Topics from GCE AS and A Level Mathematics

* Interpret scatter diagrams and regression lines for bivariate data, including recognition of scatter diagrams which include distinct sections of the population
* Understand informal interpretation of correlation.

**Investigation 4 (a)**

Investigate if there is a correlation between daily mean total cloud cover and daily mean pressure.

**The data**

Open the Excel workbook **Pearson Edexcel GCE AS and AL Mathematics data set.xlsx.**

Select the **Information** worksheet.

1. Read the information in cell **A17.**

Explain how cloud cover is calculated.

1. Read the information in cell **A19.** What units are used to measure daily mean pressure?
2. What type of variable is daily mean pressure?
3. What type of variable is daily mean total cloud cover?

The location Leuchars 2015 was randomly selected from:

Camborne 2015 Camborne 1987

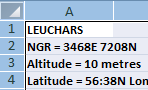
Hurn 2015 Hurn 1987

Leuchars 2015 Leuchars 1987

Leeming 2015 Leeming 1987

Heathrow 2015 Heathrow 1987.

*Select the worksheet* ***Leuchars May-Oct 2015****.*

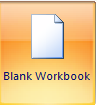
*It is difficult to analyse these data as it is presented in the dataset. The headers need to be in row 1.*

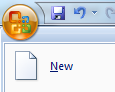
*Copy the data into a new workbook*

*Select the whole worksheet*

*Click on the small blue square in the left hand corner, this will select the whole worksheet.*

*Right click then* ***Copy***

*Open a new workbook*

*Select the* ***Office button*** *then* ***New***

*Then double click on* ***Blank Workbook***

*Select* ***A1*** *right click* ***Paste***

*Rename the worksheet*

*Double click on the tab* ***Sheet 1*** *at the bottom of the worksheet*

*Type* ***Leuchars 2015,*** *then click anywhere on the worksheet*

*Delete rows 1 – 5*

*Select rows 1 – 5 right click* ***Delete***

*Save workbook as* ***Leuchars2015***

1. What are bivariate data?

In bivariate data if one of the variables is controlled (or explains the other variable), it is known as the independent (or explanatory) variable.

A dependent (or response) variable is a variable whose value depends on the value of another variable.

The dependent variable is usually plotted on the vertical axis. Note: if you are using a regression model to predict a value, the variable for the value you wish to predict should be the *Y* variable and plotted on the vertical axis of a scatter diagram.

For bivariate data if all the points in a scatter diagram seem to lie near a straight line, there is a linear correlation between the two variables.

Pearson’s product moment correlation coefficient provides a standardised measure of linear correlation. Its value lies within −1 and +1.

**Process**

Plot a scatter diagram to investigate if there is a correlation between daily mean total cloud cover and daily mean pressure. In this case there is no controlled variable. Later in this investigation a prediction for daily mean total cloud cover will be calculated based on the regression equation for daily mean total cloud cover against daily mean pressure.

1. Which variable should be plotted on the vertical axis?

*Plot a scatter diagram in Excel*

*Open* ***Leuchars2015.xlsx’****.*

*When plotting a scatter diagram, Excel plots the variable in left hand column on the x-axis and the variable in the right hand column on the y-axis.*

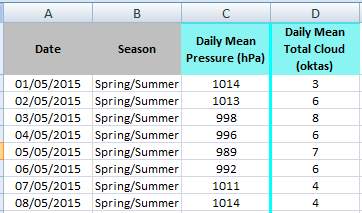
*Copy columns* ***A****,* ***I*** *and* ***K*** *into a new worksheet*

*In the worksheet* ***Leuchars2015*** *select column* ***A*** *right click* ***Copy***

*In worksheet* ***Sheet 2*** *select* ***A1*** *right click* ***Paste***

*In the worksheet* ***Leuchars2015*** *select column* ***K*** *right click* ***Copy***

*In worksheet* ***Sheet 2*** *select* ***C1*** *right click* ***Paste***

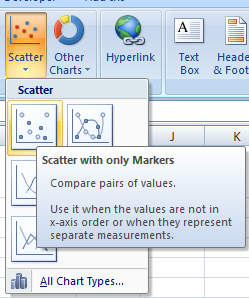
*In the worksheet* ***Leuchars2015*** *select column* ***I*** *right click* ***Copy***

*In worksheet* ***Sheet 2*** *select* ***D1*** *right click* ***Paste***

*Rename* ***Sheet 2 CloudPress***

*Select* ***B1*** *type* ***Season***

*In column* ***B*** *code the data for May, June, July and August as* ***Spring/Summer*** *and code the values for September and October* ***Autumn****. These codes will be used later.*

*Plot a scatter diagram*

*In the worksheet* ***CloudPress*** *select columns* ***C*** *and* ***D***

*Select the* ***Insert*** *tab then* ***Scatter****.*

*Select* ***Scatter with only Markers.***

*Format the diagram*

**

*Delete the legend*

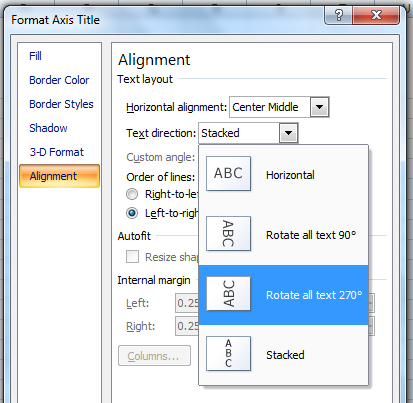
*Click on the* ***legend*** *and* ***delete***

*Add a title*

*Click on the title and type* ***Daily mean total cloud cover vs Daily mean pressure Leuchars 2015*** *then* ***Enter***

*Add a vertical axis title*

*Click on the chart then select the* ***Layout*** *tab select* ***Axis Titles*** *then* ***Primary Vertical Axis Title*** *then* ***Vertical Title*** *and type* ***Daily mean total cloud cover (oktas)*** *and* ***Enter***

**

*Change the alignment of the text in the vertical title*

*Right click on the* ***vertical title*** *select* ***Format Axis Title*** *then* ***Alignment*** *then click on the arrow next to* ***Text direction*** *click* ***Rotate all text 270o*** *and* ***Close***

*Add a horizontal axis title*

*In the* ***Layout*** *tab select* ***Axis Titles*** *then* ***Primary Horizontal Axis Title*** *then* ***Title Below Axis*** *and type* ***Daily mean pressure (hPa)*** *and* ***Enter.***

*Add a chart border*

*Right click on the* ***chart*** *select* ***Format Plot Area*** *then* ***Border Color*** *select* ***Solid line*** *open up the* ***Color*** *arrow select* ***Black*** *and* ***Close***

***Save your work***



Figure 1

1. Comment on the correlation between the two variables.
2. Give an interpretation of the correlation between the two variables.
3. Comment on the scatter.

*Find the product moment correlation coefficient in Excel*

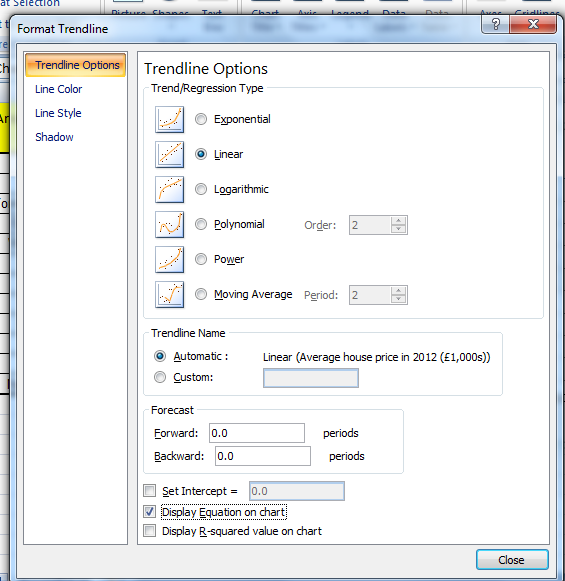
*Select* ***E2*** *type =corr*

**

**

*Select CORREL then column* ***C****, column* ***D Enter***

Correlation coefficient =

*To add a line of regression in Excel (This is called a trend line in Excel)*

*Right click on any* ***data point*** *select* ***Add Trendline***

*Select* ***Linear*** *then* ***Display Equation on Chart*** *and* ***Close***

*Drag the equation of regression (trend line) so it is clearly visible on the chart, reduce the number of decimal place values to two and format the text using the formatting in the* ***Home*** *tab.*



1. Write the line of regression using the names of the variables.
2. Interpret the gradient of the line of regression for daily mean total cloud cover against daily mean pressure.
3. Does this regression model seem to fit the data?
4. Use the regression model (equation of regression) to predict the daily mean total cloud cover for a daily mean pressure of 1030 hPa.
5. Comment on the accuracy of the prediction in question 13.

**Report**

**Investigation 4 (b)**

Plot a scatter diagram to investigate the relationship between daily mean total cloud cover against daily mean pressure for Leuchars 2015, split by season.

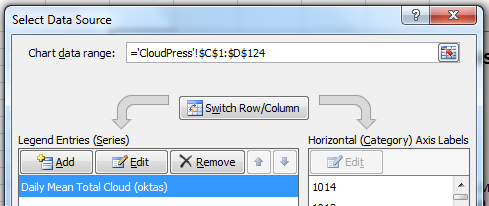
*To plot the scatter diagram daily mean total cloud cover against daily mean pressure for Leuchars 2015, split by season*

*In the worksheet* ***CloudPress*** *select* ***E1*** *type* ***Spring/Summer Enter***

*Select* ***F1*** *type* **Autumn Enter** (these will be used as series names)

*Plot a scatter diagram for Spring/Summer only, cells* ***C1:D124****.*

*Move the chart to the top of the worksheet. Either drag or* ***Cut*** *and* ***Paste.***

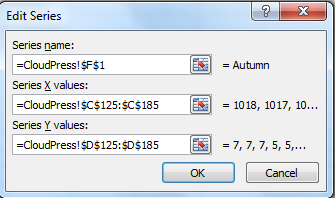
**

*Add a label for Spring/Summer*

*Right click on the* ***chart*** *select* ***Select Data***

*Select* ***Edit***

***Series name*** *select* ***E1 OK OK***

*Add Autumn data and label*

*Right click on the* ***chart*** *select* ***Select Data***

*Select* ***Add***

***Series name*** *select* ***F1***

***Series X values*** *select* ***C125:C185***

***Series Y values*** *select* ***D125:D185****(delete ={y})*

***OK OK***

*Format the chart by adding a title and labels.*



**Report**

1. Comment of the split of the data between Spring/Summer and Autumn.

**Investigation 4 (c)**

Investigate if there is a correlation between daily mean visibility and daily mean temperature. Use the same random sample i.e. Leuchars 2015.

**The data**

In the Excel workbook **Pearson Edexcel GCE AS and AL Mathematics data set.xlsx.**

Select the **Information** worksheet.

1. Read the information in cell **A12.**

Explain how daily mean temperature is measured.

1. Read the information in cell **A18.**

Define daily mean visibility.

1. What type of variable is daily mean temperature?

**Process**

Plot a scatter diagram of daily mean visibility against daily mean temperature.

1. Which variable should be plotted on the vertical axis?
2. Comment on the correlation between the two variables.
3. Explain why no line of regression should be calculated or fitted to the data.

**Investigation 4 (d)**

Investigate if there is a correlation between daily mean air temperature and daily mean pressure in Beijing May to October 2015.

**The data**

The data are provided in the Excel workbook **Edexceldataset.xlsx.**

Copy the data for Beijing 2015 into a new workbook

Rename the workbook **Beijing2015**

Delete rows 1 – 5

Save workbook as **Beijing2015**

**Process**

Plot a scatter diagram to investigate if there is a correlation between daily mean air temperature and daily mean pressure in Beijing, May to October 2015. Daily mean pressure is the explanatory variable.

1. Which variable should be plotted on the vertical axis?
2. Comment on the correlation between the two variables.
3. Give an interpretation of the correlation between the two variables.
4. Find the product moment correlation coefficient in Excel.

Correlation coefficient =

1. Add a line of regression to the plot.
2. Write the line of regression using the names of the variables.
3. Interpret the gradient of the line of regression for daily mean air temperature against daily mean pressure.

.

1. Interpret the intercept of the line of regression for daily mean air temperature against daily mean pressure.
2. Does this regression model seem to fit the data? Give a reason for your answer.

1. Use the regression model (line of regression) to predict the daily mean air temperature for a daily mean pressure of 1005 hPa.
2. Comment on the accuracy of the predicted daily mean air temperature in question 10.

1. Plot scatter diagram to show the relationship between daily mean air temperature against daily mean pressure for Beijing, May to October 2015, split by season.

**Report**