Dish M. Eldishnawy

Follow 16 Followers About

Data Analytics for Gaming: A Walk-through

Dish M. Eldishnawy May 7, 2019 · 9 min read

Background

Let me start first with a little background that I think is quite relevant. So I love gaming, Console (obviously Sony Playstation), or Mobile Phone (iPhone as a choice not obvious anymore). But other than that, I am an engineer and a data analyst, and just recently I had a company in data enrichment where I worked with leading Market insights companies such as Ipsos. In my career, I was lucky as well to work closely with Facebook in a partner company helping to scale and optimize the online marketing of leading brands & companies such as Zalando, Skyscanner, Uber, Despegar, Tinder, and many others who spend 800k+ EUR a month on online ads, including some gaming giants, so you can imagine I built a great knowledge about that field from a business perspective, but also technical & analytical.

I decided one day to combine all these together and learn how properly games are created, scaled, marketed (that I knew much already), and above all monitored and analysed. In this article (or properly series), I will share with you some insights.

Data is data, no matter what is the source

Reality is, In the past Gaming has had success before introducing game analytics. With the rise of mobile apps, the amount of possible gathered data increased massively, it's a firehose, compared to the console which dominated the past. A game is technically a mobile app or a console app. That said when it comes to data, like any other app or online source you can mine/gather data from instances/actions a user do, or even

or using an API service. Data is data at the end. Next step would be using this data to analyze and understand your service. And when it comes to gaming it's about

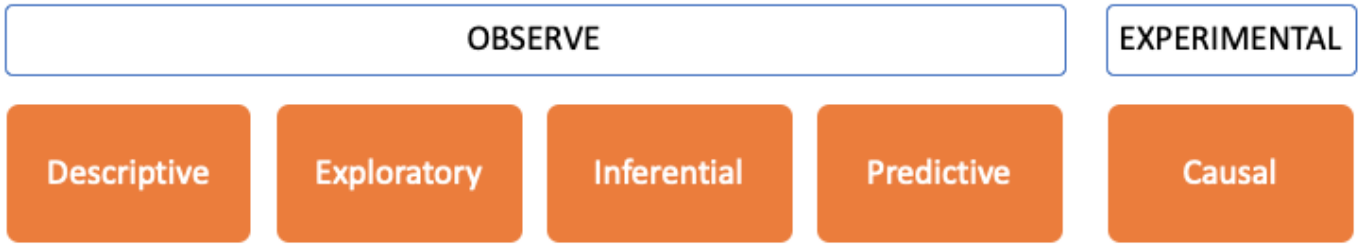
"Using data to understand your players and monetizers"

So in simple words, it's moving from shipping your game, and maybe it becomes a success, or maybe it doesn't, to knowing how your players engage with your game and using an influx of data to optimise your game in all aspects.

In the modern world, it's common nowadays that gaming companies know they need to use analytics, they instrument the events and track the actions, but integrating data into the workflow is surprisingly underused. While using external sources of data (ex. YouTube video comments, Tweets) are rarely integrated or considered by the different teams in a gaming company, or at least, it is not a common practice.

Data Analysis Types

Generally speaking, there are five different types of analysis, that knowledge doesn't really come from the Gaming industry, but from any data-intensive industry that utilises data analysis. And when I started exploring analytics for Gaming, I found that all of the different analytics still fall into these five types. The below diagram shows these five different types.



Types Of Data Analysis

You can summarise though any activity you do as either observing the data, or experimenting the data. To list some of the analysis you can do:

- Quantitive analysis (describing the data with key numbers)
- Visualization & directionality by finding correlation and exploring the data
- Use regression models (for instance Chi-Squared) and working on a sample of the data.
- Utilise the current events in predicting future events using Machine Learning.
- A/B Testing and measuring what happens to a certain variable when you change another one.

If all of that sounds quite generic, in the coming section we will discuss examples of how to apply any of these points in the Gaming industry.

Analysis in Gaming

While A/B testing can be of great help, it is not recommended to do it intensively on a game that is just launching with a limited budget, not to add that you don't really have enough data to justify experimenting with the data. As such, sticking with observing the data is the natural way to start. The easiest way is to focus on the first three types: Descriptive, Exploratory, and Inferential analysis. Which leaves out Predictive analysis that can be more technical, complicated, and data heavy. Although it is easy to fall into following the trend and kick-off with predictive analysis and predicting the future with Machine Learning or even Deep Learning, that is not the wisest thing to do, since you should spend your resources on the first three types that are indeed easy to do, yet very important for your company. So let's see some examples of these three

Descriptive Analysis

In simple words, that is to quantitatively describe your game data. A common place to start is "averages", by looking into the average revenue per user for instance. A better

business terms that means you will be able to understand more nuanced behaviour towards your data/game.

Now let's see an example. Boosts are a common tool used to generate revenue in games, where people pay money to earn a certain amount of tokens to use as boosts in the game, expecting to score more points by using a boost. The below diagram is a hypothetical analysis of the score percentiles (so 5 numbers summary in action) per Boost used in a game.

Score Percentiles by Boost Used

Just by looking into the diagram we can realise that among the three hypothetical boosts in this game, indeed 50% of the players experience a lift in their score when using a boost (a behaviour needed and good to monitor how much that lift is). An interesting finding as well can be at the high-end, where the hardcore players are that are really good at the game. The got a much higher lift when they used specifically Boost Y and Boost Z. So in simple analysis, this tells us that our hardcore players perform the best

boosts. If not, even consider in the marketing plan. In another analysis, a boost might show an overwhelming boost in the score, which means this boost needs to be revisited whether pricing-wise or fine-tuning it's effect on the gameplay.

As you can see just by simply using percentiles you can figure out pricing and targeting strategic changes that are needed to be applied.

Exploratory Analysis

It is all about finding relationships in your data that you don't necessarily know about before. Opposite to Descriptive analysis where numbers could be enough to realise findings from the data, visualizations are a key tool in an exploratory analysis. Below is yet another hypothetical histogram to show how you can explore your game data.

Monetizers by Amount Spent/Player — Revenue by Amount Spent/Player

The blue histogram shows the count of people divided by how much they spent on the game and the yellow histogram shows the same exact division, but instead of the count, it shows how much of revenue this category of players -added up- generate.

As you can see, by just looking at the visualization you can "explore" that 10% of your customers generate more than 50% of your revenue, which is a usual behaviour in a freemium game model for instance. Since this article is not specifically talking about the different types of monetizers, here is an interesting article about the top spending players "Whales" behaviour that I recommend you to read.

The message here is that exploring the data can inform you about your game, the players, and business model, and spot risk behaviour or players types to keep an eye for.

you discover statistical information about your game. It gets more into statistics and mathematics from here, and it is common to use regression models at that stage. Since you will be trying to fit a model to your data, finetuning is quite possible as well. Inferential analysis can be a good start for a gaming company before exploring the predictive analysis, where Machine Learning is more used. It is important to know though that sometimes Inferential analysis and a simple linear regression model can be enough to understand your data and infer the future, as such, a common mistake in data analysis is forcing the use of complicated ML models on what could be a much simpler case. As such wasting resources, time, and in some cases even overfitting your models. Always start by Inferential analysis.

In this section, I would like to show an example of using Inferential analysis on Retention data. By having some retention datapoints (4 in the below diagram), we can project that over time. Let's imagine that this data is coming from data gathered about a certain Cohort in our game. A cohort is a group of players that share certain characteristics within them. Below diagram shows the regression projection for 180 days.

Actual Vs Projected Retention over 180 days

how you are going to hold your players in the future, and what is projected for their retention and LTV value.

So where to start?

Now all that sounds cool, but can be overwhelming as well. So from where should I start? The key questions you should ask yourself is:

1. How many people have installed my game?
2. How many people are playing my game?
3. How much money is my game generating?

These three questions are the baby steps into game analytics. And although they sound simple, they are in the core of your game business: Acquisition, Engagement, and Monetization. That is the basics of GaaS (Game as a Service). Once you figure out or spot a problem in one of these three questions, you can then dig into more and utilise the data for your analysis.

Conclusion

It is important to keep in mind that game analysis is not a fishing exhibition, or in gaming terms, an open world side missions. Once you have a large amount of data, it is quite natural to start seeing patterns. Game analysis, or in fact any analysis should work the other way around, as such, start by exploring your current situation of acquisition, engagement and monetization and then buildup your analysis. Any analysis you do must be tied to your business goals, and support or trigger an action on your game. This way analytics will be integral in the game building and maintaining process, rather than having a role similar to accounting :). Understand the difference as well between **reporting** and **analysis**. Reporting is listing numbers: DAU (daily active users), Revenue, Conversion, ARPPU (Active Revenue Per Paying User). While analysis is about storytelling So instead of listing numbers, understand their meaning. Let's give a hypothetical example of that:

Report:

DAU= 1 Million , Conversion= 0,5% , ARPPU=6 EUR , Revenue=50k EUR

media. This was caused by a 4% decline in the **DAU**, which in turn **declined** the number of **monetizers** by 3%. The players that continued to spend kept the same **ARPPU** of 6 EUR

As you can see, the same numbers and terms are used, but a data analyst job is to tell the story of the data in hand. Making the findings actionable and meaningful business wise.

One final advice, try to keep your eyes open for 3rd party data and external sources. Mining data from YouTube, and making some NLP (Natural Language Processing) on YouTube or Reddit comments can result in finding great insights about your game. For instance, which character people prefer using, or what users hate about my game. Consider utilising external data when building a story out of your analysis.

Feel free to leave any question below, the intention of this article is mostly sharing thoughts and ideas, some of them come in fact from outside the gaming industry, but I believe fits within gaming. Have fun playing with data :)

Data Science Gaming Data Analysis