

# ASSIGNMENT 1

---

In this assignment, you will learn how to evaluate various hypotheses in cases of uncertain data. You will also learn the basics of R programming, which is one of the most popular tools for machine learning and is known to attract the highest salaries: <http://blog.revolutionanalytics.com/2014/02/r-salary-surveys.html>

You will also learn to use R for analysis of large datasets, graphing, and analysis. You can download and install R and also RStudio IDE from here: <https://www.rstudio.com/>. If you wish, you are free to do this assignment in any other programming language also, but R is highly recommended.

Please submit one zip file containing two folders – one for question 1 and another for question 2. The requirements for each question are mentioned below.

---

## **Question 1 (20 points)**

In the class, we analyzed why it's not a good idea to use ML to predict the outcome of random variables, such as coin tosses or casino games. Your friend comes up to you and says that he has discovered a system that can beat any game of chance i.e. one that involves random variables. His strategy (hypothesis) is as follows:

- Start with one unit of bet
- If you lose, double your previous bet
- If you win, bet one unit.

For example, suppose you have \$1000 to bet and one unit of bet is \$100.

\*\* Remember for a win, you double your money and for a loss, you lose your entire bet \*\*

Game	Bet	Outcome	Net Win/Loss
1	\$100	Loss	-100
2	\$200	Loss	-300
3	\$400	Win	+100
4	\$100	Win	+200
5	\$100	Loss	+100

The system above is known as the Martingale system. You can read about it here:

[https://en.wikipedia.org/wiki/Craps#Martingale\\_system](https://en.wikipedia.org/wiki/Craps#Martingale_system)

You would like to test this hypothesis on a game of craps. The rule for craps can be read here:

<https://en.wikipedia.org/wiki/Craps>.

In this assignment, you will only use the pass line bet. The code for simulating a pass line bet in R is provided in the file `craps.game.R`. So, all you need to do is implement the Martingale system using following conditions:

- You start with \$1000 initial money
- For each round, you play until you run out of money or for 10 games.
- Run simulations for 10 such rounds and report your results by filling the table below:

Round	Ending Amount	Number of games played
1	\$2000	10
2	...	...
...		

Do you think this system works? Explain in 1-2 sentences.

What to turn in:

- Your code in R or any other language. If you use another language, you would need to write the craps code in that language.
- The table as shown above and your explanation (It can be in a txt file, Word doc, or PDF)

## **Question 2 (30 points)**

2. In terms of stock market analysis, one of the most popular and recommended strategies is that of *buy and hold* [see [https://en.wikipedia.org/wiki/Buy\\_and\\_hold](https://en.wikipedia.org/wiki/Buy_and_hold) ]. This is an investment technique that recommends you to buy a stock and hold for a long time. Usually, this means till you reach retirement age. The second option is to buy and sell based on some technical indicator that recommends entry (buy) and exit (sell) points for the stock. For example, see <http://www.investopedia.com/articles/active-trading/052014/how-use-moving-average-buy-stocks.asp>. In this part you will compare these two strategies.

In this question, you will explore these two strategies using the questions below:

a. Using the quantmod package in R, download the price data for the following stock symbols starting from 1<sup>st</sup> Jan 2000 to present.

- DJIA (Dow Jones Industrial Average)
- SPY (S&P 500)
- AAPL (Apple Corp)
- BAC (Bank of America)
- NFLX (Netflix)
- PCLN (Priceline)
- AMZN (Amazon)

**\*\* Hint: You can use the getSymbols function to get the stock data \*\***

b. Plot the chart for each of the the stocks and overlay the value of Simple Moving Average 200 i.e. SMA(200). Include the plots in your report

**\*\* Hint: You can use the chartSeries function and addSMA function for overlay\*\***

c. In this last part, you will use the SIT toolbox to compare the above-mentioned trading strategies – Buy Hold and SMA Crossover. The code in R for this is provided in the file trading.R. Be sure to install the required packages by uncommenting the top lines and also change the stock symbols.

Generate the comparison chart for the above mentioned stocks:

Stock	Buy Hold		SMA Crossover	
	CAGR	Performance	CAGR	Performance
DJIA				
...				

Write a brief paragraph explaining which strategy you would choose and why?

What to turn in:

- Your code in R or any other programming language. You should include all the files that will make your code work.
- A document with the plots for part b, the table asked in part c and the paragraph asked in part c.