

## Task 2: Comparing the North Island sample with the South Island sample using boxplots (Instruction for Excel)

In this task we investigate the difference between the two samples, using boxplots.

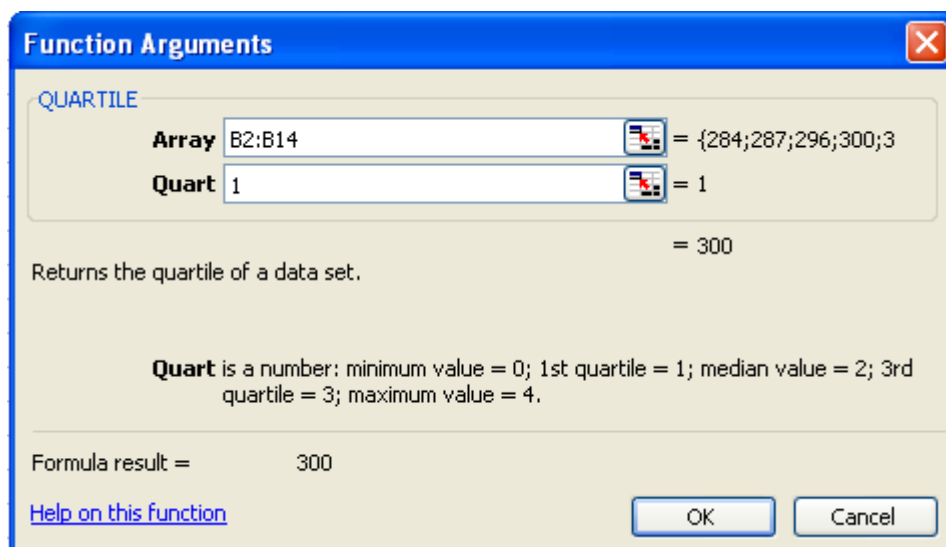
- (1) Insert a new worksheet and name it as “X1 Boxplots”.
- (2) Copy and paste the Island and X1 values from the “Dolphins data” worksheet to the “X1 Boxplots” worksheet.
- (3) Draw the following table in the “X1 Boxplots” worksheet.

Statistics	North	South
median		
Q1		
minimum		
maximum		
Q3		

- (4) Fill in the table using Excel’s built-in functions. Click on each cell in the table and choose the appropriate function in the **Statistical** category.

Use **MEDIAN** function to calculate the median values and **MIN** and **MAX** functions, the minimum and maximum respectively.

Use **QUARTILE** function for Q1 and Q3. In **QUARTILE** function, specify the values of one sample in the **Array** box and enter **1** in the **Quart** box to calculate the Q1 value. For the Q3 value, enter **3** in the **Quart** box instead.



You now should have the table shown on the following page.

	D	E	F
Statistics		North	South
median		304.00	282.50
Q1		300	276.25
minimum		284.00	251.00
maximum		319.00	304.00
Q3		313	287

- (5) Using this table, we now draw boxplots.

Select the whole table and click on the Chart Wizard icon in the Tool bar.

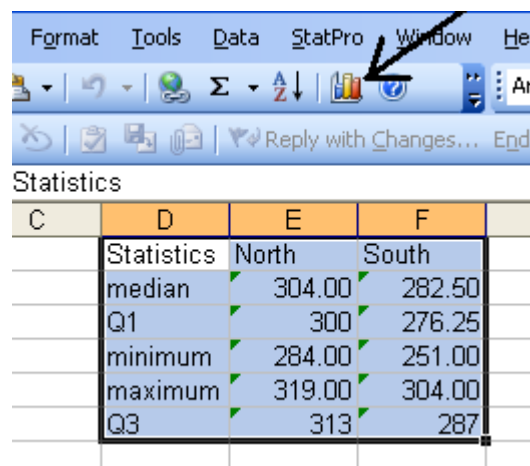
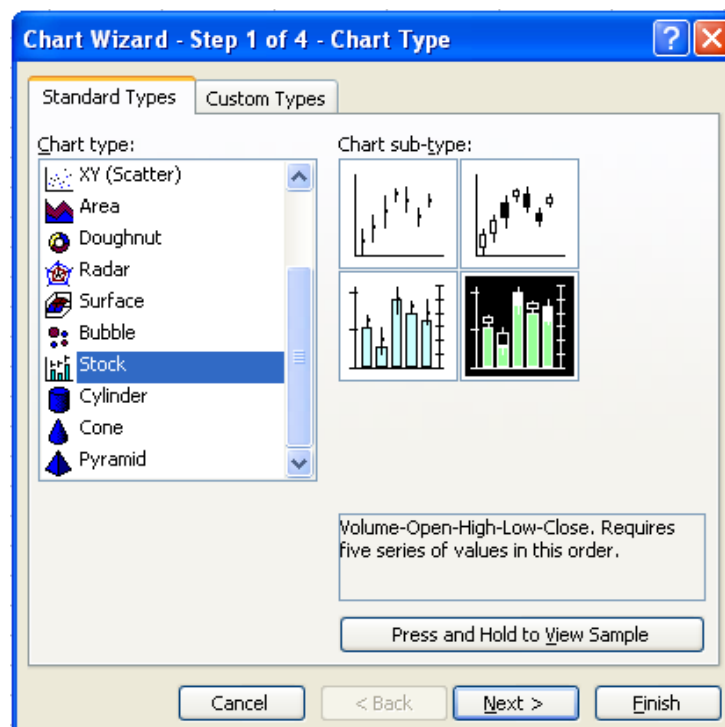
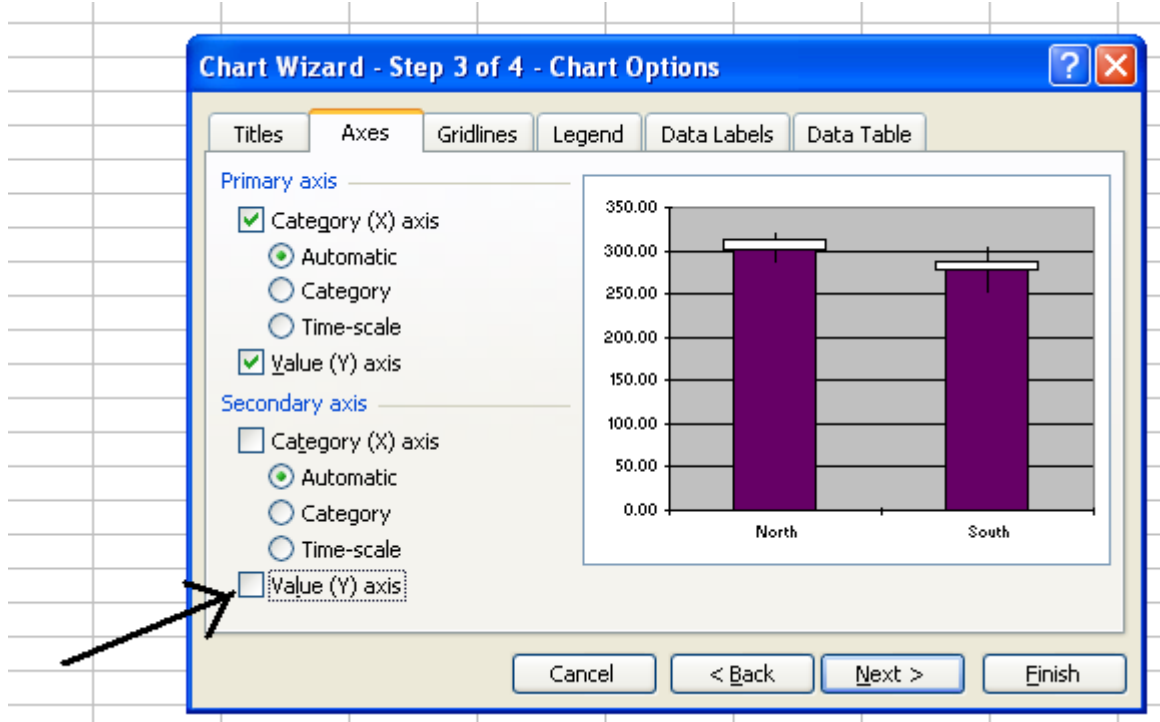


Chart Wizard dialog window should appear.

- (6) On the “Standard Types” tab, select **Stock** option as the Chart type, and then, click the fourth option in the Chart sub-type as shown below. Click **Next >**.

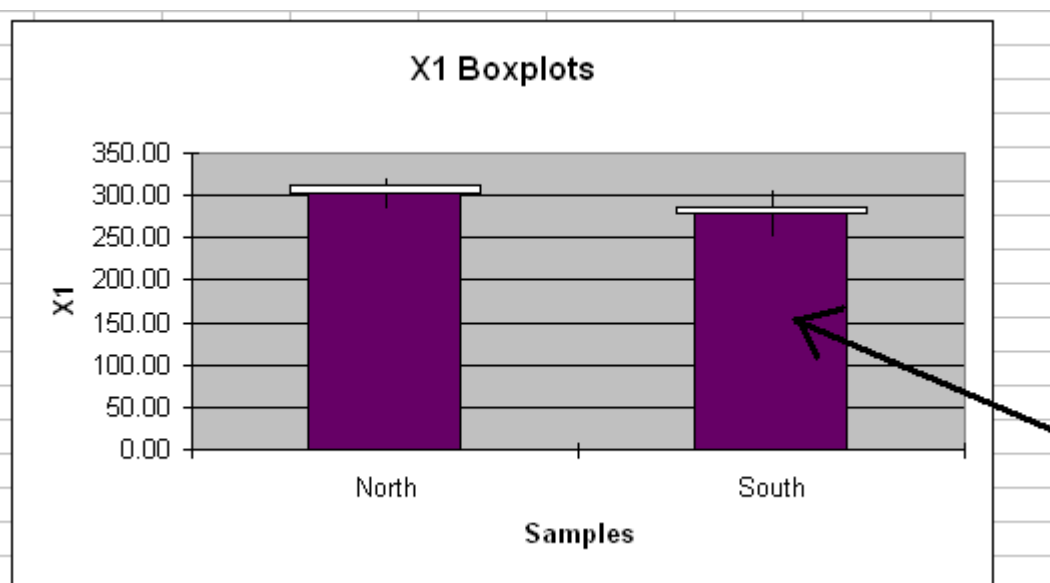


- (7) On the “Data Range” tab, click **Rows** under **Series** in, and then click **Next >**.
- (8) On the “Legend” tab, click to clear the **Show legend** check box.
- (9) On the “Axes” tab under **Secondary axis**, click to clear the **Value (Y) axis** check box.

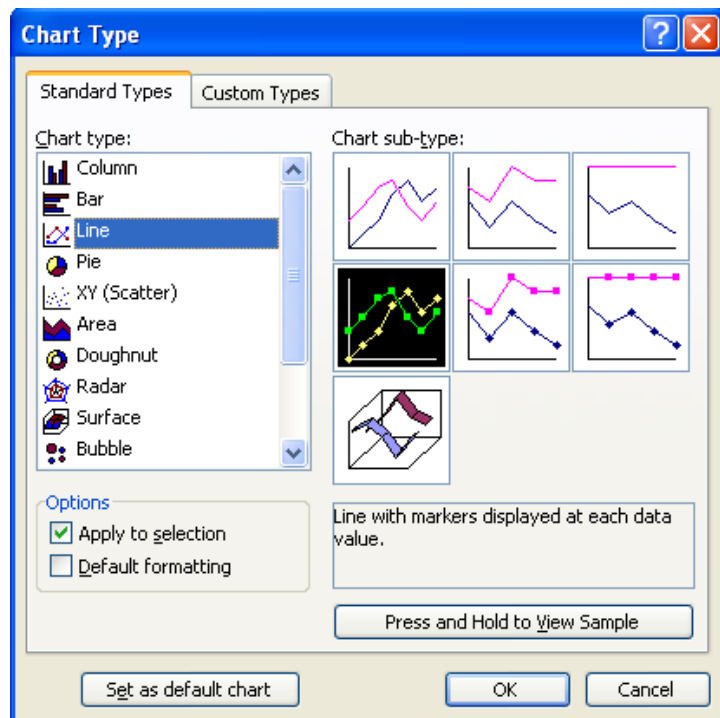


On the “Titles” tab, enter an appropriate title and axis labels. Then, click on **Finish**.

- (10) Click once on either of the coloured columns to select the “median” series.

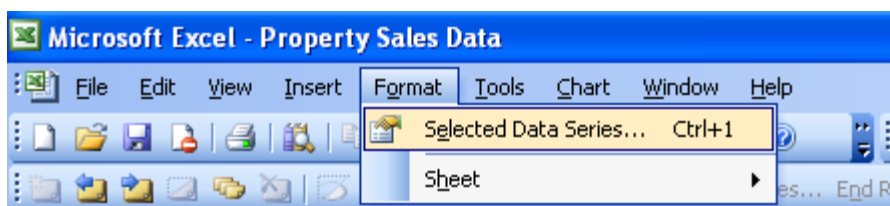
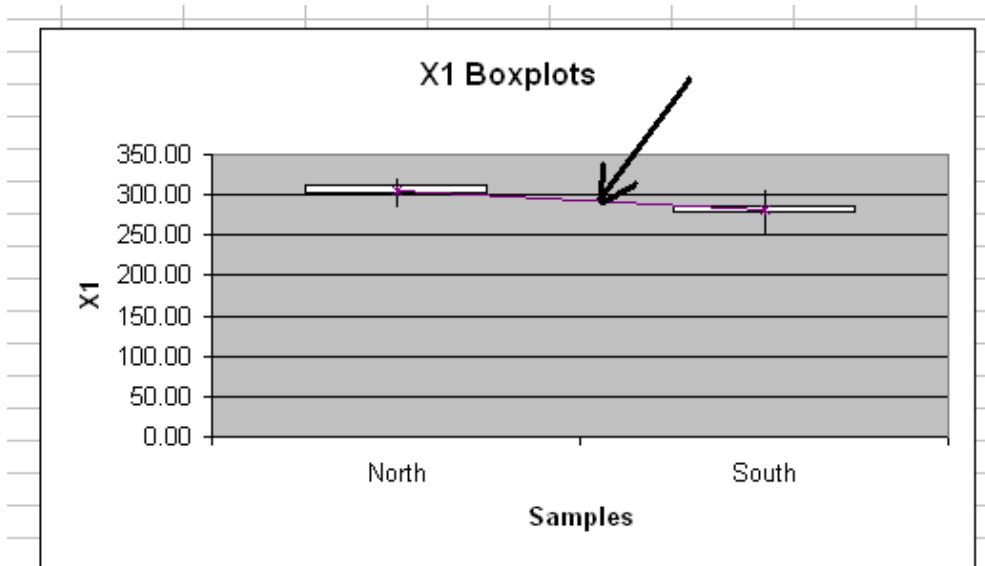


In the **Chart** menu, click on **Chart Type**, and then **Line**.

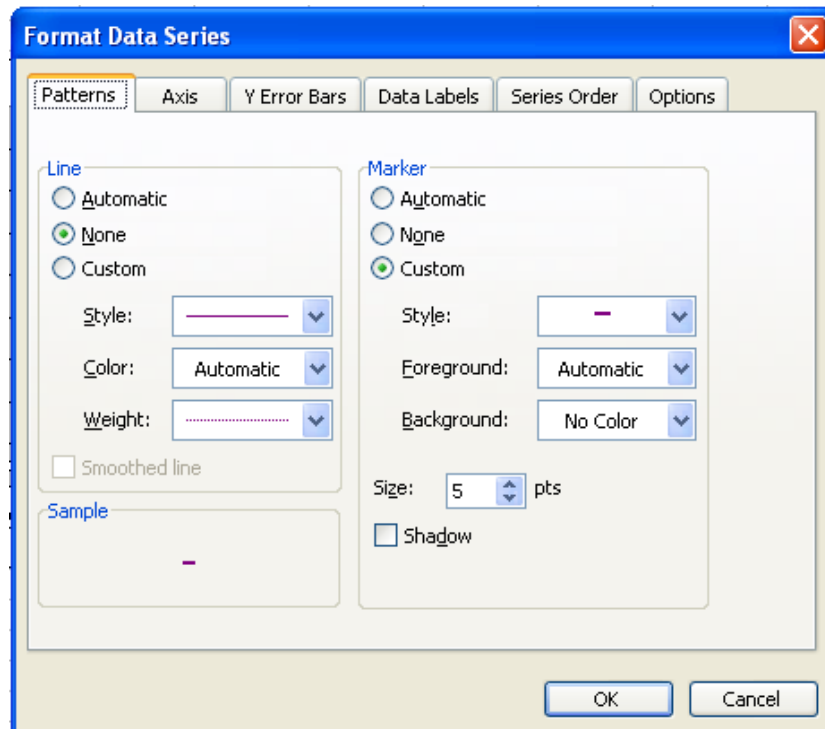


Click **OK**. A line that connects the median values should appear.

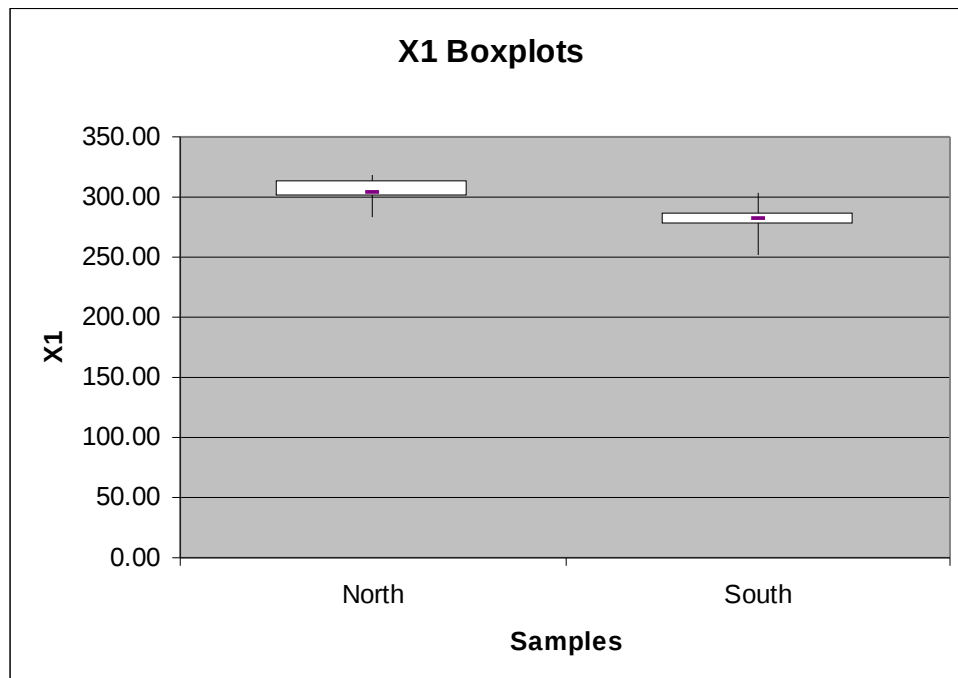
(11) Click once on the line. Then, on the **Format** menu, click **Selected Data Series**.



- (12) On the “Patterns” tab, under **Line** click on **None** and under **Marker**, **Custom**. Select “ – ” in **Style**. Then click **OK**.



We now have the following box plot.



As you have noticed, these boxplots are different from what you have learnt as boxplots. However, unfortunately, this is the best we can get from Excel, since Excel does not include a proper boxplot in its chart options.

(However, it is not an issue because many statistical software available outside school let you draw a proper boxplot by just clicking a button.)

(13) Repeat the steps (1) ~ (12) for X2, X3, X4, X5 and X6.

Congratulations! You have successfully completed this task.

**Do not forget to save all the worksheets you created, as we need to use them later for other tasks.**

**Name your worksheets appropriately, so that you can recognise them later.**