Week 2: Descriptive Statistics

Unit 2: Tabular and Graphical Methods to Describe Data

openSAP
open.sap.com

THE BEST RUN SAP

Introduction to data tabulation

Univariate Table	Bivariate Table	Multivariate Table
This is also known as a **one-way table**. It shows information on a single variable. The table reveals the counts of cases for each category of a single variable.	This is also known as a **two-way table**. This table displays counts for two variables which are cross-tabulated to examine how one variable influences the other.	This is also known as a **three-way table**. This table introduces a third variable to explain the relationship between two variables that are cross-tabulated.
An example is a **simple frequency table**.	For instance, a table which explores how a person's gender may influence their preferred choice of payment.	For instance, to examine how income level influences the relationship between gender and payment method.
	An example is a **contingency table**.	An example is a **control table**.

Univariate table

- This simple univariate table example shows dress choice for 10 ladies:

Dress Color Choice	Red	Blue	Yellow
	5	3	2

Bivariate table

- Bivariate or two-way tables are ideal for analyzing relationships between categorical variables.
- The bivariate table below shows data on the leisure activity of 50 adults, with preferences broken down by gender.

Leisure Activity	Yoga	Football	Cycling	Total
Woman	16	8	6	30
Men	2	10	8	20
Total	18	18	14	50

Multivariate table

- Often, the behavior you are analyzing is too complicated to be studied with only two variables. Therefore you will want to consider sets of three or more variables (called **multivariate analysis**).

Voting Preference	Older			Younger		
	Male	Female	Total	Male	Female	Total
	%	%	%	%	%	%
Willing to vote for a woman	43.8	56.1	49.0	44.2	55.8	52.9
Not willing to vote for a woman	56.2	43.9	51.0	55.8	44.2	47.1
	100.0	100.0	100.0	100.0	100.0	100.0
	(240)	(180)	(420)	(120)	(360)	(480)

https://www.ssric.org/trd/modules/cowi/chapter4

Introduction to data visualization

Quantitative	Qualitative
Pie chart	Bar graph
Histogram	Pareto chart
Scatter plot	Heatmap

Quantitative data – Pie charts

City	Frequency	Percent
Oakville	3	11.54%
Mississauga	4	15.38%
Toronto	7	26.92%
Whitby	1	3.85%
Markham	2	7.69%
Brampton	4	15.38%
Burlington	3	11.54%
Pickering	1	3.85%
Vaughan	1	3.85%
Total	**26**	**100.00%**
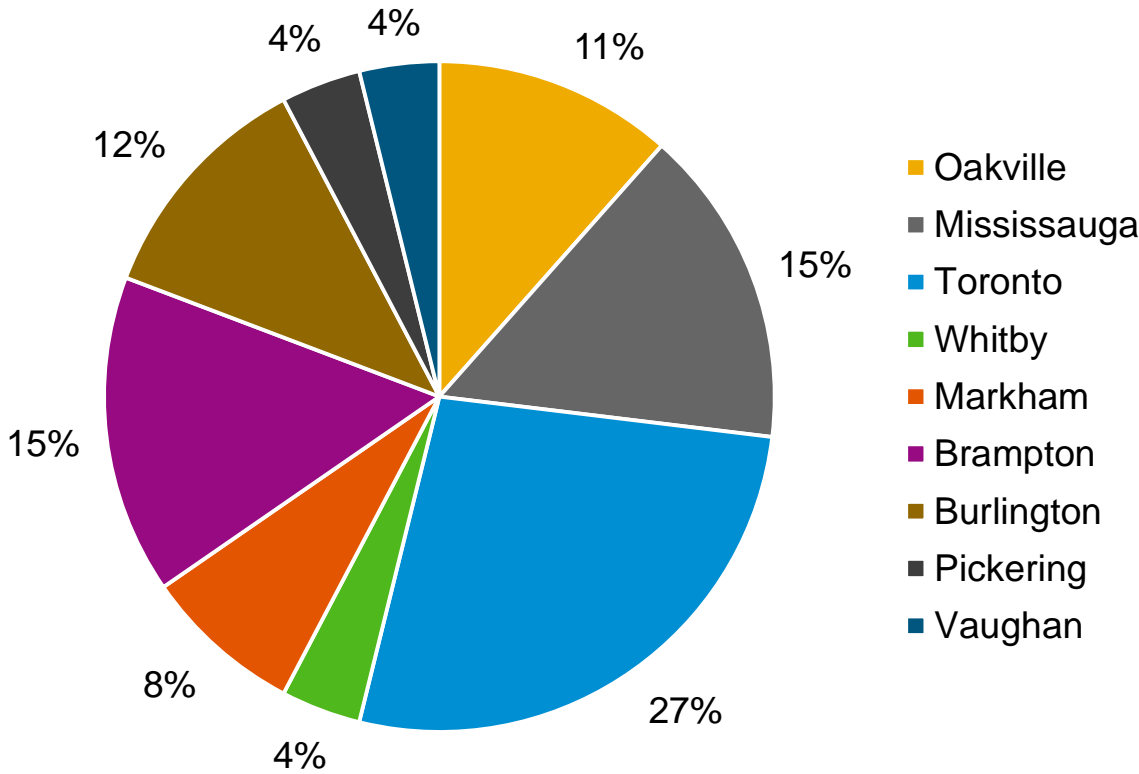
Quantitative data – Histogram

Group	Frequency	%	Cum. %
100 - 350	2	6.7%	6.7%
350 - 600	4	13.3%	20.0%
600 - 850	8	26.7%	46.7%
850 - 1100	9	30.0%	76.7%
1100 - 1350	5	16.7%	93.4%
1350 - 1600	2	6.7%	100%

Quantitative data – Determining the number of classes in a histogram
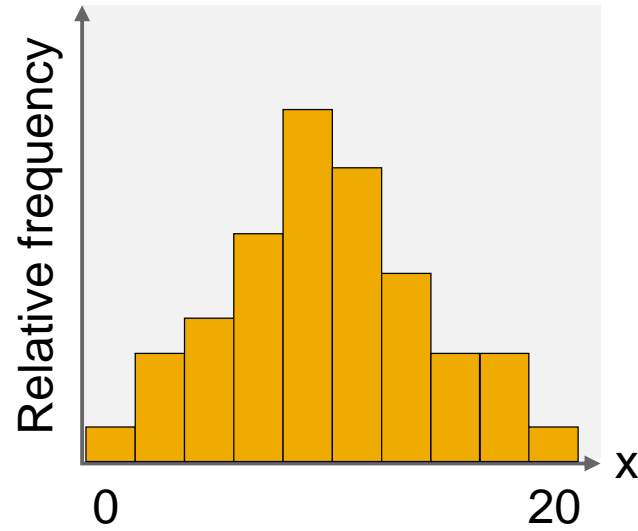
Measurement classes
a) Small data set
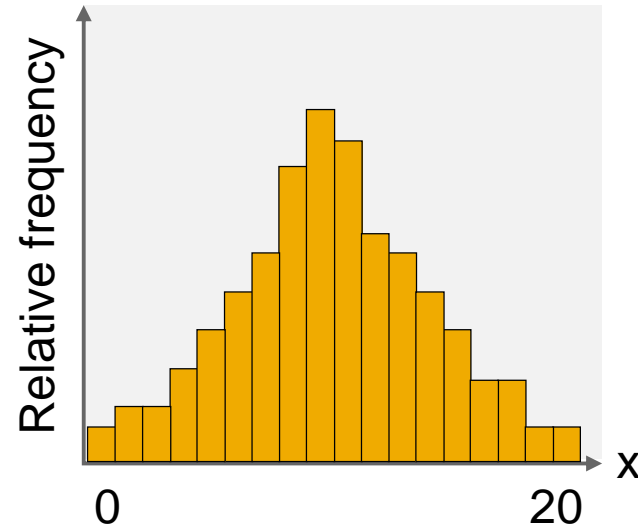
Measurement classes
b) Larger data set

Measurement classes
c) Very large data set

Number of Observations in Dataset	Number of Classes
Fewer than 25	5-6
25-50	7-14
More than 50	15-20

Quantitative data – Scatter plot
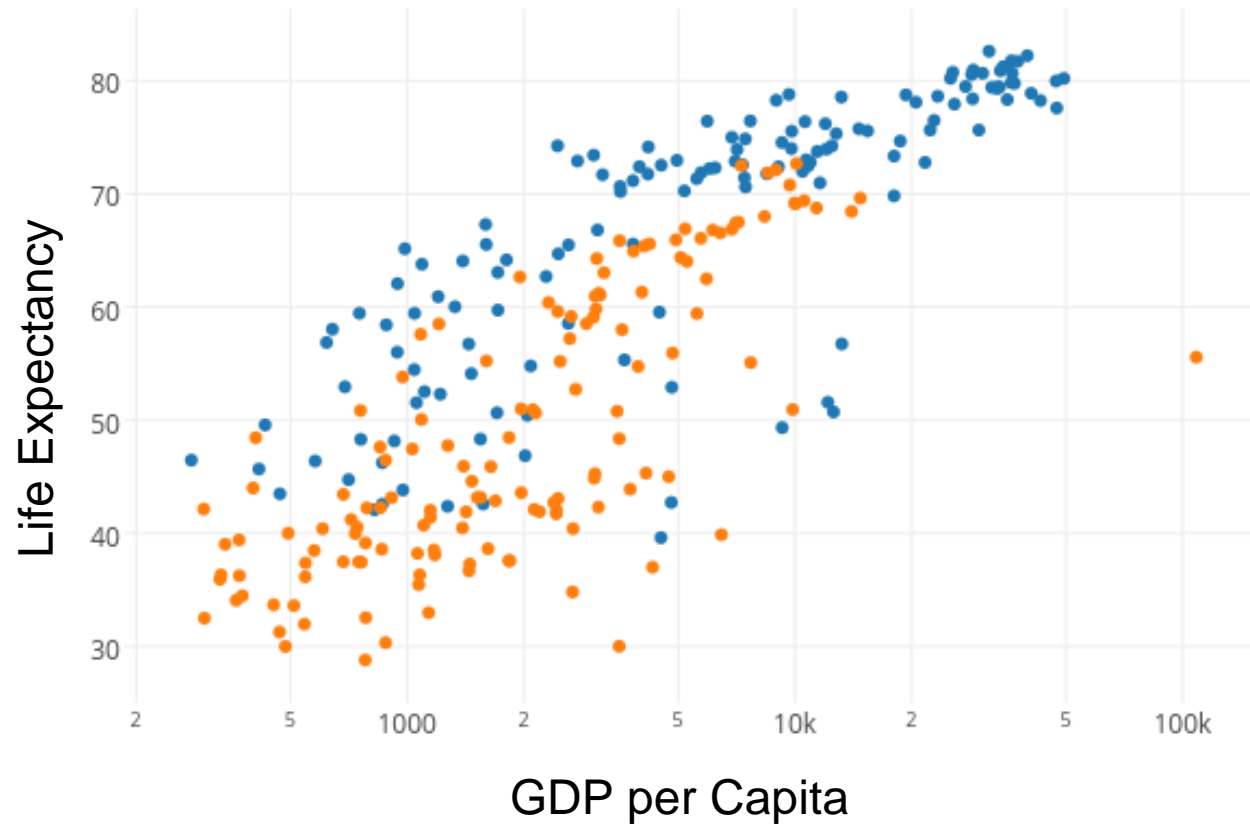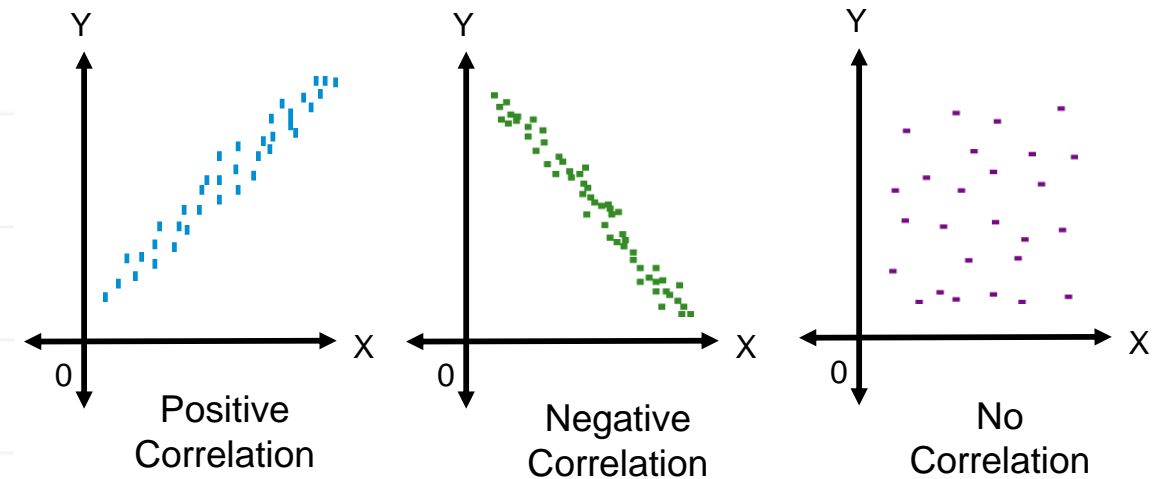
Qualitative data – Bar graph

Gaps

No Gaps
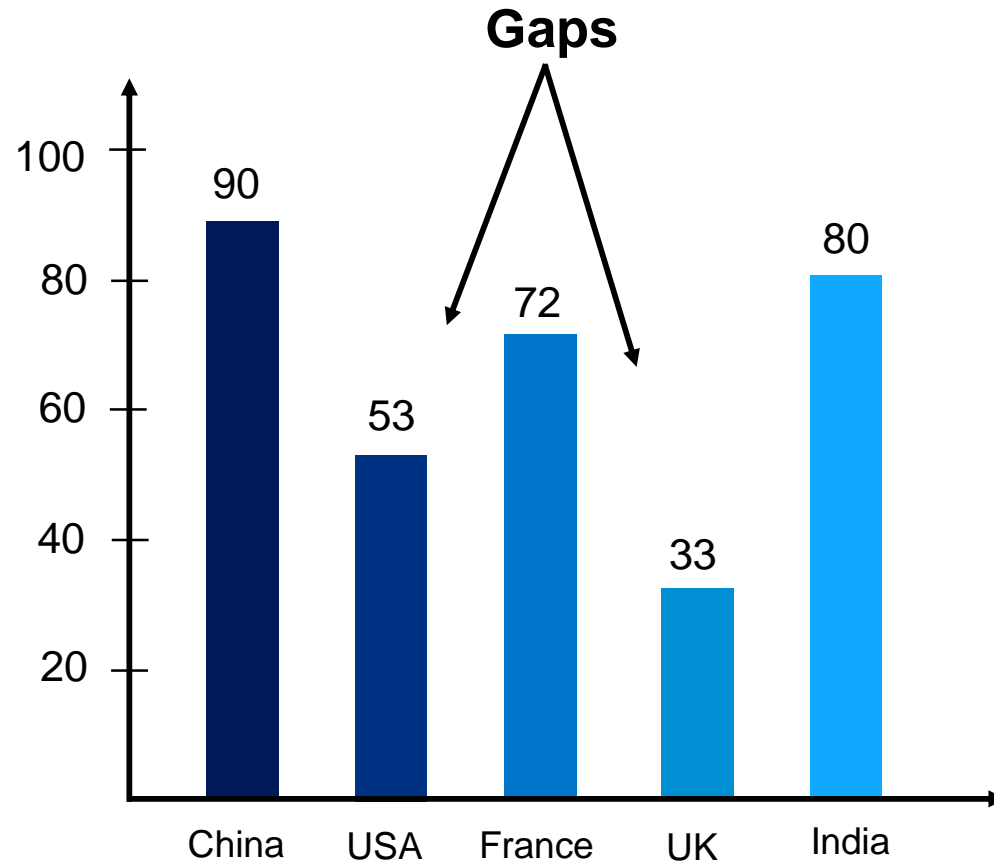
Categories

Bar Chart
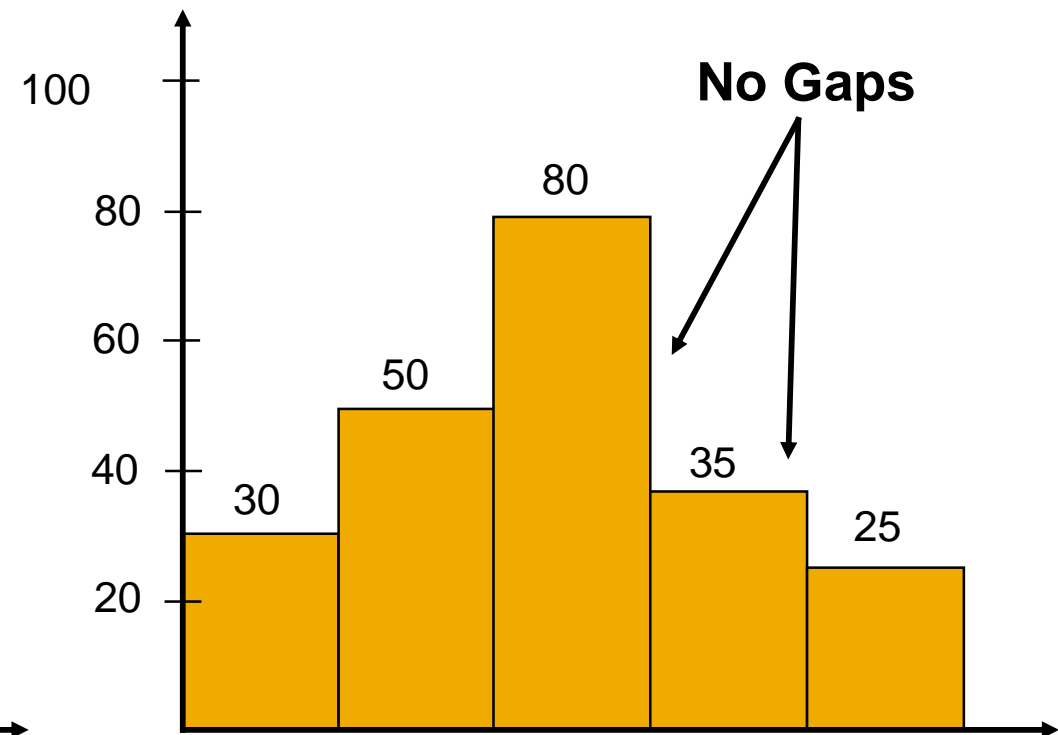
Number Ranges

Histogram

Qualitative data – Pareto chart

Category	Frequency	%	Cum. %
Canada	2	6.7%	6.7%
USA	4	13.3%	20.0%
France	8	26.7%	46.7%
China	9	30.0%	76.7%
India	5	16.7%	93.4%
UK	2	6.7%	100%
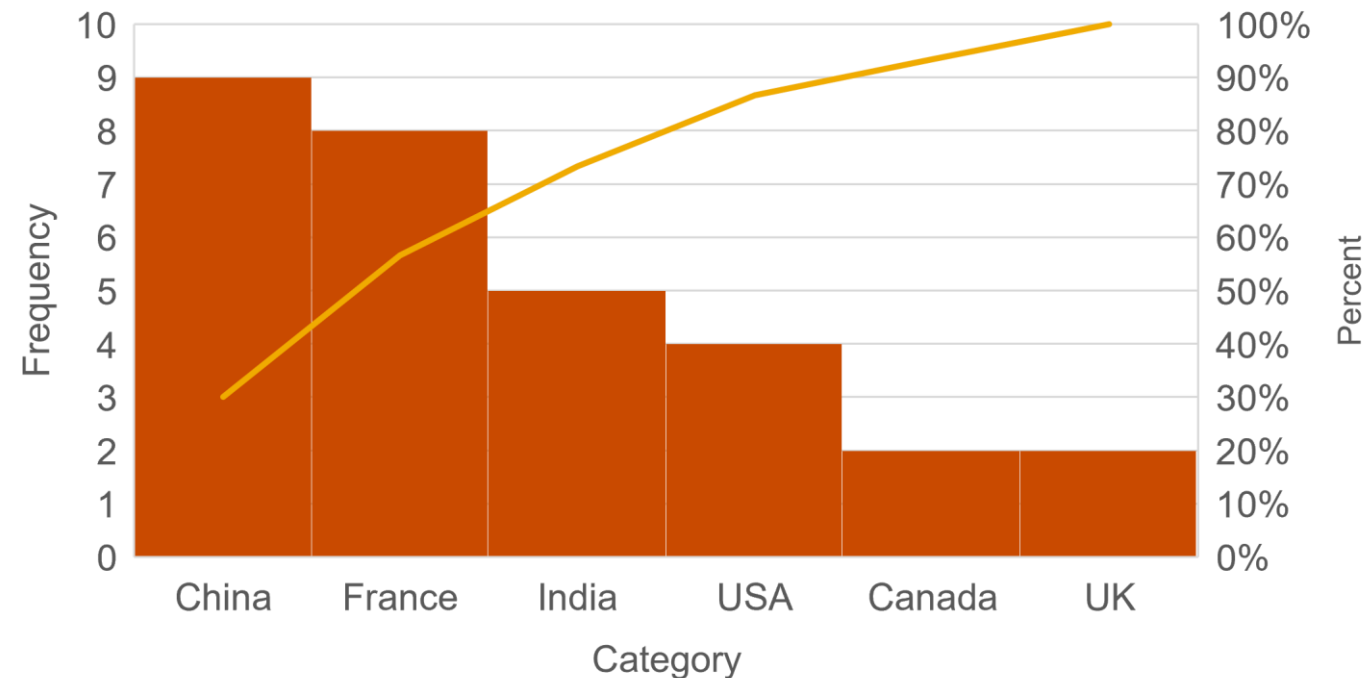
Qualitative data – Heatmap
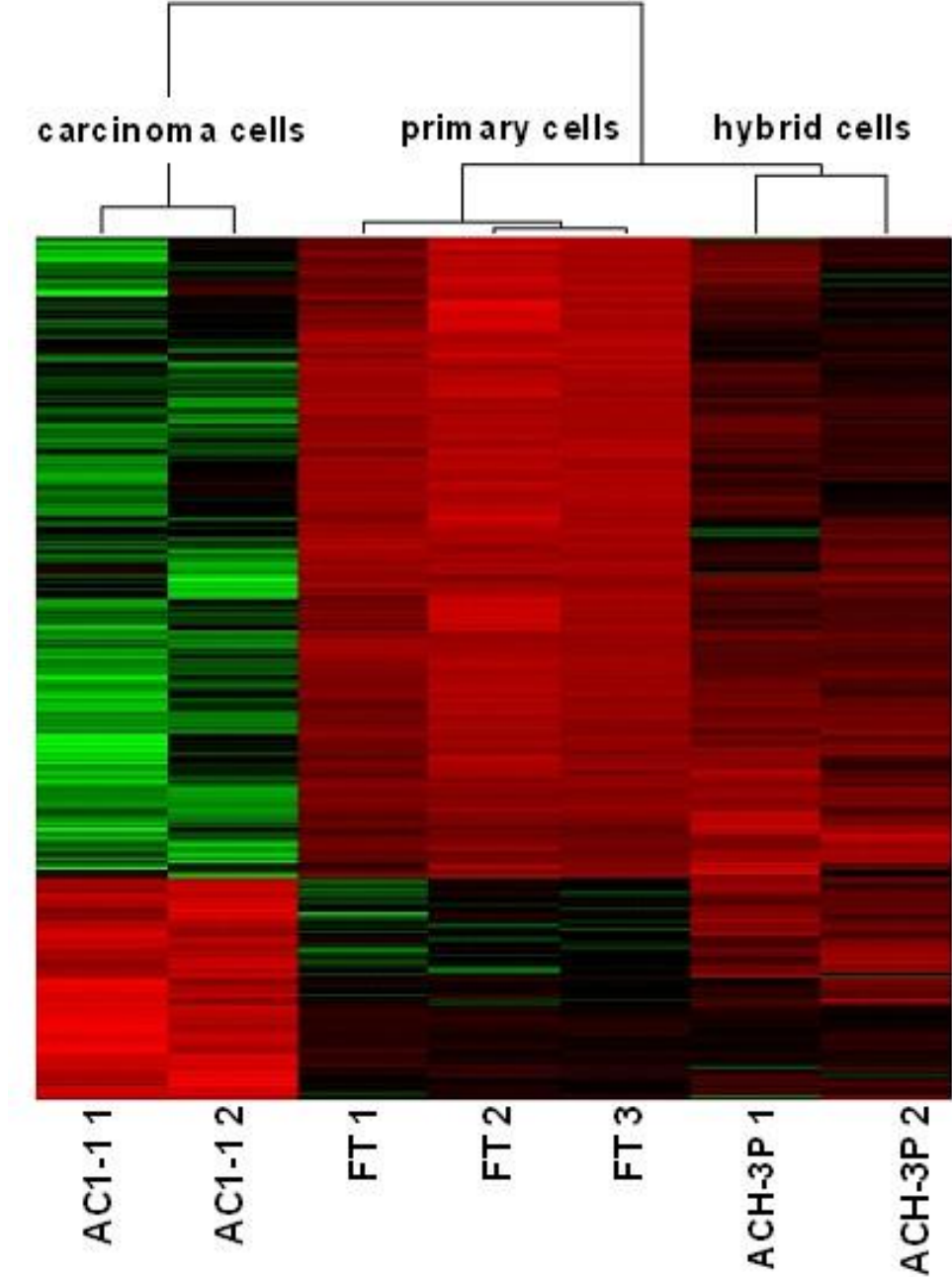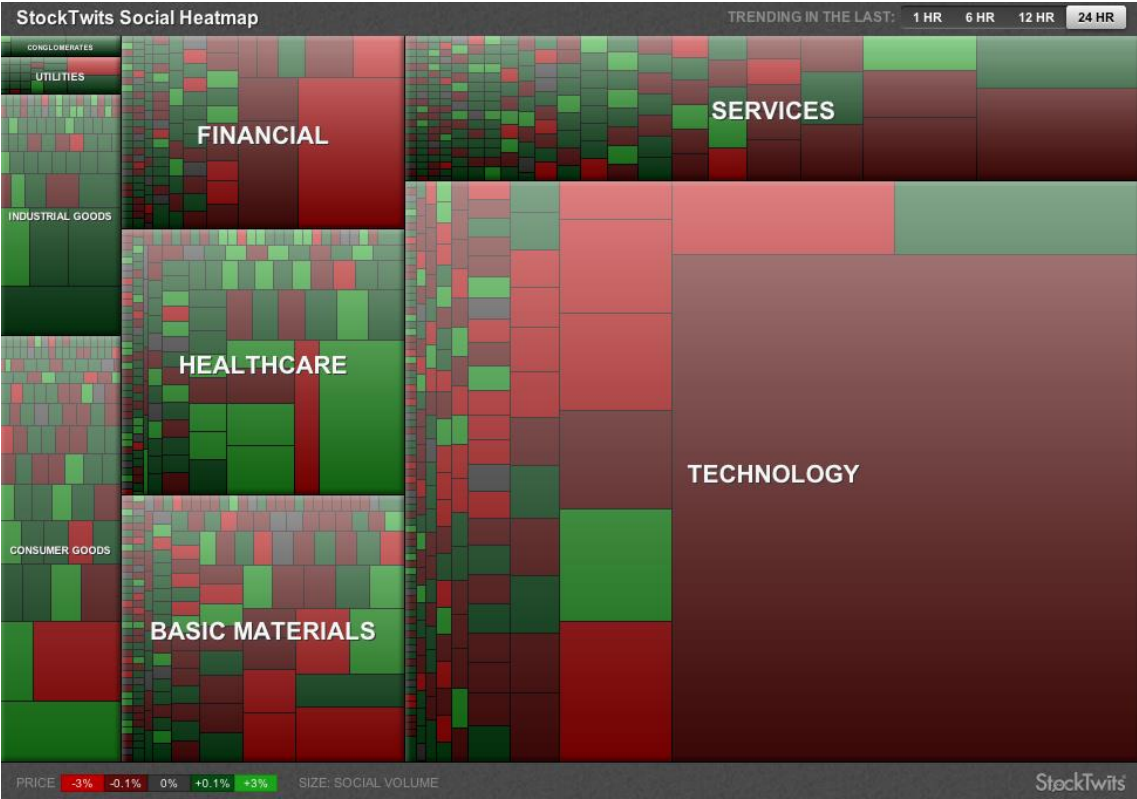
Summary

- You have learned about the different types of table you can use (univariate, bivariate, and multivariate tables) to organize and present your data.

- You have also seen which visualizations you should choose if the data is quantitative or qualitative.

Thank you.

Contact information:

open@sap.com

Follow all of SAP

www.sap.com/contactsap

THE BEST RUN SAP