Software Requirements Specification

for

Analysis of Nintendo's Advertising Investment in 2019

Version <1.0>

Prepared by <Yuda Li & Zimin Lu>

<gamesteam>

<02/26/2019>  
<02/23/2019>

Table of Contents

Table of Contents 2

Revision History 2

1. Introduction 1
   1. Problem Statement 1
   2. Intended Audience and Reading Suggestions 1
   3. Detailed Problem Definition 1
   4. Purpose and Motivation 1
   5. Problem Scope 1
   6. Project Failure 1
   7. References 1
2. Background 2

2.1 Previous Academic Approaches 2

* 1. Non-academic Approaches 2
  2. Design and Implementation Constraints 3
  3. Assumptions and Dependencies 3

1. Data Description 3
   1. Data Sources 3
   2. Independent Variables 4
   3. Dependent Variables 4

Appendix A: Glossary 5

Appendix B: Analysis Models 5

Appendix C: Issues List 5

Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Date** | **Reason For Changes** | **Version** |
|  |  |  |  |
|  |  |  |  |

# Introduction

## Problem Statement

Does Nintendo continue to strengthen advertising investment in North America?

## Intended Audience and Reading Suggestions

The audience is Nintendo, indicating that the investment in advertising will affect the sales of the product.

## Detailed Problem Definition

The impact of advertising investment on the three game platforms from 1985 to 2017 on product sales during this period.

## Purpose and Motivation

The motivation for the project is to show game platform manufacturers the impact of changes in advertising spending on product sales.

Creating a report, based on statistical models, linear models and random forest analysis of game data, comparing Nintendo and Sony and Microsoft's sales and advertising investment, and predicting the trend of Nintendo's advertising sales in 2019, planning a reasonable Nintendo company's advertising investment in 2019 Money interval.

Based on descriptive analysis and predictive analysis, the project analyzes Nintendo's advertising investment amount, which is beneficial to increase consumer groups and increase profits in the most reasonable way.

## Problem Scope

The project will analyze the specific data released by the game from 1985 to 2017.

## Project Failure

The small amount of data may cause inaccuracies in the prediction model, and some immeasurable human factors in the future will cause inconsistency with the results, that is, the investment of reasonable money for advertising but the sales growth is not obvious.

## References

* John Walker, S. 2014. Bid Data: a revolution that will transform how we live, work, and think.
* Popescu, M. 2015. Construction of economic indicator using internet searches.

# Background

## 2.1 Previous Academic Approaches

Advertising firm 360i, for example, did exactly that during the Super Bowl in 2013 with its [Oreo campaign](https://www.wired.com/2013/02/oreo-twitter-super-bowl/). After a blackout halted the game, within a matter of minutes the phrase, “Power out? No problem. You can still dunk in the dark,” was circulating around major social media channels. The campaign garnered immediate media attention and increased Oreo’s social media following.

- Andrew Medal.2017. How Big Data Analytics Is Solving Big Advertiser Problems

## Non-academic Approaches

Despite the large amount of data publicly available from the Internet searches, the opportunities for more advanced analysis are still relatively unexplored. Understanding and exploring the methods of implementing Data Mining practices in various business sectors and industries has been drawing attention of the scientific community for a few years. Each field has its challenges, and problems that need to be solved, different types of data, but also specific methods that need to be adapted to specific circumstances.

Data mining methods and techniques are used in various economic analyses to capture different economic indicators and used for various economic forecasts (Choi and Varian, 2012). The corporate world has implemented the use of Data Mining technologies to build economic indicators using the volume of Internet searches over time (Popescu, 2015). An example of two trendy brands that have invested considerably in their marketing campaigns are Coca-Cola and Pepsi, the manufacturers of two of the most popular drinks in the world.

For this Coca-Cola and Pepsi analysis, data mining for web searches was implemented to assess both companies’ brand popularity over time. From the volume of searches over time, the major events generated by each firm’s marketing activities can be analyzed to extract and exploit enormous amount of information. As one of the long-term marketing department objectives is to increase the brand popularity over time, the web searches can be used as a tool to measure the impact and results of those marketing strategies (Popescu, 2018).

Apart from increasing brand popularity and brand awareness from Data Mining technologies, volume searches for each brand, in this case Coca-Cola and Pepsi, offer the possibility to explore the geographic distribution of the searches. These internet searches can be used not only to analyze the volume searches and trends over time, but also the geographic interest for specific events, as the impact and results from these events can be measured fast.

Another type of analysis is the customers’ behavior prediction using data mining on web searches. A well-known player in this arena is Walmart, as it tries to predict what the customers will buy ahead of hurricanes. Because Walmart records and uses the data from the purchase made by each customer, they tried to predict the top sales products before a hurricane (Popescu, 2018). They found that a particular snack Pop-Tarts was bought in massive quantities before the hurricane. Walmart would fill the stocks with this product and even put the tarts near the entrance and in the strategic points so that anyone can see them. The result was an increase of sales in that period.

* John Walker, S. 2014. Bid Data: a revolution that will transform how we live, work, and think.
* Popescu, M. 2015. Construction of economic indicator using internet searches.

## Design and Implementation Constraints

The current dataset is very large and of different origins, and further importing the database is required to optimize the variables as much as possible. The structured form of the dataset in the library makes the conversion and standardization of the data convenient and easy to analyze, which helps to evaluate the different aspects of the variable process.However, moving further into the analytical process, these factors could pose an important effect on the direction of the project and its final outcome. Nonetheless, there is still more data to be collected, analyzed, and interpreted in order to determine the usefulness of the data at hand.

A possible constraint may be encountered down the road is lack of sufficient time to validate the model and get the most accurate results possible. Considering the project’s timeframe of 8 weeks, lack of sufficient time may be an impediment to thoroughly assess relationships between variables at a high level, or even not being able to fully understand the subtleties within the data.

Other possible implementation constraints are:

* Inputting categorical variables into the model based on the sub-independent variables that will be incorporated to validate the model.
* Difficulty in reducing the number of independent categorical variables
* Maintaining a 2,0000 rows minimum threshold of data post-tidying.
* Including more numerical variables into the model.

## Assumptions and Dependencies

Assumptions:

Assuming that Nintendo’s sales in North America are lower than those of Sony and Microsoft, then the advertising have an impact on Nintendo’s sales. And advertising has a strong impact on Nintendo’s sales. Nintendo needs to continue increase advertising funds to stimulate consumption growth.

Dependencies:

The data depends on the data at accurate.There must be a direct link between sales and advertising investment.

# Data Description

## Data Sources

The table data source named “games” in the database and the kaggle website, which contains 16719 game release data between 1985 and 2017.

The table data source named “games2” in the database and the statistics website, which contains 16600 game sales and marketing data between 1985 and 2017.

The table data source named “ad” in the database and the source from statista website, which contains the advertising data of three game manufacturers in 1985 1990 2015 2016 2017.

## Independent Variables

<Enumerate all of your independent variables you will use in the data analysis.>

3.1.1 Name

1. The games name
2. Data type: textual

3.1.2 Platform

1. Platform of the games release (i.e. PC,PS4, etc.)
2. Data type: textual

3.1.3 Year

1. Year of the game's release

2. Data type: numeric

3.1.4 Genre

1. Genre of the game

2. Data type: textual

3.1.5 Publisher

1. Publisher of the game

2. Data type: textual

3.1.6 Players

1. Game’s players

2. Data type: numeric

3.1.7 Critic\_Count

1. The number of critics used in coming up with the Critic\_score

2. Data type: numeric

## Dependent Variables

<Enumerate all of your dependent variables you will use in the data analysis.>

3.1.1 Sales

1. Game’s sales
2. Data type: numeric

3.1.2 Critic\_Score

1. Aggregate score compiled by Metacritic staff
2. Data type: numeric

Appendix A: Glossary

ad — advertising

Appendix B: Analysis Models

* Statistics
* Linear Regression Model
* Logistical regression analysis

- Random Forests

Appendix C: Issues List

Tidying procedures vary depending on the specific project and the model used, therefore selecting the right tidying method for the data will require a careful evaluation and may pose issues deeper into the analytical stages.