**Stats 401 final project report**

Team member： Jiyuan Wang, Luyao Wang, Peng Wang

**Introduction**

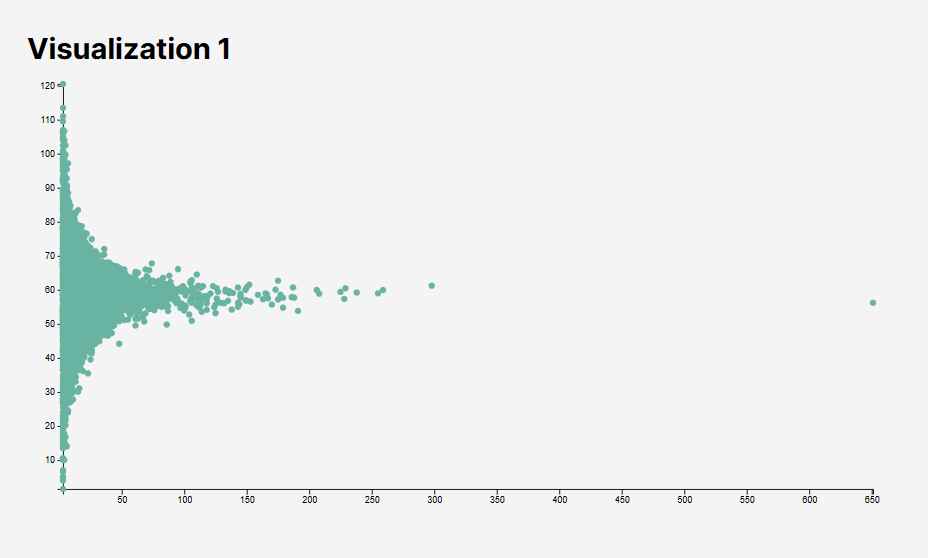
"Genshin Impact," developed by miHoYo, is a widely recognized action role-playing game that incorporates an in-game lottery system, often referred to as the "Gacha" mechanism. This system allows players to obtain characters and weapons. Originating from Japanese mobile games, the “Gacha” mechanism prompts players to use real or in-game currency for an opportunity to secure their preferred in-game items. Within the game, there are three distinct banners: “up” character, weapon, and regular banner, where players can spend a pull to make a draw. Our objective is to assess the actual probability of securing a character or a specific weapon and determine if it aligns with the odds provided by miHoYo. Additionally, we aim to understand character ownership and examine player tendencies. To present our findings, we have created an interactive website that showcases the “Gacha” history for individual players. Our final goal is to identify the usage frequency of each character during in-game activities.

**Method**

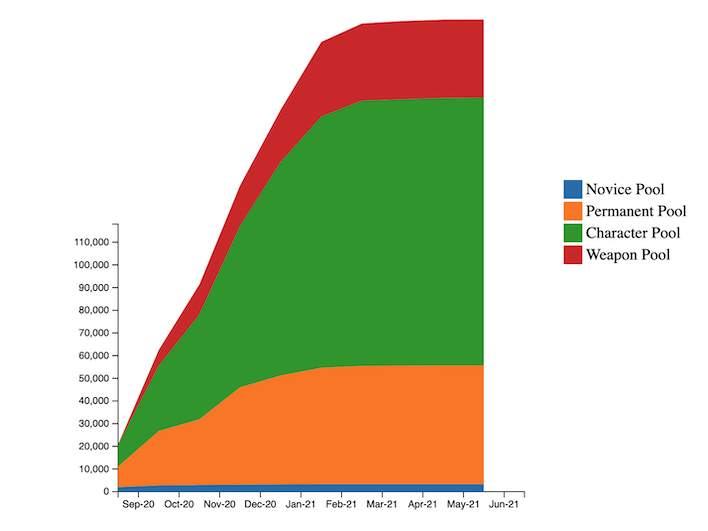
We began by sourcing an open-source dataset from GitHub. This dataset, a crowd-sourced compilation, boasts data from an impressive 15.59 million pulls, voluntarily shared by players. It provides an extensive view into player interactions with the "Gacha" system.Next, we extracted team composition statistics from a website. In "Genshin Impact", character combinations are crucial. Hence, data on team configurations is of utmost importance. These statistics shed light on frequently paired characters, unveiling both popular and potent team synergies.Lastly, we obtained the most recent "Gacha" records directly from miHoYo's Official Web API. By accessing real-time data, we can delve into the ever-changing preferences and behaviors of players. Python was employed to cleanse, filter, and analyze the datasets, ensuring discrepancies were addressed and outliers managed.Through an amalgamation of D3 coding, React application development, and front-end visualizations, we generated comprehensive visual representations of our analyses.

**Result**  
Visualization 1 displays the average number of pulls required to obtain a 5-star character, providing insight into the typical number of attempts needed to secure high-value characters. Visualization 2 is a time river chart that highlights the frequency and volume of recharge events, giving players a historical perspective on recharge activities. Visualization 3 showcases the total count of characters from open-source "Gacha" data with animations that trace the rise or decline of character availability over time. In Visualization 4, the history of "Gacha" pulls is chronicled, illustrating the average attempts for an "up" character; it allows users to switch between banners and access details of specific pulls, offering insights into pull trends and banner-specific probabilities. Lastly, Visualization 5 depicts the dynamics of team compositions in "Genshin Impact", where character icon sizes denote usage frequency, and lines between characters represent their co-occurrence in teams, aiding players in understanding character synergies, popularity, and team dynamics in the game.

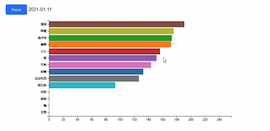
**Discussion**

****

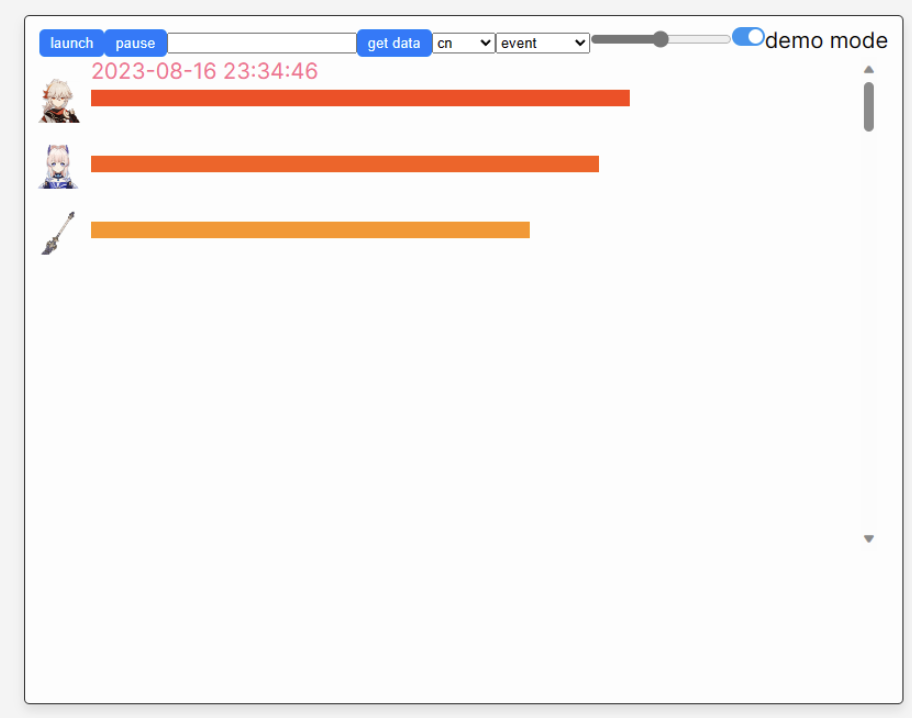
This visualization serves as a foundational metric for players to gauge their chances in the "Gacha" system. Through a scatter plot representation, players can discern the average number of pulls required to secure a coveted 5-star character. The data points clustered towards the lower end suggest that a majority of players can achieve this feat with a relatively lower number of pulls. However, the scatter towards the higher end also acts as a cautionary tale, indicating that luck varies, and some might need more attempts. This distribution can help set player expectations, manage resource allocation, and strategize their gameplay. By understanding the broader trends, players can make informed decisions, whether it's deciding to save up for future banners or spending in the current one, ultimately enhancing the in-game experience.



The visualization provides a detailed overview of recharge trends in "Genshin Impact" over a set timeframe. It employs stacked areas to depict fluctuations and growth in recharge volumes across four distinct pools: Novice, Permanent, Character, and Weapon. The X-Axis showcases the timeline, broken down by months, allowing for a granular view of activity. Meanwhile, the Y-Axis represents the accumulated recharge volume. Each pool's data is distinctly color-coded, enabling quick and easy differentiation. This choice of color scheme enhances clarity and aids in immediate visual recognition. By presenting this data in a comprehensive manner, the chart allows users to discern patterns, compare recharge behaviors across pools, and get insights into the game's overall recharge dynamics. Through its design and layout, the visualization serves as an invaluable tool for players and analysts looking to understand and dissect the game's in-depth recharge metrics.

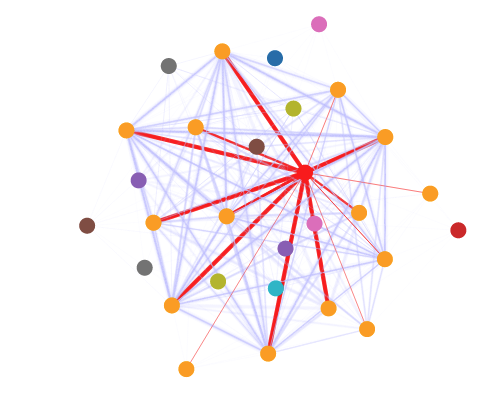
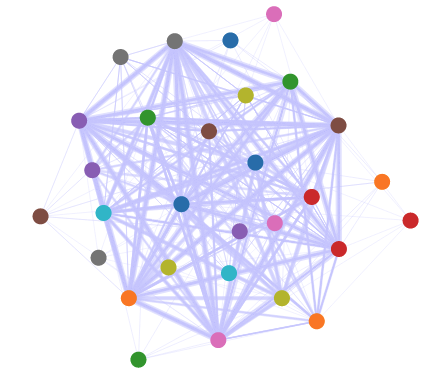


The depicted horizontal bar chart provides a comprehensive breakdown of character or item counts derived from open-source "Gacha" data. Each bar, distinctly colored, signifies a specific category, emphasizing the diversity in the game's characters or items. The X-axis offers a clear quantitative measure, likely indicating the frequency of each character or item being pulled, or perhaps their availability status within the game. The prominently displayed date, "2021-01-11," serves as a reference point, suggesting that this visual captures data from that specific day. Additionally, the presence of a "Pause" button introduces the idea of interactivity. This suggests that the chart could be part of a larger animated dataset, granting users the ability to traverse through time and observe the progression or regression of character counts. Such a dynamic feature is invaluable for tracking temporal trends. The design is both visually appealing and functional. Distinct colors for each bar ensure effortless differentiation, streamlining the user's experience in data interpretation and trend analysis.



The primary user task seems to be understanding the history of "Gacha" pulls, specifically focusing on the average attempts needed for an "up" character. By analyzing the visualization, users can discern the overall success and patterns of their pulls over time, allowing for more informed decision-making for future pulls.

Interactivity is indicated with options like "get data," "launch," and "pause." Additionally, the capability to switch between banners and access specific pull details implies a user-friendly interface. This allows for deeper dives into specific pull data and enhances the user experience by providing customized insights.



The presented visualization is a dynamic network graph representing relationships among various nodes or entities. Each node signifies an element, while the edges or links connecting them convey connections or associations. What distinguishes this visualization is its interactivity: when a user clicks on a node, it reveals all the links connected to that specific node and provides information about the importance or significance of these links in the console.Additionally, when users hover over a node, it displays the name or label associated with that node, offering an immediate way to identify and comprehend the entities within the network. This interactive approach empowers users to explore the network dynamically, gaining insights into the connections between nodes and the importance of these connections. It facilitates a deeper understanding of the underlying data and relationships within the network.However, the effectiveness of this visualization could vary based on the complexity and size of the network. Smaller networks may be easily navigated and understood through node-clicking and hovering, while larger, more intricate networks may become overwhelming to explore in this manner. Therefore, the utility of this visualization may depend on the specific context and objectives of the user.

**Conclusion**

Our study offers a comprehensive analysis of the "Gacha" system in "Genshin Impact" using a combination of extensive datasets, data processing techniques, and visualization methods. The insights generated emphasize the patterns of player interaction with the in-game lottery system, providing clarity on character availability trends, player tendencies, and the dynamics of team compositions.

**Roles**

Peng Wang: Figure2&5, report

Jiyuan Wang: Figure1&3, slides

Luyao Wang: Figure 4, dashboard.