Data Science – Data Visualization - Fundamentals

* Important charts *

Contents

1. Line graph	2
2. Bar graph/ chart	4
3. Double bar chart	
4. Box plot	
5. Scatter plots	10
6. Pie charts	12

* Important charts *

1. Line graph

- ✓ We can represent data using a line graphs.
- ✓ A line graph is formed by joining the points given by the data with straight lines.
- ✓ It is used to show the change of information over a period of time.
- ✓ This means that the horizontal axis is usually a time scale, for example minutes, days, months or years.

Example:

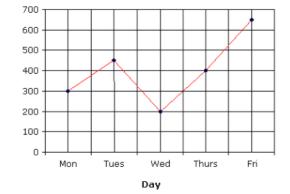
The table shows the daily earnings of a store for five days.

Day	Mon	Tues	Wed	Thurs	Fri
Earnings	300	450	200	400	650

- a) Construct a line graph for the frequency table.
- b) On which days were the earnings above \$ 400

Solution:

a)



Earnings

b) The earnings were above \$ 400 on Tuesday and Friday.

Example:

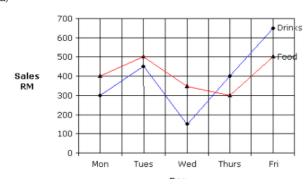
The table shows the daily sales in RM of different categories of items for five days.

Day	Mon	Tues	Wed	Thurs	Fri
Drinks	300	450	150	400	650
Food	400	500	350	300	500

- a) Construct a line graph for the frequency table.
- b) On what days were the sales for drinks better than the sales for food?
- c) What is the total earnings for food and drinks on Wednesday?

Solution:

a)



- b) Sales for drinks were better than sales for food on Thursday and Friday
- c) Total earnings for food and drinks on Wednesday is

$$150 + 350 = $500$$

2. Bar graph/ chart

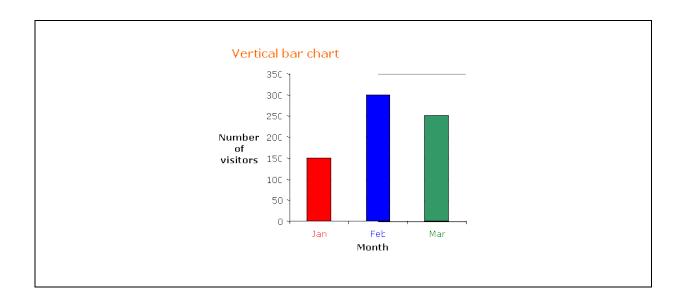
- ✓ A bar graph is a visual way to display and compare numerical data.
- ✓ It can represent the data as horizontal or vertical bars.
- ✓ The length of each bar is proportional to the amount that it represents.
- ✓ There are 3 main types of bar charts.
 - Vertical bar chart
 - Horizontal bar chart
 - Double bar chart

Example:

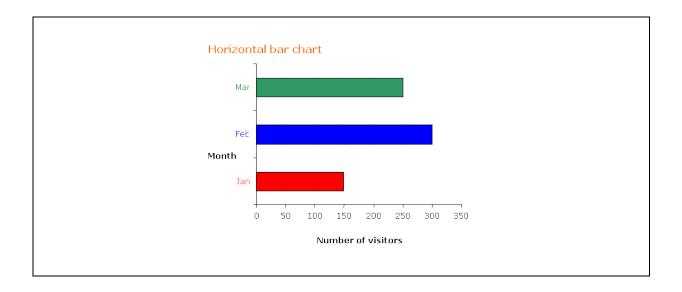
The following table shows the number of visitors to a park for the months January to March.

Month	January	February	March
Number of visitors	150	300	250

a) Construct a vertical and a horizontal bar chart for the table.



Data Science – Data Visualization - Fundamentals



3. Double bar chart

- ✓ The double bar chart is used when we want to represent two sets of data on the same chart.
- ✓ We can put the bars side by side

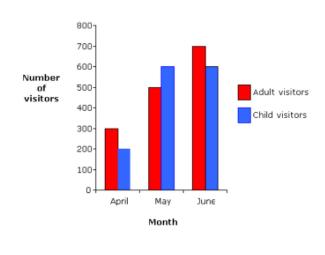
Example:

The following frequency graph shows the number of adult visitors and child visitors to a park. Construct a side by side double bar chart and a stacked double bar chart for the frequency table.

Month	April	May	June
Number of adult visitors	300	500	700
Number of child visitors	200	600	600

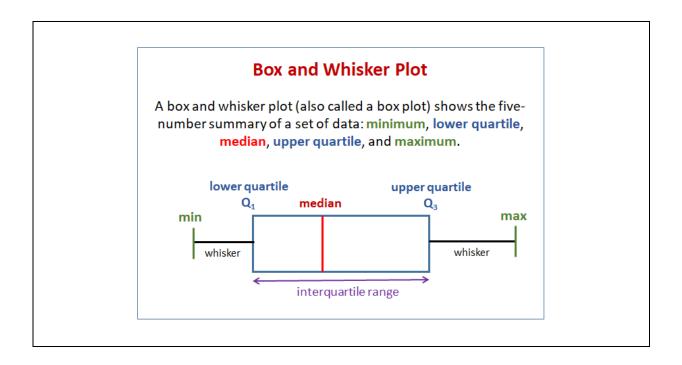
Solution:

Side By Side Double Bar Chart



4. Box plot

- ✓ A box plot (also called a box and whisker plot) shows data using the middle value of the data and the quartiles, or 25% divisions of the data.
- ✓ The following diagram shows a box plot or box and whisker plot.



Drawing a box and whisker plot

Example:

Construct a box plot for the following data:

12, 5, 22, 30, 7, 36, 14, 42, 15, 53, 25

Solution:

Step 1: Arrange the data in ascending order.

Step 2: Find the median, lower quartile and upper quartile

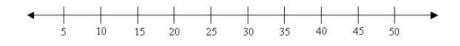
Median (middle value) = 22

Lower quartile (middle value of the lower half) = 12

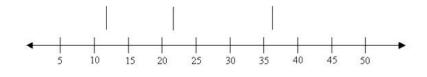
Upper quartile (middle value of the upper half) = 36

(If there is an even number of data items, then we need to get the average of the middle numbers.)

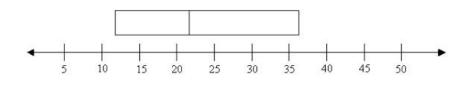
Step 3: Draw a number line that will include the smallest and the largest data.



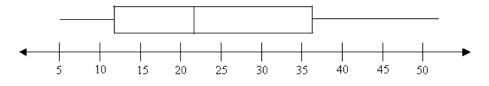
Step 4: Draw three vertical lines at the lower quartile (12), median (22) and the upper quartile (36), just above the number line.



Step 5: Join the lines for the lower quartile and the upper quartile to form a box.



Step 6: Draw a line from the smallest value (5) to the left side of the box and draw a line from the right side of the box to the biggest value (53).



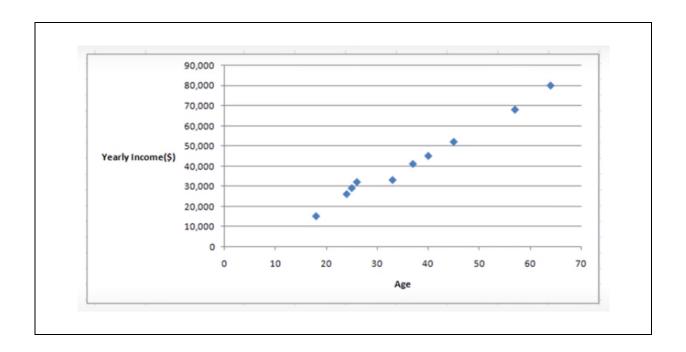
5. Scatter plots

- ✓ Scatter plots are method of graphically displaying bivariate data
- ✓ According to the given data we can find out the relationship in between age and income

Subject	Age	Yearly Income(\$)
1	18	15,000
2	25	29,000
3	57	68,000
4	45	52,000
5	26	32,000
6	64	80,000
7	37	41,000
8	40	45,000
9	24	26,000
10	33	33,000

- ✓ We have plotted Age on x-axis, Income on y axis
- ✓ So, Is there any relation between Age and Income
- ✓ Scatter plot clearly explains us as if age increases then Income also increasing

Data Science – Data Visualization - Fundamentals



6. Pie charts

- ✓ This is a special chart that uses "pie slices" to show relative sizes of data.
- ✓ It is a really good way to show relative sizes
- ✓ It is easy to see which movie types are most liked or not at a glance.

