**Research Plan and Design**

The first part of the dataset was collected from VGChartz.com, a website that tracks video game and console sales. Those unfamiliar with VG Chartz should know that their data is not 100% accurate as they "use a number of proprietary and ever-developing methods" to track their sales data. Unfortunately, I am not in any position to get this data from a more reliable source like NPD (National Purchase Diary) or Nintendo directly so it was VG Chartz or nothing. The sales data is so old that I figured a VG Chartz estimation would be good enough even if it’s a little off.

I went to the page with their [list of Gamecube games](http://www.vgchartz.com/platform/29/gamecube/), saved it to a PDF file, and then copied it to an Excel spreadsheet. This filled in the information for the variables game, release year, genre, publisher, and sales (for all regions). I looked over the entire thing and made a number of changes involving double counted games, unreleased games, incorrect publisher information, and debatable genres. Details on all of the changes can be found in the Code Book section.

Metacritic.com is a website that aggregates review scores for music, video games, TV, and movies. I used this website for the variables Metacritic score, ESRB rating, exclusivity, and multiplayer. I couldn’t figure out a way to easily copy and paste this data into like I did with VG Chartz, so I entered all of it manually.

Any game information missing from VG Chartz and Metacritic or anything I felt needed verification was researched on IGN, YouTube, and other gaming websites.

After completion of the Excel spreadsheet, all work on this project involving analysis and graphical representation of the data was done in R. The graphs were coded in R markdown files and outputted to Microsoft Word with Knitr. I’ve also included some of the R code and output indented with a “##.” Every piece of code has been saved for verification.

**Variables and Descriptions**

* **Game:** Title of the game released on the Nintendo Gamecube
* **Publisher:** Name of the company that published the game
* **Publisher 2:** Condensed recoded publisher variable
* **Release Year:** Year the game was released on the Gamecube
* **Release Year 2:** Condensed recoded release Year variable
* **Genre:** Genre the game falls under
* **Genre 2:** Condensed recoded genre variable
* **ESRB Rating:** Rating given by the ESRB
* **ESRB Rating 2:** Condensed recoded ESRB rating variable
* **Multiplayer:** Number of players that can simultaneously play the game
* **Multiplayer 2:** Condensed recoded Multiplayer variable
* **Exclusivity:** Whether or not the game released exclusively on the Gamecube
* **Metacritic Score:** Average review score collected by Metacritic
* **Sales N.A.:** Number of copies sold in North America (in millions)
* **Sales Eur:** Number of copies sold in Europe (in millions)
* **Sales Jap:** Number of copies sold Japan (in millions)
* **Sales Other:** Remainder of copies sold in other territories (in millions)
* **Sales Total:** Total number of copies sold worldwide (in millions)