Assignment #3 (G1)
Regression
Group Assignment

This assignment is a team assignment and has two parts. The regressions required for the analysis in both questions must be done in Python (not Excel), but the data construction for Q1 and the optimal price calculations in Q2b may be done in Excel if you prefer. You only need to upload one copy of the write-up per team. The uploaded file should be a pdf file that includes your code (but not the data) and should have your team identifier. **Make sure you list each of your team members on the cover page.**

Please read each question carefully and note that you will have to make assumptions (and compromises) along the way. All relevant information pertaining to this assignment is in this document.

1 Estimating Firm β's

The Capital Asset Pricing Model (CAPM) relates the expected excess return on an asset (R) to the expected excess return on the "market portfolio" (R_m). R_f is the return on a safe, risk-free investment:

$$R - R_f = \alpha + \beta(R_m - R_f)$$

where α (the intercept) is assumed to be zero in theory, but you should include it in your estimation. You will estimate β (the slope) for a few individual stocks. In this regression, $(R - R_f)$ is the dependent variable and $(R_m - R_f)$ is the independent variable.

Collect data and construct $(R - R_f)$ and $(R_m - R_f)$. Then compute estimates of β (one for each firm) for the following firms for the calendar years 1990 through 1999:

1. Home Depot (HD)
2. Apple (AAPL)
3. Verizon (VZ)
4. Cisco (CSCO)

As a write-up, fill in the tables below and use about one page of text to discuss:

1) Rationale for your choice of indexes and data frequency
2) What your findings / interpretations are on E(R), the estimates $\hat{\alpha}$ and $\hat{\beta}$, and R^2
3) Whether or not you reject the null hypotheses $H_0: \alpha = 0$ and $H_0: \beta = 0$ and at which significance level

Data Collection Choices:

Variables	Data source and/or index used	Data frequency used
R: Return for each stock HD AAPL VZ CSCO		
R_f: Return on a safe, risk-free investment		
R_m: Return on the "market portfolio"		

Estimates and p-values:

Construct	HD	AAPL	VZ	CSCO
E(R): Average Return				
$\hat{\alpha}$				
p-value for $H_0: \alpha = 0$				
$\hat{\beta}$				
p-value for $H_0: \beta = 0$				
R^2				

<CAPM Model>

See discussion on p.122 of Stock and Watson and the video "CAPM example" posted on blackboard.

<Data Collection>

For each stock, you will have to download data, choose an appropriate market index, construct excess returns, and decide on a relevant frequency of data (daily, weekly, monthly...).

You might find the following website useful in collecting data:
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
Please note that the unit of the data from Fama/French website is in % terms.

Alternatively, data can also be collected from Yahoo Finance (finance.yahoo.com) or Google Finance (https://www.google.com/finance).

The "Treasury Yield" Yahoo shows is the "annual" return rates, so if we are interested in getting a monthly return rate, we will roughly divide it by 12.

On the other hand, stock prices (including S&P 500 prices, and Wilshire 5000 prices) are current prices (not return rates). So, when you calculate the returns, you will have to calculate the returns based on the data frequency you are using.

For example, if you are interested in monthly returns, then you will look at (today's price – last month's price)/last month's price to get the monthly return. If you are interested in daily returns, then you will look at (today's price – yesterday's price)/yesterday's price.

For Rf rates, please note that I am looking at 2005-2011 in the video example, and I am using 3.8% as a fixed Rf for simplicity (just to illustrate an example to students). However, the Rf rates will differ for the period 1990-1999 and will vary over time. I recommend looking into the treasury bill (e.g., treasury yield from Yahoo finance) or Fama/French's website (the link above) where they have the monthly/weekly/daily Rf historical rates. It's totally fine (and encouraged) to use varying Rf rates for the analysis.

For Cisco, Cisco only became public in Feb 1990. So, there's no data for Jan 1990. You would just run the regression for the period where the data is available.

<Grading>
You will not be graded based on whether you have gotten the "right" metrics per se (i.e., there's no single right answer, for example, you can use either the closing price or the adjusted closing price in calculating the stock returns). Rather, I will be looking for "can you reasonably explain how you have chosen/constructed the metrics for this exercise?" If you can show your reasonable / consistent thinking process regarding how you went about collecting / constructing the dataset and conducting the analyses, that should be sufficient.

2 Computing Price Elasticities (and optimal prices)

Using the data posted on the class website (rfj_data.csv) run three simple regressions of the following type

$$Q_t = \alpha + \beta p_t + u_t$$

for each of the three brands in the data (Tropicana, Minute Maid and Private label). In this regression, Q_t (ounces sold) is the dependent variable, p_t (price) is the independent variable, and u_t represents the residual (or prediction error) term. Also, α and β are the intercept and the slope parameters, respectively, in the regression model.

(a) Using the regression results, compute own price elasticity at (average P, average Q) for each brand.

Price elasticity, ε, can be expressed as:

$$\varepsilon = \frac{\partial Q}{\partial p}\frac{p}{Q}$$

where $\frac{\partial Q}{\partial p}$ is the slope coefficient from the regression result.

(b) Assume that the per unit cost of producing an additional unit is the same across brands and is 1 cent per ounce. Based on the data provided and the estimated demand equation from above, compute the optimal price for each brand (that is, find the price point at which the profit is maximized). For this

problem, you can ignore the error term u_t when writing down the profit function. The attached "Intro to Excel Solver.pdf" or “Intro to Optimization in Python.pdf” should prove useful for setting up this problem. Students can also solve this problem analytically using calculus if you prefer.

Describe how you went about answering questions (a) and (b) and comment on your findings. Feel free to include the regression outputs if you would like. However, displaying the screenshot of the regression outputs per se will not earn points. You will need to discuss in words and comment on your findings.