UPCSE Biology –
Basic Statistic course

On drawing bar charts and histograms:

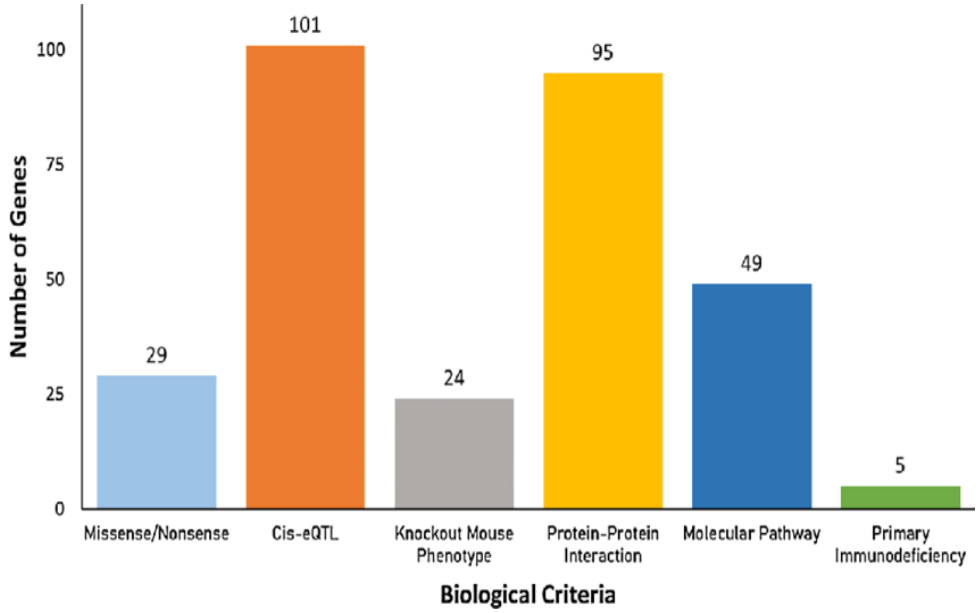
Part I: Bar charts

Introduction

- Over the next two lessons we are going to study three ways in which we can present data in visual form. These are:
  - Bar charts,
  - Histograms,
  - Scatter plots on an x-y graph.
- Histograms and bar charts look similar but are used for totally different reasons.
- All three forms of data presentation above are used for different reasons and you yourselves will have to decide which of the three to use to visualise your experimental data.

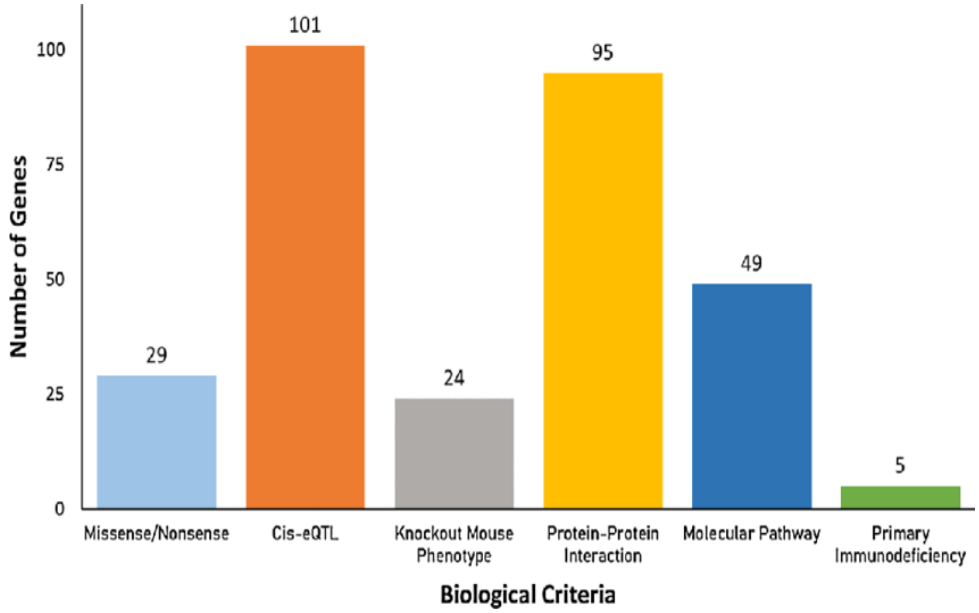
What are bar charts?

- Bar charts are one of the easiest ways to describe or present data in visual form. They act as a visual summary of data.
- They are used to compare different types of "things". More precisely we use bar charts to compare different categories of data, as shown below.
- Here we have a bar chart illustrating six different biological criteria.



What are bar charts?

- The different categories of "Biological criteria" are specified under each bar ...
- … and the height of each bar represents the number for each category.
- So we can see that there are 95 genes for the type Protein-Protein interaction, 5 genes for the type primary immunodeficiency, etc.

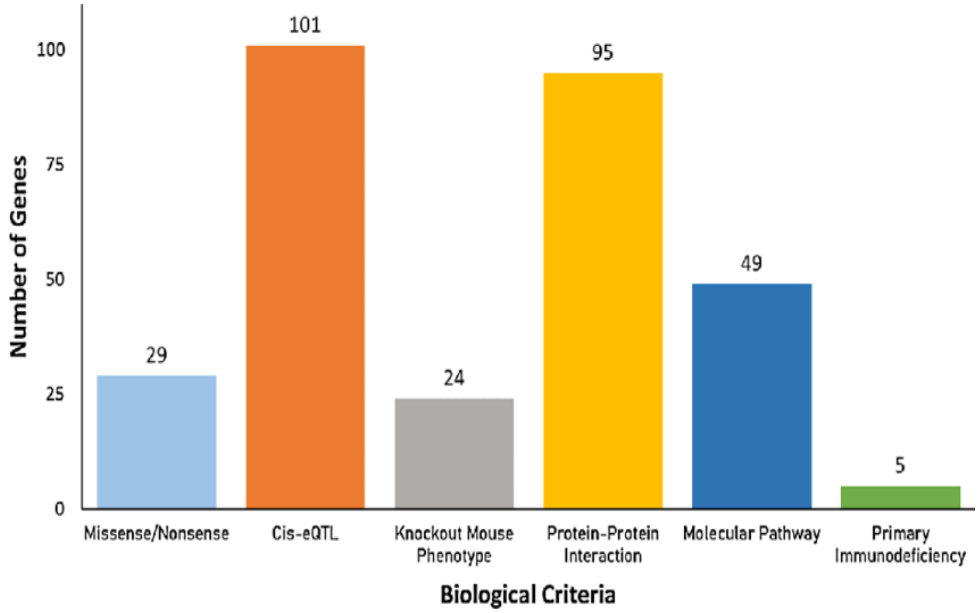


What are bar charts?

- The numerical values of the different categories are represented by the height of the rectangles.

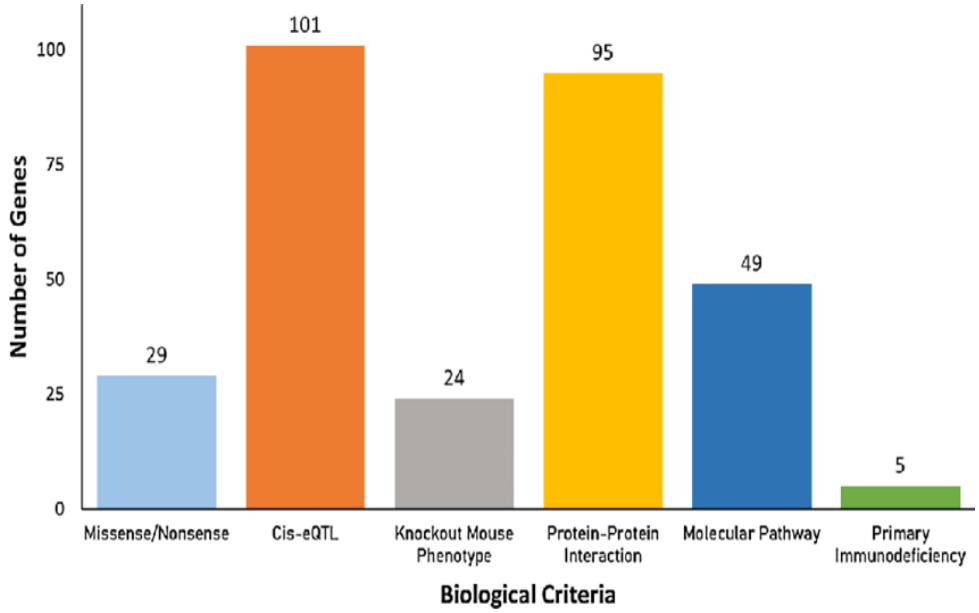
- Also, note that the width of each rectangle is not scientifically important.

- It is, however, visually important to the eye.



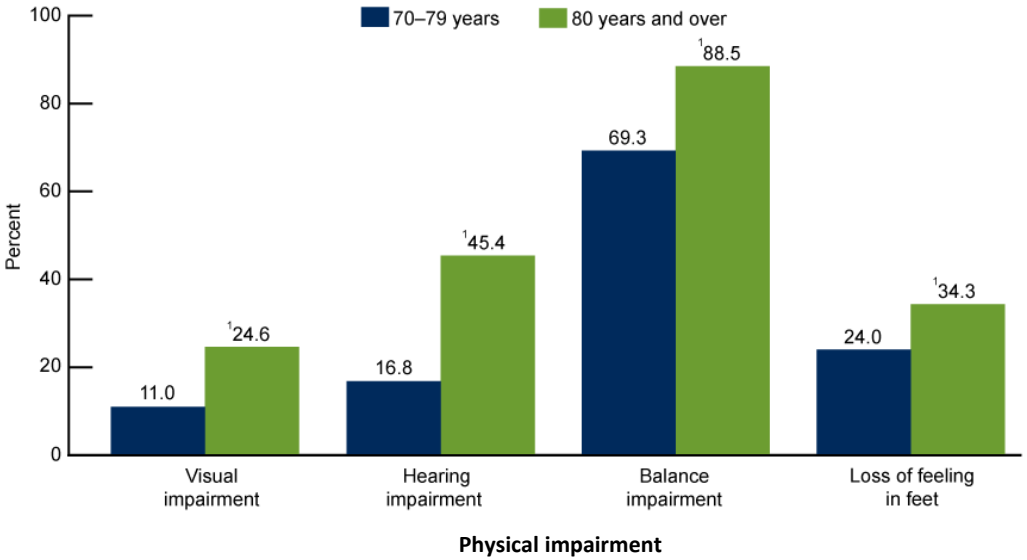
What are bar charts?

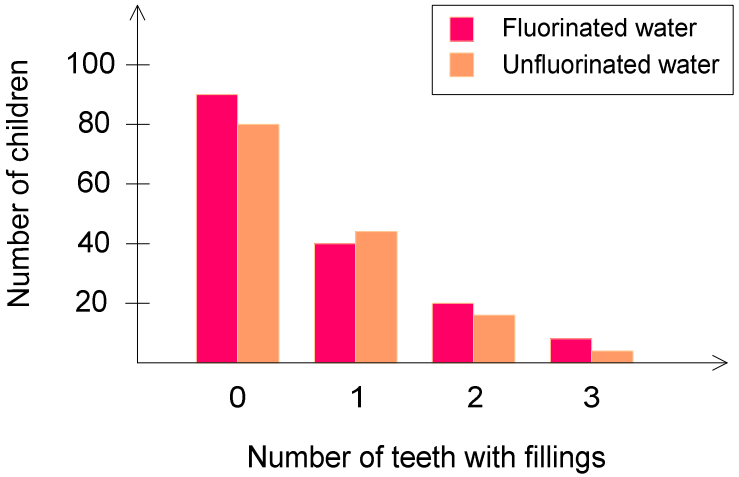
- The aim of this form of visualisation is to more easily *compare* the number of genes in each criteria.
- Hence the need for a well "formatted" chart:
  - Consistent width to each rectangle;
  - Leave a gap between bars so as to easily distinguish between these;
  - Start the vertical axis at y = 0.



Examples of bar charts

- Other bar charts are shown below.





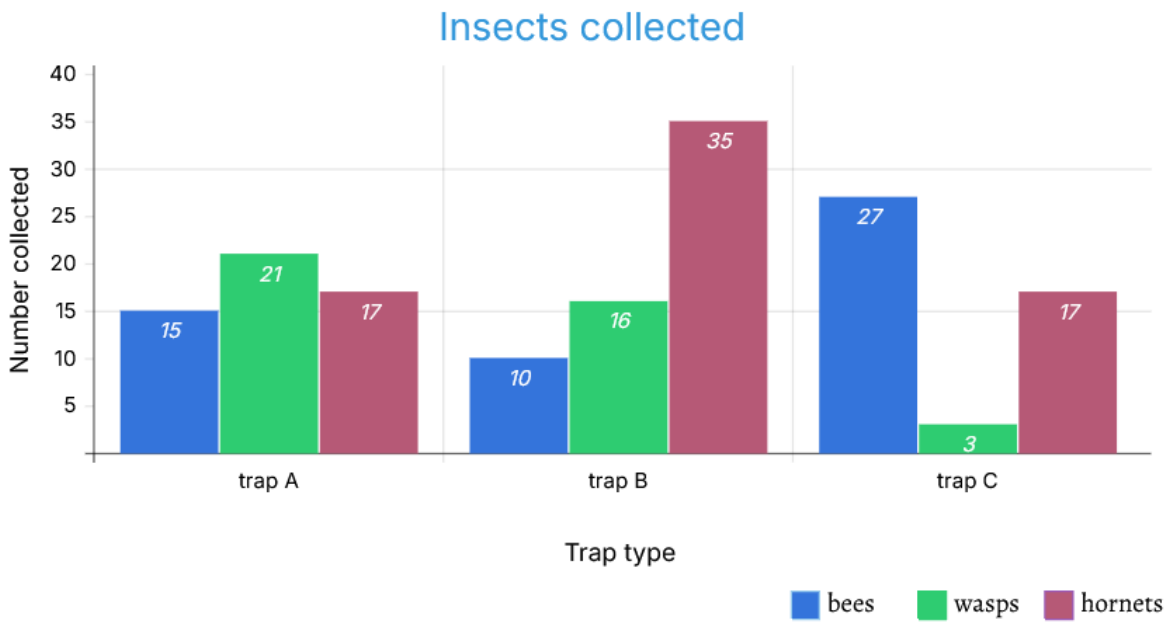
- In both of these charts we are comparing two different groups for each category: age groups and water treatment groups.

Examples of bar charts

- So, bar charts can be used not only for comparing different categories (such as different physical impairments), but also different groups within each category (such as age ranges for each physical impairment)

Examples of bar charts

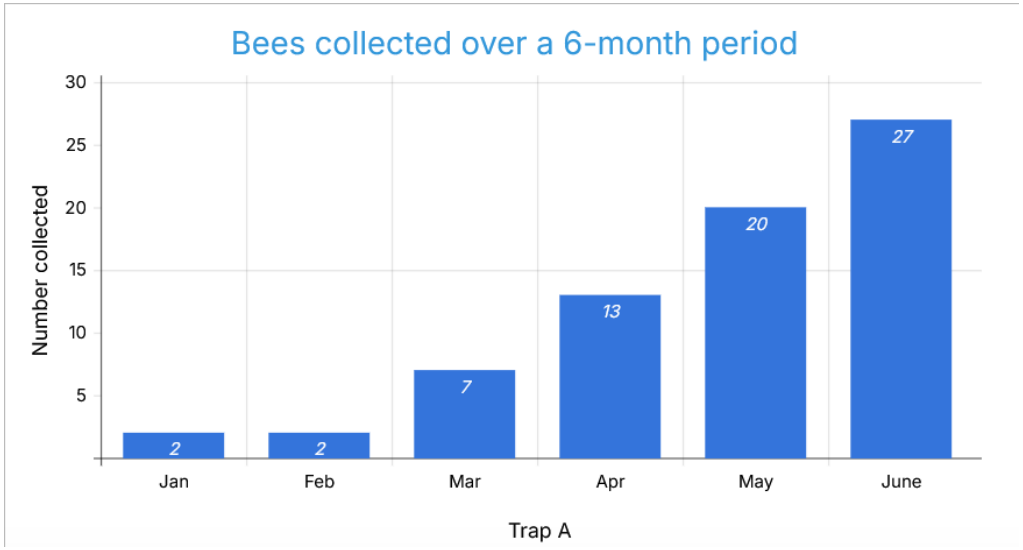
- As we have seen, bar charts are used to compare different categories of data, for example comparing the number of insects collected in three different traps



Bar chart showing the number of bees, wasps, and hornets each collected in three different traps

Examples of bar charts

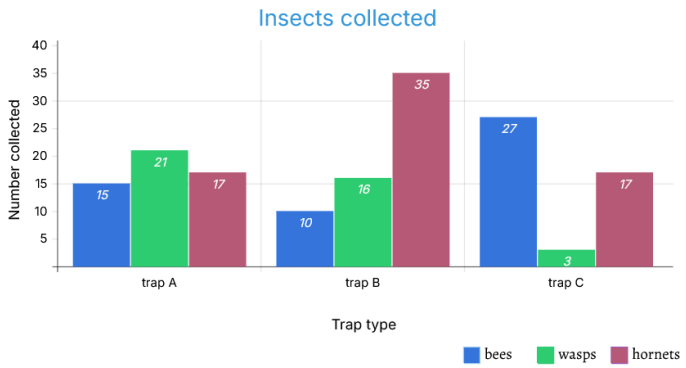
- But bar charts can also be used to track changes over discrete periods of time.
- Here the categories are no longer the traps, but the months for which one trap is used:



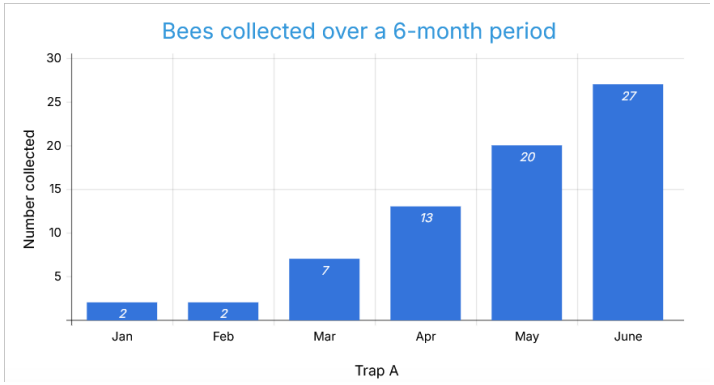
Bar chart showing the number of bees collected in trap A over a 6-month period

Examples of bar charts

- So in the first example we are comparing across traps (for one type of "animal", i.e. insects) and in the second example we are comparing across months (for one specific trap)



Bar chart comparing across traps as well as bees, wasps and hornets.



Bar chart comparing across months.

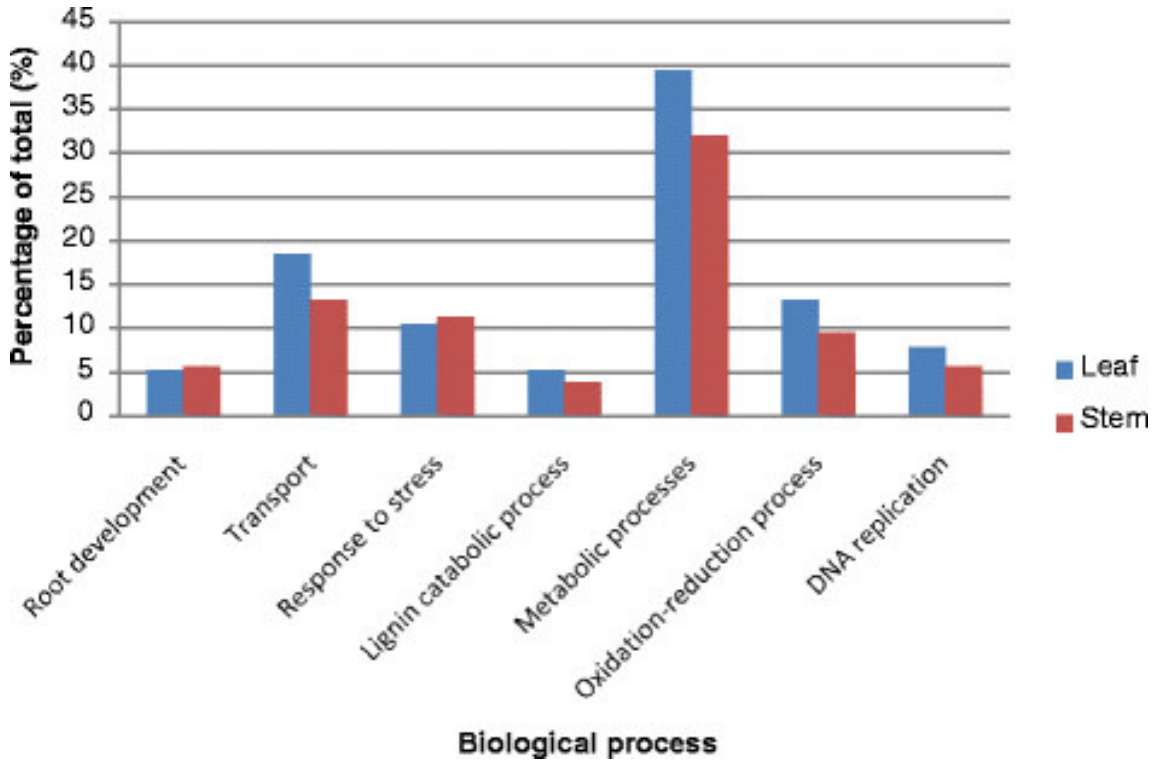
Examples of bar charts

- Bar charts are used a lot by professional biologists and biochemists for presenting and summarising data.
- The following bar charts comes from a journal paper called:

"Identification of four functionally important microRNA families with contrasting differential expression profiles between drought-tolerant and susceptible rice leaf at vegetative stage"

by Cheah, Boon Huat & Nadarajah, Kalaivani & Divate, Mayur & Wickneswari, R.. (2015), in BMC genomics. 16. 692. 10.1186/s12864-015-1851-3.

Examples of bar charts



Bar chart showing the distribution of the enrichment of biological processes between leaf and stem

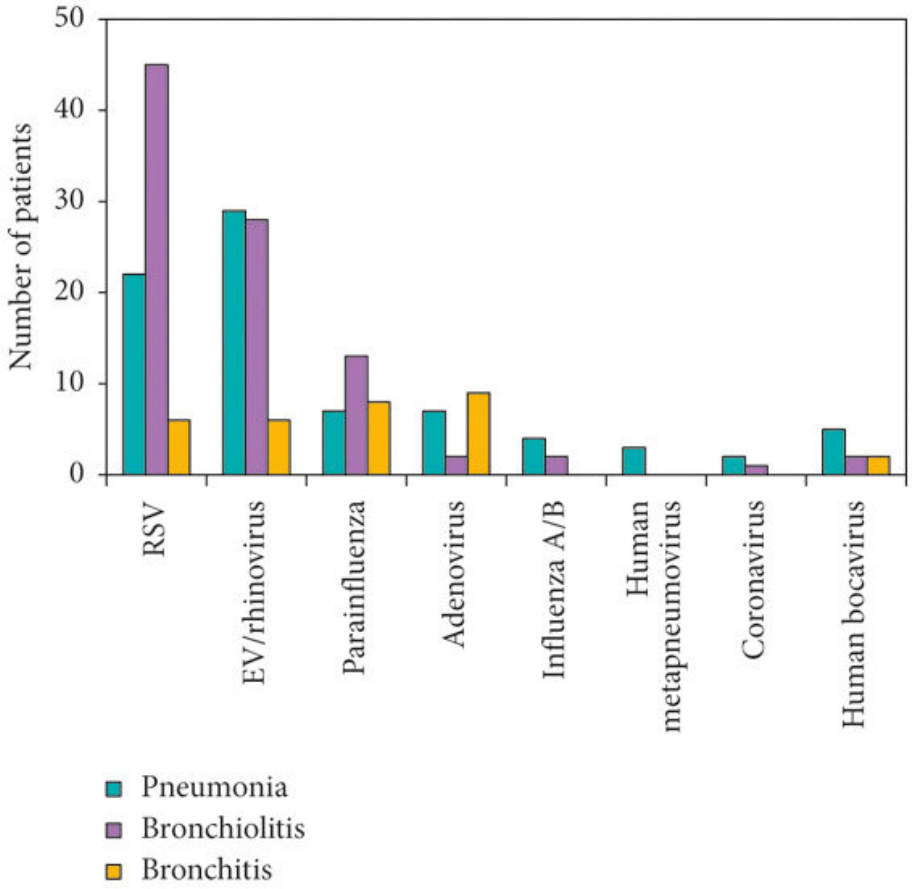
Examples of bar charts

- The following bar charts comes from a journal paper called:

"Clinical and Pathogenic Characteristics of Lower Respiratory Tract Infection Treated at the Vietnam National Children's Hospital",

by Hien, Pham & Nguyen, Phuc & Sinh, Tran & Phung, Thuy. (2020), Canadian Journal of Infectious Diseases and Medical Microbiology. 2020. 1-6. 10.1155/2020/7931950.

Examples of bar charts



The extended title of the chart is shown below. There is a problem with this title compared to what the chart is showing.
What is the problem?

Bar graph showing the respiratory viruses detected in each age group. The most abundant viruses were respiratory syncytial virus (RSV) and enterovirus (EV)/rhinovirus (Rhi), which mainly infected patients in the 2–5 months group. $*p<0.001$.

Examples of bar charts

- Examples of categories in biology for which bar charts can be used include:
  - Percentage germination of five different types of seeds over a 1 month period;
  - Bone density in children, teenagers, young adults and old adults;
  - Blood pressure in adults for different age groups

  What other categories can you think of that you could compare against each other, and present as a bar chart?

How to draw bar charts

- In one sense, bar charts are easy to draw. They simply consist of vertical rectangles spaced out along the horizontal axis.
- But it is still easy to make bad looking, or difficult to read, bar charts.
- An important point to remember is that bar charts are a visual form of data presentation and must therefore have a certain aesthetic, a certain visual appeal.

Bar charts must be clear and pleasing to the eye.

How to draw bar charts

- To achieve this make:

  1) your charts clean and clear;
  2) the bars must not be too wide or too narrow for the eye (this is an issue of aspect ratio);
  3) the space between the bars for separate categories must not be too wide or too narrow for the eye (this is an issue of aspect ratio);
  4) multiple bars within a category can be adjoined together (in this case you will have to colour each bar within that category);

How to draw bar charts

- To achieve this make:

5) label your bar chart;

6) label your horizontal and vertical axis;

7) label each category (i.e. trap A) and subcategory (i.e. bees);

8) put relevant values at the top of each bar (not all examples above do this, so how easy is it for you to read off the height of each bar?)

9) in general you should colour each bar according to its category and group within the category. BUT, see Karen and Alice for instructions when making hand-drawn bar charts.

Stacked bar charts (optional)

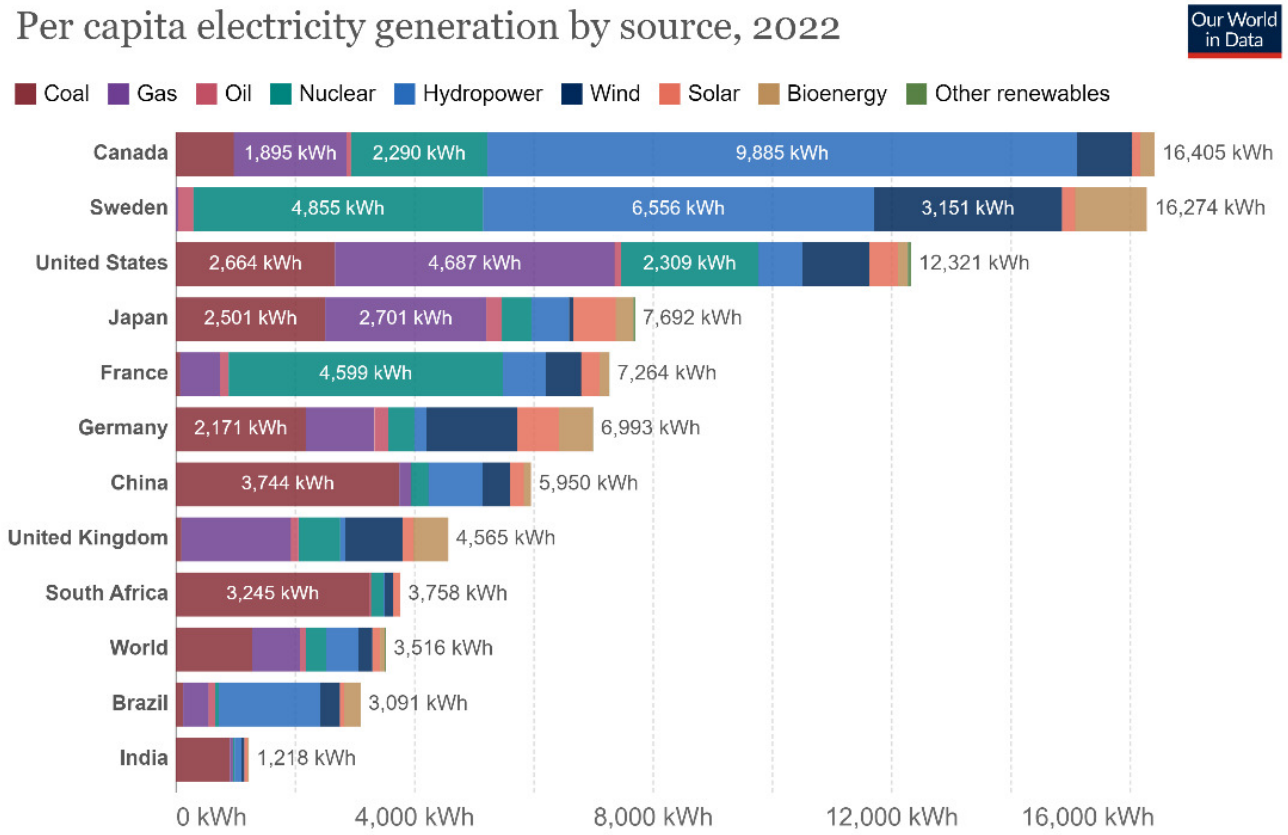
- Bar charts are useful for visually presenting a limited number of categories and groups within each category.
- But what if you have dozens of categories, each with half-a-dozen groups? This makes for a lot of bars to be drawn side by side.

Stacked bar charts (optional)

- The problem with this is that the bar chart will look far too cluttered or crowded.
- To overcome this we can draw something called a stacked bar chart.
- You will not need stacked charts for your UPCSE Biology course, since it is unlikely that you will collect data relating to so many categories and groups.
- But it is worth looking at these since they make it much easier to visually present huge amounts of data.

Stacked bar charts (optional)

- To explain the nature of stacked bar chart consider the chart here:



(from https://inforiver.com/insights/7-types-of-bar-charts-abcs-and-advantages/)

Stacked bar charts (optional)

- What we see here is that there are a lot of categories (countries) and each category has a lot of groups to it (coal, gas, oil, nuclear, etc.)
- Instead of drawing each category and group as bars side-by-side, it is much easier to stack the bars of each group.
- Canada alone has 8 different coloured bar. This means the category “Canada” would need 8 bars next to each other on the horizontal axis of a normal bar chart.
- Repeat this for the other 12 countries and we end up with a very cluttered bar chart.

Stacked bar charts (optional)

- So, here we see the reason for using stacked charts: if you have a lot of different categories (countries), each having multiple groups within them (different types of power generation), a normal bar chart (with bars side by side) would take too much space and would look too bunched up. It could therefore be visually too difficult to read.
- More importantly, the groups within a category actually denote the amount each group contributes to that category.

Stacked bar charts (optional)

- So in the case of the stacked chart above we can see more easily that, for Canada, of the total 16405 kWh of power generated, 1895kWh was generated by gas, 2290kWh was generated by nuclear energy, and 9885kWh was generated by hydropower.
- This makes the visual comparison of the proportion of each group (source of power generation) to the whole category (country) much easier.
- Then the total value of the groups should add up to the value of the category as a whole (can't be seen on the chart above).

Stacked bar charts (optional)

- As such a stacked bar chart allow us to compare totals across categories (here, countries) and simultaneously understand the values of contributing group (here, source of power generation).
- This presents a parts-to-whole view where the length of each segment is proportional to its contribution.
- It is then much easier to see the proportion of parts-to-whole than if you had bars next to each other, as in a standard bar chart (again, imagine placing side-by-side each coloured bar within each country instead of stacking the coloured bars).