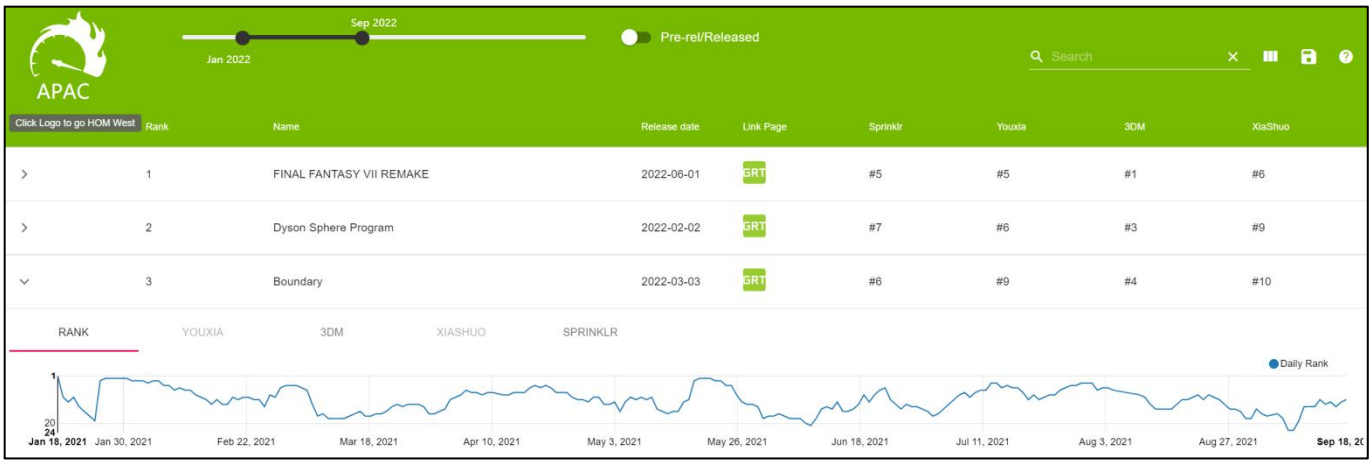
Project Hype-O-Meter APAC

1. Overview



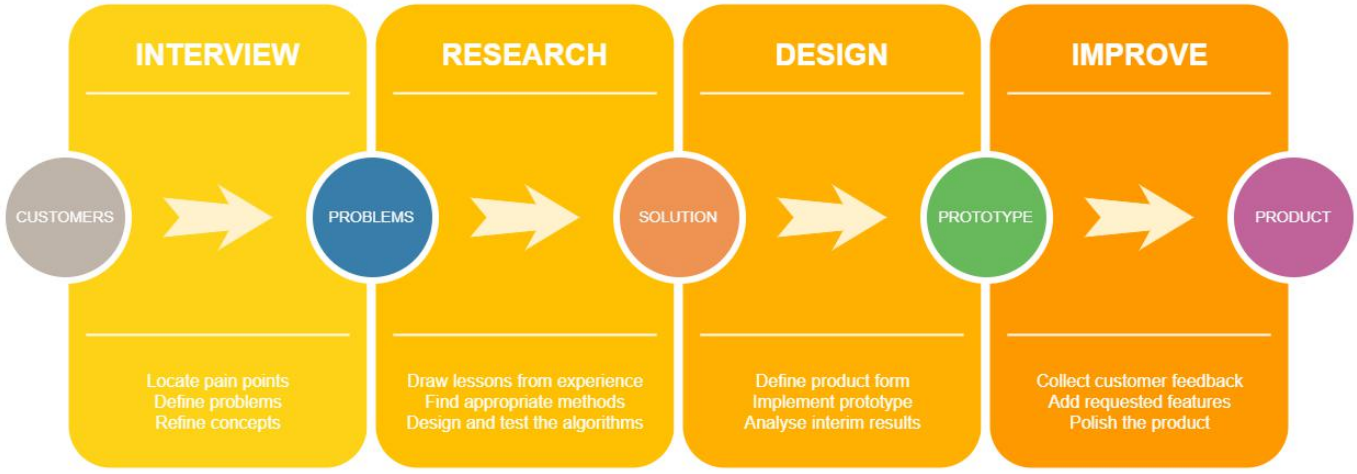
Graph 1. *The web page UI of Hype-O-Meter APAC system*

Project name: Hype-O-Meter APAC - A decision aid system for game popularity capture and prediction in APAC region
Project type: Group project with 3 engineers including myself
My role: Project owner, system design, UI design and front-end development

Project introduction:

The project Hype-O-Meter APAC aims at helping the tech marketing and game optimization teams make more accurate and efficient decisions on selecting potential popular games for further cooperation and support. It frees people from the tedious manual process of monitoring and evaluating social media and game websites by providing a solid automatic methodology for capturing and predicting the popularity of games in APAC market. It serves as a web tool which provides an intuitive and human-friendly UI for the users to get insight of the game popularity trend, and becomes a portal to connect external and internal resources together to make further operations more convenient.

2. Process of design

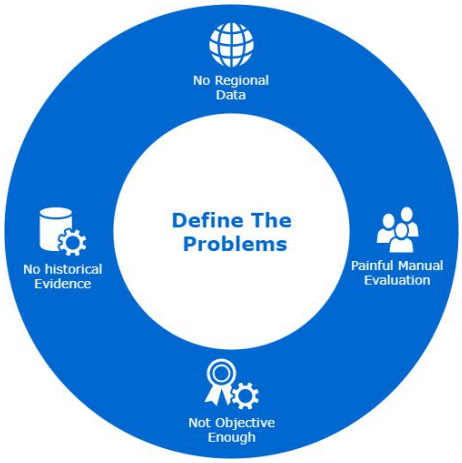


Graph 2. *The overall process of design*

2.1 Define the problems

The trend of game popularity in the Asia-pacific region, especially in China, is of great concern to the tech marketing team and game optimization team, who need make accurate and timely decisions in selecting games with potential or exiting popularity across game communities to seek for further cooperation and customer support.

But the process of obtaining insight of the popularity of various game products is always troublesome for the teams because of the complex composition of the game media and the inefficiency of the human process.



Graph 3. *The problems we concluded by interviewing the users*

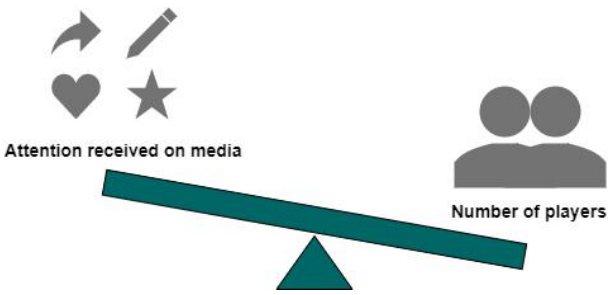
By interviewing the colleagues, we further concluded the detailed problems as below:

- The lack of regional sales information in the mainstream game platform like steam and epic leads to the difficulty of determining regional game popularity, especially for the Chinese mainland market where the preference of local players and the game/social medias are quite different from other regions, so the teams have to exploring the local game websites to get information.
- It requires a lot of manual monitoring and evaluating across multiple game medias or websites to get insights of regional game popularity trend, which is tedious and painful for people.
- It's hard to objectively evaluate how popular a game is because different game websites holds different views, thus the selection of target game are highly depending on the evaluator' s experience, which might be lack of enough data support.
- Regardless the success or failure of a target selection, the reason and historical evidence are hard to track and cannot be utilized for future reference efficiently.

So the primary target of our decision aid system is to provide solutions to solve above pain points, thus benefit the teams to make the decision making more accurate and make the decision-making process itself more human-friendly.

2.2 Refine the concept to make a clearer goal

A vital step of the system design is to determine the actual concept of "game popularity", is it the attention received by a game across the game medias or the actual number of players after the game's release?



Graph 4. *The concept of "game popularity" is determined as the game's number of players by interviewing our users*

By interviewing the colleagues from both teams, we finalized the concept of "game popularity" as the number of actual players after its release by below reasons:

- From the perspective of the tech marketing team, more players after the game's release means more sales numbers, which indicates the cooperation between companies are bringing more revenues to the developer as well as delivering a positive reputation and publicity on us.
- From the perspective of the game optimization team, more players after the game's release means more end users with our product will benefit from their optimization scheme, and the improvement of game experience will contribute to the attractiveness of our products to more potential buyers.

So the refined concept points out the actual goal for the project: automatically collecting necessary data before a game's release to predict the number of players after its retail so that both teams can select target game more accurately and timely for further operations.

As the actual total number of players are only available to the games' developers, we can only get the number of concurrent players by our data collecting system, **hereafter the concept of "the number of players" is referred to "the number of concurrent players" equivalently**.

2.3 Solution and system design

Although we abandoned the idea of attention received on game medias when interviewing the teams to refine the concept of "game popularity", it actually gave us a direction of how the teams used to do the monitoring and evaluating work to collect supportive evidence for the game's popularity.

So we've decided the main direction of how to design such a system: capturing the received attention independently across the most popular local game community or portal websites for each game, then work out a convincing method to integrate the attentions from all these sources into one combined score, then rank all the games by the final score. All the process should be handled by programs so that we can free the teams from manually doing such tedious works again.

Automating the data collection process only involves pure engineering, but how to select the proper game websites as data sources and work out an appropriate methodology to integrate all the attention received from all sources into one numeric score? More important, how to convince people that our system is reliable at predicting and capturing the popularity of games?

It requires the answering of below questions:

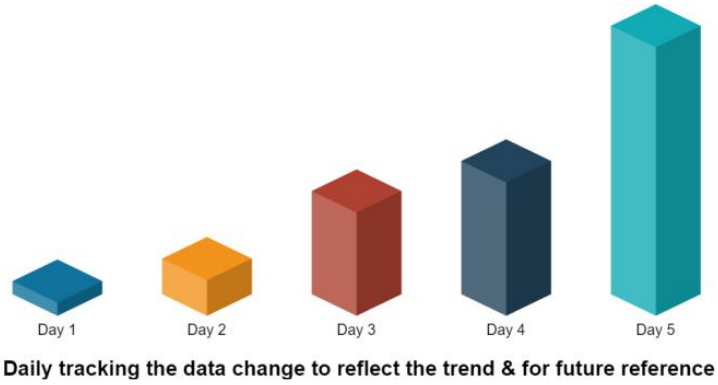
2.3.1 How to convince people that the attention received by a game is comparable?

Graph 5. *A sample of numeric data from game media or website, this website uses the number comments and the number of topics*

Human beings are usually more convinced by numbers rather than pure text description or a rank list without clear formula of how it's calculated, so the first criterion of selecting game website is that it should have explicit numeric data to indicate the attention received by the game, like the number of viewers for the game's page which can be objectively measured. Under this circumstances, games can compare between each other with no ambiguity by comparing such numeric data within the same website, thus we can get a clear view of a game's popularity in each website.

2.3.2 How to convince people that the attention received by the game can be well tracked by our system, and can be referenced effectively in the future ?

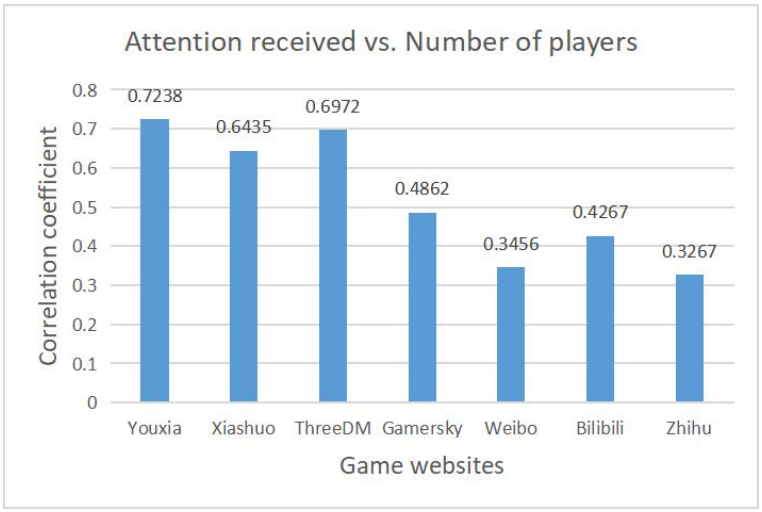
A website's content is continuously changed in the synergy of the publishers and the users' reaction, to keep track of these trending timely, we need to take snapshots periodically to save a copy of these numeric data which indicates the attention received by each game at certain moment. However, it's impractical to make our system track all the changes since they are happening every seconds, but we still need to make our data collection sensitive enough to reflect the trends.



Graph 6. *We finalized the way to track the trend of game popularity: Daily snapshot the data for each selected website, and store it for future reference*

So after discussing with our customers, we finally decided to snapshot each games' numeric data for all those suitable website at a interval of 24 hours, and the time of taking such snapshot is set to the early morning of every day's 3:00 am when most of the people in ASIA region have fallen asleep thus the changes happening on the websites should be at its minimum.

2.3.3 How to convince people that the attention received by the game before its release is indeed related to the game's number of users after its release?

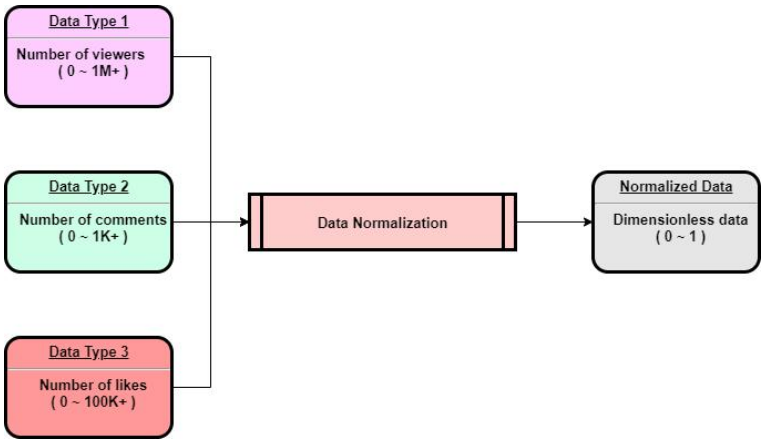


Graph 7. *The attention received on the website vs. the number of players after release for different game websites, measured by Spearman correlation coefficient*

To address this question, we conducted a correlation analysis between the 2 series of variables where x refers to a game's attention received from a website like the average number of viewers at 1 week before its release and y refers to its peak concurrent users in its first month of release, by comparing the Spearman correlation coefficient of above 2 variables across all the candidate websites, we finally get the most positively correlated 3 sources as our final data sources, all of them have a higher correlation coefficient than 0.6, which indicates that Y will high likely increase as x increases.

2.3.4 How to convince people that the attention received by the game, which presents in different format of numeric numbers, can be integrated together without introducing new problems?

Different websites indicate the attention received by a game in different way, like the number of viewers, the number of likes or the number of comments. The disparate form of numeric data cannot be calculated directly, otherwise there will be problems like the final score will tend to the websites with higher order of magnitudes in its form of data.



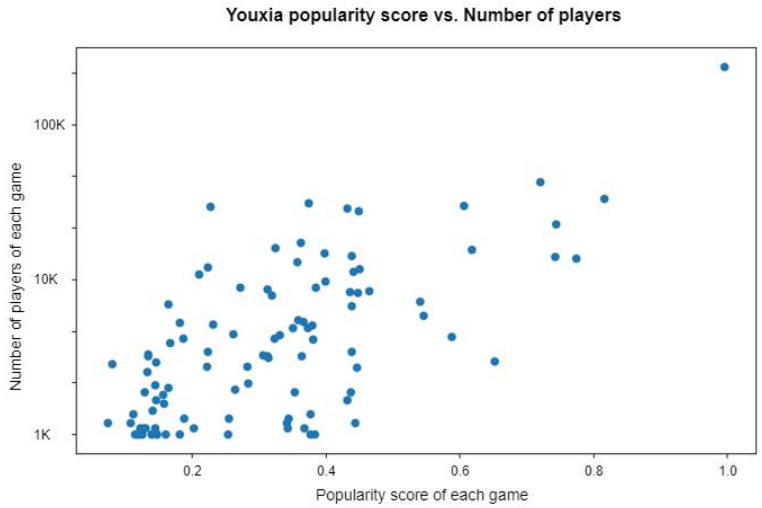
Graph 8. *By data normalization, we can integrate data from different sources together without introducing problems*

To avoid such problems, we applied data normalization to all websites' numeric data to a unified dimension. Specifically, apply the min-max normalization independently for each of our source website to normalize all the numeric data range into (0,1) as the popularity score of a game in each website. The closer a games score is to 1, the more popular it is inside the specific website. After such normalization, the attention received by a game in different websites can be mutually compared and further integrated together.

2.3.5 How to convince people that our system can reasonably weight each source's data and integrate them together to predict its number of players after release?

After we quantify the attention received by a game in a website and normalize different websites' data into dimensionless numbers within same value range, we thus define the normalized number as each source's popularity score. The next step is to integrate all the scores together to get an overall score for ranking all tracked games.

By observing the scatter diagram of each source's popularity score vs. number of players, and the correlation analysis result of the 2 variables, we noticed the obvious linear correlation between them, like a sample below from Youxia website:

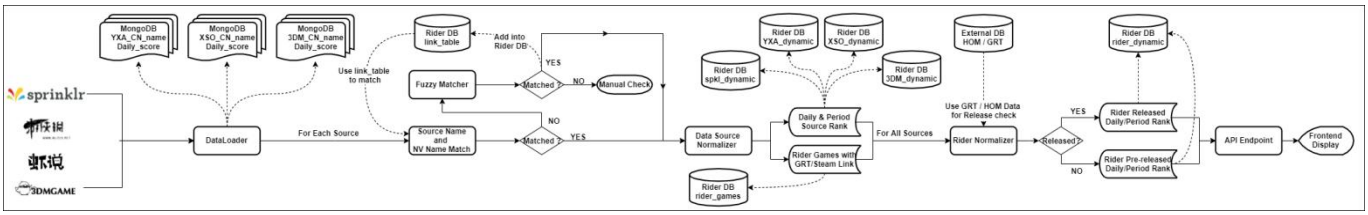


Graph 9. *Scatter diagram of 100+ games' normalized Youxia popularity score vs. The number of players which demonstrates a linear positive correlation*

So we decided to build a multiple linear regression model to integrate the popularity scores from all selected sources, each source's score will be treated as an independent variable, then we defined this aggregated score as the final popularity score. The goodness of fit of our regression model reached an adjusted R^2 score of 0.8036 which means about 80% of our data set is well explained by our model and the independent variables, which are the popularity scores from different sources.

2.3.6 The final system design and working pipeline

After answering above questions, we further contacted our users to demonstrate them the design logic of the system and successfully convinced the teams of the effectiveness of such design with all the data and analysis conclusions, the overall system architecture is constructed as below:



Graph 10. *Hype-O-Meter APAC pipeline and system design*

The fuzzy matcher mentioned in the graph is a naming matching system which provides English and Chinese name matching for game products, it helps us connect different translated Chinese game names of the same game product together to correctly integrate popularity scores from different websites.

2.4 Productization and interaction design

After proved the effectiveness of the system design, the next step is to find a appropriate way to present our solution to the users to make it an easy-to-use and human-friendly product.

As the system aimed at providing insights for better decision making, it should have frequent updates, easy access and provide convenient entries to other resources to make further operations within reach for users, thus the most suitable way of delivering it to the users is to serve it as a web tool, with a short and easy-to-remember website address.So the product design is determined as web page design which present the data produced by the system into visualizations and provide appropriate way of interaction.

We implemented the prototype of the system and continuously improved its UI design by multiple round of feedback collection from our customers, and finalized it as following:

2.4.1 Represent the results to users as intuitive as possible to serve their primary purpose.

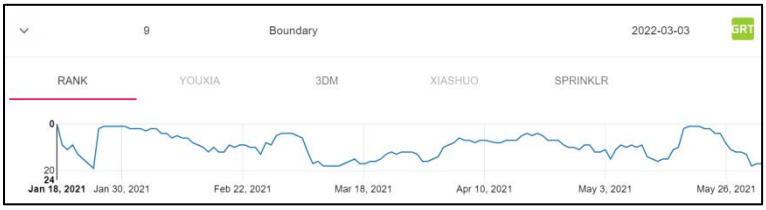
The most required feature of this system from our users is that they could easily find their target game which is popular among the players, and more specific, how popular they are and in which source it is popular, thus make further moves.

Rank	Name	Release date	Link Page	Sprinklr	Youxia	3DM	XiaShuo
1	FINAL FANTASY VII REMAKE	2022-06-01	GRT	#5	#5	#1	#6
2	Dyson Sphere Program	2022-02-02	GRT	#7	N/A	#3	#9
3	Boundary	2022-03-03	GRT	#6	N/A	#4	#10

Graph 11. *Intuitively display the results as organized table*

To clearly provide such information, the final rank list is directly displayed as a organized table in the main part of the web page, and the games' overall rank, name, release date, and their independent rank in each source are also displayed in each row, so that the users can know the overall popularity and its detailed distribution in each data source by a single glance.

2.4.2 Visualize the game's popularity trend for easier understanding



Graph 12. *Display the game's rank history in line chart to give a clearer view of the game's popularity trend*

Our users are keenly interested in a game's popularity trend, like when it received a growth of attention or when its popularity started to cease gradually. So we provided a historical line chart to demonstrate the changing of its popularity within each of the websites so that our users can get a clearer picture of the trending.

To make sure the main interface is clean and organized, we hide the line charts within a drop down area which can be unfolded by clicking the arrow icon at the beginning of each row. Besides, line charts for each sources are separately displayed in different tab within the drop down area so that the users can easily comparing between different websites.

2.4.3 Make it more convenient for high frequency operations.

By knowing the rank, users will high likely to go to the internal operation pages to make further moves like creating a report or submit a request regarding the target games they select, or navigate to the game website to get more information about the game itself.

Graph 13. *Each column of the ranks are clickable and will re-direct to the game's detail page in different websites*

For above high frequency operations, I attached corresponding hyperlinks for both internal operation page and external game websites for each game so that the users can directly click to navigate to their target websites in a new page.

2.4.4 Provide customized functions to serve different user groups.

Depending on different user groups, the users might want view the game rank in a different time range basing on the games' release dates.The tech marketing team usually want to see popular games which will be released after quite a few months, so that they can get in touch in advance to seek for opportunities of cooperation. while the optimization team

usually want to view the games which will be released in a shorter range to provide support for them in weeks before its retail.

Graph 14. *Date selector to narrow down the displayed games by filtering its release date*

Graph 15. *Release switch to switch the displayed games between released and not yet released*

So we provided date range selector along with a release switch to display the games that is within the interested date range of release date, so that different groups of users can use the tool as they wish to find their targets as quickly as possible.

2.4.5 Provide human-friendly common functions

Graph 16. *We provided useful common functions like search, column setting, data export and guidance reference link*

To make the tool more human-friendly, we also provided commonly used functions like page search, table column settings and data export. We also created detailed user guidance pages and made a reference link directly in the toolbar area of the product.

2.4.6 User-friendly localization for ease of naming comparison

Graph 17. *Hover tips to provide naming translations*

As the external game websites use translated Chinese names for the games, we provided hover tips to display the Chinese names for easier comparison and provide better user experiences for different region's customers.

2.4.7 Additional news letter service to conclude recent game trends succinctly and clearly

Top 20 most discussed Upcoming titles (To be released in 2021-09-14 ~ 2022-03-13)

Rank	Name	Chinese Name	Release Date
1	Diablo® II: Resurrected	暗黑破坏神 2 重制版	2021-09-23
2	Far Cry 6	孤岛惊魂 6	2021-10-07
3	Halo Infinite	光环：无限	2021-11-10
4	World War Z	僵尸世界大战	2021-12-26
5	FINAL FANTASY VII REMAKE	最终幻想 7 重制版	2021-12-01
6 New !	Ghostwire: Tokyo	幽灵线：东京	2022-01-11
7	New World	新世界	2021-09-28
8	Chinese Paladin 7	仙剑奇侠传 7	2021-12-31
9	Boundary	边境	2022-02-15
10	Deathloop	死亡循环	2021-09-14

Graph 18. *Newsletter service to conclude recent game trending*

To make sure our users can stay tuned for the recent game trends with minimum action required, in addition to the web tool, we also provided a news letter service which succinctly concluded recent changes in the game market with a weekly comparison of newly trending games. The news letter is automatically sent by weekly, and highlighted notable games with obvious marks.

3 The results of the project

3.1 The overall prediction accuracy and success stories

After the release of the project, we tracked 600+ games in the following year and the final prediction accuracy of the system are as below:

Confusion Matrix		Actual	
		Popular	Not-popular
Predicted	Popular	52	24
	Not-popular	9	536

Graph 19. *The overall prediction results of the system presents in confusion matrix*

Name	Release Date	HOM APAC Rank	The number of concurrent users
Cyberpunk 2077	10-Dec-20	1	1129602
NARAKA: BLADEPOINT	11-Aug-21	1	441246
Genshin Impact	15-Sep-20	1	423277
Minecraft	08-Dec-20	6	267058
Tale of Immortal	27-Jan-21	3	84686
Assassin's Creed Valhalla	10-Nov-20	1	66422
Horizon Zero Dawn Complete Edition	07-Aug-20	1	66234
Days Gone	18-May-21	10	53146
Death Stranding	14-Jul-20	2	45073
Call of Duty: Zeus (codename)	13-Nov-20	2	38395
Call of Duty: Black Ops Cold War	13-Nov-20	2	38395
Microsoft Flight Simulator	18-Aug-20	8	18929
Outriders	01-Apr-21	1	14472
NBA 2k21	04-Sep-20	4	12635
Mafia: Definitive Edition	24-Sep-20	10	12525
Watch Dogs: Legion	29-Oct-20	4	11773
Ghostrunner	27-Oct-20	8	10568

Graph 20. *Some of the popular games correctly predicted by the system*

We set the threshold of the number of concurrent players in APAC region to 5000 to divide all tracked games into popular and not-popular games, as usually a steam top 100 game will have around 7000 global concurrent players, and the results of the system is examined with real game data and proved effective enough as shown in above graphs.

3.2 Departmental recognition

After the release of the project, we successfully foresaw multiple games' regional popularity and the project is well recognized by the tech marketing and game optimization teams as it provides valuable insight of the game trending in the Chinese market, and it is now **adapted as a standard tool in the data-driven-decision-making process of the department**.