MATH216 HW 1
Due: Wednesday, February 28th at 11:59pm
Please submit an electronic copy of your assignment to both *alyford@middlebury.edu* and *tinac@middlebury.edu* using the subject line: MATH216 - HW1

This homework will investigate the *Mario Kart.csv* data set sent via email. These data contain information about selling prices of the Nintendo Wii Game *Mario Kart Wii* during the month of October 2009 on the website *www.ebay.com*. Below are descriptions of each of the variables in the data set.

ID: Auction ID assigned by Ebay.

duration: Auction length, in days.

nBids: Number of bids.

cond: Game condition, either new or used.

startPr: Start price of the auction.

shipPr: Shipping price.

totalPr: Total price, which equals the auction price plus the shipping price.

shipSp: Shipping speed or method.

sellerRate: The sellers rating on Ebay. This is the number of positive ratings minus the number of negative ratings for the seller.

stockPhoto: Whether the auction feature photo was a stock photo or not. If the picture was used in many auctions, then it was called a stock photo.

wheels: Number of Wii wheels included in the auction. These are steering wheel attachments to make it seem as though you are actually driving in the game. When used with the controller, turning the wheel actually causes the character on screen to turn.

title: The title of the auctions.

Your mission is to investigate the relationships between the variables presented here to ultimately determine how to earn the most money from selling your (pretend) version of *Mario Kart Wii*. Answer the following questions to the best of your ability using R Markdown. Questions are intentionally open-ended and often have multiple correct (and incorrect) answers. Graphs should be accompanied by explanations (where appropriate), and explanations should be accompanied by graphs (where appropriate).

1) Construct a graph showing the distribution of total selling prices for all games in this data set. Describe what you see in the graph you made.

2) Investigate any outliers observed in your graph from 1). Determine why they are such outliers, and either justify why you should remove these points from the data set or why they should remain.

3) Describe the relationship between the mean total price of the game and the number of Wii steering wheels included in the ebay auction.

4) Construct a new variable indicating whether or not the *Mario Kart* game was packaged with any number of steering wheels. Calculate the median total selling price for each combination of game condition and whether or not the package contained any number of steering wheels.

5) Construct and interpret a single graph that describes the relationship between the variables: starting price, total price, condition, and whether or not the auction used a stock photo.

6) Describe the relationship between the duration of the auction and its final selling price.

7) Describe the relationship between the number of bids (perhaps a measure of the auction's popularity) and the total selling price of the game for each combination of condition and stock photo usage.

8) Finally, if you were to sell your copy of *Mario Kart Wii* on ebay, determine the most optimal set of variables you would use to list your game. Be sure to support your answer with analyses done in this assignment.