# Assignment #1 Instructions

YOU MAY WORK TOGETHER, BUT MUST SUBMIT SEPARATE ASSIGNMENTS. SHOW YOUR WORK BY INCLUDING ALL COMPUTATIONS IN THE SPREADSHEET AND EXPLAINING YOUR REASONING IN THE REPORT. THE REPORT DOES NOT NEED TO BE AN ESSAY, BUT WRITE COMPLETE SENTENCES AND CLEARLY LABEL YOUR SECTIONS AND WHAT QUESTION YOU ARE ADDRESSING. PLEASE LABEL ANY NEW SHEETS YOU ADD TO YOUR WORKBOOK APPROPRIATELY.

Please download the spreadsheet for assignment 1.

You are an analyst for an internet retailer. The retailer has been testing out 10 different ad campaigns over the course of a year. They have chosen to place these ads on a search engine. In order to do so, they randomly select a single ad to run for each day. The retailer made bids per impression of the ad. The data shows how much they ended up paying per 1000 impressions (CPM). There are also many other columns that tracks impression to click conversion, purchases leading from clicks, and units/dollar amounts purchased. In addition to this data, you have also collected data on web traffic and purchases not generated by the ad. Assume the total purchase dollar amounts are after costs (except for the cost of advertisement - or alternatively assume the costs are 0). You are asked to evaluate the 10 different advertisement campaigns.

Instructions:

First, notice there are 2 sheets. The first sheet is the raw data. The second sheet is a copy of the first sheet for your convenience, it is provided so you can select data without accidentally messing up the raw data. Use it as you wish.

Second, notice that there are a bunch of empty yellow cells on the first sheet. Please fill in these cells using formulas to compute the appropriate measures. Note that there is a smaller table that begins on U1, there are empty cells here as well. In this smaller table, use conditional formatting to automatically bold the text in the row containing summary statistics of the best ad in terms of ROI.

Third, using pivot tables do the following:

1. Construct a table where the rows are dates, the columns are advertisements, and the computed cells are the conversion rates (Prob(click)).
   1. Using the pivot table, compute the daily weighted average conversion rate by ad and also the standard deviation of the conversion rate by ad. (Show in spreadsheet)
   2. Construct a 95% confidence interval for all the ads’ click through rates. (Show in spreadsheet)
   3. Is the best ad statistically better than the worst ad? Explain why.
   4. Is the best ad statistically better than the second best ad? Explain why.
   5. Is the worst ad statistically worse than the second worst ad? Explain why.
2. Do all of # 1 again, except, instead of conversion rates, summarize impressions.
3. Construct a table where rows are dates and the columns are days of week and the computed cells are the impression counts.
   1. Using the pivot table, compute the average and standard deviation of impressions by day of week. (Show in spreadsheet)
   2. Construct the 95% confidence interval for each day of week. (Show in spreadsheet)
   3. Are any of the days of week statistically different from the others? If so, which days? Why or why not?
4. Construct a table where the rows are months, the columns are ads, and the computed data is probability of click.
   1. Create a line graph (you can use the pivot graph option) that shows all 10 ads’ click probability over the 12 months.
   2. Describe what kind of insights this graph shows.

Fourth, regression Analysis:

1. Write a regression equation that will allow you to predict the number of ad impressions per day. (read the questions below before specifying your equation)
   1. What variables did you include as predictors, why did you pick each of these variables?
   2. Estimate this equation and report your regression results.
   3. From this regression, can you make the some conclusions about the effect of weekends (Friday and Saturday) on the impressions an ad will receive?
   4. Does traffic through direct visits affect the impressions of ads?
2. Write a regression equation that will allow you to predict the conversion rate of ads. (read the questions below before specifying your equation)
   1. What variables did you include as predictors, why did you pick each of these variables?
   2. Estimate this equation and report your regression results.
   3. From this regression, what can you conclude about the various ads?
   4. Can you say anything about the time trends for the effectiveness of ads? If so, what are some possible reasons that ads may gain or lose effectiveness over time?
   5. Does traffic through direct visits affect the conversion rates of ads?

Fifth: Write a brief recommendation on how to proceed with your online advertisement campaigns. What can you conclude about the effectiveness of the various campaigns? How should you proceed in using these campaigns? Are there any potential problems with proceeding the way you recommend that the data cannot speak to?