Hand-In Assignment #1
Descriptive Statistics of Three Asian Currencies:
January 1996 – June 1999

Fall 2001
B6014: Managerial Statistics

Preliminaries:

- Due date: Thursday, August 30, 3:00pm.
- What you need from me: The spreadsheet asian_fx.xls (available on the course web page).
- What I need from you: A brief report addressing the specific items enumerated below. You may work in groups of up to three. Each group should hand in a single report.

Suggested Schedule: The assignment has six parts, as explained below under “The Assignment.” I suggest working on them as follows.

- Monday: familiarize yourself with the data; do parts 2 and 3
- Tuesday: do parts 4 and 5
- Wednesday: do parts 1 and 6

This schedule aligns the parts of the assignment with material covered in class. It also leaves plenty of time to ask questions in case you run into any difficulties.

Purpose of this assignment: This assignment illustrates the use of some descriptive statistics in summarizing data. It is also designed to illustrate some of the capabilities of spreadsheets in working with data. The assignment uses data on exchange rates (relative to the US dollar) for the Hong Kong dollar, the Korean won, and the Thai baht. We will be comparing characteristics of these exchange rates before and after the dramatic events of the summer of 1997.
The Data: Open the spreadsheet asian_fx.xls. On the sheet labeled “Raw Data” you will see exchange rates for the Hong Kong dollar, the Korean won, and the Thai baht from January 2, 1996 to June 30, 1999. These are expressed as number of US dollars per unit of foreign currency; thus, on January 2, 1996, one Hong Kong dollar cost 0.12933 US dollars. (You can think of these numbers as prices of assets which happen to be currencies — they could just as well be the prices of a stock or commodity.) To the right of the daily exchange rates are the percentage changes in the exchange rates from each day to the next. For brevity, we will refer to these as returns. However, these returns do not reflect any interest that would have been earned by holding the foreign currency.

Scroll down the data and feel overwhelmed by it. Your task will be to summarize these columns of numbers in an informative way.

You will be comparing returns on the currencies before and after the Asian crisis of 1997. Figure 1 plots the evolution of the Thai baht and Korean won exchange rates from January 1996 to June 1999 and shows the dramatic decline in value of these currencies in 1997. It is difficult to say exactly where we should draw the lines separating “before” and “after” the crisis. I have somewhat arbitrarily decided to call everything prior to May 1, 1997 “Pre-Crash” and everything after January 1, 1998 “Post-Crash.” Thus, the shaded area in Figure 1 indicates the crash period. The sheet labeled “SplitData” groups the returns into three time periods according to these dates. Except in item 5, below, you only need to use the data labeled “Pre-Crash” and “Post-Crash.”

The Assignment: Your report will consist of a written summary supported by tables and graphs. The summary should be the first item in the report, but don’t try to write it until you have produced the tables and graphs. For examples and suggestions of how to get what you want in Excel, look at the sheet (in asian_fx.xls) labeled “Summaries” and the tips below. The results produced on the Summaries sheet apply to the full data set, whereas you will be
looking separately at the Pre-Crash and Post-Crash periods.

1. Write one or two paragraphs comparing Pre-Crash and Post-Crash properties of the exchange rates. Be specific and informative—interpret the summary measures you calculate. Make reference to tables and graphs to justify your assertions. Focus on the most salient features of the data.

2. Produce a table of summary statistics that looks something like this:

<table>
<thead>
<tr>
<th></th>
<th>Pre-Crash</th>
<th></th>
<th>Post-Crash</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>USD/HKD</td>
<td>USD/KRW</td>
<td>USD/THB</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std Dev</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Make good histograms of the USD/THB returns Pre-Crash and Post-Crash.

4. Examine the shape and normality of the Pre-Crash and Post-Crash returns by calculating the skew and kurtosis of each currency’s returns in the two periods.  

Optional: Another way of investigating skew and kurtosis is to calculate, for each currency in each period, what percentage of the returns lie more than 2 standard deviations below the mean, more than 1 standard deviation below the mean, more than 1 standard deviation above the mean, and more than 2 standard deviations above the mean. See the Excel tips below and the block labeled “Shape” on the “Summaries” sheet for an explanation of how to do this. This is a bit more involved than the rest of the assignment so it’s strictly optional — there’s no penalty for not doing it. I mention it because it is interesting statistically and as an illustration of spreadsheet capabilities.

5. Report the mean, median, standard deviation, skew, and kurtosis for the weekly and monthly average returns using the data in the sheets labeled “Weekly” and “Monthly.” Be sure to comment on differences between the daily, weekly, and monthly numbers.

6. Make a table of correlations for returns on the three currencies Pre-Crash and Post-Crash. (A table of correlations for the full time series appears in the Summaries sheet.) Make a scatter plot of returns on the Thai baht versus returns on the Korean won Pre-Crash. Repeat this with the Post-Crash data.

Excel Tips: You may want to look at the sheet in asian_fx.xls labeled “Summaries” for suggestions on doing some of the calculations. Here are a few additional tips:

- The functions AVERAGE, MEDIAN, STDEV, MIN, MAX, SKEW, KURT produce exactly what their names suggest. (STDEV produces a sample standard deviation; for a population standard deviation we would use STDEVP. We will deal with this distinction later in the course.)
In the spreadsheet `asian_fx.xls`, I have named most of the important data ranges. For example, on the Raw Data sheet, cells J4:J833 have the name USD_HKD. Thus, on the Summaries sheet, I use, e.g., the formula AVERAGE(USD_HKD) in cell D5, rather than the formula AVERAGE(J4:J833). Using range names this way makes it easier to enter and read formulas. (To name a range, highlight it and then type the name you want to use in the Name Box at the left end of the formula bar. If a range has already been named, the name will appear in the Name Box when you highlight the range. Consult Excel’s Help for details.)

Under the Tools menu you should see Data Analysis. If you don’t, then select Add-Ins under Tools and check the box for Analysis ToolPak. You should now see Data Analysis on the Tools menu. (This may require re-starting Excel.)

To produce a histogram, you can use the Histogram tool under Tools/Data Analysis. First, though, you need to create a range of bin values; these will determine the boundaries of the bins for your histogram. In choosing the bins, it is helpful to look at the minimum and maximum values of the data. However, if your data contains some extreme outliers you may want to omit these in order to make the center of the histogram more informative. In deciding how many bins to include, a very rough rule of thumb is to choose the number of bins on the order of the square-root of the number of observations. For example, with 340 observations, you would probably want to have about 15–30 bins. (You definitely want more than 5, and you definitely want fewer than 100.) Inevitably, producing an informative histogram takes some trial and error.

After you have created a range of bins, in the Histogram dialog box (under Tools/Data Analysis) you can check the box for Chart Output to get Excel to produce the histogram itself. You will want to re-format the Excel default chart to improve legibility.

There are several ways of counting the number of observations that fall, e.g., 2 standard deviations above the mean, as mentioned in part 4 of the assignment. Here is one method that seems reasonably convenient and is implemented in `asian_fx.xls`. Suppose the original data $X_1, X_2, \ldots, X_n$ has mean $\bar{X}$ and standard deviation $\sigma$. From these observations create a range of standardized values

$$Z_i = \frac{X_i - \bar{X}}{\sigma}, \quad i = 1, 2, \ldots, n.$$  

Notice that $X_i$ is more than two standard deviations above the mean (i.e., $X_i > \bar{X} + 2\sigma$) precisely if $Z_i > 2$. In fact, $Z_i$ tells us exactly how many standard deviations away from the mean $X_i$ lies. If the standardized values are in the range A1:A100, then the command COUNTIF(A1:A100,”>2”) counts the number of values in the range A1:A100 greater than two. Similarly, COUNTIF(A1:A100,”>2”) / COUNT(A1:A100) produces the proportion of values in the range greater than two. Consult Shape section of the Summaries sheet in `asian_fx.xls` for a concrete illustration.

To produce a table of correlations, you can use the Correlation tool under Tools/Data Analysis. For a single correlation, it is more convenient to use the function CORREL.