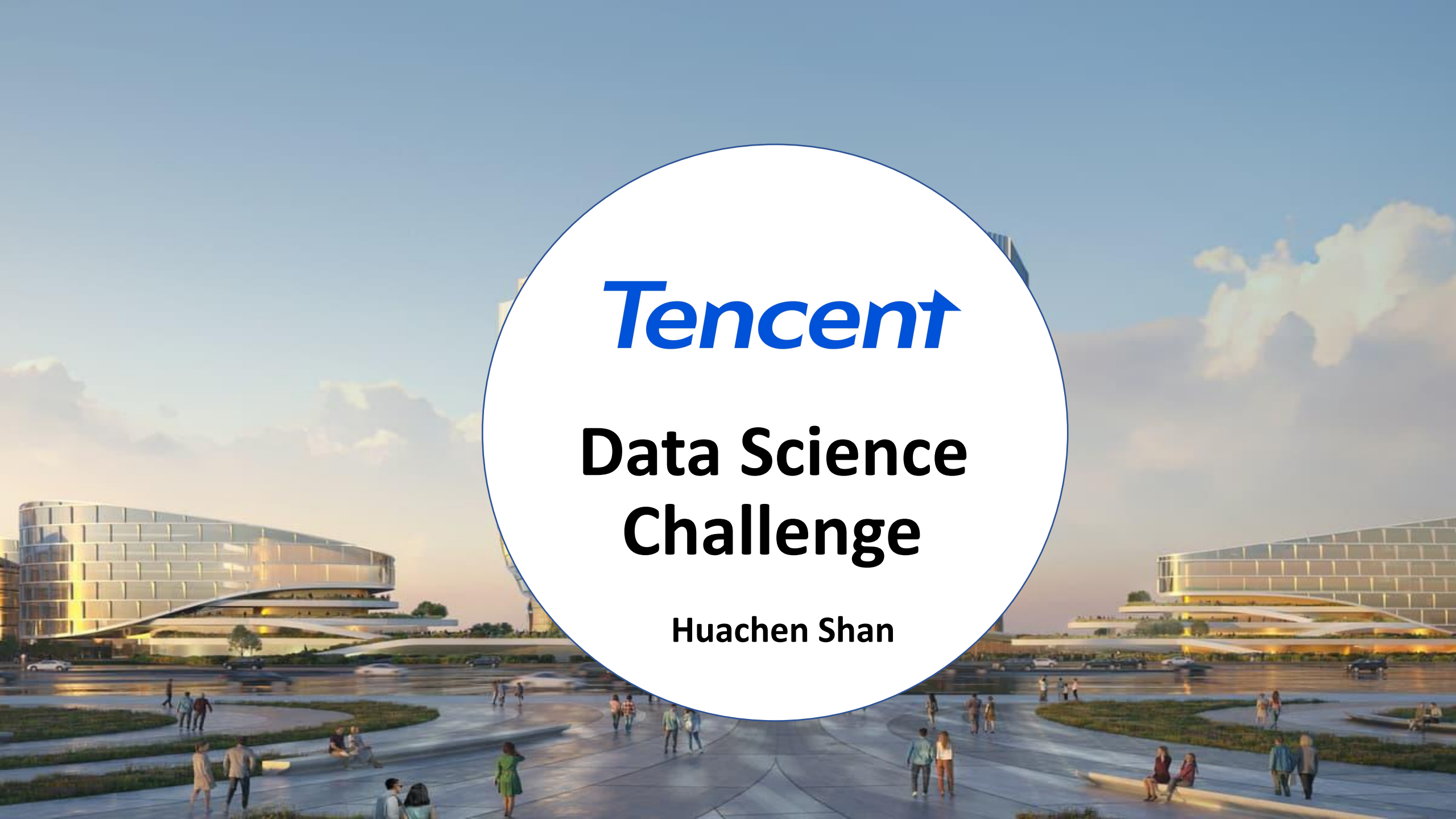


Data Science Challenge

Huachen Shan

Part 1: AB Testing Result Readout Analysis from dataset 2

Prework

Metrics Calculation Time Range:

- Relative **14** days.
- Users may have different app usage behaviors between weekdays and weekends —> Capture the weekly pattern
- **7** days evaluation period may be too short to avoid the novelty effect

Statistical Test

- **T Test**
- All metrics are means metric.
- All the data come from **two** different group. so we use **2 sample T test** to test three combinations (group A and the control group, group B and the control group, group C and the control group).

Data Cleaning

- Count the number of missing values and the proportion for each field.
- The number of missing values for each field is **0**, so we do not need to perform any additional actions for missing value deletion or filling.

Part 1: AB Testing Result Readout Analysis from dataset 2

Metrics Selection

Game Overall Engagement Metrics:

- Total **online time per user**
- Total **login count per user**

Develop Hypothesis Testing for treatment group A / group B / group C and control group

- H0: The average online time per user / login count per user of group A / group B / group C is **greater** than the average of group Control, **μControl ≤ μA / μB / μC**.
- H1: The average online time per user / login count per user of group A / group B / group C is **less than or equal to** the average of group Control, **μA / μB / μC ≤ μControl.**

Evaluation Metrics	Treatment A vs Control	Treatment B vs Control	Treatment C vs Control
Player's total online time	[-106.28, 135.19], 0.809235	[-123.9, 102.95], 0.852297	[-127.44, 113.03], 0.903805
Player's total login count	[-0.19, 0.25] , 0.778244	[-0.17, 0.17] , 0.968045	[-0.26, 0.25], 0.96386

Among all the tests, since all the p values is greater than 0.1, this means that we can not reject the null hypothesis. They are not statistically significant and The average online time per user / login count per user of group A / group B / group C is **not greater** than the average of group Control.

Part 1: AB Testing Result Readout Analysis from dataset 2

Metrics Selection

In Game Engagement Metrics:

- Total Counts **ranked top 10/per user**
- Total player's **total killing count/per user**
- Total player's **total survival time/per user**

Develop Hypothesis Testing for treatment group A / group B / group C and control group

- H0: The number of counts ranked Top 10 / Player's total killing count / Player's total survival time of group A / group B / group C is **greater** than the average of group Control, **μControl ≤ μA / μB / μC**.
- H1: The number of counts ranked Top 10 / Player's total killing count / Player's total survival time of group A / group B / group C is **less than or equal to** the average of group Control, **μA / μB / μC ≤ μControl.**

Evaluation Metrics	Treatment A vs Control	Treatment B vs Control	Treatment C vs Control
Counts ranked Top 10	[-0.08, 0.04], 0.465524	[-0.05, 0.08], 0.67184	[-0.07, 0.06], 0.932811
Player's total killing count	[-0.2, 0.24] , 0.885292	[-0.21, 0.25], 0.853033	[-0.35, 0.17], 0.495925
Player's total survival time	[-60.27, 86.45], 0.719133	**[-104.79, 8.08], 0.090652**	[-108.82, 52.66], 0.484544

Among all the tests, since all the p values is greater than **0.1** except **the p value between treatment group B and control group when evaluating player's total survival time**, this means that we can only reject the null hypothesis in that circumstance. **Only the player's total survival time of treatment B is greater than the control group.** The remaining metrics in treatment group are not statistically significant greater than the control group.

Part 1: AB Testing Result Readout Analysis from dataset 2

Metrics Selection

Monetization Metrics:

- Total player's **payment time per user**
- Total player's **payment amount per user**
- Testing the players' **willingness to pay** for the game

Develop Hypothesis Testing for treatment group A / group B / group C and control group

- H0: The average payment times / payment amount per user of group A / group B / group C is **greater** than the average of group Control, **μControl ≤ μA / μB / μC**.
- H1: The average payment times / payment amount per user of group A / group B / group C is **less than or equal to** the average of group Control, **μA / μB / μC ≤ μControl.**

Evaluation Metrics	Treatment A vs Control	Treatment B vs Control	Treatment C vs Control
Payment times	[-0.0, 0.0], 0.732421	[-0.0, 0.0], 0.173675	[-0.0, 0.0], 0.131404
Payment amount	[-0.19, 0.66], 0.264054	[-0.21, 0.24], 0.887261	[-0.18, 0.78], 0.217655

Among all the tests, since all the p values is greater than 0.1, this means that we can not reject the null hypothesis. They are not statistically significant and The average Payment times / Payment amount per user of group A / group B / group C is **not greater** than the average of group Control.

Part 1: AB Testing Result Readout Analysis from dataset 2

Question: Should Tencent launch the change?

Game Overall Engagement Metrics:

- Total **online time per user**
- Total **login count per user**
- Testing the players' **overall engagement** toward the game

In Game Engagement Metrics:

- Total Counts **ranked top 10 per user**
- Total player's **total killing count per user**
- Total player's **total survival time per user**

Monetization Metrics:

- Total player's **payment time per user**
- Total player's **payment amount per user**
- Testing the players' **willingness to pay** for the game

Takeaway

We have run an A/B test to see the average gap of different metrics between the treatment groups (A/B/C) and the control group. From the t-test, we get p-value from treatment group A, and treatment group C are much **larger** than 0.1, the p value between **treatment group B** and control group when evaluating player's total survival time is approximately **0.09**, which is **smaller than 0.1**, meaning that the player's total survival time of treatment B is greater than the control group. This indicates that the player's in treatment B is statistically significant **more engaging** than the control group. The team should launch the changes in **group B** to engage more players to play the game.

Part 2: Product Analytics and Reporting from Dataset 1

Prework

Dimension of the Test

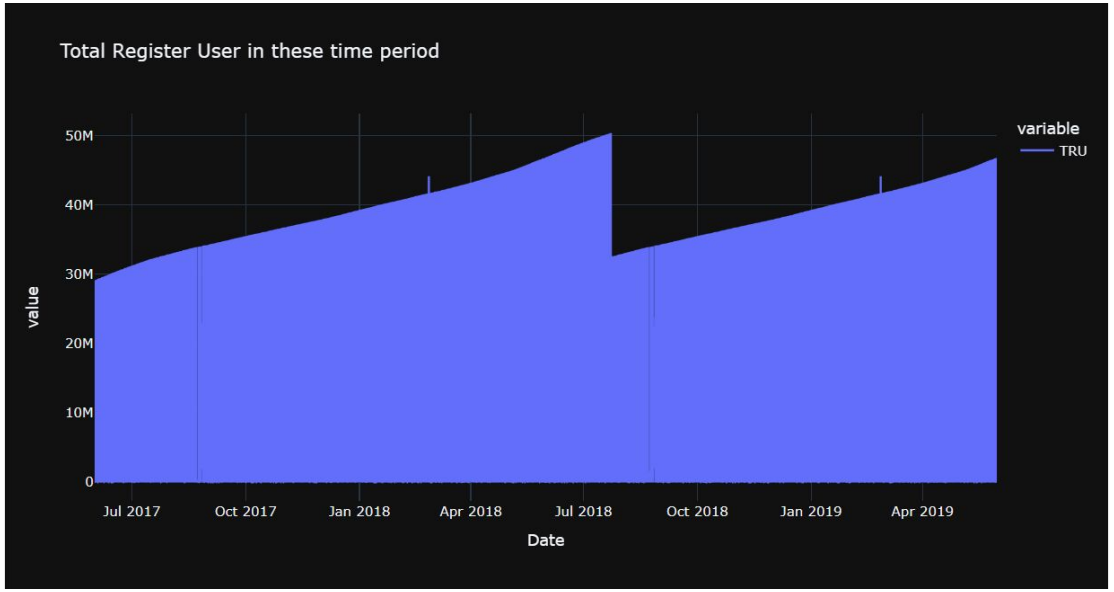
- By platform (IOS or Android)
- By country
- By time

Metrics Selection

- **Top line Metrics**
  - **Total register users** in the entire time range
  - **Total revenue** in the entire time range
  - **Daily/Weekly/Monthly average online time**
- **User Growth Metrics**
  - **Daily / Monthly / Weekly** active users
  - **New register users**
  - Average **D1** retention rate
  - Average **D7** retention rate
- **Monetization Metrics**
  - **Payer**
  - **UC total outflow**

Part 2: Product Analytics and Reporting from Dataset 2

Overall



Takeaways:

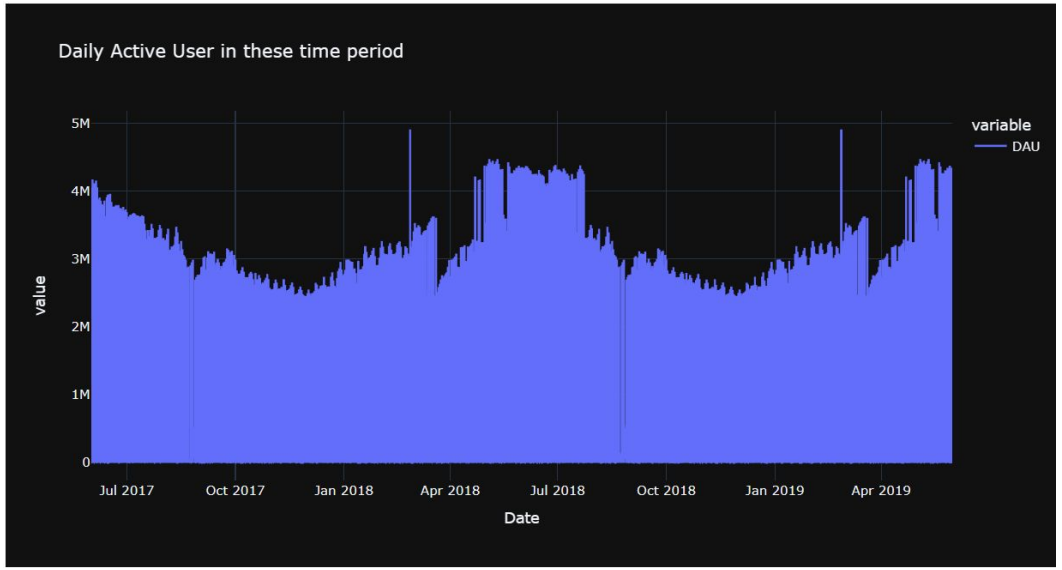
- **Within the 2 year range, from July 2017 to July 2019, the overall register user trend shows strong seasonality pattern**
  - **The second quarter of each year is the peak season.**
  - **The number of Total Registered Users each year is relatively neutral.**
  - **The number of Total Registered Users is relatively the lowest in July of each year, gradually increasing from July of the first year to July of the second year. Then it will suddenly drop to a relatively low level.**

Decompose by Android vs IOS

- **55.02% of register users are from Android; 44.97% from IOS**
- **Both the Android and the IOS users show increasing trend with seasonality pattern on the total register user.**

Part 2: Product Analytics and Reporting from Dataset 2

Overall



Takeaways:

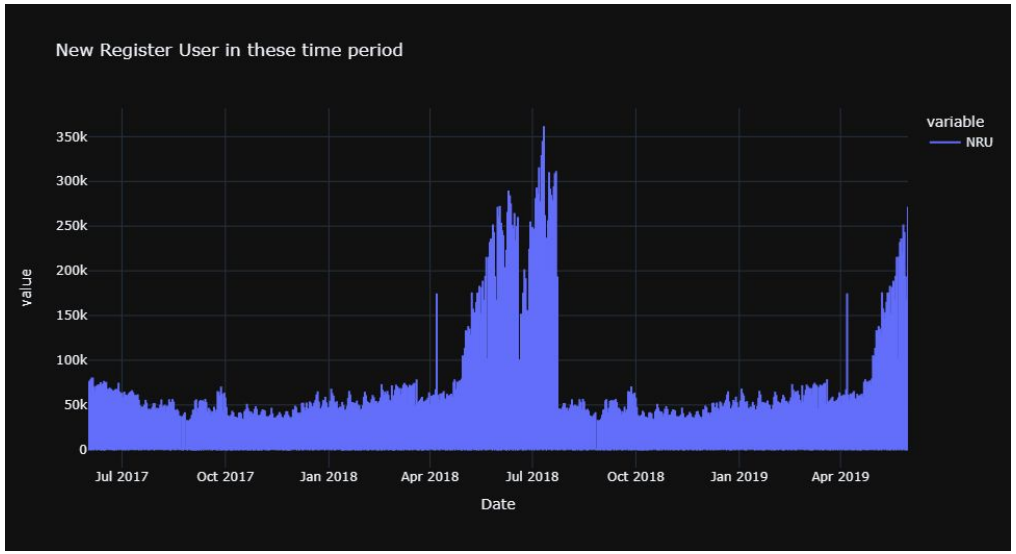
- **Within the 2 year range, from July 2017 to July 2019, the daily active user trend shows strong seasonality pattern**
  - **May to July of each year is the peak season.**
  - **The number of daily active user each year is relatively neutral.**
  - **The number of daily active user is relatively the lowest in December of each year, gradually increasing from December of the first year to July of the second year. Then it gradually decreases from July of the second year to December of the second year.**

Decompose by Android vs IOS

- **Both the Android and the IOS users show increasing trend with seasonality pattern on the daily active user.**

Part 2: Product Analytics and Reporting from Dataset 2

Overall



Takeaways:

- **Within the 2 year range, from July 2017 to July 2019, the new register user's trend shows strong annually pattern**
  - **July of each year is the peak month.**
  - **The number of new register user each year is relatively neutral.**
  - **The number of new register users is relatively the highest (about 350K) in July of each year. Then it will suddenly drop to a relatively low level (about 50K). The number of New Register User will almost remain the same from July of the first year to April of the second year. Then it will experience a steady increases from 50K in April of the second year to 350 K in July of the second year.**

Decompose by Android vs IOS

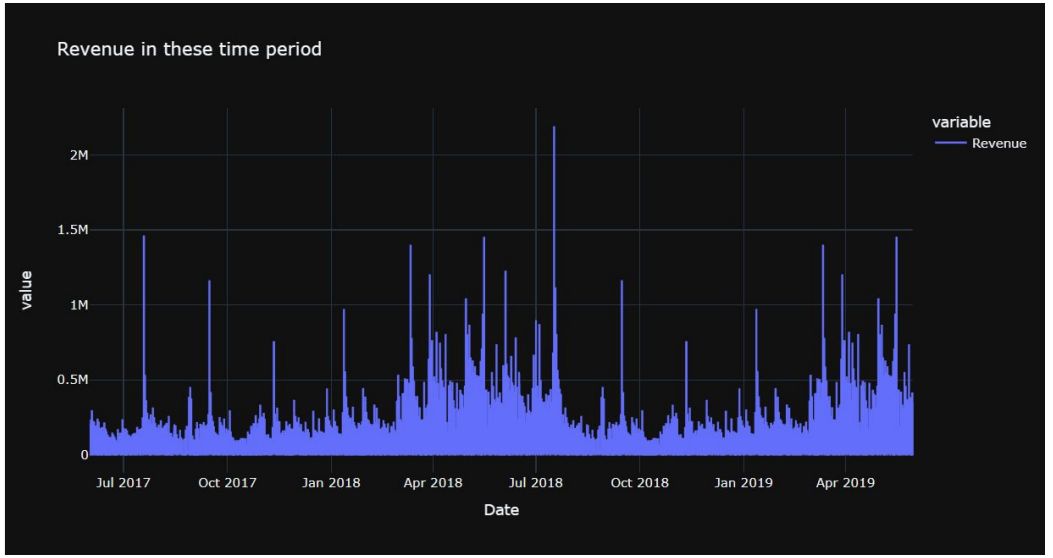
- **Both the Android and the IOS users show increasing trend with seasonality pattern on the new register user.**

Insights Analysis

- **We could consider drive commercialization from April to July by posting advertisement on social media. Conducting some online game activities or payment discounts for users to attract more users to engage in the game. since there is a steadily increase of new register user from April to July.**

Part 2: Product Analytics and Reporting from Dataset 2

Overall



Takeaways:

- **Within the 2 year range, from July 2017 to July 2019, the revenue's trend shows strong seasonality pattern**
  - **May to July of each year is the peak season.**
  - **The revenue of each year is relatively neutral.**
  - **The daily revenue is relatively higher (over 0.5M) from April to July of each year, it is relatively lower (less than 0.5M) in the remaining year. The daily revenue will almost remain the same (less than 0.5M) from August of the first year to April of the second year. Then it will experience a noticeable increases to over 0.5M from April to July**

Decompose by Android vs IOS

- **Both the Android and the IOS users show increasing trend with seasonality pattern on the new register user.**

Insights Analysis

- **The noticeable increase of revenue from April to July could be due to an increase in market demand from April to July, and the marketing campaigns for the game being more successful in this quarter. In response to the significant increase in revenue, the company could consider the following business decisions: Reinvestment: Reinvest a portion of the revenue into the research and development of game props and the improvement of game performance, or Market Expansion: Increase the marketing budget to expand market share and consolidate the company's market position.**

Part 3: Clustering/prediction models based on the attached dataset 2.

Question: Should Tencent launch the change?

Game Overall Engagement Metrics:

- Total **online time per user**
- Total **login count per user**
- Testing the players' **overall engagement** toward the game

In Game Engagement Metrics:

- Total Counts **ranked top 10 per user**
- Total player's **total killing count per user**
- Total player's **total survival time per user**

Monetization Metrics:

- Total player's **payment time per user**
- Total player's **payment amount per user**
- Testing the players' **willingness to pay** for the game

We first define hypothesis of the Hopkins statistical test.

- **H0**: The dataset is **not uniformly distributed** (contains meaningful clusters).
- **H1**: The dataset is **uniformly distributed** (no meaningful clusters).

Based on the result, the result is **0.98** which is smaller than 0.99. It is applicable for us to use clustering.

Part 3: Clustering/prediction models based on the attached dataset 2.

Question: Should Tencent launch the change?

Game Overall Engagement Metrics:

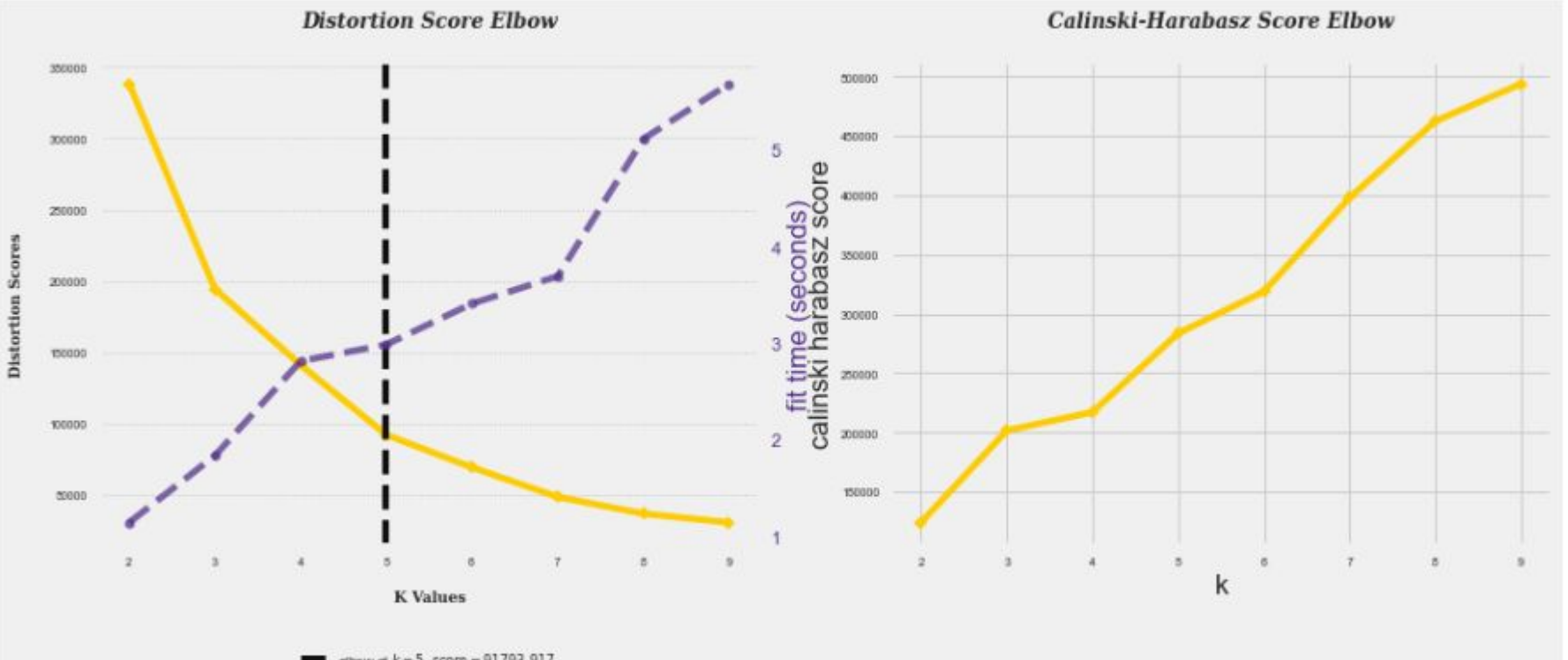
- Total **online time per user**
- Total **login count per user**
- Testing the players' **overall engagement** toward the game

In Game Engagement Metrics:

- Total Counts **ranked top 10 per user**
- Total player's **total killing count per user**
- Total player's **total survival time per user**

Monetization Metrics:

- Total player's **payment time per user**
- Total player's **payment amount per user**
- Testing the players' **willingness to pay** for the game



K values selection:

- Based on the results of the elbow method and Calinski Harabasz score above, it can be concluded that the **best clustering number** for the K-Means algorithm is **4 clusters**.