Software Design Specification

for

Analysis of Nintendo's Advertising Investment in 2019

Version <1.0>

Prepared by <Yuda Li & Zimin Lu>

<gamesteam>

<03/03/2019>  
<02/26/2019>

Table of Contents

Table of Contents ii

Revision History ii

1. Introduction 1
   1. Goals 1
   2. Statement of scope 1
   3. Model context 1
   4. Major constraints 1
2. Data design 1
   1. Data sources 2
   2. Internal data structure 2
   3. Variable Description 2
      1. Independent Variables 2
      2. Dependent Variables 2
   4. Pre-design Analysis 2
   5. Tidying procedure 3
   6. Database description 3
3. Model Architecture 3
   1. Type of Model 3
   2. Training set 4
   3. Testing set 4
4. Approach 4
   1. Implementation Details 4
   2. System requirements 4
5. Testing Strategy 4
   1. Expected response 4
   2. Performance bounds 4
   3. External review and validation 4
6. References 5

Appendix A: Glossary 5

Appendix C: Issues List 5

Revision History

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

# Introduction

As the game develops better and better, the competition is more and more motivating. The change project studied the competition between the three famous game platforms and compared their sales to help determine whether Nintendo needs to strengthen its promotion in North America.

## Goals

Establish a predictive model, compare sales, find links between sales and advertising, and determine if Nintendo needs to increase advertising spending in North America

## Statement of scope

Given the limited eight-week time to meet the business sponsor's wishes, it has been determined that the 1985-2017 data is sufficient to build and implement the model. In order to improve the accuracy of the model, including making reasonable recommendations for Nintendo advertising. However, future human-made unknown variables have an impact on the results of the study, and it is difficult to obtain preliminary data needed to evaluate this additional variable, which makes it infeasible. In addition, it may introduce defects into the model rather than increase its effectiveness.

The preliminary sample database consists of data tables of around 20,000 rows, and sales and marketing related data will be determined as model variables.

## Model context

A key part of the project involves creating models that verify the results it will generate. The team will create and test the model by implementing the programming language R. Further in-depth analysis of the process, test the validity of the model by inputting different sub-variables, and improve the accuracy of the model.

## Major constraints

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.

# Data design

Use the main independent variables and dependent variables established, such as the sales of three gamers in North America, the number of players, and the proportion of different game platforms.

## 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.

## Internal data structure

Internal data is historical sales data and product data, at this project they are all structure data.

## Variable Description

### Independent Variables

3.1.1 Players

1. Game’s players
2. Data type: numeric

3.1.2 Advertising investment

1. Game’s Advertising investment
2. Data type: numeric

3.1.3 Platform

1. Game’s publish platform

2. Data type: textual

### Dependent Variables

3.1.1 Sales

1. Game’s sales
2. Data type: numeric

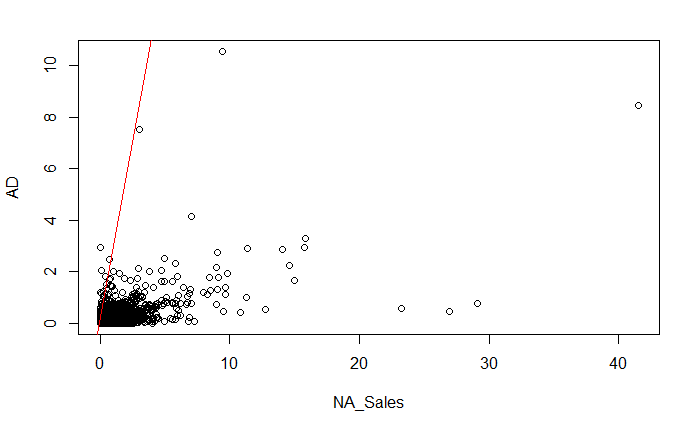
3.1.2 Ad

3. Game’s advertising

4. Data type: numeric

## Pre-design Analysis

## Tidying procedure

Because the research is aimed at the promotion of sales in North America, the European region does not consider it. The raw data needs to be cleaned up in r.

## Database description

The project's spreadsheet contains specific data on game releases from 1985 to 2017. The collected sample data was collected as a CSV file from a variety of sources, for a total of approximately 30,000 data. It includes the game name, distribution platform, number of players, sales volume, advertising investment, etc. These variables are summarized by the MySQL database and connected to Rstudio, followed by a screening test.

# Model Architecture

The initial model is based on a CSV file formatted in Excel. In the data preparation phase, a preview of the various arguments contained in the database is completed to determine the final arguments used.

## Type of Model

* Linear regression:In this model, the data and advertising input data sold in North America are used to draw a scatter plot, and then 80% of the data is used for predictive analysis.

## Training set

80% data of NA sales and ad

## Testing set

20% data of NA sales and ad

# Approach

## Implementation Details

Computer languages and versions: R version 3.4.3

Platform: x86\_64-apple-darwin15.6.0 (64-bit)

Running under: macOS High Sierra 10.13.6

libraries: tidyvers

## System requirements

minimum hard drive space：23.7M

network bandwidth：Y

RMA amont：456.8M

CPU speed：1.2%/CPU

# Testing Strategy

## Expected response

Linear regression graph with 90% correlation accuracy between ad and NA sales.

## Performance bounds

Performance bounds are subject to the limitations of the hardware of our execution machine:

RAM- 8GB

Hard drive – 80GB

Processor – Intel i5

## External review and validation

<What is the process of externally validating the model?>

# References

Appendix A: Glossary

Ad ——advertising

Appendix C: Issues List