

STATS5029P: Statistics Project and Dissertation

**#34: Does playing Pokémon Go increase physical activity?**

Steven Kwok

18/11/2021

# Content

**0. Abstract1**

**1. Introduction1**

**2. Literature Review 2**

2.1 Background Information2

2.2 Research Objectives3

**3.** **Methodology 4**

3.1 Data Description and Processing4

3.2 Model Selection9

**4. Analysis Results13**

**5. Discover and Discussion 16**

**6. Conclusion and Limitations18**

**7. References19**

**8. Appendix22**

# 0. Abstract

Pokémon Go is a well-known augmented mobile game, seems to be useful for increasing physical activity among users. However, some have argued that the effects of Pokémon Go on the amount of sport engage in are not direct or stable. In this research, we examine the relationship between Pokémon Go application usage and the amount of physical activity of users, as well as the players’ preferences. We also identify factors that affect the level of physical activity, and the factors associated with participants' attitudes towards sports. The whole study was processed using a polynomial linear regression model, scatter plots and box plots. Finally, we discuss interpretations and causes of those findings, as well as provided suggestions for future research and implications for the public.

# 1. Introduction

Pokémon Go is a globally popular augmented reality (AR) mobile game that affects the behaviour of a vast number of players. In accordance with arguments presented by several researchers (like Gunther (2016)), I hypothesized that Pokémon Go could increase levels of physical activity. To test this, I used a polynomial linear model to study the effects of Pokémon Go on physical activity. I considered several factors, including the amount of app usage and; the age, education level, and attitude towards sports of 981 participants in America. Based on this research, above hypothesis was not supported: players preferred participating in an app-related activity instead of general physical activity. Attitude towards physical activity was a key factor contributing to the negative effect of Pokémon Go on levels of physical activity. The results of the analysis and implications for the public and future research are discussed in the following sections of the report.

The study was completed in the following steps: First, we present some background information on sports, public health, AR technology and Pokémon Go. We then study the effects of Pokémon Go on levels of physical activity, considering several aspects: Frequency of app usage, the participation preferences of players; other factors related to the amount of physical activity and the effects of the attitudes towards physical activity (with the discovery of factors related to participants' attitudes). Before conducting the data analysis, we grouped highly-correlated variables by the mean of each record. The entire analysis was conducted using a polynomial linear regression model. We applied stepwise selection method (setting [AIC] as the variable selection criteria) to obtain the best model. After presenting our observations, we discuss the causes and meaning of the results, provide suggestions to the public, and identified the limitations of the study and avenues for future.

# 2. Literature Review

## 2.1 Background Information

Sports have been proven to improve the immunity system while training the muscles and respiratory system (Ornulf Seippel. 2006). However, people living in urbanized societies are prone to living sedentary lives. By Buraimo et al. (2011), approximately half of those surveyed did not participate in any kind of sports for various reasons, including job characteristics, changes in entertainment preferences, and long working hours. With lower levels of physical activity, the risk of obesity, depression, high blood pressure, as well as a series of cardiovascular diseases rise rapidly, which is harmful for public health. Quinn (2021) found that a positive attitude can encourage people to engage in sports more frequently. Augmented reality video games, which are perceived as a more interesting ways of exercising, have been considered to solve health issues. Augmented reality is an extended version of virtual reality (VR). It uses visual devices, including eyeglasses, monitors, and smart devices, to combine virtual elements with the real world (Fisher. 2021). By overlapping on and tracking real-world objects, AR objects seem to occupy the same space. Besides visualization, an AR system can also include sounds and tactile elements, providing a new version of the world. Augmented reality technology can be applied in various types of applications, such as maps and games, and mobile games are particularly common and attractive to users. In this study, we focus on Pokémon Go, a popular AR mobile game.

Pokémon Go, developed by Niantic Inc., is a well-known mobile game that was released in 2016 on both Apple Store and Google Play (Reilly, 2017). It combined the application with GPS signals and real-world map. Players of Pokémon Go can locate, catch (also obtain ingredients for training Pokémon), hatch (players walk around 2 to 10 km to obtain a Pokémon), and train virtual creatures, Pokémon. Those Pokémon can be used for battling and gym controlling (Webster, 2015). Pokémon Go uses a map and camera to display the virtual spots, such as PokeStops, gyms, and activity location points. (Smith, 2017). Pokémon Go is a celebrated application, with 632 million downloads and 147 million monthly active players. Due to its popularity and attractiveness to users, it is plausible that Pokémon Go facilitates behavioural changes that might affect public health (Dillet, 2016). If a correlation is found between Pokémon Go and users’ levels of physical activity, this will represent a new way of increasing engagement in sports, and fostering positive attitudes towards physical activity.

Multiple pieces of research have found that Pokémon Go affects levels physical activity of most participants (Gunther, 2016; Wong, 2017). One study found that users’ level of physical activities increased by approximately 25%, and physical activity levels increased across genders, ages, and weight statuses. In general, physical activity levels of players who were inactive originally increase sharply (Gunther, 2016). Some studies, however, suggested that Pokémon Go dose not directly improve the physical health. One study found that despite positive effects in the first period, players' physical activity levels drop sharply, indicating that the positive effects are not sustainable (Allana et al., 2016). This may be related to players’ motivation. Another study identified three kinds of motivation among players: health, social, and immersion. Although players motivated by health showed a significant increase in physical activity levels, the effects of social motivation and immersion were limited. However, time spent outdoors increased significantly (Lukas et al., 2017). Based on the above arguments, I hypothesize that Pokémon Go positively, but indirectly, affects physical activity levels. To test this assumption, we have conducted a study on the relationships between Pokémon Go and physical activity. If Pokémon Go does increase levels of physical activity, more AR mobile games can be developed to improve public health. The objectives are presented in the "Research Objectives" section.

## 2.2 Research Objectives

Pokémon Go was not designed for public health improvement originally. That is the reason for the confusion about the relation between Pokémon Go and the amount of physical activity. For truth discovery, the study was conducted based on four aspects: The relation between frequency of app usage and amount of physical activity, firstly, is the main focus. Theoretically, the more the app usage, the higher the opportunity for players to do physical activity as most of the app-related activity required walking (e.g. catching Pokémon, or turning Pokestop). Following the application usage aspect, we discussed the problem in players' characteristics. Since some researchers, like Alessandro (2017) argued that Pokémon Go players tend to join the game-related physical activity, instead of physical activity in general. The aim is to discover the existence of relations between Pokémon Go players and the amount of general physical activity; The level of physical activity can be affected by various factors, including motivation, education level, and gender. We, thus, wanted to locate variables associated with the number of physical activities. Last but not least, we examined the effects of the attitude towards physical activity caused by gender or educational level, for explaining the effects on the amount of physical activity of attitude towards physical activity.

# 3. Methodology

# 3.1 Data Description and Processing

The data was obtained from a study in America, obeying the code of ethics of the world medical association (Declaration of Helsinki) for studies using humans as data. Amazon Mechanical Turk (MTurk), an internet-based platform offering an online participant pool, was applied for data collection (Buhrmester et al., 2011; Paolacci & Chandler, 2014). The original data contains 999 records, described by 31 variables. Before processing data, we examine the number of missing values (which is 0), as well as filter out records by an attention filtering variable.

A variable for filtering out non-focus participants as the Mechanical Turk experiment was applied (removed after filtering). Despite the convenience as well as limitless of time and location, the Mechanical Turk experiment cannot guarantee that participants are paying attention as the survey was completed online (Jennifer Jacquet, 2011). Ensuring only data from focusing-on-survey participants were collected, a variable that acted as an attention filter was used. If failed choosing "Disagree" in this question, the records will be removed due to being classified as non-focus records. The number of remaining records were 981. After primary data cleaning, we transformed all columns into integer scores, according to the level of each variable (mentioned in the following paragraph). This subjective assigning method is plausible for applying interval scale and the concept of distance (Chaowei Yang, 2014). The identification number was just used for representing the number; the surveying date contains no information as just recording date of survey submission; IP address of a computer did not affect participants’ behaviour; the last variable, representing participants’ behaviour, was marked cannot be used. Thus, twenty-eight variables were applied for further grouping and analysis.

Despite age (discrete variable) and gender (nominal variable), all the others are ordinal variables. Both "Frequency of App Usage" and "How often sharing on social media" are ordinal data, anchored with the scale from 1 = "never" to 7 = "very often". Former accessing the extent of players using Pokémon Go per month, while the latter demonstrates the frequency players share their achievements on social media. Those variables mentioned above were treated as independent variables, while the remaining variables will be grouped by row mean, under the result of Cronbach's alpha. Cronbach's alpha, also known as alpha reliability, is a measure for assessing the strength of internal consistency, of several items or variables. The alpha score was calculated by correlating the score for every item with the total score for related observations, following the comparison of the variance of individual item scores (Cronbach Lee, 1951). The formula was shown in **formula 3.1**:

Where is the number of scale items, is the variance associated with item , and is denoted as the variance associated with the observed total scores (Chelsea Goforth, 2015). Following the *Rule of Thumb*, if the alpha score is between 0.7 and 0.8, the grouping process is plausible (Stephanie Glen, 2021). With the alpha score mentioned below (format: ), we grouped variables having strong internal consistency by mean of each instance. Cortina (1992) defined that grouping is acceptable if the alpha score is larger than 0.7, recommended in many pieces of research as it is uninterrupted. In supplement, the group-by-mean method (Underhill L.G, 1998) was applied as we preferred grouping variables, without missing much information (like median) (Akhihesh Ganti, 2021) or altering the scale (Daniel McNeish & Melissa Gordon Wolf, 2020). The detailed variables grouping process, as well as the alpha score, were mentioned below.

Formula 3.1: Alpha Score

There are 12 variables, grouped as one variable denoted as the attitude of participants towards general physical activity (). (Scale of all questions were from 1 = "completely disagree" to 7 = "completely agree"). Players' physical behaviour was assessed in two aspects, recency, and frequency (). The first three items for measuring recency of participants’ physical activity were “When was the last time you had (1) a walk for more than 30 min/ (2) had a run/ (3) had a bike ride to get some exercise?”. (The scale for those questions is 1= “more than one month ago”, 2= “about four weeks ago”, 3= “about three weeks ago”, 4= “about two weeks ago”, 5= “about one week ago”, 6= “during the last week” and 7= “yesterday”.) For measuring frequency, the following three questions were adopted: “How many times have you had (1) a walk for more than 30 min/ (2) had a run/ (3) had a bike ride to get some exercise during the last month? with the scale from 1=” never" to 7=" every day". Both former and latter were transformed as one variable, . The remaining three variables, representing participants behaviour related to Pokémon Go, were used for assessing participants’ physical behaviour relating to Pokémon Go (). Questions represented by these three variables were “How many times have you walked more than 30 min/ had a run/ had a bike ride with the intent of searching for Pokémon Go during the last month?” (anchored with 1= “never”, 2= “two times”,3= “from three to five times”,4= “from six to eight times”, 5= “from nine to eleven times”, 6= “from twelve to fourteen times” and 7= “every day”). Those variables were grouped as one variable representing participants’ general physical activity. Back of grouping, the new data set contains eight variables and 981 records. Before model selection, we look at the summary (**table 3.2**) for detailed information.

According to **table 3.2**, there were 981 records and 8 variables. By the pattern demonstrated in **figure 3.2**, we observed that Pokémon Go related behaviour was unusually correlated with age, attitude towards physical activity, and the amount of physical activity. This is possible that there were curve-linear relations between variables, with themselves or others. **Figure 3.3**, additionally, showed that relations exist between Gender and three variables, like education, attitude towards sports, and the amount of physical activity. **Figure 3.4** also proved that education level has positive relations with age and attitude towards physical activity. For modelling the interactions between variables, we apply the polynomial regression model, mentioning details in the following session.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| (n=981) | Min | 1st Qu. | Median | Mean | 3rd Qu. | Max |
| age | 18.00 | 25.00 | 30.00 | 32.55 | 37.00 | 74.00 |
| education | 1.000 | 5.000 | 8.000 | 6.889 | 8.000 | 11.000 |
| Gender | 1.000 | 1.000 | 1.000 | 1.378 | 2.000 | 2.000 |
| Attitude | 3.000 | 5.167 | 5.417 | 5.362 | 5.667 | 6.500 |
| PhysicalActivity | 1.000 | 2.333 | 3.000 | 3.244 | 4.167 | 7.000 |
| PokemonGo\_AppUsage | 1.00 | 1.00 | 1.00 | 2.45 | 4.00 | 7.00 |
| social\_sharing | 1.000 | 1.000 | 1.000 | 1.611 | 1.000 | 7.000 |
| PokemonG0\_Relate.Behaviour | 1.000 | 1.000 | 1.000 | 1.492 | 1.667 | 7.000 |

Table 3.2: Summary of Grouped Data

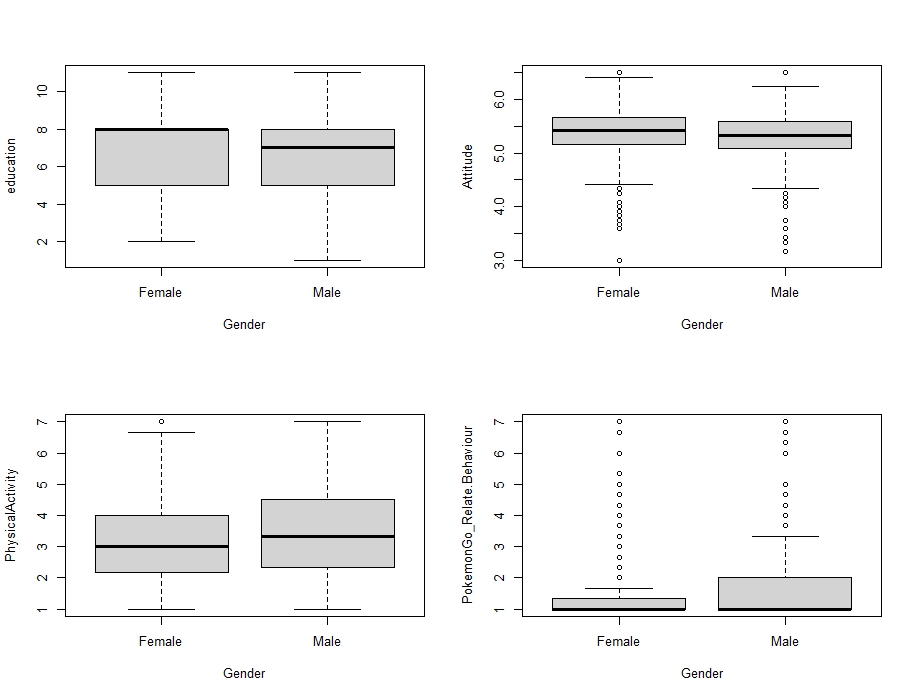


Figure 3.4: Boxplots of Gender vs 4 other variables



Figure 3.3: Scatter plot of data distribution

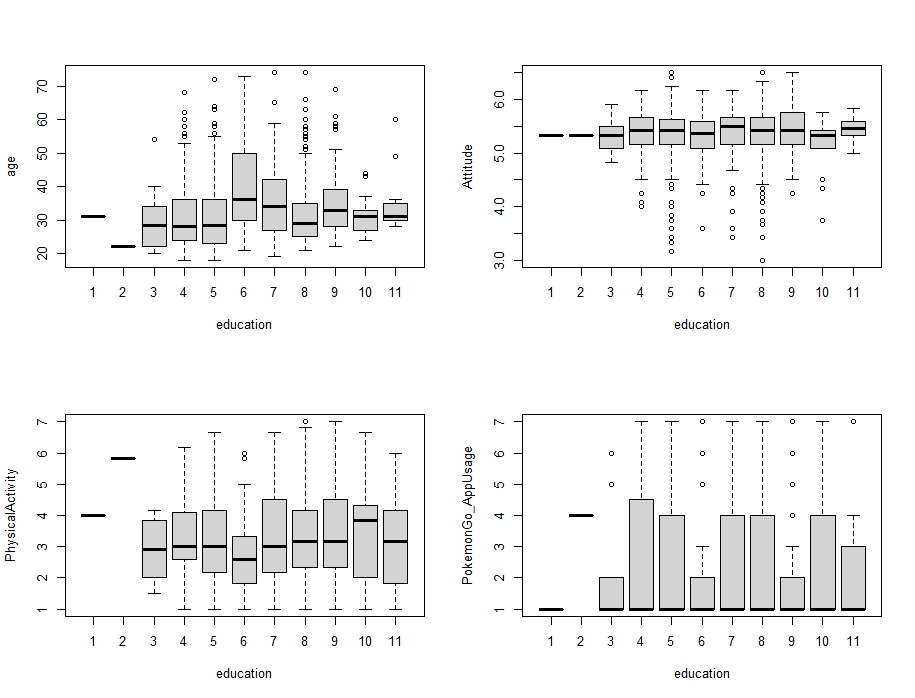


Figure 3.5: Boxplots of Education level vs 4 other variables

# 3.2 Model Selection

Polynomial linear regression model, a linear regression model with a degree of coefficient of more than one, is a model combining interactions between variables (Abhigyan, 2020). Polynomial regression can model the non-linear relationship between dependent and independent variables by adding polynomial terms to linear regression, such as the square of a variable (Agrawal. 2021). Polynomial regression was selected as the most suitable approximation of relations between variables can be provided. Some variables, in reality, have correlations with others or themselves. Simple linear regression, however, cannot model these interactions, leading to large error and inaccurate relationship estimation. With polynomial terms, we built up a better model for relations observations. A vast range of models, also, can be applied in the polynomial regression model, including linear, Gamma, and Poisson. The curvature of a polynomial regression model, moreover, is flexible (Pant. 2019). We can fit a wide range of positive values. Due to accuracy and flexibility, we select the polynomial regression model. A linear model was applied as it is easy for interpretation, as well as an explanation. There are four assumptions for the polynomial linear model. The behaviour of a response variable, to begin with, can be explained by an additive relationship (both linear and curvilinear are plausible) between a response variable and several explanatory variables. Assume that the relations between a response variable and explanatory variables are linear or curvilinear. The independent variables, additionally, should be independent of each other. Last but not least, the errors must be independent and normally distributed, following a mean zero and constant variance (Abhigyan. 2020). We examine the following assumption after obtaining the best model by the stepwise selection method.

After developing a full linear model (showed and explained in **formula 3.10**), stepwise selection, using AIC as criteria, was applied to select the best model. Stepwise regression is a step-by-step iterative and automatic model selection approach (Adam Hayes, 2021), based on backward regression and combining with forward. With this approach, we can re-examine the importance of variables, as well as correct the misleading caused by backward selection. For instance, it is plausible that a variable, removed in the backward selection, is included in the first stage of the forward selection method. Applying the stepwise selection method, we can include that variable again, for obtaining the best model (R. R. Hocking, 1967). McElreath (2016) revealed that Akaike Information Criteria (AIC) is a well-known information criterion, for evaluating the data-fitting performance of a model. The formula was shown in **formula 3.9**, where k represented the number of parameters in a model; L denoted as the likelihood of a model. The model with the smallest AIC is the best as describing the greatest amount of information with the smallest amount of variables (Bevans, 2021). According to **figure 3.11**, the relationships between age and amount of physical activity were differed by the education, justified that there was interaction between education level and age. **By figure 3.12**, the relationships between attitude towards sports and amount of physical activity were also differed by the education, justified the interaction between attitude towards sports and education level. It is necessary for checking the model assumption, to guarantee that model can be applied. The assumptions were examined by all four plots in **figure 3.7**, the explanation and examination of assumption plots were written below.

According to the plot, “Residuals vs Fitted Values” in **figure 3.7**, the pattern of residuals is not obvious, suggesting that the assumption of linear or curvilinear is acceptable. The residuals spread equally around the zero line, proved that the error terms have the same variance. Outliers, additionally do not exist as no residual standing away from the pattern (Department of Statistics Online Programs, 2018); Although having a light tail, Normal Q-Q plot suggested that the dependent variables, inside the model, are normally distributed (Ford, 2015). In the Scale-Location plot, Since the red line is approximately horizontal across the plot, with no clear pattern. In this way, the spreading of the residuals is random, as well as in the neighbourhood of equal for all fitted values. (Zach, 2020). Observing Residuals Vs Leverage, the last plot in **figure 3.7**, no points are affecting the trend much. In this way, there are no outliers. Last but not least, as the majority of players are American (Clement, 2021), and the surveying processing were not restricted on species, religions, or researchers’ favour. Thus, we assumed all records are independent. Based on the graphs in **figure 3.7**, the polynomial linear regression model, demonstrated in the summary, was applied for studying the relations between Pokémon Go and the amount of physical activity. The detailed final model was demonstrated in **table 3.6 and formula 3.8**.

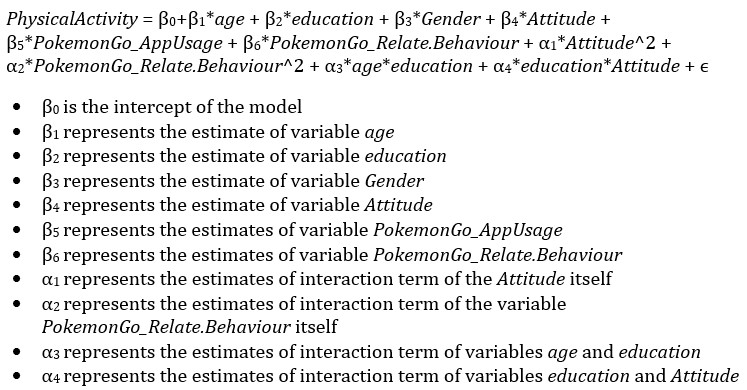
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Coefficients | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 6.136942 | 3.010388 | 2.039 | 0.04176 |
| age | 0.012821 | 0.015683 | 0.818 | 0.41382 |
| education | -0.279905 | 0.284092 | -0.985 | 0.32474 |
| Gender | 0.278323 | 0.084595 | 3.290 | 0.00104 |
| Attitude | -1.761775 | 1.013901 | -1.738 | 0.08260 |
| PokemonGo\_AppUsage | -0.198716 | 0.035205 | -5.645 | 2.17e-08 |
| PokemonGo\_Relate.Behaviour | 0.973577 | 0.207318 | 4.696 | 3.03e-06 |
| I(Attitude^2) | 0.155076 | 0.097228 | 1.595 | 0.11104 |
| I(PokemonGo\_Relate.Behaviour^2) | -0.046845 | 0.029368 | -1.595 | 0.11101 |
| age\*education | -0.003554 | 0.002289 | -1.553 | 0.12079 |
| education\*Attitude | 0.085761 | 0.051815 | 1.655 | 0.09822 |

Table 3.6: Summary of Final Model (formula was shown in formula 3.8)

Formula 3.9: formula of AIC score



Figure 3.7: model assumption plots



Formula 3.8: equation of Final Model with Explanation

# 

# 

Formula 3.10: Full model with explanation

# 

Figure 3.11: Amount of Physical Activity by Age and Education Level

# 

Figure 3.12: Levels of Physical Activity by Attitude towards Physical Activity and education level

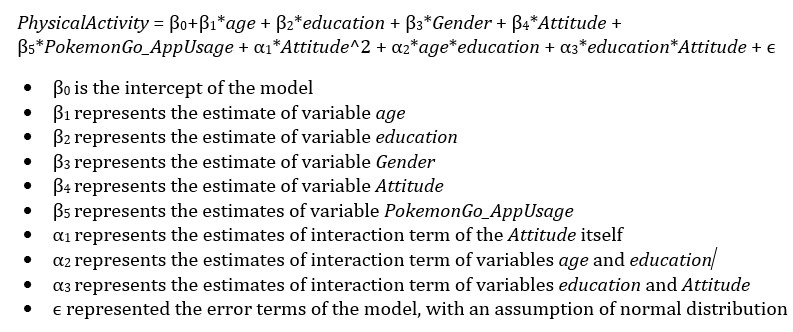
# 4. Analysis Results

With the polynomial regression model constructed before, we answered the questions mentioned in the “research objectives” session. In accordance with the summary (**Table 3.6)**, the estimate for the number of app usage is -0.198716, with a variation of approximately 0.0002460, meaning that a 1% increase in App usage lowered the 0.198716 level of the amount of general physical activity. In contrast with the expectation, playing Pokémon Go negatively affected the amount of physical activity. This phenomenon can be related to the variablerepresenting the Pokémon Go related behaviour. Unlike variable represented the usage of Pokémon Go application, **table 3.6** showed that the estimate of Pokémon Go related behaviour was 0.973577, possibly suggested activities related to Pokémon Go increased the number of physical activities, rather than the frequency of playing Pokémon Go. For explaining and examining this result, the factor of Pokémon Go related behaviour was removed from the model, and the summary was shown in **Table 4.1.** According to **table 4.1**,theestimated values of Pokémon Go application usage became positive (0.051826) after eliminating the effects of Pokémon Go related behaviour. This represented the activities associated with Pokémon Goacted as a suppressor of the amount of physical activity if fixed other factors. It, thus, was plausible that the positive effects of Pokémon Go app usage on the amount of physical activity restricted by Pokémon Go related activity. Despite the variables mentioned before, there were more variables related to the amount of physical activity.

**Table 3.6** manifested that the age and genderof a playerincrease the amount of physical activity, while education level and attitude towards sports negatively affect physical behaviour. There was, additionally, a correlation between age and education level, forming a new variable representing the interaction between age and education level. Education level and gender, in reality, is related to the attitude of participants towards physical activity***.*** Under **figure 4.3**, the average attitude score of participants, with the first three education levels, were lower than the participants accepting higher education levels. This represented that participants with higher education held a more active attitude towards physical activity. Moreover, the formation of a new variable presenting the interaction between education level and participants' attitude towards physical activity (shown in **table 4.1**), proved the correlation between education level and attitude towards physical activity. Besides education level, **Figure 4.3** represented that females have a more positive attitude towards physical activity (with a large area of overlapping), compared with males. (Note: in the variabledenoting gender of participants’, a female was labelled as "1", while the male was denoted as “2”.) Therefore, both gender and education level correlated with participants’ attitude towards sports, proving that attitude towards physical activity is a key factor. In contrast to the above situation, males have a higher amount of physical activity (represented by the positive estimates in **table 3.6**), being plausible that the attitude towards physical activity is not positively affecting the amount of physical activity. **Table 3.6** showed that the estimated value of the variable representing the attitude of participants towards sport was negative (-1.761775). Although seemed unreasonable, a more positive attitude towards physical activity reduced the amount of physical activity. Considering the interaction between education level and Attitude towards sports, we observed positive relations with the amount of physical activity, which was explained later.

Table 4.1: Summary of Model without effects of Pokémon Go related behaviour

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Coefficients | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 7.070780 | 3.010388 | 3.164619 | 0.0257 |
| age | -0.001712 | 0.016454 | -0.104 | 0.9171 |
| education | -0.016170 | 0.298011 | -0.054 | 0.9567 |
| Gender | 0.392655 | 0.088170 | 4.453 | 9.43e-06 |
| Attitude | -2.156115 | 1.060194 | -2.034 | 0.0423 |
| PokemonGo\_AppUsage | 0.051826 | 0.021254 | 2.438 | 0.0149 |
| I(Attitude^2) | 0.230975 | 0.101574 | 2.274 | 0.0232 |
| age\*education | -0.001736 | 0.002289 | -1.553 | 0.12079 |
| education\*Attitude | 0.085761 | 0.051815 | 1.655 | 0.09822 |



Formula 4.2: model without effects of Pokémon Go related behaviour



Figure 4.3: boxplot for relations between education and Attitude



Figure 4.4: boxplot for relations between Gender and Attitude

# 5. Discover and Discussion

Pokémon Go is a popular AR mobile game, causing the revolution of mobile games by combing AR technology with the mobile game (Lopez German 2016). Players of this game can catch and hatch Pokémon, a virtual creature. Those activities require walking for a certain distance or arriving at a specific location. With this characteristic, some researchers like Kamboj and Krishna (2016) claimed that Pokémon Go is an effective mobile game for rising the walking time, following the improvement of public health, including obesity. However, Gabbiadini & Greitemeyer (2018) declared that the effects caused by Pokémon Go are restricted to the activities related to the application, similar to my opinion mentioned before. As mentioned in the “Analysis Results” session, the amount of app usage (with estimate -1.761775) reduced the amount of general physical activity. The interpretation is Pokémon Go did not lead players more favour in physical activity, echoed with Baranowski and all the others (2012) that there are no obvious relationships between frequency of playing video games and amount of general physical health. There are several possible reasons for this phenomenon. The first one is the game design. For instance, some players hatched the Pokémon eggs when driving (with a speed lower than 10km per hour.) (Ayers et al., 2016), or took public transport for catching a Pokémon. In this way, walking can be unnecessary. Another reason is the preference of participants. Unfortunately, "analysis results" suggested that players were more willing to join the app-related activity, instead of general physical activity. Comparing **table 3.6**with **table 4.1**, we have discovered that the estimated values of the amount of app usage became positive if removing the factor of Pokémon Go related activity, demonstrating that app-related activity causes huge effects on the amount of physical activity. We surmise with confidence that players were willing to join an application-related activity, instead of general physical activity. In this way, we discovered that Pokémon Go cannot increase the amount of physical activity directly as the effects are limited. The effects will disappear due to the altering of playing methods, as well as the reduction of players.

The effects of Pokémon Go on the amount of physical activity, however, is unstable and inconsistent. The format of activity, to begin with, can be altered with accidents. Due to the COVID-19 pandemic, the entire game was changed for indoor playing (Maher 2020). For example, players have not required to hatch Pokémon through walking. Players, indeed, bought tools-in-game for hatching Pokémon automatically. Players, additionally, used “Incense", a tool for attracting Pokémon, to catch Pokémon without travelling. Last but not least, players can join Raid Battle (activity for catching rare Pokémon) without reaching a Gym. Those changes lowered the requirements for walking outside. Despite the company's policy, it is commonly known that most mobile games cannot attract a large number of users forever. Bratuskins (2018) also claimed that the lifespan of the mobile game became shorter. For the above reasons, it is unstable and unsustainable using mobile games for public health enhancement. Others, including Attitude, age, education level and gender are in consideration for discovering solutions optimizing public health.

Besides app usage and the amount of participation in app related activities, both attitudes, education level, gender and age are factors related to the amount of physical activity. Attitude, in reality, is a key factor, but not in the way expected previously. **By Figure 4.3**, we proved that females have a slightly more positive attitude towards physical activity on average, maximum and minimum. Females had a more positive attitude towards various areas of sports, compared with males (Tomik, 2008). This phenomenon could be related to education (according to **figure 3.4**, the average education level of females was higher than that of males). This result, however, was not absolute. In **figure 4.3**, we perceived a considerable area of overlapping when comparing the box plots of males and females. In some cases, males had a more positive towards sports (Koca & Demirhan, 2004). Thus, Gender is moderately related to the attitude towards physical activity. Besides sex, we observed the effects on the amount of physical activity in two aspects: age and education, as well as discussing relations between attitude towards sport, and the amount of physical activity.

Per **table 4.6**, the age was also interacted with the education level, demonstrating that education level is correlated with the age of participants, as visualized in **figure 3.5.** The appearance of the above trend is related to the free education policy in most of the countries. Usually, the level of education instilled was following the age of a person. Despite the age and gender of a player mentioned before, education level is also correlated with participants’ attitude towards sports, proved by the positive trend presented in **figure 4.2**. **Formula 3.8** revealed that there was an interaction between education level and attitude towards sports. However, the relationship between attitude towards physical activity, and the amount of exercise was negative. We concluded that attitude towards physical activity is not related to the amount of physical activity, unlike the opinions from Araújo and Dosil (2015). Although cannot be explained clearly, it is reasonable that this phenomenon can be related to other factors besides knowledge. Most people know the benefits of sports. They, however, did not exercise due to several reasons, including a long working hour or pandemic. Although attitude is not related to the amount of physical activity. By **table 3.6**, we discovered that the variable representing the interaction term between education level and attitude towards sport is positively correlated with the amount of general physical. It is possible that merely the increase of both education level and attitude can increase the amount of physical activity. This is associated with the amount of leisure time inactivity. Higher education level was related to the lower odds of leisure-time physical inactivity (Piirtola, et. al., 2016). However, rising the education level is not enough. **Table 3.6** education level negatively correlated the number of sports if other factors were fixed. Based on the positive value of the interaction term between education level and attitude towards sports, the government should not only increase the funding of tertiary education but also instil a positive attitude towards sports in students.

**6. Conclusion and Limitations**

Sport is essential for public health. However, the proportion having exercise regularly in the United Kingdom was limited. Pokémon Go is a well-known AR mobile game, with a huge number of players. I believed that Pokémon Go is serviceable for public health due to the claims from multiple pieces of researches, whereas some researchers argued that the effects caused by Pokémon Go were indirect and unsustainable. Discovering methods for public health improvement, we studied the relationships between Pokémon Go and the amount of general physical activity. The entire study was processed in four aspects: The relations between amount of app usage and the amount of physical activity; the preference of Pokémon Go players; other variables related to the amount of physical activity, as well as the relationships between the attitude and two variables, gender and education level. Constructing required variables, we grouped some series of variables by mean of each instance, applying Cronbach's alpha as an internal correlation observation method, as well as evidence for the variables-grouping process. After variables grouping, the polynomial linear model, a linear model that allows interactions between variables, was applied for discovering relations between variables. Stepwise selection methods using AIC as selection criteria were used for model selection. The final model was written in **formula 3.8**. With this model, we found some interesting facts.

To begin with, Pokémon Go cannot directly increase the amount of physical activity because participants focused on app-related activities. In this way, Pokémon Go is not a good method of improving public health due to the instability and unsustainability of mobile games. Attitude towards physical activity was a key factor as both education level and gender were related to the amount of physical activity, with a high correlation between education level and age. The positive attitude towards sports, nevertheless, could not increase the amount of physical activity. Focusing on the interaction between education level and attitude towards sport, we suggested instilling a positive attitude towards sports in students and reducing education fees for public health improvement. There are several limitations during the study, being improved in future. To begin with, the population of the dataset was from America. The statistics from Clement (2021), however, showed that there was a considerable number of players in Great Britain, Japan, Sweden and Canada. It is well known that there are many differences between countries, including culture and education system. Biases possibly exist if only observing data from players in America. The study should also have hosted in other countries in future. 999 records were used in this study. However, there are more than eight hundred thousands of active users in America, not to say the whole world. For future study, increasing the population was recommended. This dataset only contained values from questions inside the survey. There is a risk that participants forget the number of times playing Pokémon, or lie on the survey due to shame. The future study is suggested including participants’ data inside the application. Last but not least, figure 3.7 demonstrated that there was a hidden pattern in the fitted values versus residual plot. We can apply more kinds of models, including Poisson, negative-binomial or neural network model.

**7. Reference**

* Agrawal Raghav. (2021). All you need to know about Polynomial Regression. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/07/all-you-need-to-know-about-polynomial-regression/
* Araújo, A. T., & Dosil, J. (2015). The influence of attitudes toward physical activity and sports. Motriz: Revista de Educação Física, 21, 344-351.
* Ayers, J. W., Leas, E. C., Dredze, M., Allem, J. P., Grabowski, J. G., & Hill, L. (2016). Pokémon GO—a new distraction for drivers and pedestrians. JAMA internal medicine, 176(12), 1865-1866.
* Baranowski, T., Abdelsamad, D., Baranowski, J., O’Connor, T. M., Thompson, D., Barnett, A., ... & Chen, T. A. (2012). Impact of an active video game on healthy children’s physical activity. Pediatrics, 129(3), e636-e642.
* Bradfield & Underhill. (2004). IntroSTAT. 2nd. Department of Statistical Sciences, University of Cape Town. p.181-183
* Bratuskins Martins. (2018). Mobile game lifespan is shortening, but there’s a way to extend it.. Medium. https://medium.com/monetizr/mobile-game-lifespan-is-shortening-but-theres-a-way-to-extend-it-d7e96561b008
* Buraimo Babatunde, Jones Helen & Millward Peter. (2011). "Adult participation in sport: Analysis of the Taking Part Survey". department for culture, media, and sport. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\_data/file/137986/tp-adult-participation-sport-analysis.pdf
* Clement J.. (2021). Number of daily active users (DAU) of Pokémon Go via iPhone in selected countries in January 2021. https://www.statista.com/statistics/604551/pokemon-go-daily-active-users-in-europe/
* Cortina Jose. (1992). What Is Coefficient Alpha? An Examination of Theory and Applications. American Psychological Association, Inc. p.101
* Cronbach, Lee J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika. Springer Science and Business Media LLC. 16 (3): 297-334
* Department of Statistics Online Programs. (2018). “4.2 - Residuals vs. Fits Plot”. The Pennsylvania State University. https://online.stat.psu.edu/stat462/node/117/
* Eysenbach Gunther. (2016). Influence of Pokémon Go on Physical Activity: Study and Implications". Journal of Medical Internet Research. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5174727/
* Fisher Tim. (2021). What Is Augmented Reality?. Lifewire. https://www.lifewire.com/augmented-reality-ar-definition-4155104
* Ford Clay. (2015). Understanding Q-Q Plots. The University of Virginia. https://data.library.virginia.edu/understanding-q-q-plots/
* Gabbiadini, A., Sagioglou, C., & Greitemeyer, T. (2018). Does Pokémon Go lead to a more physically active life style? . Computers in Human Behavior, 84, 258-263.

# Ganti Akhilesh. (2021). "Median". Investopedia. https://www.investopedia.com/terms/m/median.asp

# Glen Stephanie. (2021). Cronbachâ’s Alpha: Simple Definition, Use and Interpretation. Statistics How to. https://www.statisticshowto.com/probability-and-statistics/statistics-definitions/cronbachs-alpha-spss/

# Goforth Chelsea. (2015). Using and Interpreting Cronbachâ’s Alpha. The University of Virginia Library. https://data.library.virginia.edu/using-and-interpreting-cronbachs-alpha/

# Hayes Adam. (2021). Stepwise Regression. Investopedia. https://www.investopedia.com/terms/s/stepwise-regression.asp

# Hocking, R. R. (1976). A Biometrics Invited Paper. The Analysis and Selection of Variables in Linear Regression. Biometrics, 32(1), 1â“49. https://doi.org/10.2307/2529336

# Jacquet Jennifer. (2011). The pros & cons of Amazon Mechanical Turk for scientific surveys. Scientific American. https://blogs.scientificamerican.com/guilty-planet/httpblogsscientificamericancomguilty-planet20110707the-pros-cons-of-amazon-mechanical-turk-for-scientific-surveys/

# Kamboj, A. K., & Krishna, S. G. (2017). Pokémon GO: An innovative smartphone gaming application with health benefits. Primary care diabetes, 11(4), 397-399.

# Koca, C., & Demirhan, G. (2004). An examination of high school students' attitudes toward physical education with regard to sex and sport participation. Perceptual and motor skills, 98(3), 754-758.

# Kostoulas Achilleas. (2013). On Likert scales, ordinal data and mean values. Achilleas Kostoulas. https://achilleaskostoulas.com/2013/02/13/on-likert-scales-ordinal-data-and-mean-values/

* Lopez German. (2016). Pokémon Go, explained. Vox. https://www.vox.com/2016/7/11/12129162/pokemon-go-android-ios-game
* Maher, C. (2020). How COVID-19 transformed Pokémon GO into “Pokémon stay-at-home”. ArsTechnica. Last modified July, 13.
* McElreath Richard. (2016). Statistical Rethinking: A Bayesian Course with Examples in R and Stan. CRC Press. p.189
* McNeish, D., & Wolf, M. G. (2020). Thinking twice about sum scores. Behavior research methods, 52(6), 2287-2305.
* Piirtola, M., Kaprio, J., Kujala, U. M., Heikkilä, K., Koskenvuo, M., Svedberg, P., ... & Ropponen, A. (2016). Association between education and future leisure-time physical inactivity: a study of Finnish twins over a 35-year follow-up. BMC Public Health, 16(1), 1-11.
* Reilly Luke. (2017). Pokémon GO Coming to Smartphones. IGN Entertainment. https://www.ign.com/articles/2015/09/10/pokemon-go-coming-to-smartphones
* Seippel, Ø. (2006). The meanings of sport: fun, health, beauty or community?. Sport in Society, 9(1), 51-70.
* Sengupta Somak. (2020). Gamma Distribution Explained | What is Gamma Distribution? . Great Learning. https://www.mygreatlearning.com/blog/gamma-distribution/
* Smith, C. (2017). 80 amazing Pokemon Go statistics. RetrievedNovember 17, 2017, from https://expandedramblings.com/index.php/pokemon-go-statistics/.
* Stacy, E. W. (1962). A generalization of the gamma distribution. The Annals of mathematical statistics, 1187-1192.
* Tomik, R. (2008). Adolescent attitudes towards sport depending on school level, gender and school sports club membership. Journal of Human kinetics, 20(1), 121-130.
* Webster Andrew. (2015). "With Pokémon Go, Nintendo is showing that it takes mobile seriously". The Verge. https://www.theverge.com/2015/9/10/9300101/pokemon-go-nintendo-mobile-games
* Wong, F. Y. (2017). Influence of Pokémon Go on physical activity levels of university players: a cross-sectional study. International journal of health geographics, 16(1), 1-12.
* Zach. (2020). How to Interpret a Scale-Location Plot (With Examples). Scatology. https://www.statology.org/scale-location-plot/

**8. Appendix**

