

LEANOhio

Transforming the Public Sector

Green Belt

Graphical Displays of Data

SIMPLER. FASTER. BETTER. LESS COSTLY.

Name	Years of Service	Height	Job Category	Reason Late for Work

IT
 HR
 General
 Services/Facilities
 Accounting/Finance
 Legal
 Communications
 Other

Six Sigma Detectives

Six Sigma is about:

- Embracing the detective in all of us and using that data to lead team decisions
- Proposing theories and using data to prove or disprove those theories
- Asking the right questions!



Mantra

Pictures are worth a thousand words

- Viewing the data graphically helps to confirm how much *practical* significance exists in what you've found.

- **Graph it.....**

–Then STAT IT!!!!



Using Graphical Displays

- In Six Sigma and Lean - we use graphical displays to let the data tell the story and lead the way
- There are many ways to display data

Graphical Methods

- Stem and Leaf
- Dot Plots
- Bar Charts
- Pie Charts
- Scatter Diagrams
- Histogram
- Box Plot
- Normal Probability Plots

Stem and Leaf Plot

SIMPLER. FASTER. BETTER. LESS COSTLY.

LEANOhio

Steam and Leaf Plot Data Collection

- Years of Service with the State of Ohio

Stem-and-Leaf

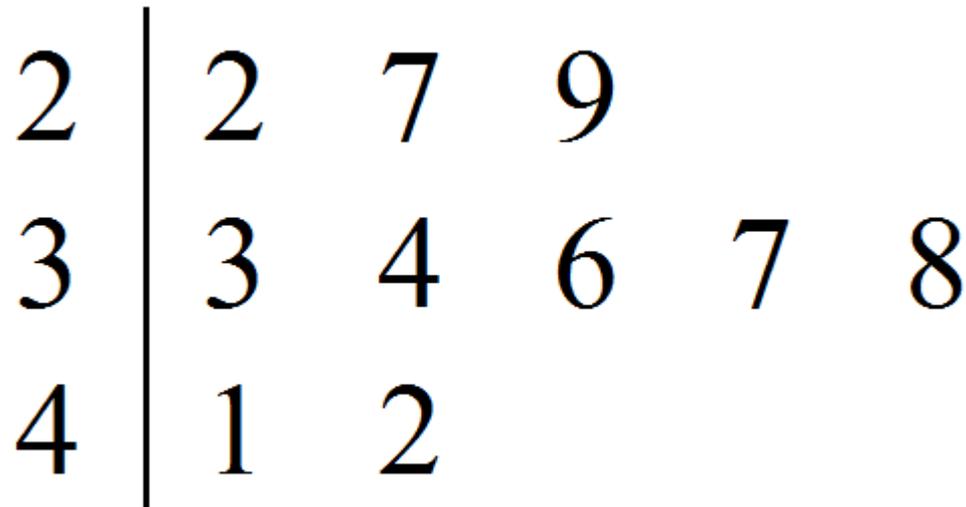
- Stem-and-Leaf Plots: is a type of histogram that separates each number into a stem (all numbers but the last digit) and a leaf (the last digit)
- Visually shows relationships between variables

Stem-and-Leaf: Benefits

- Plots are easy and quick to construct
- Shows shape and distribution
- Visually compact
- Convenient to use
- Displays both variable and categorical data sets
- Allows for data to be read directly from the diagram

Stem-and-Leaf

- For example if you had these numbers:
22,34,42,27,29,38,41,37,33,36.
- The stem-and-leaf plot would look like this:



Stem and Leaf Example

- For example if our data consisted of these numbers:

- 7
- 12
- 14
- 21
- 24
- 27
- 31

0			
1			
2			
3			

Stem and Leaf Example

- For example if our data consisted of these numbers:

- 7
- 12
- 14
- 21
- 24
- 27
- 31

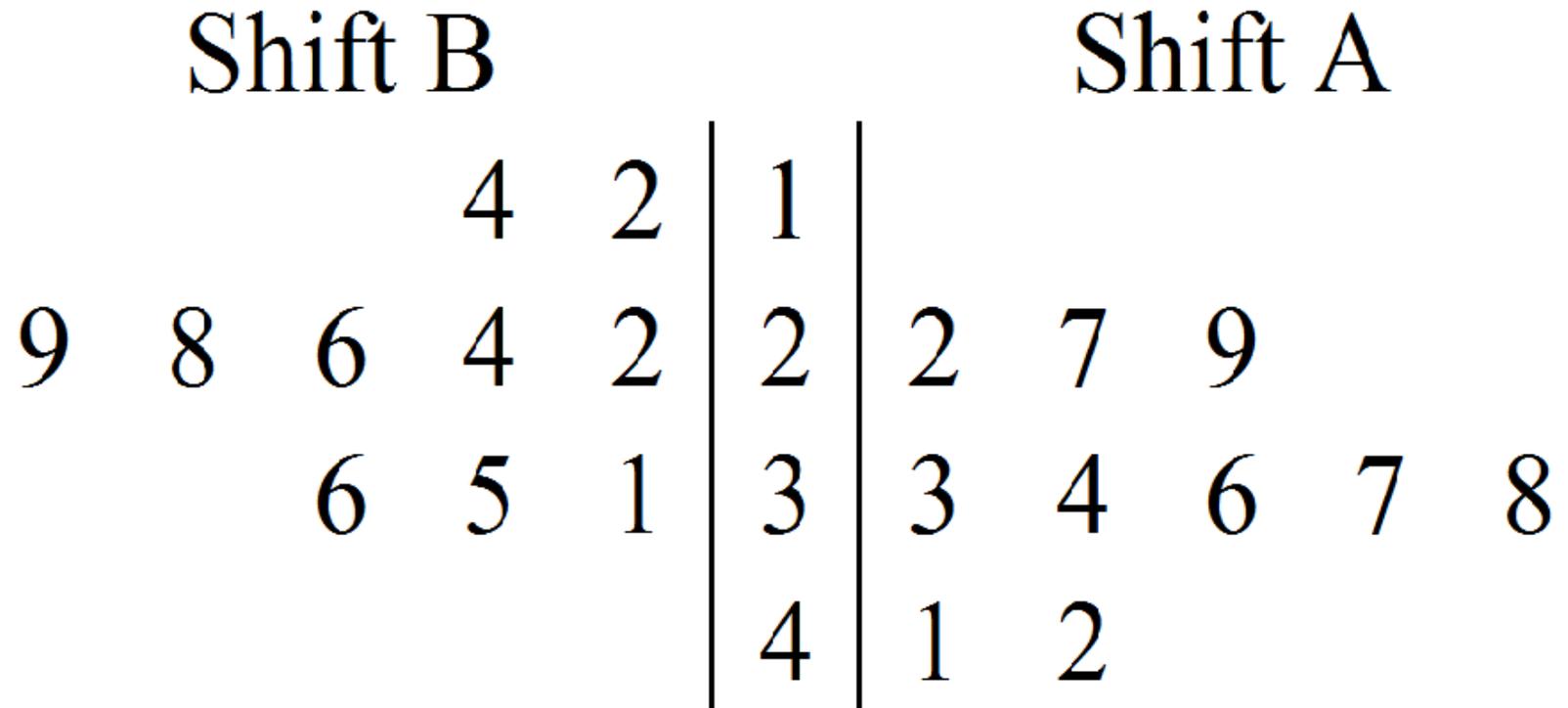
0	7		
1	2	4	
2	1	4	7
3	1		

NOTE: it is easier to put the numbers in order
(excel can do this for you)

You can even do Double Stem and Leaf Plots

- Lets say we had two Shifts and we wanted to compare the time it took to complete their applications.
- Shift A: 22,34,42,27,29,38,41,37,33,36
- Shift B: 12, 14, 22, 24, 26, 28, 29, 31, 35, 36

Double Stem and Leaf Plot



Stem and Leaf

- In groups of four: please use flip chart paper and create a Stem and Leaf Plot for the years of service data

Dot Plots

SIMPLER. FASTER. BETTER. LESS COSTLY.

LEAN**Ohio**

Dot Plots

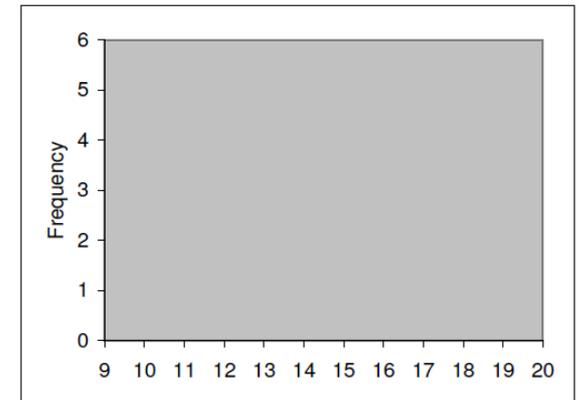
- Dot plot: A horizontal graphical display, similar to a histogram, showing distributional characteristics of a single variable
- Plots data as dots (one dot represents one observation) above a horizontal axis that covers the entire range of observations
- Dot plots are especially useful for comparing several groups of data and identifying clusters, gaps, and outliers

Dot Plots: When and Why

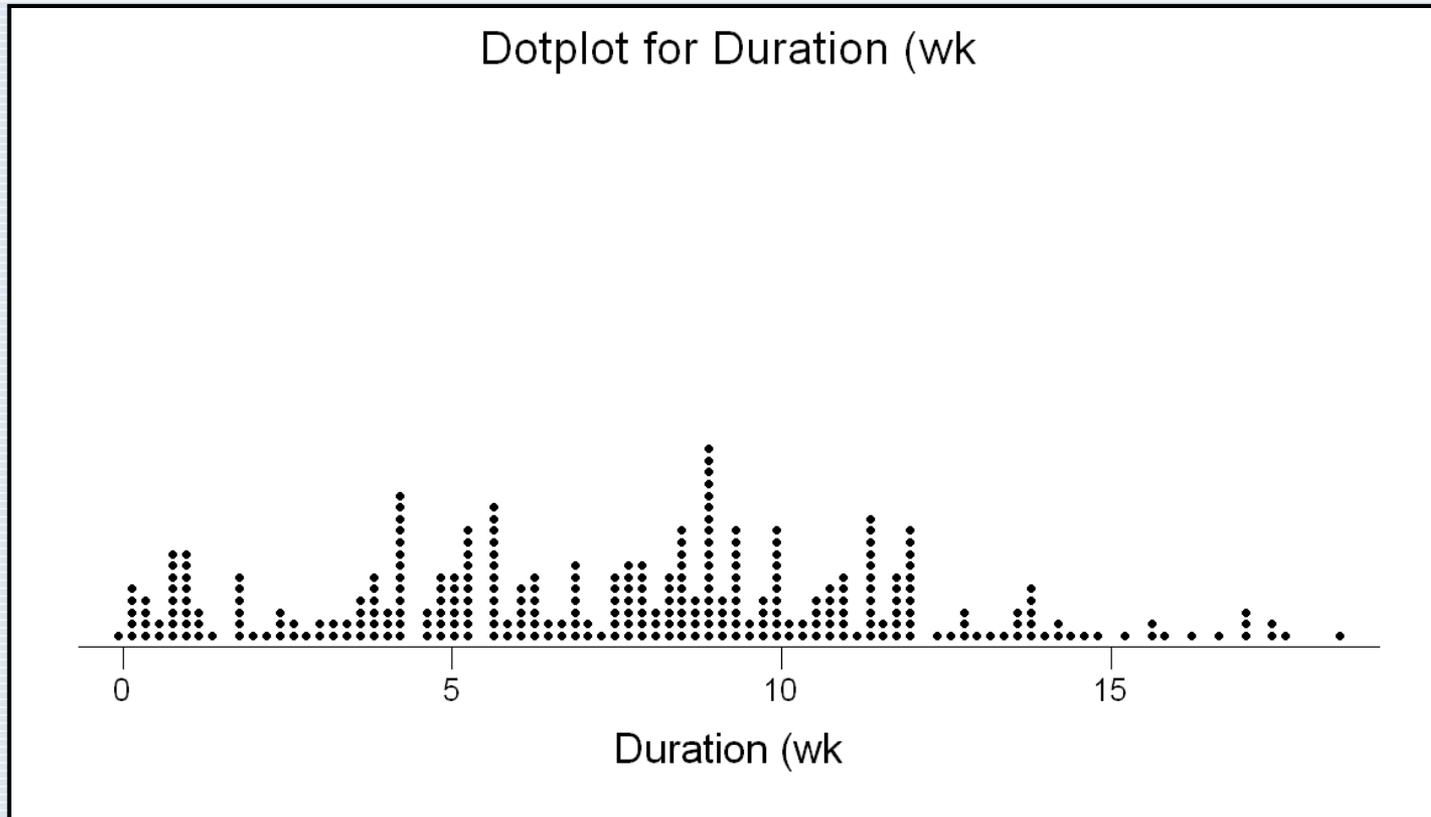
- Great way to see data points and quickly compare two or more sets of data, without performing a single calculation
- Dot plots are easily constructed
- Dot plots look very similar to histograms, but are easier to construct and can be more valuable for generating clues into potential causes

Dot Plots Procedure

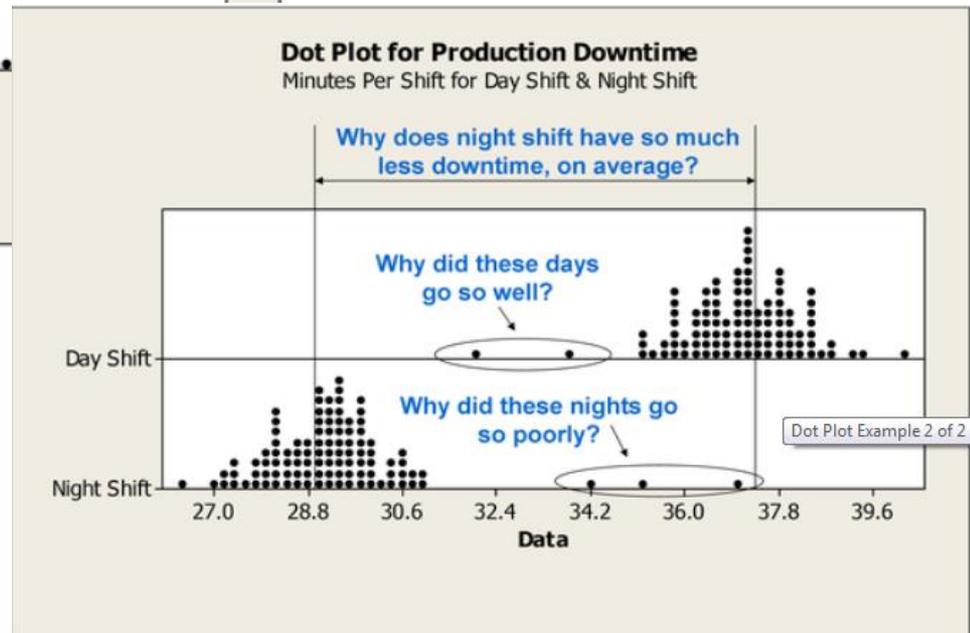
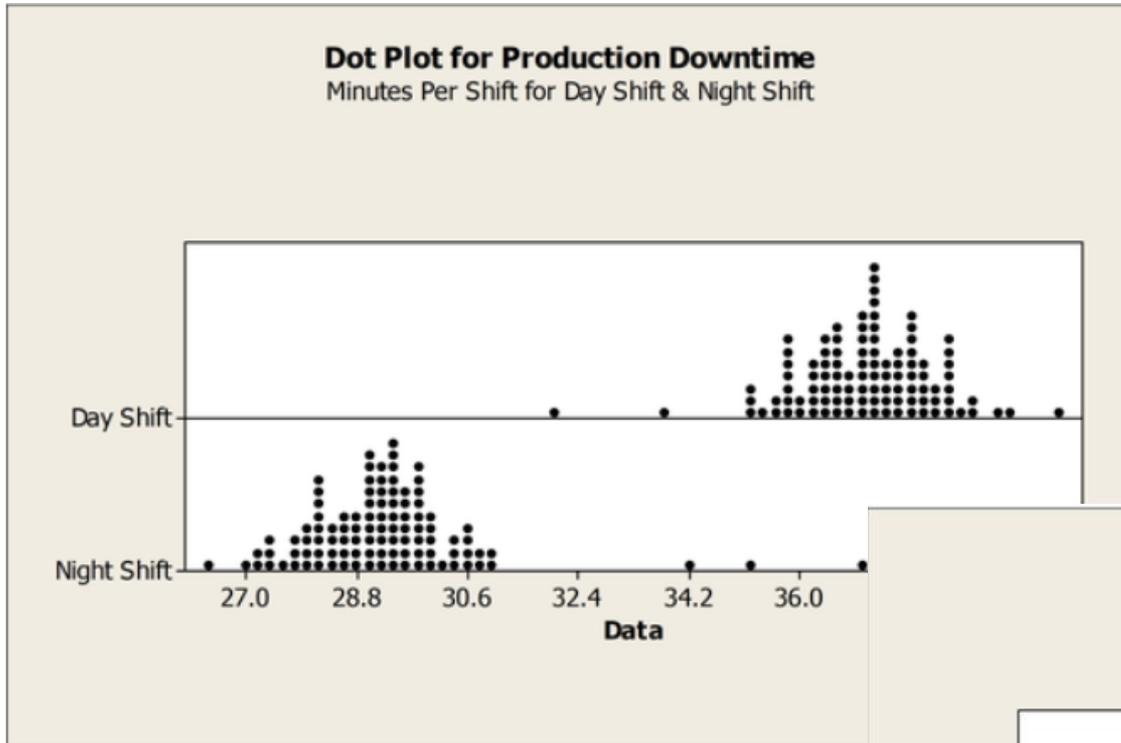
- Find your smallest and largest data point
- Make a horizontal axis that starts at one less than your smallest data point, ends at one more than your largest data point and is equally spaced
- Attach to that axis a vertical axis that has equally spaced steps that start at 0 and end with at least the number of times the mode appears
- Place one dot for each data point



Dot Plot Example



Dot Plot Examples



Where can I use Dot Plots

- Compare two forms (error rates)
- Compare regions (ODOT – Safety Violations)
- Compare Departments (Work output)
- Compare how many phone calls are handled (am vs. pm)

DOT Plot – Data Collection

- Green Belt Class Height
 - Measure your height and add it to the data sheet
 - Create a Dot Plot of the Class Height

Bar Charts, Pie Charts, and Line Graphs

SIMPLER. FASTER. BETTER. LESS COSTLY.

LEANOhio

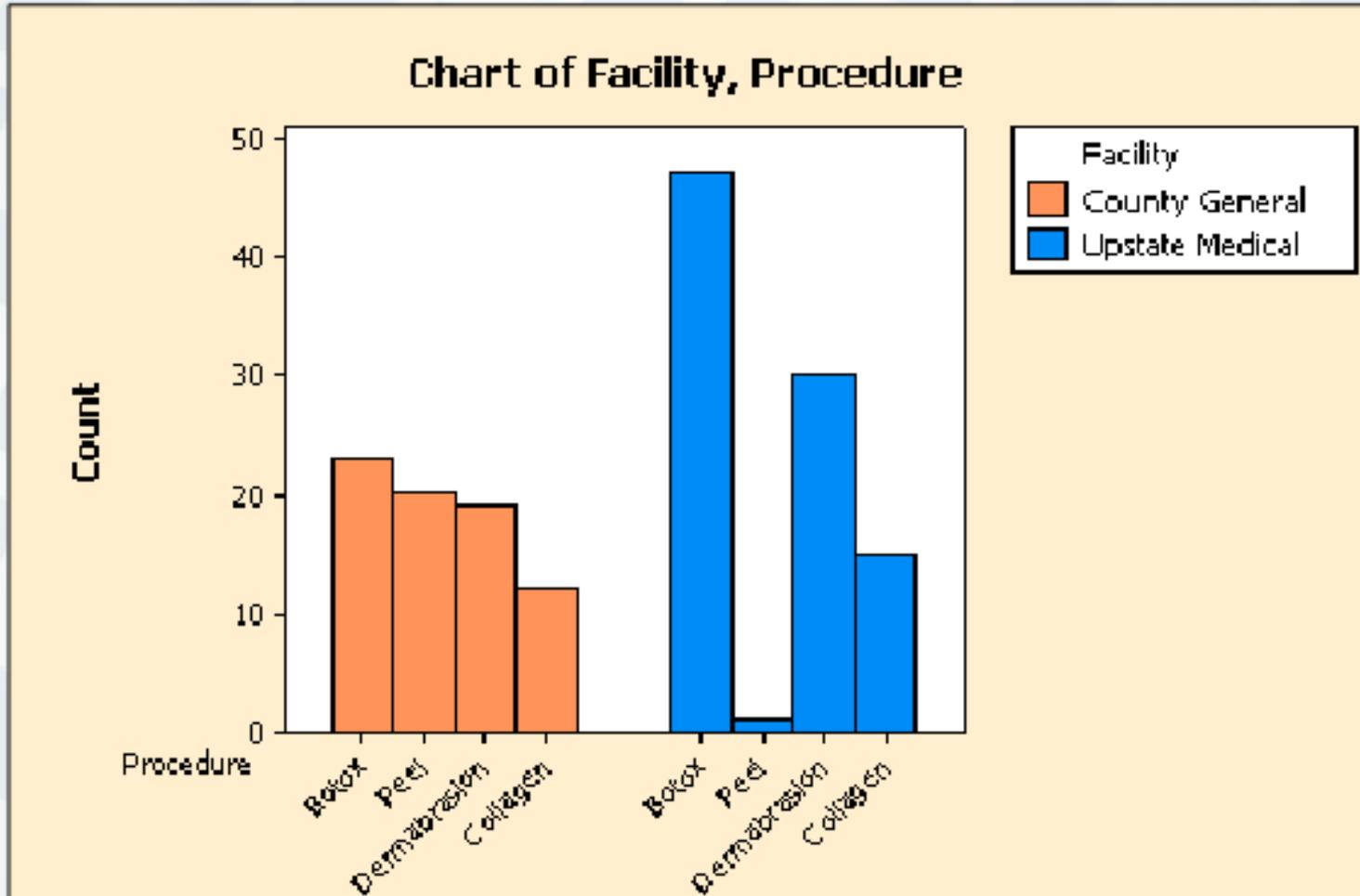
Bar Charts

- A bar chart is used to graphically summarize and display the differences between groups of data.

How to Construct a Bar Chart

- A bar chart can be constructed by segmenting the range of the data into groups or bins
- The vertical axis of the bar chart is labeled 'Frequency' (the number of counts for each bin)
- The horizontal axis of the bar chart is labeled with the group names of your response variables
- You then determine the number of data points that reside within each bin and construct the bar chart
- The groups are defined by the user

Bar Chart



Bar Chart Data Collection

- Current Job Function
 - IT
 - HR
 - General Services/Facilities
 - Accounting/Finance
 - Legal
 - Communications
 - Other

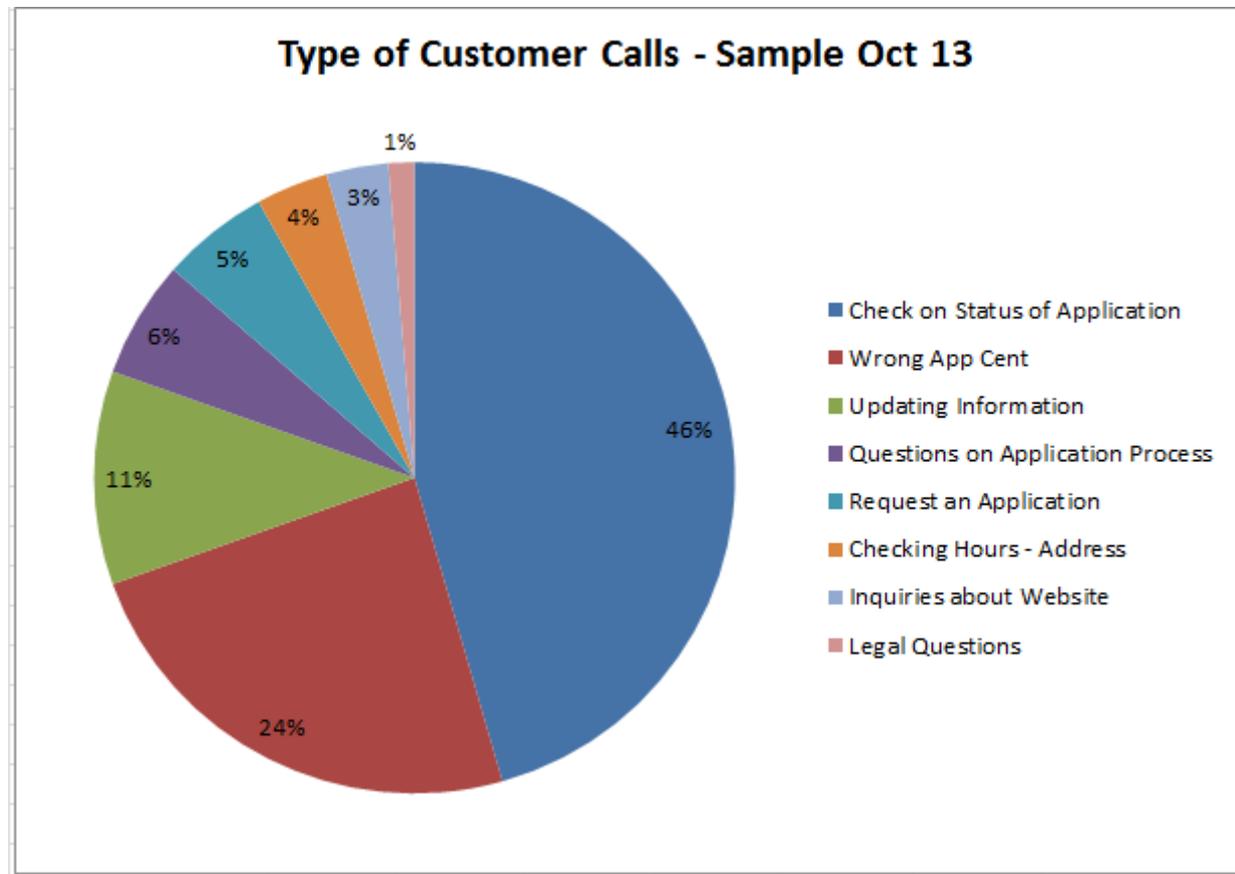
Pie Chart

SIMPLER. FASTER. BETTER. LESS COSTLY.

LEANOhio

Pie Chart

- This information can also be shown in a Pie Chart: Create a Pie Chart of Job Functions Data



Histograms

SIMPLER. FASTER. BETTER. LESS COSTLY.

LEANOhio

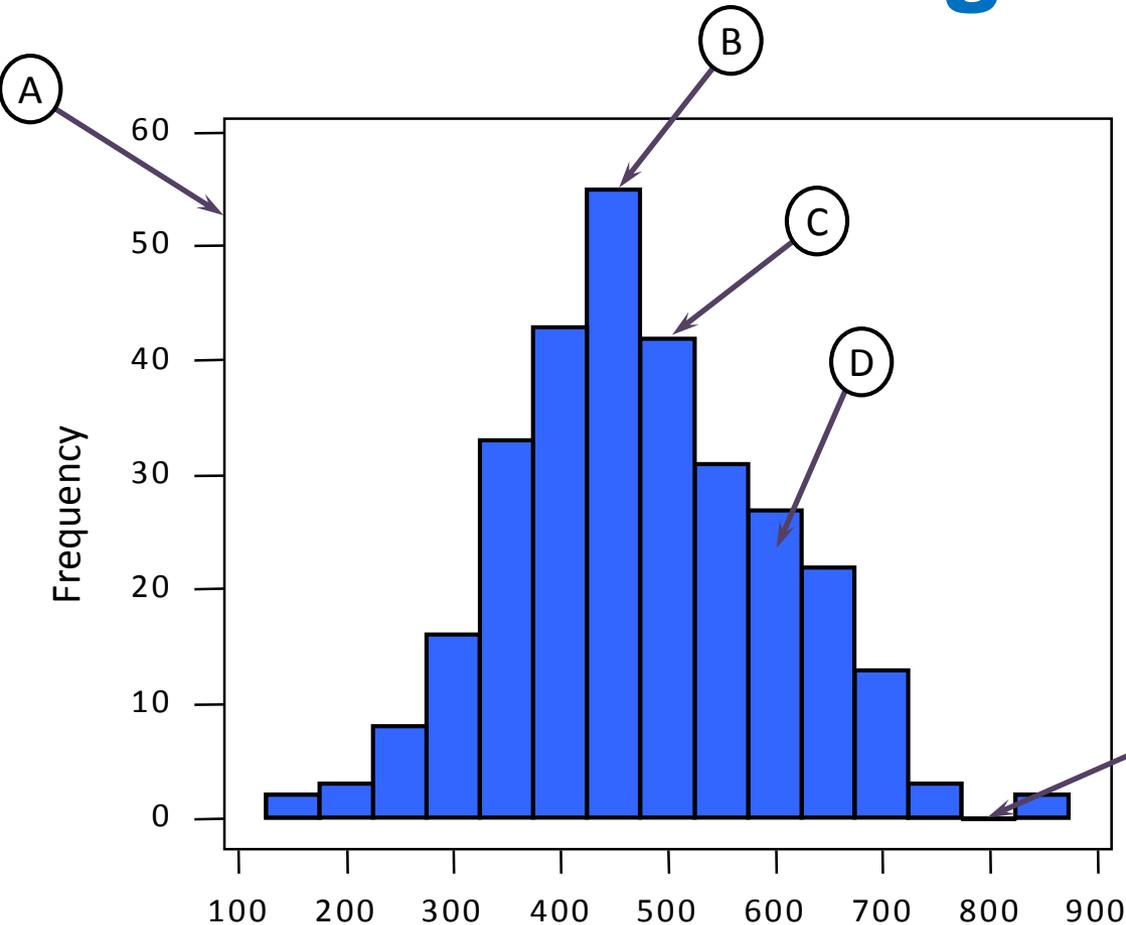
Histograms

- Histograms show frequency distribution
- A horizontal axis is the measurement scale used
- Adjacent boxes are constructed so that the length (or height) of the box represents the frequency for each class
- Histograms are commonly used for summarizing, analyzing, and displaying data

Histograms: Why

- An efficient graphical method for describing the distribution of data
 - Data divided into groups called ‘classes’ or “bins”
 - Number of data points within each class are counted
 - Bars are drawn for each class
- Used to assess
 - Measures of central tendency
 - Variation in the data
 - Normality
- Requires a minimum of 30 data points

Histograms



A. Vertical Axis - Frequency or the percentage of data points in each Class

B. Modal Class - Class with the highest frequency

C. Frequency - Number of data points found in each Class

D. Class - Each bar is one

Class, or interval

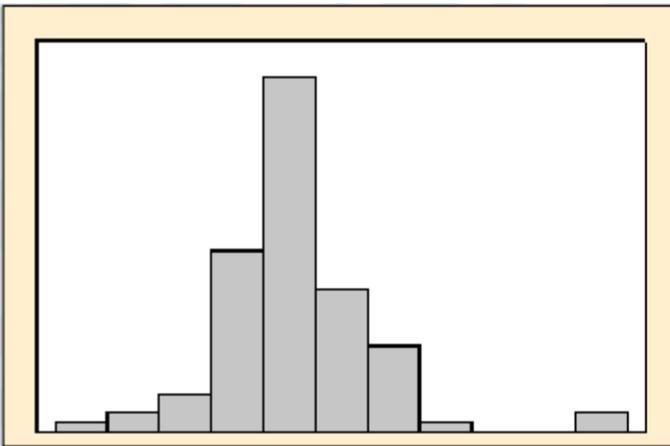
E. Horizontal axis - Scale of measure for the characteristic

Histograms: Benefits

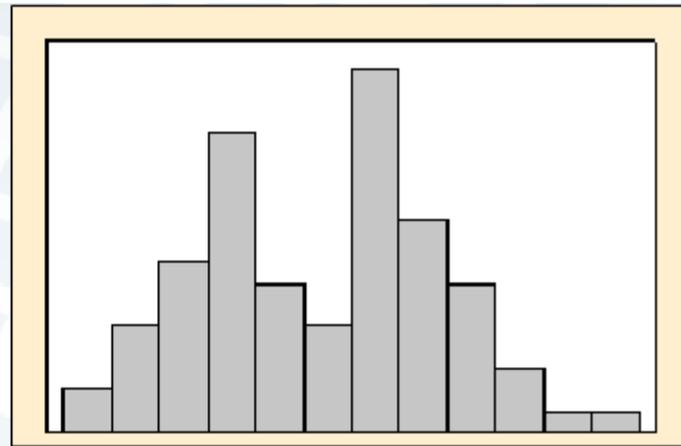
- Provide clearer and more complete picture of data patterns
- To communicate the distribution pattern
- To make decisions about where to focus improvement efforts
- Repeated use of histograms allows a person to see if a change has occurred over time

Reading Histograms

Histogram Shape can tell us about the data – is it Unimodal or Bimodal?



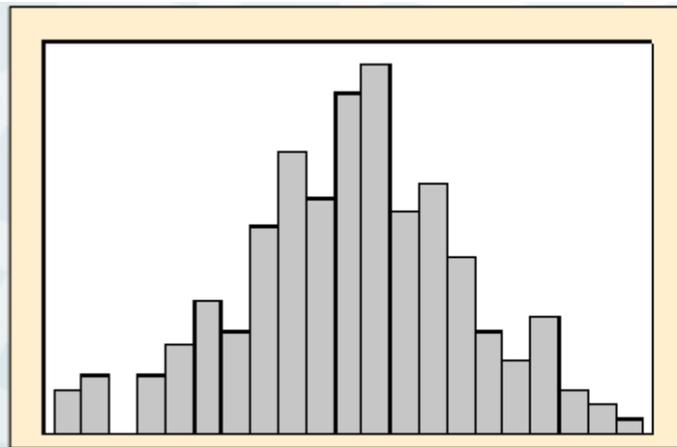
Unimodal



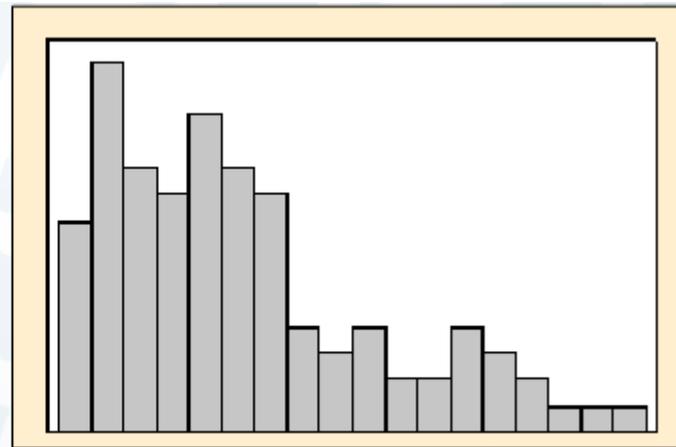
Bimodal

Reading Histograms

Is the data Symmetrical or Skewed?



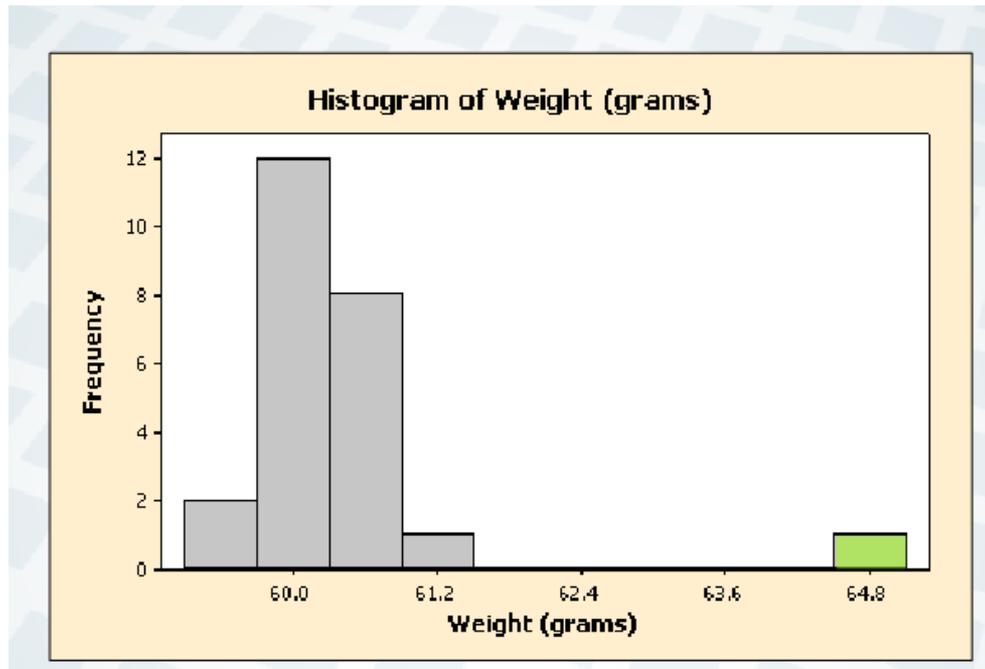
Symmetrical



Right-Skewed

Reading Histograms

Histograms can show Outliers



Outliers are values that are extremely high or low relative to other values in the data.

Histogram - Bins

- Juran's Quality Control Handbook provides these guidelines for the number of bars and state they are not "rigid" and should be adjusted when necessary

Number of Data Points	Number of Bins
30-50	6
51-100	7
101-200	8
201-500	9
501-1000	10
1000 +	11-20

Histogram

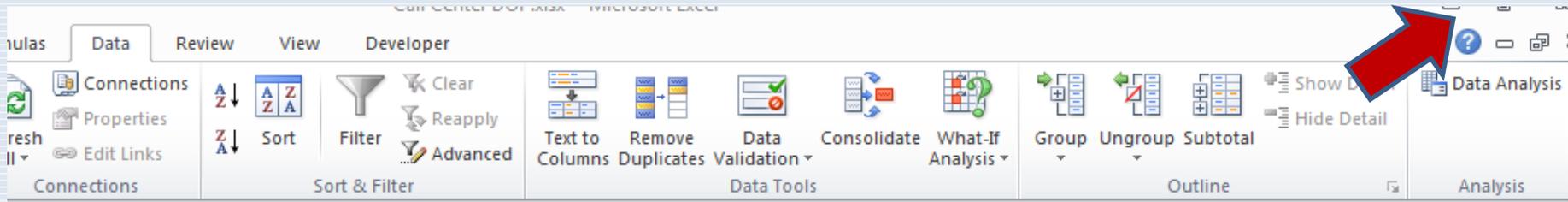
Create a Histogram with the Height Data:

Bins:

- 50-54.9
- 55-59.9
- 60-64.9
- 65-69.9
- 70-74.9
- 75-79.9
- 80-84.9

Data Analysis Pack

- Excel has Data Analysis Add In under the Data Tab



- If not showing – can be added on:
 - Click File Tab
 - Click Options
 - Add-Ins
 - Analysis ToolPak

Data Analysis Histogram

- Data Analysis has a Histogram tool. If the bins are left empty – Excel will calculate the bins!
- OR you can calculate the bins by hand:
 - Take Highest Number in Data Set - Minus Lowest number of Data Set.
 - Take that number divided by (Number of bins recommended for your data set– 1). This will give you the width of your bins.

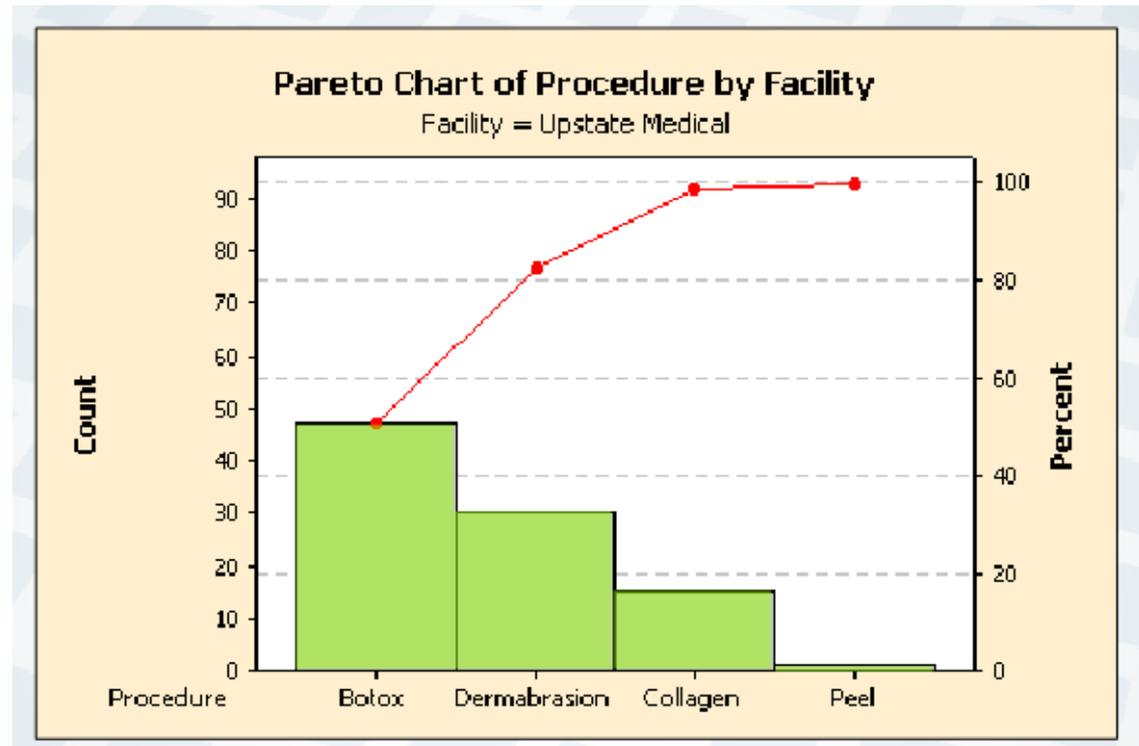
Pareto Chart

SIMPLER. FASTER. BETTER. LESS COSTLY.

LEANOhio

Pareto Chart

- Pareto Chart is a fancy bar graph.



A **Pareto chart** is a bar graph with the bars sorted in order of decreasing frequency. It is used to identify the largest opportunity for improvement.

Pareto Chart

- The lengths of the bars represent frequency or cost (time or money), and are arranged with longest bars on the left and the shortest to the right
- In this way the chart visually depicts which situations are more significant

Pareto Chart: When to Use

- When analyzing data about the frequency of problems or causes in a process
- When there are many problems or causes and you want to focus on the most significant
- When analyzing broad causes by looking at their specific components
- When communicating with others about your data

Pareto Chart: Procedure

- Decide what categories you will use to group items
- Decide what measurement is appropriate: frequency, quantity, cost or time
- Decide what period of time the Pareto chart will cover
- Collect the data
- Subtotal the measurements for each category
- Determine the appropriate scale for the measurements you have collected
- Mark the scale on the left side of the chart

Pareto Chart: Procedure

- Construct and label bars for each category
- Place the tallest at the far left, then the next tallest to its right and so on
- If there are many categories with small measurements, they can be grouped as “other”

Pareto Chart: Procedure

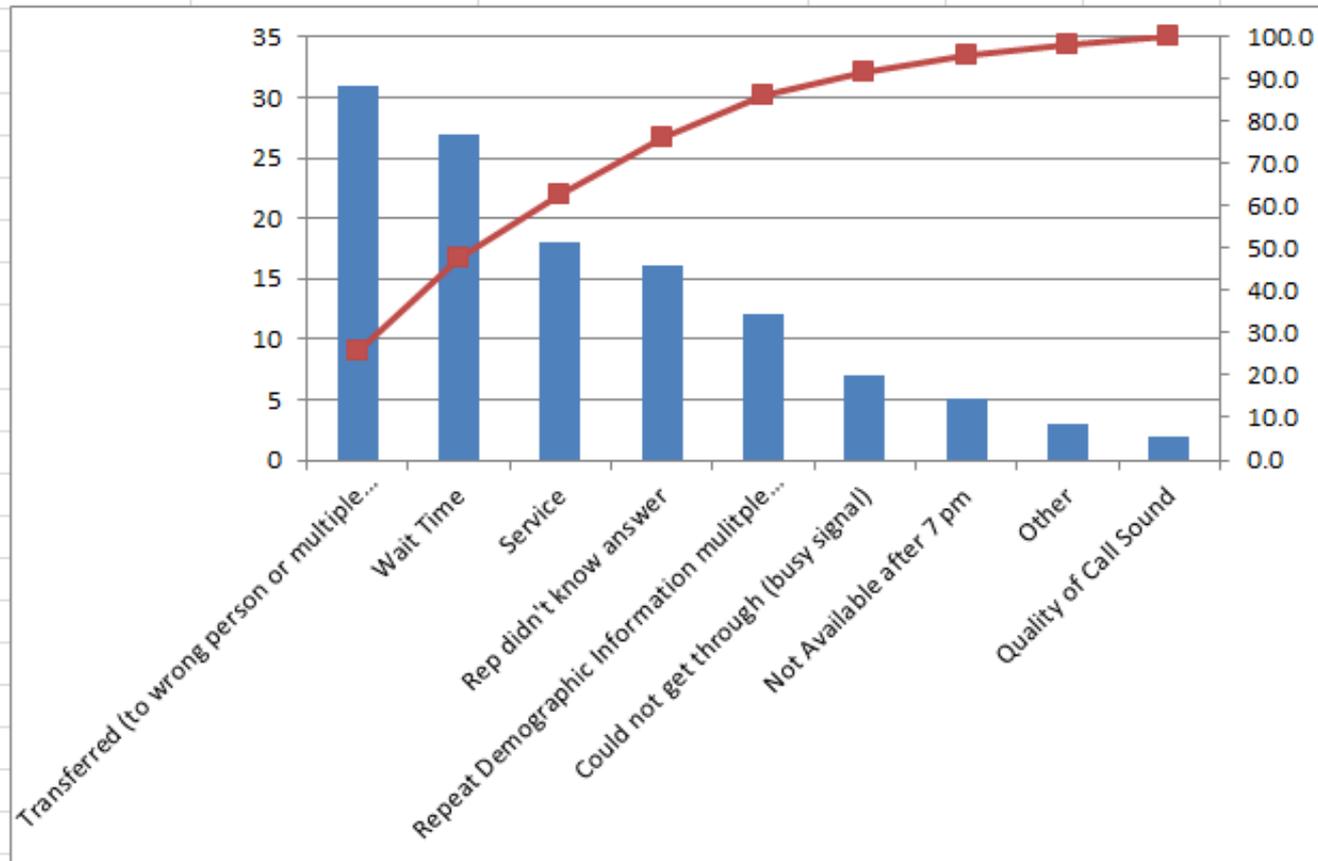
These steps are optional but are useful for analysis and communication:

- Calculate the percentage for each category: the subtotal for that category divided by the total for all categories. Draw a right vertical axis and label it with percentages. Be sure the two scales match: For example, the left measurement that corresponds to one-half should be exactly opposite 50% on the right scale.
- Calculate and draw cumulative sums: Add the subtotals for the first and second categories, and place a dot above the second bar indicating that sum. To that sum add the subtotal for the third category, and place a dot above the third bar for that new sum. Continue the process for all the bars. Connect the dots, starting at the top of the first bar. The last dot should reach 100 percent on the right scale.

Pareto Chart: Data Collection

- What is the number one reason you are late for work?
- Create a Pareto Chart based on the Class Data

Pareto Chart Example



Scatter Plot

SIMPLER. FASTER. BETTER. LESS COSTLY.

LEAN**Ohio**

Scatter Plot

- Scatter Plot illustrates the relationship between two variables
- The dots on the scatter plot represent data points
- Graph pairs of continuous data, with one variable on each axis, to examine the relationship between them

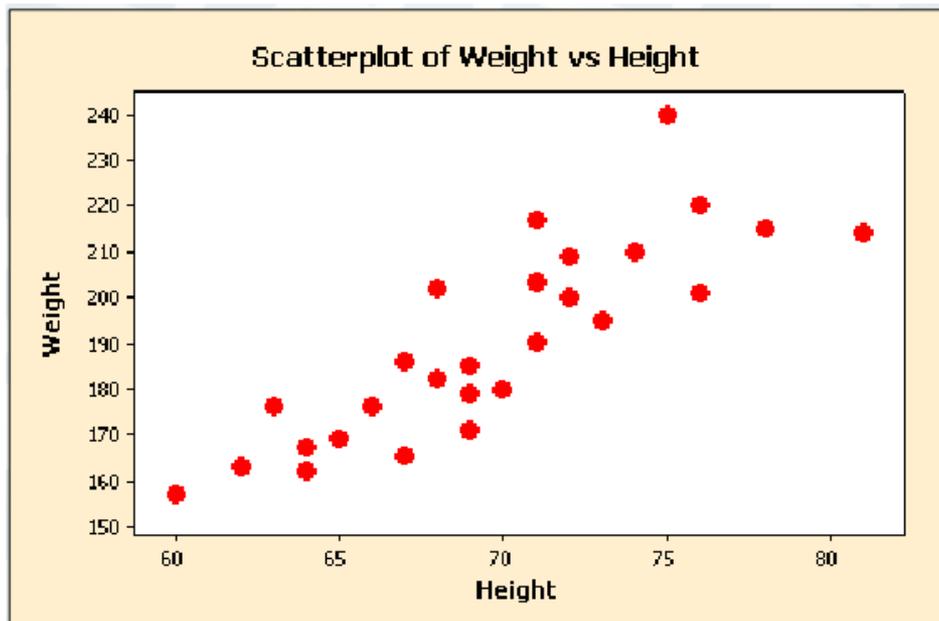
Use Scatter Plots: When

- You believe that two variables may be related
- Evaluating paired continuous data
- Attempting to identify potential root causes of a problem by relating two variables

Scatter Plots

Here is an example of a Scatter Plot:

- Consider the relationship between height and weight for men
- Generally speaking, the taller the man is, the more he weighs showing a positive relationship



Scatter Plot

- Is there a correlation between Height and Years of Service?
- Create a Scatter Plot using the class data
- Report Out

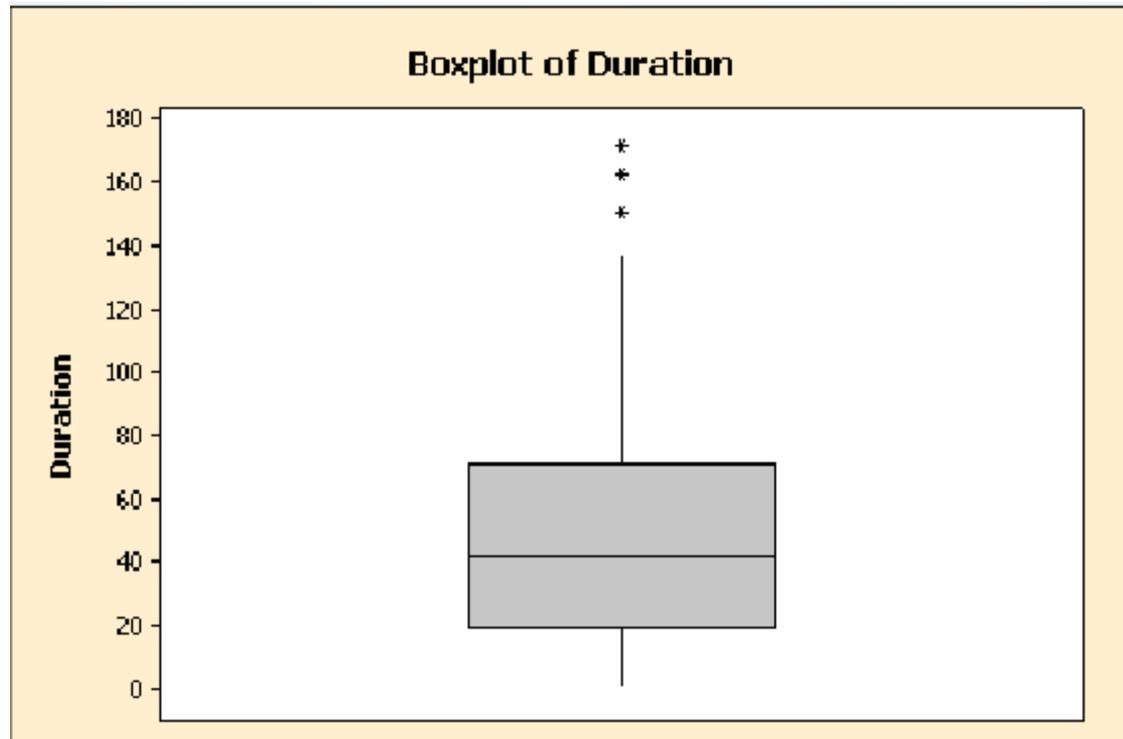
Box Plots

SIMPLER. FASTER. BETTER. LESS COSTLY.

LEANOhio

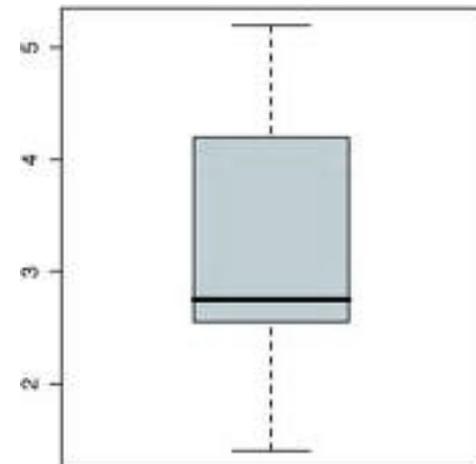
Box Plot

- Box Plots: (box-and-whisker plots) They are particularly useful for showing the distributional characteristics of data such as center and spread

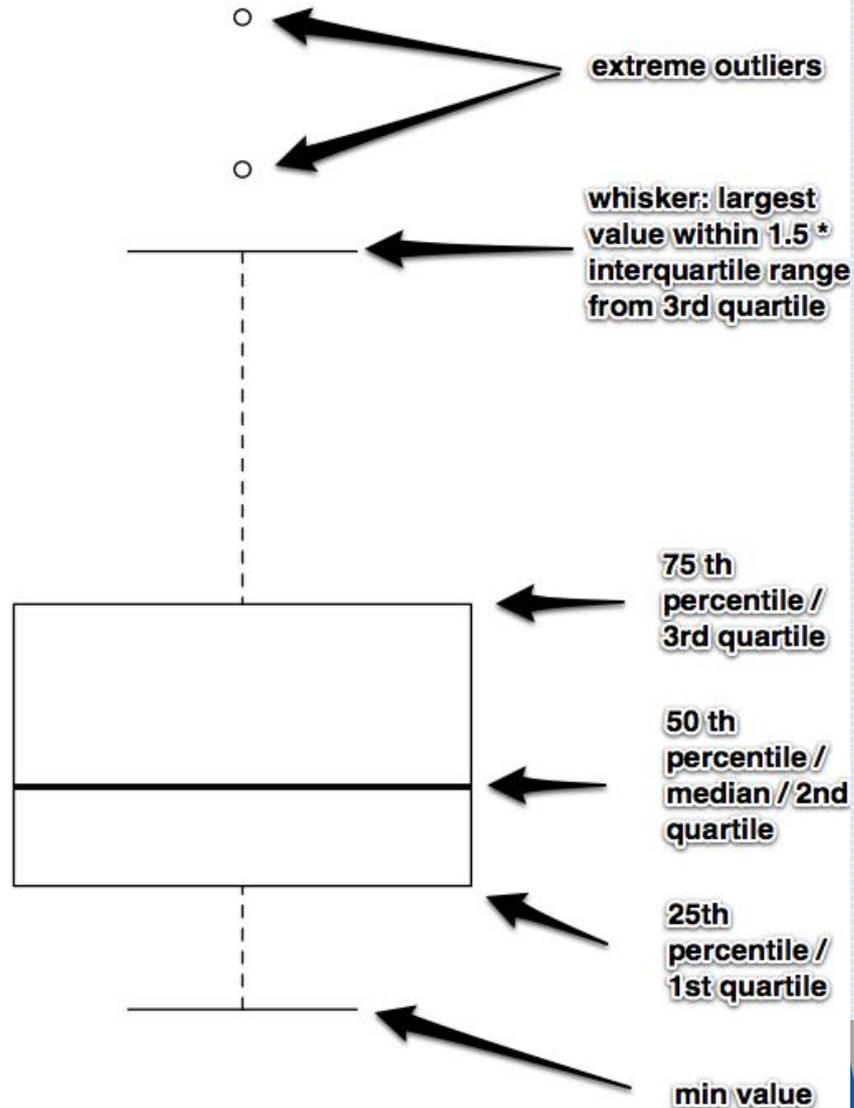


Box Plots

- A box plot consists of a box, whiskers, and outliers
- A line is drawn across the box at the median
- The bottom of the box is at the first quartile (Q1), and the top is at the third quartile (Q3) value
- The whiskers are the lines that extend from the top and bottom of the box to the adjacent values



Boxplot Example



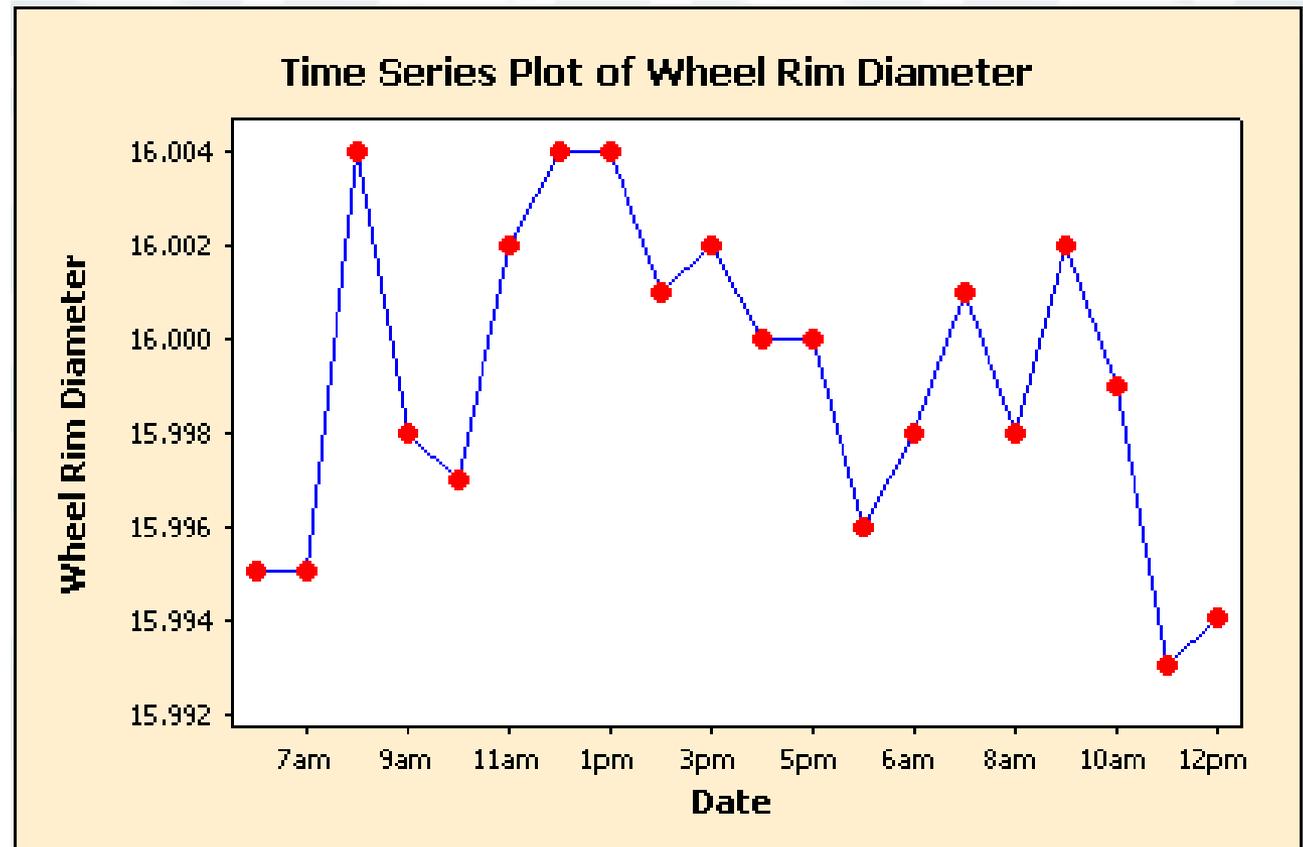
Run Chart

SIMPLER. FASTER. BETTER. LESS COSTLY.

LEANOhio

Run Charts

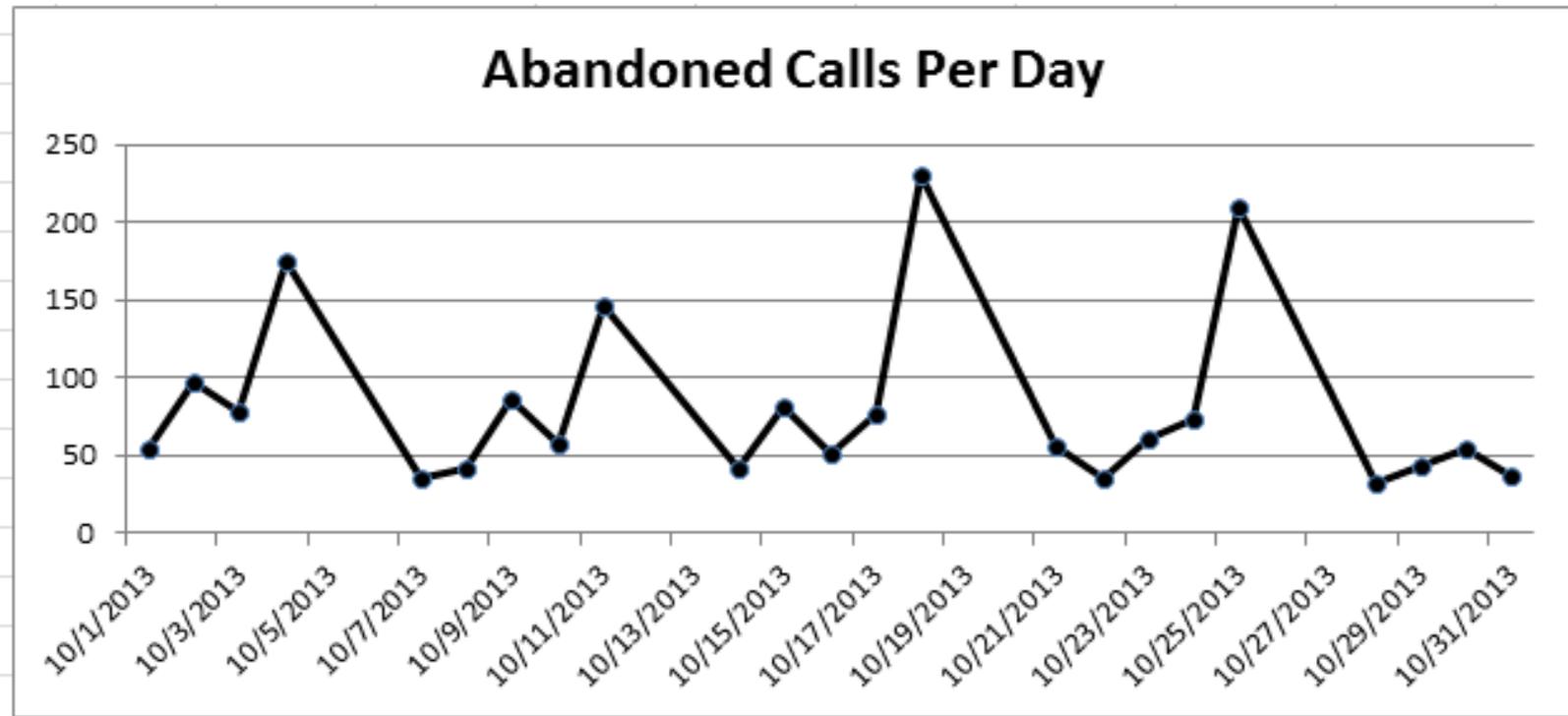
- Run Charts displays numeric data in time order.
- Used to show patterns over time



Run Chart Example

- Create a Run Chart for Class Height (graph in seat order)

Run Chart Example



Questions?

Creating Graphs in Excel Directions

- This section is dedicated to creating the various graphs in excel – To be honest you may want to get You Tube videos for Graphs!

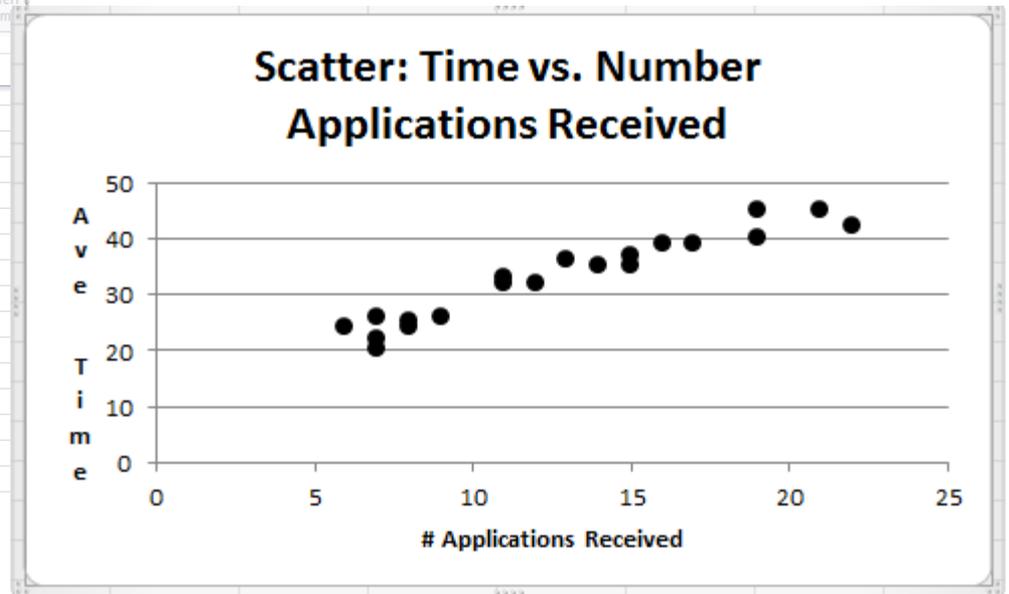
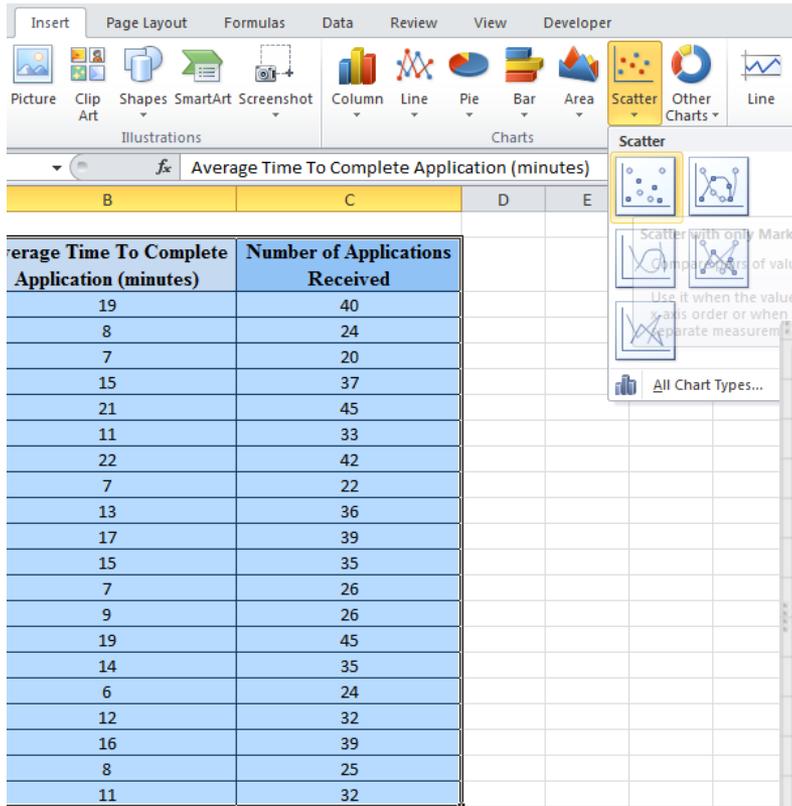
Dot Plots in Excel

- Steps to complete in Excel
 - Open/enter your dataset
 - Sort the data from smallest to largest
 - Add a Frequency row to the right of the data
 - Count how many occurrences there are for the first number and put that number of occurrences in the frequency column for that number
 - On the first occurrence of the 2nd number, Add = `if(A3=A2,B2+1,1)` in the Frequency column.
 - Drag and drop this value through the entire dataset
 - Go to Insert, Scatterplot and select both columns of data
 - Clean up the graph by deleting the gridlines, the scaling on the left, the legend, and by minimizing the graph from the bottom

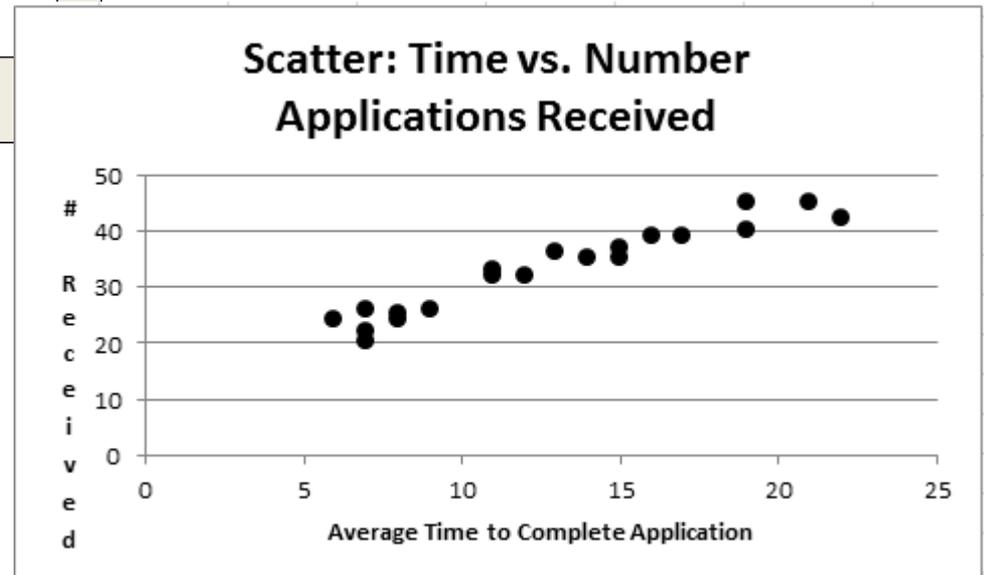
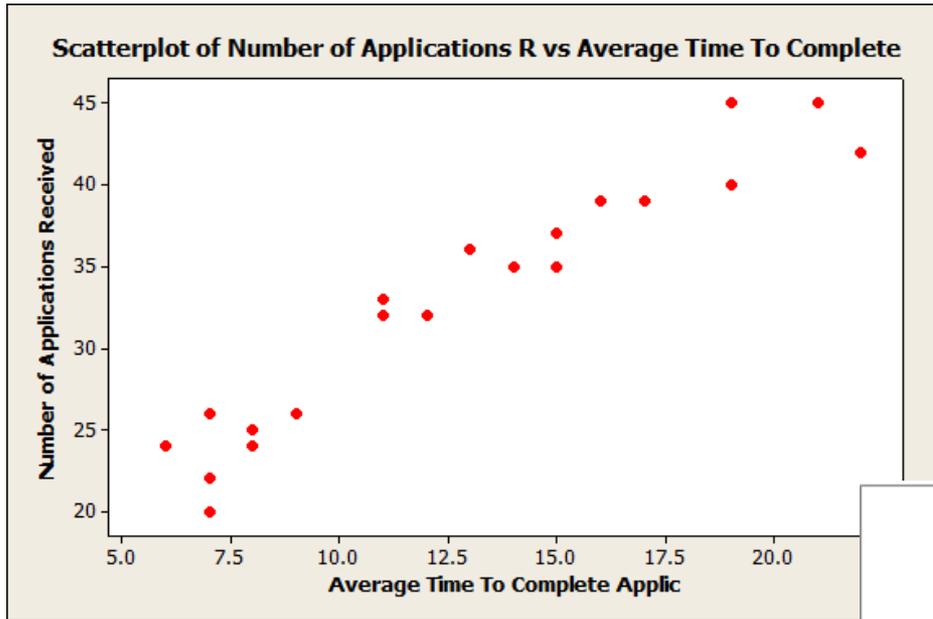
Scatter Diagram in Excel

- Input Data into Excel (Two titled columns)
- Highlight Data
- On Insert Menu – In the Charts Section Select the Scatter Option (Scatter with Markers Only is the one I typically use).

Scatter Diagram in Excel



Scatter Plot in Minitab and Excel



Scatter Diagram in Excel

- Under the tab Scatter Diagram – use the data to create a scatter diagram.

Histogram in Excel

SIMPLER. FASTER. BETTER. LESS COSTLY.

LEANOhio

Creating a Histogram in Excel

- Place Data in a column
- Use formula $k=1 + \text{Log}_2N$
- k is the number of bins
- N is the total number of data values
- This formula is easy to do in excel.
- Once you have the number of bins (for 30 samples you would have 6 bins). Find the range of the data.
- Take high value – the low value and divide by your number of bins. (for our example it was 3.5) to get the size of your bins.

Creating a Histogram in Excel

- Take the lowest value of your data set and add the bin size. Repeat this for all the bins to get your ranges.
- For our example the six ranges were:
 - 8-11.5
 - 11.5-15
 - 15-18.5
 - 18.5-22
 - 22-25.5
 - 25.5-29

Creating a Histogram in Excel

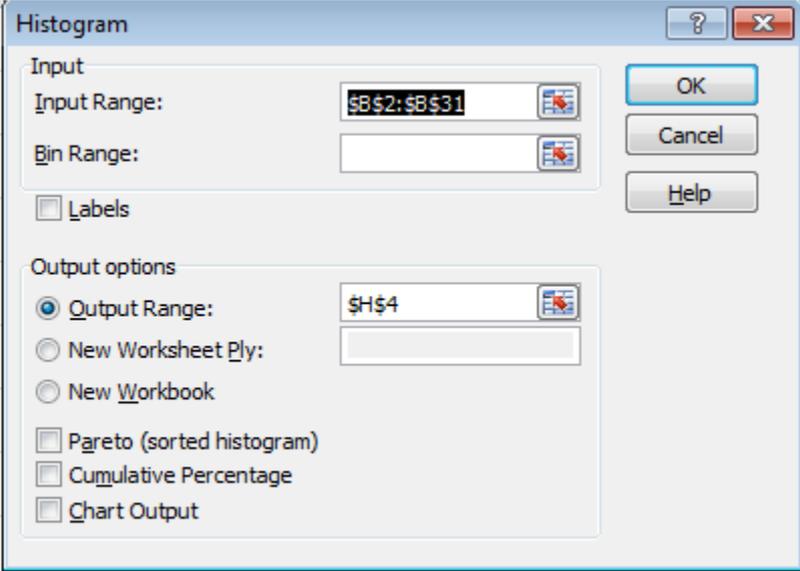
- Put your bins into excel using the higher data point

B	C	D
Histogram Data		Bins
29		8
10		11.5
20		15
27		18.5
21		22
15		25.5
15		29
19		

Data Analysis Pack

- In excel under Data Tab click the Data Analysis Pack
- Select Histogram and select OK
- In the Input Range – Select your Data

Histogram Data	
29	
10	
20	
27	
21	
15	
15	
19	
22	



The screenshot shows the 'Histogram' dialog box in Excel. The 'Input Range' is set to '\$B\$2:\$B\$31'. The 'Bin Range' is empty. Under 'Output options', 'Output Range' is set to '\$H\$4'. The 'Labels' checkbox is unchecked. Under 'Output options', 'Output Range' is selected with a radio button. The 'Pareto (sorted histogram)', 'Cumulative Percentage', and 'Chart Output' checkboxes are also unchecked. The 'OK', 'Cancel', and 'Help' buttons are visible on the right side of the dialog box.

Data Analysis Pack

- In the Bin Range – Select your Bin numbers
- In Output – Select where you would like the data to be displayed

Bins			
8			
11.5			
15			
18.5			
22			
25.5			
29			

Histogram

Input

Input Range: \$B\$2:\$B\$31

Bin Range: \$D\$2:\$D\$8

Labels

Output options

Output Range: \$H\$4

New Worksheet Ply:

New Workbook

Pareto (sorted histogram)

Cumulative Percentage

Chart Output

OK

Cancel

Help

This is the output

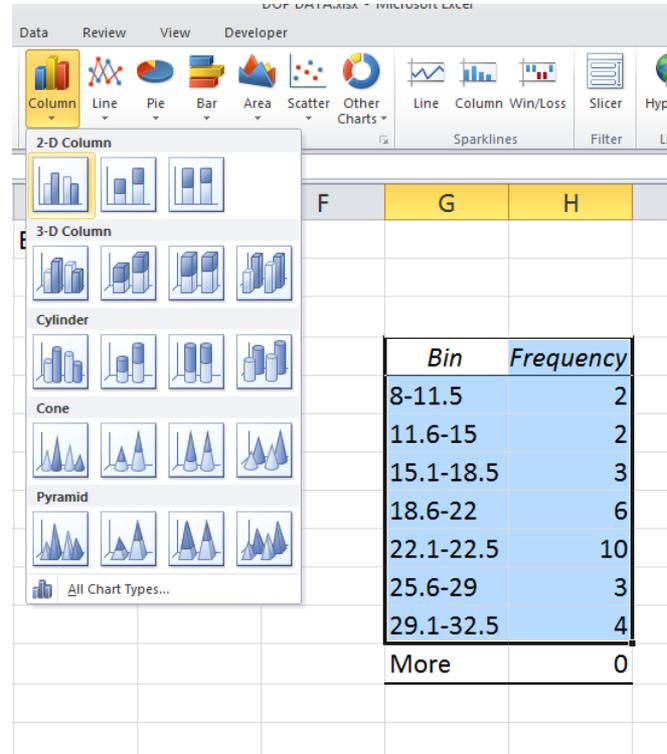
<i>Bin</i>	<i>Frequency</i>
11.5	4
15	3
18.5	6
22	10
25.5	3
29	4
More	0

Modify the Bin numbers to get it ready to create a Histogram!

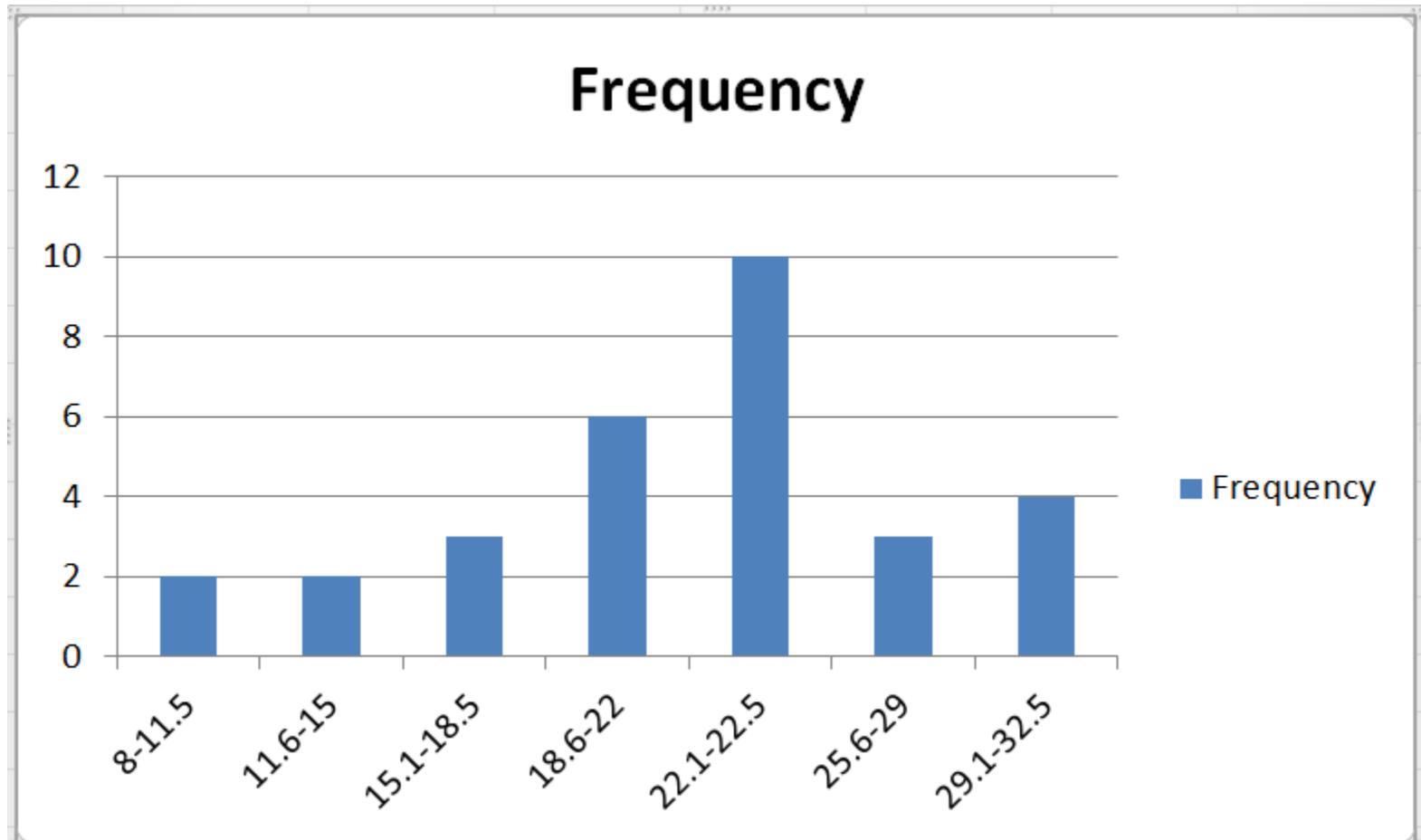
<i>Bin</i>	<i>Frequency</i>
8-11.5	4
15.1-18.5	3
18.6-22	6
22.1-22.5	10
25.6-29	3
29.1-32.5	4
More	 (Ctrl) 0

Histogram

- Select the data
- And insert a 2-D column Chart

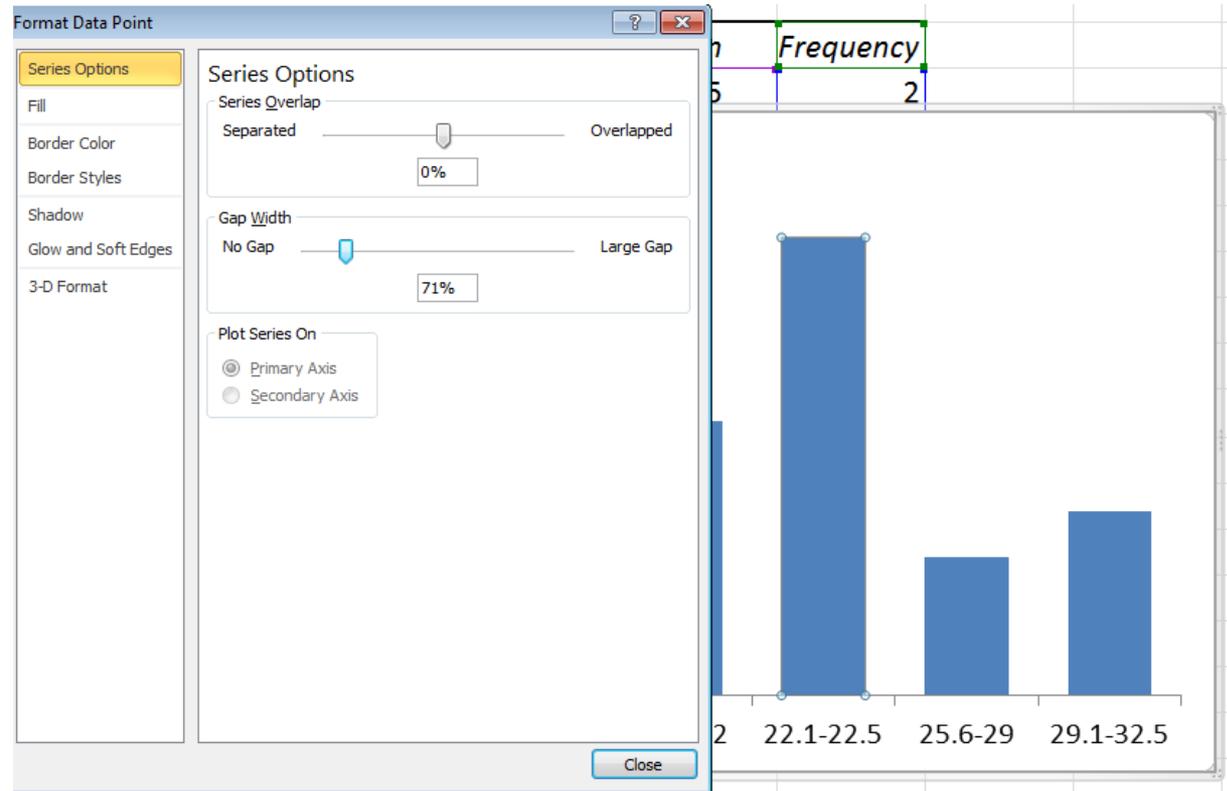


This is what you get – VERY CLOSE TO A HISTOGRAM, BUT NOT QUITE!



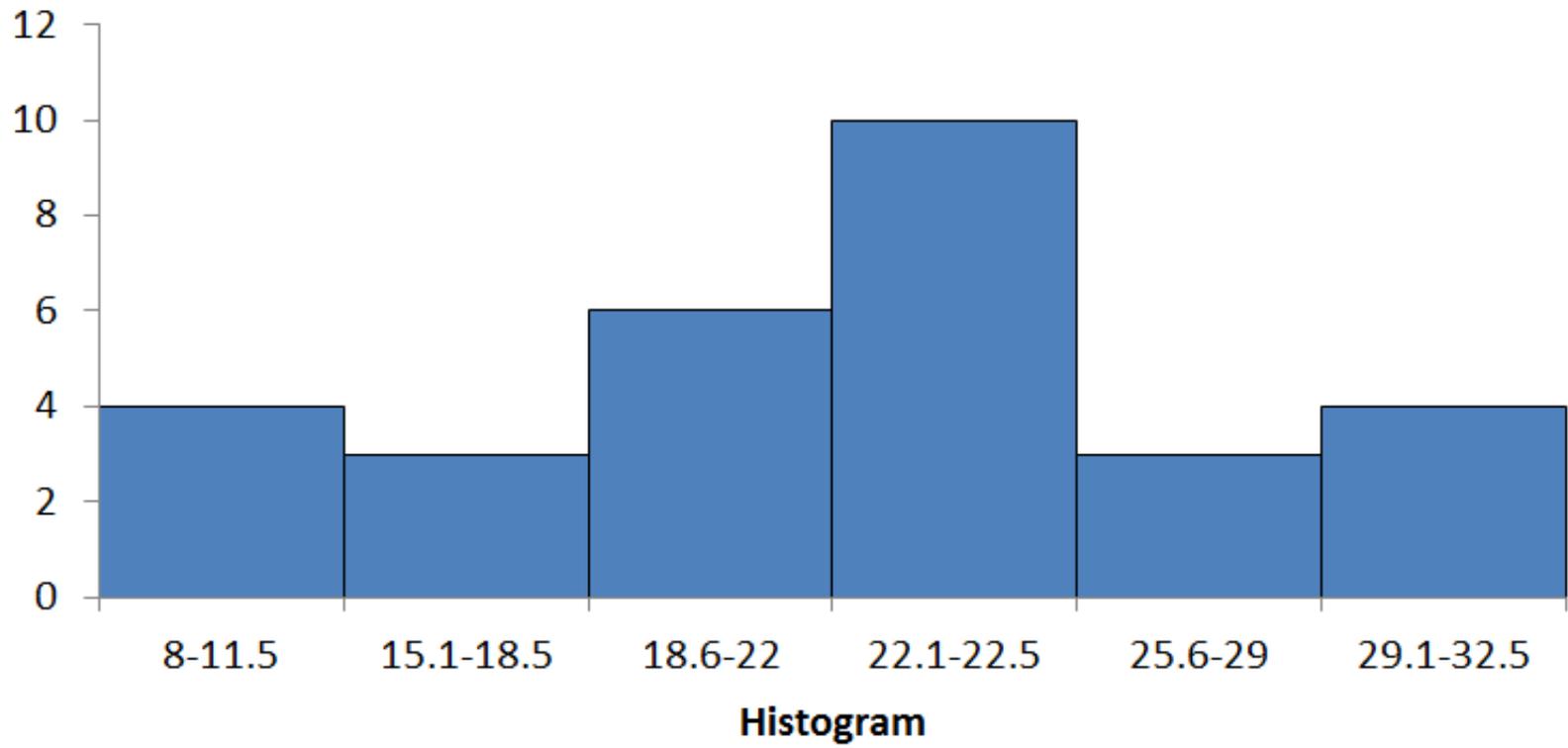
Histogram in Excel

- Remove the Title and the Series!
- Select the Bars, Right Click and select Format Data Point
- Under Series Options reduce Gap to 0.



Make it look Pretty

Histogram Example



Pareto Charts

SIMPLER. FASTER. BETTER. LESS COSTLY.

LEANOhio

Pareto Chart in Excel

- Enter Data in Excel
- Rank Order Data (So Highest at Top and Lowest at Bottom)

Error on Application Form	Number Returned
Approval Signatures Missing	23
Contact Information Missing	17
Incorrect Form Completed	10
Past Deadline	16
Missing Supporting Documents	287
No Budget Plan	9
No Audit Plan	7
Incorrect Coding	155
Incorrect Region	42

Error on Application Form	Number Returned
Missing Supporting Documents	287
Incorrect Coding	155
Incorrect Region	42
Approval Signatures Missing	23
Contact Information Missing	17
Past Deadline	16
Incorrect Form Completed	10
No Budget Plan	9
No Audit Plan	7

Pareto Chart in Excel

- Calculate the Cumulative Number for the data set.
- Add a column and title it “Cumulative Number”.
- Pull over Top Number
- Use an equation to add top number to next number.
- Pull equation down to get all the Cumulative numbers.

Error on Application Form	Number Returned	Cumulative Number
Missing Supporting Documents	287	287
Incorrect Coding	155	=C8+D7
Incorrect Region	42	
Approval Signatures Missing	23	
Contact Information Missing	17	
Past Deadline	16	
Incorrect Form Completed	10	
No Budget Plan	9	
No Audit Plan	7	

Error on Application Form	Number Returned	Cumulative Number
Missing Supporting Documents	287	287
Incorrect Coding	155	442
Incorrect Region	42	484
Approval Signatures Missing	23	507
Contact Information Missing	17	524
Past Deadline	16	540
Incorrect Form Completed	10	550
No Budget Plan	9	559
No Audit Plan	7	566

Pareto Chart in Excel

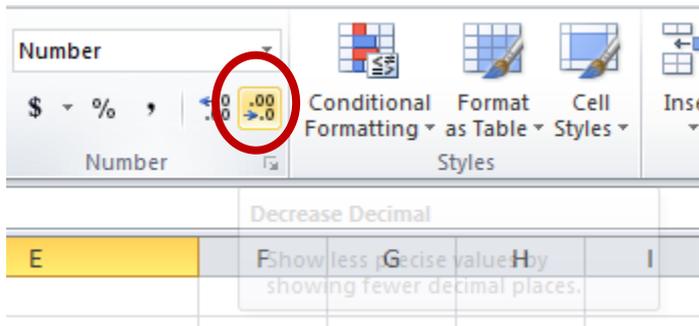
- Now calculate the Cumulative Percentage for the data set.
- Add column and title “Cumulative Percentage”
- Use Equation to take first data point divided by the Total Cumulative Number * 100 to get a percentage.
- Note that I added a dollar sign before the D and 15 in order to get the absolute value and pull the equation down.
- Pull the equation down to get Cumulative percentage

Error on Application Form	Number Returned	Cumulative Number	Cumulative Percentage
Missing Supporting Documents	287	287	<code>=(D7/\$D\$15*100)</code>
Incorrect Coding	155	442	
Incorrect Region	42	484	
Approval Signatures Missing	23	507	
Contact Information Missing	17	524	
Past Deadline	16	540	
Incorrect Form Completed	10	550	
No Budget Plan	9	559	
No Audit Plan	7	566	

Pareto Chart in Excel

- Clean up Data Set.
- Reduce data percentages to one decimal point.

Error on Application Form	Number Returned	Cumulative Number	Cumulative Percentage
Missing Supporting Documents	287	287	50.7067138
Incorrect Coding	155	442	78.0918728
Incorrect Region	42	484	85.5123675
Approval Signatures Missing	23	507	89.5759717
Contact Information Missing	17	524	92.5795053
Past Deadline	16	540	95.4063604
Incorrect Form Completed	10	550	97.1731449
No Budget Plan	9	559	98.7632509
No Audit Plan	7	566	100.0000000



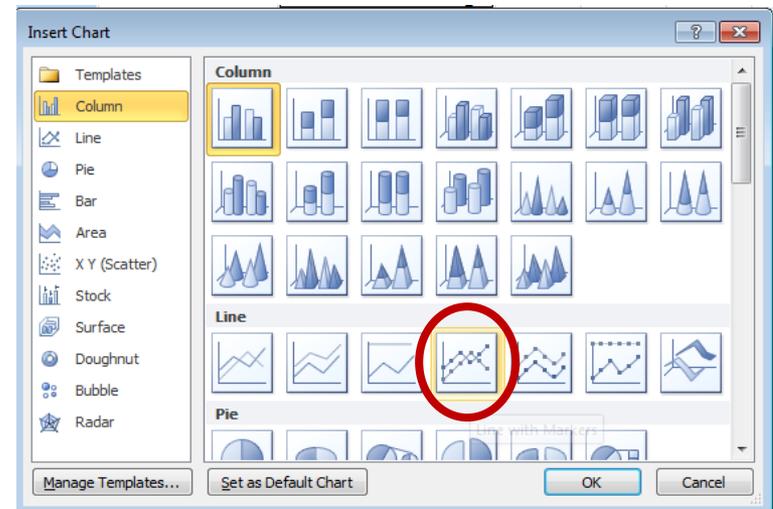
Cumulative Percentage
50.7
78.1
85.5
89.6
92.6
95.4
97.2
98.8
100.0

Pareto Chart in Excel

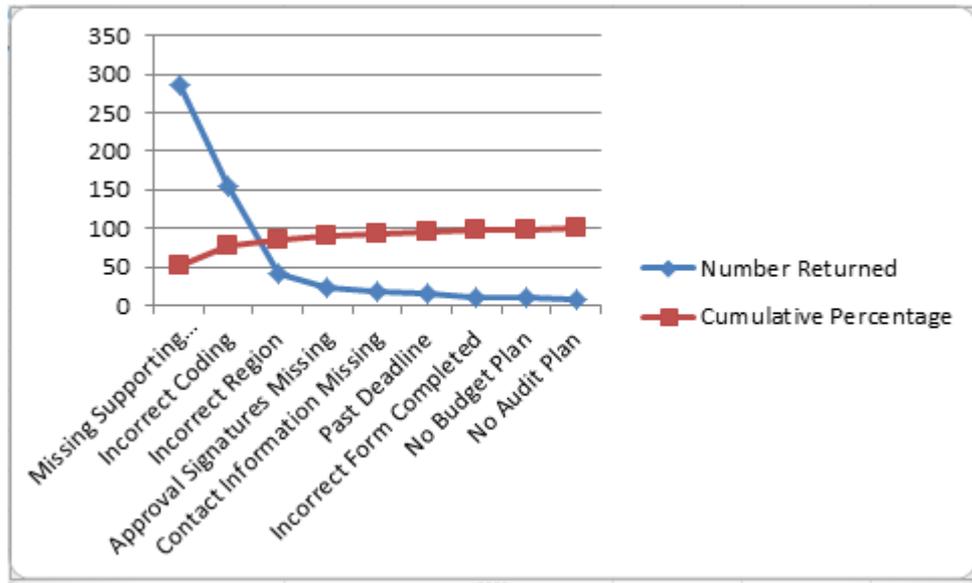
- Select Data for Chart Creation: Column A,

Error on Application Form	Number Returned	Cumulative Number	Cumulative Percentage
Missing Supporting Documents	287	287	50.7
Incorrect Coding	155	442	78.1
Incorrect Region	42	484	85.5
Approval Signatures Missing	23	507	89.6
Contact Information Missing	17	524	92.6
Past Deadline	16	540	95.4
Incorrect Form Completed	10	550	97.2
No Budget Plan	9	559	98.8
No Audit Plan	7	566	100.0

- Insert a Line Chart
- (Line with Markers)

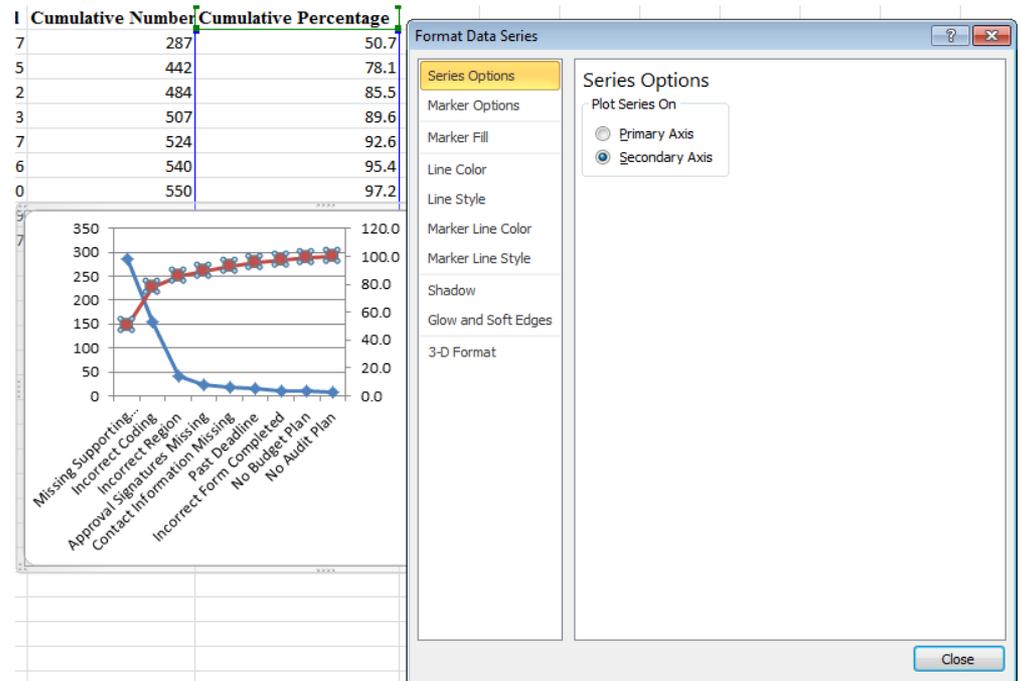


Pareto Chart in Excel



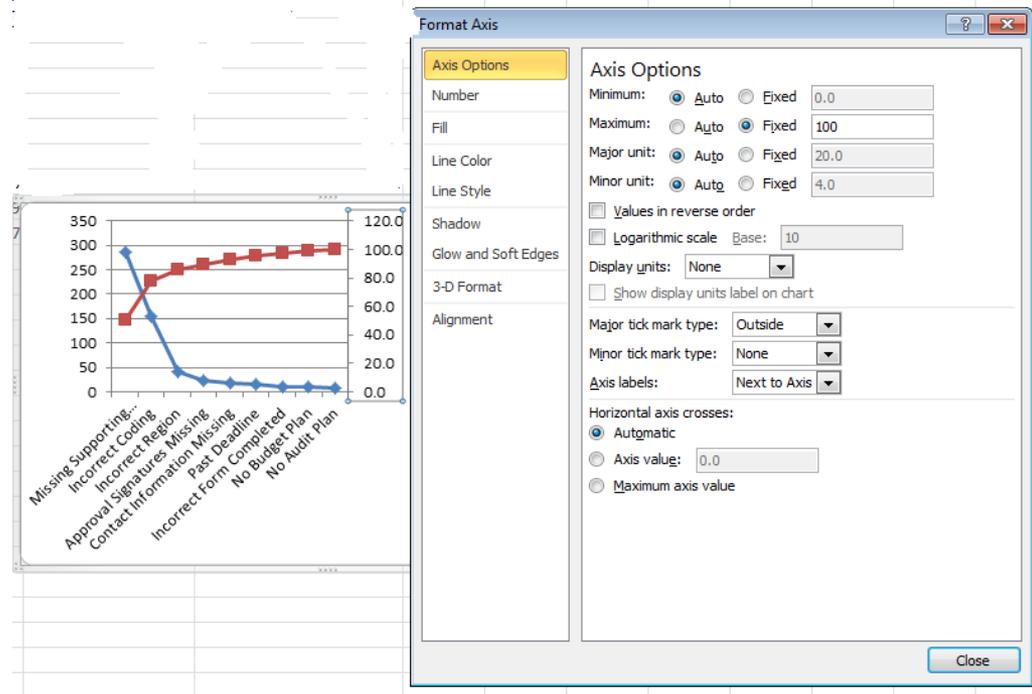
Pareto Chart in Excel

- Now Format the Graph!
- Select the Cumulative Percentage data set on the graph, right click and select format data, Under series options – select Secondary Axis.

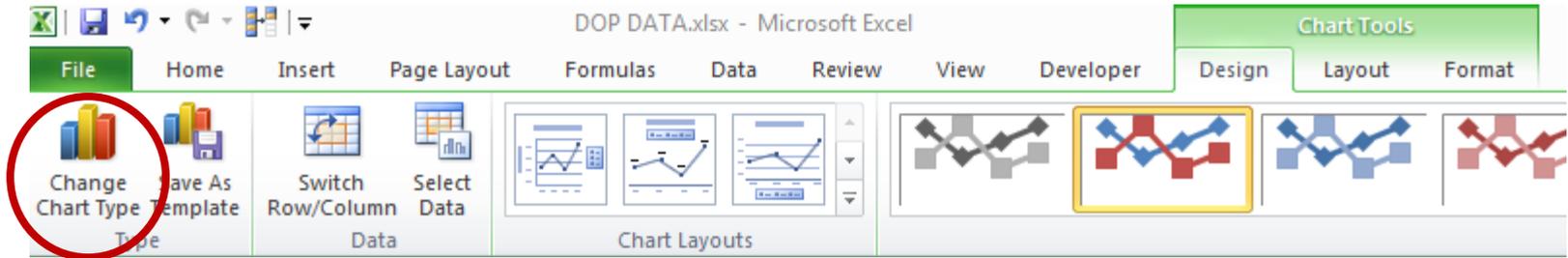


Pareto Chart in Excel

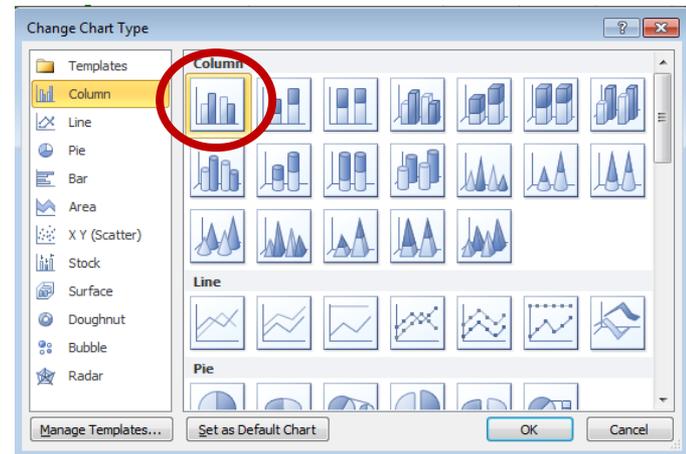
- To show 100% (not 120) Right Click on Axis and select format axis, Change the Maximum from Auto to Fixed and change the value to 100.
- Click Close.



Pareto Chart in Excel



- Change to a Bar Chart.
- Select the Blue Data Set.
- Under the Chart Tools Design Tab – Select Change Chart Type.
- Select the Bar Chart Option



Questions

SIMPLER. FASTER. BETTER. LESS COSTLY.

LEANOhio