

# Presenting Data and Infographics

## Introduction

We are going to look at various ways of representing data by using your data to produce an Infographic. An Infographic provides a quick means of presenting information, data or knowledge in a clear understandable way using a combination of illustrations, graphs, tables and text.

## Representing Data

Why should we use charts to represent data? While adding raw data to a table helps to organise and present the data it is hard to interpret any meaning or gain any useful insight. Using charts we can quickly start to see what the data might be telling us, further, we can use this so that their most significant message is immediately evident...that the point we want prove or the story we want tell is understood.

## Florence Nightingale 1820-1910

Florence Nightingale was the founder of modern nursing and a celebrated social reformer of her day. A less well known fact about Nightingale is that she was also a statistician who pioneered the visual presentation of information.

Nightingale displayed a gift for mathematics at an early age and had a keen interest in statistics eventually going on to become the first female member of the Royal Statistical Society in 1858 (the RSS was founded 24 years previous in 1834.) As a young woman during the 1840s Nightingale was determined to devote her life to the service of others and, against the wishes of her family, took up nursing, going on to work hard to educate herself in the art and science of nursing.



Nightingale came to worldwide prominence in 1854 when she led women volunteer nurses, to the Crimea where Britain and her allies had been fighting the Russian Empire for control of territories belonging to the Ottoman Empire, the conflict became known as the Crimean War (1853-1856). Nightingale became motivated to get involved upon hearing of the horrific conditions for the wounded soldiers, this news, much like her subsequent fame, owed much to the modern way the conflict was covered by the media...new technological advances, such as the telegraph, meant that first hand reports from the front line got back to Britain quickly and on a daily basis.

During the war Nightingale noticed that far more soldiers died from illnesses brought on by poor hospital conditions; such as typhus, typhoid, cholera and dysentery, than from battle wounds and began to use statistical methods to analyse the mortality and medical data collected and make changes to the way injured soldiers were cared for. After the war realising that many people would find traditional statistical reports difficult to decipher Nightingale began to display this information graphically, developing a form of pie chart we now know as a polar area diagram.

## Interesting facts about the Crimean War

- War between Russian and Britain, France, the Ottoman Empire (Turkey) and the Kingdom of Sardinia.
- Much of the fighting took place on the Crimean Peninsula on the northern coast of the Black Sea. An area located in modern-day Ukraine although since 2014 this is in dispute between Russia and Ukraine.

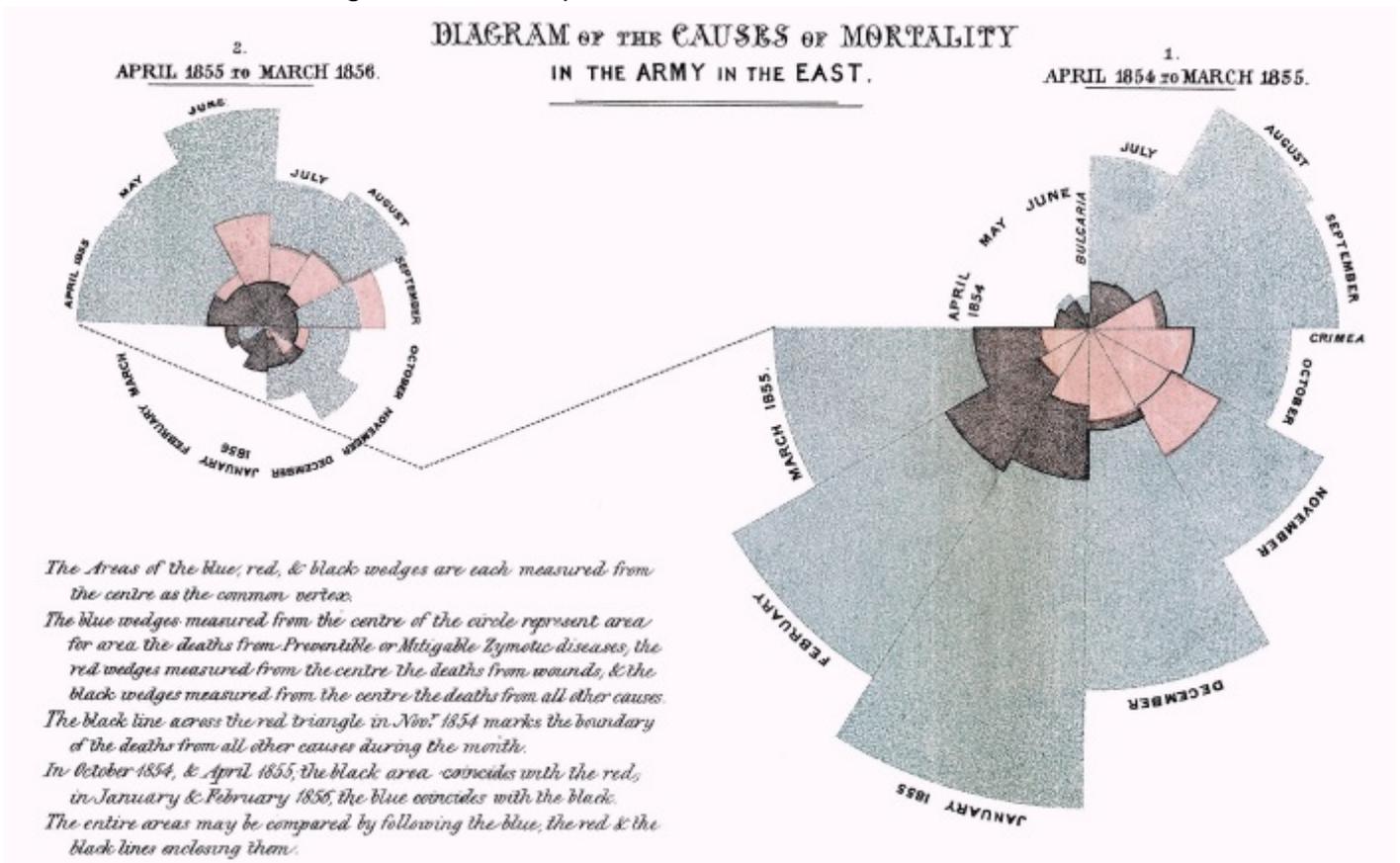
- The Battle of Balaclava (1854) is famous for the ‘Charge of the Light Brigade’ in which the Light Cavalry Brigade, commanded by the Earl of Cardigan, charged a heavily defended part of the battlefield, over 70% of the men lost their lives.
- Peace talks began in 1856 and the war came to an end with the signing of the Treaty of Paris.
- The cold conditions led to the naming of two items of clothing, the Balaclava and the Cardigan.
- New technology played a big part, including the telegraph, photography, trains and steam warships.

## An early Infographic

Sometimes referred to as the Nightingale Rose Diagram, take a look at the polar area diagram (nowadays we might call it a circular histogram) produced by Nightingale. A modern legend has been added but otherwise it is how it was when it was first published.

What do you notice? If you saw this for the first time what questions would you have? Deaths from wounding and other causes combined are still far smaller than deaths from illness. After viewing this you might want to delve further, categorising the deaths and producing further graphs of the results. You can tell, roughly, the month major battles took place and we can even make some basic assumptions about the weather, it is clear to see that the bitterly cold winter had an impact on both deaths from other causes and on overall deaths.

This diagram is an early example of an Infographic: text and graphics deliver a clear message using complex information in an interesting and engaging way that most people could understand without the need for long or detailed explanation.



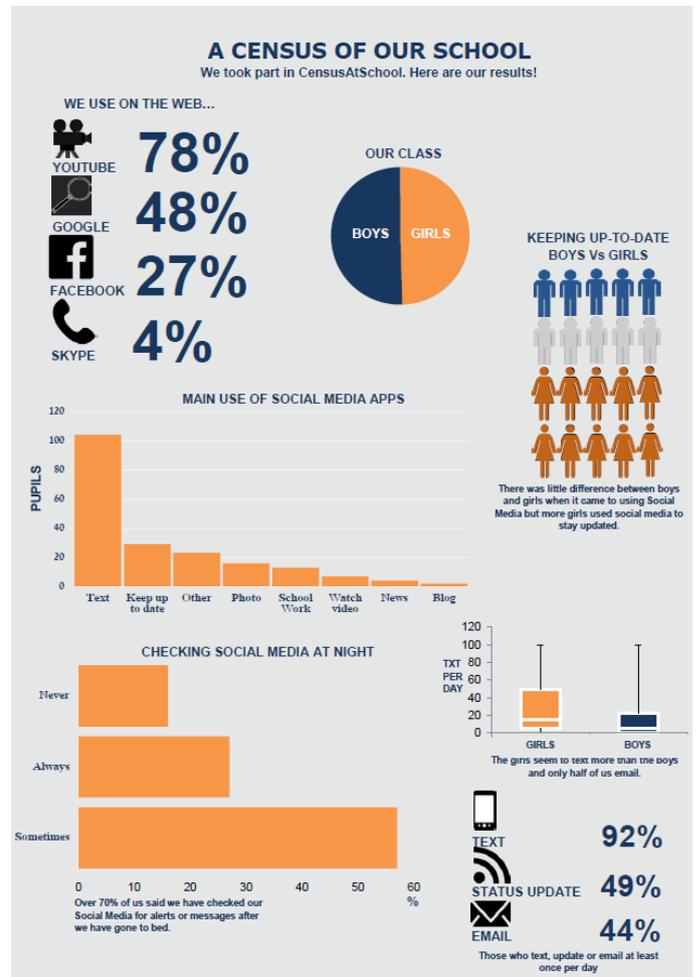
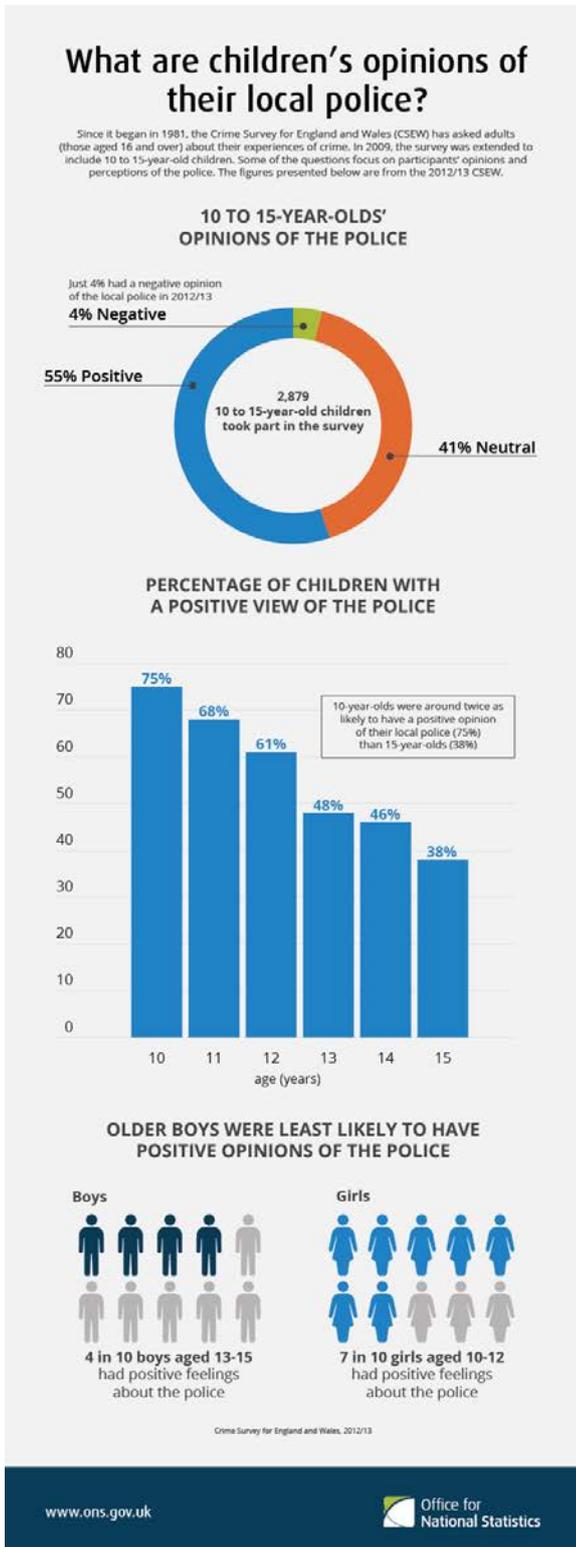
### Legend

- Death from illness
- Death from other causes
- Death from wounds sustained in battle

Here are some useful links about representing data and Infographics:

- <http://www.informationisbeautiful.net>
- <https://theidpblog.files.wordpress.com/2013/10/infographic-guidelines-v1-0.pdf>
- <https://en.wikipedia.org/wiki/Infographic>
- <http://news.bbc.co.uk/1/hi/programmes/newsnight/8901376.stm>

Below is a modern Infographic. On the right Word and Excel has been used to tell a story using CensusAtSchool data, later on we will be producing an Infographic using the returned questionnaire results.



**Figure 1 - left:** Children's opinions of their local police in England and Wales in the 2012/13 Crime Survey for England and Wales infographic. Produced by the UK Office for National Statistics and distributed under a Creative Commons licence.

**Figure 2 - above right:** This Infographic has been produced using all the data available from a single school. We would call this a self-selecting population as not every pupil will have taken part, you must always be careful about any statements you make or conclusions you draw for this reason.

## Taking a Census

A Census includes everyone and is a simple collection of facts. The Office for National Statistics (ONS), who organise the National Census every 10 years (for England and Wales), is not interested in opinions but gathers information in order to plan for the future. They provide information on the number of people in a country (did you know the current UK population is around 64.5 Million), where they live and work and the houses they live in.

**Q.** Do you know when the last National Census for England and Wales took place?

**A.** The last Census of England and Wales took place in 2011, which means the next one is in 2021.

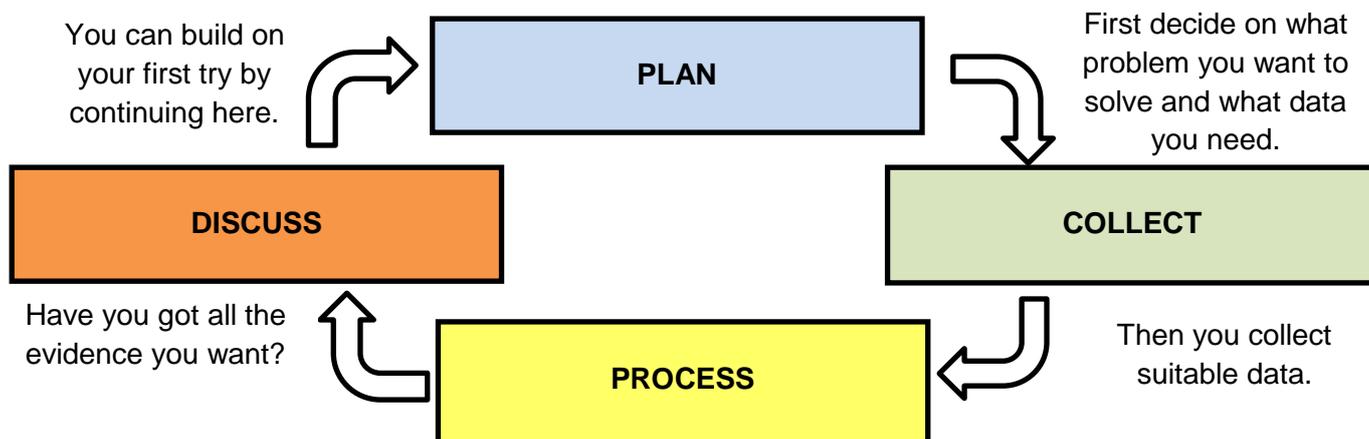
**Q.** Collecting data is time consuming and expensive so why is it so important to carry out a National Census, is it worth it?

**A.** It is important to know how many people live in the country so that the Government can work out what types of services are needed now and in the future. With statistics from the census, central and local government can allocate funds, decide future policy and plan important services.

## CensusAtSchool

Instead of collecting data about the UK population CensusAtSchool collects data from and about your school population, remember in statistical terminology a 'population' could refer to a population of people, animals, plants, cars, or in fact of anything else.

CensusAtSchool follows the Problem Solving Cycle; you will find this a useful method when tackling problems reliant on working with data, can you complete the diagram below.



### PLAN

The questionnaire used to collect your data includes some measurement questions and some themed questions. When you completed the questionnaire which questions interested you the most? Is there a bigger question you could answer using the returned information?

These questions form part of the planning process, this would be the stage at which we make our hypothesis, a statement that may turn out to be true or false. The process of identifying and testing hypotheses is known as **Hypothesis Testing**.

## COLLECT

The questionnaire has been the means of collecting the data on this occasion. Remember, this has come from a self-selecting population...your class.

What does this mean? We cannot say all of the data collected by CensusAtSchool represents a true picture of all children in the UK. In fact in relation to your class we can go further and say that it does not necessarily represent children in your school either. Therefore, we would call this returned data a biased sample.

Can you give a practical example where assuming your class data was representative of all UK children would be unwise?

.....

.....

.....

.....

## PROCESS

Once you have your data back you will need to process the data. What do you think this involves first?

- You should clean the data, what will you do with missing records, spelling mistakes and incorrect answers e.g. text where a number should be.
- You will also need to look at outliers, are these genuine responses or an error? The current questionnaire allows a respondent to enter "999" for text messages per day...if someone gave this answer should you accept it?
- You may need to format responses also. Are dates in the correct format are measurements to the correct decimal place.

Once data is cleaned you can start to plot it, this is the best way to gain some insight into what the data might tell us, use your plan to guide you.

Remember the type of data will dictate how you plot it, why not go along the columns in the returned spreadsheet and describe each variable. Is it:

- **Qualitative** – data is a quality i.e. colour
- **Quantitative** – data is a quantity i.e. a number.

Quantitative data can be divided again into two types.

- **Discrete** – every possible value can be listed i.e. types of pet.
- **Continuous** – any data where you cannot possibly list all values i.e. measurements.

Pairing data, comparing pairs of things can be insightful, for example, you could compare height and arm span. You can pair mixtures of qualitative and quantitative data.

## Practical Things to Try

You may need to construct a frequency table for one (or more) column(s) you are interested in. COUNTIF can be useful here (e.g. `=COUNTIF(A2:A251,"M")` ) if you need to set more criteria COUNTIFS can also be useful (e.g. `=COUNTIFS(A2:A251,"M",B2:B251,"Yes")` ).

Refer to the code sheet if you are unsure what an entry means, in many cases it will be self-evident e.g. "M" = Male or 'Y' = Yes, but this is not always the case.

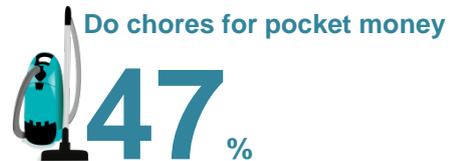
## DISCUSS

Once you have a few interesting plots discuss your findings, keep an eye out for:

- **Measures of average** i.e. mean, median or mode,
- **Measures of spread** i.e. the range, interquartile range or standard deviation of the data.
- Is the data **symmetrical** or **skewed**?

Now decided what you like to add to your Infographic? Remember the Infographic is about telling a story you only need one good graph in combination with some text and illustrations to get your point across.

A nice easy way of displaying information in an Infographic is to use percentages. Count up the instances in a column, calculate the percentage and add it to your Infographic.



Ratios are another great way to add illustrations, work out the percentage and then convert it to a ratio.

