BIVARIATE ANALYSIS AND DATA PRESENTATION

THE WARM-UP

Imagine a teacher has given a quiz, and one student receives a mark of 9. Is 9 a good score on the quiz? What other information would help us put that number into context? Did some students do very well and others struggle, or were the scores all fairly similar?

Now, imagine the test is given in the morning, and each student records how many hours they slept the night before. How do the test results compare for students and their amount of sleep?

In next week's class, we'll consider ways to present this information in ways that will help to answer these questions and help us to draw conclusions about the results. Before then, let's review some of the tools we'll use to do that.

Let's start with the data. We can look at the test scores alone. The set of scores is **univariate** data – or one variable. The test scores stand alone:

Score	6	3	9	11	2	15	5	3	8

Pairing the scores with the hours of sleep for each, we get **bivariate** data – or two variables:

Score	6	3	9	11	2	15	5	3	8
Hrs Sleep	7	4	7.5	6	6.5	7.5	6	4	5.5

We'll examine only the univariate data in this exercise.

One clue about the student's score of 9 is the total possible marks on the test. 9 out of 10, and 9 out of 20, imply different success. But what if 9 of 20 is the highest score in the class? Or if 9 of 10 is the lowest? Where does 9 fall within the class? Are the scores closely grouped, or are they spread out?

The first two numbers we'll use to investigate these questions are the **minimum** and the **maximum**, the lowest and highest values in the group.

Next, we'll need to mark the centre of the group. You may recall discussing *mean*, *median,* and *mode* as useful ways of doing this. *Mean* is the average value, and *mode* is the most common value in the group. But for the charts next week, we will use the **median**.

The median is simply the middle value in order of size; it has an equal number of values above it and below it. If there are an odd number of values in the group, then the median is the middle value. If the group has an even number of values, then the median falls between the two middle values. In this case, calculate the median as the *mean* of those two middle values.

Two examples are on the next page:

Example 1		Example 2	
Values	6 3 9 11 2 15 5 3 8	Values	4 1 12 10 7 5 9 6
Group size	9	Group size	8
Half of group size	4, rounded from 4.5	Half of group size	4
Values in order	| 2 3 3 5 | 6 | 8 9 11 15 | 4 values 4 values	Values in order	| 1 4 5 6 | | 7 9 10 12 | 4 values 4 values
Minimum	2	Minimum	1
Maximum	15	Maximum	12
Median	6	Median	(6 + 7) / 2 = 6.5

Now, the groups have been divided into halves. For one chart, we'll divide each half again, using the same process. The result will give us four **quartiles** of the group. Here is the result for Example 1:

Example 1, part a		Example 1, part b	
Values	2 3 3 5	Values	8 9 11 15
Group size	4	Group size	4
Half of group size	2	Half of group size	2
Values in order	| 2 3 | | 3 5 |	Values in order	| 8 9 | | 11 15 |
Median	(3 + 3) / 2 = 3	Median	(9 + 11) / 2 = 10

The median in part a, or the top of the first **quartile** of the group is called **Q1**. The original median, from above, is the top of the first half, which is made up of the bottom two quartiles. It is also called **Q2**.

The median in part b, then, is **Q3**, the top of the third quartile. The top of the fourth quartile is the **maximum**, already discussed.

These five numbers – the **minimum**, **Q1**, **Q2**, **Q3**, and **maximum** – make up the **five-number summary** of the data set. Next week, we will look at how to use these numbers, and other techniques, to make graphs that will help analyse and communicate the results.

As an exercise, you're welcome to complete the analysis for Example 2, and to complete a third example from the beginning. Exercises are on the following page, and answers are provided on a separate sheet.

EXERCISES

Example 2, completion:

Example 2	
Values	4 1 12 10 7 5 9 6
Group size	8
Half of group size	4
Values in order	| 1 4 5 6 | | 7 9 10 12 | 4 values 4 values
Minimum	1
Maximum	12
Median	(6 + 7) / 2 = 6.5

Example 2, part a		Example 2, part b	
Values		Values	
Group size		Group size	
Half of group size		Half of group size	
Median		Median	

Five-number group: Min = Q1 = Q2 = Q3 = Max =

Example 3:

Find the five-number group for the following group of values:

Example 3	
Values	15, 2, 56, 23, 16, 45, 30, 36, 22, 41, 8, 53, 27, 44
Group size	
Half of group size	
Values in order	
Minimum	
Maximum	
Median	

Example 3, part a		Example 3, part b	
Values		Values	
Group size		Group size	
Half of group size		Half of group size	
Median		Median	

Five-number group:

Min = Q1 = Q2 = Q3 = Max =