stcp-samuels-2

**Using PivotTables in Excel**

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

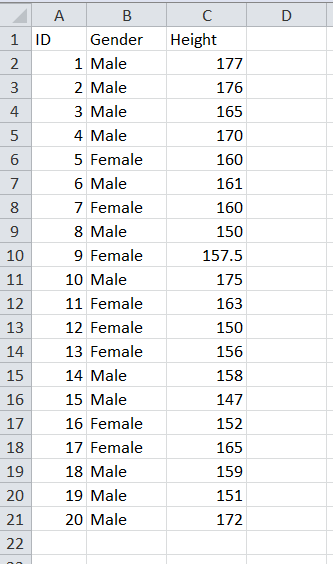
## What kind of variables: Categorical or scale/continuous

## 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:

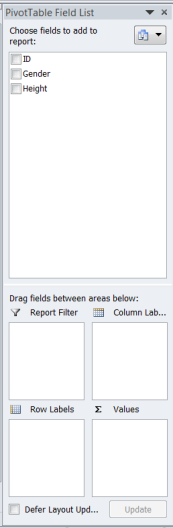
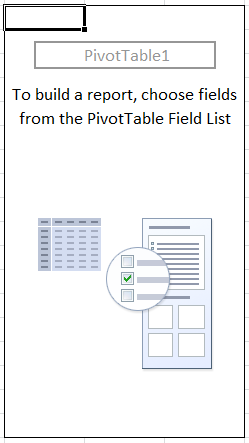
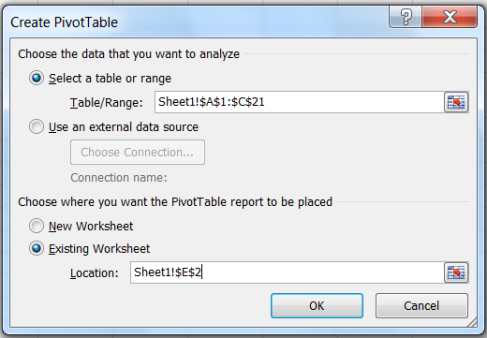
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ID | Gender | Height |  | ID | Gender | Height |
| 1 | Male | 177 |  | 11 | Female | 163 |
| 2 | Male | 176 |  | 12 | Female | 150 |
| 3 | Male | 165 |  | 13 | Female | 156 |
| 4 | Male | 170 |  | 14 | Male | 158 |
| 5 | Female | 160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

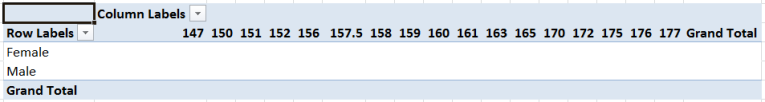
We wish to describe this data as a frequency distribution and a percentage frequency distribution split by gender.

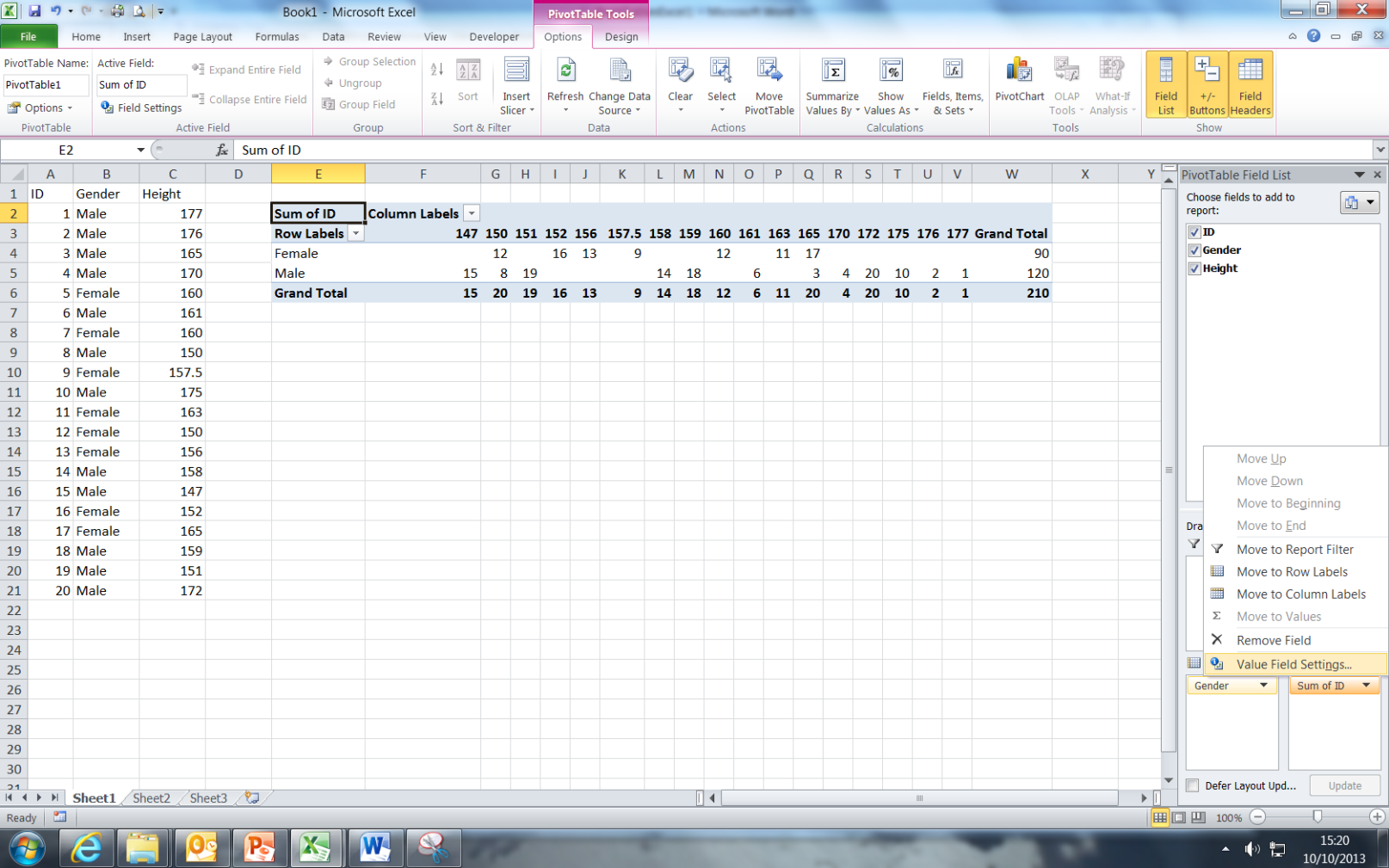
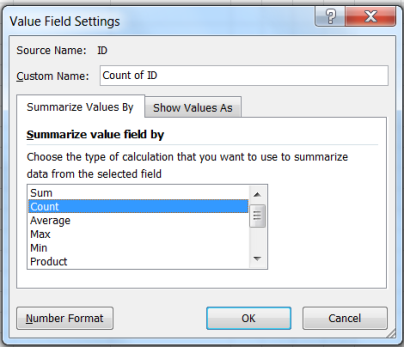
# Creating a Frequency PivotTable and Chart

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**.

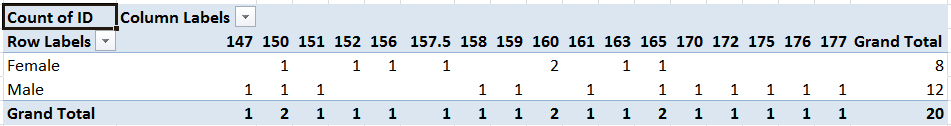
Select any of the cells in the dataset (e.g. Cell A1) then select Insert – PivotTable. This opens the Create PivotTable dialogue box as shown on the right. Select Existing Worksheet and the cell E2 as shown below on the right. This creates a blank PivotTable box and a PivotTable Field List as shown below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

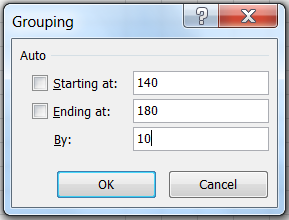
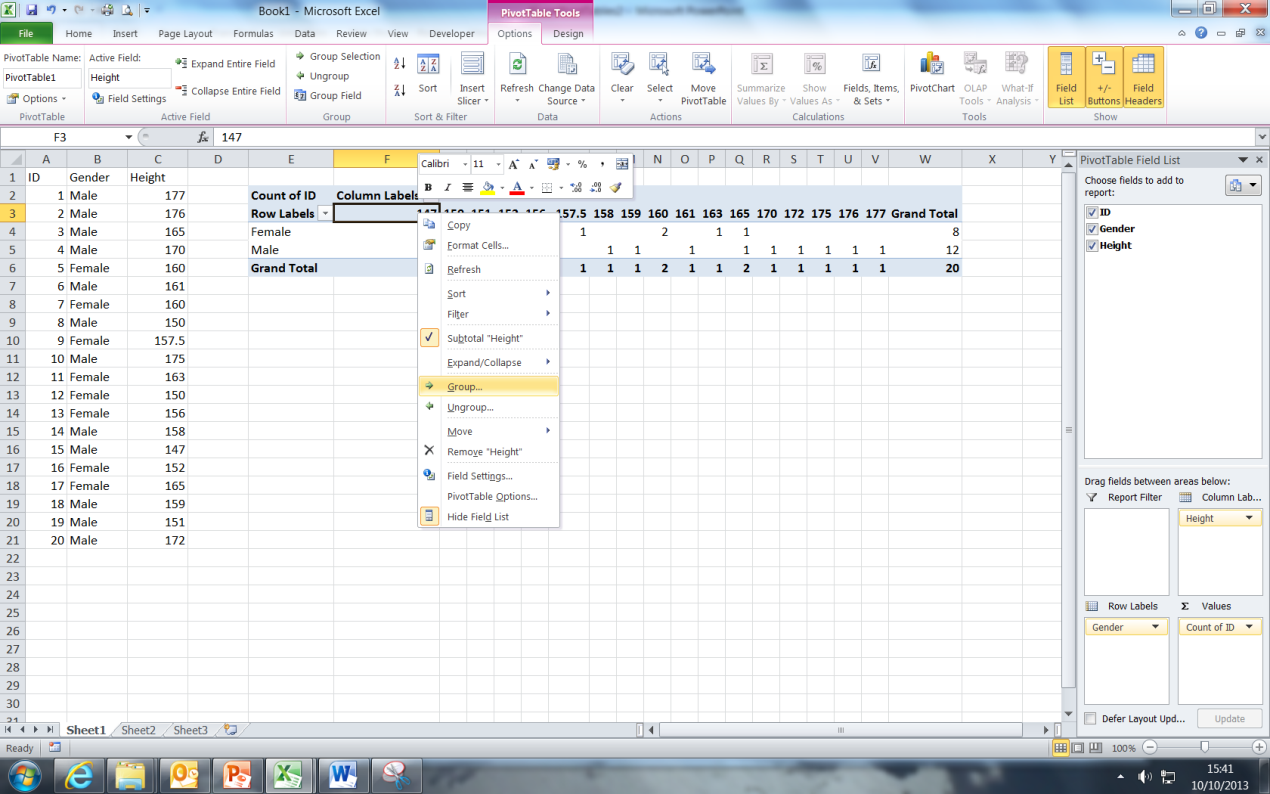
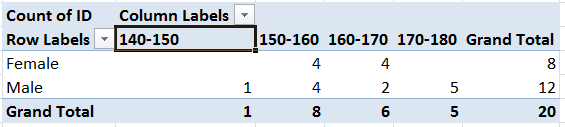


1. Drag the *Gender* field from the PivotTable Field List to the Row Labels list. Drag the *Height* field from the PivotTable Field List to the Column Labels list. This should create an empty PivotTable like the one shown below:

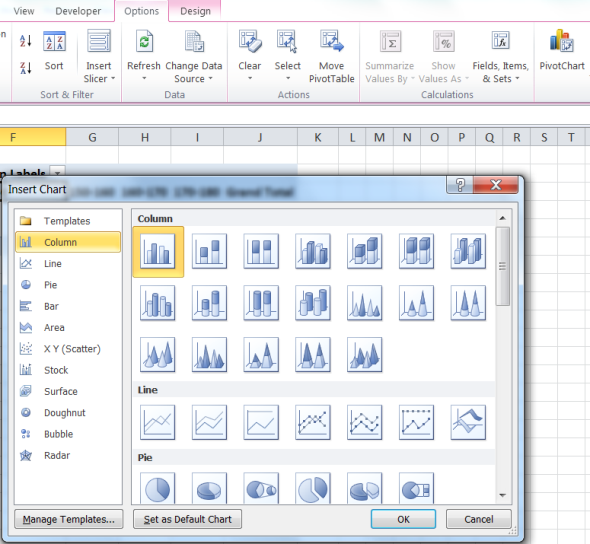


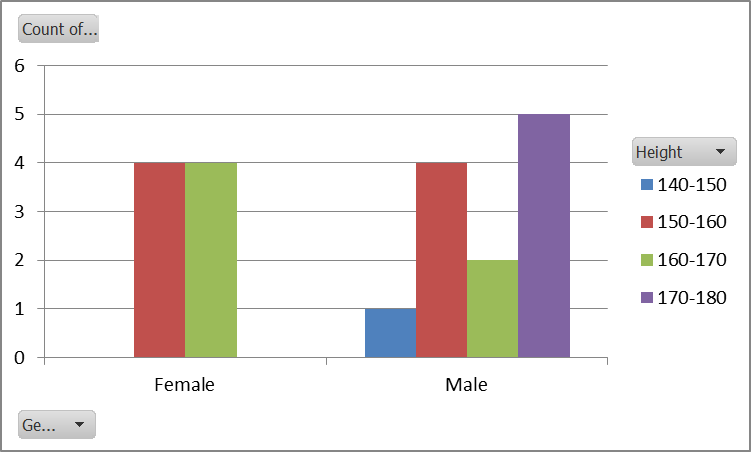
1. Drag the *ID* field from the PivotTable Field List to the Values list. In the Values box, select the triangle next to Sum of ID (ringed), then the option Value Field Settings as shown on the

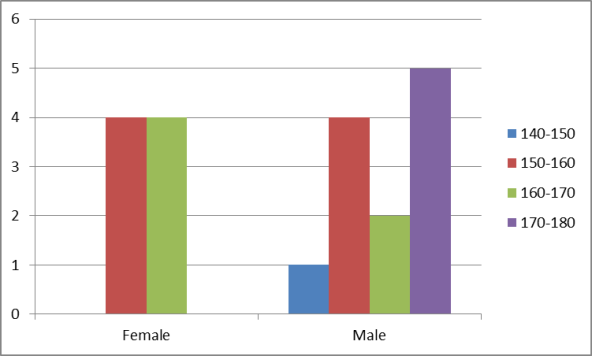
left, then select Count from the dialogue box as shown above. This should change the PivotTable to look like this:

1. 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:

**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.

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

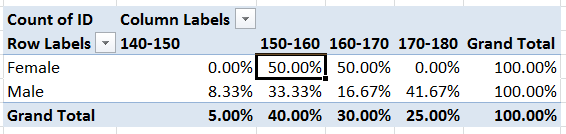
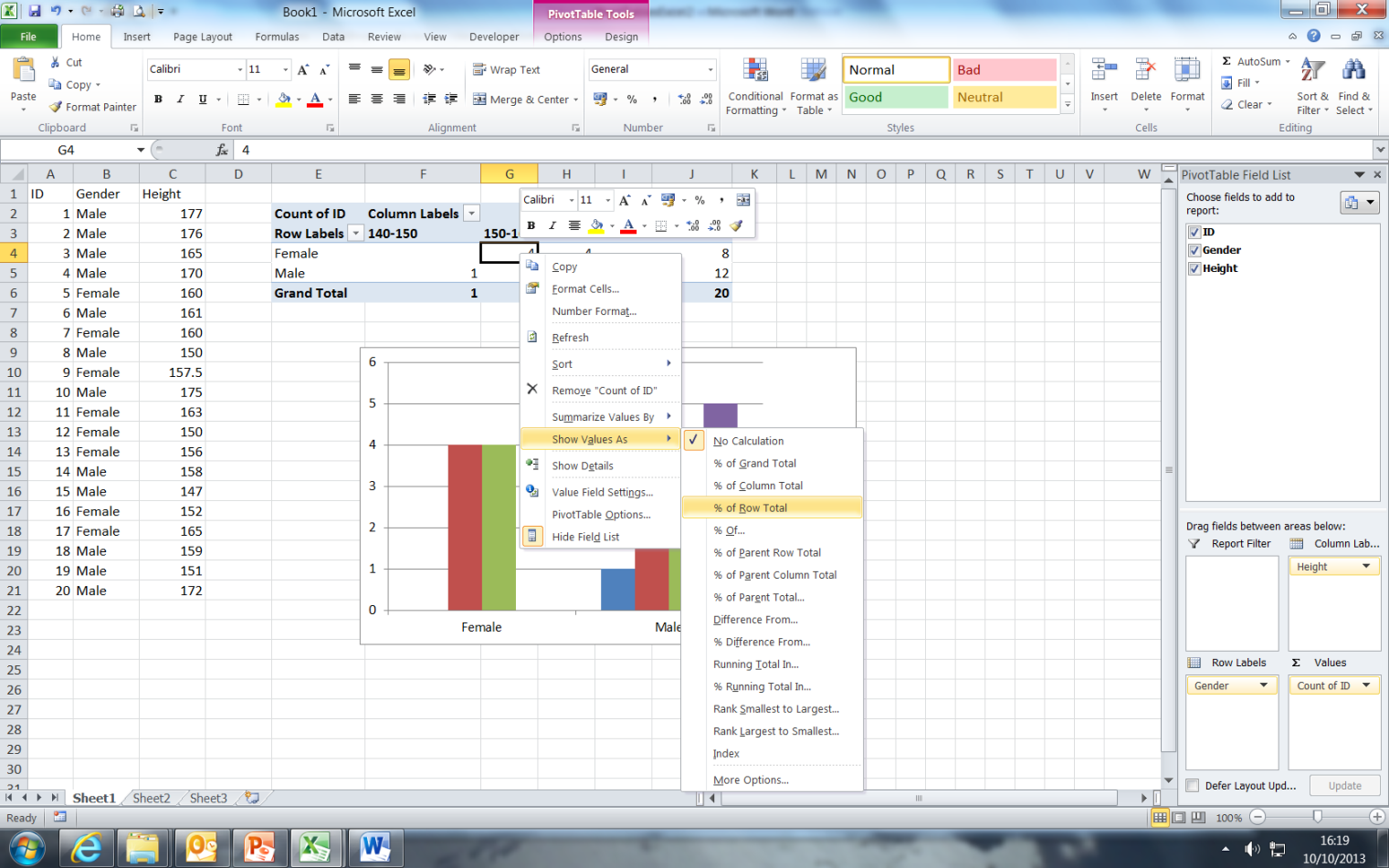
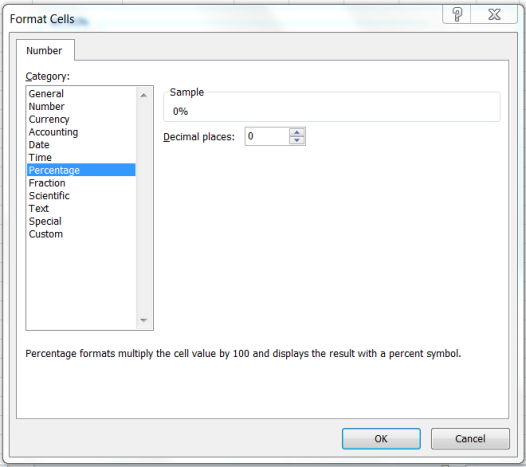


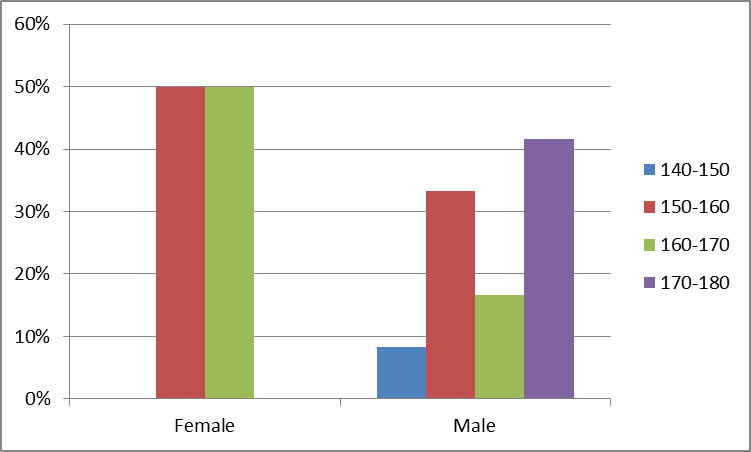


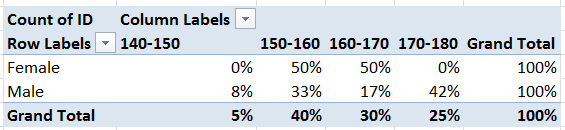
1. Right click on the *Height* Field Button then select Hide all Field Buttons on Chart option shown. This makes the chart look like this:

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

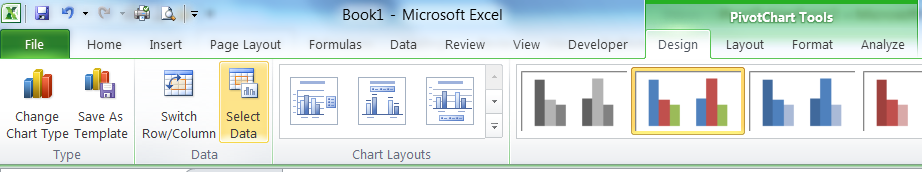
# Creating a Percentage Frequency PivotTable and Chart

1. 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:
2. 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:



## Note:

1. Percentage frequency tables and charts are useful when the group sizes are different
2. If there are several groups, column percentages may also be useful and can be calculated in a similar way
3. 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)