

ASSOCIATIONS BETWEEN TWO NUMERICAL VARIABLES

This is where we use a Scatterplot.

The table below shows the Hours Worked against the % of population going to University for 9 countries.

% University	26	20	36	1	25	9	30	3	55
Hours Worked	35	43	38	50	40	50	40	53	35

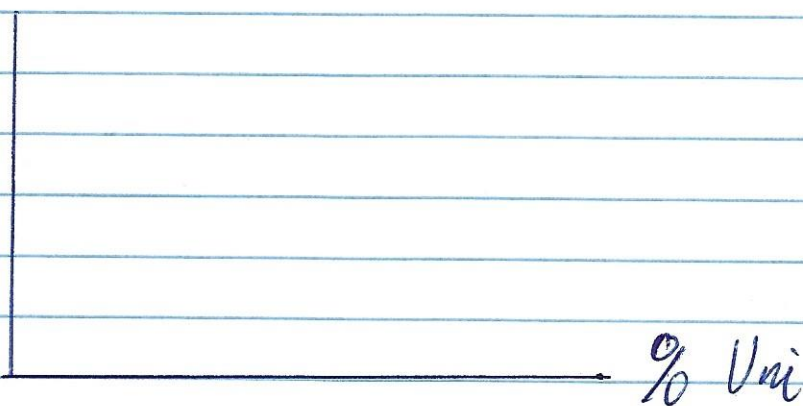
Here the EV is % Uni and the RV is Hours Worked. i.e. Does the % at Uni associate with hours worked.

It is usual that the EV is the first row in the table and RV the second.

When constructing the scatterplot:

The EV goes on the Horizontal axis (x-axis)
The RV goes on the Vertical axis (y-axis)

Hours
Worked.

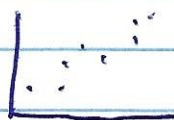


Handout: ~~Which Graph to Use~~
Scatterplot on the classpad.
- uses data from example above.

Interpreting Scatter Plots.

When we look at Scatter Plots we are interested in:

* Direction

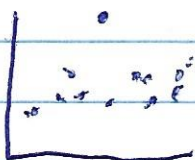


Positive

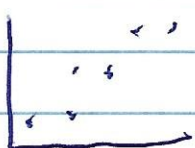


Negative

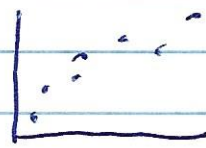
* Outliers



* Form



Linear



Non-Linear

* Strength

- Subjective: Strong, Moderate, Weak
- Pearson's Correlation Coefficient (r).

Handout: Scatterplot Patterns + Pearson

Note: The Pearson Correlation Coefficient works on the assumptions:

- * the variables are numeric
- * the association is linear
- * there are no outliers

If the data does not fit these requirements the coefficient can be misleading.

Questions: Ex 3D & 3E Questions on Workplan.