ME 101
Computer Demonstration – CD 1

*Introduction to*

*Engineering Data Analysis*

*Using Spreadsheets*
Computer Demonstration

CD 1 – Introduction to Spreadsheets

Purpose: To introduce the spreadsheet application program as an engineering tool.

Goals:
1. Open a spreadsheet program and create a new spreadsheet.
2. Enter and organize data into a spreadsheet.
3. Create charts and plot data using a line, pie or XY charts.

Assignment: See the end of this handout for the problems for this CD.

Introduction to Spreadsheets as an Engineering Tool

Computer spreadsheet programs, such as Visicalc, Excel, Lotus 1-2-3, and gnumeric, have become indispensable as computer tools. Originally used for financial work, spreadsheet applications are now used across a whole range of fields, engineering notwithstanding. Spreadsheets can be used for both data analysis and numerical simulation, including the creation of simple graphs. Spreadsheets are often limited in their power, but they are a simple tool for use in many engineering applications and will prove to be helpful throughout your engineering career. To get started, make sure your PC has a spreadsheet program, preferably Microsoft Excel, since that is the program the CD lessons are derived from.

A version of Excel should be on most of the UK campus PCs found in any of the campus microlabs. Suggested labs include the computer labs on the ground floor of the RGAN Building and on the second floor of the Civil Engineering building.

The Spreadsheet Environment

First, open up Microsoft Excel. This should be under the Start Menu and it can be in several places, either alone or with other MS programs. If you can't find it, use the Run command under the Start button and type Excel.exe as shown below. This should bring up Excel if it is located on your machine.
How a Spreadsheet Works

A computer spreadsheet is based on the paper version of the same name used by accountants for many years. The spreadsheet is divided into rows and columns and this creates a gridded pattern of cells. Each cell is an individual entity and is denoted by where it is located, that is, the row and column designation.

An example is shown below:

<table>
<thead>
<tr>
<th>Name</th>
<th>Calculus</th>
<th>Physics</th>
<th>Shop</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.Einstein</td>
<td>75</td>
<td>96</td>
<td>72</td>
</tr>
<tr>
<td>T.Edison</td>
<td>64</td>
<td>85</td>
<td>98</td>
</tr>
<tr>
<td>R.Feynman</td>
<td>99</td>
<td>94</td>
<td>83</td>
</tr>
<tr>
<td>B.Nye</td>
<td>81</td>
<td>88</td>
<td>93</td>
</tr>
</tbody>
</table>

The organization of the data (in this case, grades) allows one to easily see the relationships between them. With a quick glance, you can see that Edison got the lowest grade in Calculus, but the highest grade in shop. It is also easy to sum and average the data, thus finding the overall grade for each student or finding the average grade for each course.

This is the primary benefit of the spreadsheet - data organization and manipulation. The most useful thing about a computer spreadsheet is that unlike a paper spreadsheet, if you have to make changes, you don't have to recalculate the values yourself. The computer does all the work for you. This will come in useful if you make a mistake or if you want to see how changing values affect various parameters. Thus, a spreadsheet can also be used as a simulation tool.
The Spreadsheet Environment

When Excel opens, you should have an environment that looks much like the one shown on the previous page. Different versions of Excel have different appearances; the details are not important, however. The basic operations should be similar from version to version and even across platforms and different spreadsheet programs.

The following items are denoted above:

**Menu Bar** - The menu bar contains the list of menus for all Excel commands. The default commands are found here.

**Toolbar** - Various toolbars are available, but the standard and formatting toolbars are shown. These are shortcuts for commands found in the menus.

**Formula Bar** - This displays the value or formula of the selected cell in the workbook window. This can be a label (text), a number, or a formula.

**Column & Row Labels** - Headers for columns and rows.

**Sheet Tabs** - These are used to change from different spreadsheets; you can have multiple sheets in a single workbook.

**Spreadsheet/Workbook Window** - The current working spreadsheet.

Become familiar with each element by using the mouse and clicking on each in turn.

Data Entry

Data entry is accomplished several ways. The simplest is to click on a cell, which highlights both that cell and the formula bar, and then type what you want entered in the cell. Hitting “enter” will record the value in the cell and shift the active cell to the cell immediately below. A column letter and row number identifies each cell. For example, C4 is the 3rd cell across and the 4th cell down.

Data entry is demonstrated below.
Data accepted by Excel is alphanumeric in nature. That is, it can be either numbers or text and Excel automatically tells the difference. Excel denotes which of these two it thinks it is by flushing the entry left in the case of text or flushing it right in the case of numbers. (Entering a formula is accomplished by starting with an equal sign, = which will be covered later.)

The spreadsheet to the right contains both numeric and text cells. The entries in cells A1 and B1 serve as column labels or titles for the respective data sets below to make data identification easier. Excel automatically flushes the labels left while the data are flushed right. The column headers will be used when making graphs to make it easier to identify what you are trying to plot.

Creating Graphs

There are many types of graphs that one can create in Excel, but we will focus on only a few. The most often used graph in engineering is the XY plot, where two variables are plotted against each other, but we shall also see use of the line plots, pie charts and others as well. Excel calls any type of plot, whether it is a bar graph, surface plot or any other method of visually displaying data, a chart.

To insert a chart in Excel, click on Insert tab on the menu bar, and then select chart type: Insert >> Chart

1. **Highlight the values** you wish to plot - this includes column titles if any. (If you don't select the region you want to plot, you'll be asked to enter it by hand. Hit Cancel and start over. It is much easier this way. If you did select the region you want to use and you are asked to verify, hit the Next button.)

2. **Select the insert tab and then type of chart** you wish to use. You can choose from the following chart types: Column, Bar, Line, Pie, XY (Scatter), Area, Doughnut, Radar, Surface, Combination, and others.

3. **Select the type of subchart** you wish to use. This will determine the details of your chart, such as whether to use points or lines (or both), linear or logarithmic axes, filled or unfilled areas, etc. This is dependent upon what message your plot needs to convey.

4. **Finish** the details. At this point you may enter the chart and axis titles, change tick marks, grid lines, etc. In Excel, you can modify the chart format by clicking on the chart and then clicking on the “Chart Tools” tab on the menu bar.
The above procedure will vary from version to version, but the general process stays the same. Select the region and choose your chart. The best way to do this is through practice.

**Creating a Line Chart**

The *line chart* plots values in the spreadsheet against row numbers or alternatively, your own text labels. This is different from the XY chart because the x-axis doesn't scale; that is, the values on the x-axis are equally spaced regardless of their respective values.

In the example below, we wish to show graphically the number of different launch vehicles (i.e., rockets) for various nations. The data is tabulated below and we want to enter it into Excel.

<table>
<thead>
<tr>
<th>Country</th>
<th>China</th>
<th>ESA</th>
<th>India</th>
<th>Israel</th>
<th>Japan</th>
<th>USA</th>
<th>Russia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>7</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>15</td>
<td>13</td>
</tr>
</tbody>
</table>

First, enter the data just using the number of launch vehicles versus row number. Your spreadsheet should resemble the spreadsheet below.

Now, insert a line chart using the steps outlined earlier. First, highlight the cells A1 through A8 using the mouse. Your final chart should look something like the one shown here with row numbers on the x-axis and number of launch vehicle systems plotted on the y-axis.

Since the purpose of a graph is to convey information, this plot doesn't do a very good job. While it might work if you want to examine the relationship quickly and on your own, it wouldn't do a very good job at conveying the information to others.
To improve this plot, add the x-axis titles to the data. You can reenter the data if you like, but an easier way would be to add a single column so you can add the country names to the left of the data. To do this, click on the A1 cell and insert a column using the Insert >> Columns command on the menu bar. The information in the A column should be shifted to the right under B and the A column should now be empty.

Creating a Line Chart and Making It Better

Now you can enter in the label information in the A column. Once finished, highlight the cells A1 through B1 and click on the bold B in the toolbar (or Ctrl B). This will bold face the column headings and will make your spreadsheet data easier to identify.

Your spreadsheet should now resemble the one to the right.

Replot this data. First, highlight the cells from A1 to B8 and insert a line chart. This time, add some x- and y-axis labels as well as a plot title. When finished, double clicking on different parts of the chart allows you to format that particular aspect of the chart.

Experiment until you come up with your own chart scheme.

This is the final line chart of the data. Compare it to the first unformatted attempt and note how much easier it is to interpret the data.
Creating Pie Charts

*Pie charts* are popular in the business community because of the way to display data as part of a whole. As an engineer, you will often need to communicate with people who are not engineers, such as managers, marketers, and prospective customers. Thus making data easy to visualize will be an asset.

Using the same set of data for your line chart, make a pie chart using the same steps. You can choose to make a 2-D or 3-D pie chart; the 3-D chart conveys the same information but makes it a little flashier. It is best not to get too complicated if you're really interested in getting a point across. You can choose from several subsets of pie charts based on how you want the data labeled. It makes no difference which one you choose because you can always change it later through formatting commands.

The pie chart to the right shows the percentages of the world launch vehicle systems. Both the patterns and colors have been edited to customize it. The labels show the percentages as automatically calculated by Excel from the number of systems for each country divided by the total number of systems.

Your final pie chart can be edited to change the labels, color schemes, patterns, and even how the slices are positioned. Experiment with your pie chart always asking yourself the question, "what message am I trying to get across?"
Creating XY (Scatter) Plots

The XY or scatter plot is the most common type in engineering. XY plots are used to compare the relationships between two variables, namely an independent variable (usually on the x-axis) and a dependent variable (usually on the y-axis).

Consider the following equations,

\[ y = 2x \quad \text{or} \quad y = x^2 \]

In both cases, \( y \) is a function of \( x \); this is expressed mathematically as \( y = f(x) \). If you change the value of the independent variable \( x \), you are also changing the value of the dependent variable \( y \). (In reality, both values are dependent upon one another, that is, we could easily say \( x = 0.5y \) for the first equation. But we are assuming that the value of \( x \) is directly under our control while the value of \( y \) is not.)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Volts</td>
<td>Amps</td>
</tr>
<tr>
<td>2</td>
<td>6.97</td>
<td>0.051</td>
</tr>
<tr>
<td>3</td>
<td>5.96</td>
<td>0.044</td>
</tr>
<tr>
<td>4</td>
<td>4.95</td>
<td>0.038</td>
</tr>
<tr>
<td>5</td>
<td>3.98</td>
<td>0.032</td>
</tr>
<tr>
<td>6</td>
<td>3.03</td>
<td>0.025</td>
</tr>
<tr>
<td>7</td>
<td>1.91</td>
<td>0.018</td>
</tr>
<tr>
<td>8</td>
<td>1.02</td>
<td>0.012</td>
</tr>
<tr>
<td>9</td>
<td>0.5</td>
<td>0.008</td>
</tr>
<tr>
<td>10</td>
<td>0.2</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Now consider Ohm’s law;

\[ V = IR \quad \text{or} \quad I = V/R \]

This equation states that the value of the voltage potential across a circuit is dependent upon the current in the circuit, times the circuit’s resistance. Let us assume that \( R \) is constant. Then, \( V \) varies linearly with \( I \) and vice-versa. Given a resistor of constant value in the laboratory, we can alter the voltage across that resistor and measure the current. This will allow us to test the validity of Ohm’s law.

Enter in the experimental data as shown above in your spreadsheet. Both voltage and current have been measured with voltage as the variable input; thus, it should be on the x-axis.

Highlight the region from A1 to B10 and use the Chart Wizard to create a chart. This time select XY (Scatter) plot. When selecting a type of XY plot, make sure you choose the one with no lines; that is, choose a chart with points only.

Creating XY Plots from Experimental Data

Various formats of XY charts can be used. Since we are interested in plotting experimental data, we want to refrain from using lines to connect the points since this may lead to false interpretations of the data. Thus we would like to plot points only, as given by the various formats below.
In this instance, selections 1, or 4 would be acceptable. Option 3 or 5 actually replaces the points with a line fit. Never use this option, especially when graphing experimental data!

The data is shown plotted below with an XY plot. With the naked eye you should be able to make out that the data is not quite linear. There are slight variations in the data as \( V \) increases and there appears to be a large nonlinearity near the origin.

![Ohm's Law Measurements](image)

### Making a Plot with Multiple Data Sets

A XY plot is shown below with 2 different data sets. The points are the measurements given on the previous page while the line is a data set formed from the Ohm's law equation using

\[
I = \frac{V}{R} \quad \text{where} \quad R = 137 \text{ Ohms}
\]

(Varying the value of \( R \) will shift the slope of the curve up or down.)

To create the plot below, enter another set of data called **Theory** in column C using Ohm's law to fill in the value of the current. If you know how to use formulas in Excel, you can use them here; otherwise, enter in the data by hand. Create a plot similar to the one below, using the formatting commands to change the data points to a line. You can format almost anything on this plot by double clicking on the object in question.
Note that it was a little clunky entering in the data for the theory by hand. (Of course, the simplest way to do it would be to just enter the end points. It is a straight line after all.) Since these values are from a formula, it would make more sense for us to enter the formula with variables and let Excel calculate the actual values. This too is easy to do.

**Formula Entry**

Entering a formula into an Excel spreadsheet is very simple. You would enter it as you would any other item (text or data) by clicking on the cell and then typing, but you start with an equals sign, =. This tells Excel that what follows is a formula.

For example, if you want to multiply two columns together, say A1 and B1, you would enter into the cell of your choice, =A1*B1, as shown below.

Other basic math operations include division /, exponentiation ^, addition +, subtraction -, and percentages %. More complex operations are available using functions such as squareroot SQRT() where the cell you wish to operate on is placed within the parentheses.
If you want to multiply all of the values in the two columns A and B, you have two choices. One is to enter in the values by hand, one cell at a time; very tedious if you have many cells. The other is to use the fill command or drag the cell information. The fill takes the values in the first cell and fills down the formula into the cells you have highlighted. It automatically changes the formula cell indicators to give the correct values. This operation is shown below. Try it.

Examine each cell in column C. Note how the fill command changed each row number to the correct one. Try the same procedure, except highlight cells C1 through G1 and use the fill right command. Now examine those cells. What happened? The column letters changed instead. This would be useful if you were summing the values of many columns, for example.

Lastly, if you want to keep certain rows or columns constant in a fill command, preface the value with a $ sign. For example, A$1 will change A but not 1 and $A1 will change 1 but not A while $A$1 will fix both indicators. Experiment with this to see how it works.

Arranging Data into Bins for Histogram Plotting

We wish to analyze the data to see if we can determine if there is a trend in the data. To do this we must first arrange the data into bins. Excel has a function that will help us with this. It is called Frequency.

Frequency is an array function. You've already used an array function, though you may not realize it. The Chart function is an array function because it takes an array of numbers, or several arrays as its input. The arrays were represented in Excel as columns of data. The Frequency function takes two arrays as inputs, and also makes an array as its output.

The first array Frequency takes as input is the data we already have in the spreadsheet. The second array is called the bin array. The bin array tells frequency how many bins to make and what ranges to use. The frequency function assumes the data can have any value from negative infinity to positive infinity, and we wish to divide this range of possible values up into segments. If you wanted to divide it up into five segments you would need four numbers in your bin array.
Let us explore an example using the *frequency* function.

If your bin array had the numbers 0, 50, 100, 150 then the frequency would count how many data values were less than 0, how many were between 0 and 50, how many were between 50 and 100, how many between 100 and 150 and how many were greater than 150. That's five bins defined by four numbers.

So it is more accurate to say that the output of this frequency function has five bins as shown below with ranges less than or equal to 0; greater than 0 but less than or equal to 50; greater than 50 but less than or equal to 100; greater than 100 but less than or equal to 150; and finally greater than 150.

In order to perform the function in Excel on this example, perform the following steps.

2. Enter into the formula bar `= frequency (`. 
3. Highlight the data cells A2 through A7 and press `F2` then enter a `,` 
4. Highlight the bin array cells B2 through B5 and press `F2`. Enter a closing `)` 
5. Finish by pressing `CTRL SHIFT ENTER` at the same time.

Let us explore an example using the *frequency* function.

If your bin array had the numbers 0, 50, 100, 150 then the frequency would count how many data values were less than 0, how many were between 0 and 50, how many were between 50 and 100, how many between 100 and 150 and how many were greater than 150. That's five bins defined by four numbers.

So it is more accurate to say that the output of this frequency function has five bins as shown below with ranges less than or equal to 0; greater than 0 but less than or equal to 50; greater than 50 but less than or equal to 100; greater than 100 but less than or equal to 150; and finally greater than 150.

In order to perform the function in Excel on this example, perform the following steps.

2. Enter into the formula bar `= frequency (`. 
3. Highlight the data cells A2 through A7 and press `F2` then enter a `,` 
4. Highlight the bin array cells B2 through B5 and press `F2`. Enter a closing `)` 
5. Finish by pressing `CTRL SHIFT ENTER` at the same time.

Let us explore an example using the *frequency* function.

If your bin array had the numbers 0, 50, 100, 150 then the frequency would count how many data values were less than 0, how many were between 0 and 50, how many were between 50 and 100, how many between 100 and 150 and how many were greater than 150. That's five bins defined by four numbers.

So it is more accurate to say that the output of this frequency function has five bins as shown below with ranges less than or equal to 0; greater than 0 but less than or equal to 50; greater than 50 but less than or equal to 100; greater than 100 but less than or equal to 150; and finally greater than 150.

In order to perform the function in Excel on this example, perform the following steps.

2. Enter into the formula bar `= frequency (`. 
3. Highlight the data cells A2 through A7 and press `F2` then enter a `,` 
4. Highlight the bin array cells B2 through B5 and press `F2`. Enter a closing `)` 
5. Finish by pressing `CTRL SHIFT ENTER` at the same time.

Let us explore an example using the *frequency* function.

If your bin array had the numbers 0, 50, 100, 150 then the frequency would count how many data values were less than 0, how many were between 0 and 50, how many were between 50 and 100, how many between 100 and 150 and how many were greater than 150. That's five bins defined by four numbers.

So it is more accurate to say that the output of this frequency function has five bins as shown below with ranges less than or equal to 0; greater than 0 but less than or equal to 50; greater than 50 but less than or equal to 100; greater than 100 but less than or equal to 150; and finally greater than 150.

In order to perform the function in Excel on this example, perform the following steps.

2. Enter into the formula bar `= frequency (`. 
3. Highlight the data cells A2 through A7 and press `F2` then enter a `,` 
4. Highlight the bin array cells B2 through B5 and press `F2`. Enter a closing `)` 
5. Finish by pressing `CTRL SHIFT ENTER` at the same time.
Homework Assignment

Complete the following problems:

1. Create and print a copy of the Ohm's law graph that compares the measurements and the predicted curve completed in the tutorial.

2. Make a pie chart of the data below for the composition of the earth.

<table>
<thead>
<tr>
<th>Element</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxygen</td>
<td>45.6</td>
</tr>
<tr>
<td>Silicon</td>
<td>27.3</td>
</tr>
<tr>
<td>Aluminum</td>
<td>8.4</td>
</tr>
<tr>
<td>Iron</td>
<td>6.2</td>
</tr>
<tr>
<td>Calcium</td>
<td>4.7</td>
</tr>
<tr>
<td>Magnesium</td>
<td>2.8</td>
</tr>
<tr>
<td>Sodium</td>
<td>2.3</td>
</tr>
<tr>
<td>Potassium</td>
<td>1.8</td>
</tr>
<tr>
<td>Hydrogen</td>
<td>1.5</td>
</tr>
<tr>
<td>Titanium</td>
<td>0.6</td>
</tr>
</tbody>
</table>

3. Download a spreadsheet containing the winning times for the Kentucky Derby. Make a plot of this data of year versus winning time. The plot should be as meaningful and easy to comprehend as possible (XY scattered plot is a good choice). Make sure the graph is easy to read.

4. Plot the derby data as a histogram plot using the frequency function indicating the number of winners for winning times between 1:59 and 2:10. See the “histogram info.pdf” for additional information on histograms, if needed.

*This assignment is due at the beginning of class one week from the day it is assigned.

For each problem, print the graph as well as the spreadsheet you used to create it.