merges the `retention_table` and `cohort_size` based on the cohort month.
  - Calculates metrics like average transaction value for each cohort during their first month and in subsequent months, besides making it easy to calculate retention rate afterward.
- **Purpose**:
  - To have a comprehensive dataset that tracks the behavior and activity of each cohort over time.

2.2.2 Data Dictionary

You can find further explanations about the fields of the final tables below:

`month`**:** It's the first month that the users/traders interacted with the protocol

`total_users`: Total Number of active users

`swap_count`: total number of swaps

`vol_usd`**:** trading volume in USD generated

`fee_usd`: volume of fee paid by traders. It also represents the volume of revenue emitted to pay the liquidity provider of the Uniswap pool `avg_tx_value`**:** the average value of swaps in USD

`avg_gas_price`**:** average gas price of swap transactions

`interast_rate` `:` It's the effective federal funds rate targeted by the FED

`new_traders`: Number of new traders for that month.

`new_traders_percent`: Percentage of new traders.

`old_traders_percent`: Percentage of old traders.:

`stickiness`: Measure of user retention.

`Cohort_ID`: Number of months after the first month of interaction

`<column_name>_cohort`: They are the columns that contain the variables for users of each cohort along with months after the first month of interaction

3. Report & Data Analytics 💹

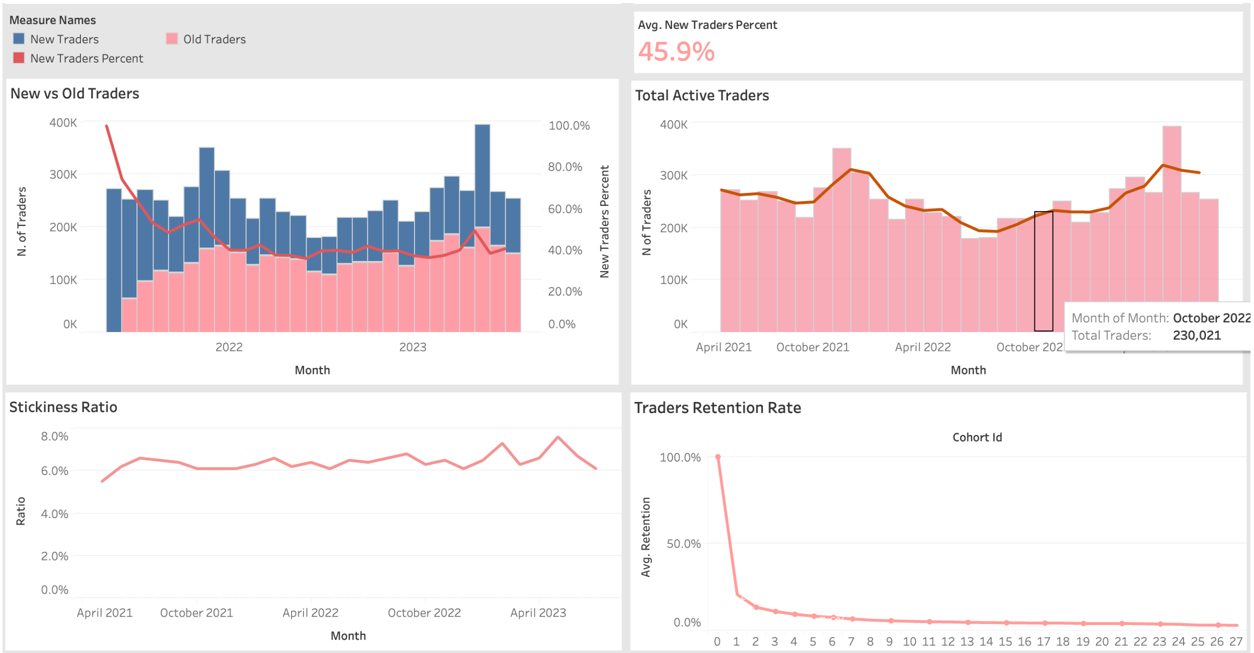
The main analytics are displayed in the **Tableau Dashboard.** Take a look at this link.

> https://public.tableau.com/views/Uniswapv3TraderBehavious/Dashboard?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link

Hence, We can divide the analyses in the 3 parts

3.1 Time Series Analysis

3.1.1 User Activity and Retention



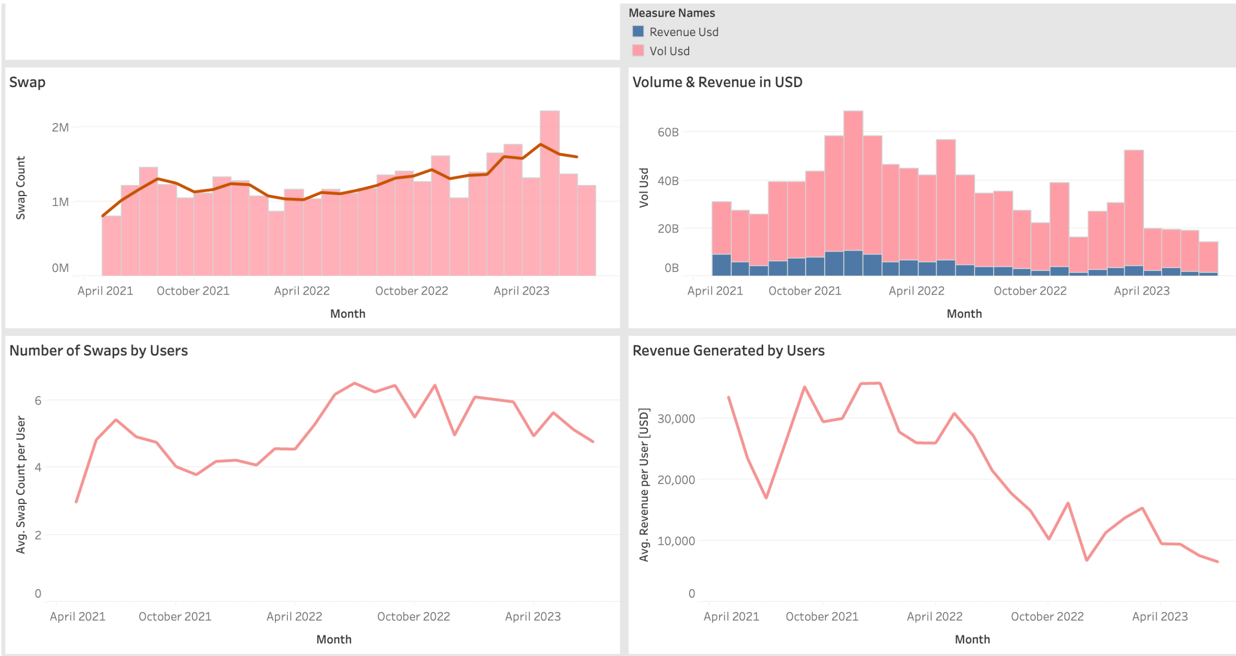
As we observe, there's a consistent number of active users engaging with the protocol over time. Furthermore, the steady influx of new traders suggests that approximately **46%** of all traders are newcomers, indicating that around **50%** of traders are returning to use the protocol. This could be interpreted as a favorable monthly retention rate.

This observation is further reinforced by examining the Stickiness Ratio (DAU/MAU). The Daily Active Users to Monthly Active Users ratio evaluates a product's "stickiness" by contrasting the number of unique daily users with the unique users who engage at least once in a given month. A high DAU/MAU ratio implies that a substantial portion of the product's users are interacting with it daily.

While a **low stickiness ratio** (DAU/MAU) might suggest infrequent use, it can coexist with a **high month-over-month (MoM)** retention rate. This scenario often arises in applications designed for sporadic use. In such cases, overall user retention can be more indicative than daily engagement. This pattern is fitting for DEXs, given that only a few traders engage in token swaps daily.

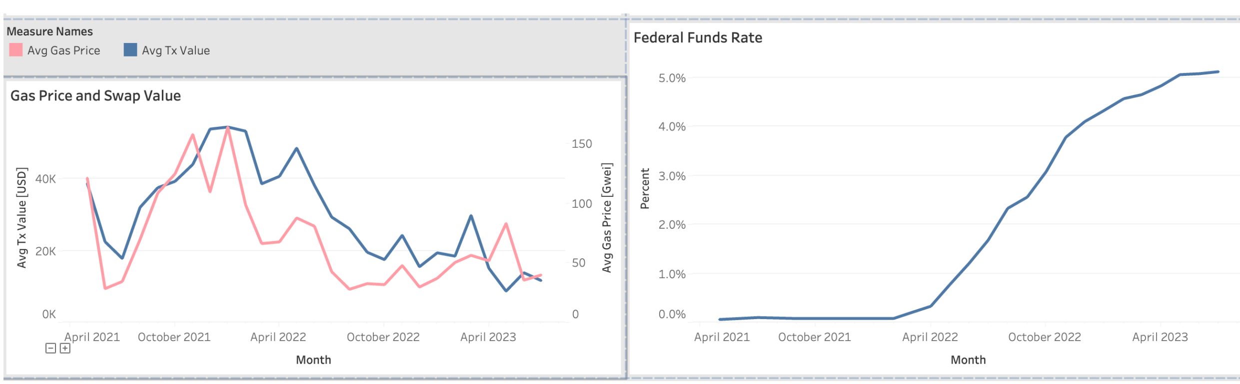
The Traders Retention Rate vs. Cohort ID chart shows that customers from specific cohorts tend not to return frequently in the long run. It's essential to note that the New vs. Old Traders metric classifies any recurring user as old (existing). Meanwhile, the retention by cohort ID represents the average retention of cohorts based on the number of months since their first interaction.

3.1.2 Trade Activity and Revenue



The **increased swaps per user** over time, especially from mid-2022 onwards, indicates that traders, on average, are conducting more trades. This could be due to various reasons such as increased market volatility, better platform features, or increased trader confidence.

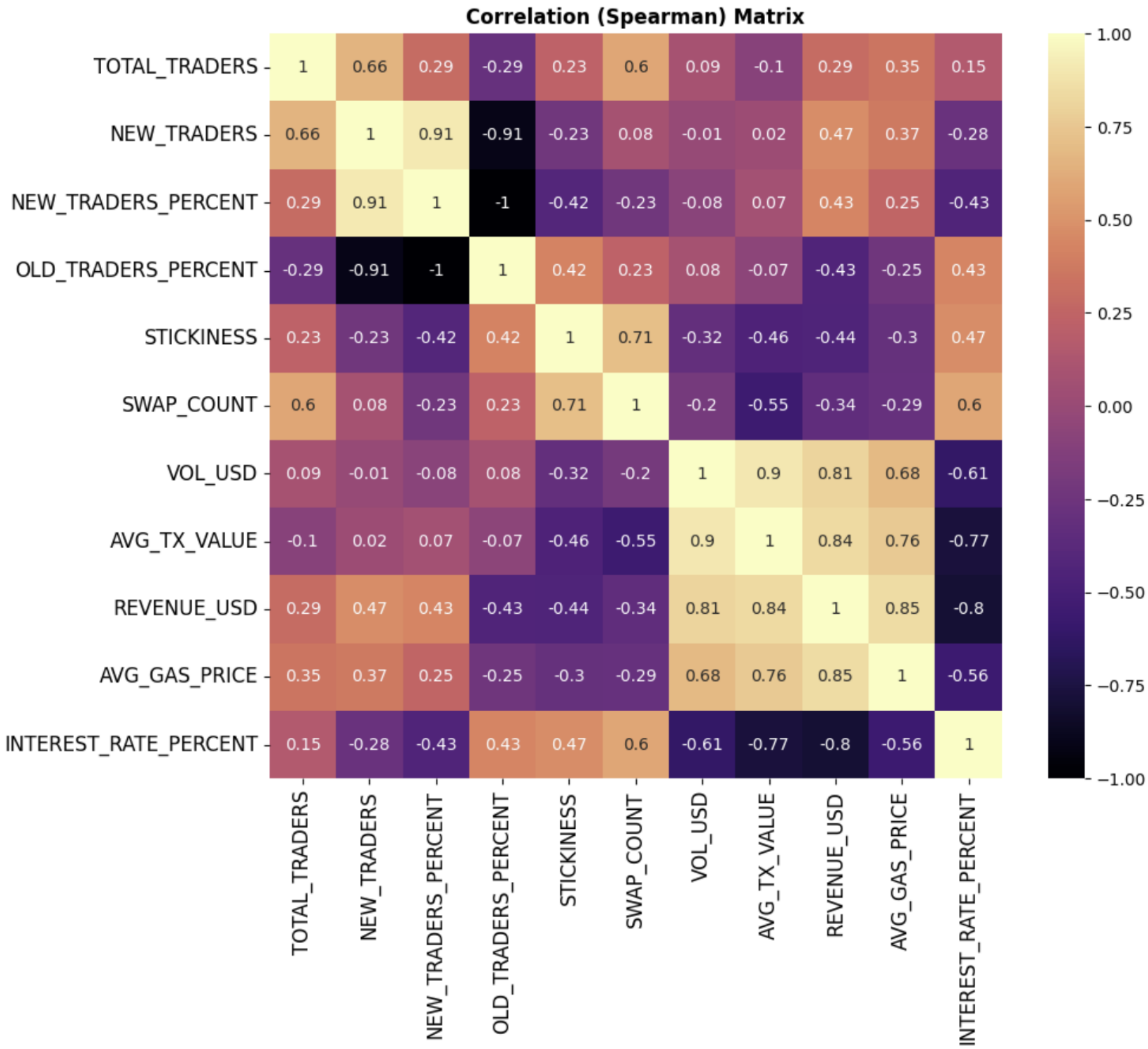
The **average volume per user** had its highs and lows, but the declining trend in the more recent months suggests that while traders might still be conducting swaps, the average value of these swaps has decreased. This could be indicative of traders being more cautious or the presence of more smaller retail traders in the platform as opposed to larger institutional traders.



The drop in average transaction value around June 2021 coincides with a significant drop in gas prices. This could indicate that when gas prices are lower, more users might engage in transactions, even if they're of a smaller value. When gas prices are high, only transactions that are of sufficiently high value (to justify the gas cost) might be carried out. On the other hand, when gas prices are lower, we might see a broader range of transaction values, including smaller ones.

In addition, an increase in the interest rate could signal a reduction in liquidity. Alternatively, other financial instruments might become more interesting than the crypto market under these conditions.

3.2 Correlation Analysis



A correlation matrix can give us valuable insights and clues for hypotheses previously raised

I want to receive insight of correlation between gas price/interest rate and other variables:

`avg_gas_price`:

1. **Revenue (REVENUE_USD)**: There's a strong positive correlation (**0.85**) between `avg_gas_price` and `REVENUE_USD` This suggests that as gas prices increase, the trading volume and fees in USD tends to increase. This may seem counterintuitive at first, as higher gas prices might deter small traders. However, it could indicate that when gas prices are high, only high-value transactions occur, contributing to a higher overall volume.

2. **New Traders**: The correlation is positive (**0.37**), suggesting that more new traders might join when gas prices are high. This is a bit surprising and could be due to other market factors.
3. **Stickiness Ratio (Retention)**: There's a negative correlation (**-0.30**) between gas price and stickiness. As gas prices rise, the retention seems to decrease slightly. This is expected as higher gas prices could deter users from making frequent transactions.
4. **Transaction Value (AVG_TX_VALUE)**: A strong positive correlation (**0.76**) is observed. As mentioned earlier, this could indicate that when gas prices are high, only significant transactions take place.

`Interest_rate`:

1. **Revenue (REVENUE_USD)**: The correlation is strongly negative (**-0.80**), indicating that as interest rates rise, the trading volume and fee paid (revenue collected) tend to decrease. This could be because higher interest rates might mean less liquidity in the market or other financial instruments becoming more attractive than crypto trading.
2. **New Traders**: There's a negative correlation (**-0.28**), indicating fewer new traders might join when interest rates are high.
3. **Stickiness (Retention)**: Positive correlation (**0.47**) suggests that users might become more sticky (i.e., continue using the platform) as interest rates rise. This could be because existing users have already invested in the ecosystem and continue trading regardless of external factors.
4. **Transaction Value (AVG_TX_VALUE)**: There's a strong negative correlation (**-0.77**) between interest rates and average transaction value. This could suggest that as interest rates rise, the average value of transactions decreases, possibly due to reduced liquidity or traders being more cautious.

3.3 Cohort Analysis

From the visualizations of cohort matrix in the Dashboard, we can observe the following trends:

1. **Retention by Cohort**

Retentation Rate - Cohort Matrix

Cohort Id

Month of Month	Total Users	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
May 2021	272,106	24.2%	17.5%	16.2%	14.3%	16.0%	15.6%	12.7%	9.9%	7.4%	7.8%	6.9%	6.8%	5.7%	4.9%	5.7%	5.3%	5.1%	5.7%	4
June 2021	187,319	26.9%	17.2%	13.2%	12.7%	12.5%	10.8%	8.4%	6.5%	7.2%	6.2%	6.2%	5.4%	4.7%	5.5%	5.1%	4.9%	5.5%	4.3%	4
July 2021	171,543	24.2%	13.0%	11.2%	10.6%	9.3%	7.3%	5.6%	6.3%	5.4%	5.4%	4.4%	4.2%	4.8%	4.4%	4.2%	4.7%	3.6%	4.0%	4
August 2021	133,181	20.9%	15.6%	14.3%	12.3%	10.1%	7.3%	7.5%	6.6%	5.9%	4.9%	4.3%	4.8%	4.3%	4.3%	4.8%	3.6%	4.2%	5.3%	5
September 2021	106,122	23.8%	19.0%	16.0%	12.8%	9.2%	9.7%	8.2%	7.4%	6.0%	5.0%	5.7%	5.2%	5.1%	5.7%	4.4%	4.9%	6.0%	6.0%	5
October 2021	143,658	25.0%	15.3%	11.6%	8.3%	8.3%	7.0%	6.2%	4.6%	4.0%	4.5%	4.2%	4.0%	4.4%	3.3%	3.8%	4.8%	4.7%	4.1%	4
November 2021	191,688	20.2%	12.8%	8.8%	8.6%	7.4%	6.3%	4.8%	4.0%	4.4%	4.0%	4.0%	4.4%	3.4%	3.7%	4.8%	4.5%	3.9%	4.8%	2
December 2021	141,362	20.9%	12.7%	11.9%	10.1%	8.3%	6.0%	5.2%	5.4%	4.8%	4.7%	5.2%	4.1%	4.4%	6.3%	5.5%	4.6%	5.7%	3.5%	3
January 2022	101,114	21.2%	16.1%	13.1%	10.3%	7.1%	6.2%	6.3%	5.5%	5.2%	5.7%	4.4%	4.9%	6.3%	5.8%	5.0%	6.1%	3.7%	3.3%	2
February 2022	85,537	23.4%	15.4%	11.8%	8.1%	7.0%	7.5%	6.6%	6.1%	6.2%	4.8%	5.1%	6.3%	5.9%	5.0%	6.0%	3.8%	3.3%	2.7%	
March 2022	107,666	19.2%	13.1%	8.5%	7.2%	7.6%	6.9%	6.2%	6.3%	5.0%	5.2%	6.3%	6.0%	5.1%	5.9%	3.8%	3.3%	2.6%		
April 2022	85,506	19.4%	10.3%	8.5%	8.8%	7.5%	7.0%	7.0%	5.3%	6.0%	7.4%	6.7%	5.6%	6.4%	4.0%	3.5%	2.9%			
May 2022	82,263	16.1%	11.2%	10.9%	9.9%	8.0%	8.2%	5.9%	6.2%	6.9%	7.3%	5.6%	6.5%	4.3%	3.6%	2.8%				
June 2022	64,097	19.3%	14.0%	11.5%	9.5%	9.6%	6.8%	6.9%	7.6%	8.0%	5.7%	6.3%	4.8%	4.2%	3.2%					
July 2022	71,568	21.7%	14.2%	10.7%	10.3%	7.3%	7.5%	7.4%	7.4%	5.6%	5.7%	4.6%	4.0%	3.2%						
August 2022	86,207	22.4%	14.2%	12.3%	8.5%	8.8%	8.8%	8.8%	6.6%	7.0%	5.2%	4.8%	3.9%							
September 2022	84,548	19.7%	13.5%	8.8%	8.4%	8.1%	8.3%	6.2%	6.4%	5.0%	4.2%	3.3%								
October 2022	95,936	18.0%	10.5%	9.3%	8.6%	8.2%	6.7%	6.9%	5.1%	4.5%	3.5%									
November 2022	98,316	17.1%	12.0%	10.5%	10.1%	7.1%	6.8%	5.3%	4.8%	3.7%										
December 2022	82,859	17.8%	11.7%	11.2%	7.5%	6.7%	5.4%	4.8%	3.7%											
January 2023	84,744	22.3%	15.6%	10.6%	9.6%	7.1%	6.4%	5.0%												
February 2023	99,129	21.0%	12.2%	11.3%	7.4%	6.4%	5.1%													
March 2023	110,018	16.0%	11.1%	7.9%	6.3%	4.8%														
April 2023	106,376	29.4%	12.8%	9.8%	7.1%															
May 2023	193,681	18.5%	10.5%	6.5%																
June 2023	101,712	16.8%	8.6%																	
July 2023	103,039	16.2%																		

Cohort Id
0 27

SUM(Retention)
1.0% 29.4%

- Most cohorts show a significant drop in retention in the second month. This is common in many platforms where the initial excitement or motivation to use a service decreases after the first interaction.
- Older cohorts (earlier months) tend to have better retention compared to newer ones. This could indicate stronger user loyalty in the initial user base or market dynamics that have affected newer user behavior.

2. **Fee USD by Cohort**:

Revenue Generated - Cohort Matrix

Cohort Id

Month of Month	Fee Usd	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
May 2021	9.12B	9.12B	3.79B	1.58B	1.71B	2.00B	1.92B	2.08B	1.67B	1.61B	0.81B	0.71B	0.52B	0.61B	0.32B	0.19B	0.25B	0.17B	0.16
June 2021	2.14B	2.14B	2.11B	2.99B	2.87B	2.57B	2.75B	3.04B	1.94B	1.29B	0.94B	0.99B	1.38B	0.79B	0.62B	0.51B	0.50B	0.35B	0.38
July 2021	0.88B	0.88B	1.02B	1.10B	0.80B	0.97B	1.02B	0.65B	0.20B	0.61B	0.68B	0.93B	0.56B	0.44B	0.36B	0.35B	0.26B	0.28B	0.12
August 2021	0.82B	0.82B	0.91B	0.66B	0.79B	0.68B	0.62B	0.30B	0.33B	0.27B	0.26B	0.15B	0.09B	0.09B	0.06B	0.05B	0.12B	0.05B	0.06
September 2021	0.86B	0.86B	0.70B	0.69B	0.54B	0.44B	0.20B	0.23B	0.17B	0.20B	0.15B	0.11B	0.09B	0.05B	0.03B	0.08B	0.03B	0.04B	0.07
October 2021	1.49B	1.49B	1.61B	2.00B	1.46B	0.27B	0.19B	0.16B	0.18B	0.08B	0.05B	0.08B	0.05B	0.08B	0.12B	0.04B	0.06B	0.06B	0.09
November 2021	1.64B	1.64B	0.96B	0.67B	0.31B	0.35B	0.31B	0.20B	0.11B	0.05B	0.05B	0.04B	0.03B	0.05B	0.02B	0.03B	0.05B	0.06B	0.04
December 2021	1.02B	1.02B	0.84B	0.43B	0.45B	0.42B	0.26B	0.18B	0.09B	0.07B	0.05B	0.03B	0.05B	0.02B	0.04B	0.07B	0.07B	0.04B	0.11
January 2022	0.86B	0.86B	0.42B	0.34B	0.25B	0.25B	0.11B	0.08B	0.09B	0.08B	0.05B	0.08B	0.05B	0.09B	0.14B	0.15B	0.07B	0.08B	0.04
February 2022	1.74B	1.74B	1.76B	1.31B	1.44B	1.30B	0.81B	0.75B	0.59B	0.21B	0.24B	0.08B	0.18B	0.30B	0.37B	0.11B	0.18B	0.13B	0.08
March 2022	0.73B	0.73B	0.38B	0.36B	0.21B	0.10B	0.09B	0.07B	0.06B	0.09B	0.03B	0.06B	0.08B	0.11B	0.05B	0.07B	0.03B	0.02B	0.02
April 2022	0.48B	0.48B	0.28B	0.13B	0.09B	0.10B	0.08B	0.05B	0.07B	0.02B	0.04B	0.06B	0.06B	0.04B	0.09B	0.03B	0.02B	0.02B	
May 2022	0.48B	0.48B	0.27B	0.11B	0.12B	0.08B	0.06B	0.07B	0.03B	0.05B	0.06B	0.07B	0.04B	0.05B	0.02B	0.01B	0.01B		
June 2022	0.54B	0.54B	0.91B	0.77B	0.44B	0.23B	0.27B	0.06B	0.10B	0.18B	0.14B	0.09B	0.07B	0.06B	0.04B	0.01B			
July 2022	0.16B	0.16B	0.14B	0.09B	0.08B	0.13B	0.04B	0.05B	0.08B	0.07B	0.04B	0.06B	0.04B	0.03B	0.02B				
August 2022	0.30B	0.30B	0.37B	0.37B	1.09B	0.21B	0.33B	0.24B	0.22B	0.14B	0.22B	0.19B	0.13B	0.06B					
September 2022	0.19B	0.19B	0.13B	0.13B	0.04B	0.07B	0.09B	0.10B	0.05B	0.06B	0.03B	0.02B	0.01B						
October 2022	0.13B	0.13B	0.14B	0.04B	0.06B	0.08B	0.11B	0.08B	0.10B	0.08B	0.08B	0.05B							
November 2022	0.45B	0.45B	0.21B	0.27B	0.36B	0.46B	0.23B	0.08B	0.03B	0.03B	0.02B								
December 2022	0.10B	0.10B	0.09B	0.10B	0.08B	0.04B	0.05B	0.03B	0.02B	0.02B									
January 2023	0.21B	0.21B	0.45B	0.51B	0.22B	0.16B	0.04B	0.03B	0.02B										
February 2023	0.30B	0.30B	0.19B	0.09B	0.11B	0.03B	0.03B	0.02B											
March 2023	0.57B	0.57B	0.30B	0.21B	0.09B	0.15B	0.10B												
April 2023	0.23B	0.23B	0.42B	0.13B	0.09B	0.06B													
May 2023	0.69B	0.69B	0.22B	0.10B	0.04B														
June 2023	0.20B	0.20B	0.15B	0.15B															
July 2023	0.19B	0.19B	0.17B																
August 2023	0.14B	0.14B																	

SUM(Fee Usd Cohort)
0.00B 9.12B

- The fees (which can be considered as revenue for the protocol) tend to be the highest during the first month of user interaction with the protocol (COHORT_ID 0).
- There's a sharp decline in the trading volume for subsequent months for all cohorts. However, there's some variation in trading volume across different cohorts, suggesting that some months attracted more high-value traders than others.
- It's hard to overlook the significant volume contributed by users who joined the protocol in its first 10 months, especially when compared to the revenue generated by users who joined Uniswap in 2023

3. **Swap Count by Cohort**:

N of Swps- Cohort Matrix

		Cohort Id																	
Month of Month	Swap Count	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
May 2021	805.0K	805.0K	559.5K	441.5K	343.3K	274.2K	263.7K	255.4K	214.9K	172.6K	117.6K	131.7K	109.7K	114.4K	97.9K	85.4K	107.2K	108.0K	94.5K
June 2021	659.5K	659.5K	425.7K	256.6K	191.0K	160.8K	155.2K	151.4K	116.1K	86.1K	107.2K	92.6K	134.4K	116.5K	123.6K	119.7K	135.1K	98.0K	105.5K
July 2021	591.9K	591.9K	283.2K	157.0K	114.1K	102.8K	95.8K	68.7K	46.0K	84.4K	70.2K	103.8K	91.7K	101.6K	95.6K	98.5K	76.1K	84.8K	53.2K
August 2021	348.7K	348.7K	156.8K	113.7K	101.6K	89.9K	75.6K	50.2K	61.4K	52.7K	48.4K	43.9K	35.8K	38.6K	34.8K	32.7K	45.9K	31.0K	37.2K
September 2021	263.3K	263.3K	127.5K	107.4K	97.8K	68.8K	49.1K	59.1K	46.8K	49.4K	42.2K	41.6K	41.7K	32.3K	27.1K	39.5K	26.2K	34.7K	46.1K
October 2021	329.2K	329.2K	190.9K	154.3K	116.1K	61.0K	64.7K	54.4K	58.9K	49.5K	47.6K	50.5K	38.1K	32.9K	36.9K	23.0K	29.0K	35.8K	33.2K
November 2021	412.1K	412.1K	171.5K	110.8K	74.3K	83.4K	69.9K	55.1K	45.0K	41.0K	43.7K	41.6K	37.6K	46.0K	26.6K	31.2K	42.3K	38.5K	31.4K
December 2021	302.7K	302.7K	130.0K	87.7K	96.5K	87.5K	74.2K	65.7K	49.7K	45.7K	42.4K	37.9K	45.5K	29.4K	35.5K	43.5K	39.2K	31.8K	59.8K
January 2022	210.6K	210.6K	81.0K	69.2K	56.1K	49.7K	41.1K	39.0K	47.0K	49.1K	40.9K	57.9K	38.4K	57.2K	70.3K	58.1K	41.9K	48.4K	31.9K
February 2022	221.8K	221.8K	163.1K	117.3K	120.9K	124.4K	113.5K	117.0K	102.6K	69.6K	87.6K	36.9K	65.7K	82.5K	71.9K	38.7K	46.5K	34.2K	26.4K
March 2022	238.5K	238.5K	91.5K	78.7K	63.9K	58.5K	58.5K	51.4K	47.8K	62.8K	36.0K	42.0K	44.7K	42.9K	29.3K	36.7K	20.5K	16.9K	12.8K
April 2022	187.1K	187.1K	83.7K	61.4K	55.2K	55.1K	51.3K	41.5K	52.0K	27.1K	34.1K	37.7K	32.4K	24.2K	38.9K	21.6K	18.3K	13.2K	
May 2022	194.0K	194.0K	78.4K	59.9K	58.8K	53.7K	39.9K	47.1K	26.1K	33.3K	33.5K	30.4K	22.1K	28.8K	16.7K	14.3K	10.9K		
June 2022	185.8K	185.8K	136.0K	120.3K	93.9K	65.7K	77.4K	36.3K	46.0K	55.1K	48.8K	28.0K	27.9K	22.0K	15.1K	8.4K			
July 2022	191.8K	191.8K	118.8K	80.0K	60.5K	69.8K	37.4K	41.3K	40.4K	35.7K	24.5K	29.0K	23.4K	18.6K	12.5K				
August 2022	234.8K	234.8K	156.6K	120.4K	180.2K	71.4K	87.2K	79.8K	69.9K	47.0K	66.2K	47.4K	37.7K	25.8K					
September 2022	234.3K	234.3K	102.8K	81.4K	44.9K	53.9K	52.2K	48.7K	31.7K	36.7K	21.5K	18.0K	14.6K						
October 2022	236.3K	236.3K	107.6K	53.7K	57.9K	68.7K	90.6K	63.6K	72.6K	55.2K	42.0K	27.3K							
November 2022	269.1K	269.1K	110.0K	105.1K	104.8K	116.1K	61.8K	50.1K	29.6K	25.1K	20.0K								
December 2022	195.5K	195.5K	109.5K	80.3K	54.8K	31.1K	33.1K	22.9K	18.6K	15.2K									
January 2023	228.7K	228.7K	150.7K	127.9K	67.2K	77.1K	38.0K	27.6K	22.7K										
February 2023	263.9K	263.9K	138.5K	69.9K	79.3K	36.7K	28.6K	22.5K											
March 2023	322.7K	322.7K	138.0K	99.0K	54.0K	48.0K	37.0K												
April 2023	268.9K	268.9K	239.7K	90.2K	66.0K	46.3K													
May 2023	729.5K	729.5K	218.9K	106.1K	59.2K														
June 2023	295.0K	295.0K	146.1K	101.3K															
July 2023	265.4K	265.4K	103.7K																
August 2023	201.9K	201.9K																	

SUM(Swap Count Coho...

8.4K 805.0K

- The number of swaps (transactions) is also highest in the first month of user interaction. This indicates that users are most active immediately after joining.
- The subsequent months show a decline in activity, but there are still some variations across cohorts.

3.4 Bonus: User Clusters

We employed the **K-means Clustering** algorithm to create a custom user cluster, aiming to identify distinct group characteristics. Highlighting these characteristics could aid the protocol in attracting and retaining strategic users. The implementation of k-means was broken down into the following steps:

1. Feature Engineering (data preparation). Here, it was used the dataset **univ3_metrics_by_user.csv**
2. Determining the Number of Clusters
3. Applying K-means Clustering

You can find the detailed code under the section "K-means Clustering Application" in the file `main.ipynb` within the provided repository: https://github.com/LimaRods/Uniswap-User-Behaviour/blob/main/main.ipynb.

The group discovered are:

1. **Cluster 0**:
    - **Swap Count**: Low (around 26 swaps on average)
    - **Volume in USD**: Low (around $698,744 on average)
    - **Revenue in USD**: Low (around $106,664 on average)
    - **Average Transaction Value**: Low (around $15,980 on average)
    - **Average Gas Price**: Moderate (around 62 on average)
2. **Cluster 1**:
    - **Swap Count**: Very Low (around 1 swap on average)
    - **Volume in USD**: Very Low (around $55,182 on average)
    - **Revenue in USD**: Very Low (around $12,656 on average)
    - **Average Transaction Value**: Moderate (around $27,547 on average)
    - **Average Gas Price**: Extremely High (around 5,378 on average)
3. **Cluster 2**:
    - **Swap Count**: Extremely High (around 54,127 swaps on average)
    - **Volume in USD**: Extremely High (around $4.43 billion on average)
    - **Revenue in USD**: Extremely High (around $558 million on average)
    - **Average Transaction Value**: High (around $167,220 on average)
    - **Average Gas Price**: Moderate (around 106 on average)
4. **Cluster 3**:
    - **Swap Count**: Low (around 10 swaps on average)

- **Volume in USD**: Moderate (around $60.25 million on average)
- **Revenue in USD**: High (around $2.74 million on average)
- **Average Transaction Value**: Extremely High (around $7.69 million on average)
- **Average Gas Price**: Moderate (around 55 on average)

4. Recommendation and Conclusion 🎯

Given these observations described and data discussed extensively in topic 3, we can derive insights and propose recommended actions such as:

4.1 From Cohort, Correlation and Time-Series Analysis

1. **Focus on Onboarding and Initial Experience:**
   - Given the sharp drop-off after the first month, the protocol should investigate its onboarding process. Improving the initial user experience, offering tutorials, or providing incentives might help retain more users after their first interaction.
2. **Engagement Initiatives in Subsequent Months**:
   - Given the significant decline in volume, revenue, and transaction count following the initial month of user interaction, it's imperative for the protocol to implement engagement strategies to boost user retention. Possible initiatives could include promotional offers, loyalty rewards, a points system to recognize frequent traders or those providing high volume, and other user incentives.
3. **Focus on High-Value Cohorts**:
   - Some cohorts (like the May 2021 cohort) seem to have a higher average transaction value in the subsequent months. The protocol should analyze the characteristics of these cohorts to understand what attracted these high-value users and replicate similar strategies in the future.
4. **Optimize Gas Prices**:
   - High Ethereum gas prices can be a deterrent for users to transact frequently. The protocol should consider strategies to optimize gas fees or offer alternatives like Layer 2 solutions or integrations with other low-fee chains.
5. **Educational Content:**
   - Considering the influence of macroeconomic factors on the crypto market, it's evident from our observations that an increase in interest rates can result in reduced liquidity

and significantly lower trading volumes. Educating users about market cycles and the effects of macroeconomic conditions on volatile markets like crypto can be beneficial. The Web3 realm is no longer isolated from real-world economic dynamics.

Additional Data for Deeper Analysis:

1. **User Demographics**: Understanding where the users are from, their age, and other demographic factors can provide insights into which cohorts are more valuable and why.
2. **User Feedback**: Direct feedback from users can help understand their pain points, needs, and reasons for reduced activity in subsequent months.

4.2 From the Clustering:

Focus on Cluster 0:

- Users in Cluster 0 have a moderate average transaction value (around $15,980) with a relatively low number of swaps (26 swaps on average). This suggests that they might be trading more substantial amounts less frequently.
- By focusing on this cluster, the protocol can aim to increase the frequency of swaps for these users. Since they already have a higher average transaction value, getting them to trade more frequently could be beneficial.
- Strategies could include:
  - **Trading Education**: Offering educational content or webinars on trading strategies, benefits of frequent trading, etc.
  - **Incentives**: Providing fee reductions or other incentives for users who increase their trading frequency.

Secondary Focus on Cluster 2:

- Cluster 2 represents the "whales" or high-volume traders. While it's tempting to focus heavily on them due to their high volume and revenue generation, it's also essential to diversify the user base. Over-reliance on a few high-volume traders can be risky.
- Strategies:
  - **Personalized Features and Services**: Retaining them is crucial. Personalized services, dedicated support, or even offering advanced trading tools/features could be strategies to ensure their continued engagement.

5. References 📚

- https://uniswap.org/
- https://defillama.com/
- **Flipside LiveQuery:** https://flipsidecrypto.xyz/livequery
- **A beginner's guide to cohort analysis:** https://www.appcues.com/blog/cohort-analysis#:~:text=with cohort analysis.-,What is cohort analysis%3F,of people with shared characteristics.
- **K-means Custering**: https://scikit-learn.org/stable/auto_examples/cluster/plot_cluster_iris.html#k-means-clustering