**Introduction**

The effect of the global pandemic of COVID-19 has altered a great deal of cultures internationally and the lifestyles of everyone worldwide. Changes brought about by recommendations made by health professionals and quarantines instituted by governments has drastically altered the ways we interact with one another as we work and play. One particular belief is that the pandemic has caused an increase in the playing of video games, especially in conjunction with movements such as #PlayApartTogether along with various activities and rewards by many games across the world to encourage the social distancing recommended to slow and halt the pandemic.

I have seen this very effect in my own home where my wife – who is adamantly not a gamer – has been drawn into frequent playing of the Sims on her phone as a method of dealing with the extra time in seclusion. Over the course of my lifetime, I have watched serious video game playing evolve from a “geeky” pursuit to a mainstream competitive event as culture and society has embraced the products of the information and data age. While this gradual shift in perception over time has done much to improve the acceptance of video gaming as a vibrant hobby of many, the effect of such a major event as this pandemic pushing increasingly new numbers toward this leisure activity may very well dwarf the previously observed growth.

How much has actually changed with video game playing, however? Initially using data provided by the SteamSpy API, I plan to compare the actual numbers produced by Valve Corporation’s digital gaming service of Steam to learn what differences and changes have occurred in this year compared to the results over time – in particular, since the data’s beginning starting in March 2009. I will be comparing such things as game playing time, purchases, numbers of owners, and the like to see what differences can be seen over time to test the hypothesis that there actually has been an increase in the playing of video games during the course of 2020 compared to prior years.

My primary documentation, as mentioned above, will the SteamSpy API, using the referenced information presented from here to pull the above referenced data: <https://steamspy.com/api.php>. While this is my planned initial data source, I do foresee the possibility, if not probability, of expanding into other sources of data in order to provide a more complete test of the hypothesis through other platforms, such as GOG and console games; this will primarily depend on how much the data for the Steam platform can present a clear picture concerning the results and overall conditions of gaming play has changed in this first half of the year.

**Preliminary Requirement**

As mentioned above, the primary plan for the data is to pull it from the SteamSpy API – this will be done through the use of Python with the urllib and requests modules. From there I plan to perform statistical tests such as logistical regression, *z*-tests for the difference between two population means, two-sample *t* tests, and multifactor analysis of variance to determine if the hypothesis that the pandemic has caused an increase in the amount of video game playing is accurate within a reasonable confidence interval. In order to help improve the understanding of the data, data visualizations based on time – such as line and bar charts – will also be implemented to help demonstrate the differences for all audiences, not only those that have an understanding of data analysis or statistics. Additionally, highlighting the data for the year 2020 in these visualizations will also help to bring attention to the particular changes which have occurred.

While the initial data is fairly straightforward, metadata will need to be implemented to help expand the data’s information and clarity. For example, the differences between the totals off all the data from March 2009 to the present and the totals of the last couple of months will help to provide an overview of the gaming data for the past few years prior to the latest developments. Data provided by additional sources will help to clarify this information further.

Since the data provided by the SteamSpy API is a dynamic, live extraction, multiple pulls of that data and/or the acquiring of historical data pulls provided by other analysts publicly online will likely also be required. Thus, this data will need to be prepared in a clear manner that differentiates the time factors of when the data was received in order to properly display the chronology of the information for comparisons.

Testing and evaluation will be performed on the basis of the individual months in 2020 to the same months in years past in order to minimize any additional factors that could account for increased variability in the data, such as holidays or seasonal sales. In addition, all such evaluations will be performed on a basis of the either same game or genre being compared against itself, or as a comparison of all games in the specified time periods against themselves. Additional research may be required to determine if any unique factors – such as the release of a highly anticipated game – has any effect on the results at specific periods of time as well as methods to counteract those issues.

**Expected Results**

By implementing the above goals and practices to reduce or eliminate bias and outliers, I hope to provide the most accuracy possible to determine the veracity of the hypothesis. While I will be attempting to maintain a level of skepticism and detachment from the data, I believe that the null hypothesis – which states that there are no variations in the sample means between the data collected for 2020 and the data collected for years prior – will be rejected in favor of the alternate hypothesis, which states that there has been an increase in gaming in 2020 compared to past years, indicated by an increase of time played, purchases, and owners of video games through the Steam platform.

**Execution and Management of Project**

The purpose of this project is to get a glimpse of the effect which the coronavirus has had on the mannerisms of the population of the world beyond the blatantly obvious effects of health, social interaction, and overall digital presence. As mentioned previously, the perception of gaming as a pastime has grown – along with many other “geeky” pursuits such as comics, roleplaying games, and the pursuit of mathematics and science – from one which was only enjoyed by social pariahs to achieving fame and interest on par with that of professional sports.

The primary risks for this project are the influences of variables which have not been accounted for in the evaluations and biases inherent in the data. As mentioned previously, I plan to use research and same-month comparisons to counteract any unknown or additional variables. As for biases in the data, due to the fact that the information pulled will be pure data which is not affected by collection processes which are directed by perceptual biases. My hope is that this minimization of these issues will produce the best accuracy.

The influence of this current time of international crisis, I believe, has increased this perception greatly. Therefore, the execution of this project is geared toward increasing the understanding of the influence that gaming has had upon the world, especially due to the fact that many more are making use of this leisure activity and interacting with a new digital world – a world that has long accepted differences of race and culture through the application of diversity and variety (which even goes so far as to add green, blue, red, orange, purple, and more to the binary skin colors of black and white) and encourages thought and introspection through the use of puzzles of both an intellectual and philosophical nature. With encouragement around the world to #PlayApartTogether, societies around the world have begun to connect in ways that they never have before by erasing the lines that have divided us historically.