RAINBOW SIX SEIGE: EXPLORATORY DATA ANALYSIS

USING UBISOFT VELVET SHELL DATA SET

ABSTRACT

R6 is a widely played multiplayer shooting video game in which players take on the roles of attackers or defenders. Using the R6 velvet shell data set, we're attempting to uncover hidden trends and patterns in operators, maps, game types, and other variables in order to identify the greatest and worst operators, as well as the finest weapons and attachments. This in-depth understanding would also aid gamers and game developers in comprehending the operator's and map's strengths and limitations.

BHARATHPRASAD SIVAPRASAD

BS330

Bs330@student.le.ac.uk

Table of Contents

BHARATHPRASAD SIVAPRASAD
BS330
Bs330@student.le.ac.uk

PROJECT CONTEXT

Rainbow Six Siege is a tactical shooter with a focus on meticulous planning, collaboration, and finely calibrated tactical play. Since its initial release in 2015, the game has grown significantly, with additional levels, operators, weapons, and seasonal events. PvE and PvP are the two main cornerstones of the game. PvE is divided into two modes: 'Scenarios,' in which you improve your abilities in live-fire drills, and 'Terrorist Hunt,' in which you team up with friends to neutralise AI-controlled targets and defuse high-threat situations. PvP is based on tactical 5-v-5 team play and is used in a variety of situations in which the Red and Blue Teams compete to defuse bombs, rescue hostages, or secure important areas. (Playstation, n.d.)

For this particular project, we have considered to analyse the player vs player (PvP) mode. The game's principle is simple: two teams of five players compete in each game, one as the Attackers and the other as the Defenders. The Attackers, as the name implies, is tasked with breaching, clearing, and securing the target. Their gadgets, which include a ready supply of ballistics, reconnaissance equipment, and crowd control weaponry, are tailored to this effect. Defenders must keep the objective by fortifying critical places throughout the map and reinforcing the building's entrances. Armour and medic abilities to keep allies standing, as well as intel gathering and denial devices to prevent enemy reconnaissance, are among their gadgets. Depending on the game option, attackers are given a goal, which might range from rescuing a prisoner to planting and detonating a bomb in a structure. Each game consists of a best-of-five, or seven games played on a random map, with each round concluding when the Attackers' objective is fulfilled within a timeframe, time runs out, or all members of one side are eliminated. (Ubisoft, n.d.)

The main purpose of this project is to help both the players and the developers to find useful insights. The players may find a particular operator to be more useful for a particular game mode. They can choose which operator is more powerful and agile when compared to other operators for a particular map. Maps in R6 are diverse and different in every aspect. Therefore, there is a chance for a quick operator to perform better in certain maps which have similar properties. We can compare the operators based on their kill/death ratio. A newbie player into the world of R6 can use this data analysis project to find which operators are mostly used by the more experienced players in the R6 community. So, by using this project, Newbie players can play the game much more easily and become a professional in less time. Developers may also use this project paper to improve the game and the game’s overall gameplay. For example, they can check whether some operators are overpowered compared to other operators by comparing each one kills to death ratio or whether the real-life players are picking a particular operator much more often. Developers can also use the project to improve a particular operator which seems to be picked less often. In R6 some maps favour Attackers while some favour Defenders. This creates an imbalance in power for each map and makes it difficult for players. For example, for a defender team, it is easy to win games involving hostage rescue operations mode. This is because of the fact that the Attackers have to kill Defenders along with the objective to protect and rescue the hostages. So, keeping the hostage alive brings burden to the Attackers. From a different point of view, this difficulty turns to be exciting for some but not much for others. So, the developers can choose to keep the difficulty or alter the map attributes in a way to balance the mode for both Attackers and Defenders. For a player, who likes to play as an Attacker in Hostage Rescue operations, can choose the best combo of operators to carry out the operation. So, the scope of R6 data analysis project is extensive and useful for both players and developers.

BHARATHPRASAD SIVAPRASAD
BS330
Bs330@student.le.ac.uk

PROPOSED QUESTIONS

1. Who are the best operators for both Attacker and Defender roles in each map?

   Players can select the best operators in each map and play better by identifying who they are. Because of their experience accumulated from many rounds, professional and experienced players tend to select specific operators for each map. However, inexperienced players may lack the knowledge necessary to select the appropriate operators based on the map. We aim to make things easier for them by providing them with accurate information.

2. Which are the best attachments for primary weapons?

   For each operator, certain attachments tend to offer better results. By attaching accessories to each weapon, attributes such as precision, accuracy, and recoil can be increased. Although these attachments may be a matter of personal preference for each player, they usually create similar game results. We strive to figure out which attachments are the best for each weapon and advise gamers on which to utilise while playing.

3. What are the main reasons for victory?

   There could be a major reason for the triumph in each round. It might be because the attackers defused the bomb or eliminated all of the opponents. In each map, we aim to uncover the best reasons to win. This analysis's useful case is to guide players on how to choose the best victory condition. For example, in the bomb game mode, for a specific level, it may be more effective to go for the kill rather than attempting to diffuse the bomb first, possibly due to map design. We're trying to figure out if there's a link between the win condition and the map.

4. Are there any operators who are incomparably better than the others?

   The developers may benefit more from this analysis than the players. Because if we can uncover any operators which exceptionally better than the others, the developers can change the design of that particular operator to balance the whole gameplay. We may or may not detect any irregularities in the operator design in this hypothetical problem.

SOURCES OF DATA

Geoffroy Mouret, a Ubisoft R6S Development Team Game Intelligence Analyst, shared datasets containing various metrics from in-game data on June 6, 2017. The dataset is a data dump from R6 Year 2 Season 1 game which is "Operation Velvet Shell" (Ubisoft, n.d.). It can be accessed from https://www.ubisoft.com/en-us/game/rainbow-six/siege/news-updates/2fQ8bGRr6SlS7B4u5jpVt1/introduction-to-the-data-peek-velvet-shell-statistics.

Along with this data, he disclosed that Ubisoft analysts collaborate with a variety of teams and play an essential role in the game's overall development. The Development Team chose to disclose certain in-game data as a method to encourage members of the community to do their own data analysis, based on the intriguing numbers they discovered.

Three datasets are given in the official Ubisoft website. The analysts can use any of the three datasets according to their expertise in data science. Below is the name of the file, the expertise level required, and the website link.

1. dataDump_s5_summary_operator_loadout.csv

    - Expertise level required – Amateur
    - Size of the file – 39MB
    - Major attributes – Platform, skill rank, operators and role, primary and secondary weapon, wins, kills, deaths, picks
    - Link address - http://static2.cdn.ubi.com/pxm/RainbowSix/Data_Dump/dataDump_s5_summary_operator_loadout.zip

2. dataDump_S5_summary_objectives.csv

    - Expertise level required – Professional
    - Size of the file – 767MB
    - Major Attributes – Platform, Game mode, Maps, Skill rank, Operators and role, wins, kills, deaths, picks etc
    - Link address - http://static2.cdn.ubi.com/pxm/RainbowSix/Data_Dump/datadump_S5_summary_objectives.zip

3. dataDump_S5.csv

    - Expertise level required – Expert
    - Size of the file – 19.3 GB
    - Major attributes – Platform, game modes and maps, win role, reason for victory, operators and roles, weapons, weapon attachments, kills, deaths, picks, wins etc
    - Link address - http://static2.cdn.ubi.com/pxm/RainbowSix/Data_Dump/datadump_S5.zip

We will be utilizing all the three datasets for this project according to the complexity of the question we tend to answer. From the initial analysis of the datasets, the dataDump_S5.csv tends to be extensive and contains much more useful data. But because of the reason that the file is huge, we might only be using this data set for very complex analysis. For simpler analysis queries, it would be better to use the other two datasets as they are smaller in size and process faster to yield results.

PROPOSED TOOLS

We want to complete three critical processes in this project: extract and load the data, analyse the data (which includes data cleansing and EDA), and visualise the results using statistical graphs and/or dashboards. A database, such as MySQL, or a data processing framework, such as Apache Spark, can be used for the first step (extract and load the data). The dataset's size (which is 19GB) necessitates a lot of processing power, thus it can't be used directly. Apache Spark can analyse big data sets quickly and distribute data processing jobs across numerous computers, either alone or in conjunction with other distributed computing tools. The speed of Apache Spark is an advantage over any other solution. Spark also supports APIs in Scala, Java, and Python, with support for other languages such as R planned in future versions. (Apache Software Foundation, n.d.)

Python is the programming language we use to process the data. We may either shrink the enormous dataset to a manageable size and process it locally with Jupyter Notebook, or we can use the PySpark API (which combines Python API within Apache Spark) to perform comparable tasks. We use built-in data libraries like pandas to process the data, and we can also generate basic statistical graphs using ggplot2 or matplotlib, which are Python data visualisation frameworks. (Selvaraj, 2020)

The third and final stage is to visualise data and, if necessary, develop dashboards. Microsoft PowerBI, a free data visualisation product aimed largely at business intelligence, can be used to accomplish this. Histograms, pie charts, heat maps, box plots etc can be used to visualise the comparison between several attributes.

References

Apache Software Foundation, n.d. *Overview - Spark 3.2.1 Documentation.* [Online]
Available at: https://spark.apache.org/docs/latest/

Playstation, n.d. *Tom Clancy's Rainbow Six Seige - Beginner's Guide.* [Online]
Available at: https://www.playstation.com/en-gb/games/tom-clancys-rainbow-six-siege/a-beginners-guide-to-rainbow-six-siege/

Selvaraj, N., 2020. *A Beginner's Guide to Data Analysis in Python.* [Online]
Available at: https://towardsdatascience.com/a-beginners-guide-to-data-analysis-in-python-188706df5447

Ubisoft, n.d. *INTRODUCTION TO THE DATA PEEK: VELVET SHELL STATISTICS.* [Online]
Available at: https://www.ubisoft.com/en-us/game/rainbow-six/siege/news-updates/2fQ8bGRr6SlS7B4u5jpVt1/introduction-to-the-data-peek-velvet-shell-statistics

Ubisoft, n.d. *Learn | Tom Clancy's Rainbow Six Seige | Ubisoft(US).* [Online]
Available at: https://www.ubisoft.com/en-us/game/rainbow-six/siege/game-info