

Introduction to Junior Statistics

Data and the PPDAC cycle

Name: _____

Created by Elizabeth Sneddon

Data

What does data look like?

We need to think about how we collect the data, what question we ask, what answers people might give, and how to record the answers.

We also need to think about how many people to ask. This is called our **sample size**.

There are several ways to collect data:

- By carrying out a survey
 - This is when you ask people some questions.
- By doing an experiment
 - This is when you test or measure people or objects, like in science.
- By taking a census.
 - This is where you ask every person in the country a bunch of questions.

Survey Exercise:

Ask 4 people the following questions, and record the information in the table below.

- 1) How old are you?
- 2) Are you male or female?
- 3) What did you eat for dinner last night?
- 4) How many songs have you listened to today?

Student	Age	Gender	Dinner	Songs
1				
2				
3				
4				

Asking students questions like this is called a survey.

Describe what a survey is, in your own words.

A survey is _____

Experiment Exercise:

Take the measurements below from 4 people, and record the information in the table below.

Student	Length of thumb	Length of little finger
1		
2		
3		
4		

Measuring students like this is called an experiment.

Describe what an experiment is, in your own words.

An experiment is _____

Census Exercise:

A census is where we collect information from every person living in New Zealand. We do this every 4 years.

The government gets information such as:

- How many young people we have in NZ
- How many elderly people we have in NZ
- How many people live in one house
- How many hours people work each week

The government uses this information to help it decide things like:

- Where to build new schools (if there are a lot of young children in one area, they will need a school)
- How many hospital do we need
- Do families need more financial or medical assistance

Describe what a census is, in your own words.

A census is _____

Data Type Exercise:

Look at your data on pages 3 and 4.

Think about the answers – are they a measurement (e.g. height), a count (e.g. how many people have a cell phone in our class), or a group (e.g. ethnic or cultural group).

Circle the data type for each variable in the table below.

Variable	Type of data
Age	Measurement Counting Group or category
Gender	Measurement Counting Group or category
Dinner	Measurement Counting Group or category
Songs	Measurement Counting Group or category
Length of thumb	Measurement Counting Group or category
Length of little finger	Measurement Counting Group or category

Different data

There are 4 types that we need to tell the difference between:

- Summary (Univariate data)
- Relationship (Bivariate data)
- Comparison (Multivariate data)
- Time series (data collected over time)

We need to know what investigation problem is, because the data, graphs and calculations we do for analysis, and the conclusions will be different for each data type.

Variable

A variable is a collection of values (measurements, counts or groups) that are all about the same thing.

Example

If we measure the height of all students in our class, then the height is the variable and it has a lot of values (or measurements).

Problem

Summary Problems

- This has one variable, and could be a measurement or count.
- **For example:** height.
- The investigation problem looks for a summary of the variable.
- **For example:** I wonder what the typical height of Year 9 students at Aorere College is?

Relationship Problems

- This has two variables. They must both be measurements.
- **For example:** height and age.
- This investigation problem looks for a relationship between two variables.
- **For example:** I wonder if there is a relationship between an Aorere College students' height and age?

Comparison Problems

- This has two variables. One is a group and the other is a measurement.
- **For example:** height and gender.
- This investigation problem compares the two groups to see if there is a difference in the measurements.
- **For example:** I wonder if there is any difference between the heights of boys and girls at Aorere College?

PPDAC cycle

Copy the PPDAC cycle into your notes here. This is the core of all statistical investigations.

Exercise:

Classify each of the following as summary, relationship or comparison questions. Circle your answer.

Classify	Question
C / S / R	I wonder what the typical height of Junior students is?
C / S / R	I wonder what are typical right foot lengths for Junior boys?
C / S / R	I wonder what the arm span tends to be for Junior students?
C / S / R	I wonder if girls tend to have a longer right foot length than boys?
C / S / R	I wonder what the most popular sport played is?
C / S / R	I wonder if boys tend to have longer arm spans than girls?
C / S / R	I wonder what the favourite subject for Junior students is?
C / S / R	I wonder if there is a relationship between wrist and neck circumference for Junior students?

C / S / R	I wonder how heavy school bags tend to be for Junior students?
C / S / R	I wonder if right handed students prefer art subjects compared to left handed?
C / S / R	I wonder if boys have had their current phones for longer than girls?
C / S / R	I wonder what are the different regions that these students live in?
C / S / R	I wonder if there is a relationship between how heavy school bags tend to be and how old students are?
C / S / R	I wonder what are typical ways that students carry their bags to school?
C / S / R	I wonder if the boys tend to have larger wrist circumferences than girls at Aorere College?
C / S / R	I wonder what the typical neck circumferences are for Junior students?
C / S / R	I wonder where Junior students tend to go for a holiday?

This data is for the following exercises:

Student	Gender	Age	Height (cm)	Length of arm span (cm)	Main way of travel to school*	Time taken to get to school (min)	Did <i>most</i> at lunchtime *
1	male	12	163	163	walk	Less 10	Ran
2	female	14	155	155	bus	20 - 30	Sat
3	female	12	155	155	walk	10 - 20	Ran
4	male	10	141	144	motor	Less 10	Ran
5	female	14	163	164	motor	20 - 30	Walked
6	male	9	144	144	bus	30 plus	Walked
7	female	13	164	165	bus	30 plus	Sat
8	female	14	158	118	motor	10 - 20	Sat
9	female	14	166	162	bus	10 - 20	Sat
10	female	10	143	138	motor	10 - 20	Walked
11	male	11	149	144	bike	Less 10	Ran
12	female	9	140	140	motor	10 - 20	Ran
13	male	9	127	128	walk	10 - 20	Ran
14	male	13	163	163	motor	10 - 20	Ran
15	female	13	150	147	walk	10 - 20	Ran
16	male	11	146	125	bike	Less 10	Ran
17	male	13	165	154	motor	Less 10	Walked
18	female	12	159	159	motor	Less 10	Walked
19	female	15	160	156	walk	30 plus	Stood
20	male	13	168	175	walk	Less 10	Ran
21	female	15	170	175	motor	Less 10	Sat
22	female	9	132	130	motor	Less 10	Ran
23	male	14	174	182	motor	Less 10	Ran
24	female	12	150	150	bus	30 plus	Stood

*** Questionnaire wording**

Main way to travel to school options: walk, motor vehicle, bus, bike, other.

What you did most at lunchtime options: sat down, stood around, walked around, ran around or played

Summary Questions

Summary questions are ones where there is only one variable. It could be measurements, counts or groups.

Example:

I wonder what the average height of these students is.

Exercise:

Look at the data provided and generate as many different summary questions as you can.

Relationship Questions

Relationship questions are ones where both variables are measurements.

Example:

I wonder if there is a relationship between the height and arm span of students.

Exercise:

Look at the data provided and generate as many different relationship questions as you can.

Comparison Questions

Comparison questions are ones where one variable is a measurement and the second variable is a group.

Example:

I wonder if there tends to be a difference in the time it takes to get to school for students who walk compared with students who catch the bus.

Exercise:

Look at the data provided and generate as many different comparison questions as you can.

Other Questions

If there are any questions that go beyond the data, or do not fit into one of the other 3 categories, place them here.

Example:

I wonder if students who went to school using a car could have used the bus.

I wonder if the results would be different for our class.

Exercise:

Any other questions that you might generate from the data.

Analysis

Summary graphs

- This has only one variable – a measurement or count.
- This could be a stem and leaf plot, bar graph, dot plot, or box and whisker plot.

Relationship graphs

- This has two measurement variables.
- This is a scatterplot.

Time series graphs

- This has data that is collected over time.
- This is a line graph.

Comparison graphs

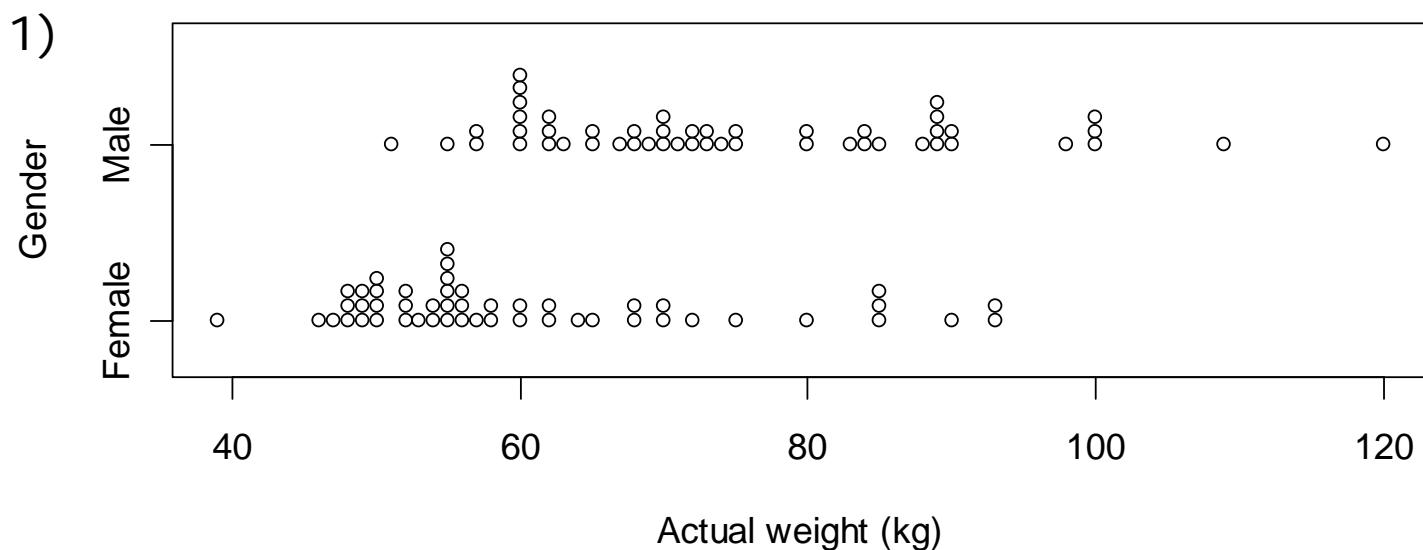
- These compares measurements between two groups.
- This could be back-to-back stem and leaf plots, comparison dot plots, or comparison box and whisker plots.

Exercise:

For each of the graphs below, state what type of graph it is (see the list to choose from below), and what type of problem it would be investigating.

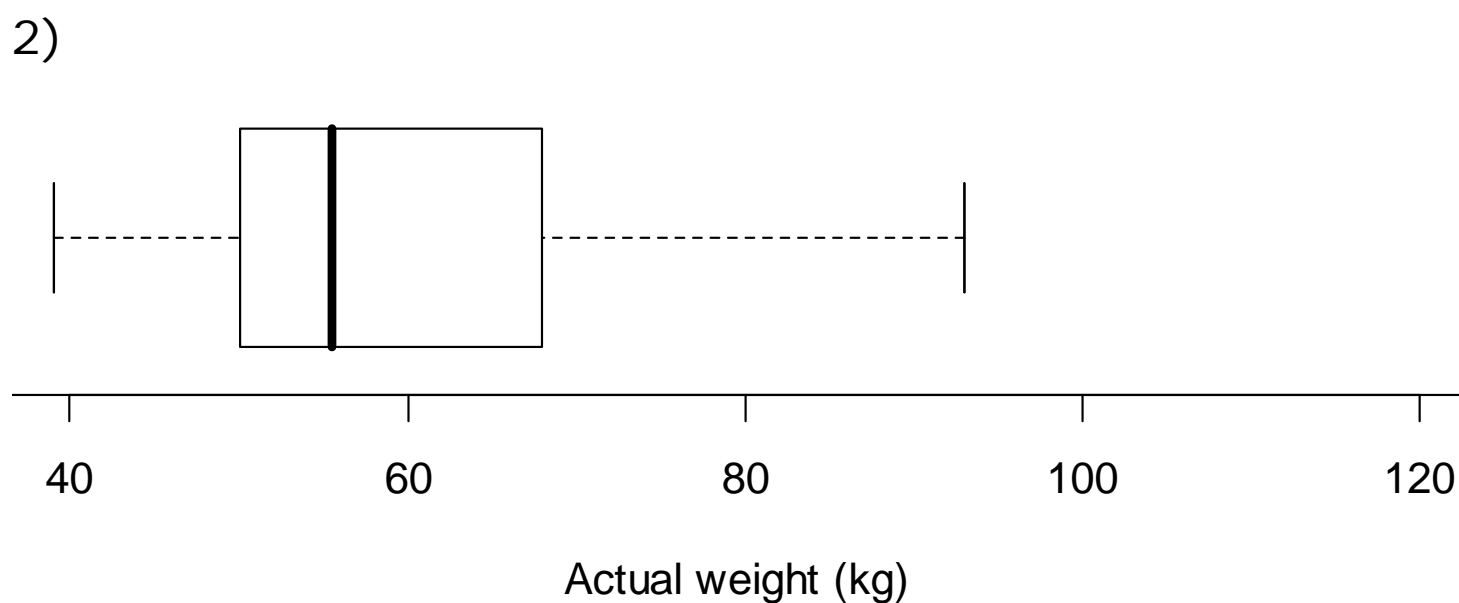
Graph types:

Investigation problem	Graph types
Summary problem	<ul style="list-style-type: none">• Stem and leaf plot• Bar graph• Dot plot• Box and whisker plot
Relationship problem	<ul style="list-style-type: none">• Scatterplot
Time series problem	<ul style="list-style-type: none">• Line graph
Comparison problem	<ul style="list-style-type: none">• Back to back stem and leaf plot• Comparison dot plot• Comparison box and whisker plot



Graph type: _____

Investigation problem: _____



Graph type: _____

Investigation problem: _____

3)

Actual weights of university students (kg)

Females		Males
9	3	
99988876	4	
8876665555554432220000	5	1577
88542200	6	0000002223557889
5200	7	00012233455
5550	8	00344589999
330	9	008
	10	0009
	11	
	12	0

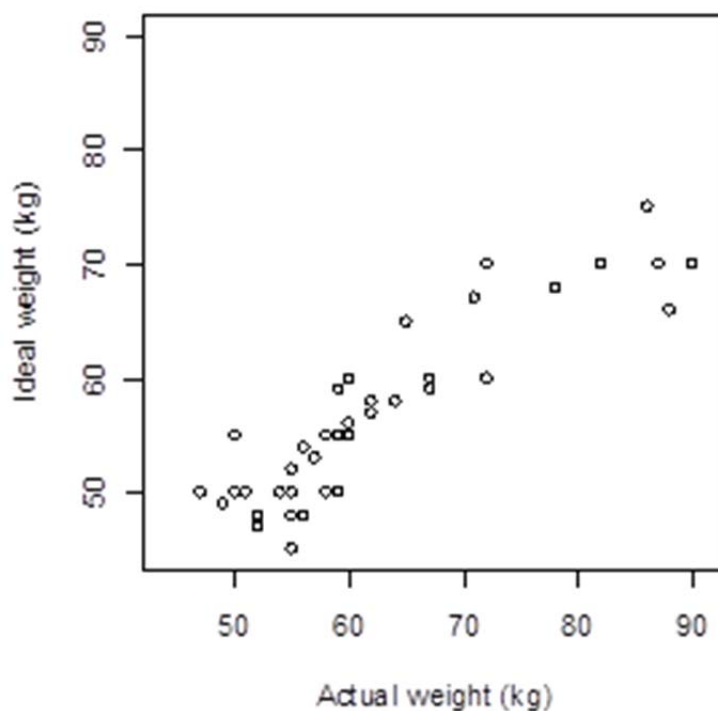
The stem unit is 10kg

Graph type: _____

Investigation problem: _____

4)

Female ideal and actual weights



Graph type: _____

Investigation problem: _____

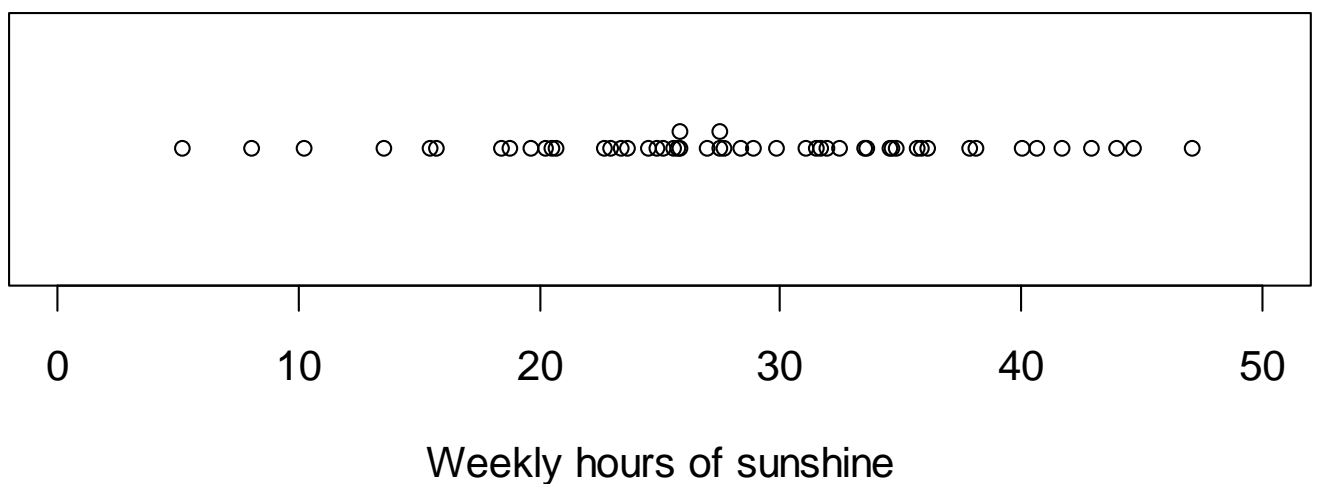
5)



Graph type: _____

Investigation problem: _____

6)

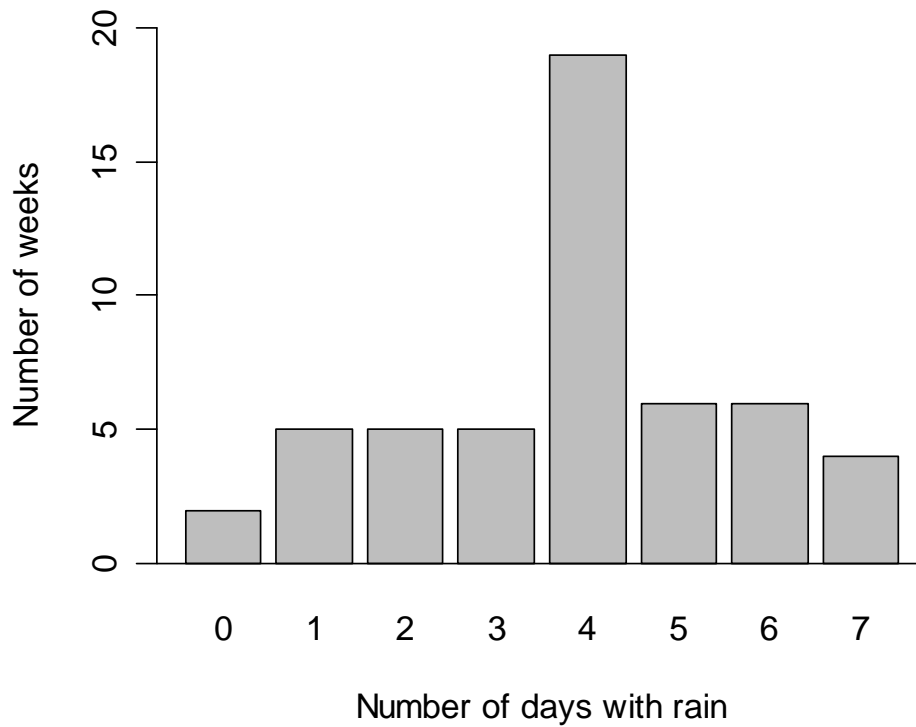


Graph type: _____

Investigation problem: _____

7)

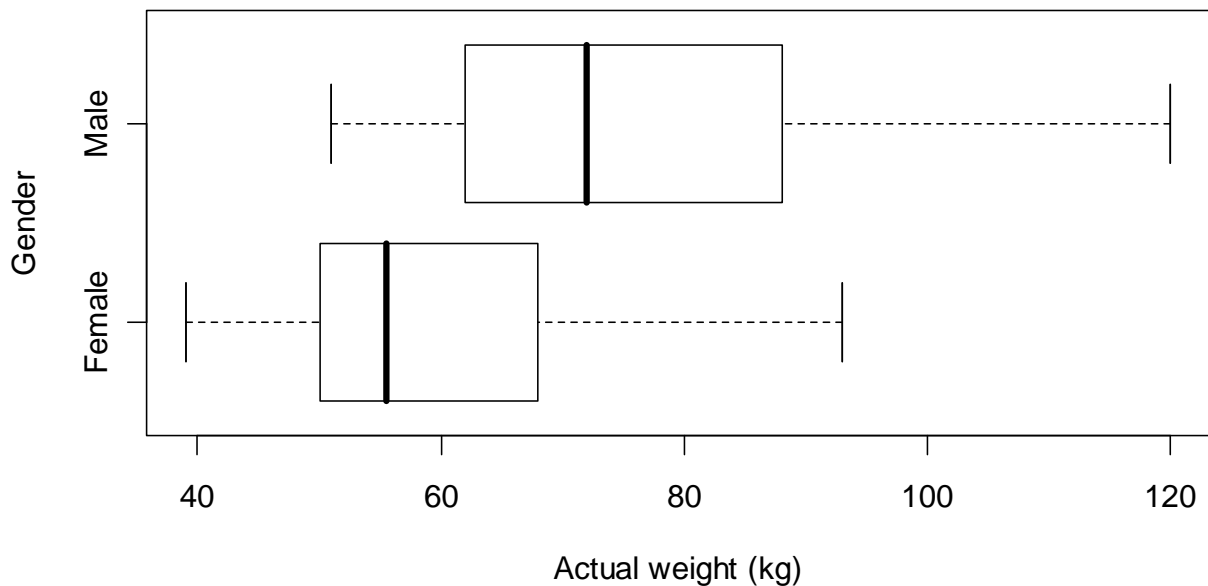
Rainy days per week, Auckland, 2006



Graph type: _____

Investigation problem: _____

8)



Graph type: _____

Investigation problem: _____

9) **Actual weights of male university students (kg)**

5		1577
6		0000002223557889
7		00012233455
8		00344589999
9		008
10		0009
11		
12		0

The stem unit is 10kg

Graph type: _____

Investigation problem: _____

Now we will look at the PPDAC cycle in detail for each of the different investigation types.

- Summary (Univariate)
- Relationship (Bivariate)
- Comparison (Multivariate)
- Time Series

There are separate booklets for each of these investigation types.