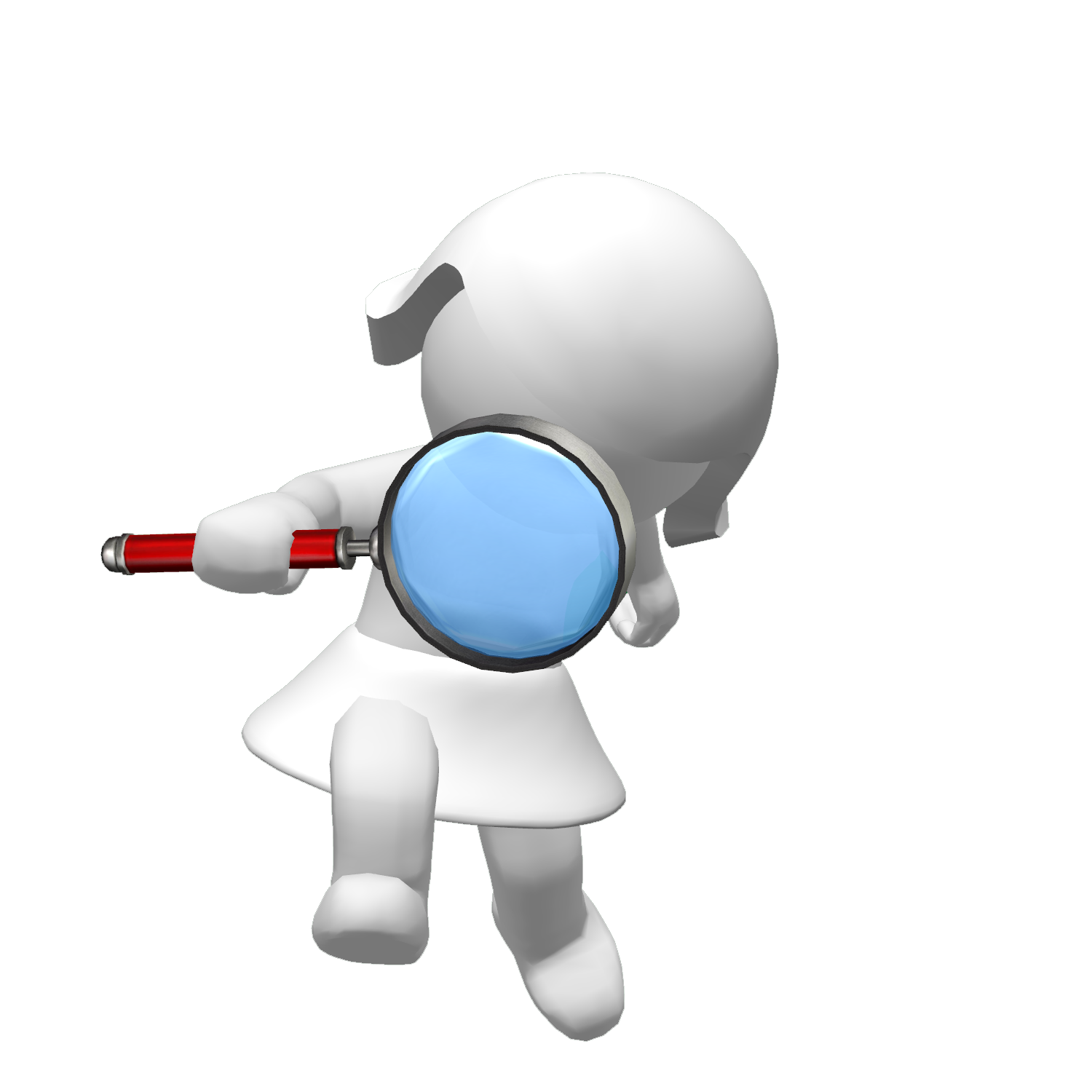


Level 1 Introduction to Data Patterns

##### 

## Name:



**By Liz Sneddon**

## PE Example – whole class

#### Investigation question:

I wonder how long Year 11 OSC students tend to do a wall sit for?

#### Hypothesis / Prediction:

What do you predict?

#### Plan:

1. Students need to pair up.
2. Each pair of students need a timer – either on a phone, or stopwatch.
3. Choose who is going to do the wall sit first, and who is going to use the timer.
4. Student doing the wall sit:
   1. Place your back against the wall.
   2. Place your feet approximately shoulder width apart, facing forwards.
   3. Slide your back down the wall until your back and upper legs form a right angle.
   4. Adjust your feet until your lower legs and upper legs form a right angle.
   5. Place your arms against the wall and beside your body.
5. Student with the timer:
   1. Start the timer when the student is in position.
   2. When the student changes position (e.g. drops to the floor, slides down the wall, or move out of the 90 degree position), stop the timer.
6. Swap the student doing the wall sit and timer, repeating steps 4 and 5.
7. Record the data on the spreadsheet provided.

#### Data:

Follow the instructions and collect data from all students in the class.

Record the data on spreadsheet provided by your teacher, and on the table below.

|  |  |  |  |
| --- | --- | --- | --- |
| **Age** | **Gender** | **Maths teacher** | **Wall sit time (seconds)** |
| 14 / 15 / 16 | Male / Female | AN / CK / HM / SI / SN |  |

#### Observations:

We want data that is as consistent and accurate as possible. What differences did you notice when students were doing the wall sit?

What other factors do you need to control?

### 

### 

#### Explorations:

Think about the data you collected. What investigation questions could you ask of the data? What comparisons could you make? What relationships might exist?

### 

### 

### 

## Social Media Example – whole class



#### Exercise:

Watch the Social Media video ([**http://tiny.cc/8lpopy**](http://tiny.cc/8lpopy)). What information did you notice?

### 

### 

#### Investigation question:

I wonder how long Year 11 OSC students tend to spend on Facebook after school?

#### Hypothesis / Prediction:

What do you predict?

### 

#### Plan / Data:

Login to a device, and fill in the following Google Form: [**http://tiny.cc/vi2mpy**](http://tiny.cc/vi2mpy)

The data will be recorded automatically on a Google Sheets.

#### Observations:

Talk to others around you. What experiences do you have with social media?

### 

### 

### 

#### Explorations:

What other investigation questions could you ask of the data? What comparisons could you make? What relationships might exist?

### 

### 

### 

### 

## Bungee jumps – whole class

#### Investigation question:

I wonder how weight affects a bungee jump?

#### Hypothesis / Prediction:

What do you predict?

### 

#### Plan:

1. Students need to work in groups of 2 - 3.
2. Collect the equipment: elastic, measuring tape, plastic bags and 10 marbles.
3. Explore how different weights affect the length the bungee travels.
4. Record the data.

#### Data:

Record the data on spreadsheet provided by your teacher and on the table below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Type of elastic** | **Maths teacher** | **Length of elastic (mm)** | **Number of marbles** | **Distance (cm)** | **Weight** |
| Narrow / Wide | AN / CK / HM / SI / SN |  |  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

#### Observations:

What did you notice? What other factors might affect a bungee jump?

### 

### 

### 

#### Explorations:

What other investigation questions could you ask of the data? What comparisons could you make? What relationships might exist?

### 

### 

### 

### 

### 

## Sphero Racing Example – whole class



#### Investigation question:

I wonder how long the Sphero’s tend to take to complete the obstacle course?

#### Hypothesis / Prediction:

What do you predict?

### 

#### Plan:

1. Students need to pair up.
2. One student needs a timer – either on a phone, or stopwatch. The other student needs to collect a sphero, and connect their phone. There are instructions on how to connect the sphero on the wall in LC4.
3. In pairs, each student will race the sphero around the obstacle course, while the other student will time using a timer.
4. Each student will take a second and third turn to race the spheros around the obstacle course.
5. Record the data.

#### Data:

Record the data on spreadsheet provided by your teacher and on the table below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Age** | **Gender** | **Maths teacher** | **Obstacle course 1st race (seconds)** | **Obstacle course 2nd race (seconds)** | **Obstacle course 3RD race (seconds)** |
| 14 / 15 / 16 | Male / Female | AN / CK / HM / SI / SN |  |  |  |

#### Observations:

What differences did you notice when students raced the spheros on the obstacle course? What factors do you need to control?

### 

### 

### 

#### Exploration:

What investigation questions could you ask of the data? What comparisons could you make? What relationships might exist?

### 

### 

### 

### 

### 

Investigative Process

The PPDAC cycle is the core of all statistical investigations.

##### 

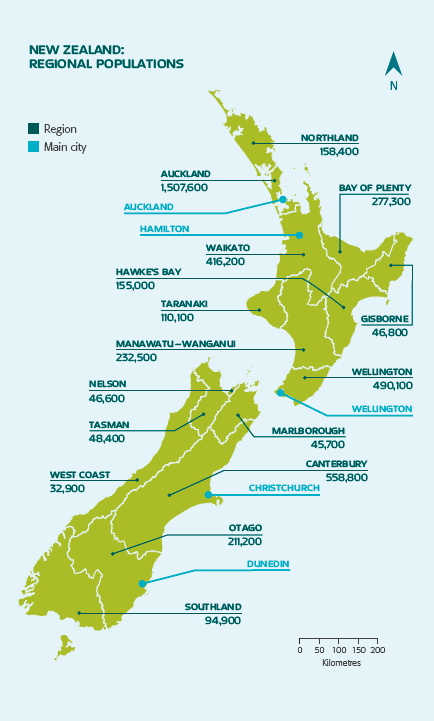
Problem

We start with an investigation question about a population. We often have a hypothesis or prediction of what we expect to find.

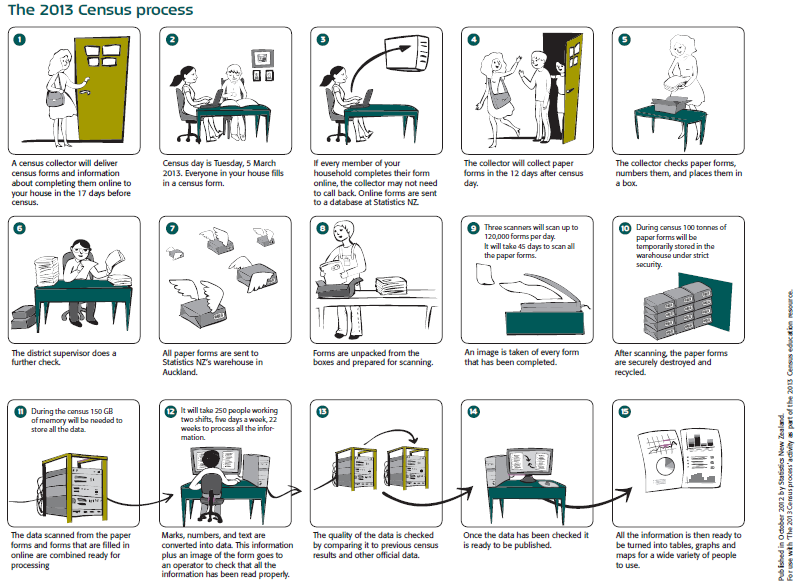
A **population** is all the individual members or items that make up a group.

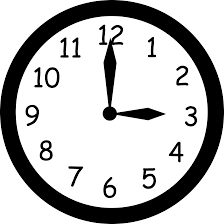
A **sample** is a group of individuals (or items) selected from the population.

##### 

A **census** is a study that attempts to measure every unit in a population.

* The government collects data every 4 years.
* It helps government, councils and businesses to plan for the future.
* The last census in 2013 cost over $100 million dollars.
* It took more than 6 months to collect the data from every person in NZ (over 4.2 million people).





The reasons we take samples are:

* It takes a long **time** to do a census.
* It costs a lot of **money** to collect that much data.

#### Exercise:

1. Describe the population of Ormiston Senior College (be specific).

### 

1. Why don’t we sample the entire population?

### 

1. For the 4 investigations the class did (PE, Media, Bungee and Sphero), decide if your data came from a Census or a Sample. Then describe the population of these investigations.

|  |  |  |
| --- | --- | --- |
| **Investigations** | **Census or Sample?** | **Describe Population** |
| PE – Wall sit |  |  |
| Social Media |  |
| Bungee |  |
| Sphero Racing |  |

4. How accurate is the information from a census compared with a sample? Explain.

### 

### 

5. What does the government use the data collected from the census for? Explain.

### 

### 

Use the following scenarios to identify populations and samples.

6. A beverage company wanted to see if people in the United States liked their new logo. Which choice best represents a population?

A. A selection of logo artists.

B. Every person in the United States.

C. A selection of shoppers from different states.

D. 3,800 children age 5 - 15

7. A musician wanted to see what people who bought his last album thought about the songs. Which choice best represents a sample?

A. Every person who bought the album.

B. A selection of people who didn't want to buy the album.

C. 250 girls who bought the album.

D. A selection of 3,294 people who bought the album.

8. A gaming website wanted to find out which console its visitors owned. Which choice best represents a population?

A. Visitors to the 3DS section.

B. All of the website visitors.

C. Visitors to the PS4 section.

D. Visitors who are on the website for more than 5 minutes.

9. Before a nationwide election, a polling place was trying to see who would win. Which choice best represents a sample?

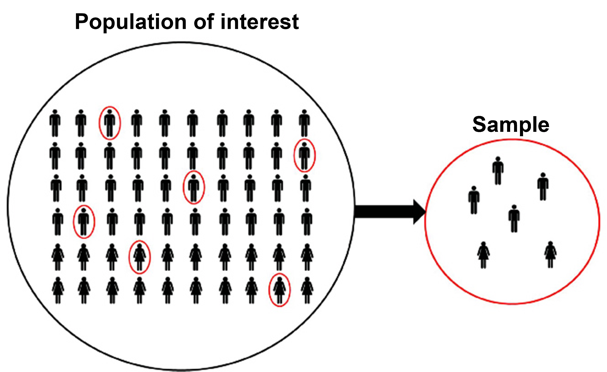
A. A selection of voters over age 50.

B. A selection of male voters.

C. A selection of voters of different ages.

D. All voters



Plan

Data needs to be collected by taking a sample. The sample data will allow us to make estimates about the population without needing the time, money and effort to collect a census.

# Sampling

Samples are selected **randomly** so the characteristics of the sample are typical (**representative**) of the population.

A **random sample** means that each member of the population has the same chance of being selected.

A **biased sample** is not typical of the population. It has a bias for particular members.



A representative sample is a group of people who have been selected **randomly**, so that there is a mix of characteristics in the sample that match the population.

Characteristics may include: a mix of genders, ethnicities, socio-economic status, eye colour, sporting preferences, etc.

### **Example**

If I do a questionnaire with **only blue eyed** students, then I have a **biased** **sample**. This means I do not have any information about people with other coloured eyes (e.g. brown, green, grey, etc), so my data does not represent the population of all people, only the people with blue eyes.

# Sample size

We want to take a big enough **sample size**, so that the data is **accurate** enough to represent the population.

The more data we have, the greater the accuracy of our results.

Use the following rules of thumbs:

* For **discrete (count)** data use a sample size of at least **50**.
* For **continuous (measurement)** data use a sample size of at least **30**.

This is the sample size requirement for **each** group, and for comparisons, the group sample sizes **do not** need to be the **same**.

#### Exercise:

1. Collect a bag of data from your teacher. Take a random sample of 10 males and 10 females.
2. It is important for our samples to be randomly selected. Why?

### 

### 

1. Would a sample of 30 students from Ormiston Senior College be representative of the population of NZ? Explain why/why not.

### 

1. Circle the words that complete the sentences below.
   1. Smaller sample sizes take a **shorter / longer**  time to collect data, but are **more / less** accurate.
   2. Larger sample sizes take a **shorter / longer** time to collect data, and are **more / less**  accurate.
2. Mrs Sneddon is going to survey 35 girls and 40 boys at OJC to investigate their use of ipads at home. Is this a representative sample? Explain.

### 

1. OSC has aout 650 students. The school wants to do a survey on the use of phones by students. For the methods below, state if they are biased or representative samples, and explain why.
   1. Interview all students in a Year 11 class.

### 

* 1. Interview every 20th student leaving school at the end of the day.

### 

* 1. Ask for 40 volunteers to fill in a questionnaire.

### 

* 1. Interview 40 students at a sports game on Wednesday afternoon.

### 

Data

We’ve been collecting data, and now we need to understand what methods we used and what the different data types are.



Data can either be from an **observational study** or an **experimental** **study**.

An **observational study** is where the population is observed without any interference by the investigation.

An **experimental study** is where the investigator randomly assigns people into one of two groups, controlling all other conditions.

#### Example:

The data that you have collected in all 4 Investigations, is **observational** data.

# 

# Data Collection Methods

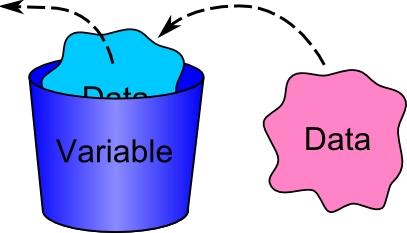
Observational data can be collected in several different ways:

#### Exercise:

For the investigations you trialed, decide what data collection method was used (Observation, Interview, Questionnaire, Database).

|  |  |
| --- | --- |
| **Investigations** | **Data Collection Method** |
| PE – Wall sit |  |
| Social Media |  |
| Bungee |  |
| Sphero Racing |  |

# Variables and Data



A **variable** describes a characteristic of an individual from the population. The characteristic changes or varies from one individual to another.

**Data** is collected when the values of variables are recorded for individuals.

#### Example:

Here is the PE spreadsheet:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **First Name** | **Last Name** | **Age** | **Gender** | **Maths teacher** | **Wall sit time** |
|  |  | 14 / 15 / 16 | Male / Female | AN / CK / HM / SI / SN |  |
|  |  | 14 / 15 / 16 | Male / Female | AN / CK / HM / SI / SN |  |

Each **row** is a set of **data** belonging to a student.

Each column is a variable.

In this example there are **6 variables**:

* First name,
* Last name,
* Age,
* Gender,
* Maths teacher, and
* Wall sit time.

#### Exercise:

For each of the spreadsheets, name the variables.

Here is the Social Media spreadsheet:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Age** | **Gender** | **Devices** | **Social Media Platforms** | **Computer Time** | **TV Time** | **Game console** | **Phone Time** |
| 14 / 15 / 16 | Male / Female | Cellphone / Own computer / Family computer / None of the above | Facebook / Twitter / Instagram / Snapchat / None of the above |  |  |  |  |

**Variables:**

### 

Here is the Bungee spreadsheet:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **First Name** | **Last Name** | **Type of elastic** | **Maths teacher** | **Length of elastic (mm)** | **Number of marbles** | **Distance (cm)** | **Weight** |
|  |  | Narrow / Wide | AN / CK / HM / SI / SN |  |  |  |  |
|  |  | Narrow / Wide | AN / CK / HM / SI / SN |  |  |  |  |

**Variables:**

### 

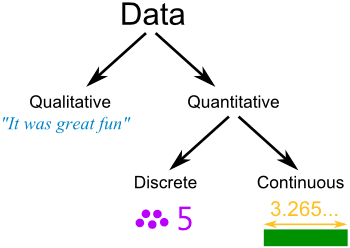
Here is the Sphero Racing spreadsheet:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **First Name** | **Last Name** | **Age** | **Gender** | **Maths teacher** | **Obstacle course 1st race** | **Obstacle course 2nd race** | **Obstacle course 3RD race** |
|  |  | 14 / 15 / 16 | Male / Female | AN / CK / HM / SI / SN |  |  |  |
|  |  | 14 / 15 / 16 | Male / Female | AN / CK / HM / SI / SN |  |  |  |

**Variables:**

### 

# Data Types

**Qualitative (groups) variables** are characteristics, that cannot be described by numbers e.g. gender, ethnicity, apple variety.

**Quantitative (numerical) variables** are characteristics described by numbers e.g. height, age, number of apples, weight. Numerical variables are either **discrete** or **continuous**.

**Discrete variables** (whole numbers), values obtained by counting.

**Continuous variables** (measurement), values obtained by measuring.

#### Example:

Here is the PE spreadsheet:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **First Name** | **Last Name** | **Age** | **Gender** | **Maths teacher** | **Wall sit time** |
|  |  | 14 / 15 / 16 | Male / Female | AN / CK / HM / SI / SN |  |
|  |  | 14 / 15 / 16 | Male / Female | AN / CK / HM / SI / SN |  |

Data Types

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **First Name** | **Last Name** | **Age** | **Gender** | **Maths teacher** | **Wall sit time** |
| Qualitative | Qualitative | Quantitative  Discrete | Qualitative | Qualitative | Quantitative  Continuous |

#### Exercise:

For each of the spreadsheets, name the data types.

1. Social Media spreadsheet:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Age** | **Gender** | **Devices** | **Social Media Platforms** | **Computer Time** | **TV Time** | **Game console** | **Phone Time** |
| 14 / 15 / 16 | Male / Female | Cellphone / Own computer / Family computer / None of the above | Facebook / Twitter / Instagram / Snapchat / None of the above |  |  |  |  |

Data Types

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Age** | **Gender** | **Devices** | **Social Media Platforms** | **Computer Time** | **TV Time** | **Game console** | **Phone Time** |
|  |  |  |  |  |  |  |  |

2. Bungee spreadsheet:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **First Name** | **Last Name** | **Type of elastic** | **Maths teacher** | **Length of elastic (mm)** | **Number of marbles** | **Distance (cm)** | **Weight** |
|  |  | Narrow / Wide | AN / CK / HM / SI / SN |  |  |  |  |

Data Types

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **First Name** | **Last Name** | **Type of elastic** | **Maths teacher** | **Length of elastic (mm)** | **Number of marbles** | **Distance (cm)** | **Weight** |
|  |  |  |  |  |  |  |  |

3. Sphero Racing spreadsheet:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **First Name** | **Last Name** | **Age** | **Gender** | **Maths teacher** | **Obstacle course 1st race** | **Obstacle course 2nd race** | **Obstacle course 3rd race** |
|  |  | 14 / 15 / 16 | Male / Female | AN / CK / HM / SI / SN |  |  |  |

Data Types

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **First Name** | **Last Name** | **Age** | **Gender** | **Maths teacher** | **Obstacle course 1st race** | **Obstacle course 2nd race** | **Obstacle course 3RD race** |
|  |  |  |  |  |  |  |  |

## Cleaning data

Look for the following issues:

* Data entry mistakes
* Incorrect units
* Missing data

BUT, you cannot change/delete data unless you KNOW that it is a mistake.

If you are CERTAIN the data is wrong, then make the cell blank (or enter a 0).

#### Exercise:

Find any data that doesn’t make sense and highlight the values. Add a comment stating what correction (if any) you would make.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Gender** | **Age** | **Country of birth** | **Languages spoken** | **Height** | **Right foot** | **Arm span** | **Index finger** | **Corrections / changes** |
| girl | 14 | Russia | 1 | 149 | 220 | 115 | 5 |  |
| boy | 11 | NZ | 1 | 141 | 22 | 142 | 65 |  |
| girl | 14 | NZ | 2 | 175 | 255 | 176 | 81 |  |
| girl | 13 | NZ | 1 | 162 | 25 | 64 | 80 |  |
| girl | 1 | NZ | 1 | 158 | 25 | 163 | 97 |  |
| girl | 12 | NZ | 1 | 164 | 28 | 1 | 80 |  |
|  | 13 | NZ | 2 | 166 | 26 | 180 | 100 |  |
| girl | 12 | cookisl | 1 | 154 | 23 | 156 | 49 |  |
| girl | 14 | NZ | -1 | 170 | 26 | 1 | 70 |  |
| girl | 12 | India | 1 | 144 | 21 | 153 | 65 |  |
| girl | 11 | Kiribati | 2 | 155 | 24 | 158 | 90 |  |
| girl | 10 | NZ | 1 | 147 | 23 | 151 | 8 |  |
| boy | 14 | NZ | 1 | 170 | 28 | 1 | 90 |  |
| girl | 12 | NZ | 2 | 0 | 24 | 158 | 70 |  |
| girl | 12 | NZ | 1 | 150 | 22 | 157 | 60 |  |

## Data displays

Here are some common data displays.

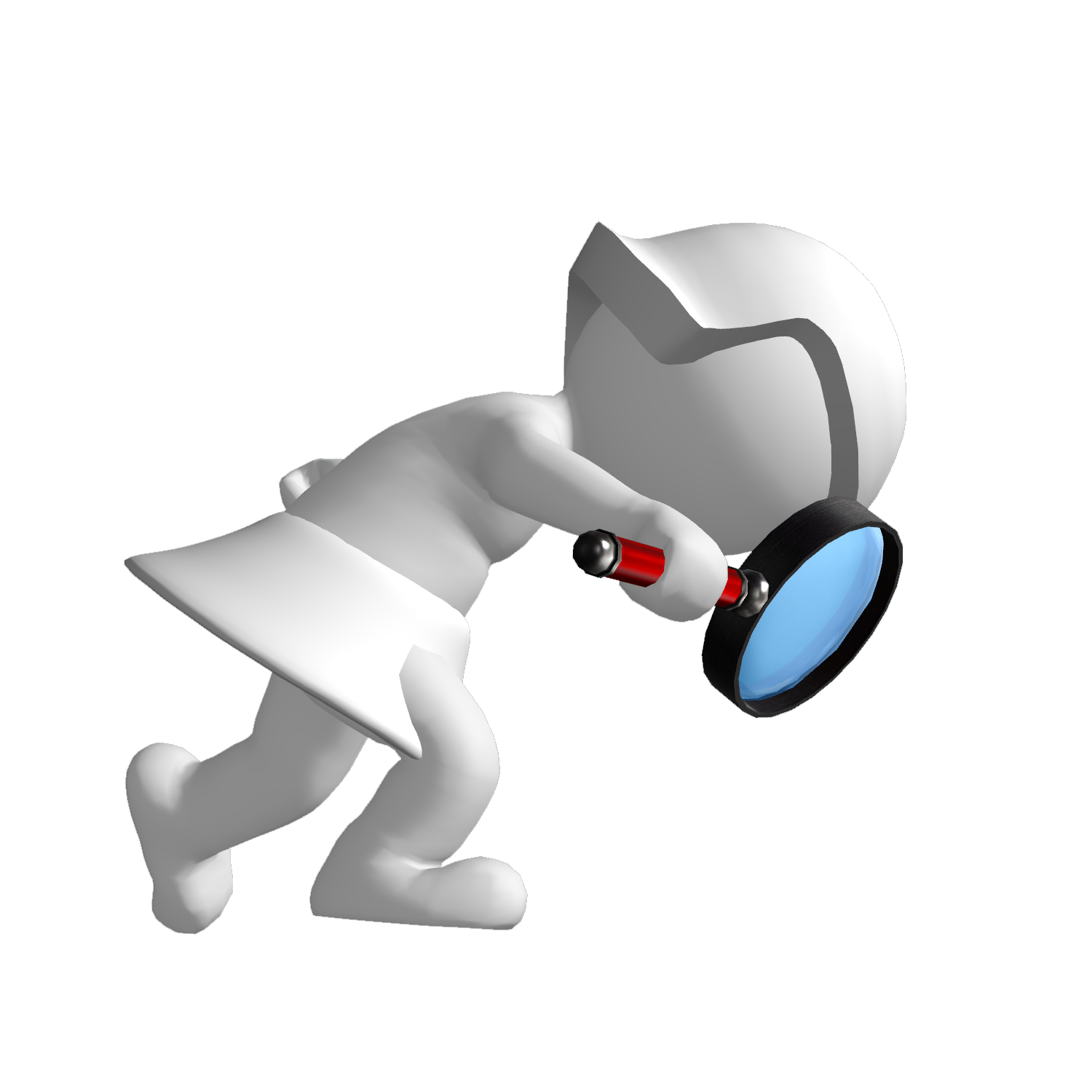
|  |  |  |
| --- | --- | --- |
| **Graph types** | **Summary Example** | **Comparison Example** |
| Stem and leaf plot |  |  |
| Bar graph |  |  |
| Histogram |  |  |
| Dot plot |  |  |
| Box and whisker plot |  |  |
| Scatterplot |  |  |
| Line graph |  |  |
| Time series graph |  |  |



## Data Investigations

There are 4 types of investigations 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.

## 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 OSC 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 OSC students’ height and age?

**Comparison Problems**

This has **two** variables. One is a **group (qualitative)** 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 boys tend to be taller than girls at OSC?

#### Exercise:

Classify each of the following as summary, relationship or comparison questions. Circle your answer. Then state the variable(s).

|  |  |  |
| --- | --- | --- |
| **Classify** | **Question** | **Variable(s)** |
| 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 OSC? |  |
| 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? |  |

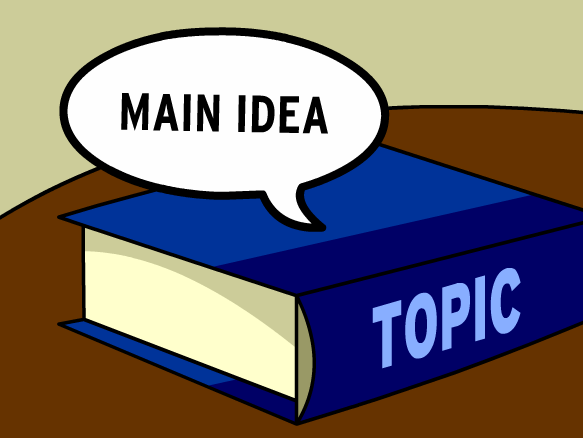
This data is for the following exercises:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Student | Gender | Age | Height  (cm) | Length of arm span  (cm) | *Main* way of travel to school\* | Time taken to get to school (min) | Did *most* at lunchtime\* |
| 1 | male | 12 | 163 | 163 | walk | 10 | Ran |
| 2 | female | 14 | 155 | 155 | bus | 15 | Sat |
| 3 | female | 12 | 155 | 155 | walk | 20 | Ran |
| 4 | male | 10 | 141 | 144 | motor | 6 | Ran |
| 5 | female | 14 | 163 | 164 | motor | 25 | Walked |
| 6 | male | 9 | 144 | 144 | bus | 34 | Walked |
| 7 | female | 13 | 164 | 165 | bus | 37 | Sat |
| 8 | female | 14 | 158 | 118 | motor | 12 | Sat |
| 9 | female | 14 | 166 | 162 | bus | 18 | Sat |
| 10 | female | 10 | 143 | 138 | motor | 14 | Walked |
| 11 | male | 11 | 149 | 144 | bike | 8 | Ran |
| 12 | female | 9 | 140 | 140 | motor | 16 | Ran |
| 13 | male | 9 | 127 | 128 | walk | 19 | Ran |
| 14 | male | 13 | 163 | 163 | motor | 11 | Ran |
| 15 | female | 13 | 150 | 147 | walk | 17 | Ran |
| 16 | male | 11 | 146 | 125 | bike | 7 | Ran |
| 17 | male | 13 | 165 | 154 | motor | 6 | Walked |
| 18 | female | 12 | 159 | 159 | motor | 3 | Walked |
| 19 | female | 15 | 160 | 156 | walk | 56 | Stood |
| 20 | male | 13 | 168 | 175 | walk | 7 | Ran |
| 21 | female | 15 | 170 | 175 | motor | 8 | Sat |
| 22 | female | 9 | 132 | 130 | motor | 5 | Ran |
| 23 | male | 14 | 174 | 182 | motor | 8 | Ran |
| 24 | female | 12 | 150 | 150 | bus | 45 | 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.

### 

### 

### 

### 

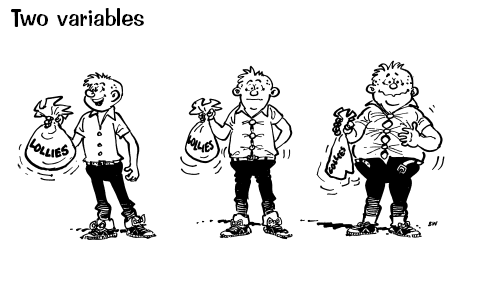
### 

### 

### 

### 

### 

**Relationship Questions**

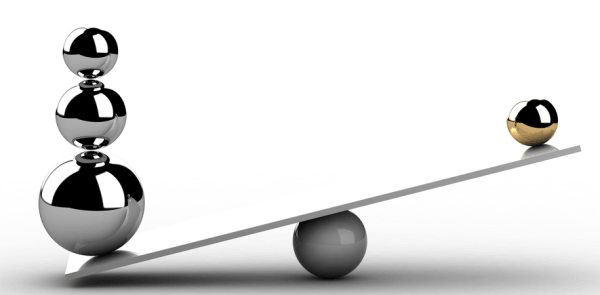
Relationship questions are ones where **both variables** are **measurements**.

#### 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 students who walk to school tend to take longer to get to school than 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.

**Graphs and Investigations**

In order to analyse the data we first need to draw graphs. Depending on the investigation type, this will determine which graph you will draw.

|  |  |  |
| --- | --- | --- |
| **Investigation type** | **Variables** | **Graph types** |
| Summary investigation | One variable   * measurement or count. | * Stem and leaf plot * Bar graph * Dot plot * Box and whisker plot |
| Relationship investigation | Two variables   * one must be a measurement. * the second could be a measurement or a count. | * Scatterplot |
| Time series investigation | Two variables   * one must be time (years, months, days, hours, etc) * the second could be a measurement or a count. | * Line graph |
| Comparison investigation | Two variables   * one must be a measurement or count. * the second must be a group. | * Back to back stem and leaf plot * Comparison dot plot * Comparison box and whisker plot |

#### 

#### 

#### 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 investigation it is.

|  |  |  |
| --- | --- | --- |
|  | Graph type: |  |
| Investigation type: |  |

|  |  |  |
| --- | --- | --- |
|  | Graph type: |  |
| Investigation type: |  |

|  |  |  |
| --- | --- | --- |
|  | Graph type: |  |
| Investigation type: |  |

|  |  |  |
| --- | --- | --- |
|  | Graph type: |  |
| Investigation type: |  |

|  |  |  |
| --- | --- | --- |
|  | Graph type: |  |
| Investigation type: |  |

|  |  |  |
| --- | --- | --- |
|  | Graph type: |  |
| Investigation type: |  |

|  |  |  |
| --- | --- | --- |
|  | Graph type: |  |
| Investigation type: |  |

|  |  |  |
| --- | --- | --- |
|  | Graph type: |  |
| Investigation type: |  |

|  |  |  |
| --- | --- | --- |
|  | Graph type: |  |
| Investigation type: |  |

## 

Go to the following website: <https://www.jake4maths.com/grapher/>

Select the **dataset** (drop down menu on the top right hand side), then choose the **Graph Type:** Pairs plot.

This gives an overview of the dataset, the variables, and the comparative graphs. If you click on any of the graphs, it will take you to that graph.

#### Exercise:

Go to NZGrapher and explore one of the following 2 datasets:

1. Diamonds.csv

2. Rugby.csv

Work out how to add a title to the graph, change the axis labels, summary statistics, regression lines, and other features.

## 

Fill in the planned completion dates, and when you have finished each Key skill, show your teacher so they can track your progress.

|  |  |  |  |
| --- | --- | --- | --- |
| **Key Skill** | **Workbook pages** | **Planned Completion Date** | **Completion Date (teachers sign)** |
| Problem | 10 - 13 |  |  |
| Plan | 14 - 16 |  |  |
| Data | 17 - 25 |  |  |
| Investigations | 26 - 36 |  |  |

## Next step

You need to know how to interpret them, analyse them and draw conclusions.

There are workbooks for the following topics:

* Multivariate Analysis
* Bivariate Analysis
* Linear Algebra

Talk with your teacher about your selection of Investigation topic.

* PE
* Media studies
* Barbie bungee
* Sphero racing