Question 1

A researcher organizes an experiment designed to test the hypothesis that students have more incidental learning if they pay attention to the meaning of something as compared to if they pay attention to its surface features. To test this the researcher gives students a list of words and asks them to cross out vowels (i.e. to pay attention only to the surface features of the words). After a period of time in which the students complete a separate task, they are asked to recall as many words from the list as they can. The next week the researcher gives the same students another list of words and asks them to identify if the each word is associated with the concept of 'living'. After a period of time in which the students complete a separate task, they are asked to recall as many words from the list as they can.

Each student's score on the two tasks is presented in the table below.

Student number	Vowels task	Meaning of life task
1	1	7
2	0	6
3	2	8
4	3	5
5	4	6
6	0	4
7	0	8
8	2	3
9	1	5
10	2	9
11	4	4
12	3	7
13	2	7
14	1	5
15	6	11
16	3	7
17	2	5
18	1	8
19	4	3
20	3	4
21	2	7
22	5	8
23	1	8
24	0	9
25	0	5
26	2	3
27	1	5
28	1	6
29	2	7
30	0	2

(a) Scatter plots are used to represent the relationship between two continuous variables. Represent this data on a scatter plot comparing each student's performance on each test. Make sure to include the features that are required in all charts (number title, axis titles etc.)

(b) Numerically, the relationship between two continuous variables can be described using the correlation co-efficient (which can be calculated used the CORREL function in Excel). What is the correlation between these two variables? How would you interpret this figure (is the correlation positive or negative? Is it weak, moderate or strong?).

(c) Correlation, (like average mean) can be distorted by outlier cases. Using the scatterplot, identify if there are any outlier (i.e. unusual) cases here. If so, remove the outlier and calculate the correlation between the two variables. Interpret your finding.

Question 2

Ignoring for a minute the question as to whether or not the distribution of scores is normal, which of the following tests would the researcher use to test the hypothesis outlined in Q.1:

- one-sample t-test
- two-sample t-test for matched pairs
- two-sample t-test with equal variance assumed
- two-sample t-test with equal variance not assumed.

State the null hypothesis and then test it using the appropriate measure, and interpret the results.

Question 3

Last week in class we conducted an experiment to test the hypothesis that students will learn more if they are interested by the subject. To test this we collected data on prior interest in sport and in history and then taught a lesson on nationalism and sport in late 19th century Ireland. After a delay in which the groups were distracted by another task, the whole class was tested on their recall of material.

The database of results is on moodle. The level of interest in sport is identified in column 4. The score is in column 5.

As we noted last week, it is often problematic if we have a large number of categories as compared to our total sample size. That is the case here so it makes sense to recode the data: Recode level of interest in sport with "low interest" being represented by a score of 1, 2 and 3 and "high interest" being represented by a score of 4 and 5).

A box plot can be used to visualize the relationship between a categorical variable (like interest) and a continuous variable (like score). Draw a horizontal boxplot which allows the results of the low and high interest groups to be compared. Ensure to label axes and the chart appropriately as always.

Question 4

Ignoring for a minute the question as to whether or not the distribution of scores is normal, which of the following tests would the researcher use to test the hypothesis outlined in Q.3:

- one-sample t-test
- two-sample t-test for matched pairs
- two-sample t-test with equal variance assumed
- two-sample t-test with equal variance not assumed.

State the null hypothesis and then test it hypothesis using the appropriate measure. Interpret the results.