

Create a Histogram

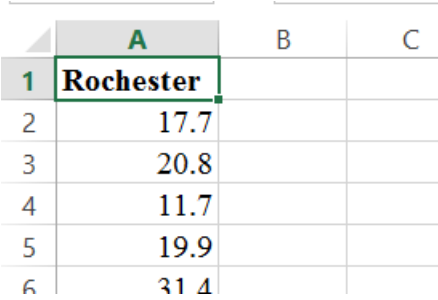
Excel Step-by-Step How-to for Windows

Excel for Mac Instructions on [page 12](#)

Instructions: Use this guide to create a histogram using Excel.

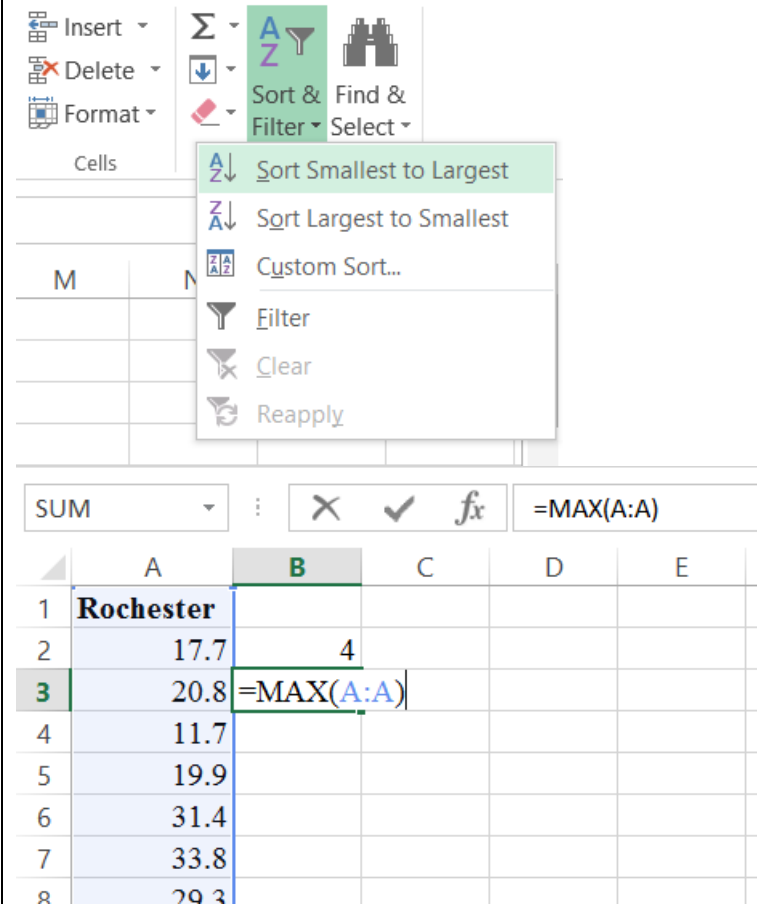
Data requirement: one variable with quantitative data.

Sample Data: yearly snowfall in Rochester.

Step	Windows Instructions + Screen Shot
1. Arrange the data you want to use into a column. If you have multiple variables, ensure that each column is a unique variable.	

2. Determine the minimum and maximum values of your data set.

If you have a large data set, you may want to use Excel to find the smallest and largest point. These values make it easier to determine the starting and ending values for your histogram. Select the column with your data and then use the “Sort” function or type `=MIN(A:A)` in a blank cell in a different column (i.e. column B) and then type `=MAX(A:A)` to get the biggest number. The data range A:A will select everything in column A.



The screenshot shows the Excel interface. The 'Sort & Filter' menu is open, displaying options: 'Sort Smallest to Largest', 'Sort Largest to Smallest', 'Custom Sort...', 'Filter', 'Clear', and 'Reapply'. Below the menu, a data table is visible with columns A through E. Column A contains the following values: 'Rochester', 17.7, 20.8, 11.7, 19.9, 31.4, 33.8, and 29.2. Column B contains the value 4 in row 2 and the formula `=MAX(A:A)` in row 3. The formula bar at the top shows `=MAX(A:A)`.

	A	B	C	D	E
1	Rochester				
2	17.7	4			
3	20.8	<code>=MAX(A:A)</code>			
4	11.7				
5	19.9				
6	31.4				
7	33.8				
8	29.2				

3. Based on your minimum and maximum values, choose an appropriate bin size for your histogram.

TIP: Selecting the right bin sizes for your histogram:



4. In a separate area of your Excel sheet, type-in your bin ranges.

A bin is the interval by which you want to sort your data. A histogram displays how many values from your data set fall into each bin (this is known as the frequency of the bin).

It is important to choose a bin size that is not too small or too large. You want the bin to be wide enough to show a pattern of distribution. If your bin is too small, it will be hard to make sense of the patterns.

In the Rochester snowfall example, the first bin would start at 5, the next one at 10, the next at 15, etc. Each bin should have the same width, or range.

	A	B	C	D	E	F	G
1	Rochester						
2	17.7		5				
3	20.8		10				
4	11.7		15				
5	19.9		20				
6	31.4		25				
7	33.8		30				
8	29.3		35				
9	15.2		40				
10	29.8		45				
11	14		50				
12	49.7		55				
13	61.3		60				
14	43.4		65				
15	11.9						

5. In the empty column next to your bin ranges, use an Excel formula to count how many values are less than or equal to each bin range.

=COUNTIF(range, criteria)

- **Range:** all the values of your data set
- **Criteria:** "<="&firstbinnumber (this criteria sub-formula indicates that you want to count all the numbers that are less than or equal to your first bin range number. In this example, the completed formula would look like this:

fx =COUNTIF(\$A2:\$A77,"<="&C2)

B	C	D	E	F	G
	5	2			
	10				
	15				
	20				
	25				
	30				

	A	B	C	D	E	F	G
1	Rochester						
2	17.7		5	2			
3	20.8		10	7			
4	11.7		15	19			
5	19.9		20	30			
6	31.4		25	41			
7	33.8		30	52			
8	29.3		35	56			
9	15.2		40	60			
10	29.8		45	63			
11	14		50	65			
12	49.7		55	64			
13	61.3		60	63			
14	43.4		65	64			

=COUNTIF(B2:B77, "<="&E9), where "E9" is populated with the number 5, which is the first bin range number. Note that your formula may call to another cell in your sheet that is not "E9." The key here is to reference the appropriate bin range number. Repeat this for each bin range.

TIP: To quickly
replicate formulas
across
cells:



First, it is important to lock the appropriate values in your formula. Locking a values makes the formula always reference the locked cell instead of changing based on location. You can lock the values in each formula by placing a "\$" in front of each coordinate. For example, by representing the formula as B\$24+C\$35, the value will always calculate using rows 24-35 no matter where that formula is placed. Note that if you also place a "\$" in front of the column letter, that will also lock the column.

Then, highlight the cell with the formula you want to copy, and drag the cursor across or down to the next cell.

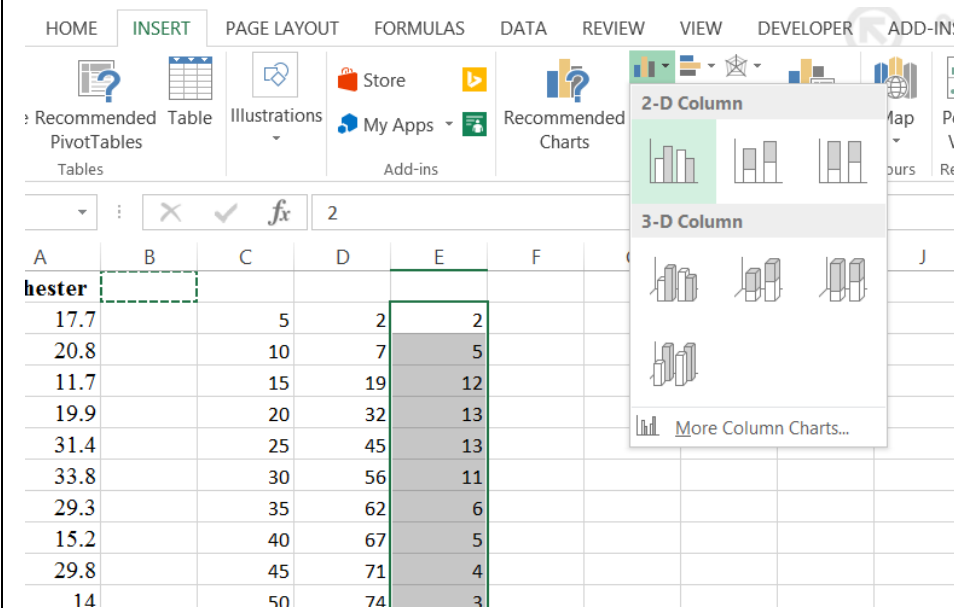
6. Generate frequency values that count how many values fall into each bin range.

In a third column next to your results from step 4, calculate the difference between each pair of bin count totals. This step is necessary because you don't want any values to be counted as falling into multiple bins. If there are 7 values less than 10, and 2 values less than 5, then there must be 5 values in the 5-10 bin. Note that the frequency of your first bin should equal whatever result you got for that bin in step 4. These values will be the histogram bar heights.

✕ ✓ <i>f_x</i>		=D3-D2				
B	C	D	E	F	G	H
	5	2	2			
	10	7	5			
	15	19	12			
	20	32	13			
	25	45	13			
	30	56	11			
	35	62	6			
	40	67	5			
	45	71	4			
	50	74	3			

7. Insert your histogram.

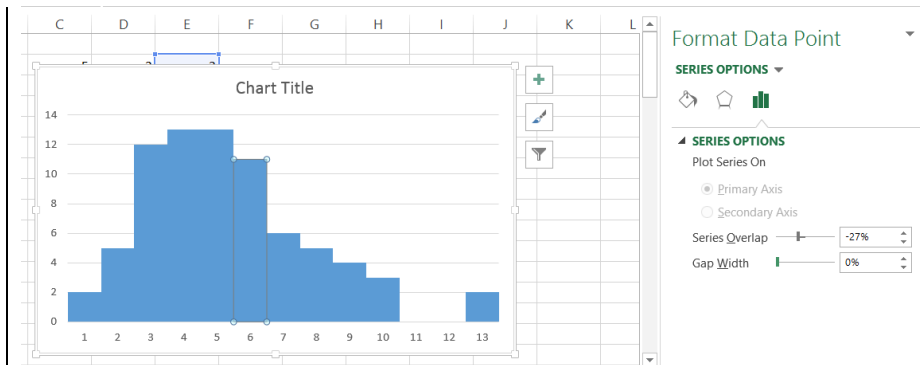
Highlight all the bin frequency values you calculated in step 5. Then, from the “insert” tab in the top toolbar, click to insert a simple 2-D bar chart.



8. Adjust the bar gap width to “0.”

To do this, double click on any of the bars of the chart to open the “Format data series” menu. Change “gap width” to “0%”.

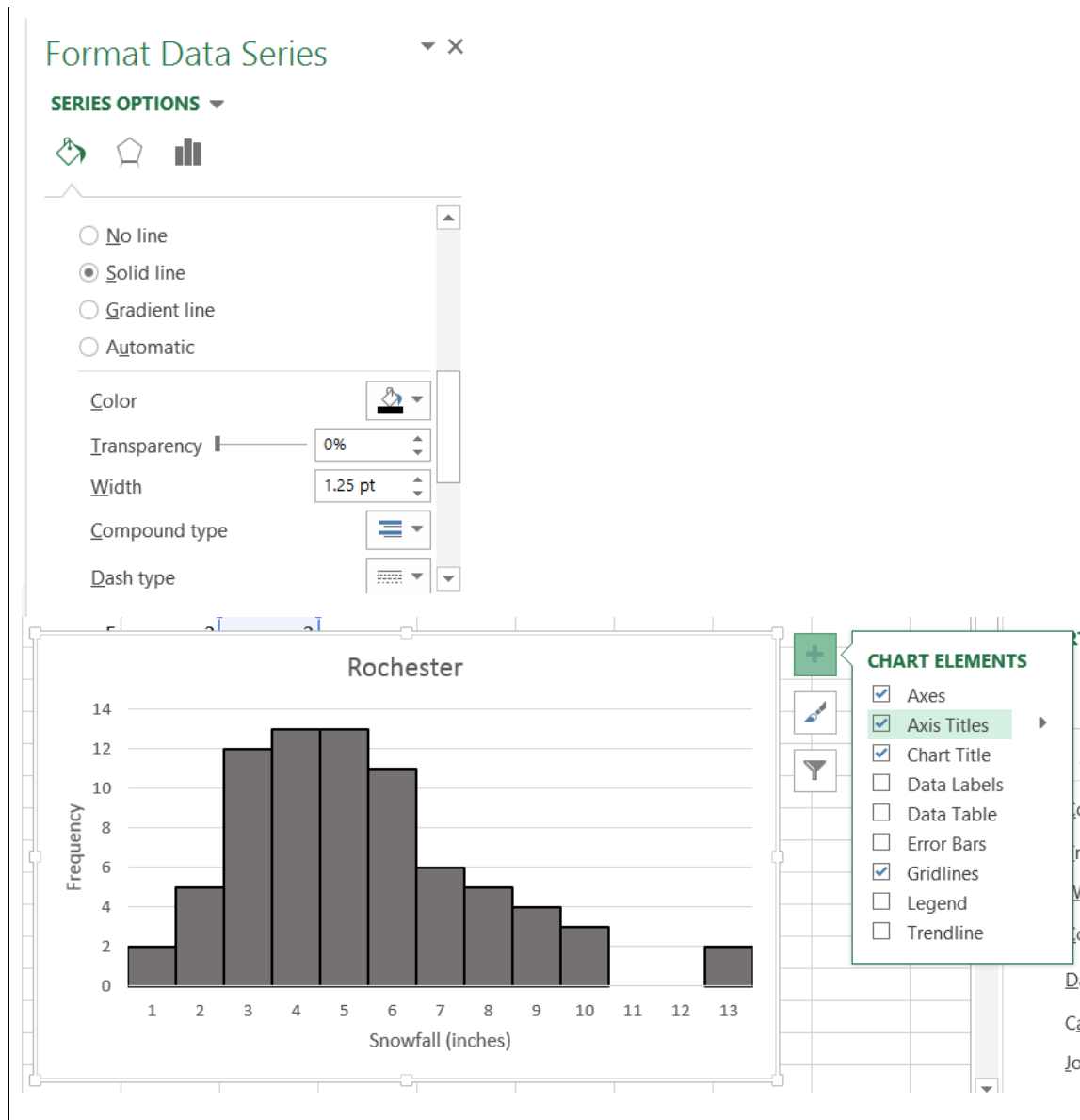
Note that histograms should never have gaps between bars since bins are continuous.



9. Format the histogram's display settings for legibility.

To do this, double click on any of the bars of the chart to open the "Format data series" menu.

- Change the "fill" color to gray
- Change the "border" to a solid black line
- Click on the "+" sign to add "axis titles" to the graph. In this example, the horizontal axis title should be "Snowfall (inches)" and the vertical axis title should be "Frequency"



TIP: To hide all the
field buttons on a
chart:



Right click on each field button and select "Hide all fields buttons on chart."

Create a Histogram

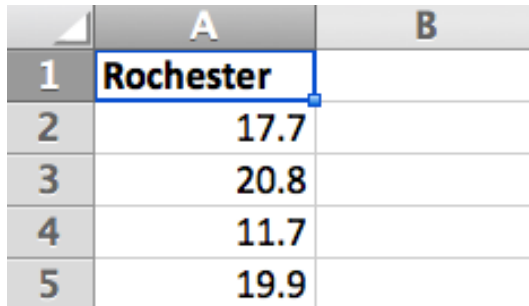
Excel Step-by-Step How-to for Mac

Excel for Windows Instructions on [page 1](#)

Instructions: Use this guide to create a histogram using Excel.

Data requirement: one variable with quantitative data.

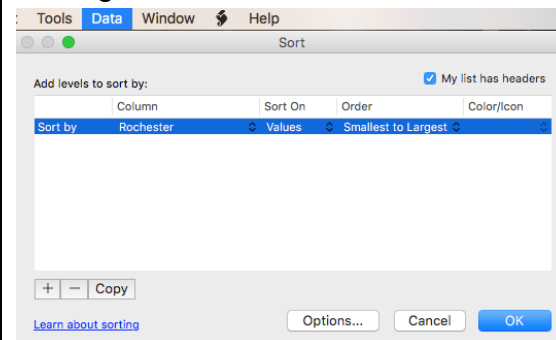
Sample Data: yearly snowfall in Rochester.

Step	Mac Instructions + Screen Shot																		
1. Arrange the data you want to use into a column. If you have multiple variables, ensure that each column is a unique variable.	 <table><tr><th></th><th>A</th><th>B</th></tr><tr><td>1</td><td>Rochester</td><td></td></tr><tr><td>2</td><td>17.7</td><td></td></tr><tr><td>3</td><td>20.8</td><td></td></tr><tr><td>4</td><td>11.7</td><td></td></tr><tr><td>5</td><td>19.9</td><td></td></tr></table>		A	B	1	Rochester		2	17.7		3	20.8		4	11.7		5	19.9	
	A	B																	
1	Rochester																		
2	17.7																		
3	20.8																		
4	11.7																		
5	19.9																		

2. Determine the minimum and maximum values of your data set.

If you have a large data set, you may want to use Excel to find the smallest point and the largest point to make it easier to determine the starting and ending values for your histogram. Type your data into a single column and then use the “Sort” function or type `=MIN(A:A)` in a blank cell in a different column (i.e. column B) and then type `=MAX(A:A)` to get the biggest number.

*For Macs: The “Sort” function is found within the Data tab. The ‘Order’ is where ‘Smallest to Largest’ can be selected.



	A	B
1	Rochester	
2	17.7	4
3	20.8	=MAX(A:A)
4	11.7	
5	19.9	

3. Based on your minimum and maximum values, choose an appropriate bin size for your histogram.

A bin is the interval by which you want to sort your data. A histogram displays how many values from your data set fall into each bin (this is known as the frequency of the bin). It's important to choose a bin size that is not too small or too large. You want the bin to be wide enough to show a pattern of distribution. If your bin is too small, it will be hard to make sense of the patterns.

TIP: Selecting the right bin sizes for your histogram:



4. In a separate area of your Excel sheet, type-in your bin ranges.

Professor input here... It's easiest to first calculate the range of your data set...

In the Rochester snowfall example, the first bin would start at 5, the next one at 10, the next at 15, etc. Each bin should have the same width, or range.

	A	B
1	Rochester	
2	17.7	5
3	20.8	10
4	11.7	15
5	19.9	20
6	31.4	25
7	33.8	30
8	29.3	35
9	15.2	40

5. In the empty column next to your bin ranges, use an Excel formula to count how many values are less than or equal to each bin range.

=COUNTIF(range, criteria)

- **Range:** all the values of your data set
- **Criteria:** “<=”&firstbinnumber (this criteria sub-formula indicates that you want to count all the numbers that are less than or equal to your first bin range number. In this example, the completed formula would look like this:

=COUNTIF(B2:B77, “<=”&E9), where “E9” is populated with the number 5, which is the first bin range number. Note that your formula may call to another cell in your sheet that is not “E9.” The key here is to reference the appropriate bin range number. Repeat this for each bin range.

SUM				
A	B	C	D	E
Rochester				
17.7	5	=COUNTIF(A2:A77, "<="&C2)		
20.8	10	COUNTIF(range, criteria)		
11.7	15			
19.9	20			
31.4	25			
33.8	30			
29.3	35			
15.2	40			

	A	B	C
1	Rochester		
2	17.7	5	2
3	20.8	10	7
4	11.7	15	19
5	19.9	20	30
6	31.4	25	41
7	33.8	30	52
8	29.3	35	56
9	15.2	40	60

TIP: To quickly replicate formulas across cells:



First, lock the values in each formula by placing a "\$" in front of each coordinate. For example, by representing the formula as $B\$24+C\35 , the value will always calculate using rows 24-35 no matter where that formula is placed. Note that if you also place a "\$" in front of the column letter, that will also lock the column.

Then, highlight the cell with the formula you want to copy, and drag the cursor across or down to the next cell.

6. Generate frequency values that count how many values fall into each bin range.

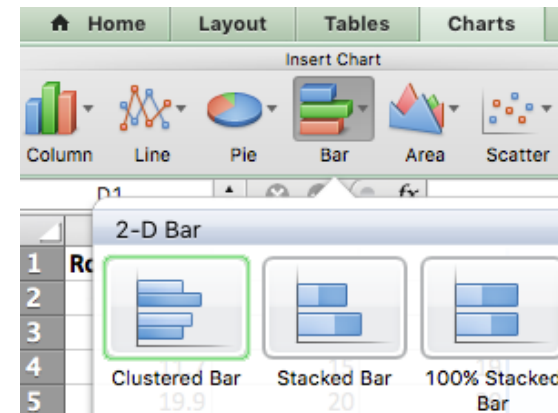
In a third column next to your results from step 4, calculate the difference between each pair of bin count totals. This step is necessary because you don't want any overlap in the numbers that fall into each bin. If there are 7 values less than 10, and 2 values less than 5, then there must be 5 values in the 5-10 bin. Note that the frequency of your first bin should equal whatever result you got for that bin in step 4. These values will be the histogram bar heights.

fx = C3-C2		
B	C	D
5	2	2
10	7	=C3-C2
15	19	
20	30	
25	41	
30	52	
35	56	
40	60	

7. Insert your histogram.

Highlight all the bin frequency values you calculated in step 5. Then, from the “charts” tab, click the “Bar” menu. Then, click “Clustered Bar” to insert a simple 2-D bar chart. A bar chart will automatically be inserted once the “Clustered Bar” button is clicked.

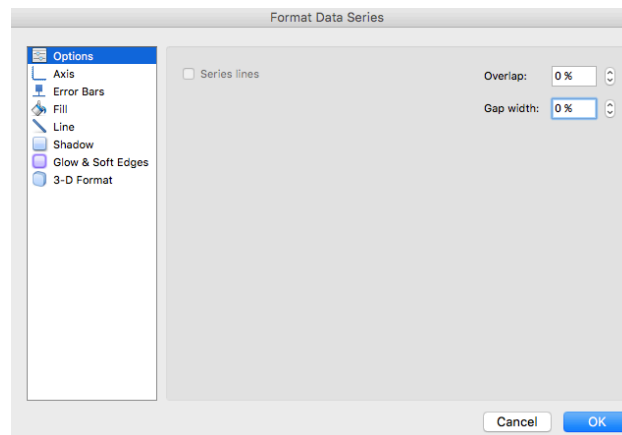
A	B	C	D
Rochester			
17.7	5	2	2
20.8	10	7	5
11.7	15	19	12
19.9	20	30	13
31.4	25	41	13
33.8	30	52	11
29.3	35	56	6
15.2	40	60	5



8. Adjust the bar gap width to “0.”

To do this, double click on any of the bars of the chart to open the "Format data series" menu. Change "Gap width" to “0%”

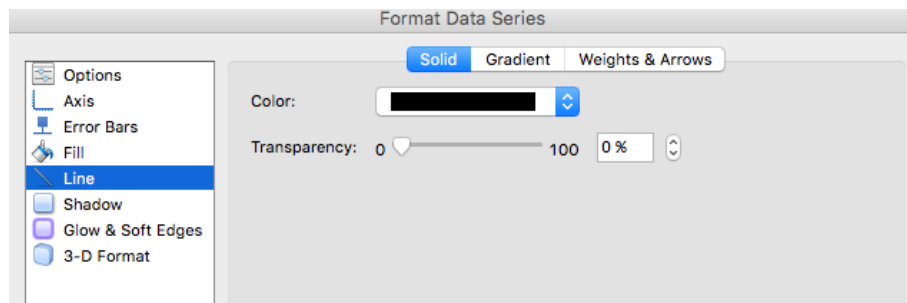
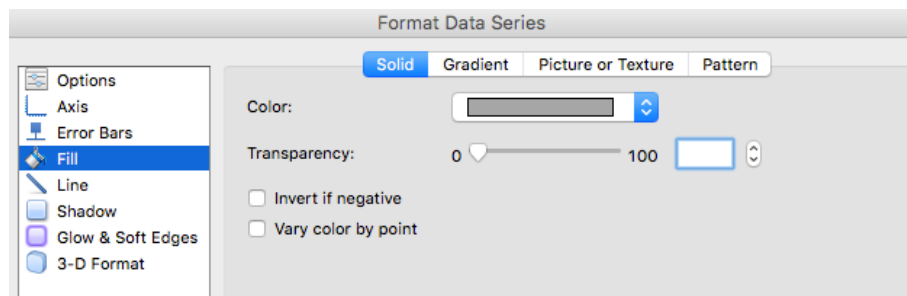
Note that histograms should never have gaps between bars since bins are continuous.



9. Format the histogram's display settings for legibility.

To do this, double click on any of the bars of the chart to open the "Format data series" menu.

- Change the "fill" color to gray
- Change the "border" to a solid black line
- Click on the "+" sign to add "axis titles" to the graph. In this example, the horizontal axis title should be "Snowfall (inches)" and the vertical axis title should be "Frequency"



TIP: To hide all
the field buttons
on a
chart:



Right click on each field button and select "Hide all fields buttons on chart."