

# community project

encouraging academics to share statistics support resources  
All stcp resources are released under a Creative Commons licence

stcp-samuels-2

## Using PivotTables in Excel

**Research question type:** Any involving descriptive statistics of single or multiple data series

**What kind of variables:** **Categorical** or **scale/continuous**

**Common applications:** Creating summary tables and charts of multiple series data sets

### Example: Student gender and height data

The heights of a group of 20 students were measured in cm as shown:

ID	Gender	Height
1	Male	177
2	Male	176
3	Male	165
4	Male	170
5	Female	160
6	Male	161
7	Female	160
8	Male	150
9	Female	157.5
10	Male	175

ID	Gender	Height
11	Female	163
12	Female	150
13	Female	156
14	Male	158
15	Male	147
16	Female	152
17	Female	165
18	Male	159
19	Male	151
20	Male	172

We wish to describe this data as a frequency distribution and a percentage frequency distribution split by gender.

### Creating a Frequency PivotTable and Chart

1. Enter the data into Excel in the form shown on the right. Notice how the cells around the data in Column D and Row 22 are blank and the column titles are in the top row. This means the cells containing the data can be identified by Excel as being separate, known as a **dataset**.

	A	B	C	D
1	ID	Gender	Height	
2	1	Male	177	
3	2	Male	176	
4	3	Male	165	
5	4	Male	170	
6	5	Female	160	
7	6	Male	161	
8	7	Female	160	
9	8	Male	150	
10	9	Female	157.5	
11	10	Male	175	
12	11	Female	163	
13	12	Female	150	
14	13	Female	156	
15	14	Male	158	
16	15	Male	147	
17	16	Female	152	
18	17	Female	165	
19	18	Male	159	
20	19	Male	151	
21	20	Male	172	
22				



Select any of the cells in the dataset (e.g. Cell A1) then select Insert – PivotTable. This opens the Create PivotTable dialogue box as shown on the right. Select Existing Worksheet and the cell E2 as shown below on the right. This creates a blank PivotTable box and a PivotTable Field List as shown below on the right.

2. Drag the *Gender* field from the PivotTable Field List to the Row Labels list. Drag the *Height* field from the PivotTable Field List to the Column Labels list. This should create an empty PivotTable like the one shown below:

Column Labels	147	150	151	152	156	157.5	158	159	160	161	163	165	170	172	175	176	177	Grand Total	
Row Labels																			
Female																			
Male																			
Grand Total																			

3. Drag the *ID* field from the PivotTable Field List to the Values list. In the Values box, select the triangle next to Sum of ID (ringed), then the option Value Field Settings as shown on the left,

then select Count from the dialogue box as shown above. This should change the PivotTable to look like this:

Count of ID	Column Labels	147	150	151	152	156	157.5	158	159	160	161	163	165	170	172	175	176	177	Grand Total	
Row Labels		1																		
Female				1		1		1						2			1		1	8
Male		1	1	1					1	1			1	1	1	1	1	1	1	12
Grand Total		1	2	1	1	1	1	1	1	2	1	1	2	1	1	1	1	1	1	20

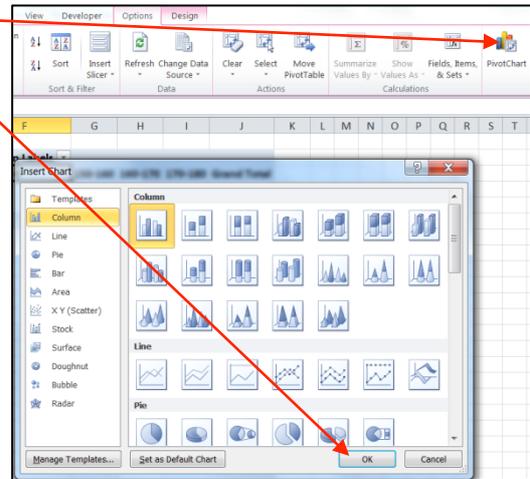
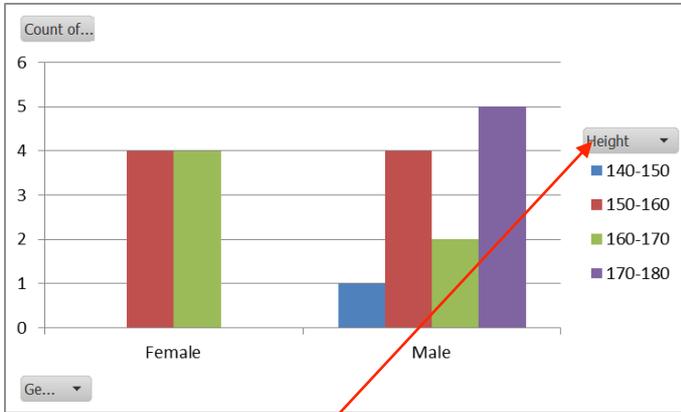
4. Right click one of the *Height* column labels (such as 147) and select Group... as shown on the left. In the Grouping dialogue box enter 140 and 180 as the start and end values as shown on the right. This creates a PivotTable with interval frequencies as shown below:

Count of ID	Column Labels	140-150	150-160	160-170	170-180	Grand Total
Row Labels						
Female			4	4		8
Male		1	4	2	5	12
Grand Total		1	8	6	5	20



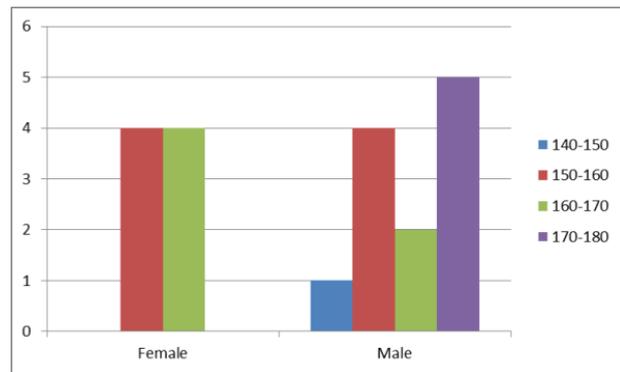
**Note:** The cut-off values are actually **less than 150** in the first group, so Student 12's height of 150cm has been counted in 150-160cm.

- Select the PivotChart button then click on OK to create the default two dimensional bar chart shown below:



- Right click on the *Height* Field Button then select Hide all Field Buttons on Chart option shown. This makes the chart look like this:

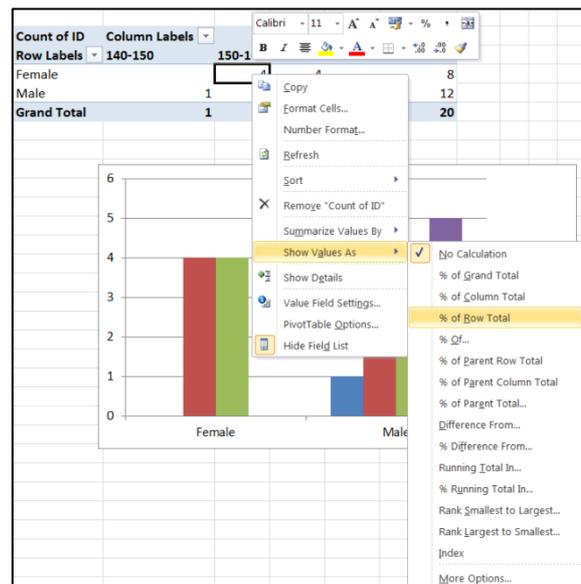
Other changes (such as chart and axis titles and changing the font) can be made using options on the Home, Layout and Format tabs.



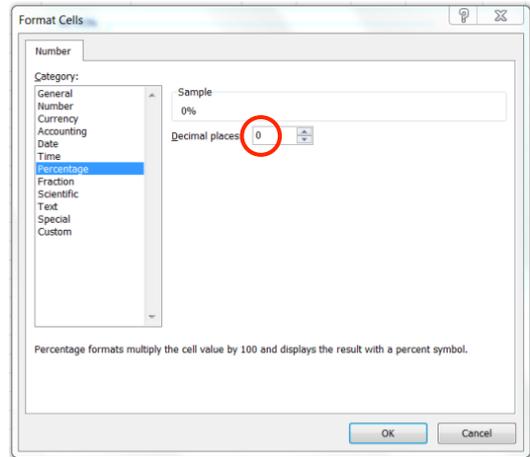
### Creating a Percentage Frequency PivotTable and Chart

- Right click on one of the frequency cells in the PivotTable, select Show Values As then select % of Row Total as shown on the right. This should change the PivotTable to look like this:

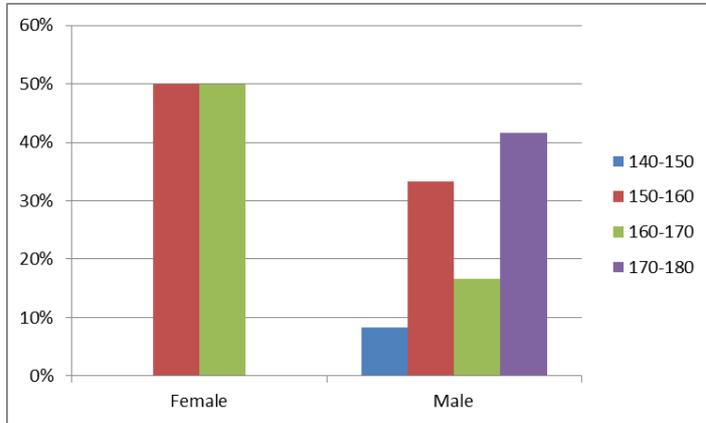
Count of ID	Column Labels	140-150	150-160	160-170	170-180	Grand Total
Female		0.00%	50.00%	50.00%	0.00%	100.00%
Male		8.33%	33.33%	16.67%	41.67%	100.00%
<b>Grand Total</b>		<b>5.00%</b>	<b>40.00%</b>	<b>30.00%</b>	<b>25.00%</b>	<b>100.00%</b>



- Right-click on a cell containing a percentage then select then select Number Format... then reduce the number of decimal places to zero using the dialogue box shown on the right:



The chart and PivotTable should then look like this:



Count of ID	Column Labels				
Row Labels	140-150	150-160	160-170	170-180	Grand Total
Female	0%	50%	50%	0%	100%
Male	8%	33%	17%	42%	100%
<b>Grand Total</b>	<b>5%</b>	<b>40%</b>	<b>30%</b>	<b>25%</b>	<b>100%</b>

**Note:**

- Percentage frequency tables and charts are useful when the group sizes are different
- If there are several groups, column percentages may also be useful and can be calculated in a similar way
- It is also sometimes useful to switch around the rows and columns in the data and see what effect this has on the table and chart using this button on the chart Design tab (select the chart first)

