ACCT421 ANALYTICS FOR VALUE INVESTING
Term 2 2020-21
(Prof Andrew Lee)

INDIVIDUAL ASSIGNMENT

(Due 11:59 p.m. Friday 19 Feb 2021)

QUESTION 1

Prior studies have suggested that stock prices tend to *rise* in the month of January and *fall* in the month of September relative to other months of the year (an effect that has come to be known as the January effect) – i.e. monthly stock returns for the month of January are, on average, significantly higher or more positive than other months, while those for the month of September are significantly lower or more negative than other months.

Required:

a) Collect data on monthly S&P500 Composite Index (ticker: ^GSPC) from December 1980 to December 2020 (40 years or 480 months). Any legitimate data source is acceptable (e.g. Yahoo Finance, CRSP etc).

b) Write Python code to:

(i) Compute monthly stock returns from the data of index prices over the period.

(ii) Compute descriptive statistics of those monthly returns.

(iii) Test the hypotheses that the returns for the month of January are higher or more positive, while those for the month of September are lower or more negative, than those for other months, using the following statistical tests:

- *t*-test of difference in mean returns between January and non-January months;
- *t*-test of difference in mean returns between September and non-September months;
- an OLS regression of monthly S&P500 returns as dependent variable and two dummy independent variables as follows:

$$R_t = \beta_0 + \beta_1 D_{Jan} + \beta_2 D_{Sep} + \varepsilon_t$$

where:
R_t = return for month t
dummy variable D_{Jan} = 1 if the return month is January and 0 otherwise
dummy variable D_{Sep} = 1 if the return month is September and 0 otherwise.

In your regression above, be sure you use heteroskedasticity-consistent standard errors as well as autocorrelation-consistent standard errors in your statistical tests of the regression coefficients.

Note: To determine the month from a '*Date*' object, you may need to use the Python method '*month*' from the '*datetime*' library – e.g.

```
import datetime as dt
dataset['month'] = dataset['Date'].dt.month
```

c) Briefly explain your statistical results and their implications for the hypotheses above. Your answer should also include an explanation of what the intercept (β_0) of the above regression implies.

QUESTION 2

Suppose you postulate that a firm's future operating profitability is positively correlated with its current operating profit margin and efficiency in utilization of its operating assets (as measured by asset turn).

Required:

a) Collect *annual* financial statement data on a sample of U.S. companies from *Compustat* for the period Dec 2003 to Dec 2019 (16 years).

   To keep your dataset manageable, exclude financial institutions ('gsector' equals 40) and exclude firms with market capitalization less than US$1 billion.

b) Test the hypothesis that operating profitability for the year is positively correlated with the previous-year operating profit margin and operating asset turnover, using the three regression models below:

   (i) *Pooled* OLS regression of current-year profitability against previous-year operating profit margin and operating asset turn, with heteroskedasticity-consistent *t*-tests on the regression coefficients, of the following form:

$$\%\Delta EBIT_{i,t} = \beta_0 + \beta_1 \frac{EBIT_{i,t-1}}{Rev_{i,t-1}} + \beta_2 \frac{Rev_{i,t-1}}{NOA_{i,t-1}} + \varepsilon_{it}$$

   (ii) *Year-by-year* OLS regression of the above form, with *t*-tests on the time-series of regression coefficients.

   (iii) *Pooled* logistic regression of current-year profitability indicator variable ($D_{i,t}$) against previous-year operating profit margin and operating asset turn, of the following form:

$$D_{i,t} = \beta_0 + \beta_1 \frac{EBIT_{i,t-1}}{Rev_{i,t-1}} + \beta_2 \frac{Rev_{i,t-1}}{NOA_{i,t-1}} + \varepsilon_{it}$$

   where (parentheses denote *Compustat* acronyms for corresponding items):

   $EBIT_{i,t}$ = Earnings before interest and tax (EBIT) of firm i for year t;
   $Rev_{i,t}$ = Total revenue (REVT) of firm i for year t;
   $NOA_{i,t}$ = Net operating assets of firm i at the end of year t;
   $\%\Delta EBIT_{i,t} = \left(EBIT_{i,t} - EBIT_{i,t-1}\right)/EBIT_{i,t-1}$;
   $D_{i,t} = 1$ if $\%\Delta EBIT_{i,t} > 0$, and 0 otherwise.

   Net operating assets = Total assets (AT) excluding financial assets *minus* Current liabilities (LCT) excluding current financial liabilities.

ASSIGNMENT FORMAT AND DELIVERABLES

The following should be submitted:

a) A report (in Word or PDF format) – please submit your answers to all questions in one single file.
b) Your Python script (in py, txt or Word format) – please submit individual Python script for each question.

Please preface every filename with your name (e.g. AndrewLee-Assignment.docx, AndrewLee-Q1python.py etc.).

CSV files of your datasets are **not** required to be submitted unless individually requested.

Keep your report short and write concisely. There is no need to write any introduction, literature review, summary or conclusion. Go straight to the point. A nice cover page is not required but be sure you indicate your full name and student ID prominently on the front page of your report.

Your report should include/discuss the following:

(i) Sources of your data
(ii) Tests conducted
(iii) Results obtained (showing all relevant statistical results in appropriate tables and/or charts).

Deadline for submission

All deliverables are to be submitted online via eLearn by **11:59 p.m. Friday 19 Feb 2021 (Week 6).**

Grading criteria

This assignment constitutes **10%** of total course marks.

The grading criteria is outline in the table below:

Deliverables	Attributes assessed	Weight (% of total Assignment marks)
Python code	• Correctness of coding. • Use of appropriate functions/methods.	**Not more than 15%**
Research and statistical methodology (including interpretation of results)	• Appropriateness of research methodology. • Appropriateness of statistical models and methods/tests used. • Correct interpretation of results. • Discussion and treatment of any biases in your study.	**Not less than 75%**
Significance of empirical results	• Whether empirical results of tests are statistically significant. • Whether investment strategy is profitable and makes money.	**Not more than 10%**

Reminder

You should submit only work done on your own, and not in collaboration with anyone else. All work submitted will be checked with plagiarism detection software. Plagiarism is a serious offence and will be severely penalized.
