Introduction to Data Visualization (I)

IT1164/IT1364/IT1564/IT1664/IT1864/IT1964

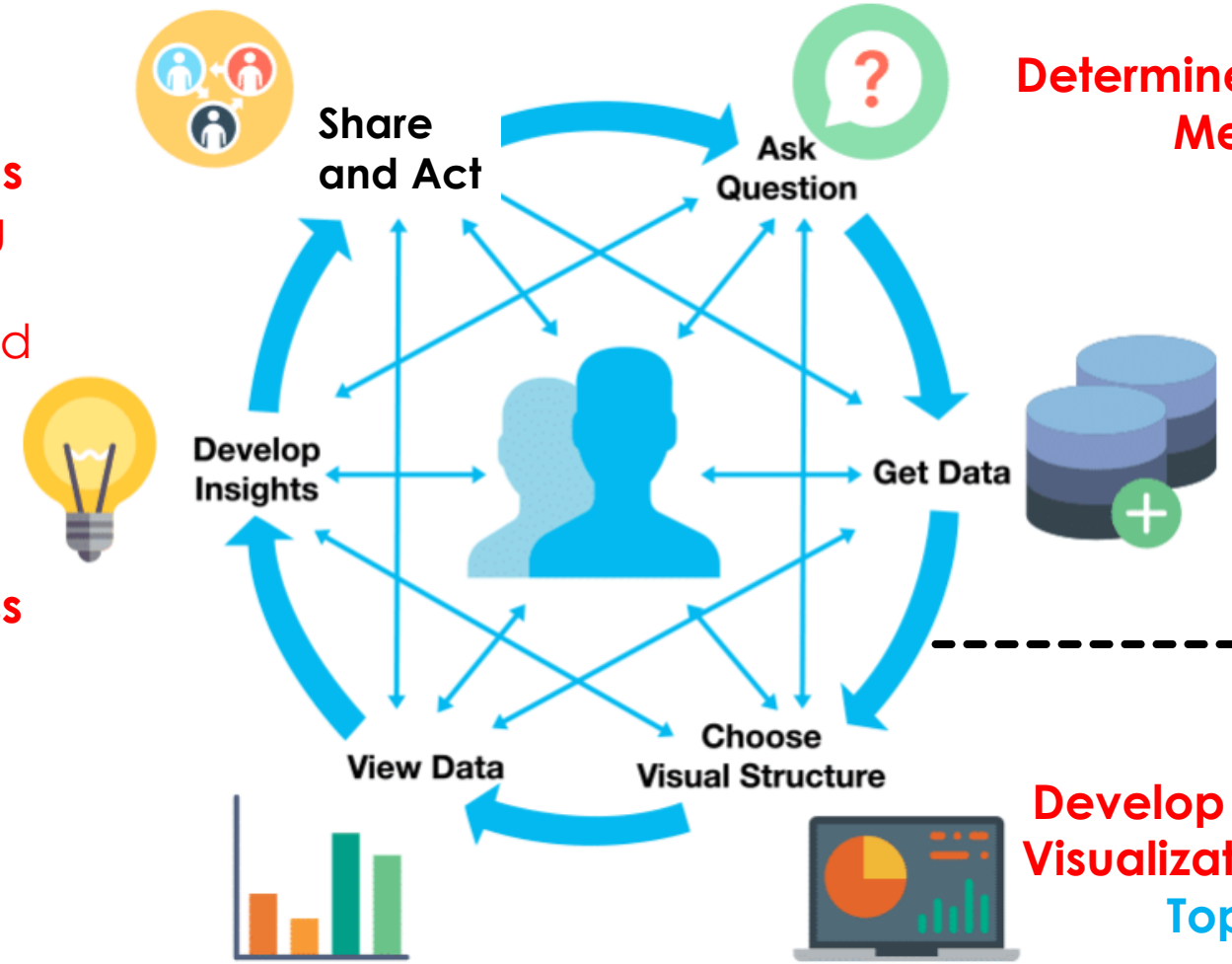
Learning Outcomes

1. Define data visualisation and explain the visualisation process
2. Describe the seven visual variables used in mapping data
3. Identify good visual representation

The Cycle of Visualisation

Publish the results for others to view.
Conduct investigations to find the underlying reason for the data trends you've identified
Topic 3

Develop Insights About Your Business



Determine What You're Trying to Measure and Why
Topic 1

Get the Data that Answers Those Questions

Pre-processing
Topic 2

Develop Your Data Visualization Design
Topic 3

Determine what you are trying to measure

- Identifying key focus or objective for your analysis is the first step before any visualisation can be done.
- The purpose of an analysis can be used to (1) understand the problem or (2) to propose a solution or both.

3 Steps process to identify key focus of analysis

Step 1 – Identify and select data fields that are 'usable' for analysis

Usable fields??

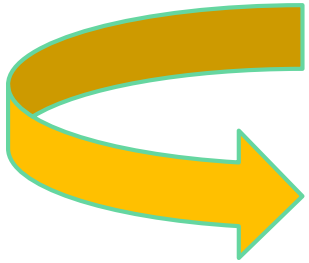
Step 2 – Formulate possible ways to analyze the data using the selected data fields.

Step 3 – Identify one or two key focus of the analysis

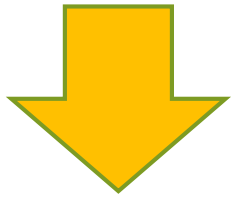
Key focus in analysis should be Concise, Specific and Measurable

What is Visualisation?

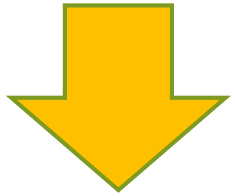
Graphical display of abstract information for data analysis and communication.



To discover and understand the patterns in our data



Making sense of complex data



Present visually to others

Basic Concepts

- Understand the data you are trying to visualise, including data size and the data preparation effort that will be required.
- Determine what you are trying to visualise and what kind of information you want to communicate.
- Know your audience and understand how it processes visual information.
- Use a visual that conveys the information in the best and simplest form for your audience.

Purpose of Data Visualisation

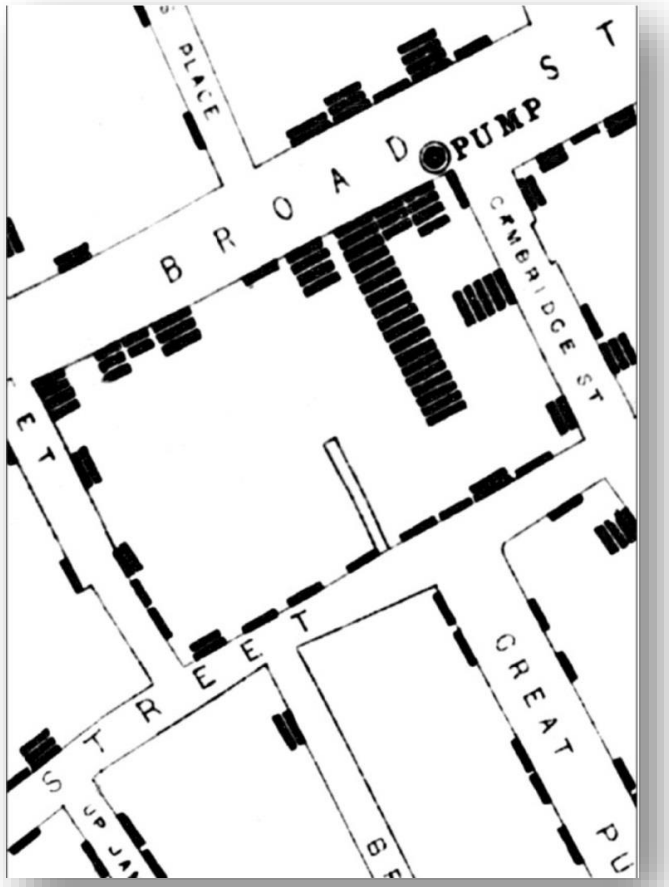
Is it a process of making beautiful graphic or image?

Analysis – Understand your data better and act upon that understanding.

Presentation – Communicate and inform others more effectively.

Infographics?

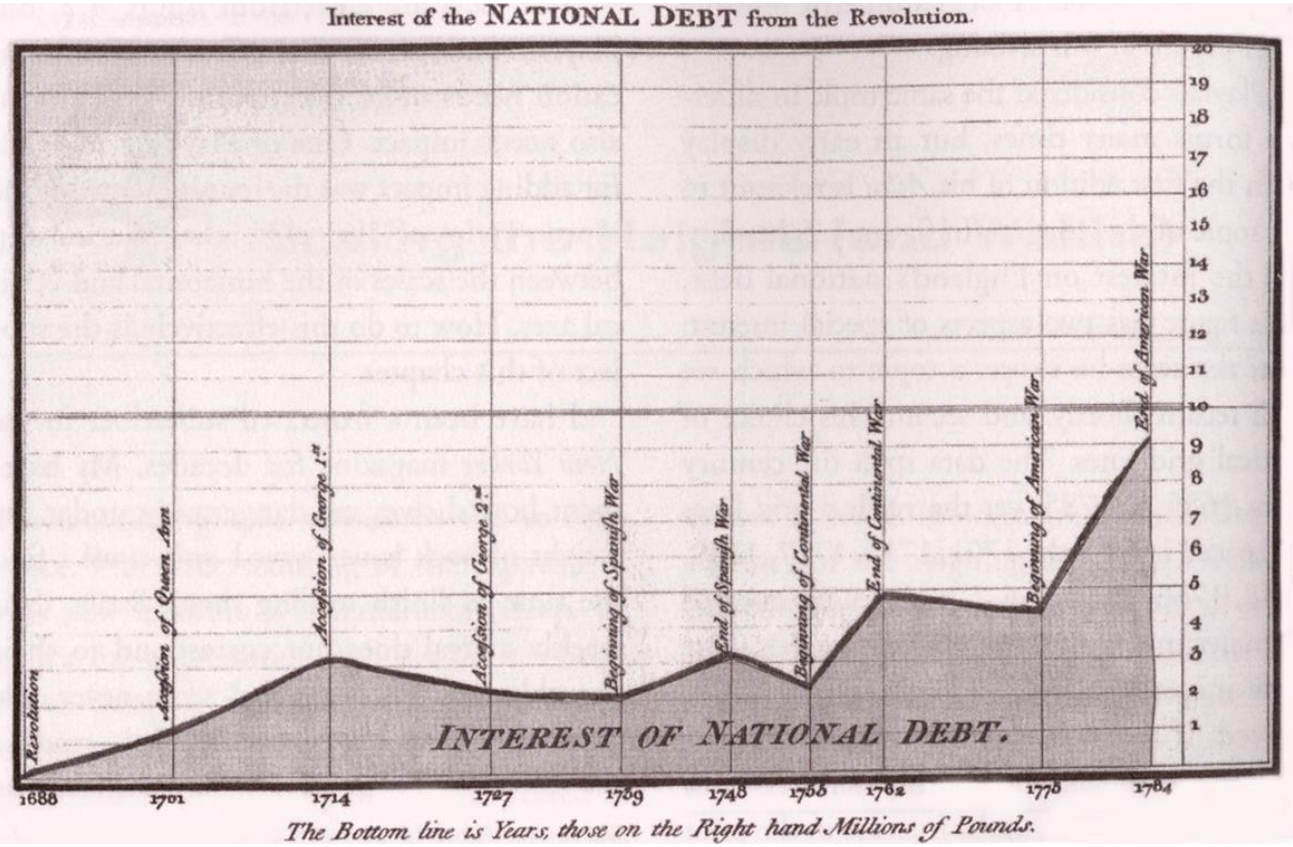
Early Visualisation



A section of John Snow's map of the deaths from cholera in London in 1663.

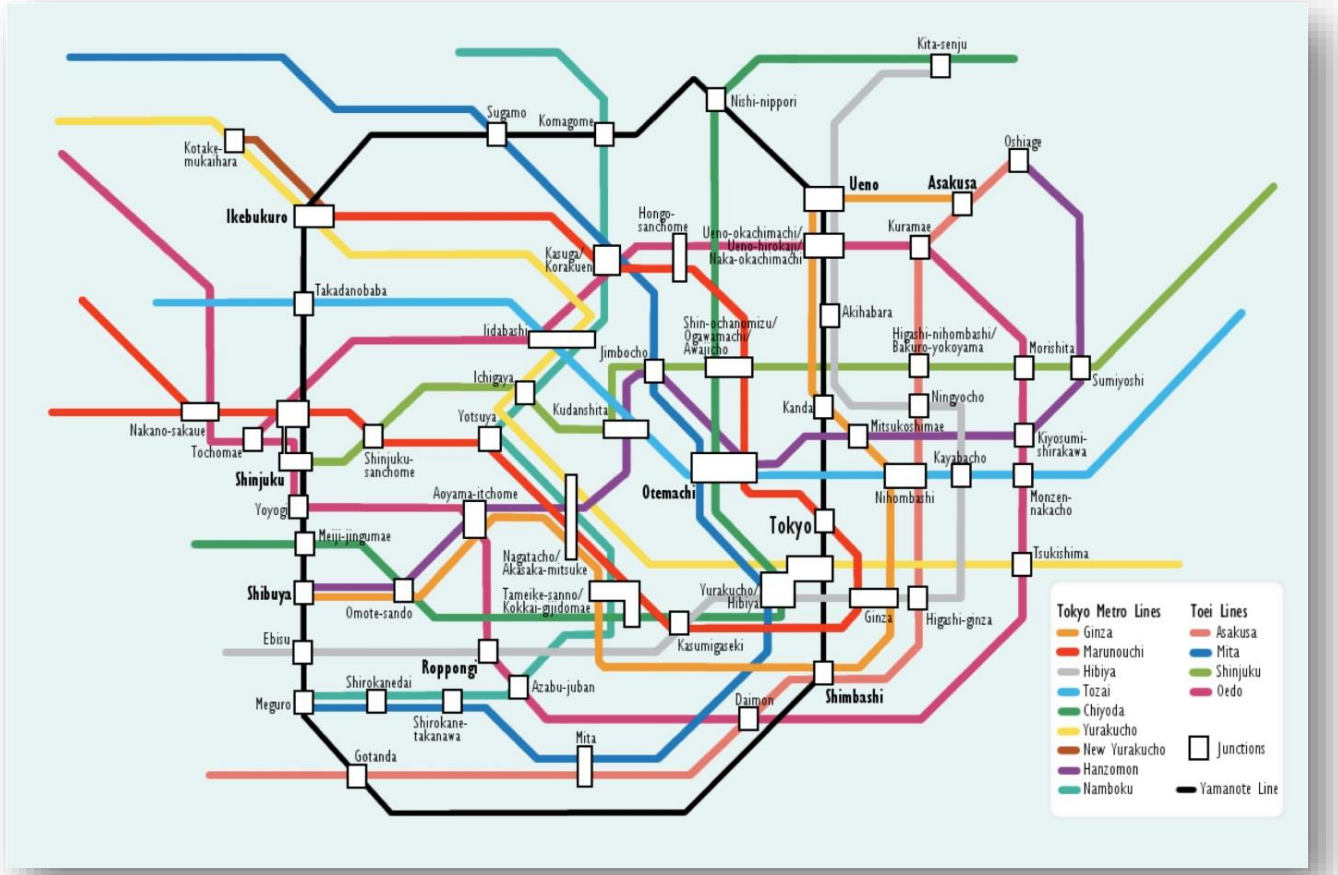
Each bar within the houses represents one deceased individual.

Early Visualisation example



William Playfair, a Scottish social scientist, is the inventor for bar graph, line graph and pie chart in late 18th century. This is one of his original graphs.

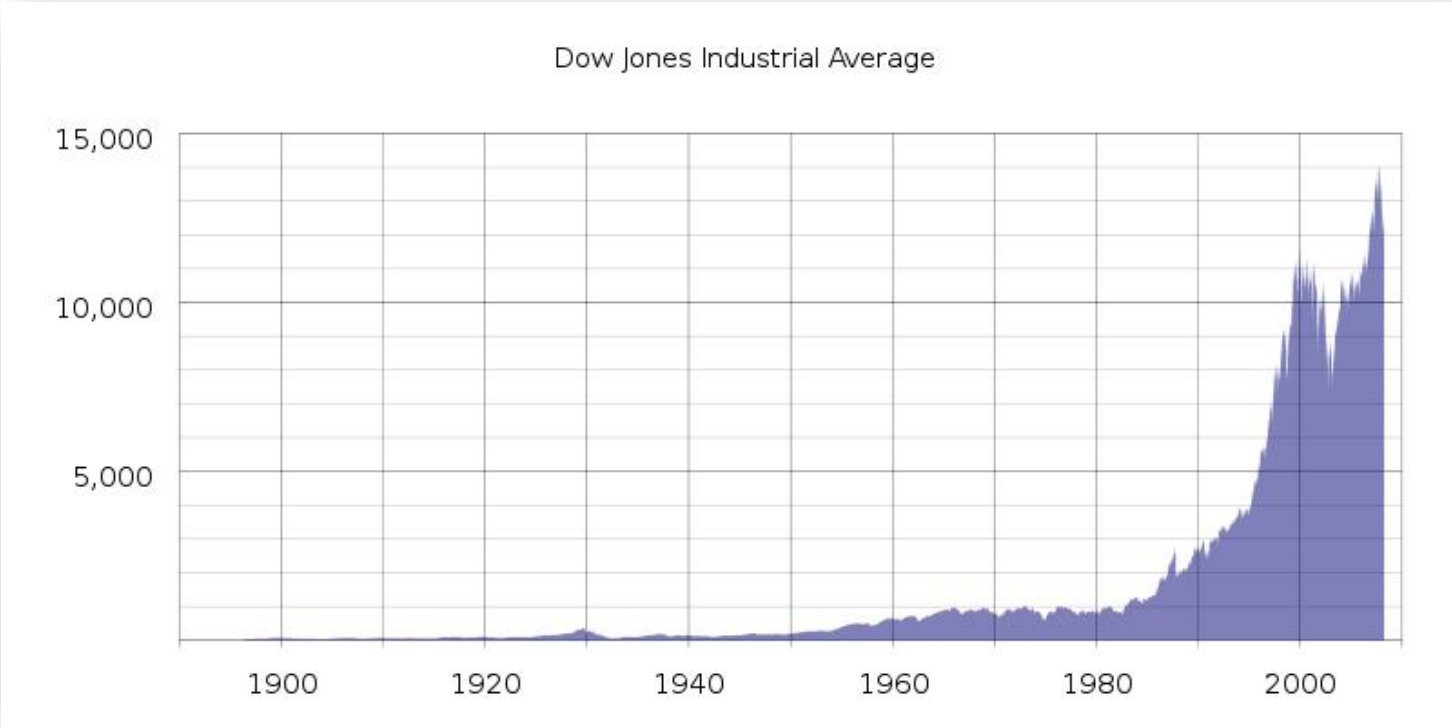
Visualisation Today



The Tokyo Underground Map

A logical representation of the metro highlighting qualitative relationships between the stops.

Visualisation Today



Dow Jones Industrial Average (DJIA) from 1900 to 2000

The Dow Jones Industrial Average is a U.S. stock index based on the weighted average of the stock prices of 30 large and actively traded U.S. companies.

Why Data Visualisation?

Importance of Data Visualization

- The way human brain processes information – Visual vs Text
- An easier and quick way to convey abstract concepts.

A picture is worth a thousand words!

Using the right charts to tell story about your data.

Through data visualization you can easily:

1. Visualise data (make sense of data, especially big data)
2. Classify and categorise data
3. Find relationship among data
4. Understand the composition of data
5. Understand the distribution of data
6. Understand the overlapping of data
7. Determine patterns and trends
8. Detect outliers and other anomalies in data
9. Predict future trends
10. Tell meaningful and engaging stories to decision makers

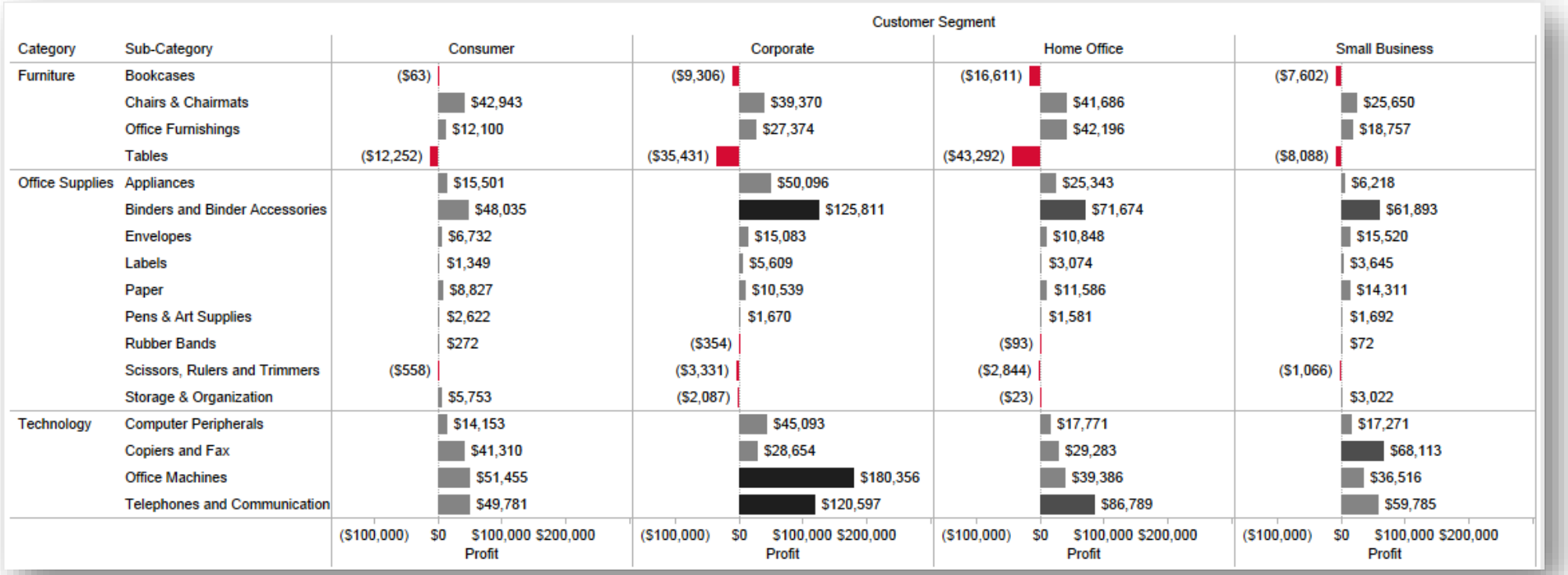
Can you "see" the most unprofitable product subcategory?

		Customer Segment			
Category	Sub-Category (group)	Consumer	Corporate	Home Office	Small Business
Furniture	Bookcases	-63.02	-9,305.76	-16,610.95	-7,602.40
	Chairs & Chairmats	42,942.97	39,370.10	41,686.28	25,650.38
	Office Furnishings	12,099.80	27,374.47	42,196.25	18,757.40
	Tables	-12,251.51	-35,430.73	-43,292.40	-8,087.89
Office Supplies	Appliances	15,501.48	50,095.94	25,343.06	6,217.58
	Binders and Binder Ac..	48,035.27	125,811.27	71,674.19	61,892.69
	Envelopes, Labels, Pa..	16,907.52	31,230.67	25,508.13	33,476.65
	Pens & Art Supplies	2,621.68	1,670.40	1,580.82	1,691.88
	Rubber Bands	271.85	-353.54	-93.12	72.14
	Scissors, Rulers and ..	-558.10	-3,330.62	-2,844.06	-1,066.47
	Storage & Organization	5,752.65	-2,086.83	-23.24	3,021.57
Technology	Computer Peripherals	14,152.79	45,092.93	17,771.05	17,270.71
	Copiers and Fax	41,310.35	28,654.48	29,283.14	68,113.50
	Office Machines	51,454.78	180,356.22	39,386.23	36,515.70
	Telephones and Com..	49,781.48	120,596.92	86,788.72	59,784.52

Can you "see" the most unprofitable product subcategory better?

		Customer Segment			
Category	Sub-Category	Consumer	Corporate	Home Office	Small Business
Furniture	Bookcases	-63.02	-9,305.76	-16,610.95	-7,602.40
	Chairs & Chairmats	42,942.97	39,370.10	41,686.28	25,650.38
	Office Furnishings	12,099.80	27,374.47	42,196.25	18,757.40
	Tables	-12,251.51	-35,430.73	-43,292.40	-8,087.89
Office Supplies	Appliances	15,501.48	50,095.94	25,343.06	6,217.58
	Binders and Binder Ac..	48,035.27	125,811.27	71,674.19	61,892.69
	Envelopes	6,731.55	15,082.58	10,848.34	15,520.13
	Labels	1,349.23	5,608.87	3,073.87	3,645.20
	Paper	8,826.74	10,539.22	11,585.92	14,311.32
	Pens & Art Supplies	2,621.68	1,670.40	1,580.82	1,691.88
	Rubber Bands	271.85	-353.54	-93.12	72.14
	Scissors, Rulers and ..	-558.10	-3,330.62	-2,844.06	-1,066.47
	Storage & Organization	5,752.65	-2,086.83	-23.24	3,021.57
Technology	Computer Peripherals	14,152.79	45,092.93	17,771.05	17,270.71
	Copiers and Fax	41,310.35	28,654.48	29,283.14	68,113.50
	Office Machines	51,454.78	180,356.22	39,386.23	36,515.70
	Telephones and Com..	49,781.48	120,596.92	86,788.72	59,784.52

We're Faster When We Can "See" Data



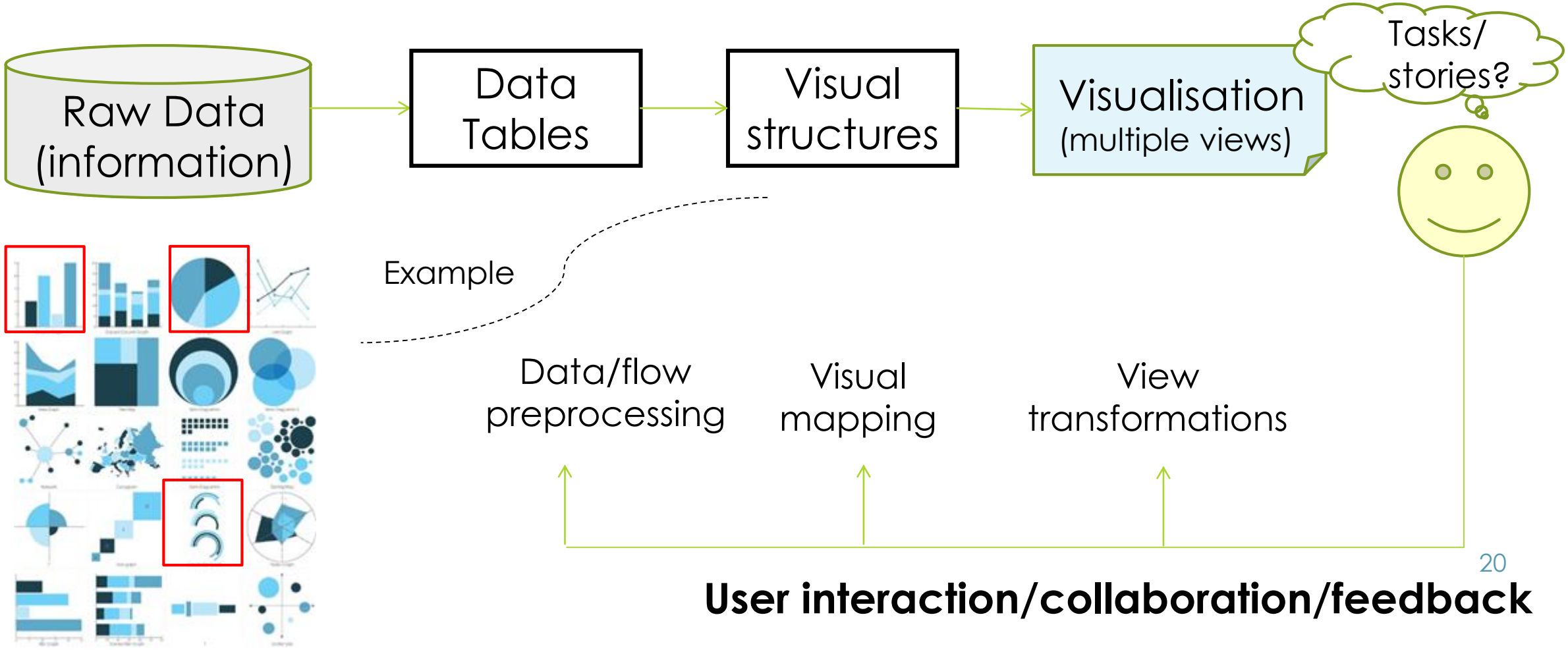
Customer Segment

Category	Sub-Category	Consumer	Corporate	Home Office	Small Business
Furniture	Bookcases	($63)	($9,306)	($16,611)	($7,602)
	Chairs & Chairmats	$42,943	$39,370	$41,686	$25,650
	Office Furnishings	$12,100	$27,374	$42,196	$18,757
	Tables	($12,252)	($35,431)	($43,292)	($8,088)
Office Supplies	Appliances	$15,501	$50,096	$25,343	$6,218
	Binders and Binder Accessories	$48,035	$125,811	$71,674	$61,893
	Envelopes	$6,732	$15,083	$10,848	$15,520
	Labels	$1,349	$5,609	$3,074	$3,645
	Paper	$8,827	$10,539	$11,586	$14,311
	Pens & Art Supplies	$2,622	$1,670	$1,581	$1,692
	Rubber Bands	$272	($354)	($93)	$72
	Scissors, Rulers and Trimmers	($558)	($3,331)	($2,844)	($1,066)
	Storage & Organization	$5,753	($2,087)	($23)	$3,022
Technology	Computer Peripherals	$14,153	$45,093	$17,771	$17,271
	Copiers and Fax	$41,310	$28,654	$29,283	$68,113
	Office Machines	$51,455	$180,356	$39,386	$36,516
	Telephones and Communication	$49,781	$120,597	$86,789	$59,785

($100,000) $0 $100,000 $200,000 Profit

What are the steps/processes involved in Data Visualisation?

Visualisation Process



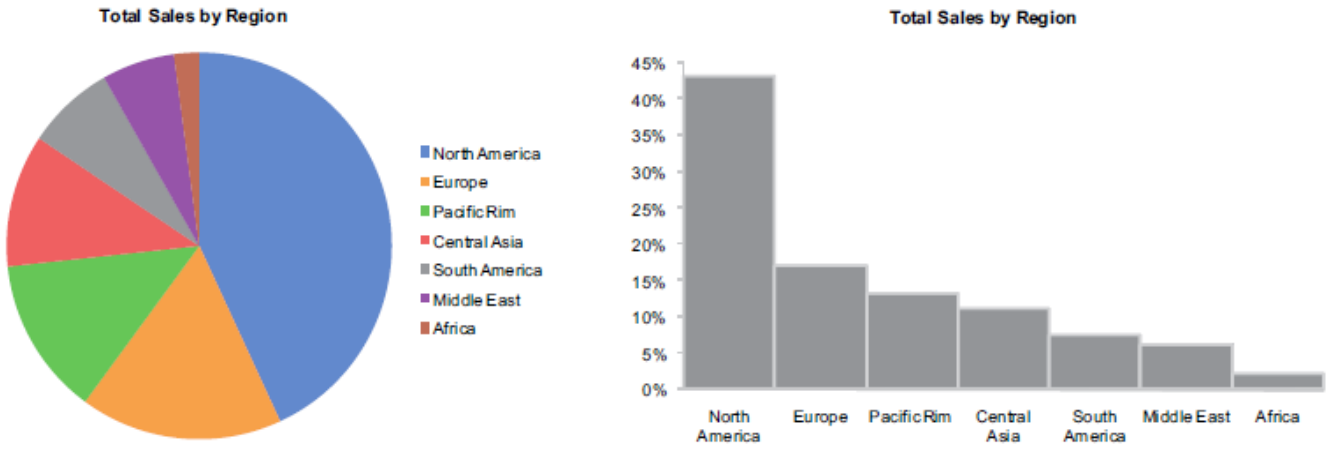
Data preprocessing/handling

- Data is mapped to fundamental data types
- Specific application data issues - missing values, errors in input, large data
  - Removal of missing data? Interpolation?
  - Using different methods to extract relevant data – CSV, JSON, XML
  - Large data may require sampling, filtering, aggregation

    Objective -> clean data -> meaningful visualisation

Visual Mapping

- Which visual representation to use?



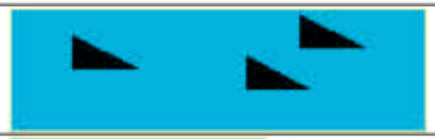
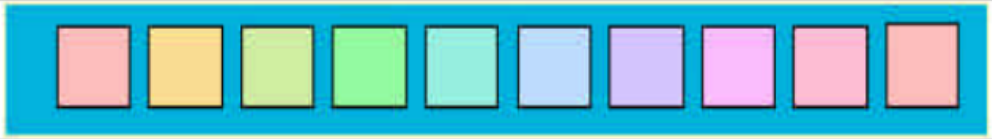
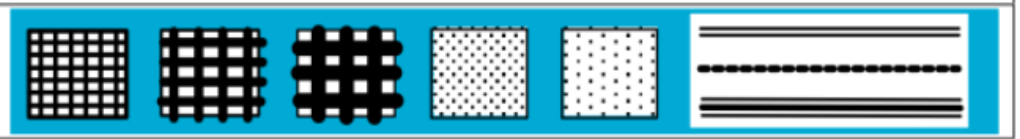
The pie chart doesn't work nearly as well as the bar graph because, to decode it, we must compare the 2-D areas or the angles formed by the slices, but we can easily compare the lengths of bars on the right.

View Transformation

- Mapping of the visual to the final presentation (dashboard, report)
- Measure by expressiveness and effectiveness
- **Expressiveness**
  - An expressive visualisation presents all the information, and only the information
- **Effectiveness**
  - A visualisation is effective when it can be interpreted **accurately** and **quickly**

Seven Key Visual Variables

Bertin's Original Visual Variables

Position changes in the x, y location	
Size change in length, area or repetition	
Shape infinite number of shapes	
Value changes from light to dark	
Colour changes in hue at a given value	
Orientation changes in alignment	
Texture variation in 'grain'	

What is
Good Visual Representation?

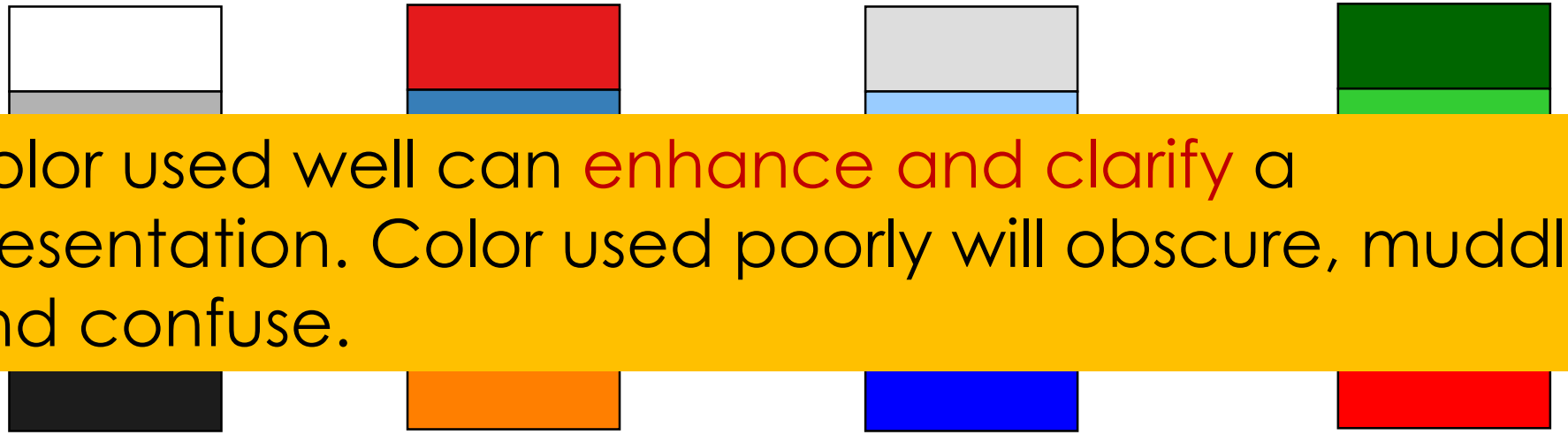
Good Visual Representation

A successful visualisation is one that efficiently and accurately conveys the desired information to the target audience.

- Suitable mapping from data to visualisation
- Ability to select and modify view
- Sufficient information density – not too much or too little
- Importance of keys, labels and legends
- Using color with care
- Importance of aesthetics

Using colors to distinguish Data

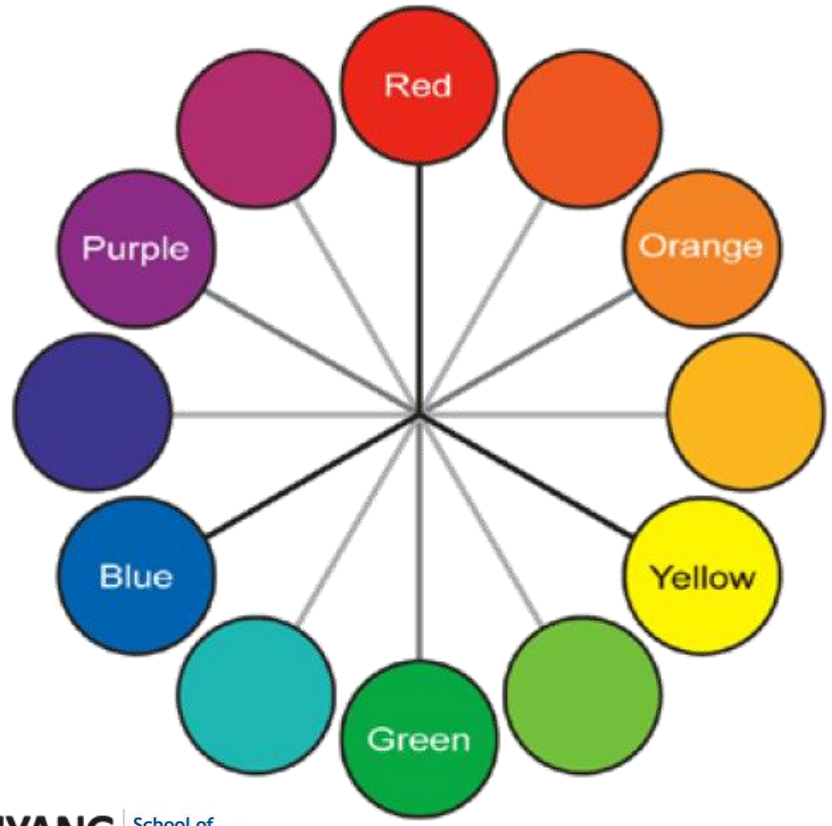
+Form	+Qualitative	+Quantitative +Ordinal	+Quantitative (Diverging)
Gray scale	Full spectral scale	Single sequence single hue scale	Double-ended multiple hue scale



Color used well can enhance and clarify a presentation. Color used poorly will obscure, muddle and confuse.

Using colors to distinguish Data

Figure – Hue circle

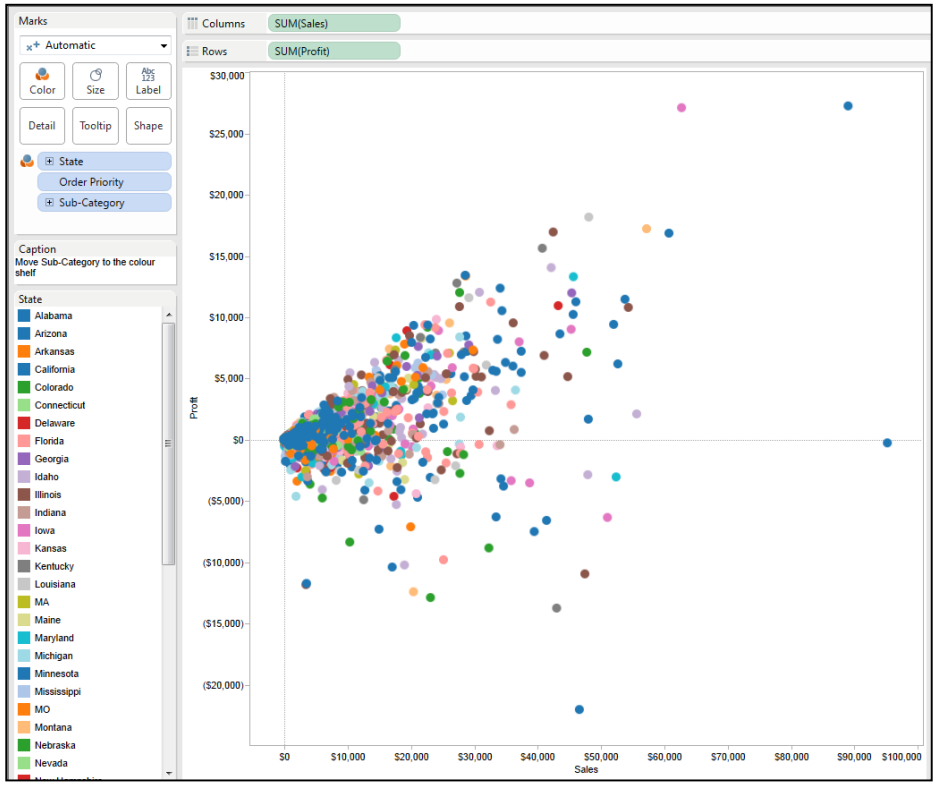


Hue is the color's name, such as red, green etc.

In any hue circle, analogous hues are close together, most simply variations of the same color name (such as red, or red-orange). Contrasting hues are on the opposite side of the hue circle

Using Colours to Distinguish Data

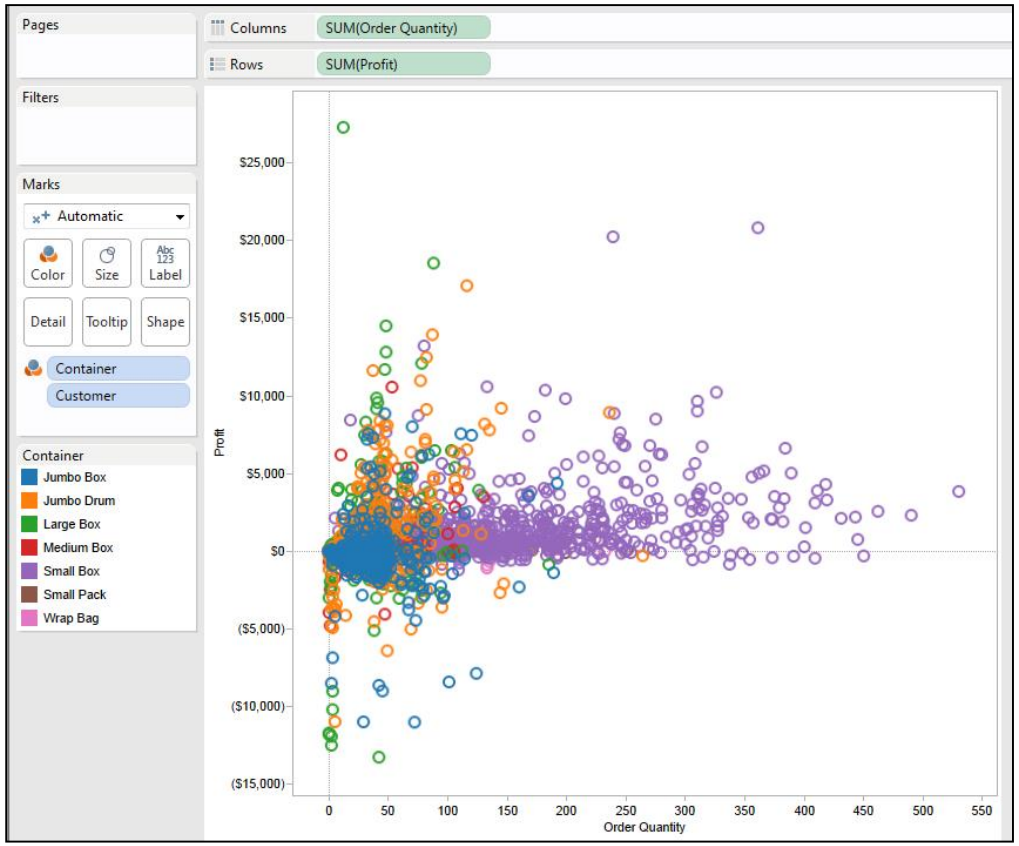
Humans can only distinguish ~8 colours



This is not helpful.

Using Colours to Distinguish Data

Humans can only distinguish ~8 colours



This is helpful.

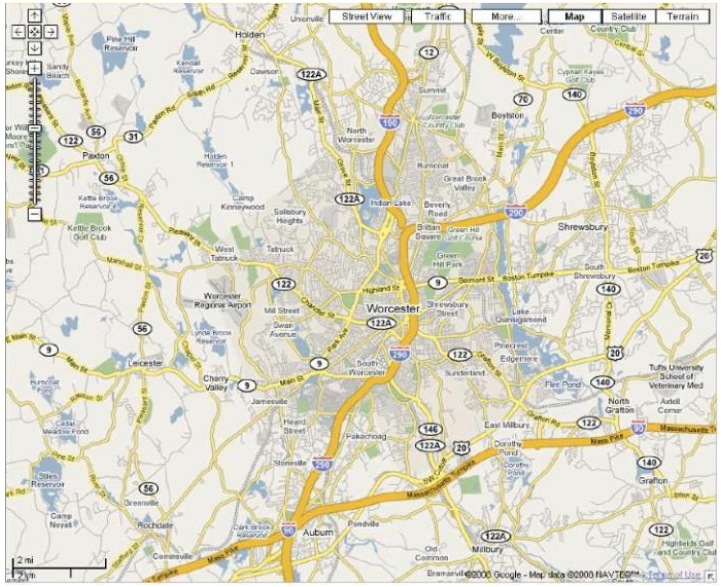
Suitable Mapping Data to Visualisation

- Mapping spatial data (longitude and latitude) to position on a map
- Mapping based on context – temperature to colour, blood pressure to height
- **Important consideration**:
  - Compatibility between scale of data field and the attribute.

    For example, ordered data attributes (e.g. age) should not be represented by un-ordered graphical attribute like shape

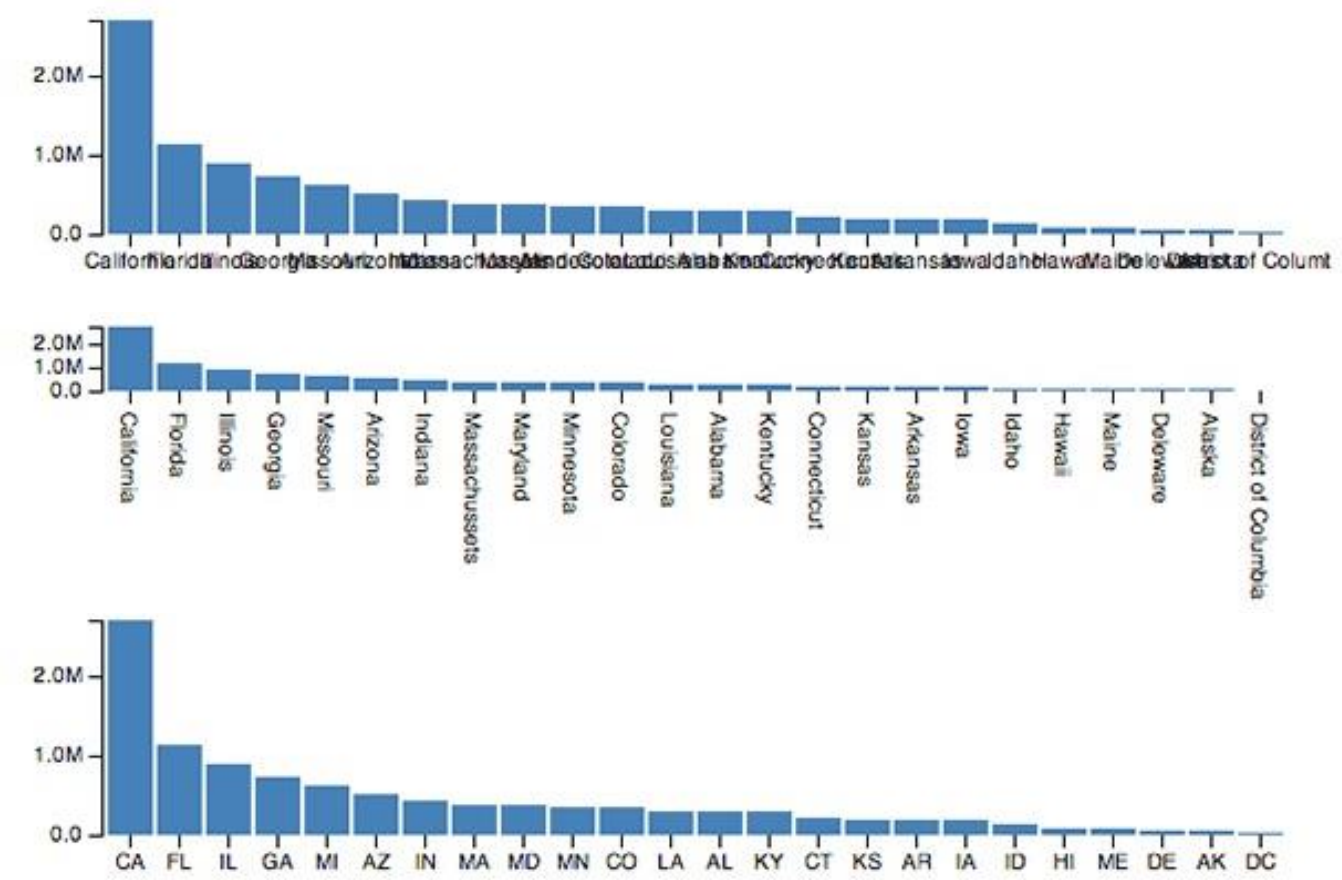
Ability to Select and Modify Views

Levels of Detail in Maps

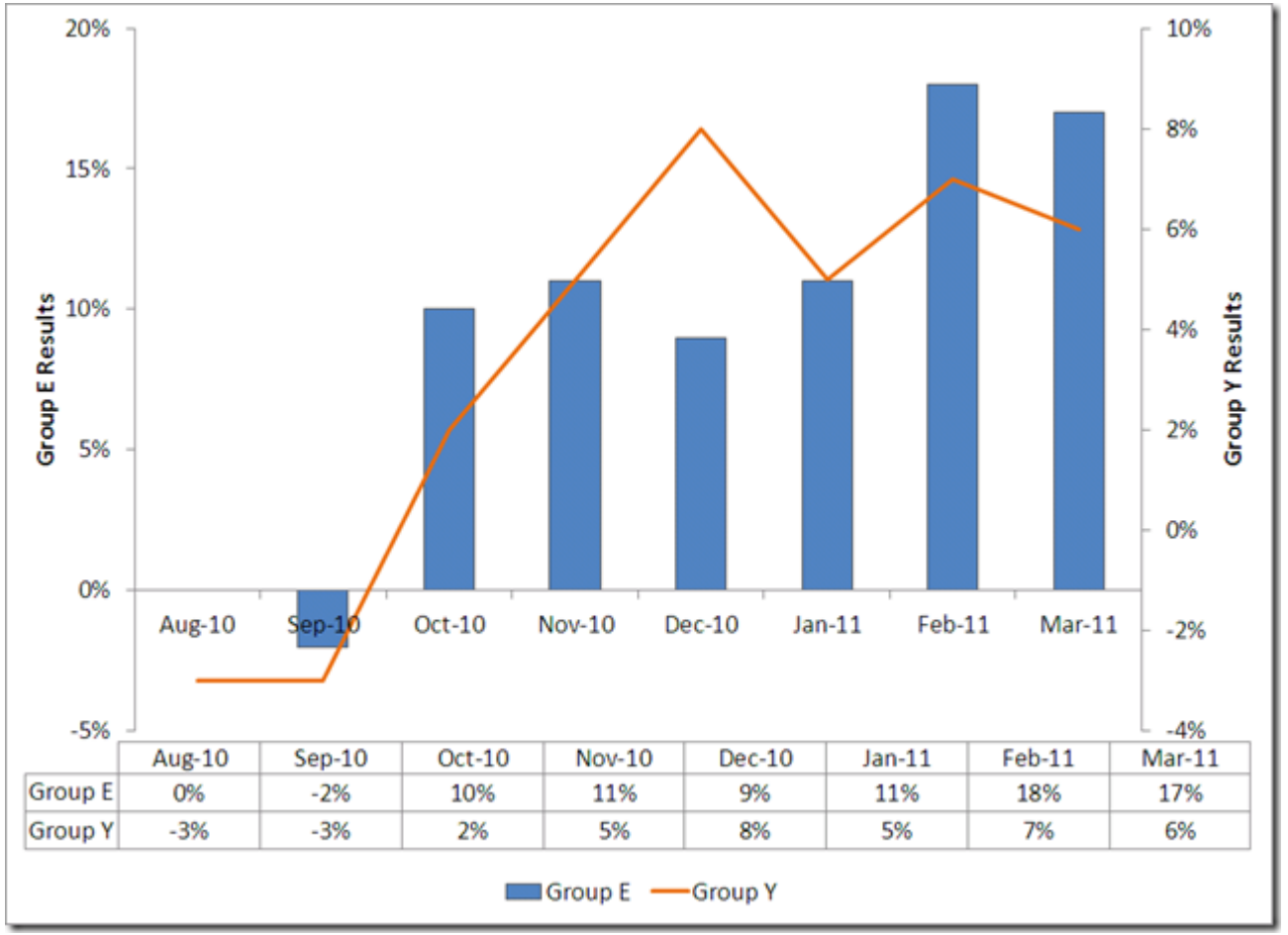


- Google Maps (http://maps.google.com/)

Importance of Keys, Labels & Legends

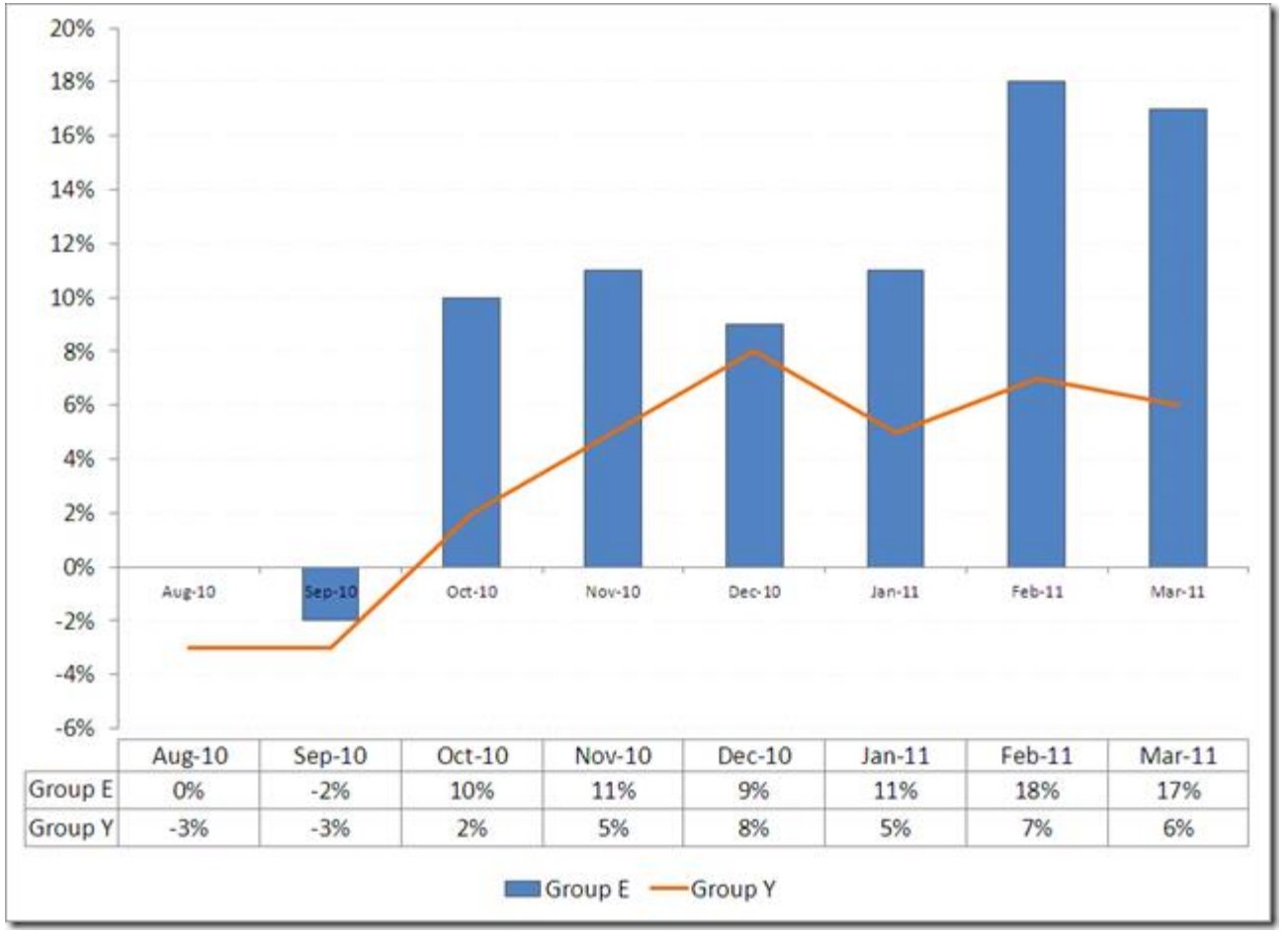


Importance of Keys, Labels & Legends



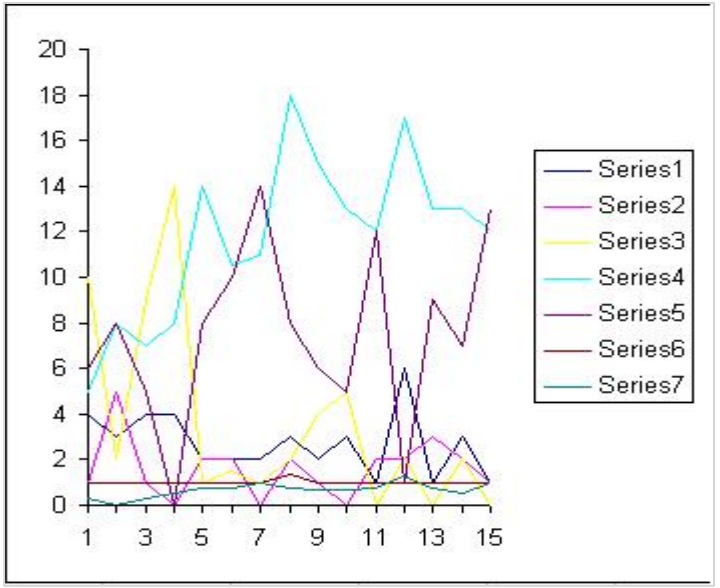
	Aug-10	Sep-10	Oct-10	Nov-10	Dec-10	Jan-11	Feb-11	Mar-11
Group E	0%	-2%	10%	11%	9%	11%	18%	17%
Group Y	-3%	-3%	2%	5%	8%	5%	7%	6%

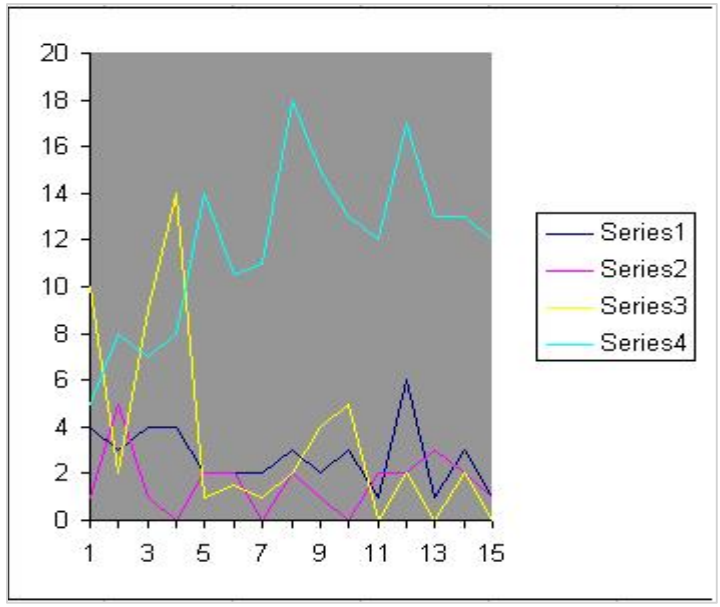
Importance of Keys, Labels & Legends



	Aug-10	Sep-10	Oct-10	Nov-10	Dec-10	Jan-11	Feb-11	Mar-11
Group E	0%	-2%	10%	11%	9%	11%	18%	17%
Group Y	-3%	-3%	2%	5%	8%	5%	7%	6%

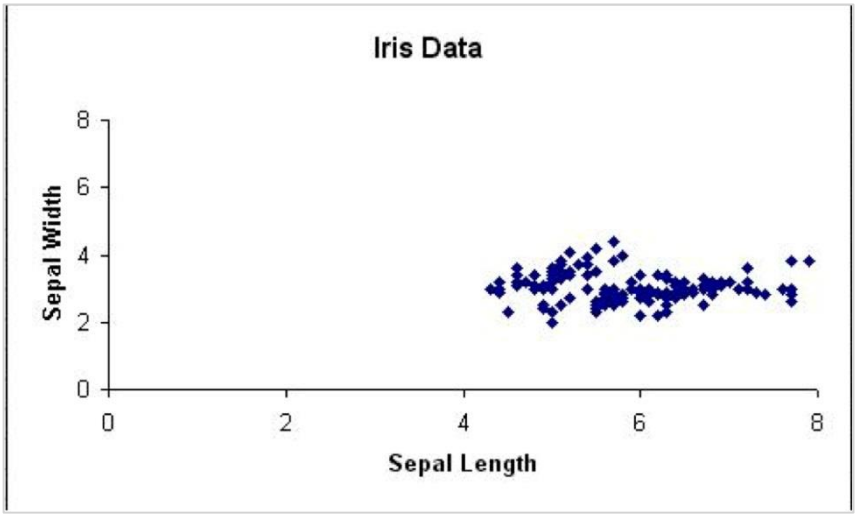
Using Colour with Care

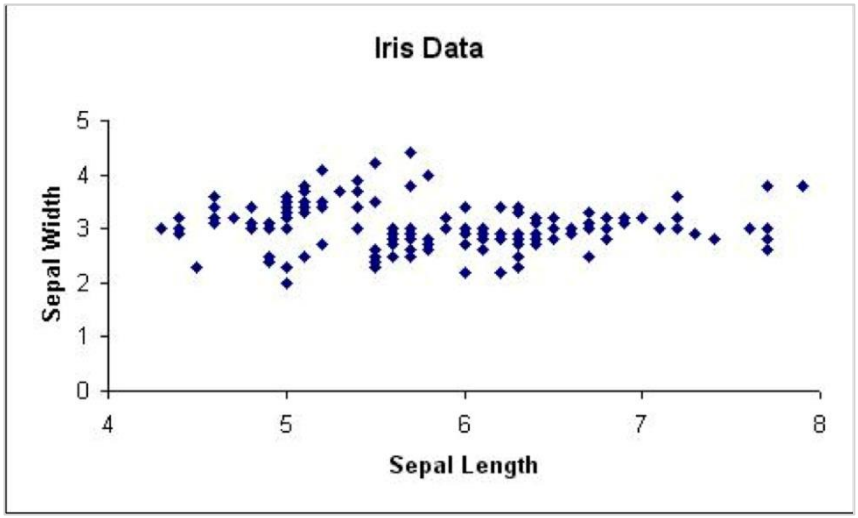




Too many colours vs. A moderate number of colours

The Importance of Aesthetics





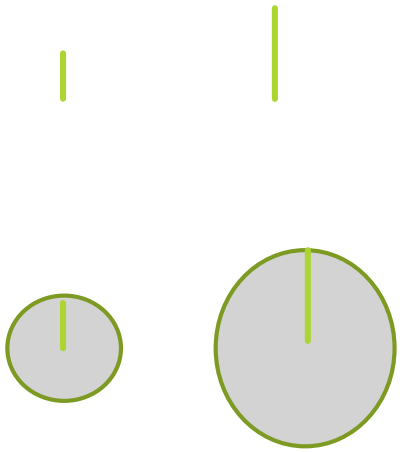
Everything to one side vs. Balanced between left and right.

Tell the Truth, No Distortion!

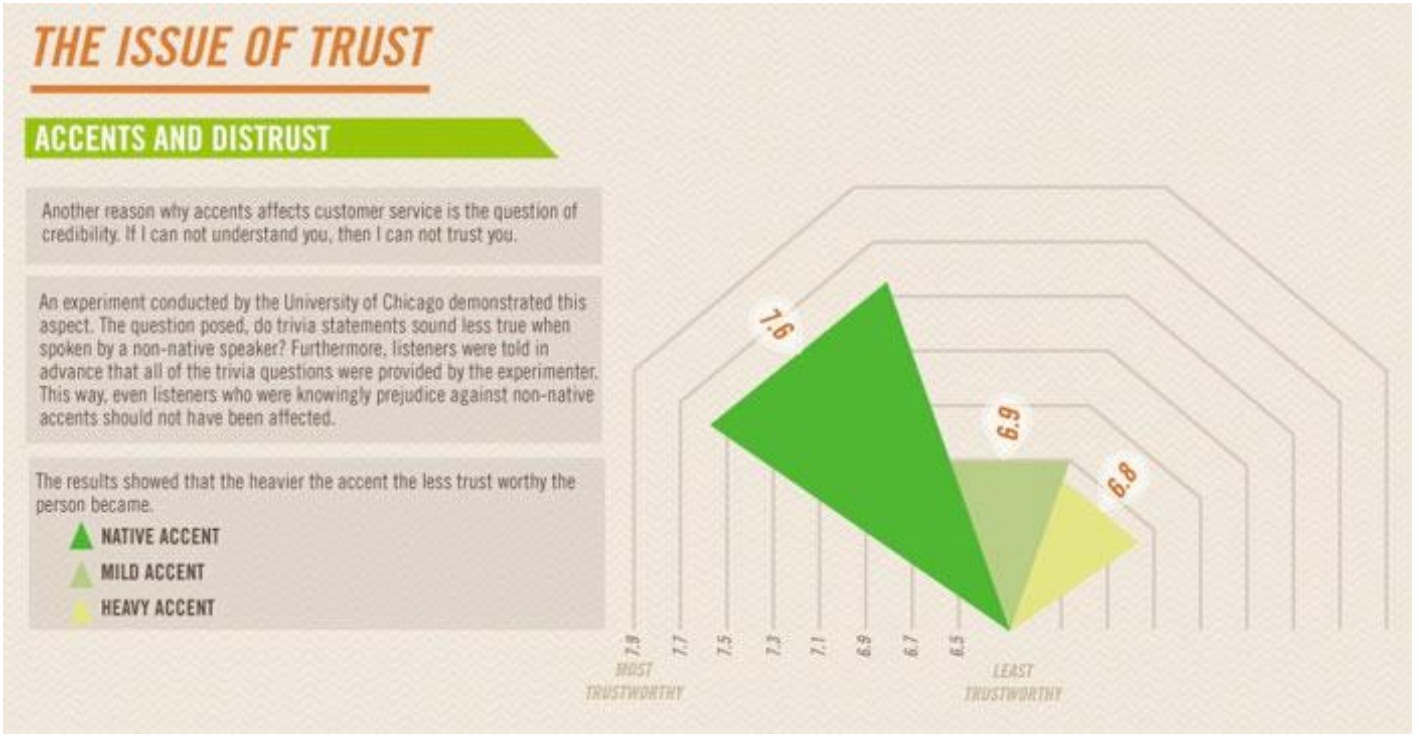
- "Three Dimensional "effects
- Nonlinear data scaling (lie factor)
- Truncated graph

"Three Dimensional" effects

Nonlinear Data Scaling



Data is linear.
Length is fine
but not area

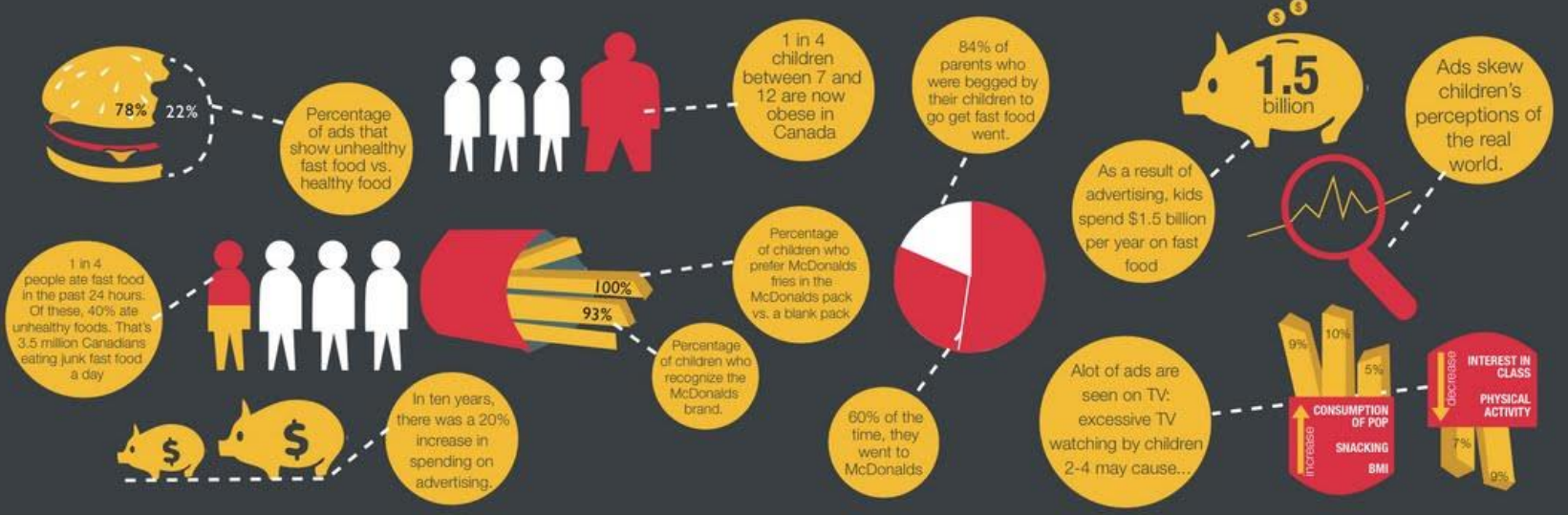


the KINGS of FAST FOOD

FAST FOOD'S ADVERTISING BLITZ TARGETS OUR FUTURE GENERATIONS: HERE'S AN UP-CLOSE LOOK AT THE GOOD, THE BAD AND THE UGLY IN THE WORLD OF FAST FOOD

FAST FOOD and their ADVERTISMENTS:

We all know that feeling we get when we see a fast food ad- your mouth salivates and suddenly, you're hungry! We also know we spend too much darn money on it, even though we know it's doing us no good... But what are the statistics and numbers concerning fast food and its devilish advertising?



WHAT CAN WE POSSIBLY DO? WHAT ARE OTHER PEOPLE DOING ABOUT IT?

It's obvious now that all of these ads are harming us, but what can we do? In Quebec, they passed a "Quebec Consumer Protection Act" (CPA) which prohibits commercial advertising to children under the age of 13, and the results speak for themselves.

SO HOW DO WE DO OUR PART TO STOP THESE CRAZY ADVERTISMENTS?

- Educate youre kids about advertising and it's effects
- Watch non-commercial TV
- Encourage savy consumer habits
- Write to Advertising Standards Canada (ASC) about print advertising and Canadian Radio-television and Telecommunications Commission (CRTC) and the Canadian Broadcast Standards Council for TV and radio ads
- Participate in "Buy Nothing Days"

FAST FOOD COMPANY

ADVERTISING BUDGET

ADVERTISMENTS VIEWED PER WEEK BY AGE GROUP

REASON PARENTS WENT

OTHER TACTICS USED TO LURE CHILDREN

- TOYS
- COMMERCIALS
- URL WEBSITES
- MASCOTS/ CELEBRITY
- SONGS/ JINGLES
- PLAYPLACES
- AFFILIATION WITH TV SHOW/ VIDEO GAME

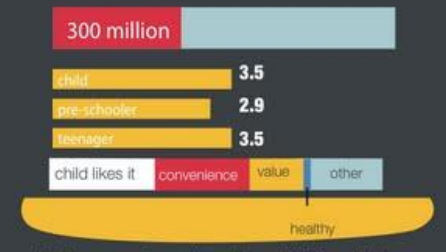
THE GOOD SUBWAY

If you must eat out, Subway is one of your best options. It advertises on the basis of truth and health, not using toys, gimmicks and lies. Even though it has a market for children, its purpose is to promote healthy eating habits, not junk!

SUBWAY; THE RIGHT WAY ✓

Subway has one of the healthiest fast food kid meals of all. With lean turkey, ham or beef, its sandwich has less fat and more protien than burgers. With juice and a fruit roll up, we give it a thumbs up.

THE BAD BURGER KING



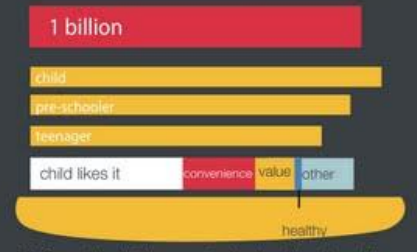
Just a mere step down from McDonalds in terms of success and unhealthiness, BK has a similar marketing strategy to McDonalds by targeting children. This means, like McD's, there's a good turn out due to kids pestering parents to go. They *are* easier on the amount of ads they show to children though, so we thank them for that.

BURGER NOT SO FIT FOR A KING

MEH

BK's kids meal is actually not all that bad. While a burger isn't the healthiest option, their apple "fries" and juice make up for it, not too far behind Subway!

THE UGLY MCDONALDS



McDonalds is the worst of all evils when it comes to bombarding our children with ads. Also, most of the ads are based on things other than the food; they get to our children with songs, games, bright colours, lies, toys etc.

UNHAPPY MEAL ✗

Children don't know any better, so of course, when given more options containing bad choices, they are more likely to order the junk.


THANK YOU