Using Excel to Create Graphs

Final Project for PRVM 881
Performance Improvement in Public Health

December 2010

Suzette Brotton
Laura Weir-Barton
Ruth Wetta-Hall
Ashley Moss

Table of Contents

Introduction 3
  Labeling 3
  Scales 3
  Data elements 4
Understanding Types of Data 5
Activating the Excel Analysis ToolPak Add-in 7
Formatting a Chart 11
Pie Chart 13
Histogram 16
Bar Chart 26
Pareto Chart 33
Run Chart 43
Control Chart 55
Radar Chart 67
Scatter Diagram 70
References 74

Introduction

Effective Performance Improvement (PI) is a data driven process. The use of data enables informed, evidence based decision making. The application of data provides means for initial analysis of how a process is currently performing, assists teams in priority setting, identifying gaps, developing measurable goals and assessing the impact of interventions designed to improve processes and outcomes.

Accurately constructed graphs are important to the PI process because they provide a visual means for rapid, easy synthesis of potentially large complex quantities of data. Quantitative information is communicated in a qualitative manner through the use of graphs (Curran, 1999). This is particularly important for team members who may be less skilled in the analysis of text data, tables or other more intricate formats. Trends, comparisons and patterns from large data sets are readily apparent in well organized, formatted graphs. However, graphs that are unclear, ambiguous or unnecessarily cluttered can make interpretation difficult and may misrepresent the data. All graphs must be constructed to present a clear, undistorted, efficient view of the true picture. To accomplish the graphic display must include clear labeling, scales and data elements (Klass, 2006).

Labeling

1. Avoid all capital, script, cursive or novelty font formats (Curran, 1999).
2. Minimize all text that does not contribute to clarification of what the data means.
3. The title should be a clear and concise description of what the data represents.
4. Avoid repetition in the axis labels. If an axis clearly represents the year or "percentage" has been used in the title, it is unnecessary to label the axis. Is an axis label is needed, further define the measure that was used, i.e. percentage, # per 1000 births.
5. Use a legend in all charts that contain more than one data series. The ideal placement for a legend is at the bottom of the graph to maximize the space used for the graphic elements (Klass, 2006).

Scales

1. The scales (either axis) identify the measure of the data.
2. Use regular intervals for scales.
3. Display time scales left to right on the x-axis

Data elements

1. If used, gridlines should be light enough to not distract from the main elements.
2. Use most of the ink in the display for the graph elements.
3. Avoid the use of three dimensional effects, they distort the data and make interpretation difficult.
4. If a bar chart is to contain more than 8-10 elements use horizontal bars versus vertical bars (Klass, 2006).
5. Use a different line style or color to clearly differentiate the different data series in a line graph.
6. When using color to represent different data elements there are colors that are not usually confused unless the reader is color blind. These are: black, gray, red, green, yellow, blue, pink, brown, orange and purple (Curran, 1999).

Understanding Types of Data

In order to choose the correct type of chart, it's necessary to understand the type of data element that you are analyzing. There are four types of data: categorical, ordinal, interval and ratio.

	Type	Definition	Example	Appropriate chart
Discrete	Categorical	data that can be separated into mutually exclusive groups (meaning that an observation in one category cannot also fall into another category)	• gender (male/female) • racial origin (Asian, Black, Native American, White) • insurance type (private insurance, Medicare, Medicaid, Champus, self-insured, uninsured)	Bar chart Pie chart Pareto chart Control chart
	Ordinal	A type of data containing limited categories with a ranking from the lowest to the highest, e.g. mild, moderate, severe. The distance between mild and moderate is not necessarily the same distance as between moderate to severe.	• Severity of disease (mild, moderate, severe) • Military rank (private, sergeant, lieutenant, captain, general) • Likert scales (1 = strongly disagree, 2 = disagree, 3 = neutral, 4= agree, 5 = strongly agree)	Radar Chart Bar chart

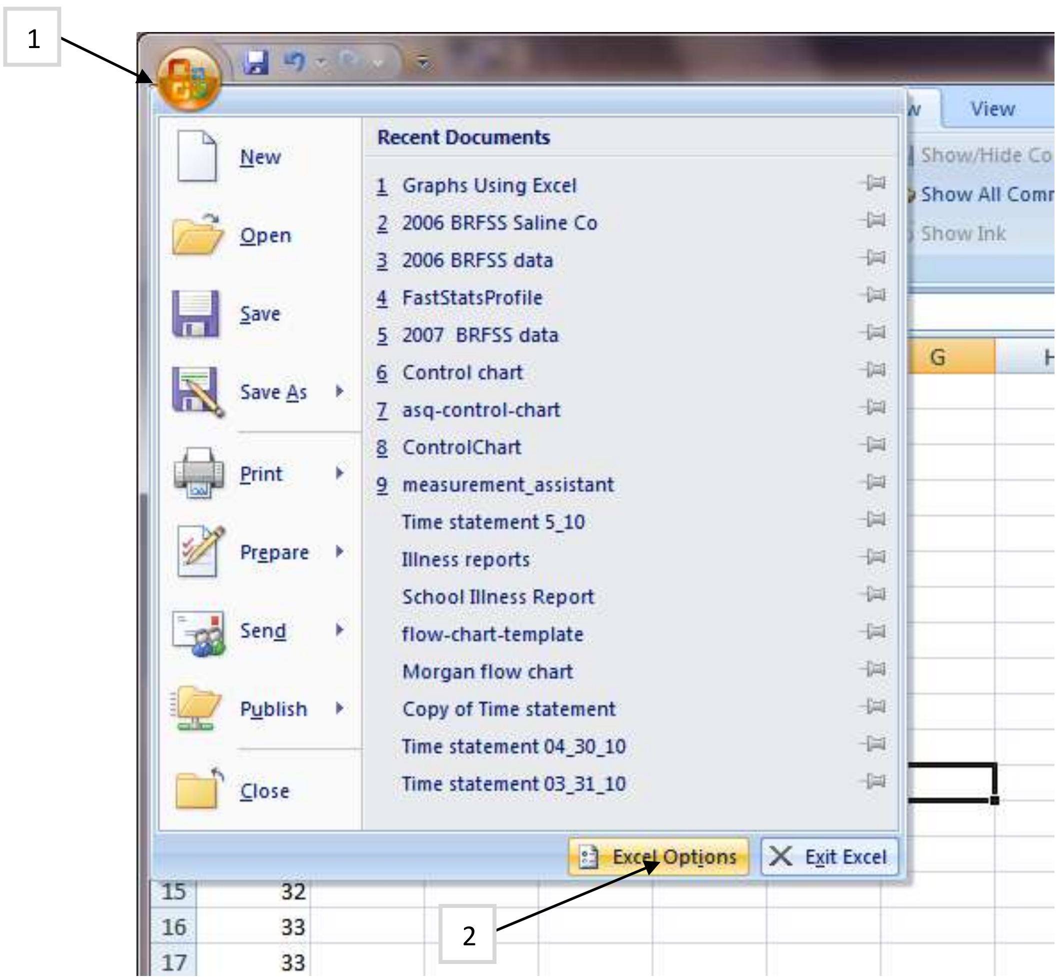
Continuous	Interval	Continuous data (equal distances between numbers on the scale, negative values can be used, but there is no “natural” zero (the start of the scale is arbitrary)	• Calendar time • Altitude • Fahrenheit temperature • Celsius temperature	• Bar chart • Pie chart • Pareto chart • Run chart • Control chart • Radar Chart • Scatterplot • Histogram • Line chart
	Ratio	Continuous data (equal distances between numbers on the scale AND there is a “natural” zero point on the scale (the start of the scale is NOT arbitrary)	• Height • Weight • Time • Kelvin temperature scale	

Getting Started

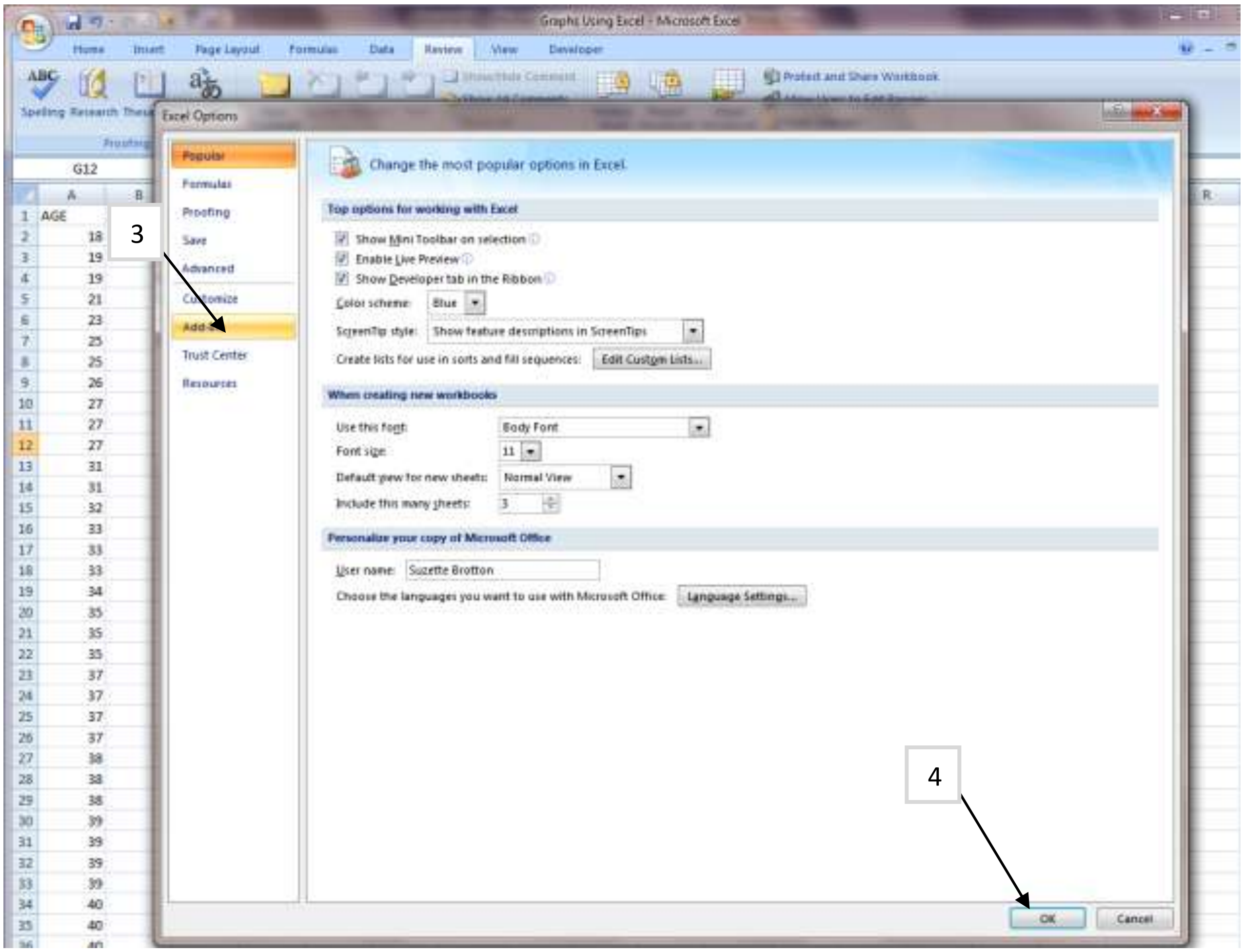
Excel provides a set of data analysis tools, the Analysis ToolPak, which can help to save steps for some statistical analyses. Activating this option may streamline some of your statistical analyses. All steps are shown with Excel 2007.

Activating the Excel Analysis ToolPak Add-in

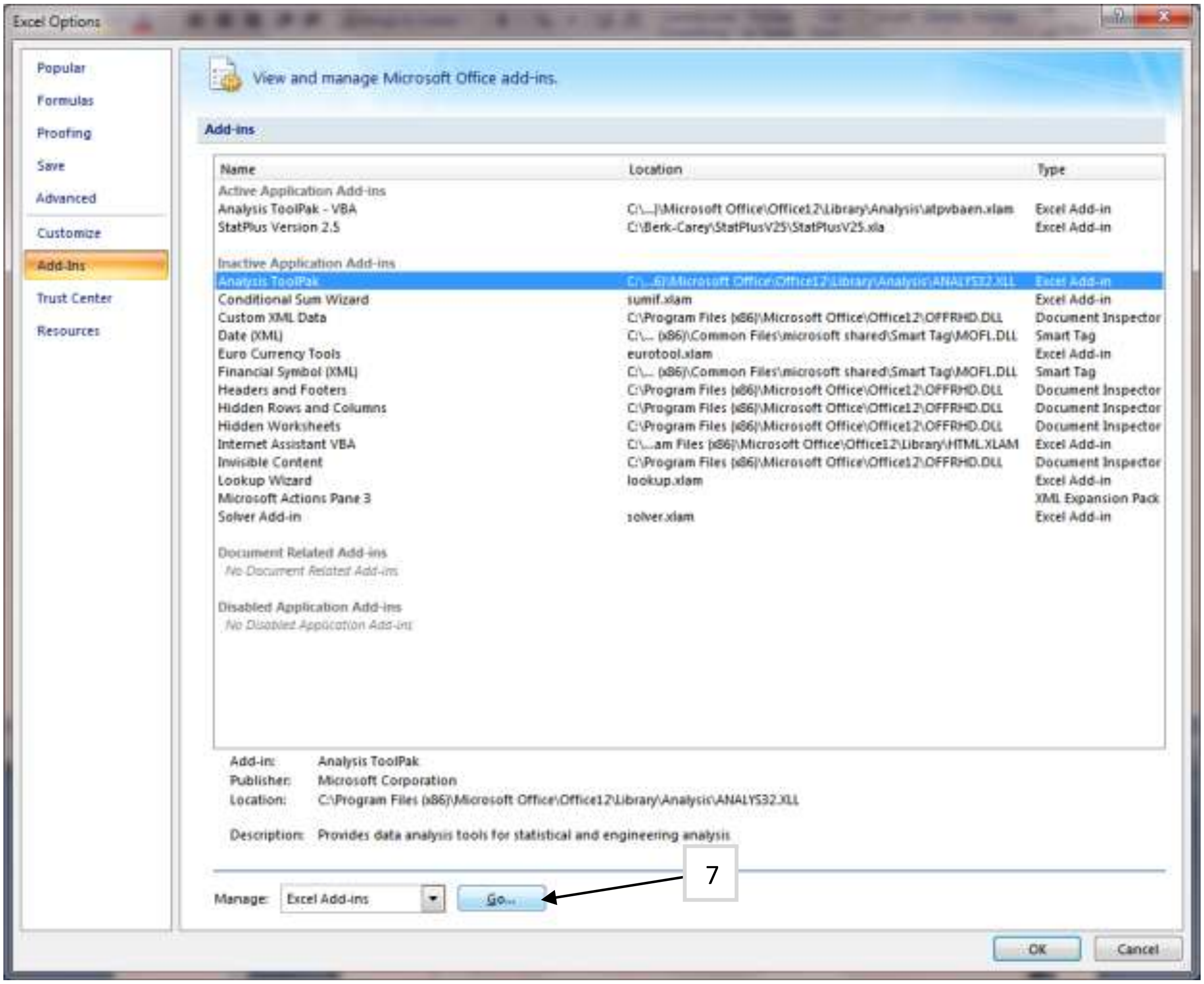
1. Click on the Microsoft icon in the upper left hand corner of the screen.
2. Click on Excel Options at the bottom of the drop down box.



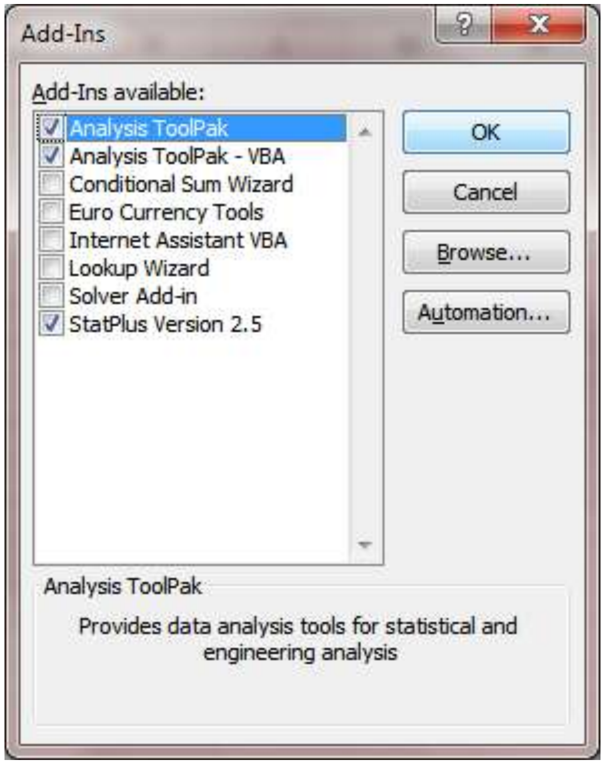
3. Click on Add Ins.
4. Click OK.



5. If the ToolPak has not been activated it will be in the list of Inactive Application Add-ins.
6. Click on Analysis ToolPak.
7. Click Go.

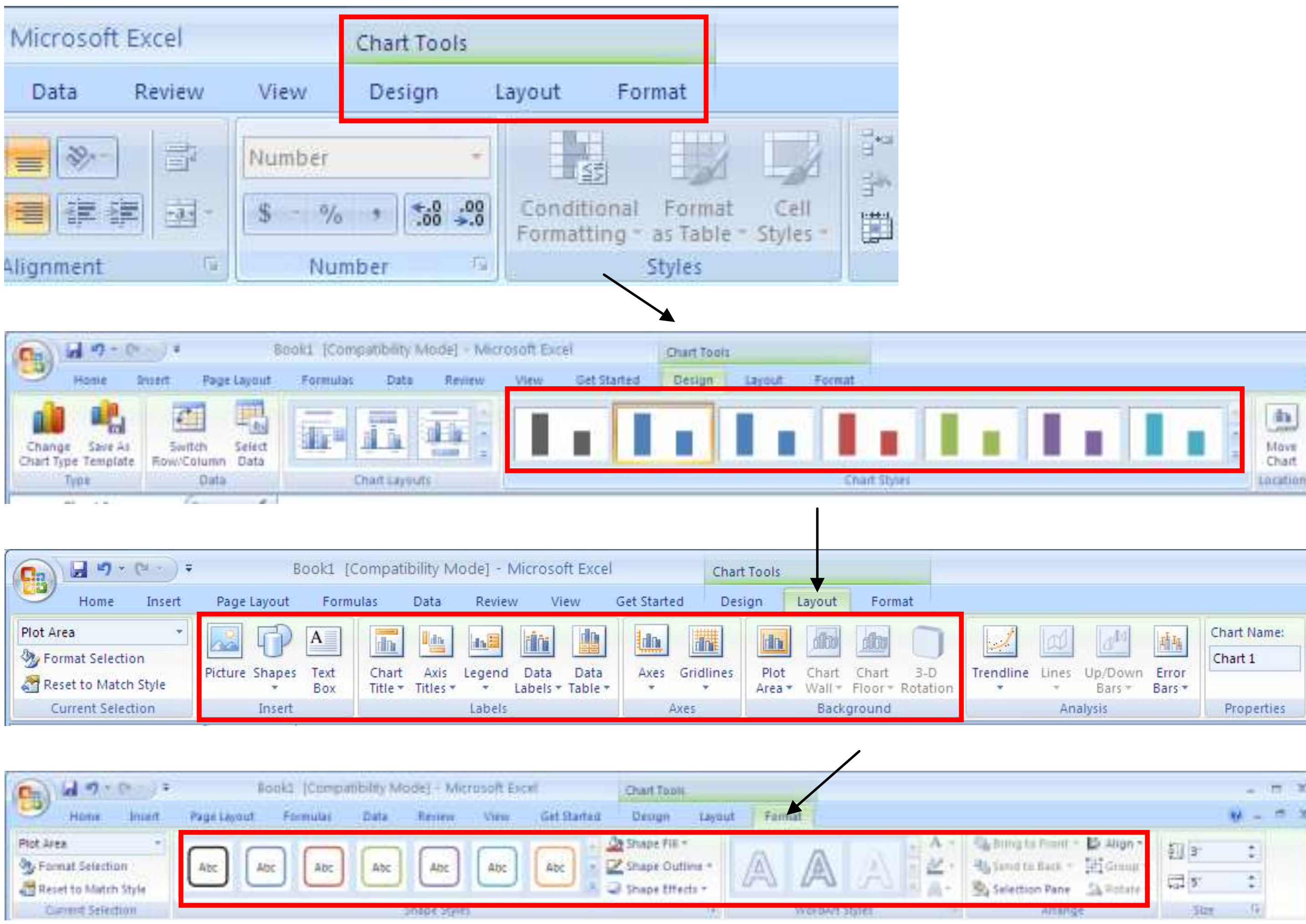


8. Check the Analysis ToolPak.
9. Click OK.

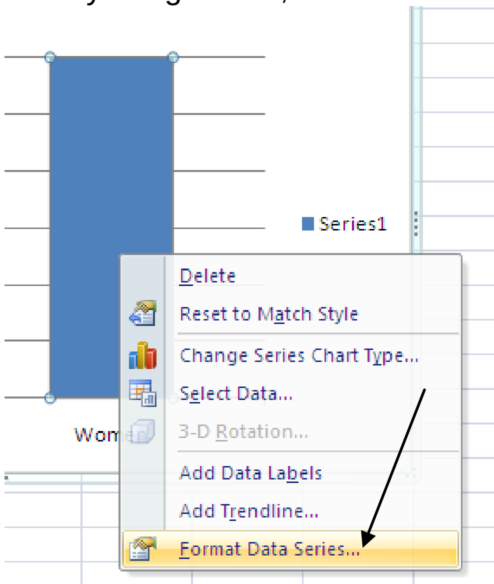


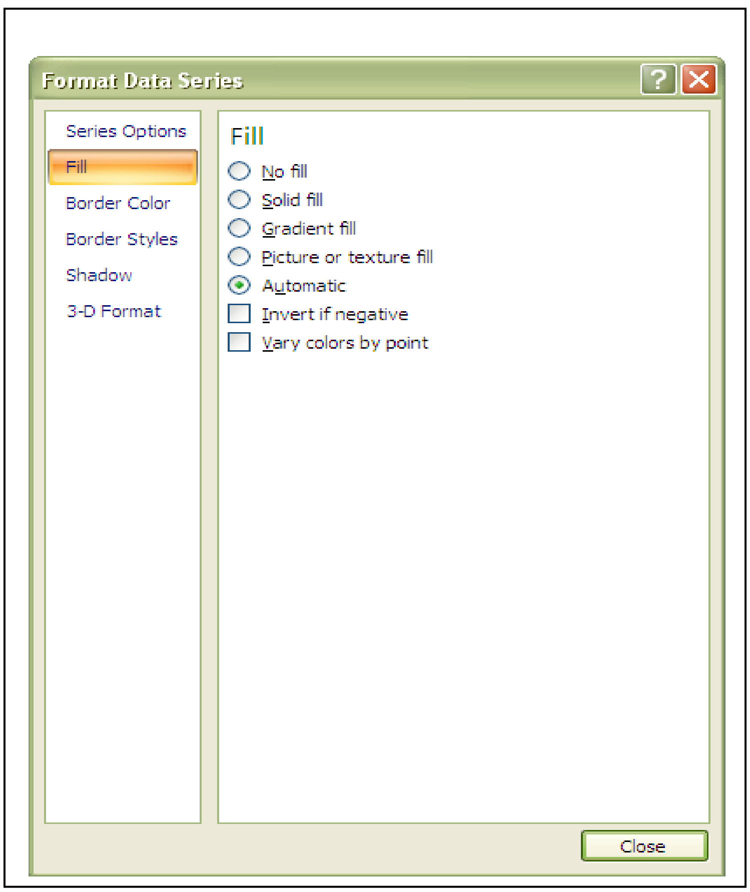
Formatting a Chart

Formatting a chart can be completed in several ways. When the chart is selected, the Format, Layout, and Design tabs will appear. Quick changes can be completed through these tabs, such as color schemes, adding chart titles, axis labels, and legends.



Formatting can also be completed by selecting an area of the chart and right clicking. Once you right click, select Format Data Series.

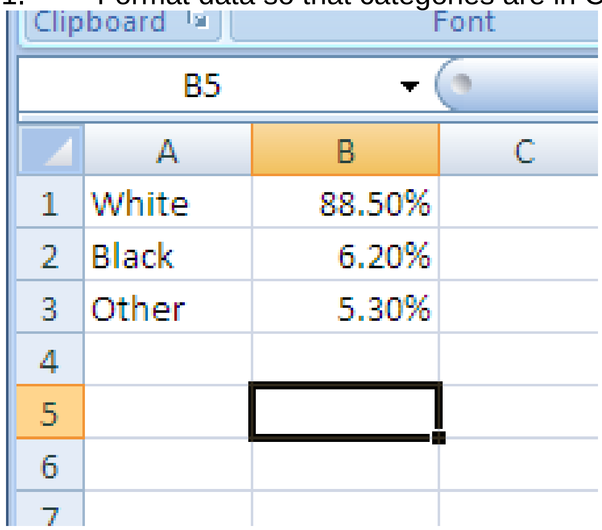




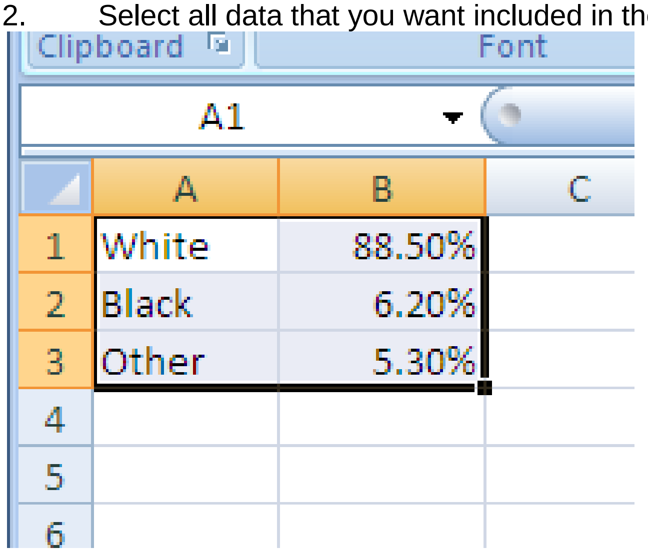
From this option, you can choose various colors for specific sections (under the Fill Option), or change border styles (under Border Options). Be sure to choose colors that reflect your project. The colors should coordinate where appropriate.

Pie Chart	
Usage	A pie chart is used to describe proportions. Pie charts work best when sections represent at least 25% of the data.
Data Types	Categorical Data
Interpretation	Read each section as a portion of the whole. Ex. White people make up 88.5% of the people in Kansas.
Creation in Excel	You will put your categories of information in one column and percentages in the column next to it. Make sure percentages are calculated before starting the pie chart.

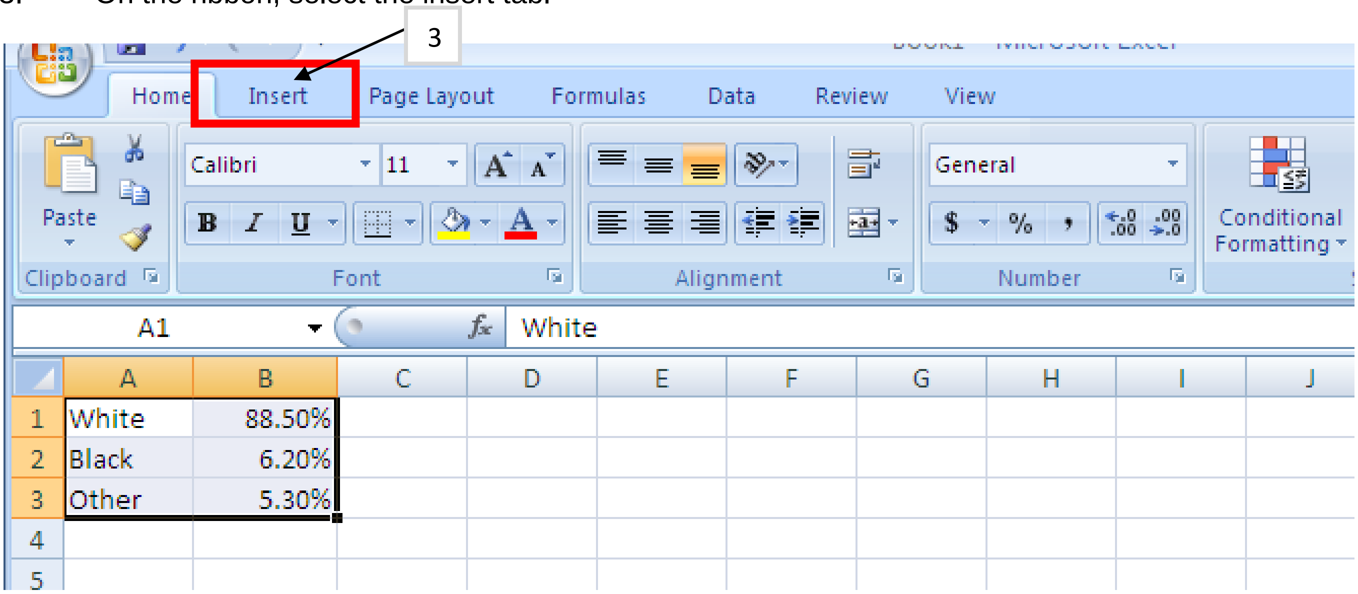
1. Format data so that categories are in Column A and percentages are in Column B.



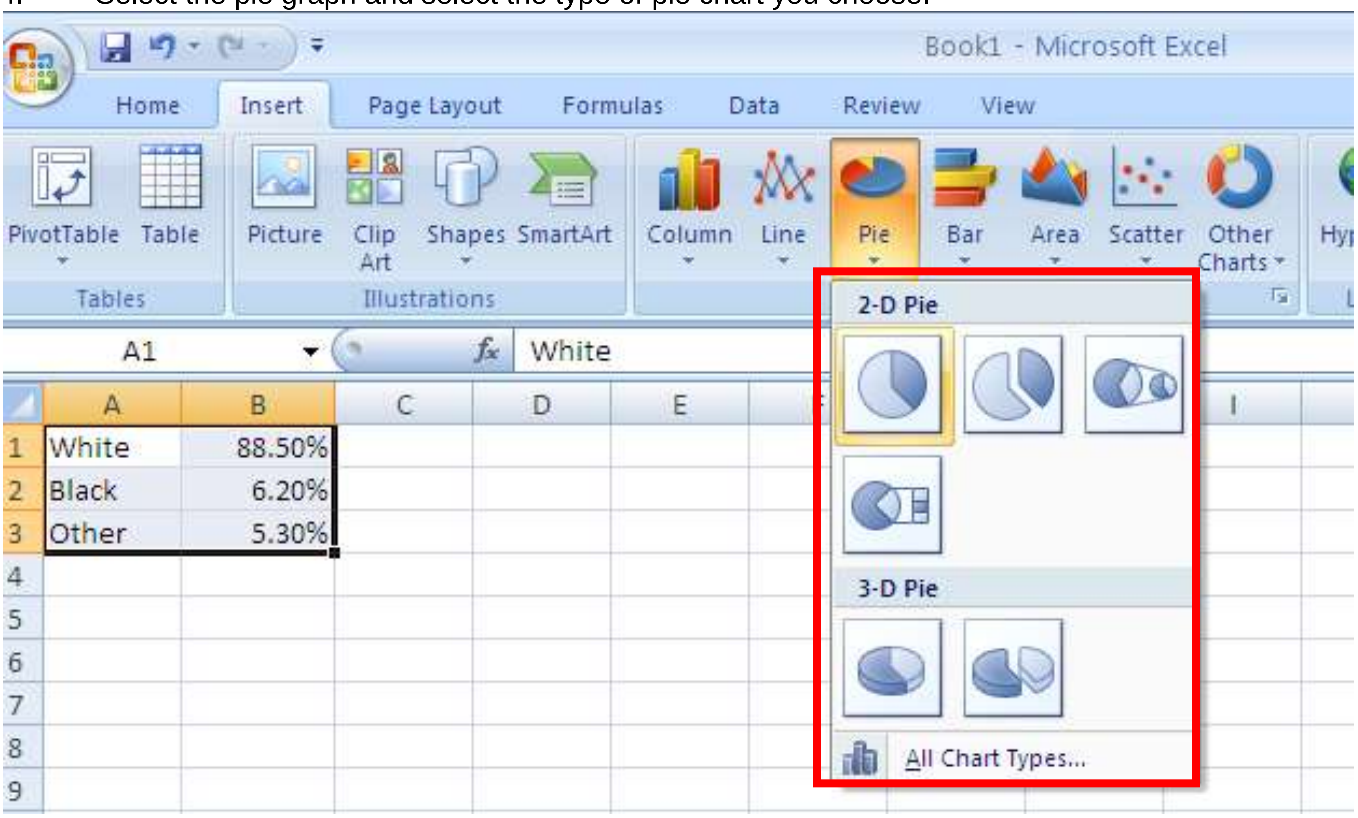
2. Select all data that you want included in the chart.



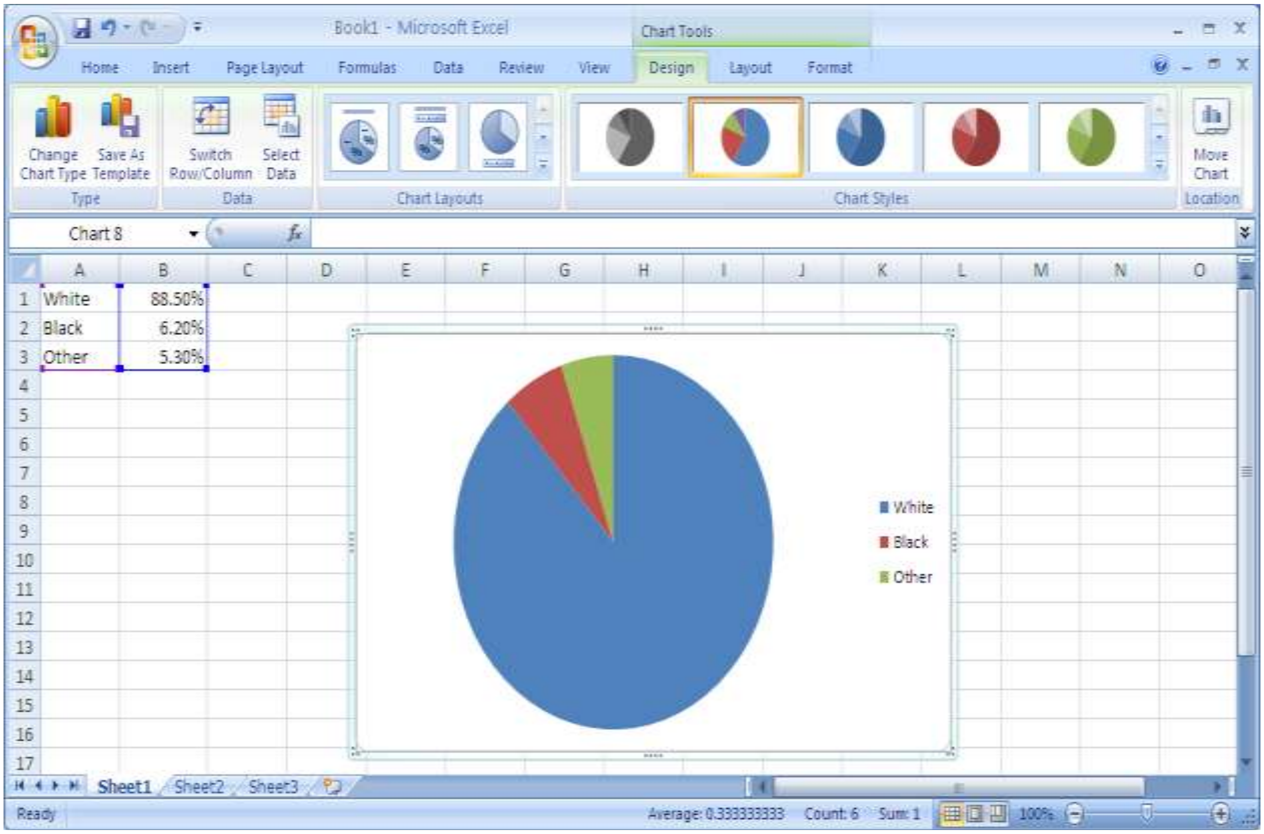
3. On the ribbon, select the insert tab.



4. Select the pie graph and select the type of pie chart you choose.



5. Click the type of pie chart you want. The chart will be automatically inserted into the sheet.

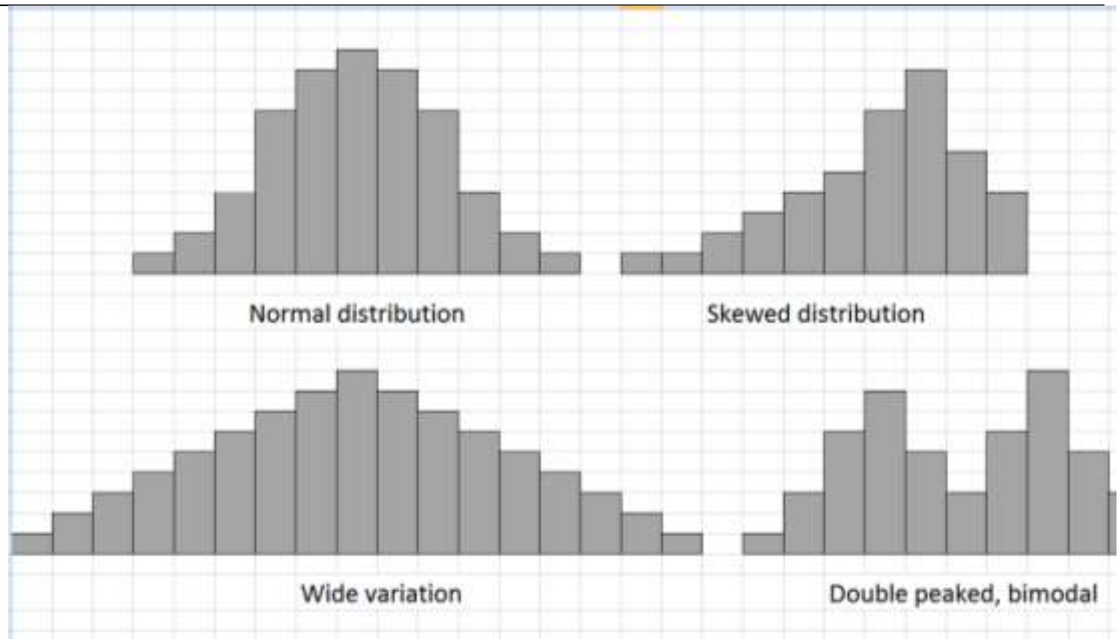


6. To add titles and labels, select the Layout Tab while the chart is selected. Titles, data labels, and legends can be added from that tab. For more on Formatting the chart, refer to the section on Formatting.

Histogram

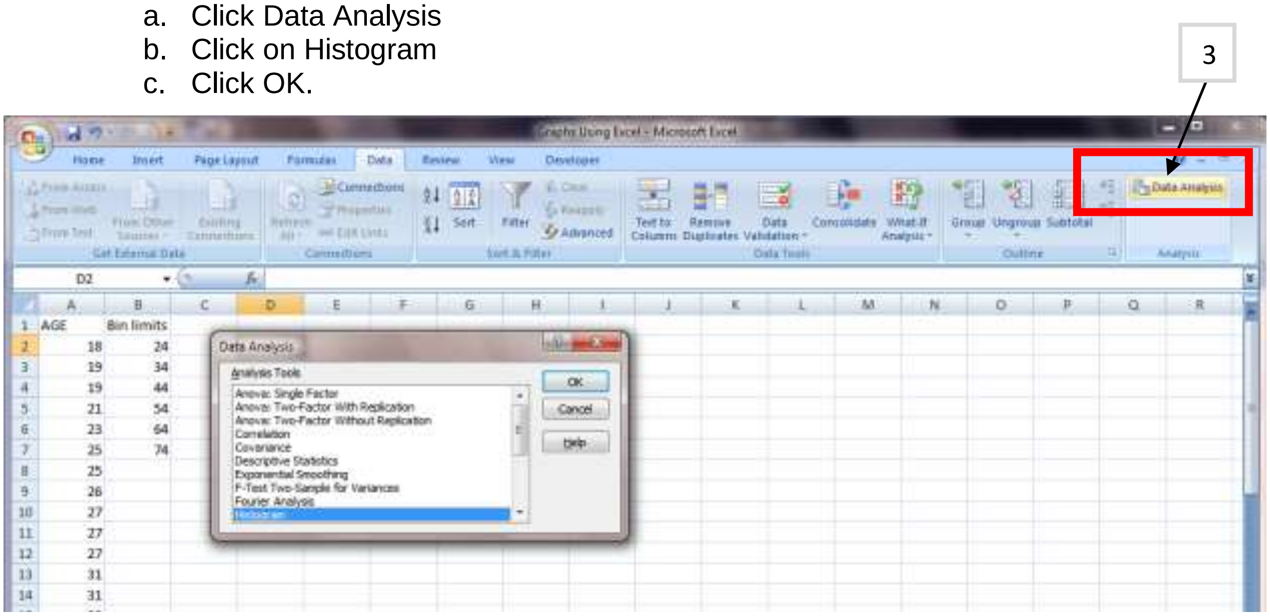
Usage	To examine the variability of the observed variable (aka central tendency), a histogram will visually display the (1) spread and shape of a group of values and (2) visually display the count (frequency) of those values within user defined ranges. A histogram may be helpful as the initial step in analysis of a problem or process. For example, if you are interested in understand the range of age within a population, a histogram will "count" the number of persons whose age falls between 0-18, 19-40, 41-60 and > 60 years
Data Types	Single variable, continuous numerical data. Examples: age, height, weight, temperature.
Creation in Excel	1. Format the data so that the numerical values populate a column. A minimum of 50 data points are recommended for the construction of a histogram (Tague, 2004). 2. Sorting the data may be helpful when determining the range of the values and trying to determine groupings, but is not required. The Excel Help menu can be used if needed for sorting. 3. Determine the range limits for the groupings of interest. The number of groups is partially determined by the number of data points in the set. Additional considerations may include the goal of the analysis and the intended audience. Too many or too few groups may preclude reliable analysis. The number and ranges can be mathematically derived or by using general guidelines (Brassard & Ritter, 2007): 4. The group ranges are plotted on the x- axis and the frequency on the y- axis. 5. The following example uses the Analysis ToolPak Add-In for Microsoft Excel 2003, 2007 and 2010. This Add-In is included in the Excel program, but needs to be activated (as described on page 7).
Interpretation	1. A normal distribution results in a bell shaped curve, with a single peak approximately centered and as many data points on one side of the center as the other. 2. Distributions that are skewed (more data points are present in lower or higher ranges) may be due to natural limitations in a process. Example: infant birth weight in a population. Skewing may also be the result of a process that is not well controlled. 3. A histogram that is wide (has more spread) indicates variability in the process. 4. The presence of more than one peak is indicative of two separate processes or sources of information. Example: a process that occurs on two different work shifts with two different results (Tague, 2004). 5. The mean is an appropriate measure of center for normal distributions. The median may be more appropriate for skewed distributions.

Number of data points	Number of groups
< 50	5-7
50-100	6-10
100-250	5-8
>250	6-11



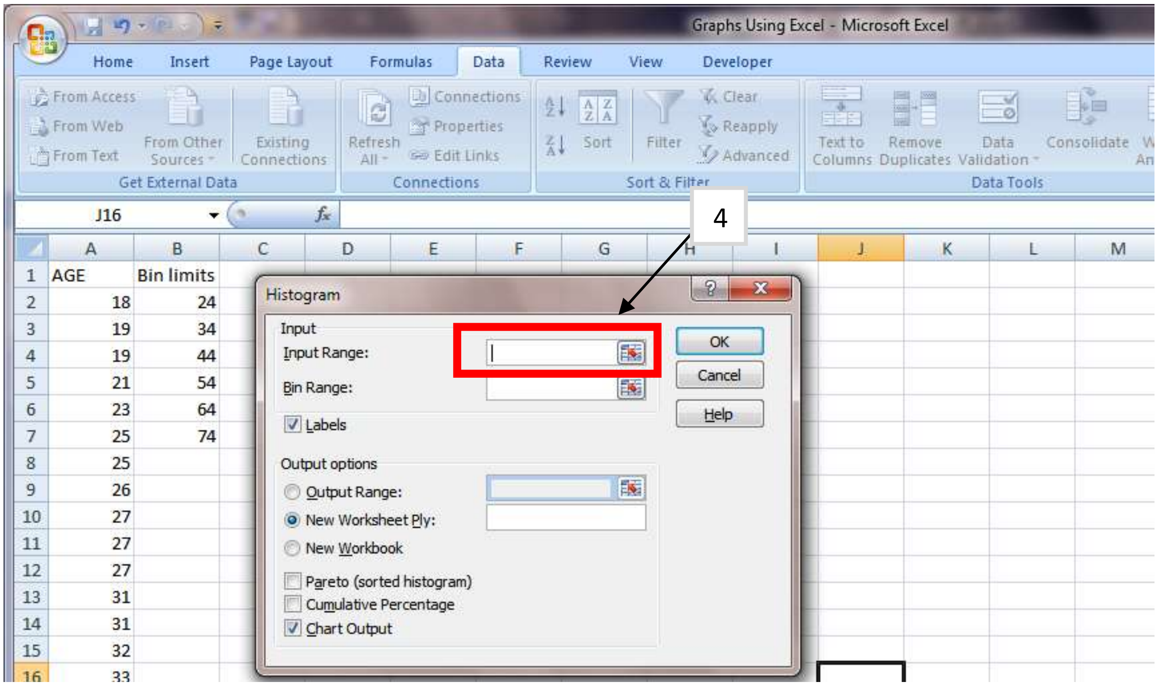
Creating the histogram

1. The values are sorted in this example. Excel will count the number of values that belong in each group and use that information as the source for the graph.
2. Enter the desired group/bin limits in rows. Enter only the upper limits for each group in this step. Group limits should be entered in continuous rank order.
3. To create the chart:
    a. Click Data Analysis
    b. Click on Histogram
    c. Click OK.

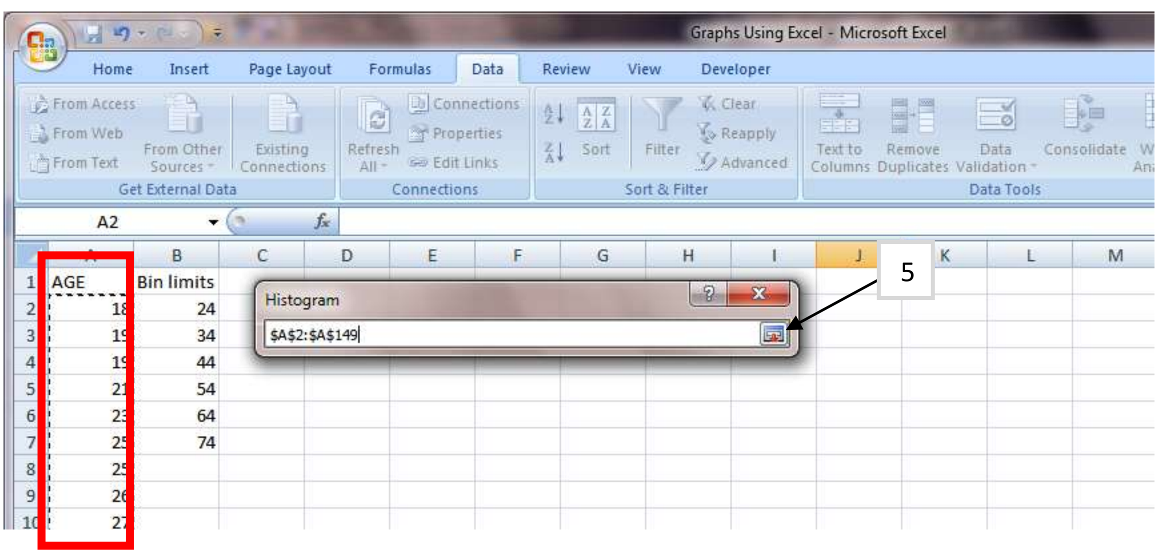


4. A dialogue box will appear allowing the selection of the data to be included in the graph using the mouse, or the keyboard. The use of the mouse for this purpose is the method

used in the next several examples. Click the button to the right of the Input Range and a second dialogue box will appear.

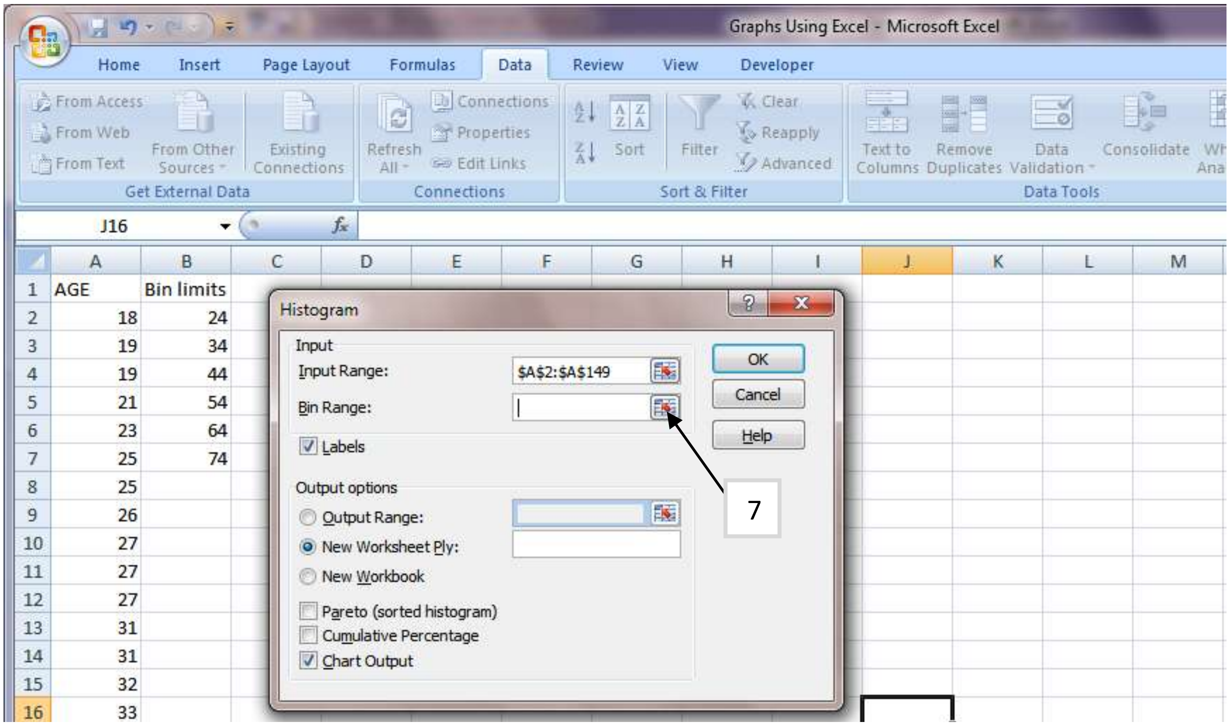


5. Using the mouse highlight the range of data to be counted into groups.

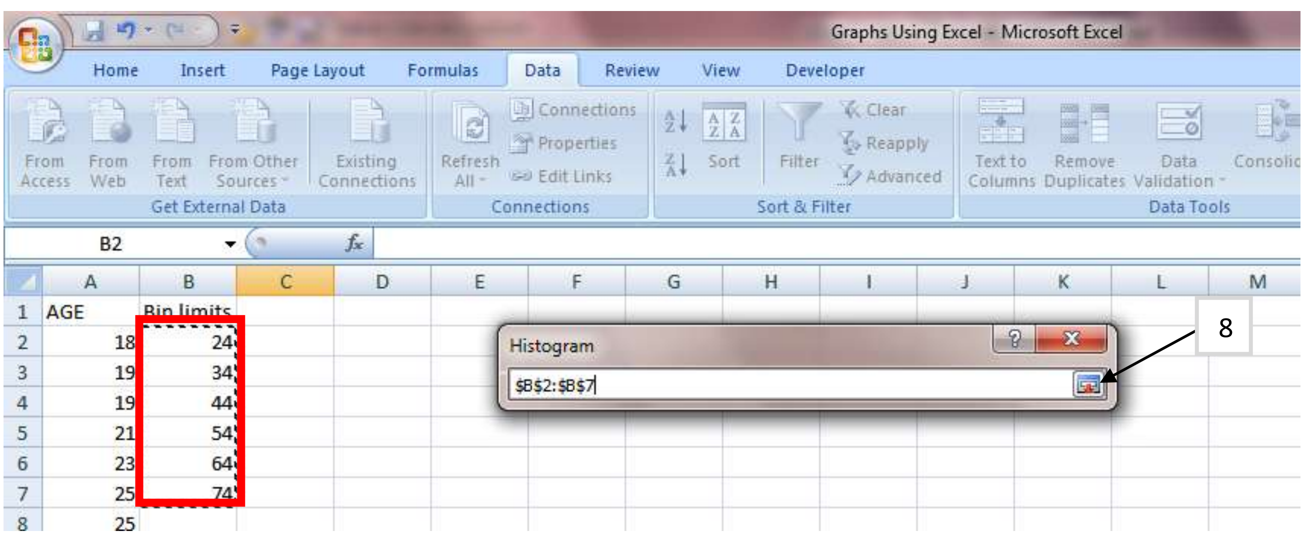


6. Press Enter after all of the applicable values have been selected. You will be returned to the histogram dialogue box.

7. To identify the range of values for the groups (bins) click on the arrow to the right of the Bin Range.



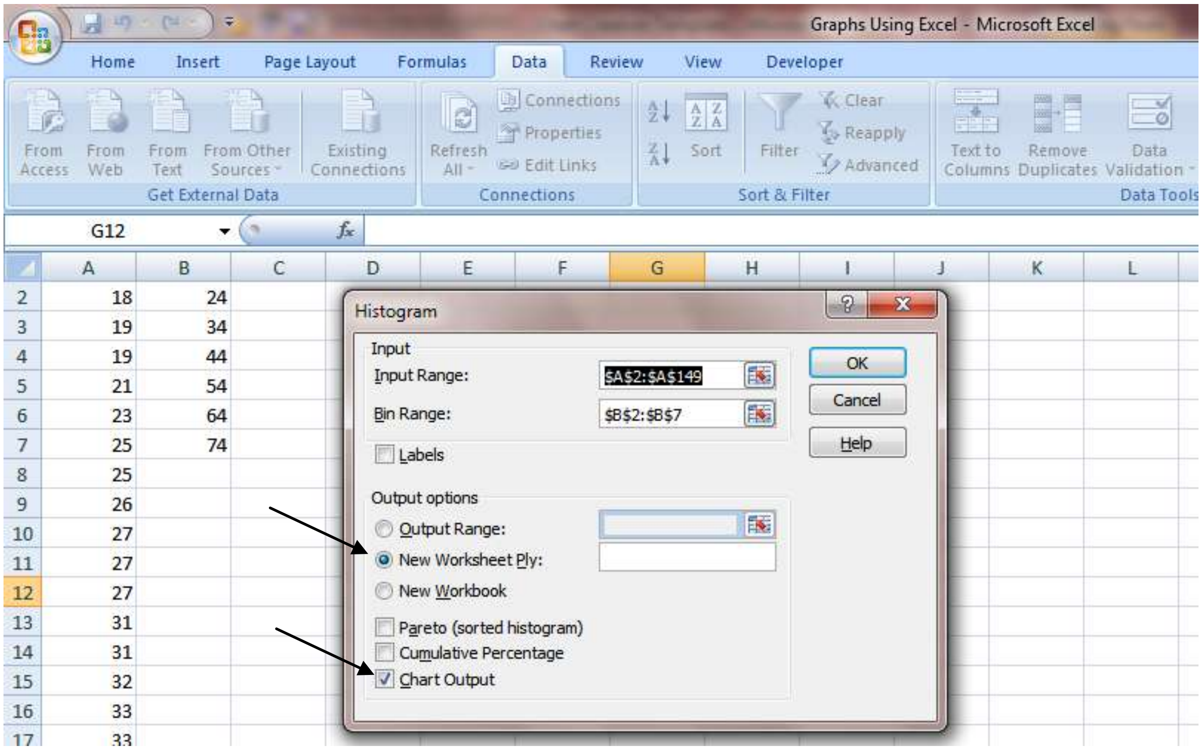
8. Using the mouse select the range of values to be used as groups.



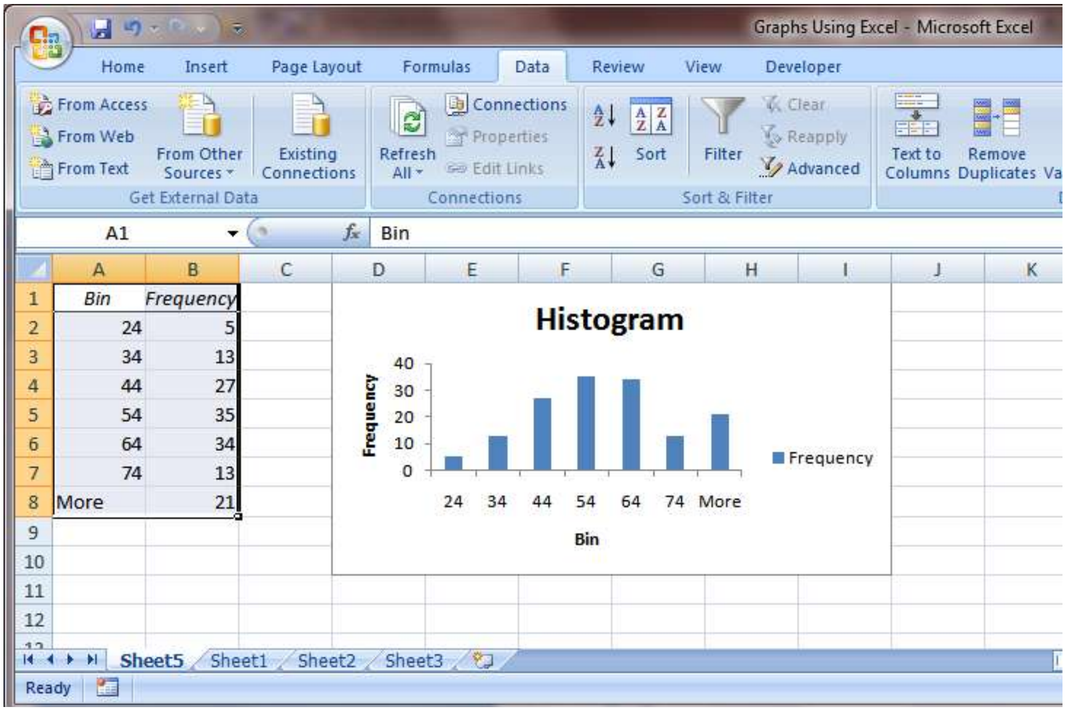
9. Press enter when the appropriate range of values has been selected. You will again be returned to the Histogram dialogue box.

10. There are three options for placement of the resulting data and chart; (1) within the original Work Sheet, (2) in a separate Work Sheet within the existing Work Book (a new tab at the bottom of the Work Book), or (3) in a new Work Book (file). For this example

the chart will be placed in a new Work Sheet within the existing Work Book. Select where the chart will be placed by clicking the applicable button.

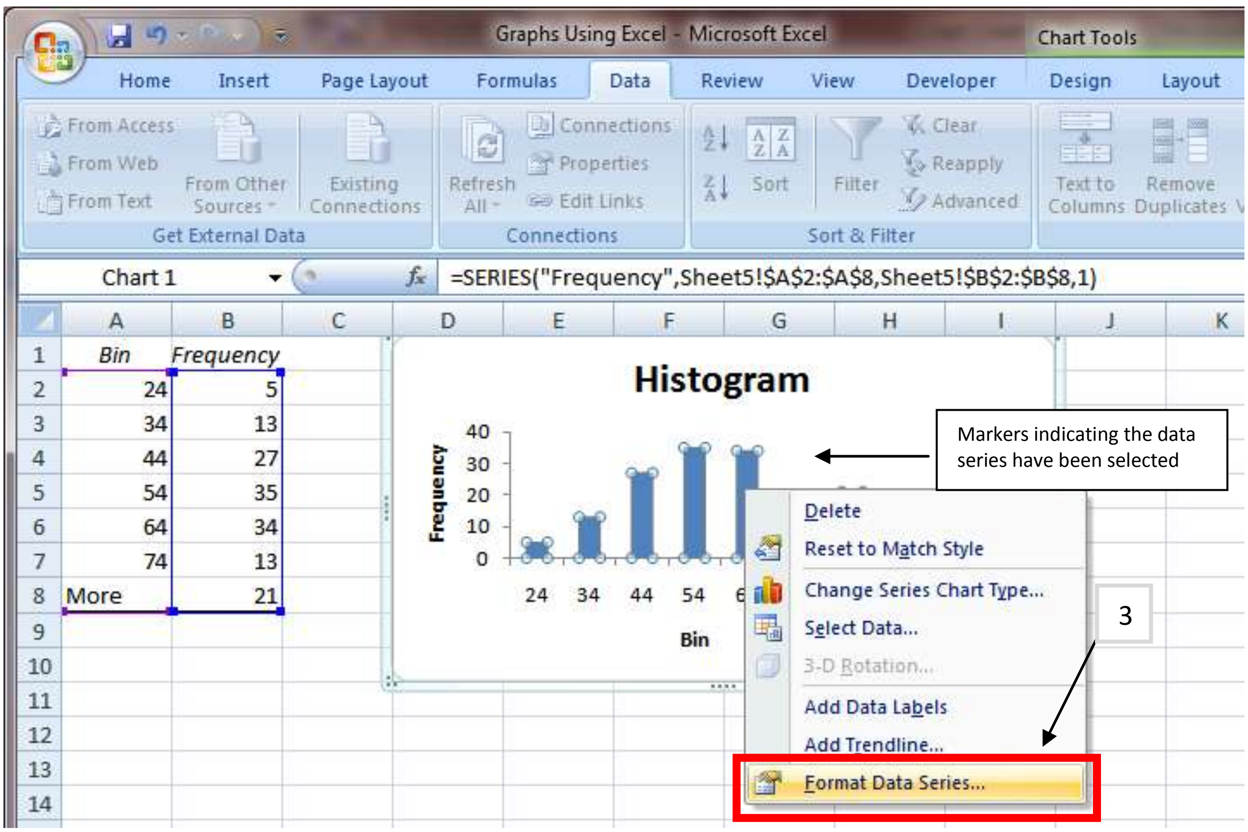


11. Click OK. The resulting data will be displayed and the histogram created.

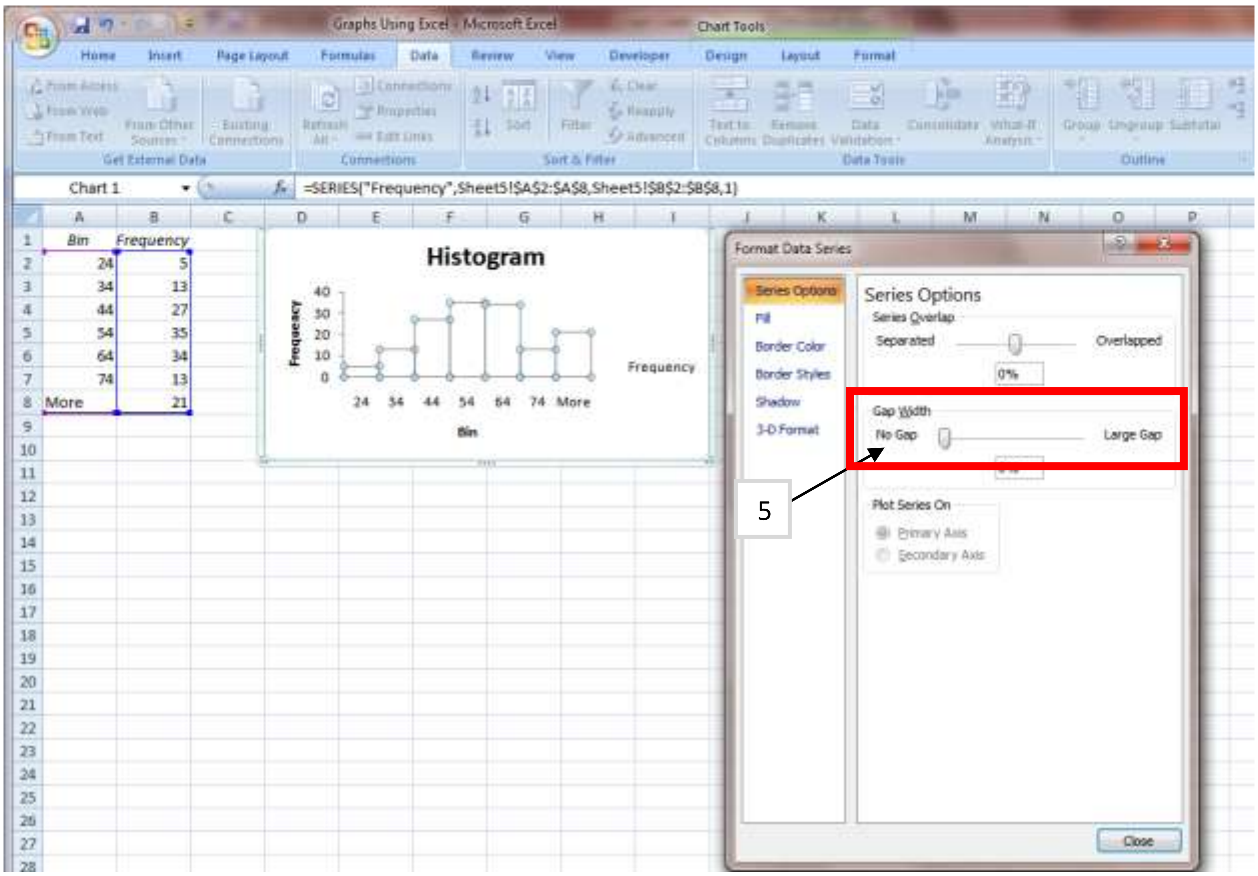


Formatting the histogram

1. Changes can be made to the chart by clicking on the chart and selecting a specific element within the chart. Histograms do not have spaces between the bars, so the resulting bars representing the data within the groups will have to be formatted.
2. Click in the chart, then specifically on the bars. Markers will appear that designate the bars (data series) as the element to be formatted.
3. Right click on “Format Data Series”

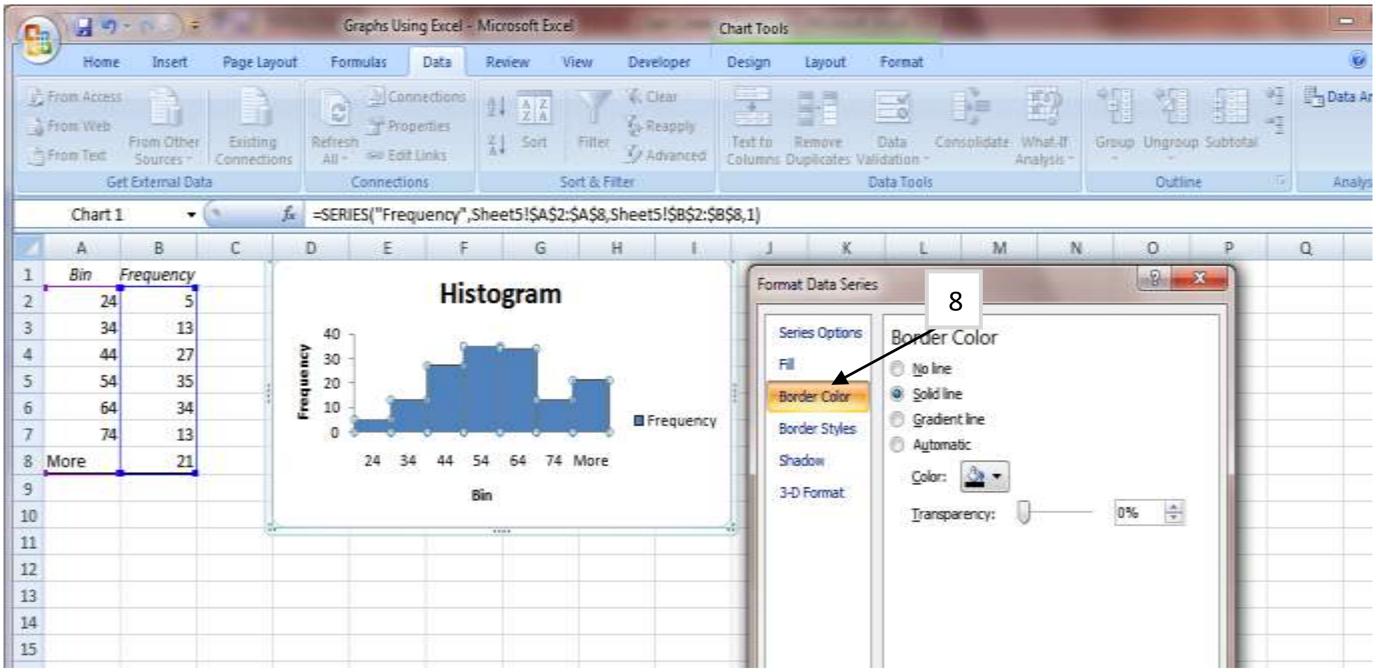


4. In the Format Data Series dialogue box highlight Series Options.
5. Move the Gap Width slider as far left as it will go toward “No Gap”. The bar color has been removed for visibility in this example. Other aspects of the data series, such as color, can be easily customized using the options found in the Format Data Series dialogue box.



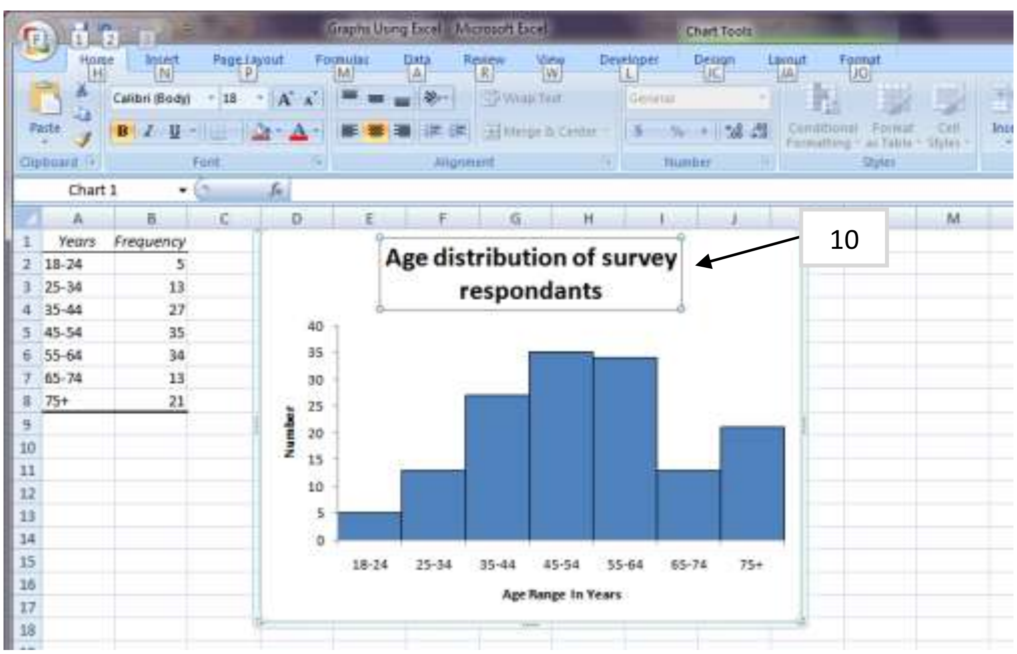
6. To facilitate visibility and differentiate the data in each range a border has been placed around each bar.
7. In the Format Data Series dialogue box select Border Color on the left.

8. Then select Solid Line and a Border Color. A color should be selected that will be visible against the bar color. For the purpose of this example black was selected.

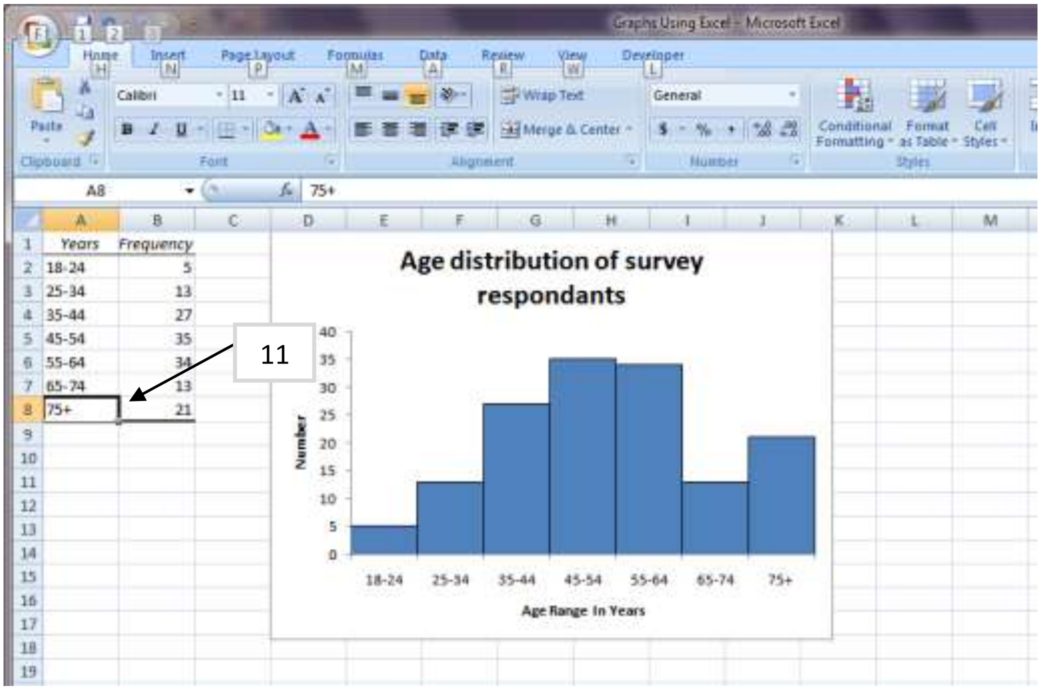


9. Exit the dialogue box by clicking on Close.

10. The chart title and axis labels are easily edited by clicking on that component of the chart and typing in the desired information.



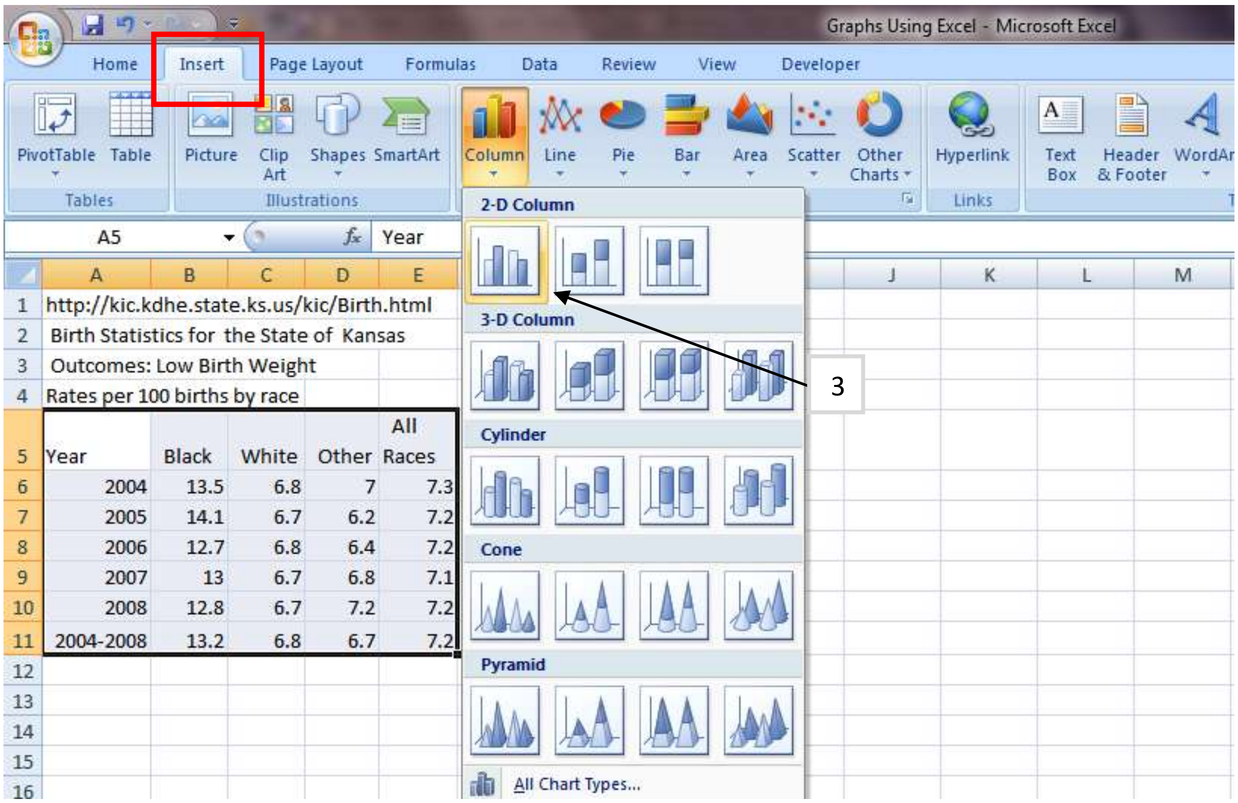
11. The ranges for the axis labels are easily edited by typing in the cells that contain that information. The chart will be updated automatically.



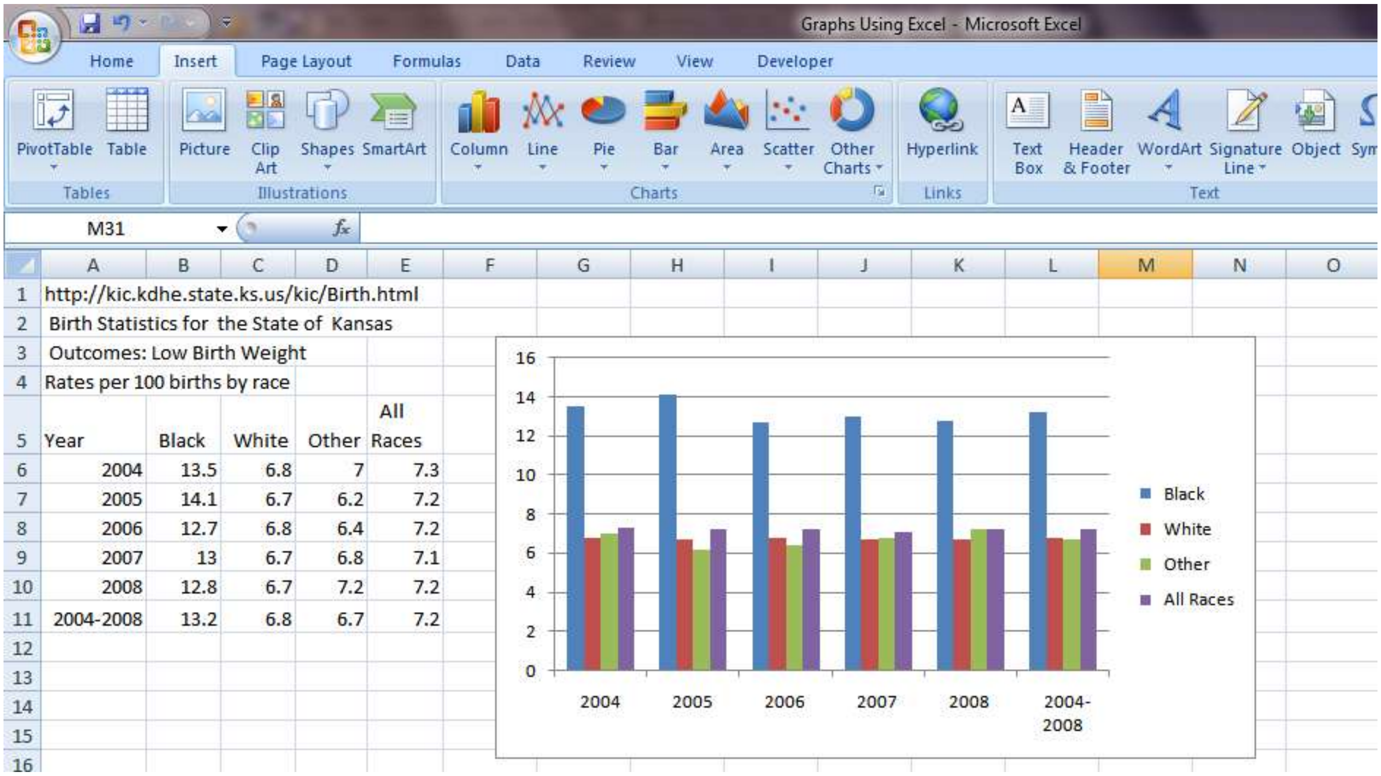
Bar Chart	
Usage	Bar charts provide a visual image for comparison of values among groups or categories of data.
Data Types	Any discrete attribute or variable data. . Examples: race, gender, marital status, department, work shift, year
Creation in Excel	1. One or more groups of data may be represented. Identify the groups of interest. 2. The bars may be vertical or horizontal. 3. Groups or categories are represented along one axis. The other axis is scaled as a reference point for the numerical values represented by the height or length of the bars. 4. Enter the data to be used for the chart as displayed in the example below. 5. Sort the data by the variable of interest. In this example the rate of black infants born of low birth weight is the variable to be highlighted.
Interpretation	Focus on the tallest and shortest bars, the size of each bar within a group, as well as the change in size of the bars over time. The birth certificate data used in the following example highlights the substantially higher rate of low birth weight infants born to Blacks as compared to other races. Additionally, it is apparent that there was little change from 2004 to 2008.

Creating the bar chart

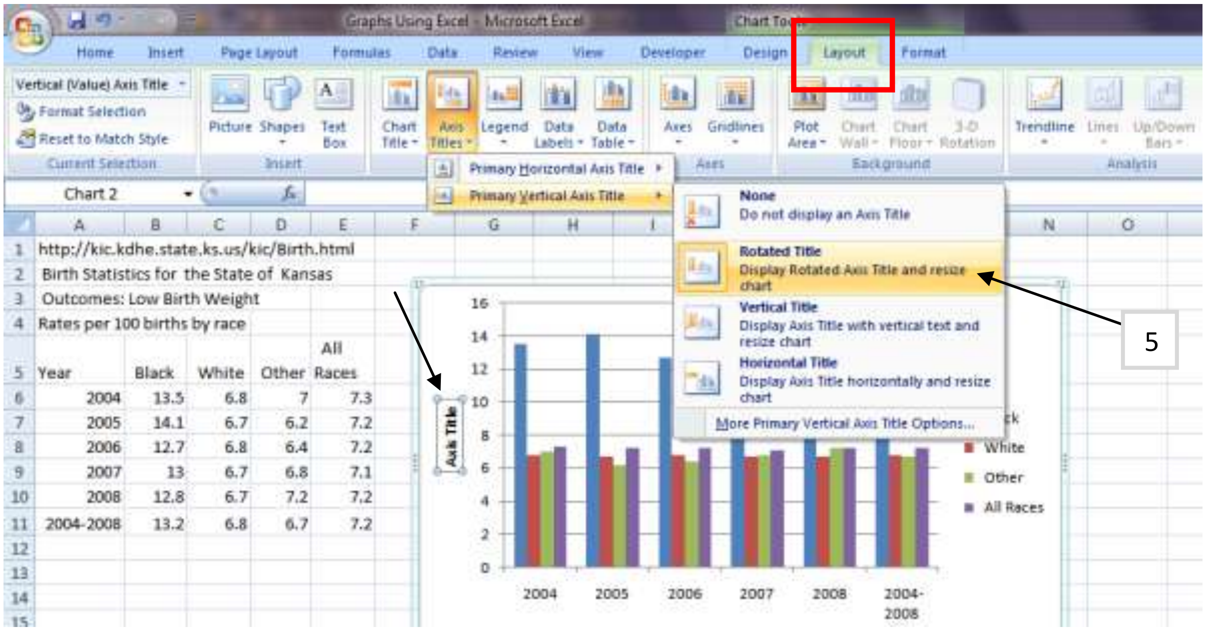
1. Highlight the data and captions for the desired chart.
2. Click on the "Insert" tab in the ribbon.
3. Click the "Column" chart in the chart types and select the desired chart type from the drop down box.



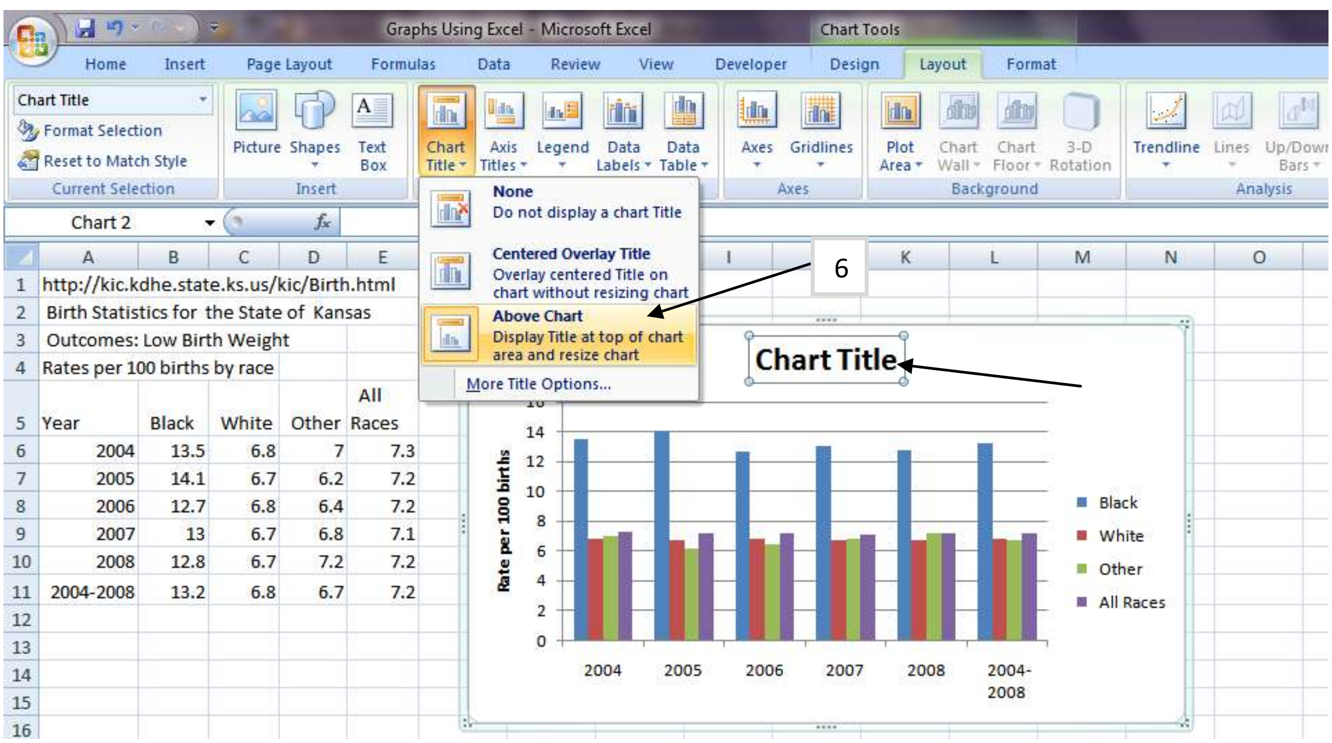
4. The chart will automatically be placed in the existing worksheet. If the columns and rows include headers and that information was included when the data for the chart was highlighted the axis intervals will be appropriately labeled automatically.



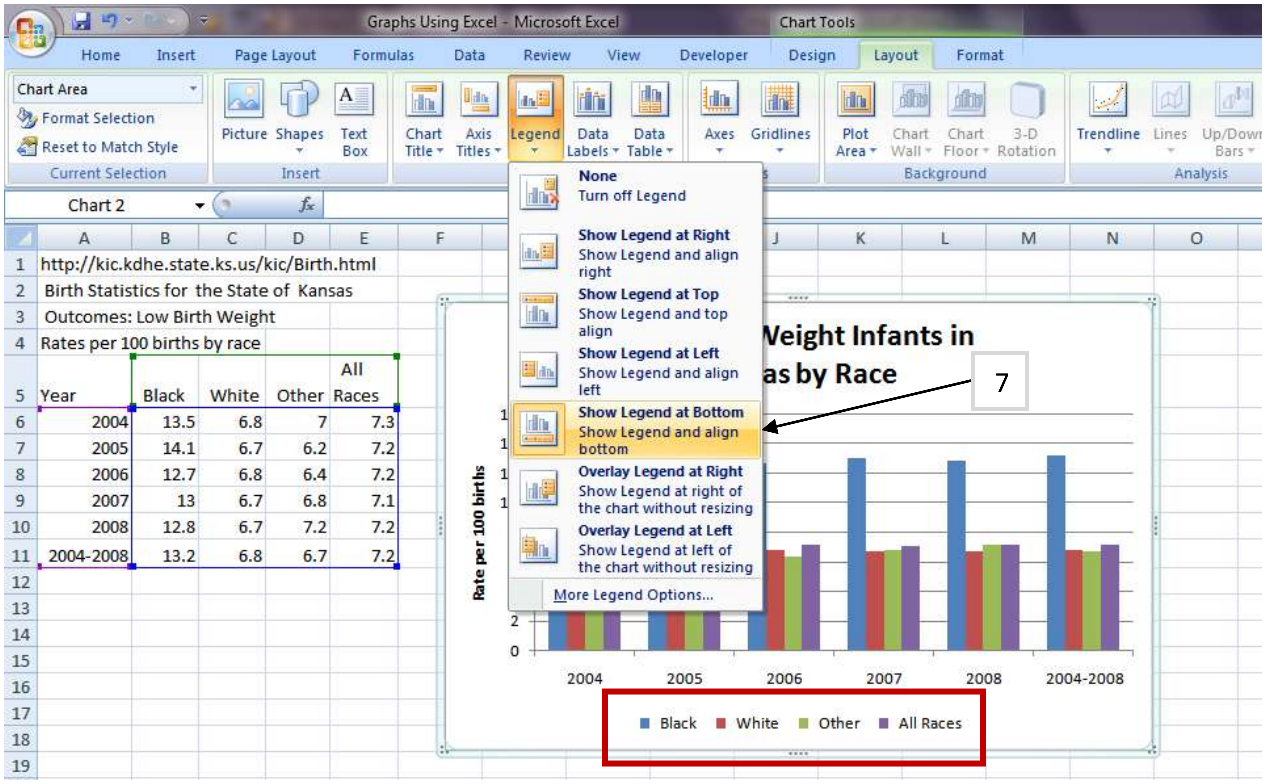
5. The chart title, axis labels and legend are added and edited by highlighting the chart and selecting the desired component from the “Layout” tab in “Chart Tools”. In this example a horizontal axis label is omitted because the category labels are self explanatory.



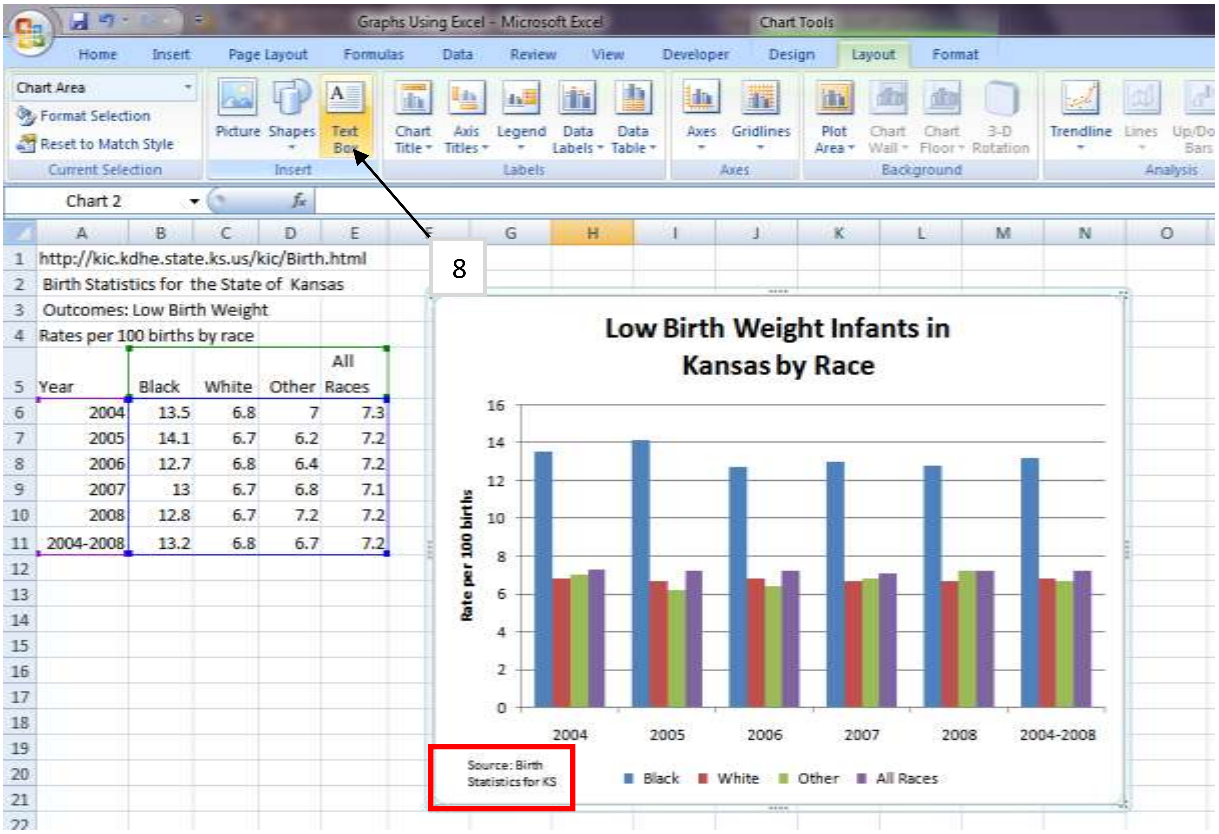
6. Similarly, a chart title is added by selecting “Chart Title” and the desired location in the Layout tab of Chart Tools. Type the desired title.



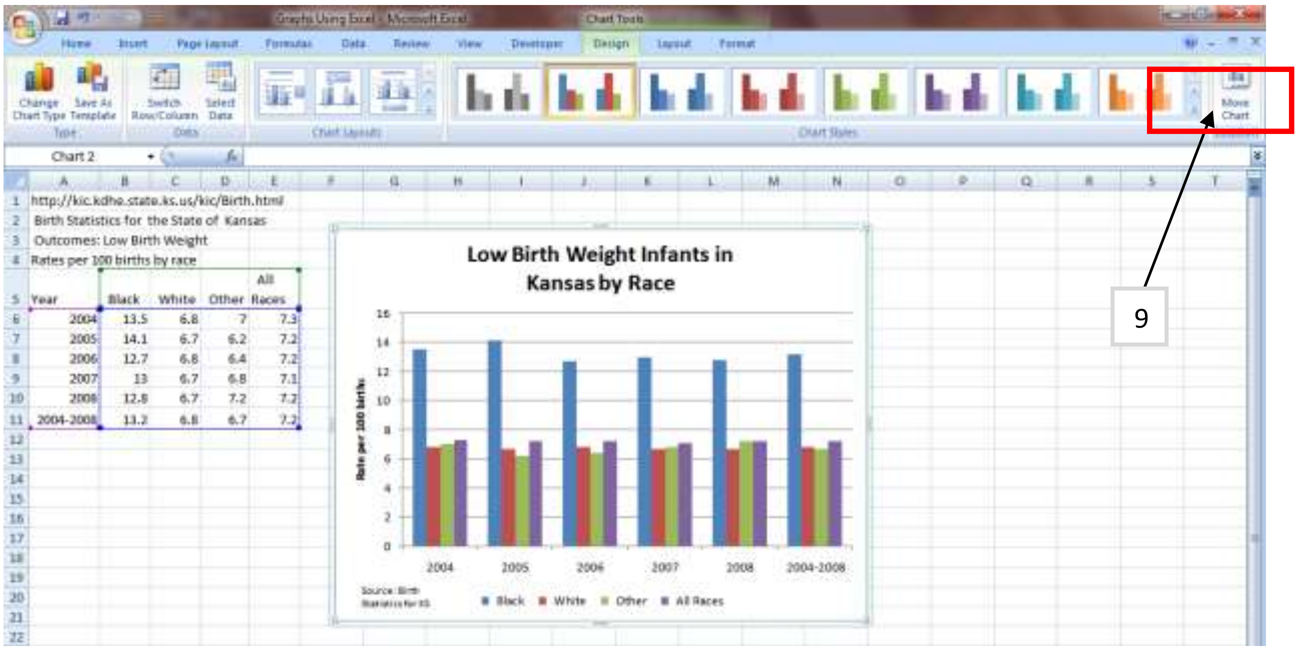
7. The legend that is generated when the chart is created is automatically placed to the right of the plot area. The preferred placement is within or below the plot area (Klass, 2006). To move it select the Legend option and the desired placement from the drop down box.



8. A text box can be added to identify the source of the data used for the chart. Select Text Box from the Layout tab in Chart Tools. The required information can be typed in to the text box.

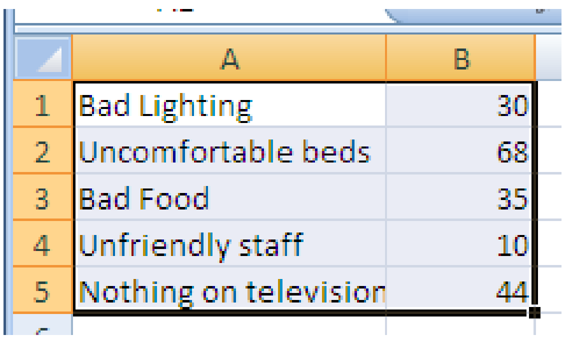


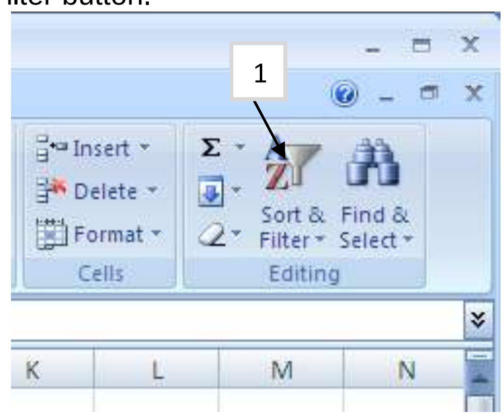
9. At any time the chart can be moved to a separate worksheet within the workbook by clicking on the chart and selecting the Design tab in Chart Tools, then selecting Move Chart. A dialogue box will appear allowing the user to select a different worksheet within the workbook as the location for the chart.



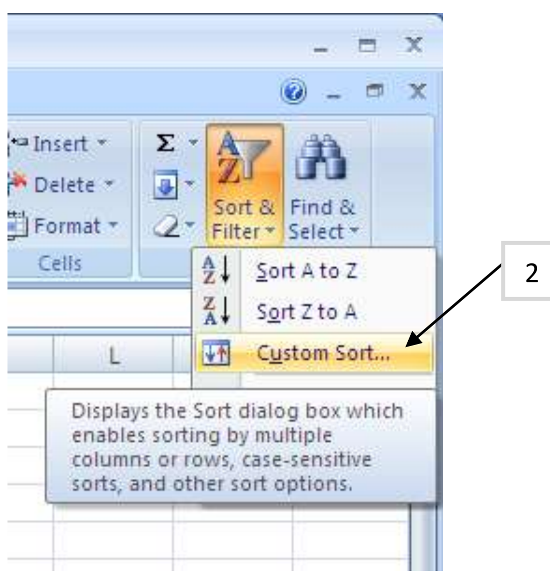
Pareto Chart	
Usage	The Pareto chart is used to identify areas which are causing the largest problems. The idea of the Pareto Chart stems from the Pareto Principle, where 80% of the trouble comes from 20% of the problems. The Pareto Chart helps to identify how many issues can be addressed with limited resources.
Data Types	Categorical data which is able to be quantified.
Interpretation	The chart not only shows the areas of largest problem, but also shows the how much of the total those problems represent. In the case below, "Uncomfortable Beds", "Nothing on Television", and "Bad Food" are the categories in which 80% of our problems are stemming from. It would be wise for the organization to focus on those three areas, as they will give the most return.

1. Arrange categorical data in order of largest to smallest. To do this, highlight both the categories and numbers. Then, click the Sort & Filter button.

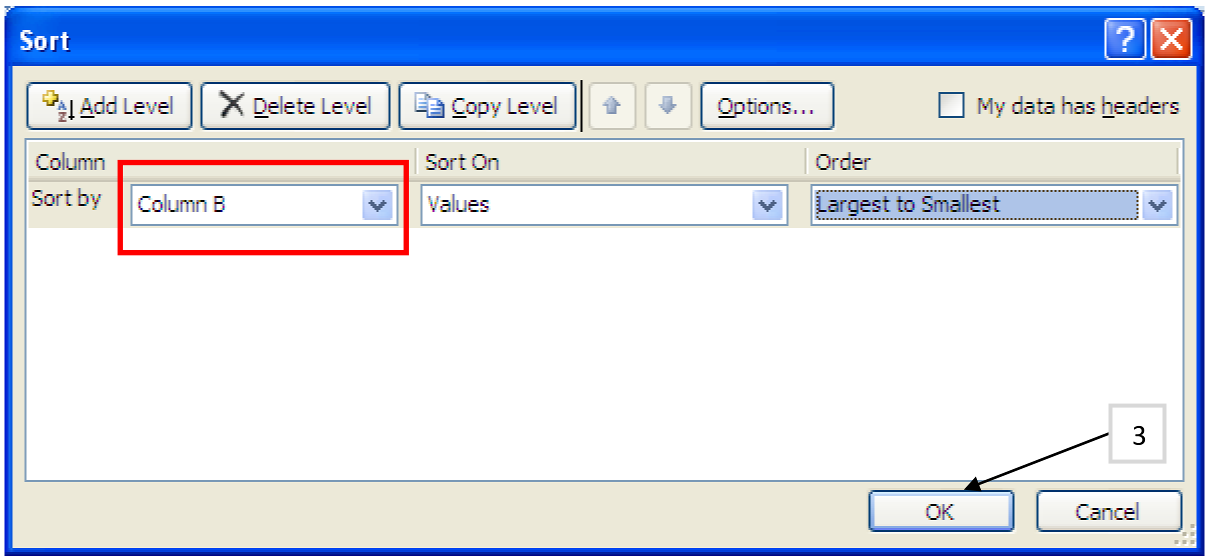




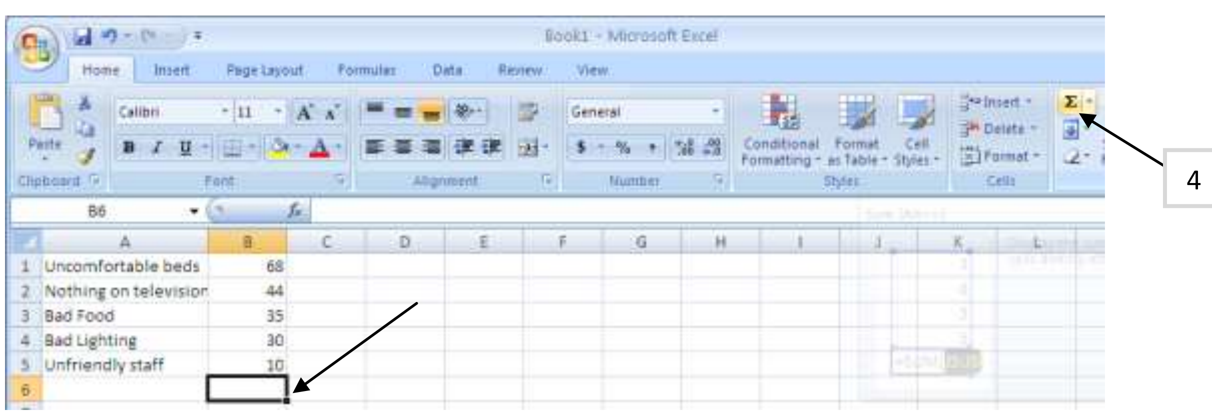
2. Once you click the Sort &Filter button, select the "Custom Sort" option.



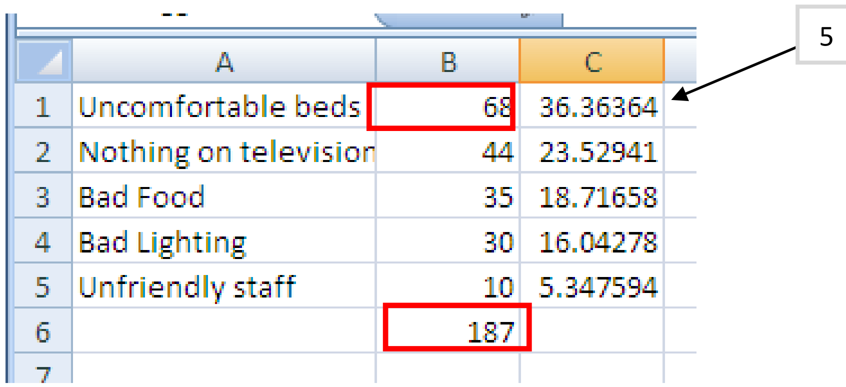
3. Sort by the column you have your numbers in. In this case, we chose column B to sort by, as our numbers are in Column B. Under the Order section, choose Largest to Smallest. Select Okay.



4. Under the numbers column, total the figures. Be sure you have selected the cell under column you want totaled. Select the auto sum button.



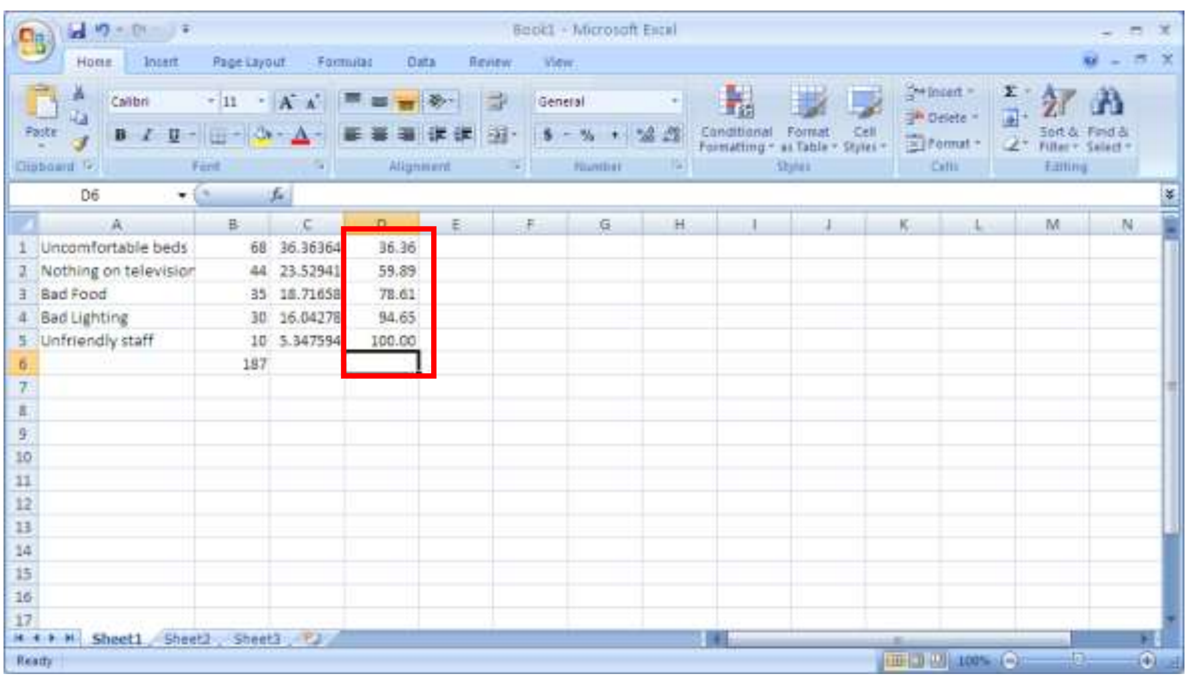
5. In column C, the percentages need to be calculated. You do this by dividing the category count by the total and multiplying by 100. In this case, your figures should match what is below. (68/187) *100 = 36.36364



	A	B	C
1	Uncomfortable beds	68	36.36364
2	Nothing on televisior	44	23.52941
3	Bad Food	35	18.71658
4	Bad Lighting	30	16.04278
5	Unfriendly staff	10	5.347594
6		187	
7			

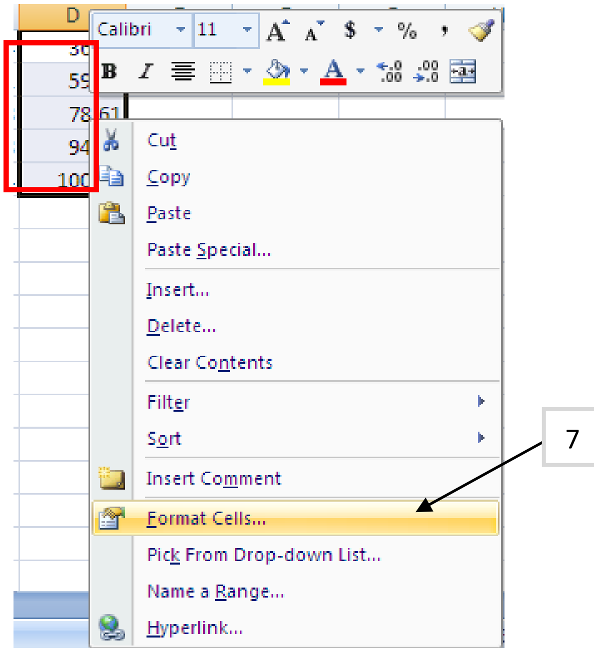
6. Once the percentages have been calculated, in column C, the percentages need to be added together in the following format:
C1 = D1
D2=D1+C2
D3=D2+C3
D4=D3+C4
D5=D4+C5

Continue this format until you have added all the categories. Your totals should reflect what is shown below.

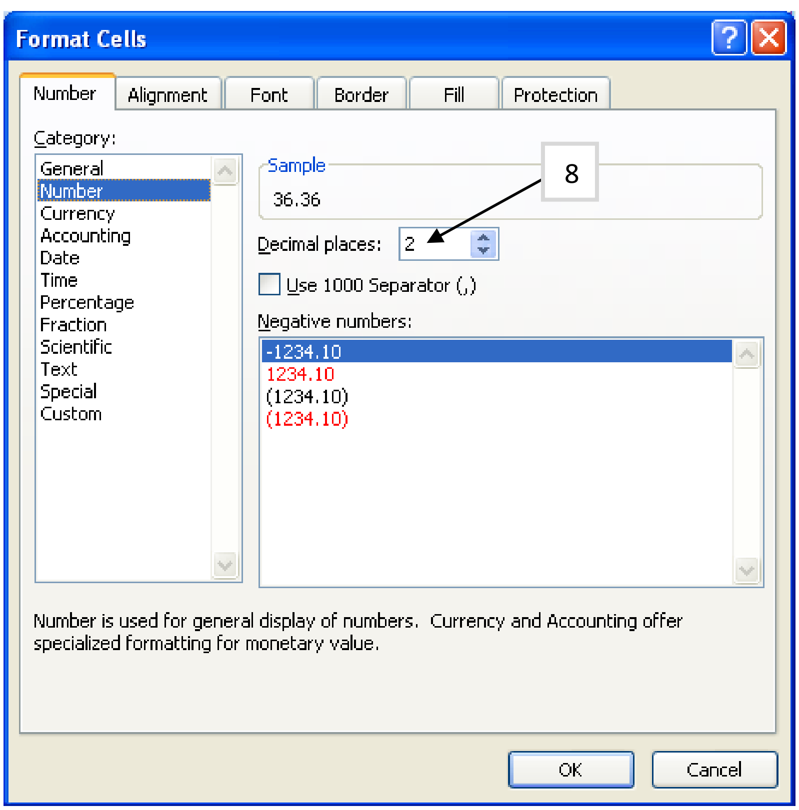


	A	B	C	D
1	Uncomfortable beds	68	36.36364	36.36
2	Nothing on televisior	44	23.52941	59.89
3	Bad Food	35	18.71658	78.61
4	Bad Lighting	30	16.04278	94.65
5	Unfriendly staff	10	5.347594	100.00
6		187		

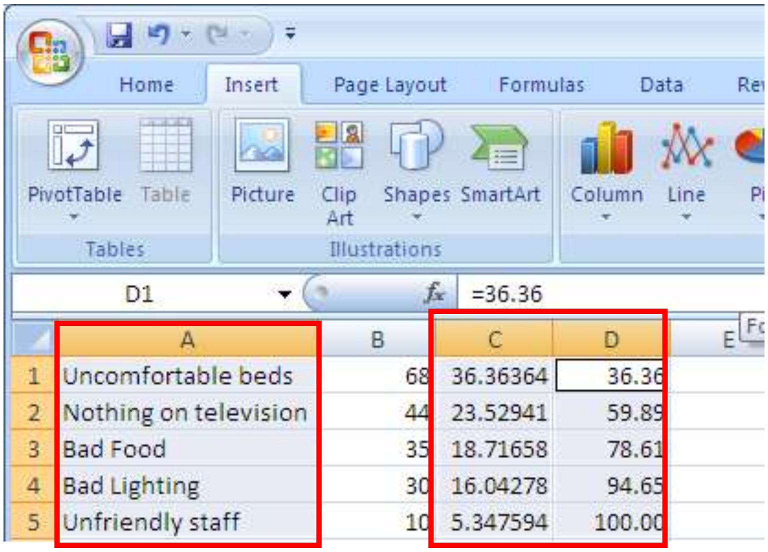
7. Format your numbers to round to 2 decimal points. Do this by highlighting the data, right click and select Format Data.



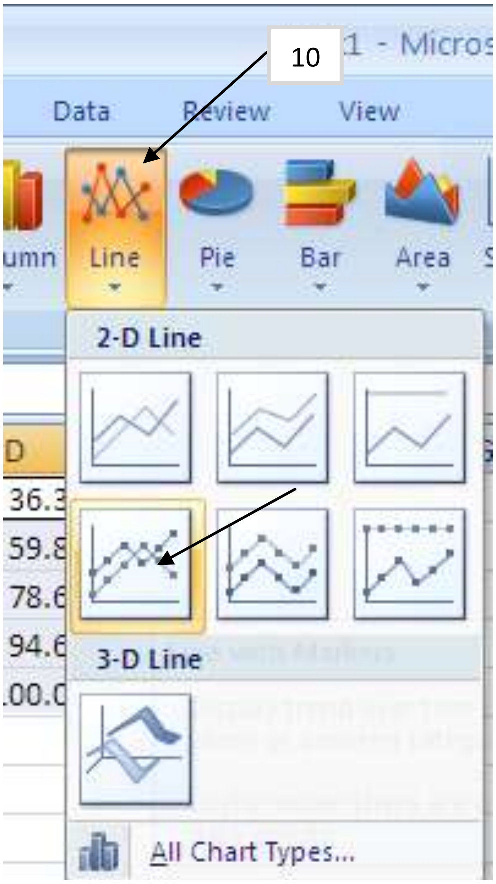
8. Select number category and set the decimal places to 2. Select okay.



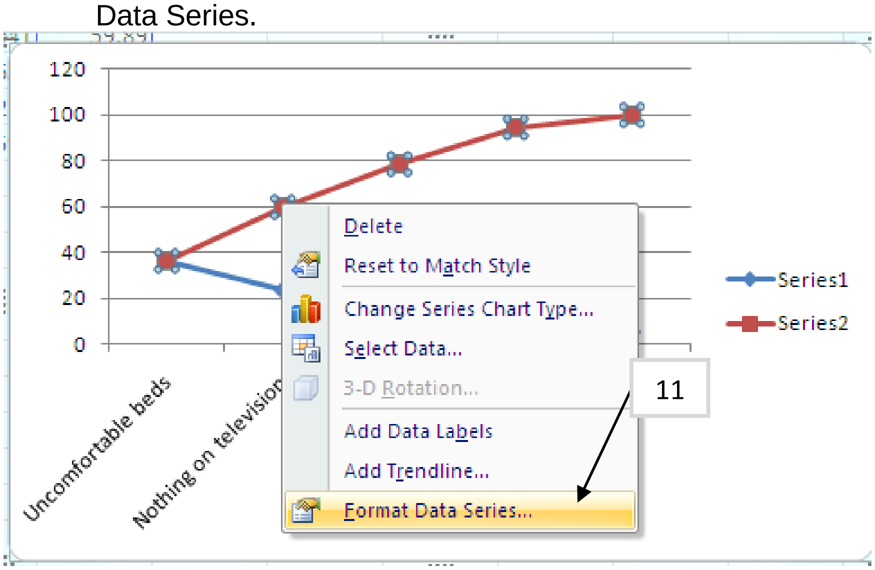
9. Highlight columns A, C, and D. Select the Insert Tab on the Ribbon.



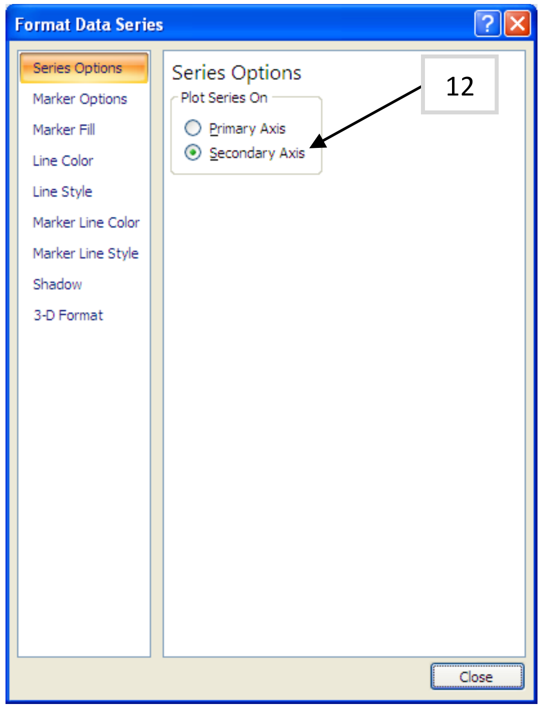
10. From the Line Graphs, select the Lines with Markers chart.



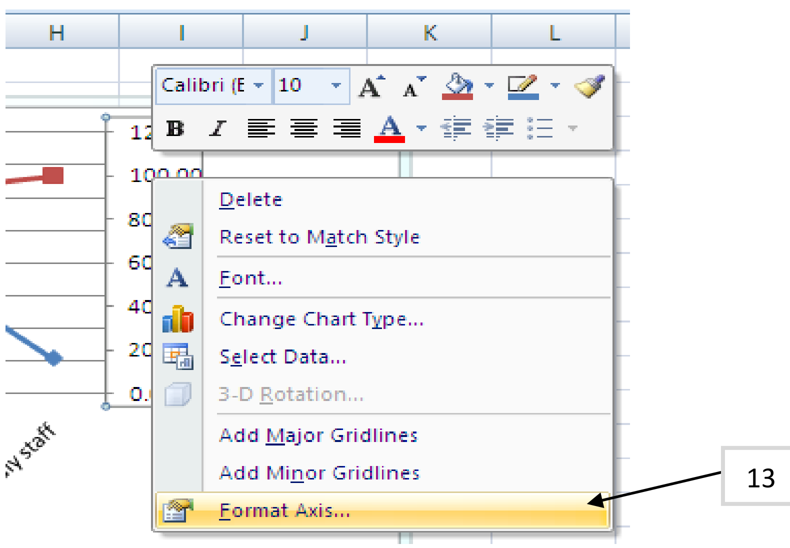
11. The chart will be inserted in the worksheet. Right click on the red line. Select Format Data Series.



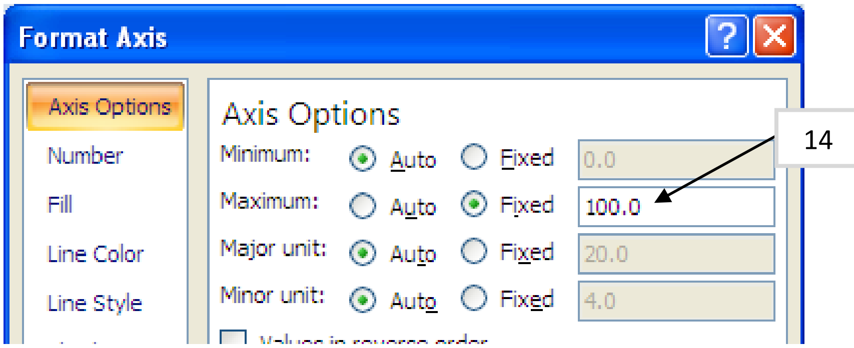
12. Under Series Options, select secondary axis. Hit close.



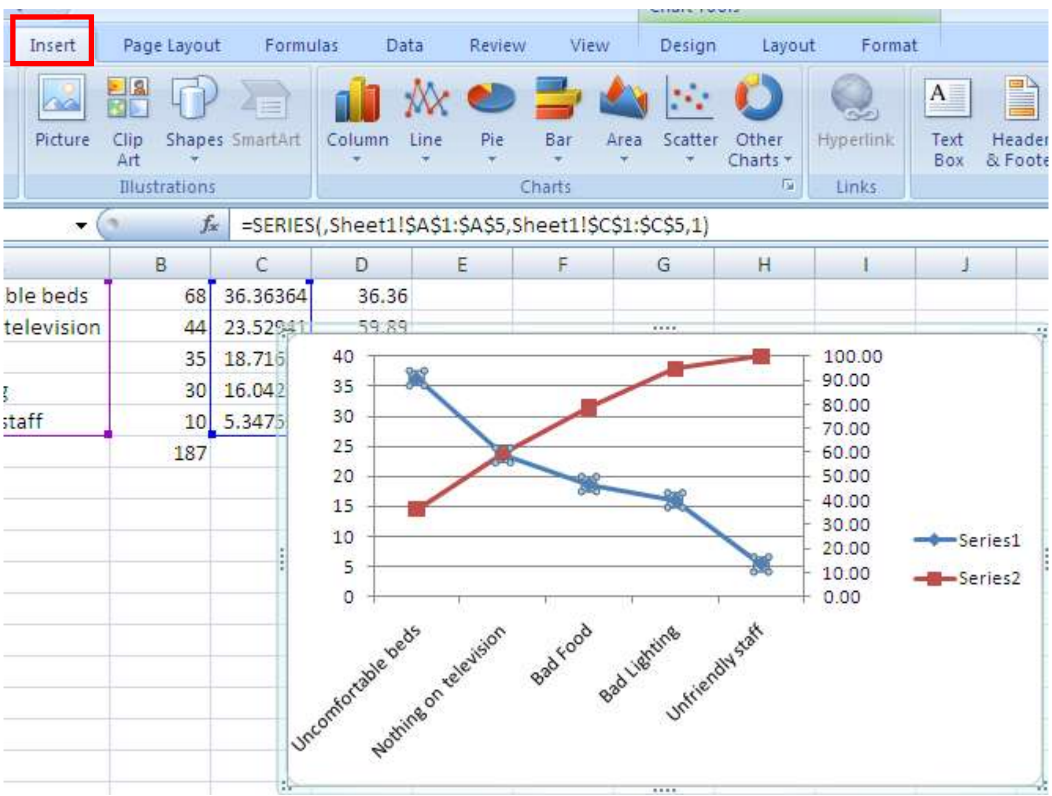
13. Once the right axis is added, right click on it and select Format Axis.



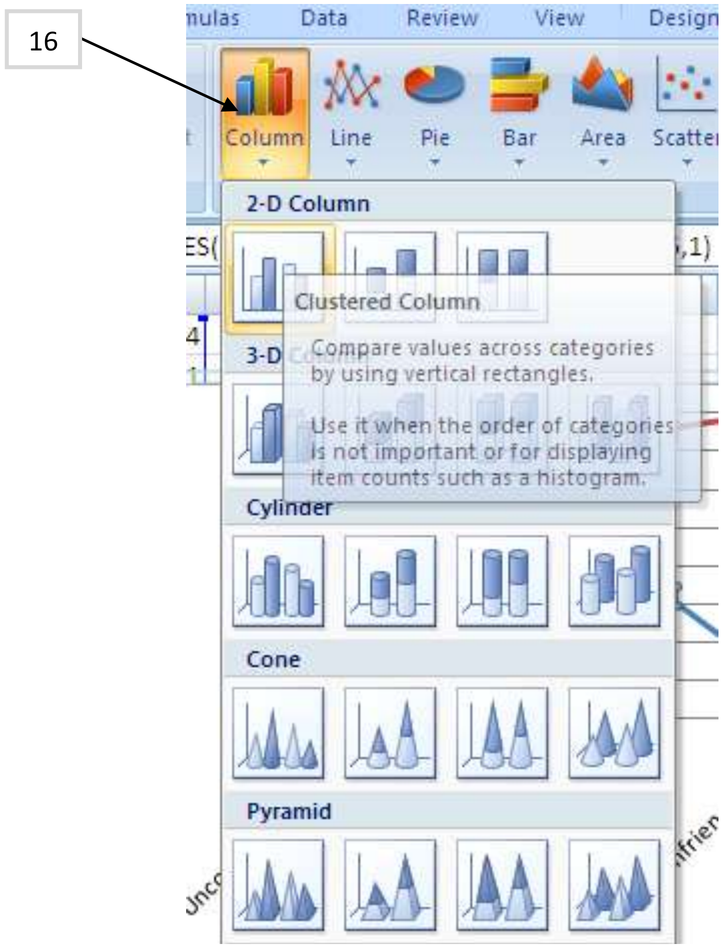
14. Change the Fixed Maximum Amount to 100.00. Select close.



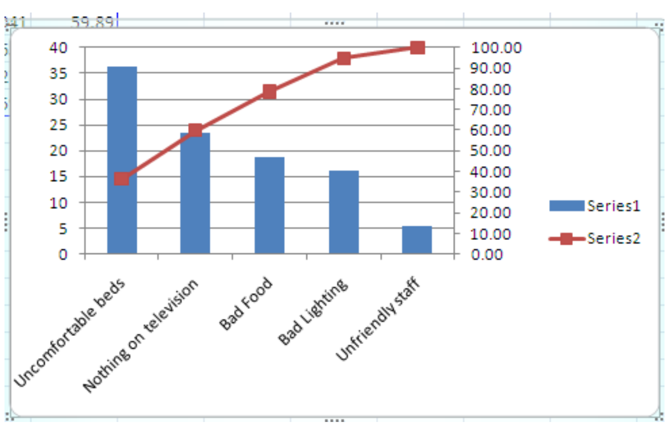
15. Select the blue line. With the blue line selected, click the Insert Tab in the Ribbon.



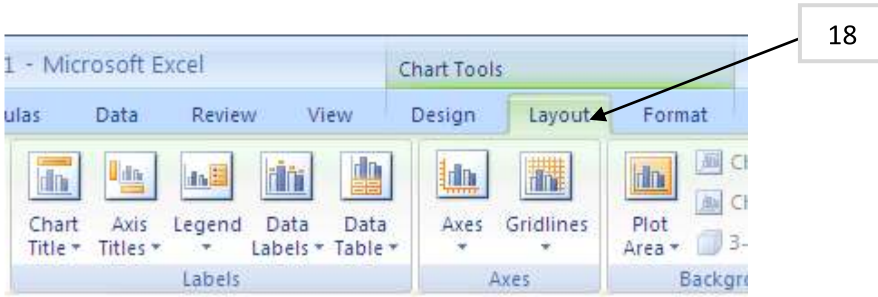
16. Select the Column Chart, and select the Clustered Column option. This will change the line to bars.



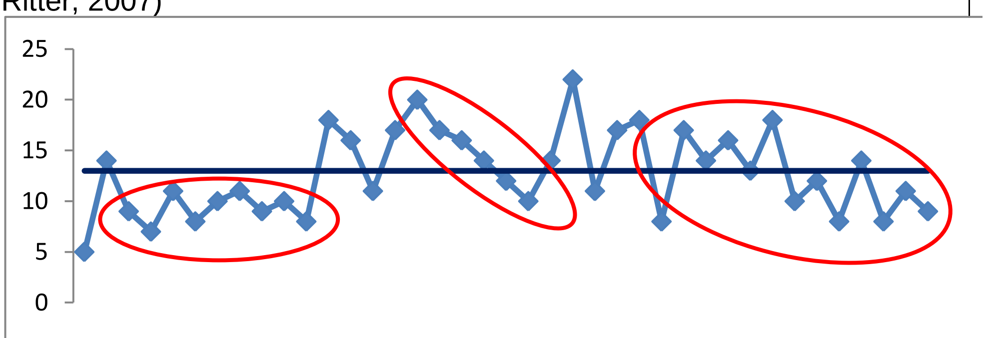
17. Once selected your chart will reflect what is below.



18. Select the legend and hit delete. This will remove the legend from the graph. Under the layout tab, chart titles and axis labels can be modified/added.

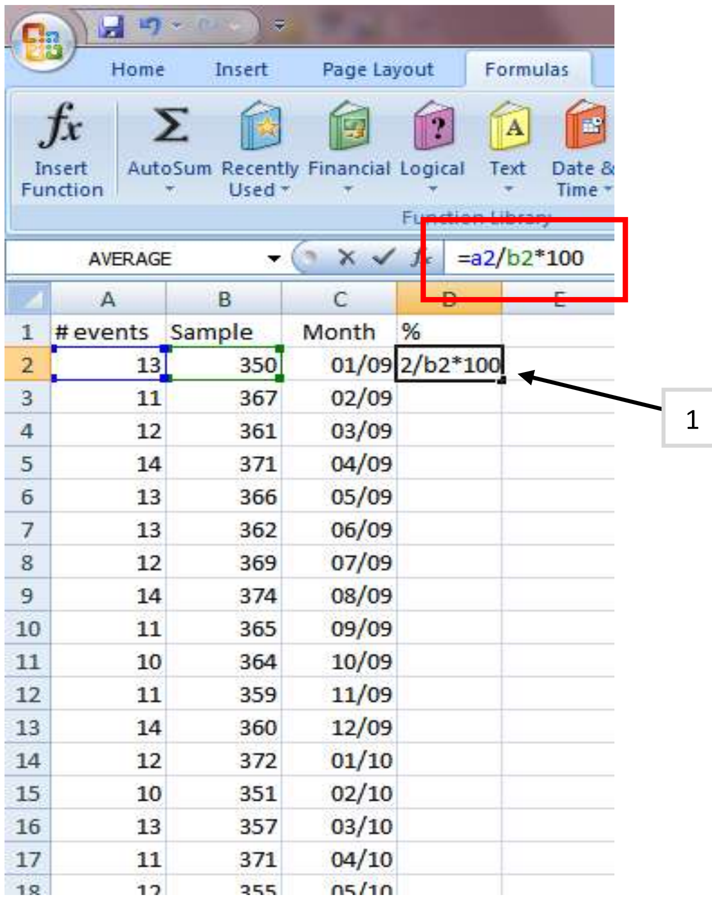


19. To customize the chart, refer to the formatting section.

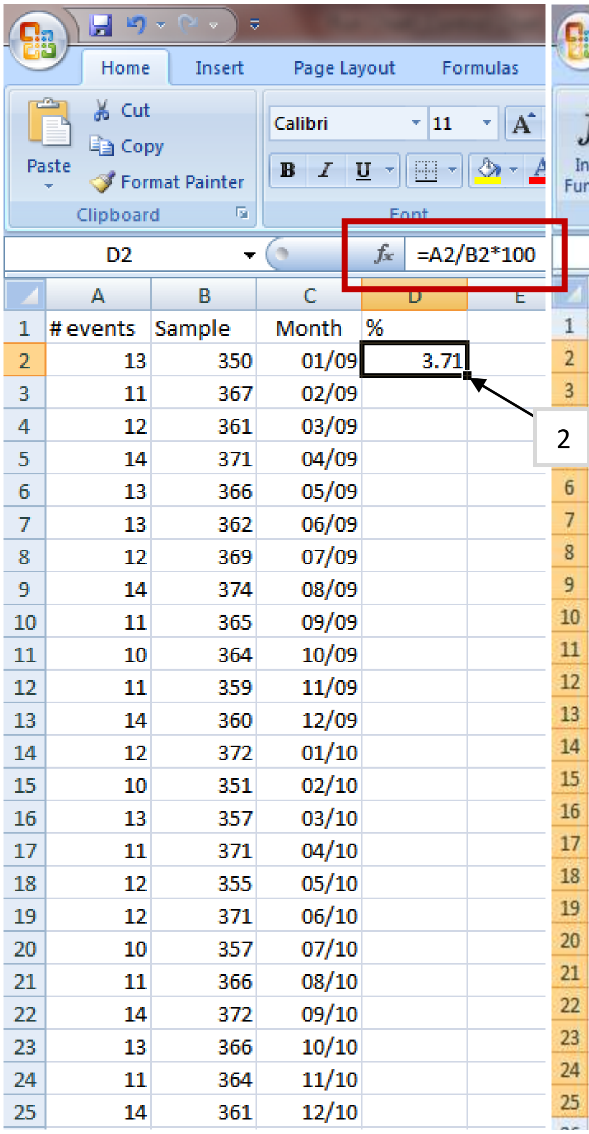
Run Chart	
Usage	Monitoring trends and/or changes in processes or events over time. Run charts are not very helpful for detecting special cause, but general rules regarding observation of data points for trends or shifts in the process may indicate further investigation is needed. A run chart is the precursor of a statistical process control chart.
Data Types	Any attribute or variable data that can be measured on a time scale. In public health practice surveillance data is an example of events that are measured over time.
Creation in Excel	1. A minimum of 20-25 data points is recommended for the construction of a run chart (Brassard & Ritter, 2007). 2. Determine a meaningful time interval for the data of interest. This may be days, weeks, months, etc. The time sequence is plotted as a consistent interval on the x-axis. 3. The measurement data is plotted in reference to the y-axis. 4. The central measure, which serves as a reference point, for a run chart is the median 5. The data elements should populate columns, with the appropriate header at the top. 6. The chart in this example uses percentages, the number of events that occurred each month (numerator) divided by the sample size (denominator) each month multiplied by 100.
Interpre-tation	1. When trending data over time it is extremely important that the data source, definitions used for the measure or sample size (denominator if calculating a proportion or rate), time interval, etc. remain the same. Any change in the data source, collection or analysis methods will render comparisons invalid. 2. Interpretation of the chart should be done by someone with underlying knowledge of the process. 3. General rules that may indicate a need for further investigation of the process include: a. 9 consecutive values above or below the center line b. 6 consecutive increases or decreases c. 14 consecutive points alternately decreasing and increasing (Brassard & Ritter, 2007) 

Performing the calculations and populating the cells needed for the chart.

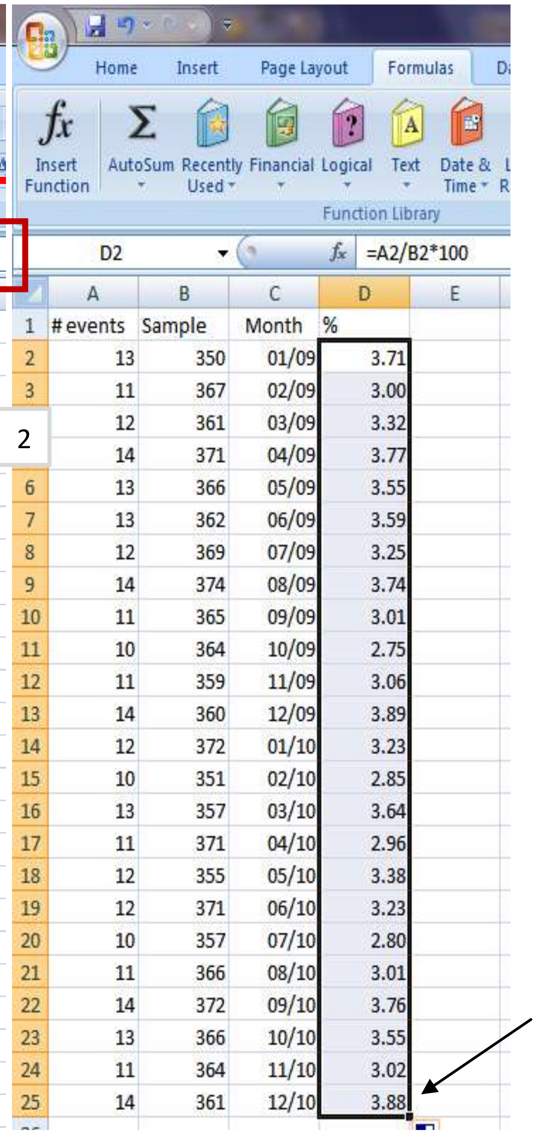
1. Column D must be populated with the percentage rate. The easiest way to do this is to highlight the cell where the value is to be placed and type the formula using the column/cell where the values are found.



2. To calculate the values for the remaining months and populate the corresponding cells, place the cursor on the bottom lower right corner of the cell and a + will appear. Click the left mouse button. Holding the left mouse button down, highlight the remainder of the column where the values are to be placed.
3. The cells will be populated with a formula that applies to that specific row of values.

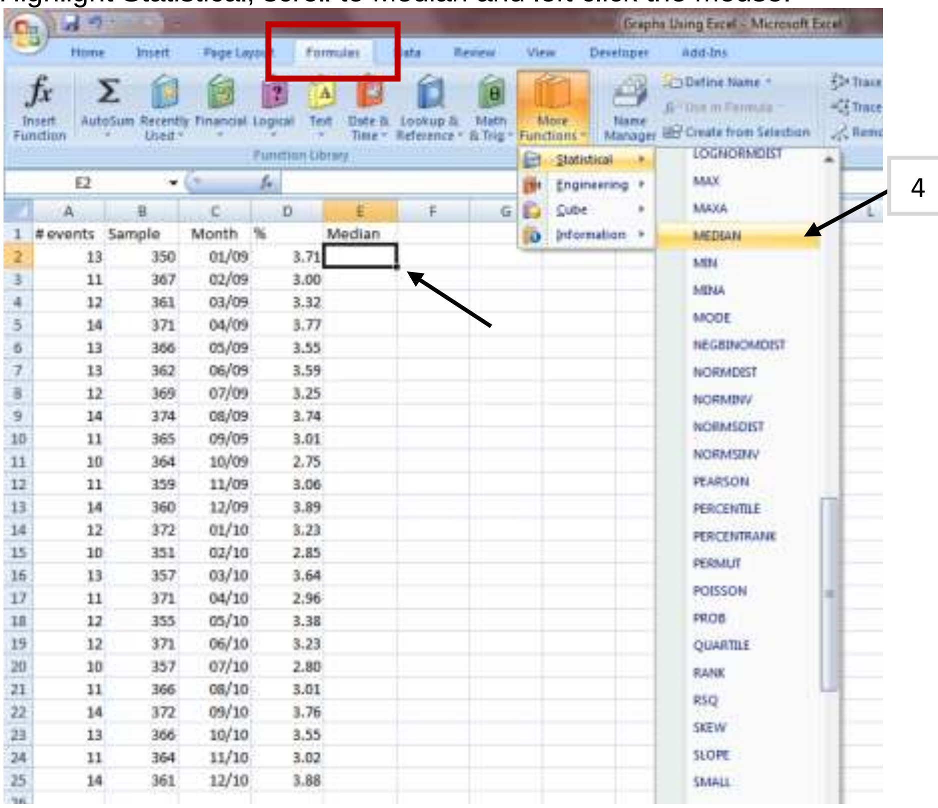


	A	B	C	D
1	# events	Sample	Month	%
2	13	350	01/09	3.71
3	11	367	02/09	
4	12	361	03/09	
5	14	371	04/09	
6	13	366	05/09	
7	13	362	06/09	
8	12	369	07/09	
9	14	374	08/09	
10	11	365	09/09	
11	10	364	10/09	
12	11	359	11/09	
13	14	360	12/09	
14	12	372	01/10	
15	10	351	02/10	
16	13	357	03/10	
17	11	371	04/10	
18	12	355	05/10	
19	12	371	06/10	
20	10	357	07/10	
21	11	366	08/10	
22	14	372	09/10	
23	13	366	10/10	
24	11	364	11/10	
25	14	361	12/10	

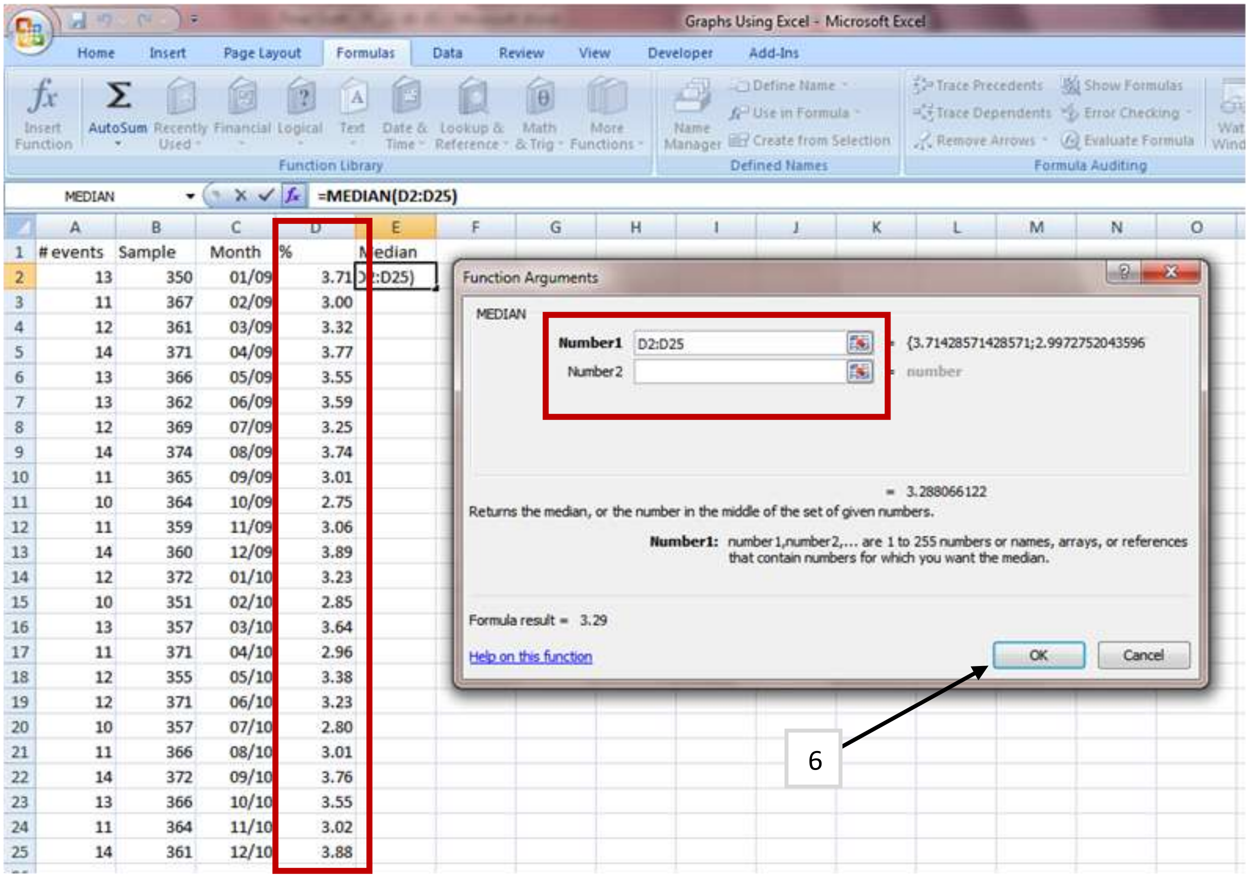


	A	B	C	D
1	# events	Sample	Month	%
2	13	350	01/09	3.71
3	11	367	02/09	3.00
4	12	361	03/09	3.32
5	14	371	04/09	3.77
6	13	366	05/09	3.55
7	13	362	06/09	3.59
8	12	369	07/09	3.25
9	14	374	08/09	3.74
10	11	365	09/09	3.01
11	10	364	10/09	2.75
12	11	359	11/09	3.06
13	14	360	12/09	3.89
14	12	372	01/10	3.23
15	10	351	02/10	2.85
16	13	357	03/10	3.64
17	11	371	04/10	2.96
18	12	355	05/10	3.38
19	12	371	06/10	3.23
20	10	357	07/10	2.80
21	11	366	08/10	3.01
22	14	372	09/10	3.76
23	13	366	10/10	3.55
24	11	364	11/10	3.02
25	14	361	12/10	3.88

4. The median value for the center line can be obtained using the statistical functions in Excel.
    a. Place the curser in the cell where you want the value placed
    b. Click on More Functions in the Formulas tab.
    c. Highlight Statistical, scroll to Median and left click the mouse.



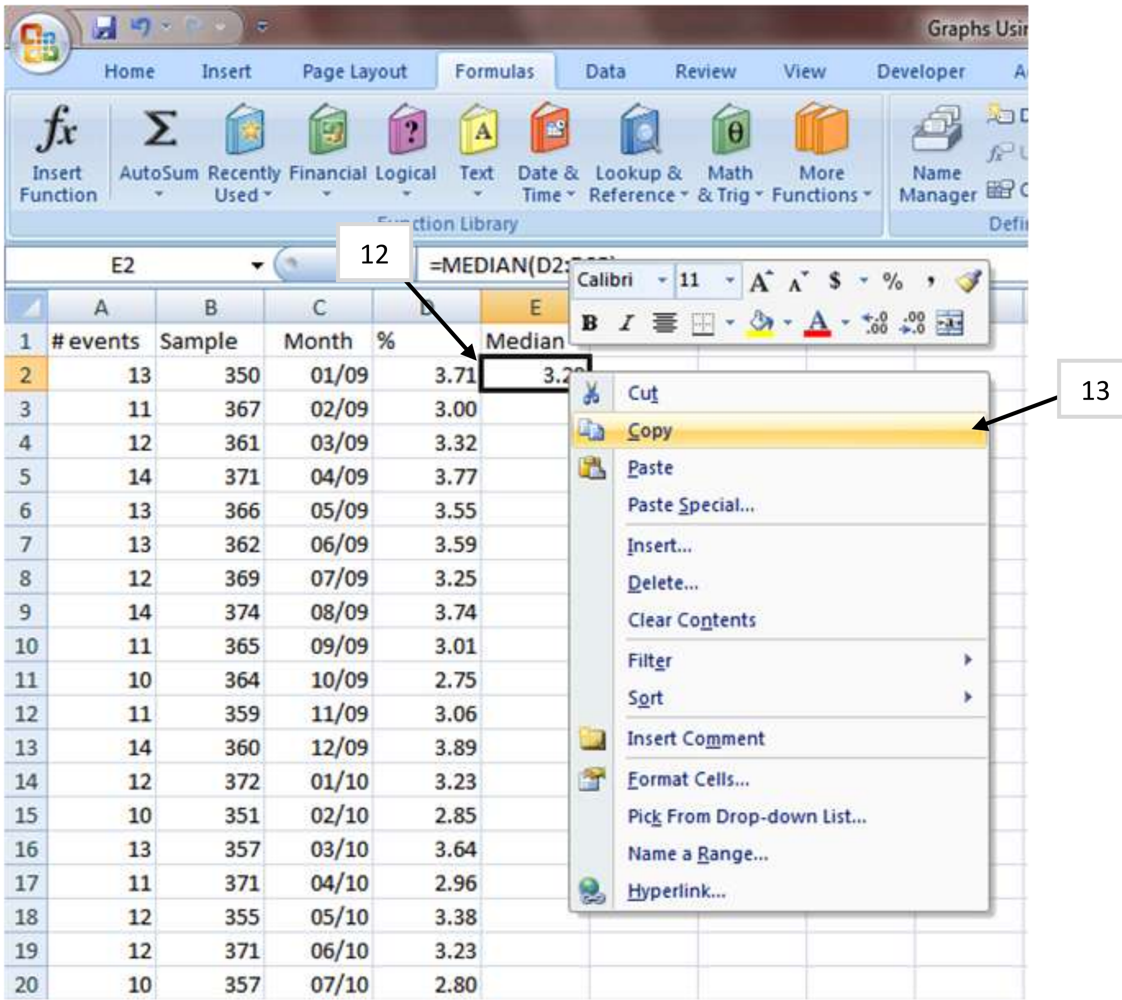
5. A dialogue box will appear asking for the range of values that will be used to obtain the median. In the box titled Number 1 type D2:D25.



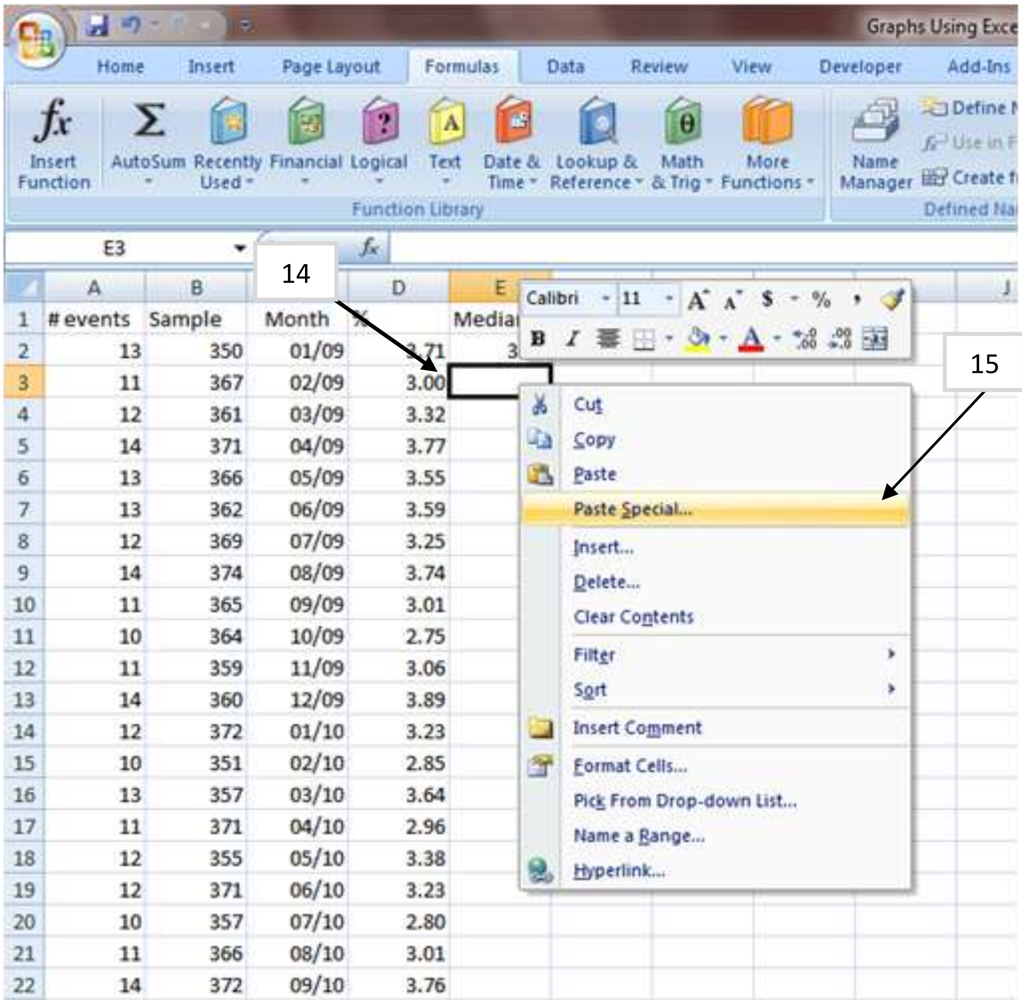
6. Click OK.
7. Alternatively, click on the arrow to the right of the box titled Number 1.
8. When a new dialogue box appears (below), click the arrow on the right.
9. Highlight the range of cells that contain the values to be used.
10. Press Enter. You will be returned to the previous dialogue box (above).
11. Click OK.

12. There are various methods to provide the data that tells the program to chart the mean line. The simplest is to populate the values in a column that will provide the value for a data series. To populate the cells that provide the data right click the mouse in the cell that contains the median value.

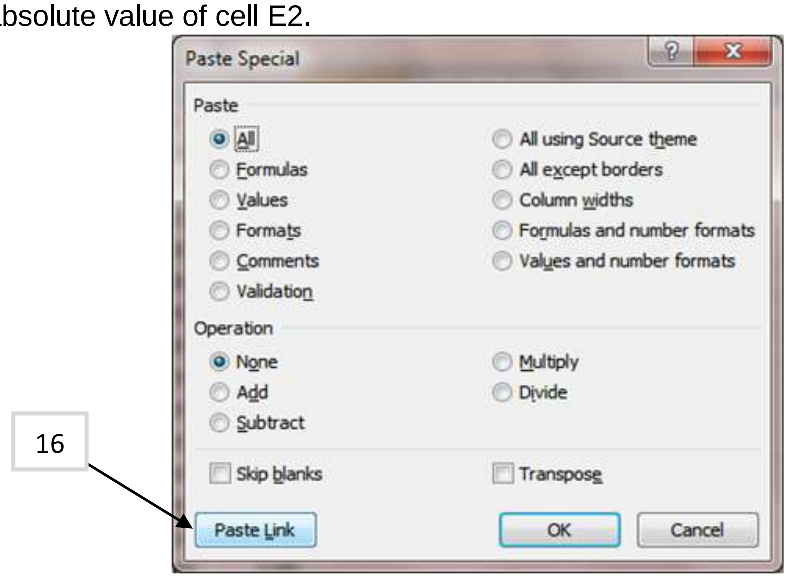
13. Click Copy.



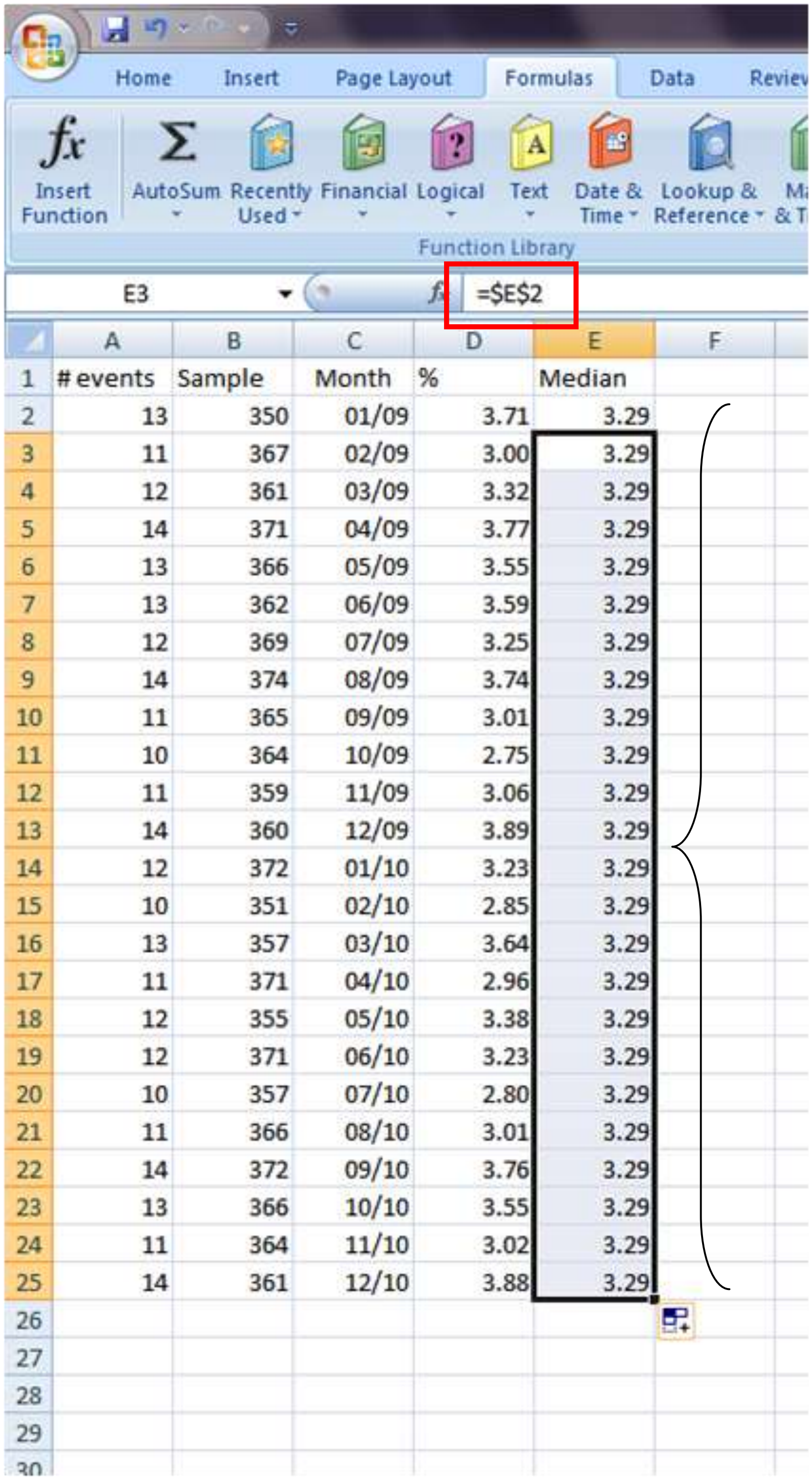
14. Place the cursor in the next cell down, cell E3.
15. Right click the mouse and left click on Paste Special.



16. Click Paste Link in the dialogue box that appears. This tells the program to populate cell E3 with the absolute value of cell E2.



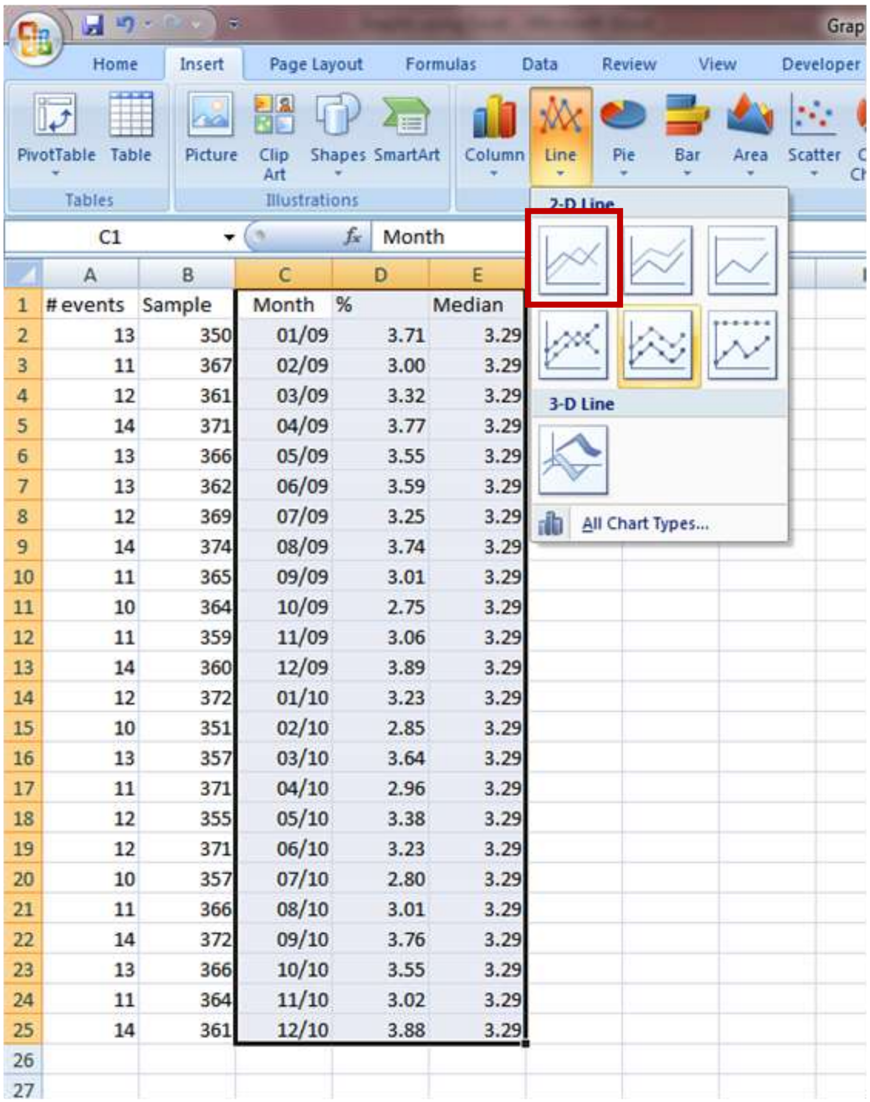
17. Fill the remainder of the cells in column E by placing the cursor on the lower right corner of cell E3, right click, hold, drag and highlight the cells to be filled.



	A	B	C	D	E	F
1	# events	Sample	Month	%	Median	
2	13	350	01/09	3.71	3.29	
3	11	367	02/09	3.00	3.29	
4	12	361	03/09	3.32	3.29	
5	14	371	04/09	3.77	3.29	
6	13	366	05/09	3.55	3.29	
7	13	362	06/09	3.59	3.29	
8	12	369	07/09	3.25	3.29	
9	14	374	08/09	3.74	3.29	
10	11	365	09/09	3.01	3.29	
11	10	364	10/09	2.75	3.29	
12	11	359	11/09	3.06	3.29	
13	14	360	12/09	3.89	3.29	
14	12	372	01/10	3.23	3.29	
15	10	351	02/10	2.85	3.29	
16	13	357	03/10	3.64	3.29	
17	11	371	04/10	2.96	3.29	
18	12	355	05/10	3.38	3.29	
19	12	371	06/10	3.23	3.29	
20	10	357	07/10	2.80	3.29	
21	11	366	08/10	3.01	3.29	
22	14	372	09/10	3.76	3.29	
23	13	366	10/10	3.55	3.29	
24	11	364	11/10	3.02	3.29	
25	14	361	12/10	3.88	3.29	
26						
27						
28						
29						
30						

Creating and formatting the chart

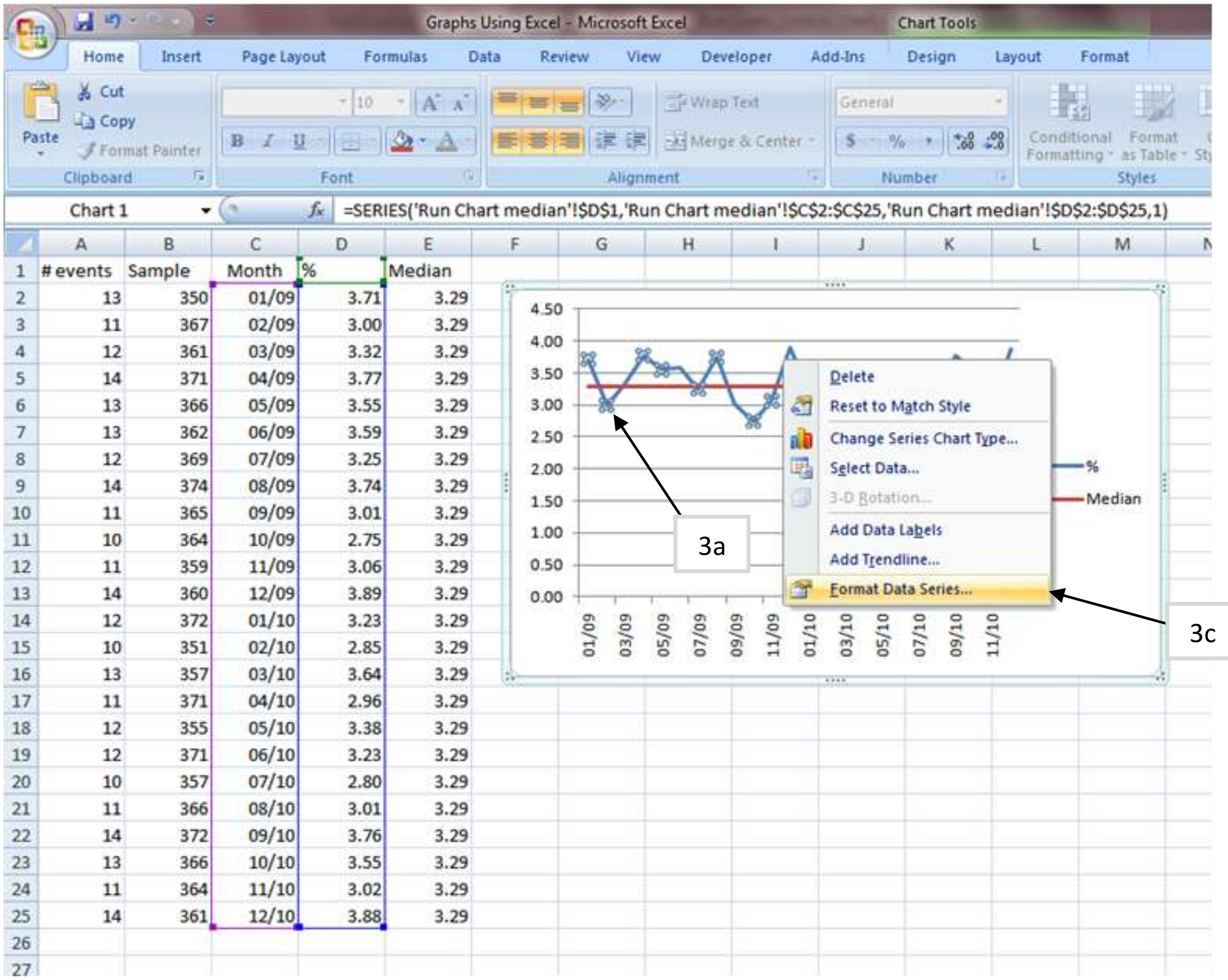
1. A run chart is called a line graph in Excel. To create the chart:
   a. Click on the Insert tab.
   b. Highlight the data and headers that will be used to create the chart.
   c. Click on the Line chart and highlight the desired type of chart. For this example a line chart without data markers is used as the foundation to build the chart.



2. The chart will automatically be placed in the existing worksheet.

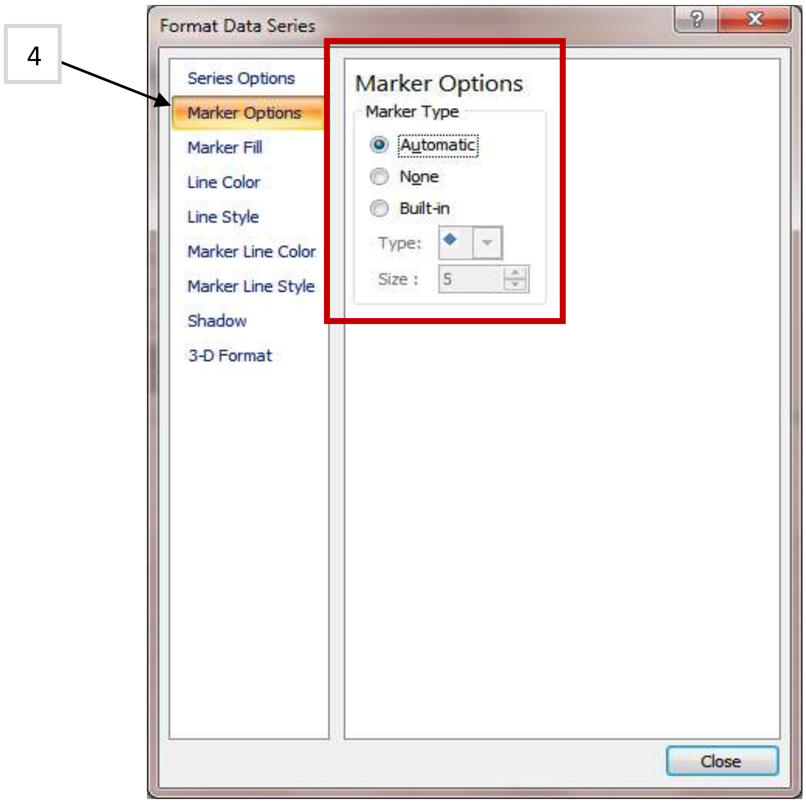
3. For the sake of visibility, it may be helpful to add markers at each data point. To do this:

a. Click on the data series. Markers will appear indicating the data series has been selected.
b. Right click the mouse
c. Highlight and left click on Format Data Series. A new dialogue box will appear.

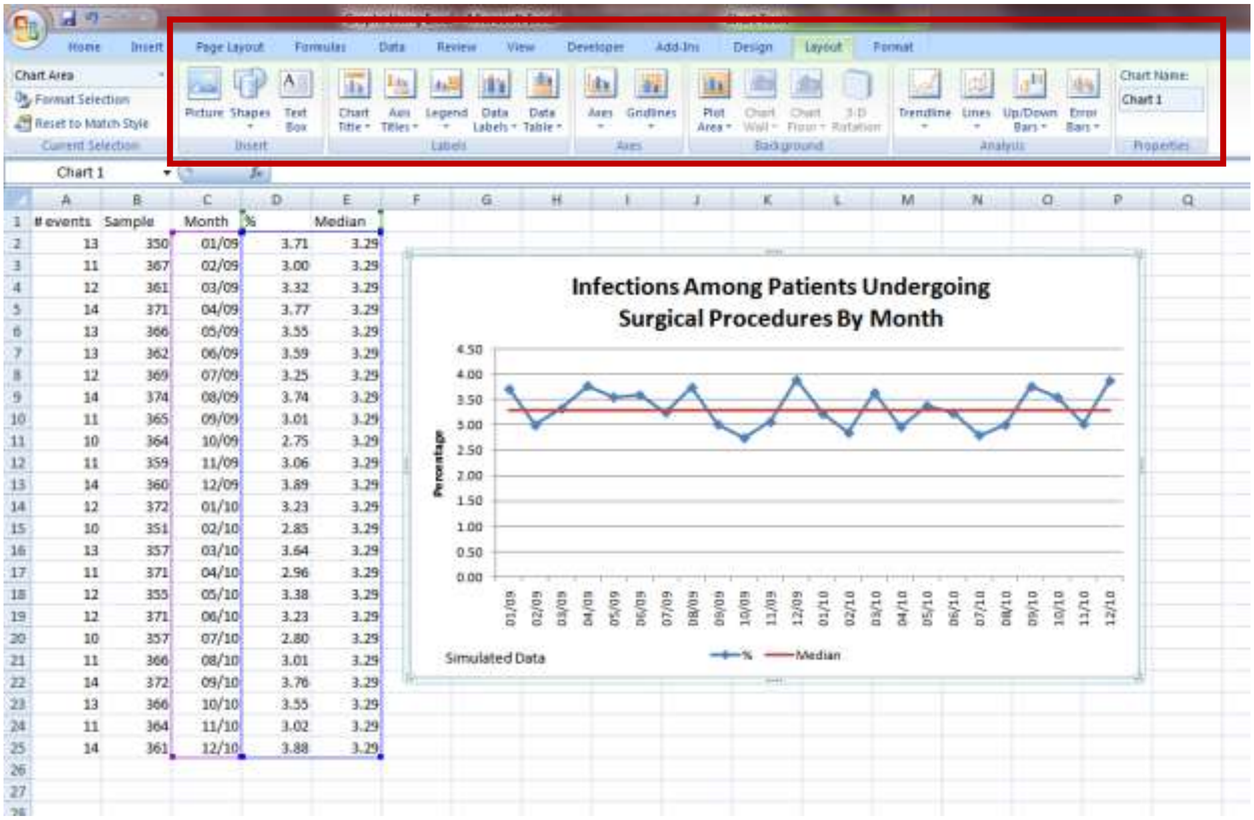


4. Select Marker Options on the left of the dialogue box that appears.

5. Select the desired marker using the radio buttons in the right side of the Format Data Series box.



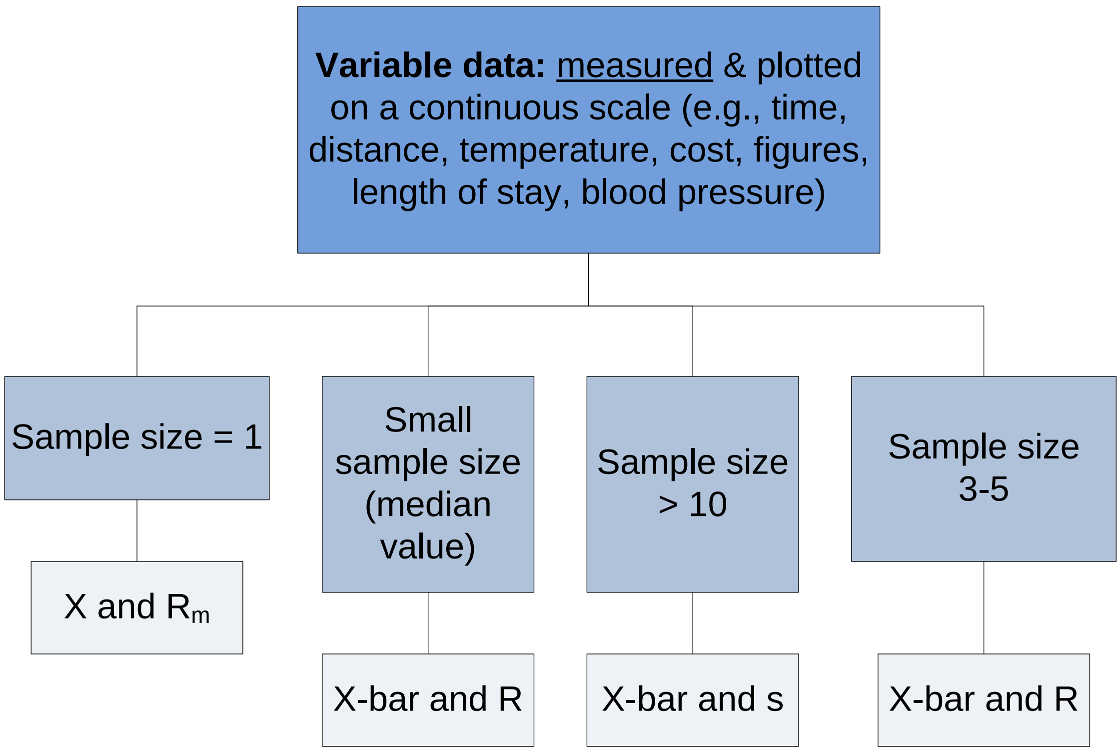
6. Format the chart axis, labels and legend using the Chart Tools, Layout tab options.
7. Add an explanatory title to the chart.

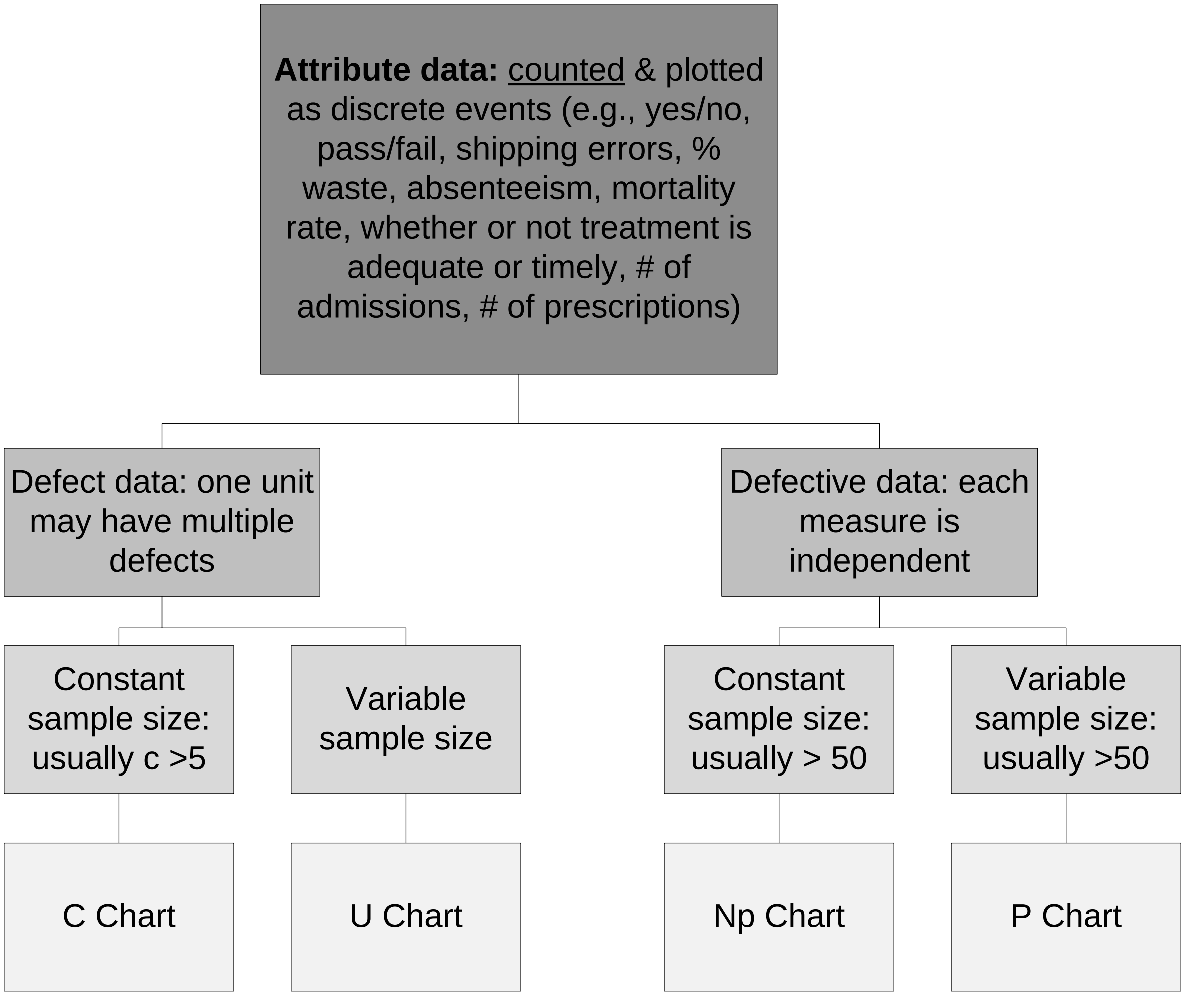

Home
Insert
Page Layout
Formulas
Data
Review
View
Developer
Add-Ins
Design
Layout
Format
Chart Area
Format Selection
Reset to Match Style
Current Selection
Picture
Shapes
Text Box
Insert
Chart Title
Axis Titles
Legend
Data Labels
Data Table
Labels
Axes
Gridlines
Axes
Plot Area
Chart Wall
Chart Floor
3-D Rotation
Background
Trendline
Lines
Up/Down Bars
Error Bars
Analysis
Chart Name:
Chart 1
Properties
Chart 1
events Sample Month % Median
13 350 01/09 3.71 3.29
11 367 02/09 3.00 3.29
12 361 03/09 3.32 3.29
14 371 04/09 3.77 3.29
13 366 05/09 3.55 3.29
13 362 06/09 3.59 3.29
12 369 07/09 3.25 3.29
14 374 08/09 3.74 3.29
11 365 09/09 3.01 3.29
10 364 10/09 2.75 3.29
11 359 11/09 3.06 3.29
14 360 12/09 3.89 3.29
12 372 01/10 3.23 3.29
10 351 02/10 2.85 3.29
13 357 03/10 3.64 3.29
11 371 04/10 2.96 3.29
12 355 05/10 3.38 3.29
12 371 06/10 3.23 3.29
10 357 07/10 2.80 3.29
11 366 08/10 3.01 3.29
14 372 09/10 3.76 3.29
13 366 10/10 3.55 3.29
11 364 11/10 3.02 3.29
14 361 12/10 3.88 3.29
Infections Among Patients Undergoing Surgical Procedures By Month
Percentage
Simulated Data
%
Median

Control Chart

Usage	A control chart builds on the run chart by adding upper and lower control limits based on statistical methods. The premise of the control chart is that the application of these methods provides guidance as to the probable cause of variation in the process being monitored. The underlying assumption is that almost all of the variation that occurs in a process is random, or expected, and that special cause variation is powerful, but occurs very rarely. However, the observation that a process is in statistical control does not mean that there is not opportunity for improvement.
Data Types	Attribute and variable. There are several different types of control charts. Selection of the appropriate chart is dependent on the type of data, sample and sample size. The calculation of control limits are specific to the type of chart used (Brassard & Ritter, 2007).
Creation in Excel	1. The calculation of control limits should be undertaken only when a process is believed to have been in control, i.e. no trends, unusual runs or evident issues. A minimum of 25 data points are recommended for the calculation of control limits. Control limits should not be recalculated unless there is a process change. 2. The industrial standard for control limits is 3 standard deviations (σ) above and below the mean. The use of 3σ for control limits provides a 99.7% probability that measures within the control limits are the result of expected variation in a process (Hansen, 2005). There is some debate about the use of 3σ for healthcare because of the complex nature of human pathology and the potential seriousness of waiting too long to investigate the root cause of an out of control signal (Sellick, 1993). The use of 2σ results in 95.4% probability that variation within the limits is random. However, there is a slight decrease in specificity, or that a random event could be mistaken for special cause (Hansen, 2003). For the purpose of this example, both 2σ and 3σ control limits are created. 3. While it is tempting to use the Excel Standard Deviation function to calculate the standard deviation for all of the monthly percentage values, this may be an over simplification of accepted statistical methods (Benneyan, 1998). For this example the standard formula is used to calculate control limits. 4. The example below creates a p-chart based on the data provided for the creation of the run chart example above. 5. Use the average sample size (n-bar) creates a static control limit for ease of creation in Excel.
Interpretation	The occurrence of measurements that signal the need for further investigation include: 1. Any single measurement outside the 3σ upper or lower control limit. 2. Eight consecutive measurements above or below the center line. 3. Three consecutive measurements in the range between 2σ and 3σ. 4. Five consecutive measurements in the range between 1σ and 2σ. 5. Patterns, runs or trends as discussed in the Run Chart example.

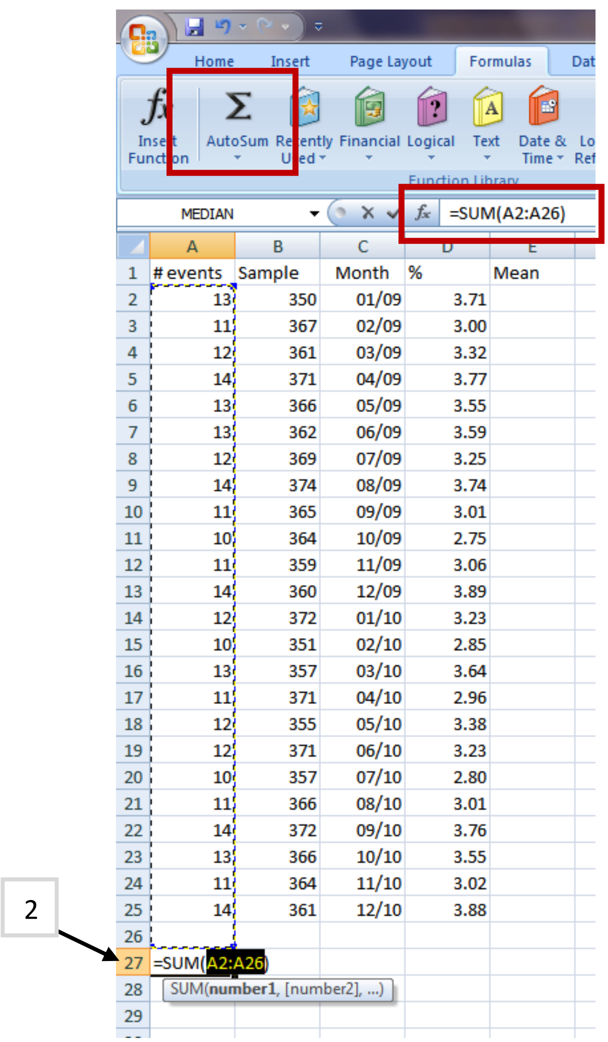
Selecting the appropriate chart (Brassard & Ritter, 2007)

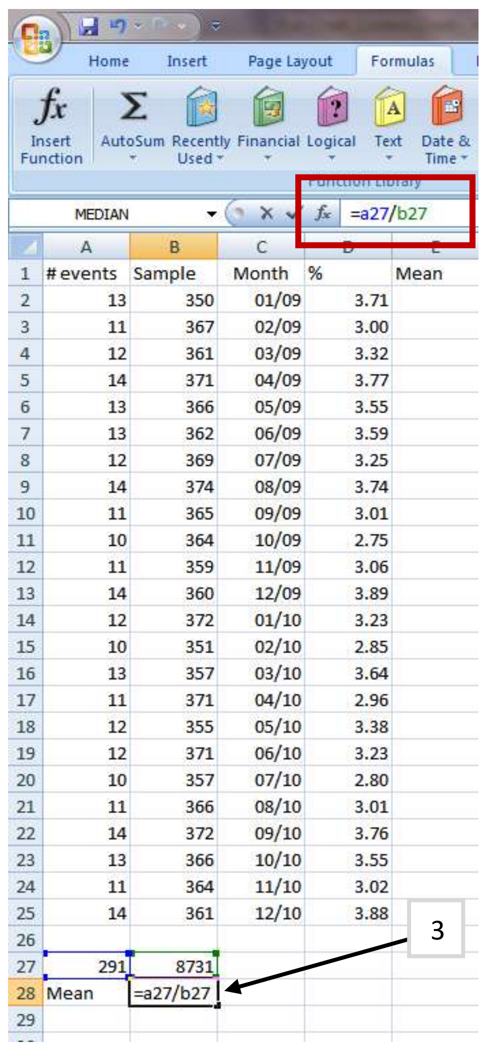



Attribute data: counted & plotted as discrete events (e.g., yes/no, pass/fail, shipping errors, % waste, absenteeism, mortality rate, whether or not treatment is adequate or timely, # of admissions, # of prescriptions)
Defect data: one unit may have multiple defects
Defective data: each measure is independent
Constant sample size: usually c >5
Variable sample size
Constant sample size: usually > 50
Variable sample size: usually >50
C Chart
U Chart
Np Chart
P Chart

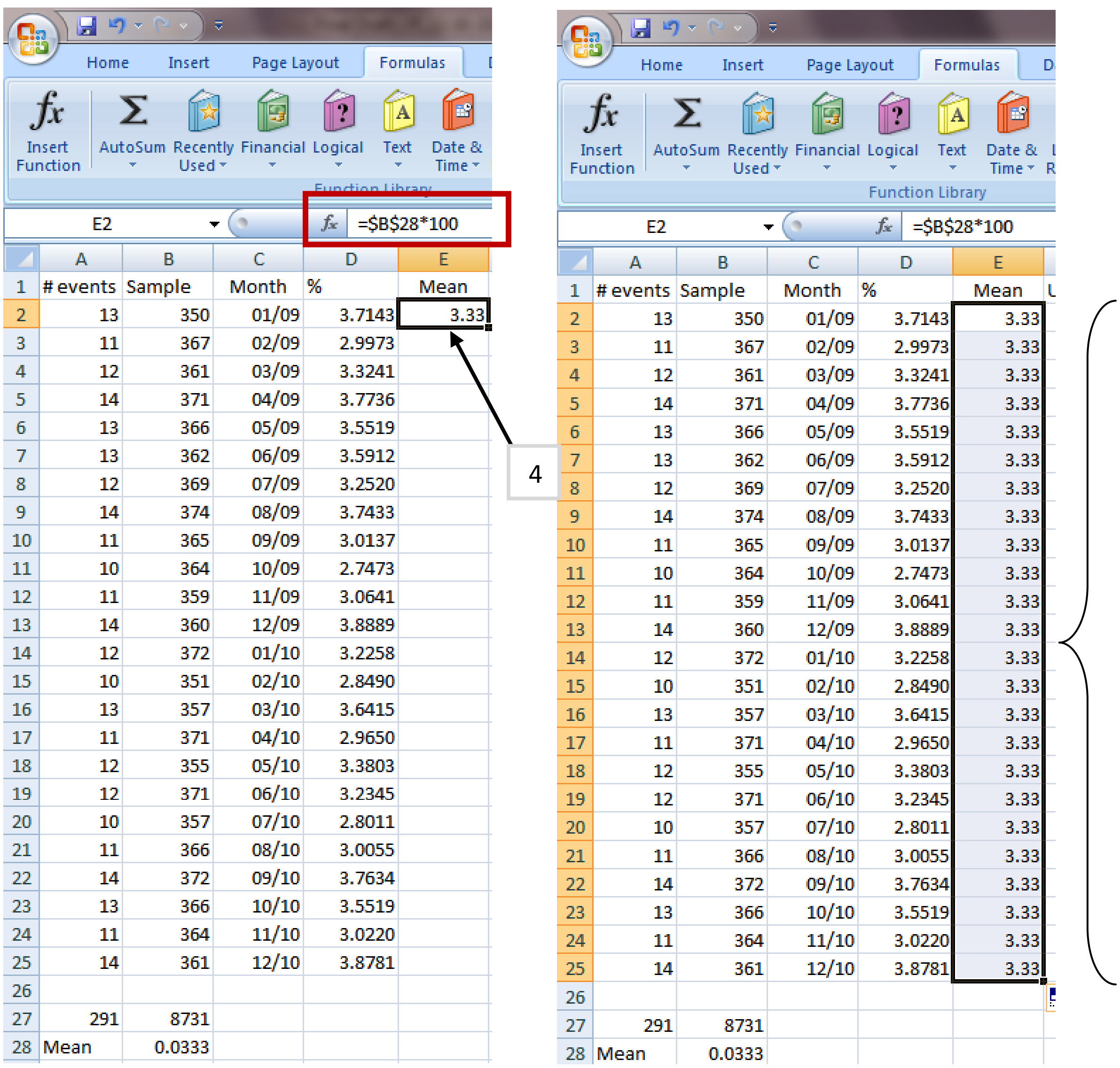
Calculating the values and populating the cells for the chart.

1. The p chart uses the mean for the center line and to calculate control limits.
2. The AutoSum function is used here to obtain the total number of events for the entire monitoring period and the total sample size. These values are used to calculate the mean.
    a. Highlight the cell below the values to be summed.
    b. Click the AutoSum function in the Formulas tab and press enter.
    c. The result will be placed in the highlighted cell.
    d. Complete this action for both the events and the samples.
3. To calculate the mean:
    e. Highlight the cell where the value is to be placed.
    f. In the formula bar type = a27/b27.
    g. Press enter.





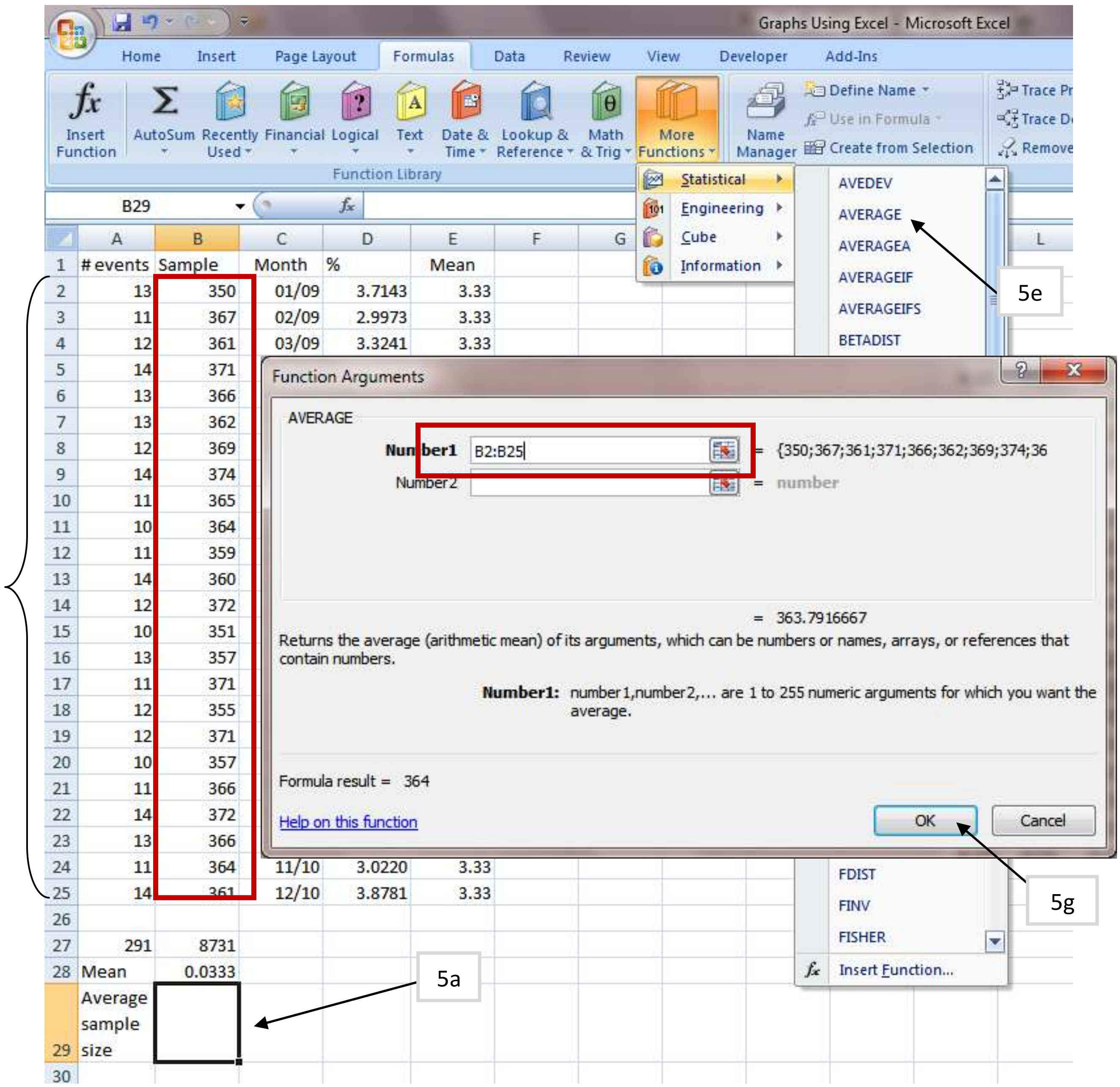
4. The value for the mean must now be placed in the column that the program will use to chart the line.
    a. Highlight the mean value (cell B28).
    b. Right click the mouse and click Copy in the drop down box that appears.
    c. Click in the cell at the top of the column that will provide the data for the mean line.
    d. Right click, highlight Paste Special and left click.
    e. Place the cursor in the f_x bar to the right of the 8 and type *100. This multiplies the absolute value of B28 by 100, changing the decimal to a percentage.
    f. T fill the remaining cells in column E left click the lower right corner of cell E2, hold the mouse button down, drag and highlight.



	A	B	C	D	E
1	# events	Sample	Month	%	Mean
2	13	350	01/09	3.7143	3.33
3	11	367	02/09	2.9973	
4	12	361	03/09	3.3241	
5	14	371	04/09	3.7736	
6	13	366	05/09	3.5519	
7	13	362	06/09	3.5912	
8	12	369	07/09	3.2520	
9	14	374	08/09	3.7433	
10	11	365	09/09	3.0137	
11	10	364	10/09	2.7473	
12	11	359	11/09	3.0641	
13	14	360	12/09	3.8889	
14	12	372	01/10	3.2258	
15	10	351	02/10	2.8490	
16	13	357	03/10	3.6415	
17	11	371	04/10	2.9650	
18	12	355	05/10	3.3803	
19	12	371	06/10	3.2345	
20	10	357	07/10	2.8011	
21	11	366	08/10	3.0055	
22	14	372	09/10	3.7634	
23	13	366	10/10	3.5519	
24	11	364	11/10	3.0220	
25	14	361	12/10	3.8781	
26					
27	291	8731			
28	Mean	0.0333			

	A	B	C	D	E
1	# events	Sample	Month	%	Mean
2	13	350	01/09	3.7143	3.33
3	11	367	02/09	2.9973	3.33
4	12	361	03/09	3.3241	3.33
5	14	371	04/09	3.7736	3.33
6	13	366	05/09	3.5519	3.33
7	13	362	06/09	3.5912	3.33
8	12	369	07/09	3.2520	3.33
9	14	374	08/09	3.7433	3.33
10	11	365	09/09	3.0137	3.33
11	10	364	10/09	2.7473	3.33
12	11	359	11/09	3.0641	3.33
13	14	360	12/09	3.8889	3.33
14	12	372	01/10	3.2258	3.33
15	10	351	02/10	2.8490	3.33
16	13	357	03/10	3.6415	3.33
17	11	371	04/10	2.9650	3.33
18	12	355	05/10	3.3803	3.33
19	12	371	06/10	3.2345	3.33
20	10	357	07/10	2.8011	3.33
21	11	366	08/10	3.0055	3.33
22	14	372	09/10	3.7634	3.33
23	13	366	10/10	3.5519	3.33
24	11	364	11/10	3.0220	3.33
25	14	361	12/10	3.8781	3.33
26					
27	291	8731			
28	Mean	0.0333			

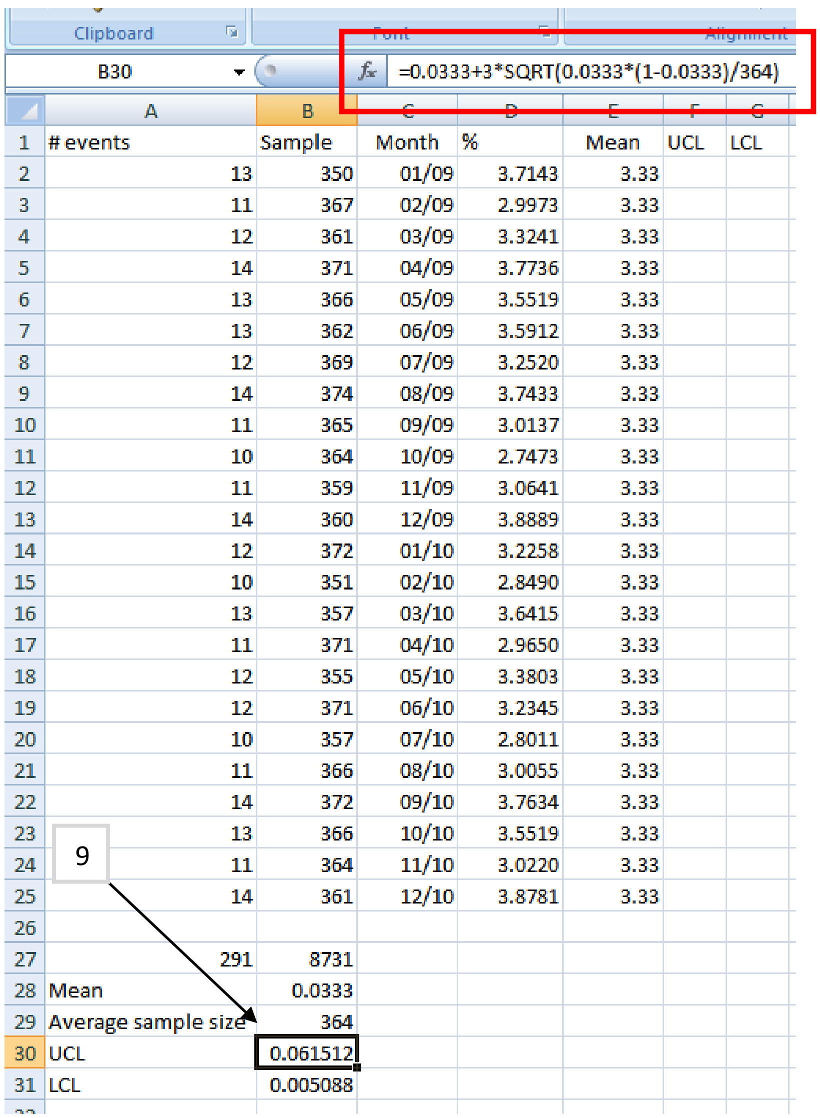
5. Now calculate the average sample size using AVERAGE in Statistical Functions.
    a. Highlight the cell where placement of the value is desired.
    b. Click on More Functions in the Formula bar.
    c. Click Statistical.
    d. Move the cursor down until AVERAGE is highlighted in the drop down box.
    e. Click on AVERAGE.
    f. Type B2:B25 in the Number 1 box in the dialogue box that appears.
    g. Click on OK when done.



6. The resulting average monthly sample size will populate the cell that was highlighted.
7. The cell location for all calculations is a matter of user preference. For this example, column A below the Mean and Average sample size are used to identify the upper and lower control limits and the calculation is performed in the adjacent cell, column B.
8. The formula for calculation of p-chart control limits is: p-bar is the grand mean, the total number of defects or events divided by the total number in the sample, n-bar is the average monthly sample size

$$\bar{p} \pm 3\sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$$

9. To perform the calculation, place the cursor in the desired cell for the value and left click.
10. Using the mouse place the cursor in the formula bar (f_x)
11. Type the formula using the values and the square root function. In Excel language, the formula will read: =0.0333+3*SQRT(0.0333*(1-0.0333)/364)
12. After the formula has been typed in the formula bar, press enter.



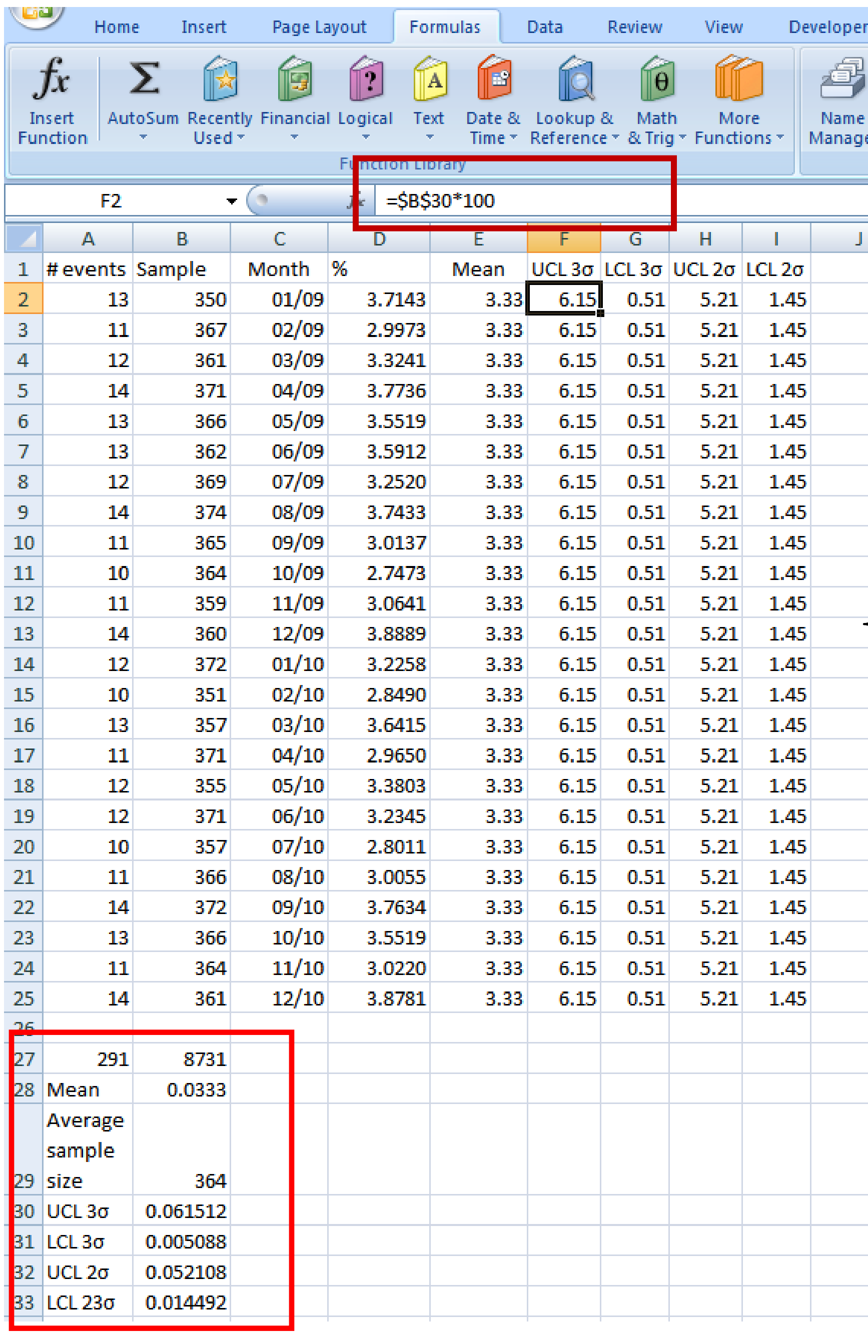
Clipboard | Font | Alignment

B30 | f_x =0.0333+3*SQRT(0.0333*(1-0.0333)/364)

	A	B	C	D	E	F	G
1	# events	Sample	Month	%	Mean	UCL	LCL
2	13	350	01/09	3.7143	3.33		
3	11	367	02/09	2.9973	3.33		
4	12	361	03/09	3.3241	3.33		
5	14	371	04/09	3.7736	3.33		
6	13	366	05/09	3.5519	3.33		
7	13	362	06/09	3.5912	3.33		
8	12	369	07/09	3.2520	3.33		
9	14	374	08/09	3.7433	3.33		
10	11	365	09/09	3.0137	3.33		
11	10	364	10/09	2.7473	3.33		
12	11	359	11/09	3.0641	3.33		
13	14	360	12/09	3.8889	3.33		
14	12	372	01/10	3.2258	3.33		
15	10	351	02/10	2.8490	3.33		
16	13	357	03/10	3.6415	3.33		
17	11	371	04/10	2.9650	3.33		
18	12	355	05/10	3.3803	3.33		
19	12	371	06/10	3.2345	3.33		
20	10	357	07/10	2.8011	3.33		
21	11	366	08/10	3.0055	3.33		
22	14	372	09/10	3.7634	3.33		
23	13	366	10/10	3.5519	3.33		
24	11	364	11/10	3.0220	3.33		
25	14	361	12/10	3.8781	3.33		
26							
27	291	8731					
28	Mean	0.0333					
29	Average sample size	364					
30	UCL	0.061512					
31	LCL	0.005088					

9

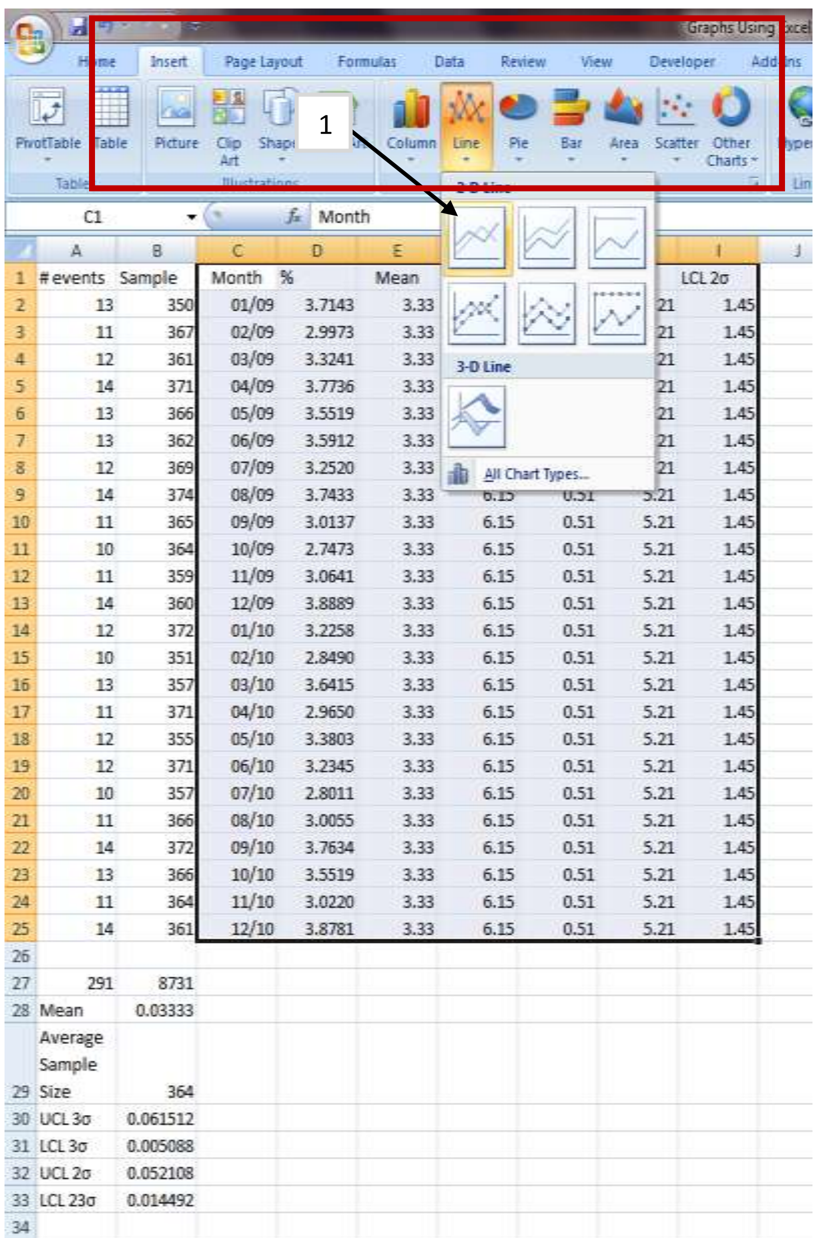
13. Populate the first cell in the column used by the chart to plot the control limit line using the Copy, Paste Special, Past Link sequence and multipy by 100 to obtain the percentage value for the 3σ upper control limit. Fill the cells below it by placing the cursor on the lower right corner of that cell and dragging down.
14. Repeat steps 9-11 for each control limit. The formula can be copied and pasted into the remaining control limit cells and edited as appropriate for each calculation. To perform the lower control limit calculation simply replace the + with a – to subtract 3σ from the mean.
15. To calculate the 2σ limits replace the 3 with 2.



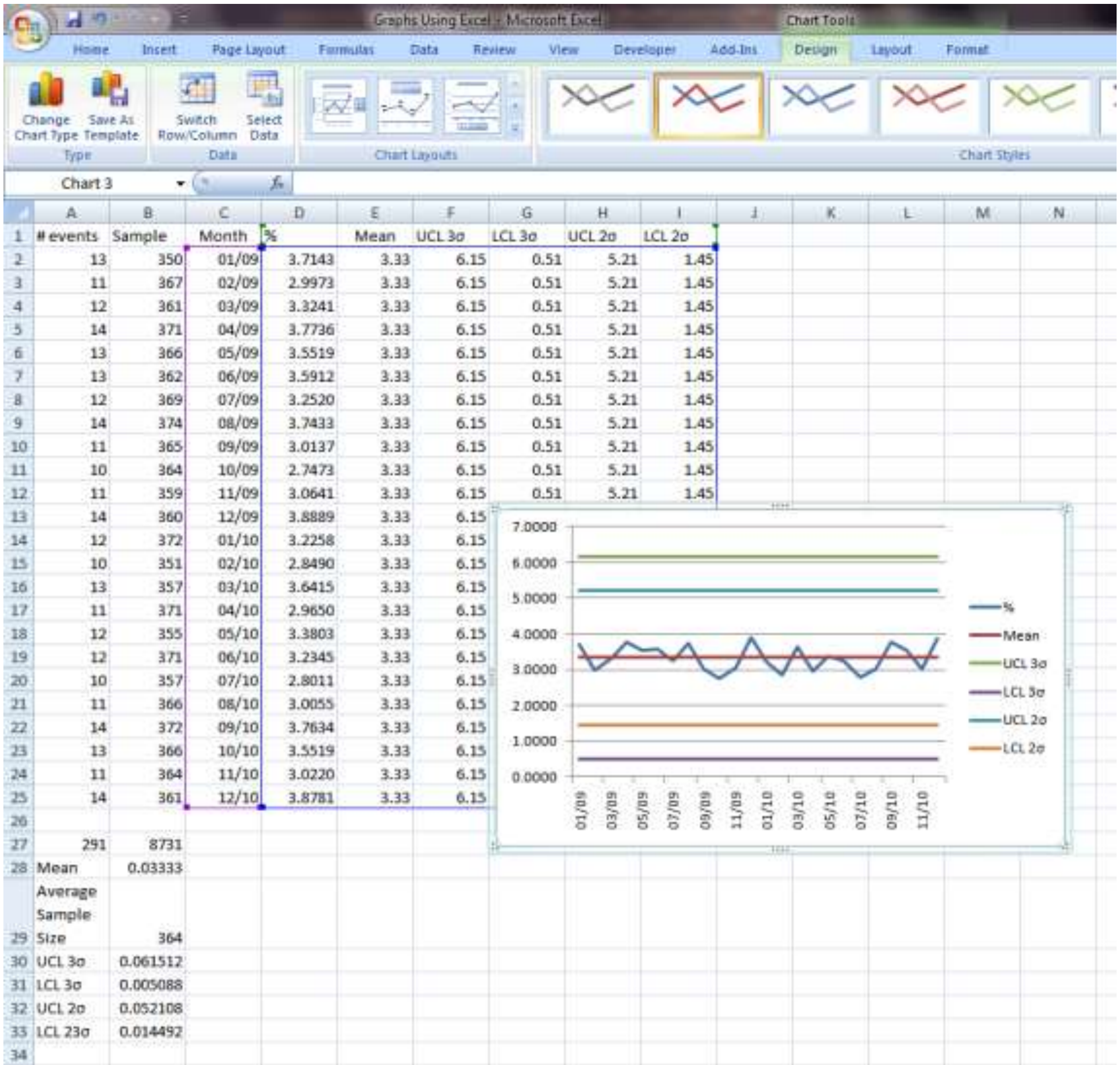
	A	B	C	D	E	F	G	H	I	J
1	# events	Sample	Month	%	Mean	UCL 3σ	LCL 3σ	UCL 2σ	LCL 2σ	
2	13	350	01/09	3.7143	3.33	6.15	0.51	5.21	1.45	
3	11	367	02/09	2.9973	3.33	6.15	0.51	5.21	1.45	
4	12	361	03/09	3.3241	3.33	6.15	0.51	5.21	1.45	
5	14	371	04/09	3.7736	3.33	6.15	0.51	5.21	1.45	
6	13	366	05/09	3.5519	3.33	6.15	0.51	5.21	1.45	
7	13	362	06/09	3.5912	3.33	6.15	0.51	5.21	1.45	
8	12	369	07/09	3.2520	3.33	6.15	0.51	5.21	1.45	
9	14	374	08/09	3.7433	3.33	6.15	0.51	5.21	1.45	
10	11	365	09/09	3.0137	3.33	6.15	0.51	5.21	1.45	
11	10	364	10/09	2.7473	3.33	6.15	0.51	5.21	1.45	
12	11	359	11/09	3.0641	3.33	6.15	0.51	5.21	1.45	
13	14	360	12/09	3.8889	3.33	6.15	0.51	5.21	1.45	
14	12	372	01/10	3.2258	3.33	6.15	0.51	5.21	1.45	
15	10	351	02/10	2.8490	3.33	6.15	0.51	5.21	1.45	
16	13	357	03/10	3.6415	3.33	6.15	0.51	5.21	1.45	
17	11	371	04/10	2.9650	3.33	6.15	0.51	5.21	1.45	
18	12	355	05/10	3.3803	3.33	6.15	0.51	5.21	1.45	
19	12	371	06/10	3.2345	3.33	6.15	0.51	5.21	1.45	
20	10	357	07/10	2.8011	3.33	6.15	0.51	5.21	1.45	
21	11	366	08/10	3.0055	3.33	6.15	0.51	5.21	1.45	
22	14	372	09/10	3.7634	3.33	6.15	0.51	5.21	1.45	
23	13	366	10/10	3.5519	3.33	6.15	0.51	5.21	1.45	
24	11	364	11/10	3.0220	3.33	6.15	0.51	5.21	1.45	
25	14	361	12/10	3.8781	3.33	6.15	0.51	5.21	1.45	
26										
27	291	8731								
28	Mean	0.0333								
29	Average sample size	364								
30	UCL 3σ	0.061512								
31	LCL 3σ	0.005088								
32	UCL 2σ	0.052108								
33	LCL 23σ	0.014492								

Creating and formatting the chart

1. To create the chart:
   a. Click on the Insert tab.
   b. Highlight the data and headers that will be used to create the chart.
   c. Click on the Line chart and highlight the desired type of chart.



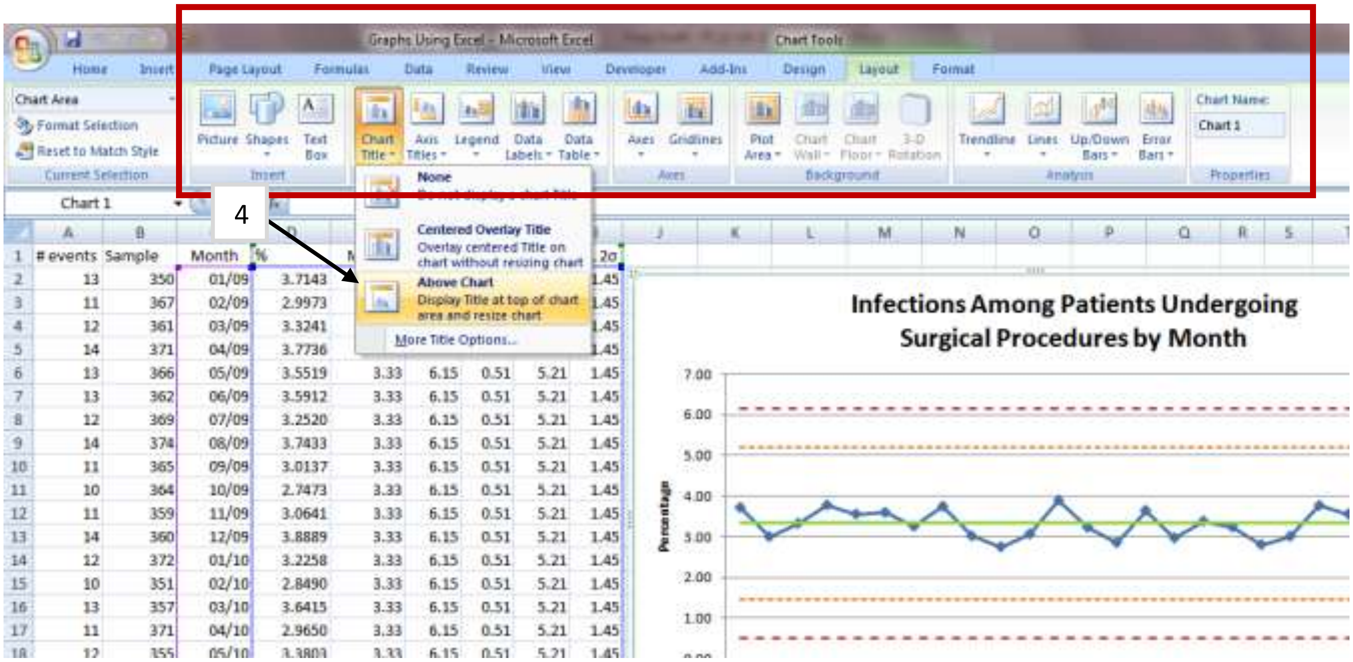
2. The chart will automatically be placed in the existing worksheet.



	# events	Sample	Month	%	Mean	UCL 3σ	LCL 3σ	UCL 2σ	LCL 2σ
2	13	350	01/09	3.7143	3.33	6.15	0.51	5.21	1.45
3	11	367	02/09	2.9973	3.33	6.15	0.51	5.21	1.45
4	12	361	03/09	3.3241	3.33	6.15	0.51	5.21	1.45
5	14	371	04/09	3.7736	3.33	6.15	0.51	5.21	1.45
6	13	366	05/09	3.5519	3.33	6.15	0.51	5.21	1.45
7	13	362	06/09	3.5912	3.33	6.15	0.51	5.21	1.45
8	12	369	07/09	3.2520	3.33	6.15	0.51	5.21	1.45
9	14	374	08/09	3.7433	3.33	6.15	0.51	5.21	1.45
10	11	365	09/09	3.0137	3.33	6.15	0.51	5.21	1.45
11	10	364	10/09	2.7473	3.33	6.15	0.51	5.21	1.45
12	11	359	11/09	3.0641	3.33	6.15	0.51	5.21	1.45
13	14	360	12/09	3.8889	3.33	6.15			
14	12	372	01/10	3.2258	3.33	6.15			
15	10	351	02/10	2.8490	3.33	6.15			
16	13	357	03/10	3.6415	3.33	6.15			
17	11	371	04/10	2.9650	3.33	6.15			
18	12	355	05/10	3.3803	3.33	6.15			
19	12	371	06/10	3.2345	3.33	6.15			
20	10	357	07/10	2.8011	3.33	6.15			
21	11	366	08/10	3.0055	3.33	6.15			
22	14	372	09/10	3.7634	3.33	6.15			
23	13	366	10/10	3.5519	3.33	6.15			
24	11	364	11/10	3.0220	3.33	6.15			
25	14	361	12/10	3.8781	3.33	6.15			
27	291	8731							
28	Mean	0.03333							
29	Average Sample Size	364							
30	UCL 3σ	0.061512							
31	LCL 3σ	0.005088							
32	UCL 2σ	0.052108							
33	LCL 23σ	0.014492							

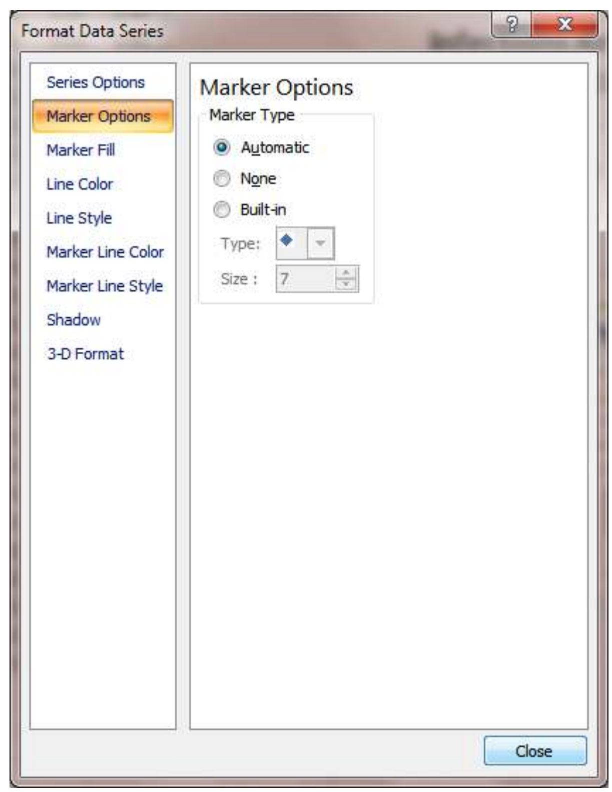
3. Right click to select the chart.

4. In the layout tab of the Chart Tools choose Chart Titles, Axis Titles, etc.. to add and format the components of the chart.

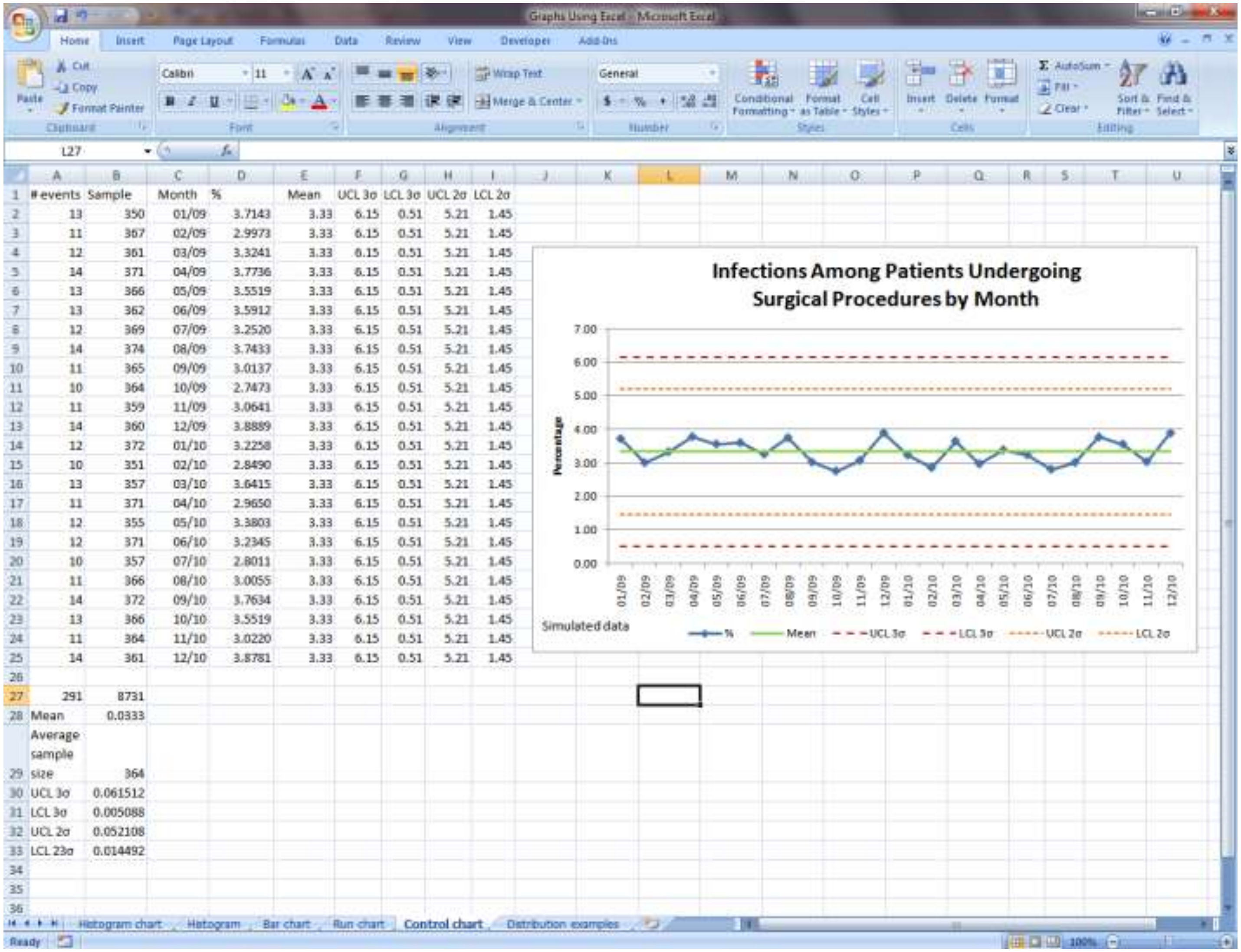


5. Right click on the primary data series line to add markers and further format the line as desired.

5



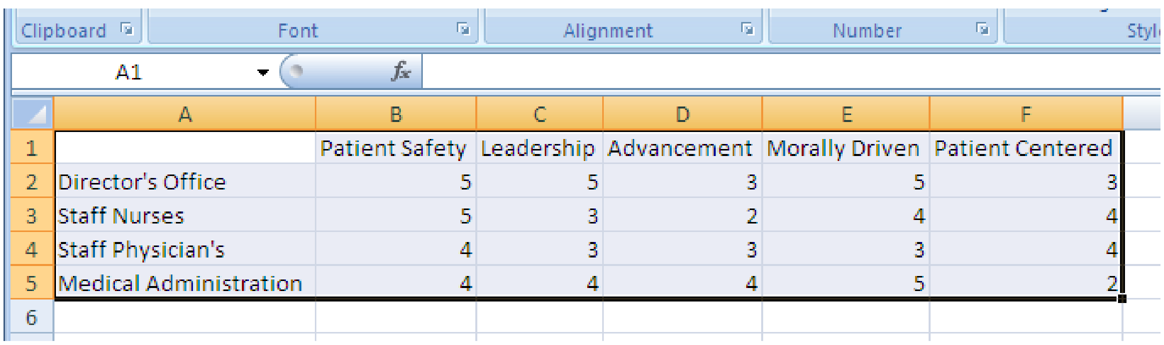
16. The control limits in the chart are formatted for type (dashed, dotted, etc.) and color using the same process.



	A	B	C	D	E	F	G	H	I
1	# events	Sample	Month	%	Mean	UCL 3σ	LCL 3σ	UCL 2σ	LCL 2σ
2	13	350	01/09	3.7143	3.33	6.15	0.51	5.21	1.45
3	11	367	02/09	2.9973	3.33	6.15	0.51	5.21	1.45
4	12	361	03/09	3.3241	3.33	6.15	0.51	5.21	1.45
5	14	371	04/09	3.7736	3.33	6.15	0.51	5.21	1.45
6	13	366	05/09	3.5519	3.33	6.15	0.51	5.21	1.45
7	13	362	06/09	3.5912	3.33	6.15	0.51	5.21	1.45
8	12	369	07/09	3.2520	3.33	6.15	0.51	5.21	1.45
9	14	374	08/09	3.7433	3.33	6.15	0.51	5.21	1.45
10	11	365	09/09	3.0137	3.33	6.15	0.51	5.21	1.45
11	10	364	10/09	2.7473	3.33	6.15	0.51	5.21	1.45
12	11	359	11/09	3.0641	3.33	6.15	0.51	5.21	1.45
13	14	360	12/09	3.8889	3.33	6.15	0.51	5.21	1.45
14	12	372	01/10	3.2258	3.33	6.15	0.51	5.21	1.45
15	10	351	02/10	2.8490	3.33	6.15	0.51	5.21	1.45
16	13	357	03/10	3.6415	3.33	6.15	0.51	5.21	1.45
17	11	371	04/10	2.9650	3.33	6.15	0.51	5.21	1.45
18	12	355	05/10	3.3803	3.33	6.15	0.51	5.21	1.45
19	12	371	06/10	3.2345	3.33	6.15	0.51	5.21	1.45
20	10	357	07/10	2.8011	3.33	6.15	0.51	5.21	1.45
21	11	366	08/10	3.0055	3.33	6.15	0.51	5.21	1.45
22	14	372	09/10	3.7634	3.33	6.15	0.51	5.21	1.45
23	13	366	10/10	3.5519	3.33	6.15	0.51	5.21	1.45
24	11	364	11/10	3.0220	3.33	6.15	0.51	5.21	1.45
25	14	361	12/10	3.8781	3.33	6.15	0.51	5.21	1.45
26									
27	291	8731							
28	Mean	0.0333							
29	Average sample size	364							
30	UCL 3σ	0.061512							
31	LCL 3σ	0.005088							
32	UCL 2σ	0.052108							
33	LCL 23σ	0.014492							

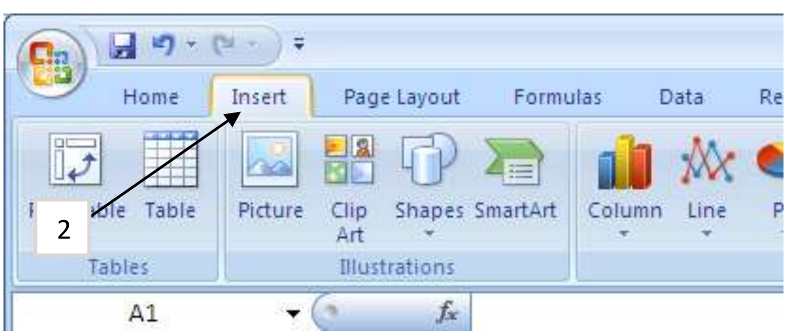
Radar Chart	
Usage	A radar chart is used to visually display areas of strengths and weaknesses and show the gaps in various areas. These are best used to show performance improvement or to evaluate and organizations situation based upon the opinions of team members.
Data Types	Ordinal or Interval data, such as data accumulated from a survey
Interpretation	Points which are furthest from the center are considered strengths. Points closer to the center are weaknesses or areas for improvement. The points on the web usually represent the areas of evaluation. The lines represent the services/areas surveyed or analyzed. In the example for this radar chart, team members were asked to rate the performance (on a scale of 1 = not well at all, to 5 excellent) of their unit on several key areas, including Director's office,

1. Format data so categories are across the top and rankings will fall into the columns. Select all

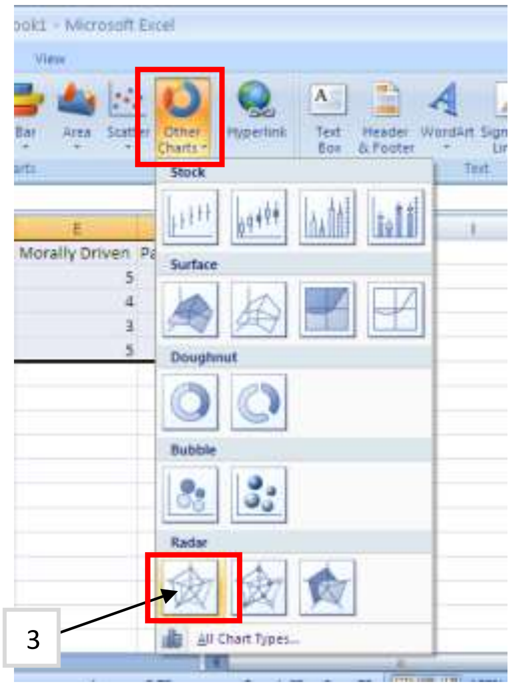


	A	B	C	D	E	F
1		Patient Safety	Leadership	Advancement	Morally Driven	Patient Centered
2	Director's Office	5	5	3	5	3
3	Staff Nurses	5	3	2	4	4
4	Staff Physician's	4	3	3	3	4
5	Medical Administration	4	4	4	5	2
6						

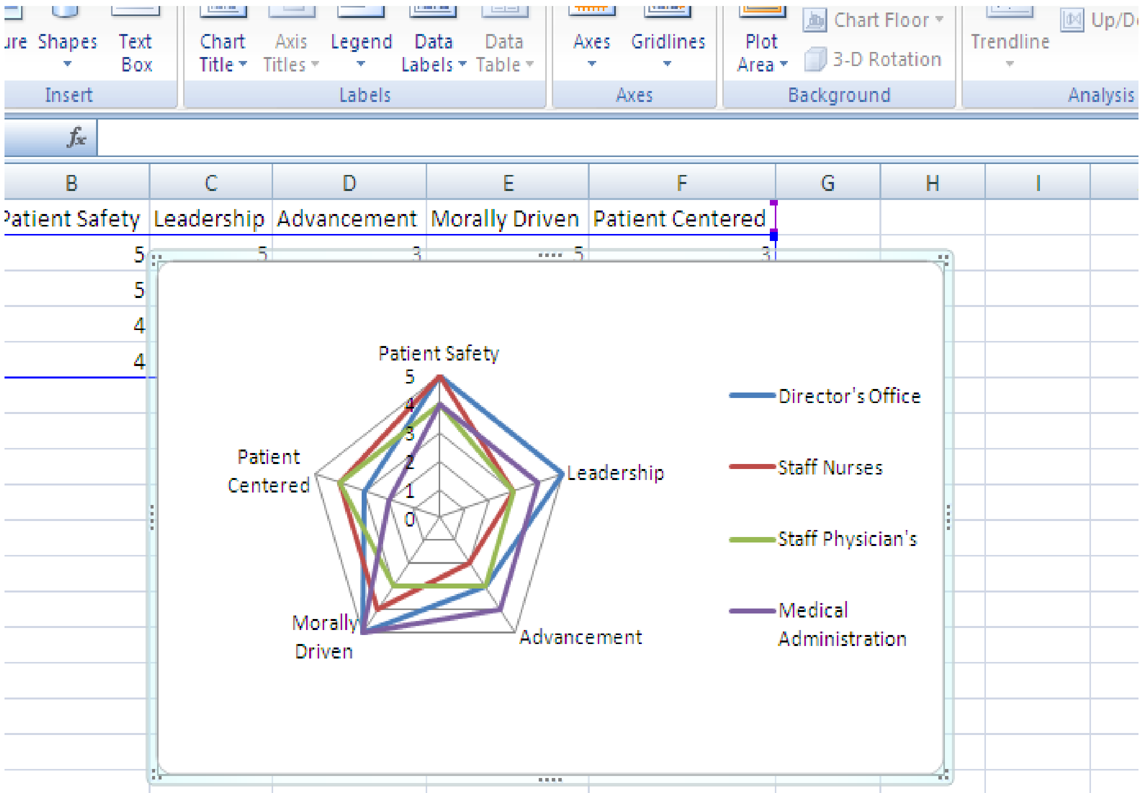
2. Click on the Insert Tab in the ribbon.



3. Click Other Charts, and select the first Radar Chart under the Radar heading. The second Radar Chart has defined points. The third choice under the heading fills in the area. For this chart, the first choice was selected because it is most easily understood.



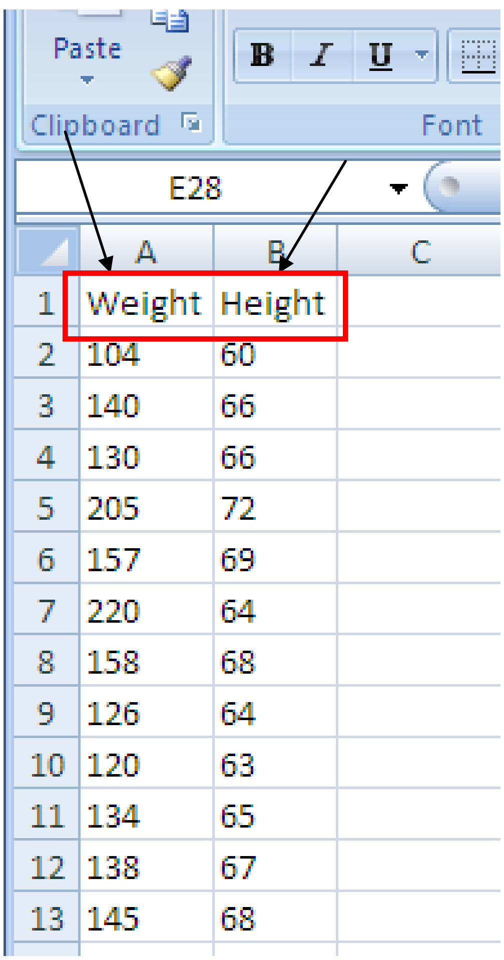
4. Once selected, your radar chart will appear. In this example, the different colored lines represent each area surveyed. The five points of the "web" represent the areas of the evaluation.



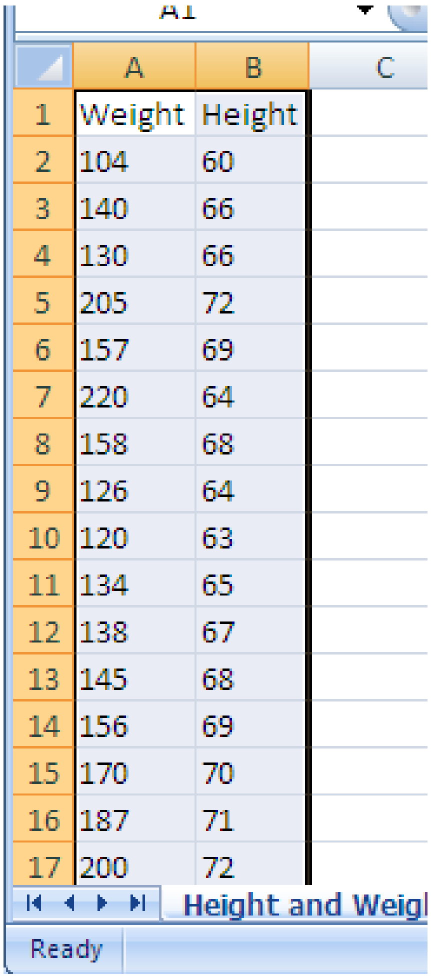
5. Refer to the Formatting section to learn how to customize your chart to your needs.

Scatter Diagram	
Usage	To display paired sets of data and evaluate for a relationship as a series of single data points.
Data Types	Interval or ratio data such as height and weight
Interpretation	If the data progress from the lower left to the upper right, there is a positive relationship. If the data progress from the upper right to the lower left, there is a negative relationship. If a "line" cannot be defined, there is no relationship of the data. When a curve is formed, there is both a positive and negative relationship. Remember, relationship does not mean causation
Creation in Excel	Data must be paired.

1. Arrange data into columns, column A representing the dependent variable, column B representing the independent variable. An independent variable is the variable of interest, whereas the dependent variable may influence the outcome (dependent variable). In this case, weight is influenced by height, so it is the independent variable.

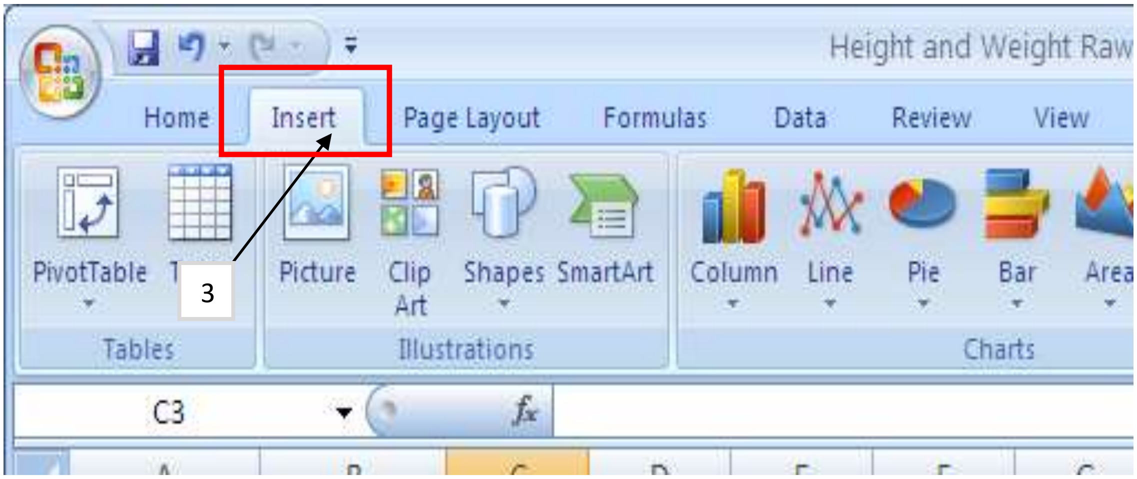


2. Select all the data.

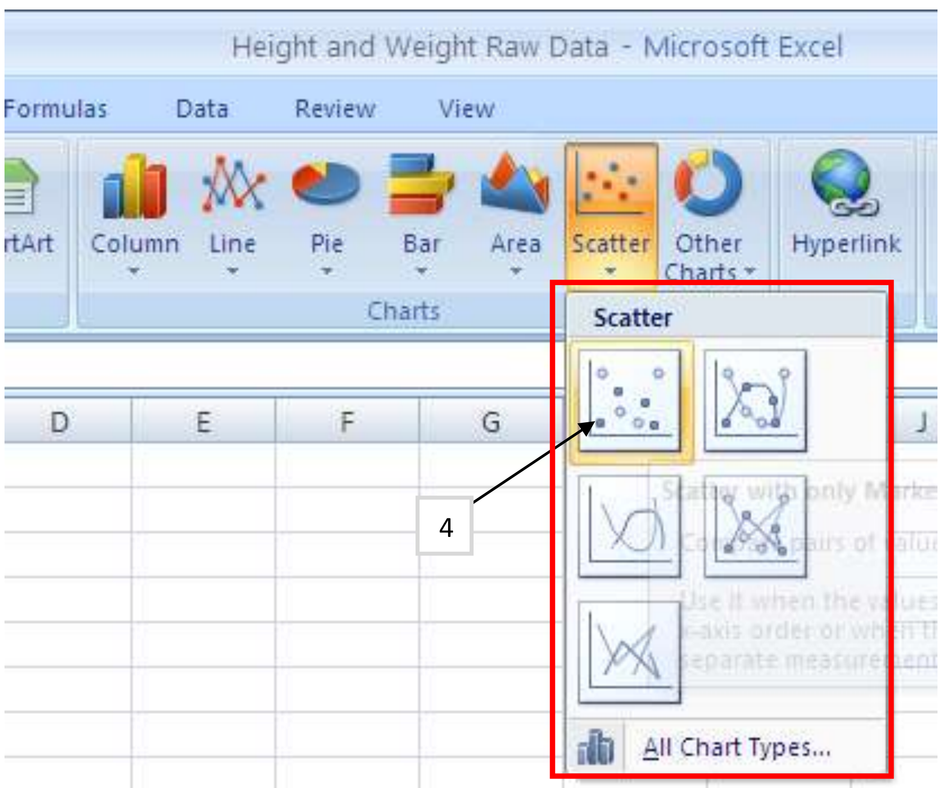


	A	B	C
1	Weight	Height	
2	104	60	
3	140	66	
4	130	66	
5	205	72	
6	157	69	
7	220	64	
8	158	68	
9	126	64	
10	120	63	
11	134	65	
12	138	67	
13	145	68	
14	156	69	
15	170	70	
16	187	71	
17	200	72	

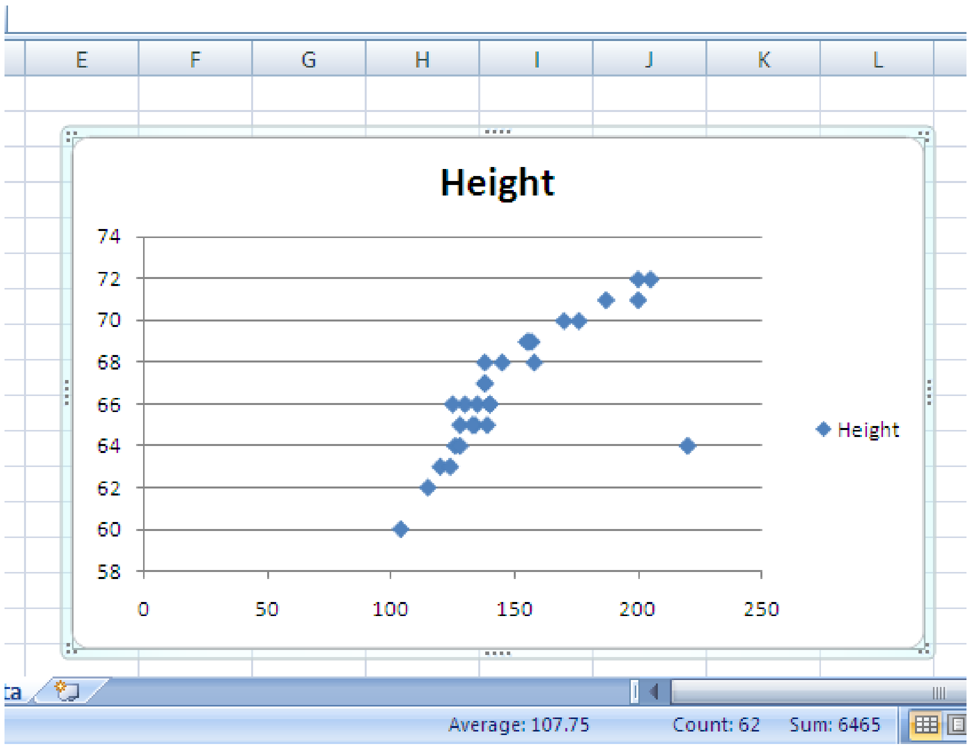
3. Click on the insert tab in the ribbon.



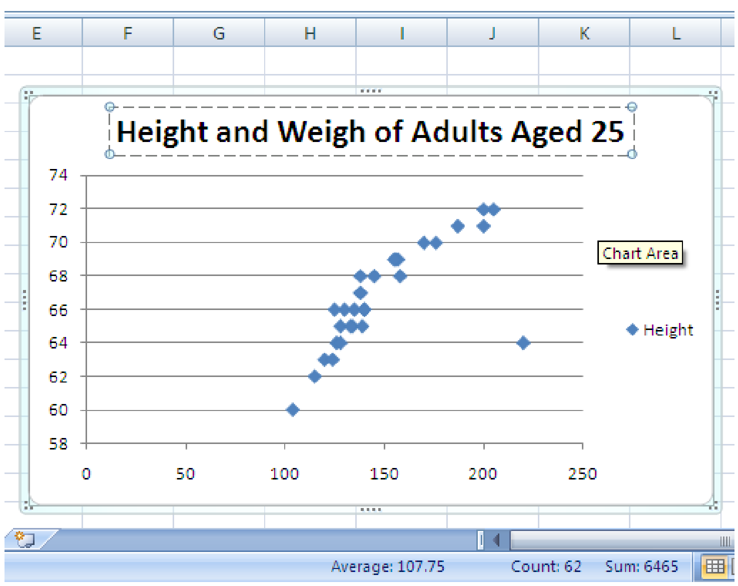
4. Select the Scatter plot chart. In the drop down, select the first chart type. The other scatter plot options contain lines; this changes the type of chart to a line graph.



5. Once selected, the scatter plot will be inserted into the worksheet.



6. Double click on the title to change it to reflect what the graph is showing.



7. Refer to the formatting section to customize your chart to your specific needs.

References

Benneyan, J. C. (1998). Use and interpretation of statistical quality control charts. *International Journal for Quality in Health Care , 10* (1), 69-73.

Brassard, M., & Ritter, D. (2007). *The Public Health Memory Jogger II.* (F. Oddo, Ed.) GOAL/QPC.

Hansen, J. P. (2003). CAN'T MISS: Conquer Any Number Task by Making Important Statistics Simple. Part 2. Probability, Populations, Samples, and NOrmal Distributions. *Journal for Healthcare Quality , 25* (4), 25-33.

Hansen, J. P. (2005). CAN'T MISS: Conquer Any Number Task by Making Important Statistics Simple. Part 8. Statistical Process Control: n, np, c, u Control Charts. *Journal for Healthcare Quality , 27* (4), 42-52.

Klass, G. (2006, July 21). *Presenting Data*. Retrieved December 3, 2010, from Illinois State University: http://lilt.ilstu.edu/gmklass/pos138/datadisplay/sections/goodcharts.htm

Sellick, J. A. (1993). The Use of Statistical Process Control Charts in Hospital Epidemiology. *Infection Control and Hospital Epidemiology , 14* (11), 649-656.

Tague, N. R. (2004). *Quality Tools*. Retrieved November 28, 2010, from ASQ: http://asq.org/learn-about-quality/data-collection-analysis-tools/overview/histogram.html