Statistical Tests on Data

This activity will help you to:

- ❑ Understand why psychologists use statistical tests
- ❑ Understand the significance of the null hypothesis
- ❑ Know what a test statistic and a critical value are
- ❑ Find critical values from a table
- ❑ Use test statistics and critical values to make judgments about data

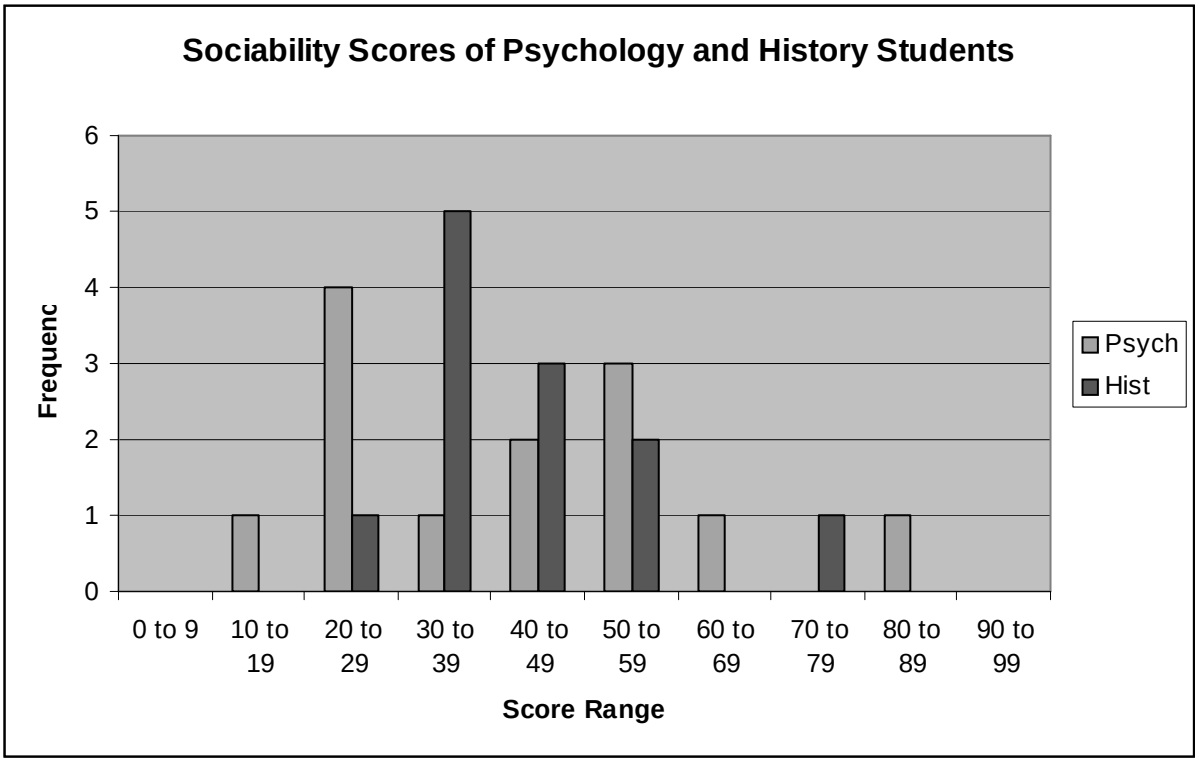
Statistical tests are used to decide whether any pattern found in a set of data is **significant** or whether, on the other hand, it was likely to be **caused by chance**.

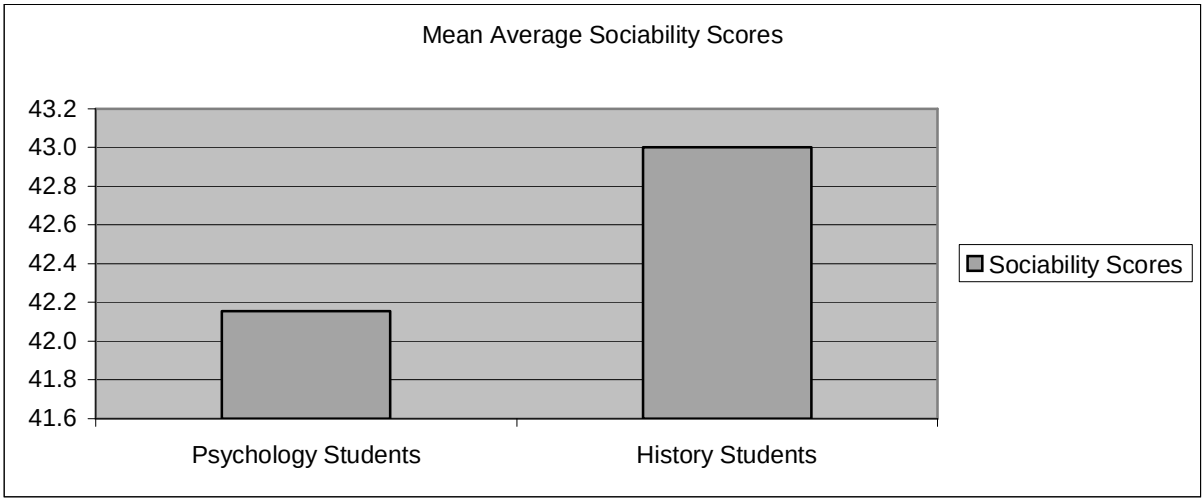
For example, suppose we suspected that Psychology students were more sociable than History students. We could devise a questionnaire to measure how sociable people are, and give it to a sample of Psychology students and a sample of History students. Here are some imaginary data for such a survey:

Sociability Scores	
Psychology students	**History students**
22	35
42	43
54	57
12	23
63	35
25	41
35	38
49	76
86	43
56	54
24	36
54	32
26	46

Have a look at these data and note below whether they seem to suggest that there is a difference in the sociability scores of historians and psychologists.

Data in the form of **raw scores** are very difficult to make sense of. When analysing data, the first step is to put them on a graph or chart.





What do these graphs seem to tell us?

The graphs seem to indicate that, on average, History students are more sociable than Psychology students. The question is, is this difference between the groups important? We have two possibilities here:

Possibility 1.

There is something about being a History student that makes you more sociable than a Psychology student.

OR

The difference between sociability scores of Psychology and History students is **not** due to chance

Possibility 2.

There is no real difference between History and Psychology students. The difference in our data occurred because of the way any two samples of people will be slightly different.

OR

The difference between sociability scores of Psychology and History students is **due to chance**.

We can use a statistical test to work out how likely it is that the difference is due to chance. But first, we must decide the point at which a result due to chance is so unlikely that the difference must be **significant.**

Here are some possibilities:

% Chance	Probability (decimal)	Explanation
50%	P=0.5	A difference that has a 1 in 2 probability of being caused by chance.
20%	P=0.2	A difference that has a 1 in 5 probability of being caused by chance
5%	P=0.05	A difference that has a 1 in 20 probability of being caused by chance
1%	P=0.01	A difference that has a 1 in 100 probability of being caused by chance
0.1%	P=0.01	A difference that has a 1 in 1000 probability of being caused by chance

Whichever value we choose is called our **minimum acceptable probability for significance** or **P-value**.

Before choosing a P-value, we need to be aware of one more thing:

- When the P-value is very large there is a danger that we will take a chance result and decide it is significant when it is not (a **type 1 error**)
- When the P-value is very small, there is a danger that we will take a significant result and decide that it was caused by chance (a **type 2 error**).

Which P-value do you think we should choose?

Once we have decided what our P-value is going to be, we need to do some calculations on the data (more of this next week). All you need to know for now is that the result of these calculations is called the **test statistic**.

The test statistic is used to tell us whether the difference between our samples is likely to have been caused by chance. In order to work this out, we need something else called a **critical value** against which we compare our test statistic.

We get the critical value from a table like this:

	Minimum Acceptable Probability for Significance				
Sample Size	**P<0.5**	**P<0.2**	**P<0.05**	**P<0.01**	**P<0.001**
5	78	90	102	114	126
6	72	84	90	102	114
7	66	78	84	90	102
8	60	72	78	84	90
9	54	66	72	78	84
10	48	60	66	72	78
11	42	54	60	66	72
12	36	48	54	60	66
13	30	42	48	54	60
14	24	36	42	48	54
15	18	30	36	42	48
16	12	24	30	36	42
17	6	18	24	30	36

For significance at the chosen level, the test statistic must equal or exceed the critical value from the table.

To obtain our critical value, we need to know:

1. The size of our samples of data
2. Our minimum acceptable P-value

What is our critical value for this investigation?

Then we need to compare the critical value with our test statistic. The statistical analysis of our data gave a test statistic of **35**. We make the comparison according to the rule at the bottom of the table.

- ❑ If the result is **significant**, we **reject** the null hypothesis – we say that the difference was **not** due to chance.
- ❑ If the result is **not significant**, we **accept** the null hypothesis – we say that the difference was due to chance.

Is the difference between our two samples significant?

Do we accept or reject the null hypothesis?

Testing for Significance

Use the table on the handout to test the significance of the following findings, then write an answer explaining whether the result is significant or not. In writing each answer…

You must:

- ❑ State whether or not the null hypothesis (i.e. the prediction that the difference will be due to chance) should be **accepted** or **rejected**.

You should:

- ❑ Explain **why**, by referring to the critical value and the test statistic.

You could:

- ❑ Comment on whether the research is likely to have made a **type 1** or a **type 2 error**.

1. A researcher found that the maths test scores of 16 students who played a musical instrument were higher, on average, than the scores of 16 who didn't. The test statistic for the difference was 38, and her minimum acceptable P-value was 0.05.

2. A researcher found that the average emotional stability scores of 10 Psychology graduates were lower than those of 10 Physics graduates. His minimum acceptable P-value was 0.2 and his test statistic was 63.

3. Using a minimum acceptable P-value of 0.01, a researcher tested the significance of a difference between the motivation scores of 8 schoolteachers and college lecturers. The test statistic was calculated to be 74.

4. A researcher compared the reading ability scores of two classes of 17 8 year-olds. Using a minimum acceptable P-value of 0.001, she calculated a test statistic of 31.

5. A researcher compared self-esteem scores in 13 students who had attended a leadership training course and 13 who hadn't. She used a minimum acceptable P value of 0.2 and calculated her test statistic to be 44.

And if you finish those…

Look at the table on the handout and describe:

1. The relationship between sample size and obtaining a statistically significant result.
2. The relationship between chosen P-value and obtaining a statistically significant result.