Introduction to Data Visualisation

- Theory

Contents

Data Visualisation
Introduction

Topics

In the following theory section we will guide you through the following topics:

- **Intro**
- **Why visualise?**
- **Classifications of Visualisations**
- **Visualisation process**

The topics covered in this section are meant to give an general overview of data visualisation and are by no means exhaustive lists or examples.

Introduction to Data Visualisation

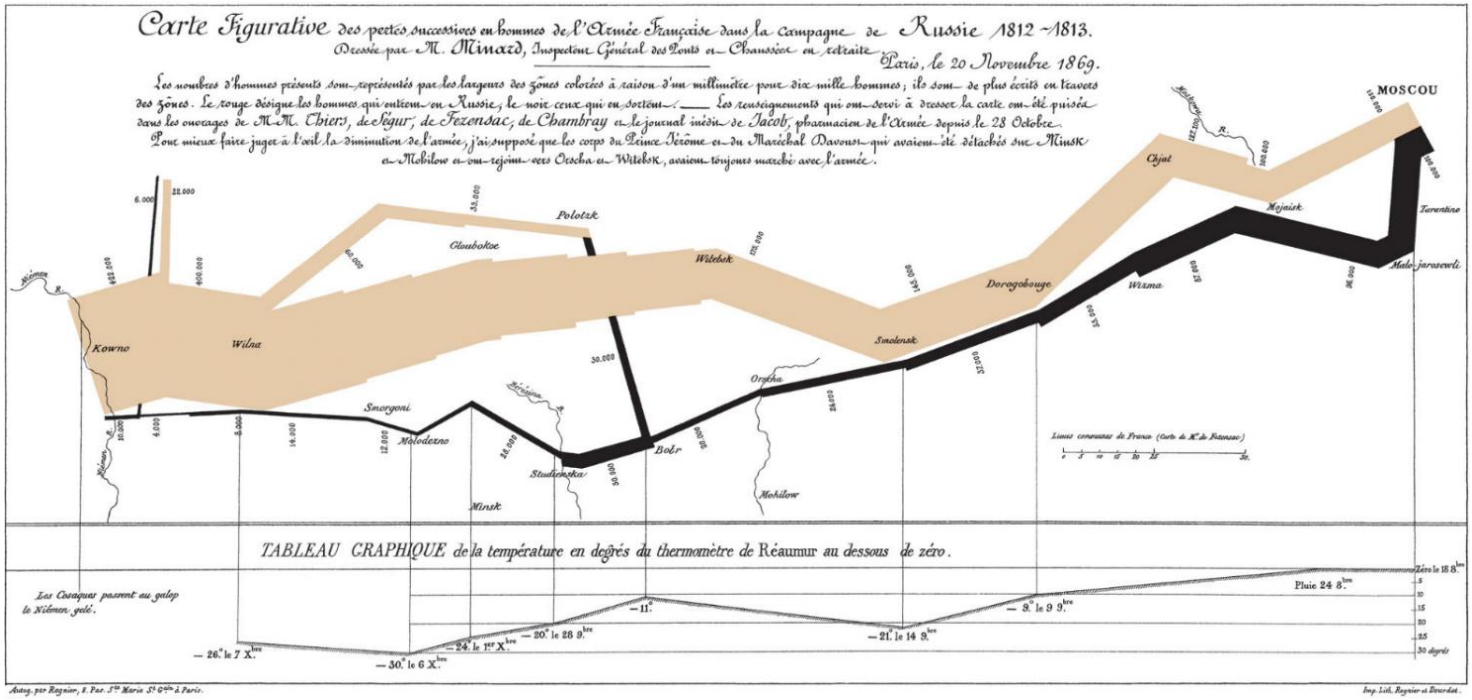
- Intro

Data Visualisation

History of Visualisation

Charles Minard's graph of Napoleon's invasion

Drawn in 1869 one of the most cited examples of statistical graphics occurred when Charles Minard mapped Napoleon's invasion of Russia. The map depicted the size of the army as well as the path of Napoleon's retreat from Moscow – and tied that information to temperature and time scales for a more in-depth understanding of the event.



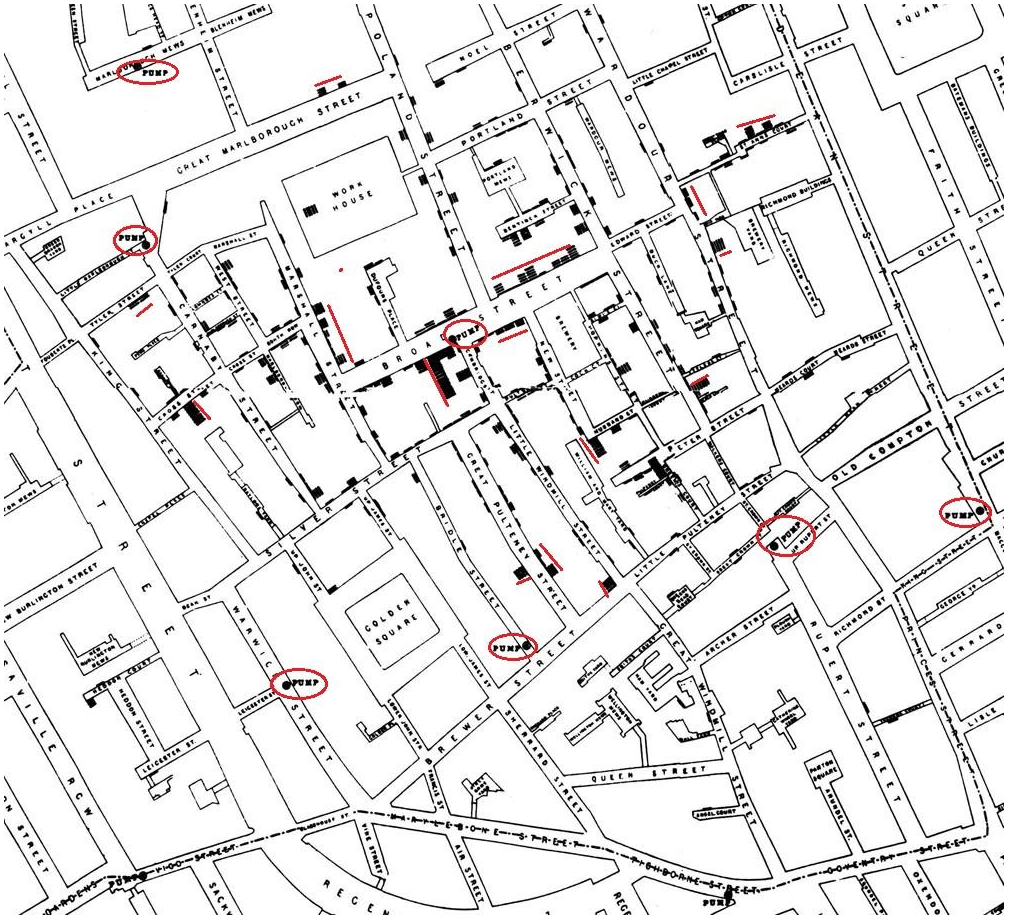
Data Visualisation

History of Visualisation

John Snow's cholera map

In 1854 doctor John Snow used a map to chart the cases of cholera in London's Soho district. He identified that the cases where clustered around water pumps and thus determined that the disease was spread through the water supply.

Snow's ground breaking study was a major event in the history of public health and geography. It is regarded as the founding event of the science of epidemiology.



Introduction to Data Visualisation

- Why visualise?

Data Visualisation

Why visualise?

What is it?

Data visualisation helps the end user to **understand** and get **insights** on the visualised data. Visualisation focuses on techniques to present data in a visual way in order to facilitate the discovery and understanding of underlying patterns, whether it's done for research, science or for decision makers, in order to:

Display

Plots, trends, timelines, etc.

Analyse

Develop and asses hypotheses

Discover errors in data

Find patterns

Communicate

Share and persuade

Collaborate

Data Visualisation

Why visualise?

Advantages

Data visualisation takes advantage of the human brain's highly evolved **visual system**. Our visual cognitive abilities mean we can quickly recognise patterns in an image.

Visualisation takes advantage of this ability to **identify**, **explore**, **interpret** and **understand** patterns within large datasets.

This is becoming increasingly important as the amount of data involved in every profession is growing exponentially by the day.

Data Visualisation

Why visualise?

To Display

There are a multitude of ways available to display data according to the needs of each circumstance.

Bar Charts	Scatter Plots	Spark Lines	Box Plots
Tree Maps	Line Graphs	Bullet Graphs	Heat Maps
Pie Charts	Flow Charts	Org Charts	etc.

Whether we seek to **aid analysis** of the data or simply to **inform** the viewer, all of the above help us display data in a visual manner.

Data Visualisation

Why visualise?

To Analyse

The outcome of the above charts and graphs helps us detect:

- **Trends**
- **Anomalies**
- **Correlations**
- **Patterns**

and eventually make:

- **Decisions**

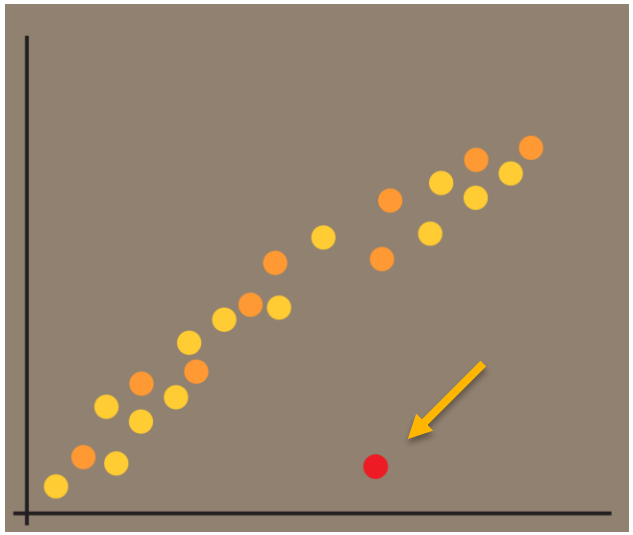
Data Visualisation

Why visualise?

To Analyse

If you need to **find** something buried inside **1,000,000,000** of data points

- Would you rather **read** through the data or **visualise** it?
- What if you don't even know **what** you are looking for?

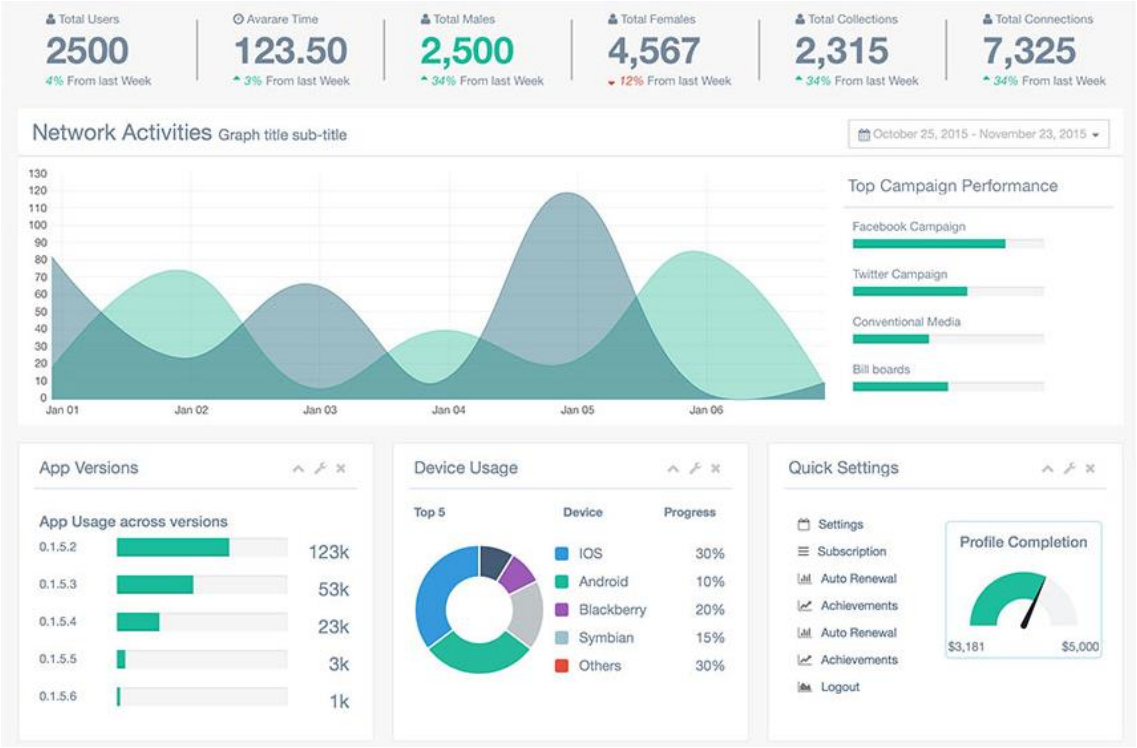


Data Visualisation

Why visualise?

To Analyse

A well-designed **dashboard** allows decision makers to analyse **massive** datasets at a glance



Data Visualisation

Why visualise?

To Communicate

- The right visualisation can **emphasize** key points, provide **context** and **engage** the audience.
- Great speakers use visualisations to **support** their ideas and make them **memorable**.
- Sometimes the visualisation **communicates** all the information **on its own**.

Data Visualisation

Why visualise?

Reasons We Visualise By Francesco D'Orazio

Francesco D'Orazio, a visualisation expert, gives a few convincing reasons why visualisation is necessary to exploit the value of data:

- Visualisation acts as an **external memory**, allowing us to take into account a greater number of variables and conduct reasoning on them.
- It allows us to **objectify abstract information** with shapes and colors to more easily compare and classify large amounts of data.
- Visualisation is perfect for providing **context and narrative** to data, thus allowing us to grasp a holistic view of a problem, not just a fraction of it.
- Visualisation allows us to represent **process**, thus we can incorporate time in spatial terms and depict transformative processes in a visual way.

Presentation by Francesco d'Orazio:

http://www.slideshare.net/Facegroup/10-reasons-why-we-visualise-data?from_action=save

Introduction to Data Visualisation

- Classification of Visualisations

Data Visualisation

Classification of visualisations

Categorisation by Andy Kirk

Andy Kirk, a data visualisation specialist who has grown to become a guru of the field proposes a categorisation of visualisations based on the intended function of the visualisation, namely:

Exploratory

Explanatory

Exhibitory

However it is important to remember that often visualisations fall into more than one of these categories.

Data Visualisation

Classification of visualisations

Exploratory visualisation

Exploratory visualisation is used when there is a big amount of data and we are **unsure of the information hidden within it**. In order to get a sense of what is hidden in the data we use a visual medium to help identify its features such as patterns, trends and outliers.

Exploration is better to start at a high level of granularity. After detecting points of interest, one may dig deeper to detect details.

Example:

Britain's Diet: http://britains-diet.labs.theodi.org/?es_p=1359956

Data Visualisation

Classification of visualisations

Exploratory visualisation

Line Graphs

Data Visualisation

Classification of visualisations

Explanatory visualisation

Explanatory visualisation is used when we already know what is in the data and need to convey and explain an insight to someone else, such as a decision maker or the general public.

Examples:

Syrian Conflict:
http://www.slate.com/blogs/the_slatest/2015/10/06/syrian_conflict_relationships_explained.html

Vaccination effects: https://www.theguardian.com/society/ng-interactive/2015/feb/05/-sp-watch-how-measles-outbreak-spreads-when-kids-get-vaccinated

Data Visualisation

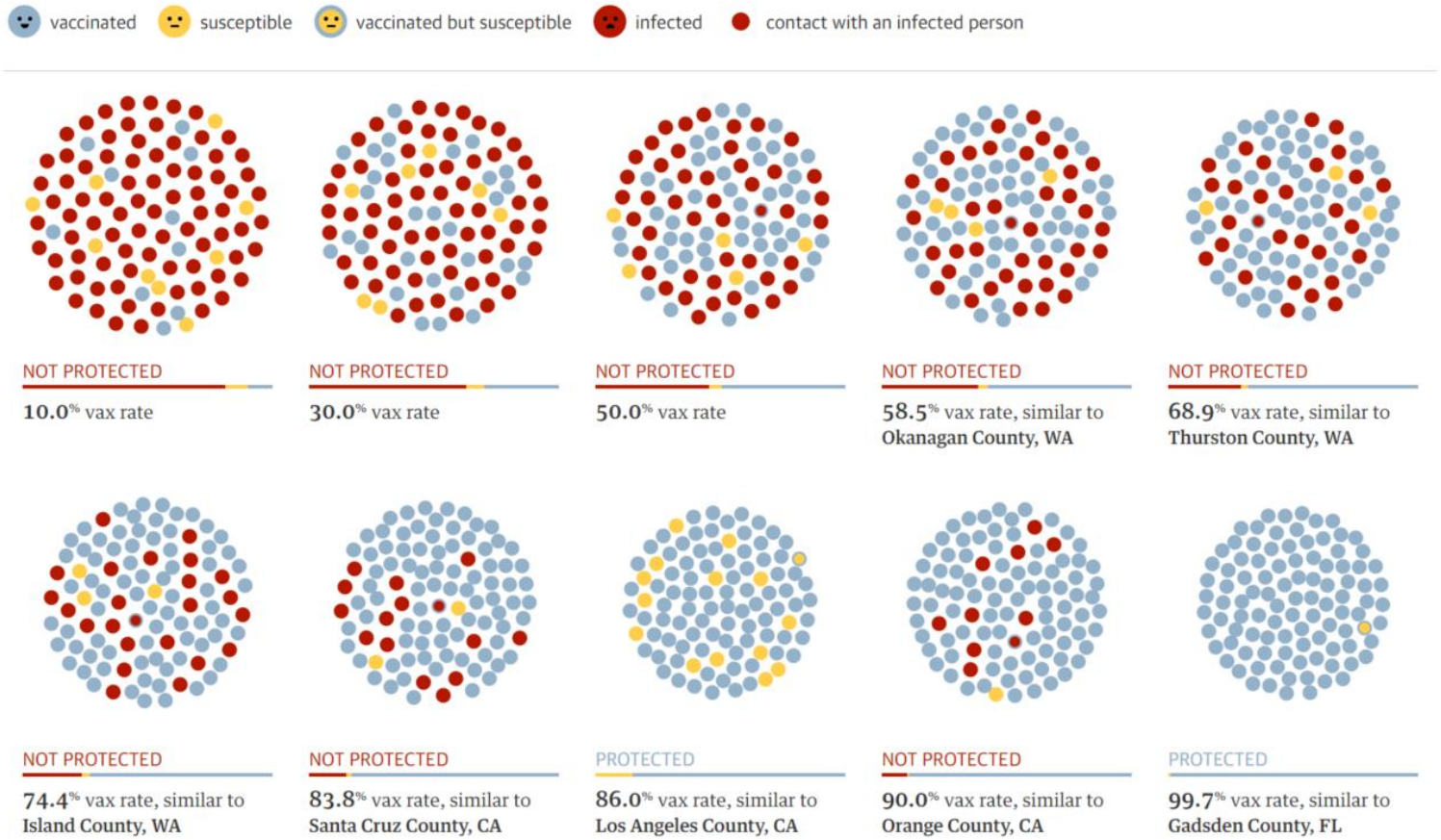
Classification of visualisations

Explanatory visualisation
Infographic

Data Visualisation

Classification of visualisations

Explanatory visualisation
Simulation



Data Visualisation

Classification of visualisations

Explanatory visualisation

Map + Relations

Data Visualisation

Classification of visualisations

Exhibitory visualisation

Exhibitory visualisation is simply the displaying of data. It may be the case in communication scenarios or in times where simply displaying the information tells a story on its own.

Examples:

NatGeo offshore wind: http://www.nationalgeographic.com/climate-change/carbon-free-world/index.html?source=carbon-free-america#map/offshoreWind/GRC

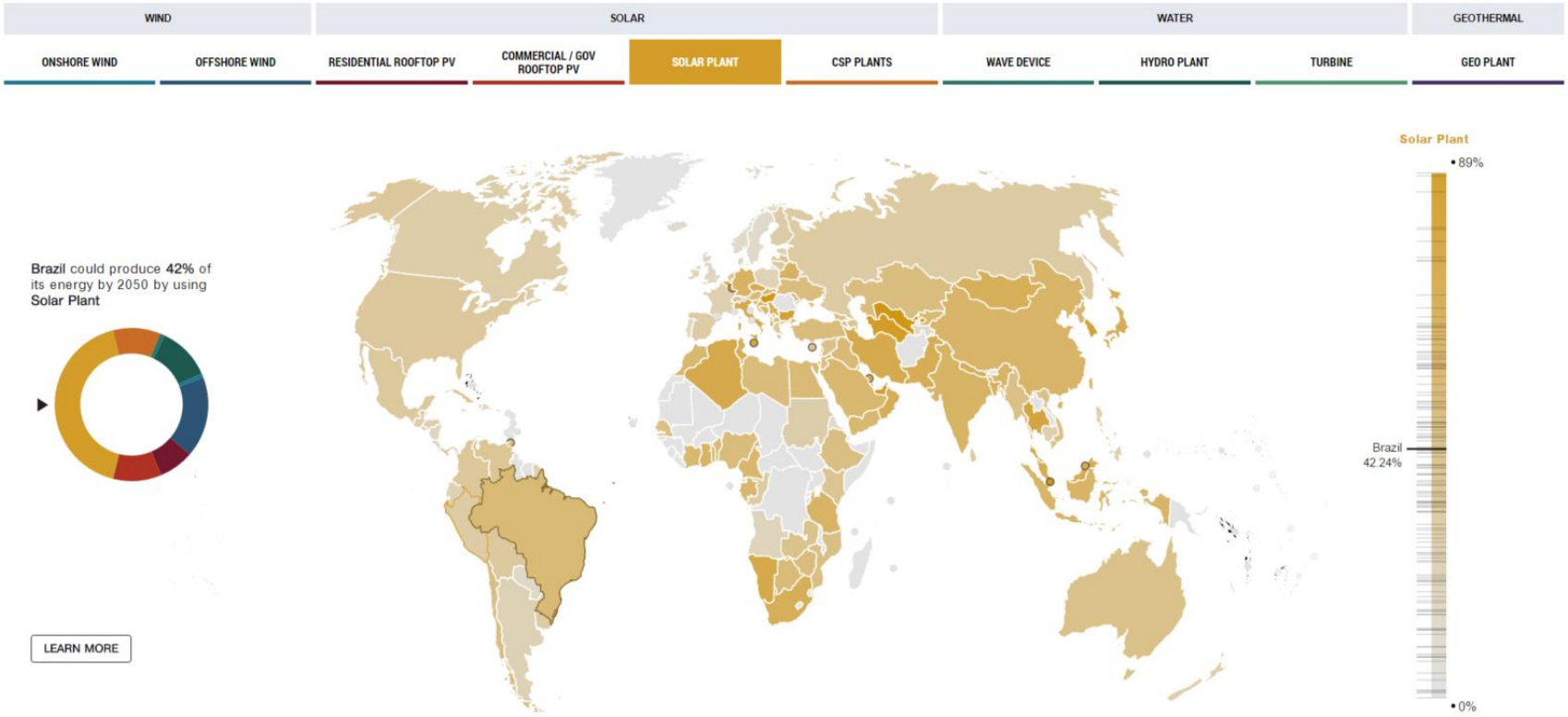
World Languages: http://www.densitydesign.org/ddfs13/afterbabylon/

Data Visualisation

Classification of visualisations

Exhibitory visualisation

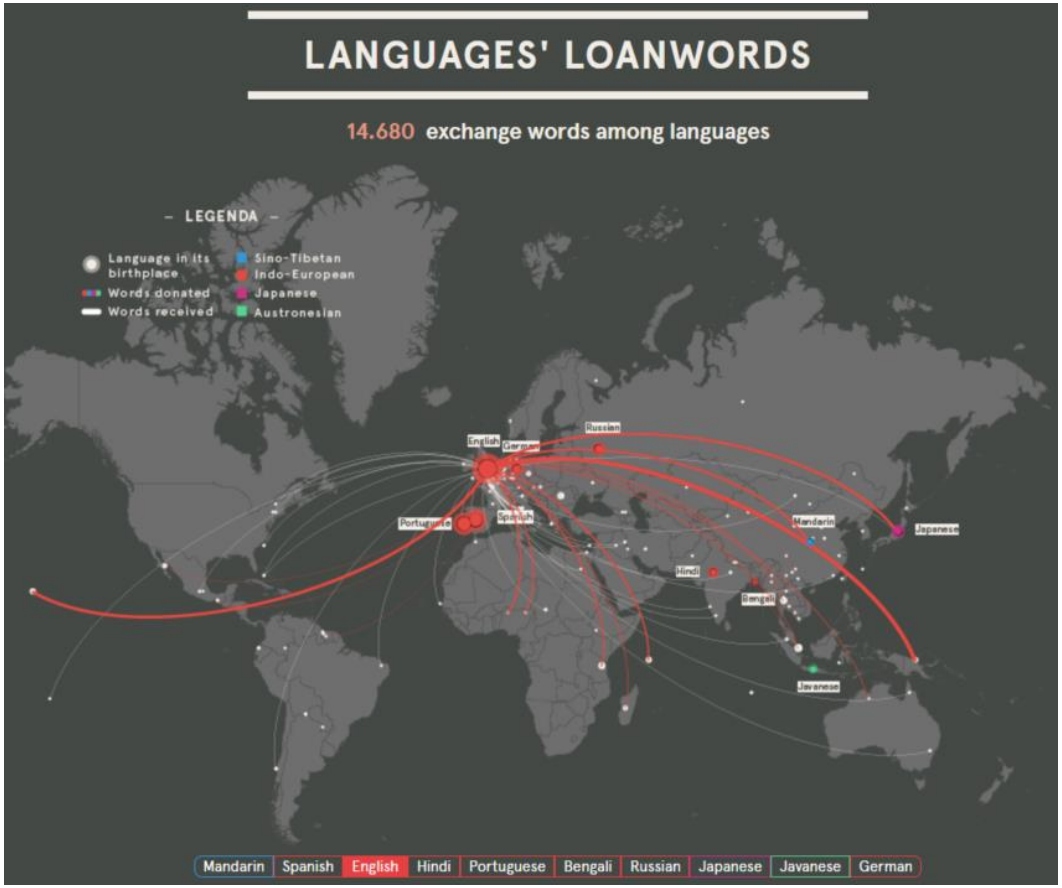
Map



Data Visualisation

Classification of visualisations

Exhibitory visualisation
Map + Relations



Introduction to Data Visualisation

- Visualisation Process

Data Visualisation

Visualisation process

Stages of the visualisation process

In order to end up with a satisfying result in data visualisation, it is important to follow a step-by-step process. There are four main stages of this process:

1. **Question formulation**
2. **Data preparation**
3. **Considering the medium**
4. **Development of a visual representation**

Data Visualisation

Visualisation process

Question Formulation

When creating a visualisation, the first step should always be to **clearly state the question to be answered**.

By being conscious of the answer we need, we can more effectively choose the data required to answer it.

A common mistake is to dive head first into all the available data and end up losing the initial goal and over-complicating a rather simple process.

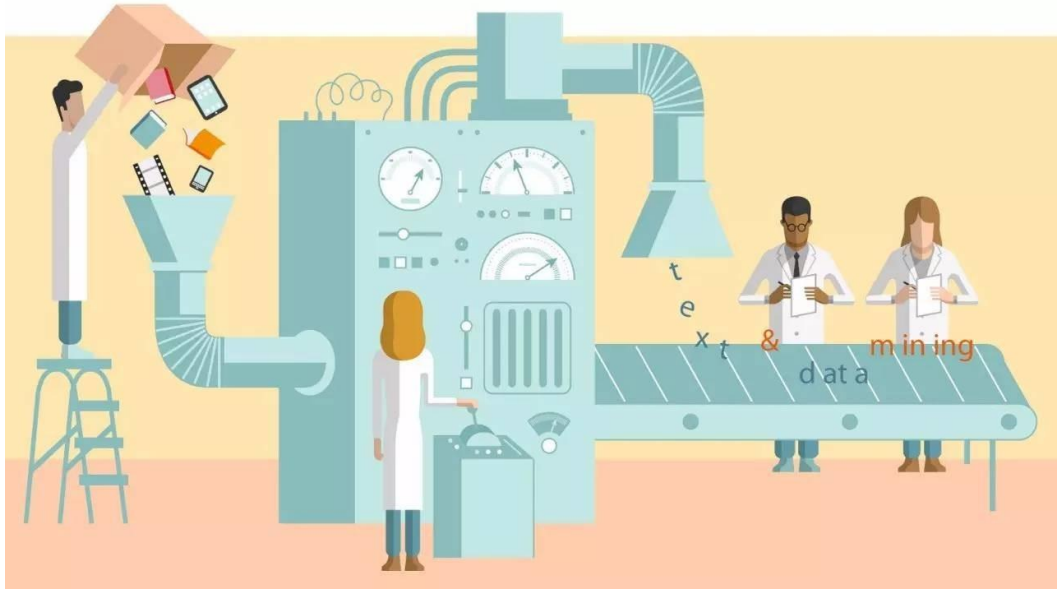
Data Visualisation

Visualisation process

Data Preparation

In this step we ensure that we have **all the data** we need and **the way we need it**. It is often the case that the most time consuming step of a visualisation project is preparing the data to be used. Preparing the data involves, among others, ensuring the required:

- Accessibility
- Validity
- Accuracy
- Relevancy
- Format
- Consistency
- Granularity



Data Visualisation

Visualisation process

Considering the Medium

Another thing to consider before moving on to create a wonderful visualisation, one must consider the medium that will be used to display it. A couple of examples could be:

- Will it be **print or digital**?
  - In print we have a set space that we can predefine, while in digital we have to take into account how it will look on different screen types and sizes.
  - Furthermore in print the colour output can be tested and set, but in digital colours can appear differently from screen to screen, and ruin the outcome.

- Will it be **static or interactive**?
  - If it is static all we need to do is design a single layout. But if it's interactive we have to consider how elements change and how this affects the information conveyed and the aesthetics of the bigger picture.

Data Visualisation

Visualisation process

Development of visual representation

After having identified our question and prepared our data we can move on to creating the visualisation. Only now is it time to decide on any issue relating to the appearance of the visualisation itself.

The optimal visualisation design depends on two factors, primarily: the **message** to be conveyed to the audience, meaning the question to be answered, and the **variables** to be shown.

Data Visualisation

Visualisation process

Good visualisation is about making good decisions

To make the best decisions you need to be familiar with all your **options** and aware of the things that will influence your **choices**.

Data Visualisation

Visualisation process

Good data visualisations are trustworthy

Communicating with numbers is, in many ways, just like communicating with words. You **make decisions** about what to emphasize and what to downplay, and about how to convey a full understanding of the subject at hand.

Christopher Ingraham, The Washington Post

Data Visualisation

Visualisation process

Things to consider

The first things to **consider** when visualising data are:

- **What do we want to present?**
- **Which variables?**
- **What types of variables?**
- **Which properties?**
- **Which aspects?**
- **What is the range of the variable values?**
- **What medium will display our visualisation?**

Data Visualisation

Types of data

That is why we must understand the type of data at hand.

Quantitative Data

- Data dealing with numbers
- Data can be measured
- Continuous data deriving from measurements or processed data such as means, variances etc.
- Data can be categorical or continuous
- E.g.: length, volume, speed, etc.
- Sub-categories of numerical data are:
  - ***Shape*** data
  - ***Coordinate*** data

Qualitative Data

- Data dealing with descriptions
- Data can be observed but not measured
- Data can be categorical only
- Consider quantifying the data
- E.g. colors, mood, appearance, etc.

Data Visualisation

Types of data– Example

Geospatial Data

- Data that can be characterised by a specific location in space and time
- Information about a physical object that can be represented by numerical values in a geographic coordinate system
- Describes the location, size and shape of an object

E.g. Coordinates:

Lat: 51.0543, long: 3.7174

N 51 3'16", E 3 43'3"

Network Data

- Represents hierarchical or non-hierarchical relationships and interactions between entities, examples include:
  - Graphs representing relationships between entities (e.g., FB friends);
  - Interactions (e.g., communication traces in social networks);
  - And hierarchies (e.g. taxonomies).

E.g. LinkedIn data

Textual Data

- Data consisting of text
- Usually analysed to produce:
  - Text categorisation;
  - Text clustering;
  - Concept and pattern extraction;
  - document summaries,
  - And sentiment analysis etc.
- Has to be "mined" before analysis

E.g. Poetry text

Shakespeare (sonnet 18)

"Shall I compare thee to a summer's day? Thou art more.."

Data Visualisation

Types of visualisation by technique

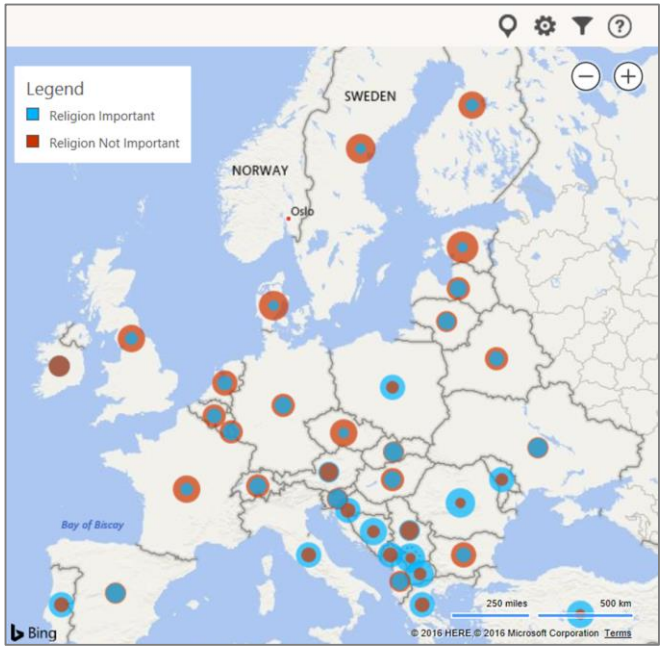
After understanding our data, we need to decide on the most effective way to visualise it for our needs.

Maps

Display data according to spatial relations, representations showing how data is distributed spatially.

Select data -> Insert -> Bing Maps

Country	Religion Important	Religion Not Important
Estonia	16%	84%
Sweden	17%	82%
Denmark	19%	80%
Norway	21%	78%
Czech Republic	21%	75%
United Kingdom	27%	73%
Finland	28%	70%
France	30%	69%
The netherlands	33%	67%
Belgium	33%	58%
Bulgaria	34%	62%
Russia	34%	60%
Belarus	34%	56%
Luxembourg	39%	59%

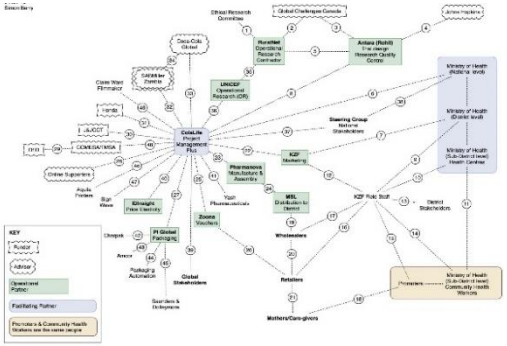


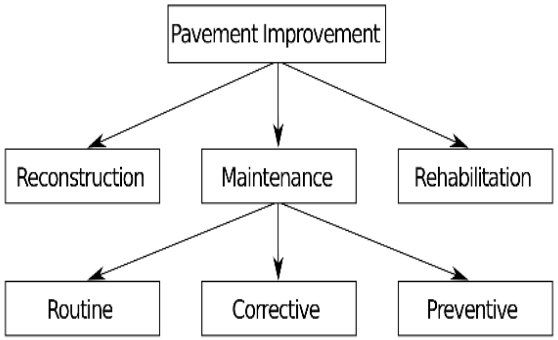
Data Visualisation

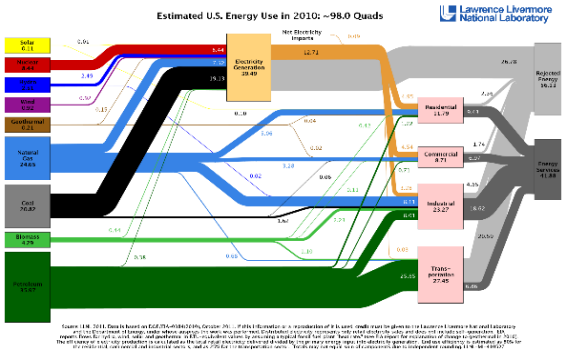
Types of visualisation by technique

Relations and dependencies

Relationships and dependencies between entities can be represented by using nodes (representing data elements) and links (representing relationships or dependencies).







Some of the content on this slide may not be done natively in Excel

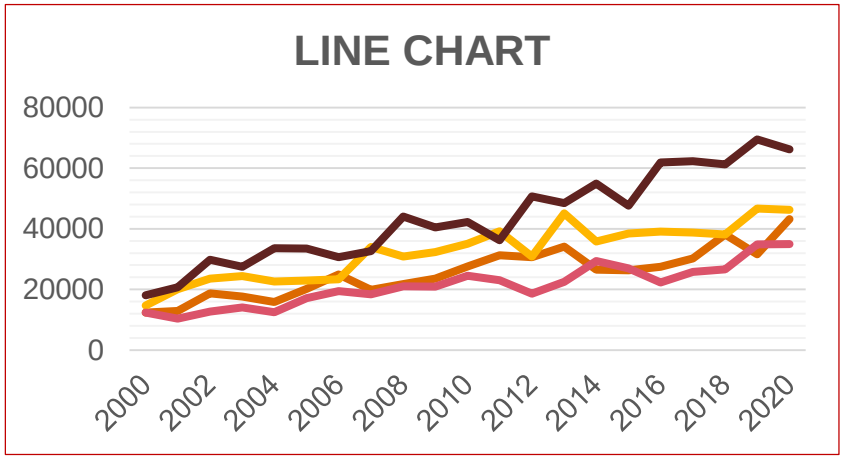
Data Visualisation

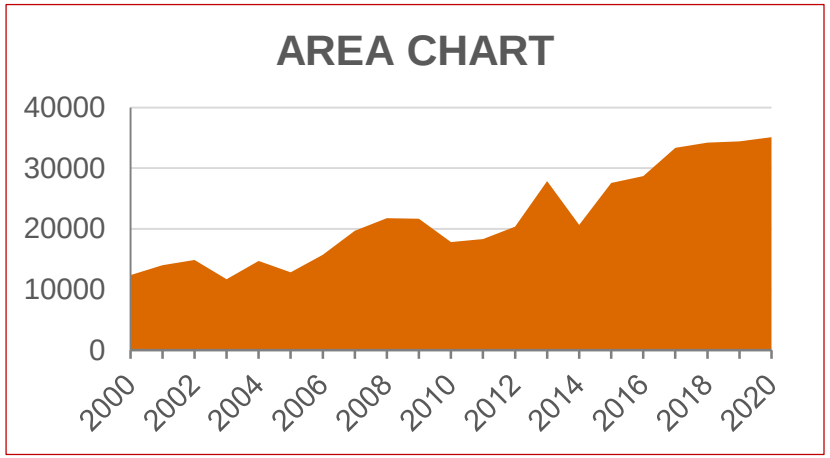
Types of visualisation by technique

Line Graphs

Represent the relation between two or more variables as a single line or area.

Select Data -> Insert -> Charts -> ...





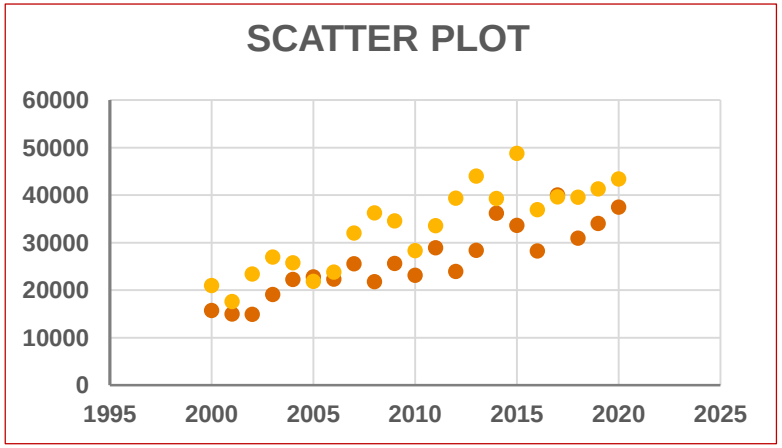
Data Visualisation

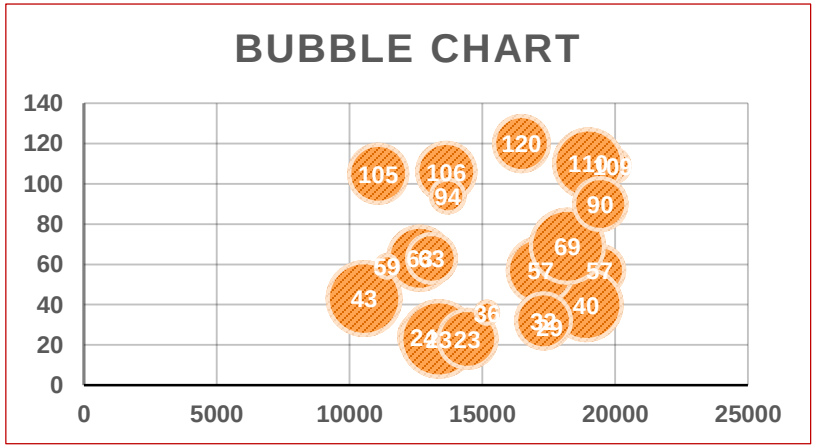
Types of visualisation by technique

Dots or Bubble graphs

Are a graphical display of data, representing the relation between two or more variables using dots or bubbles.

Select Data -> Insert -> Charts -> ...





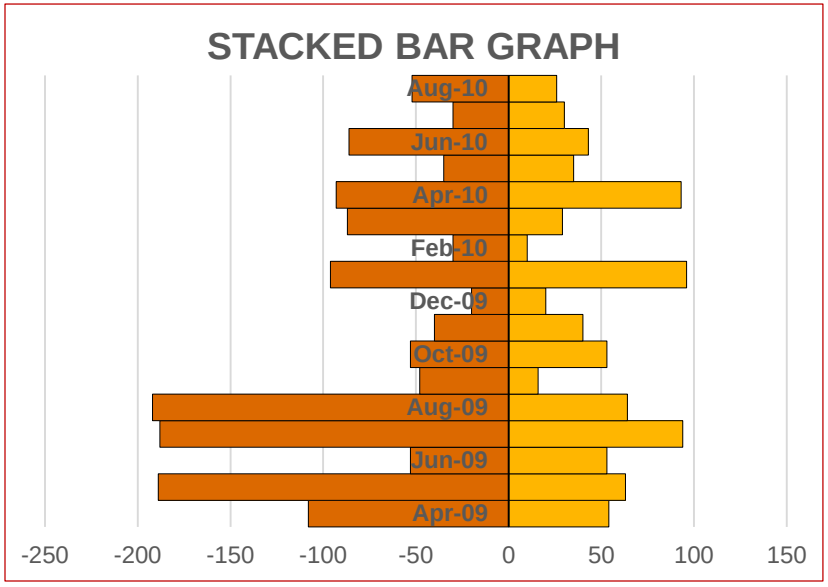
Data Visualisation

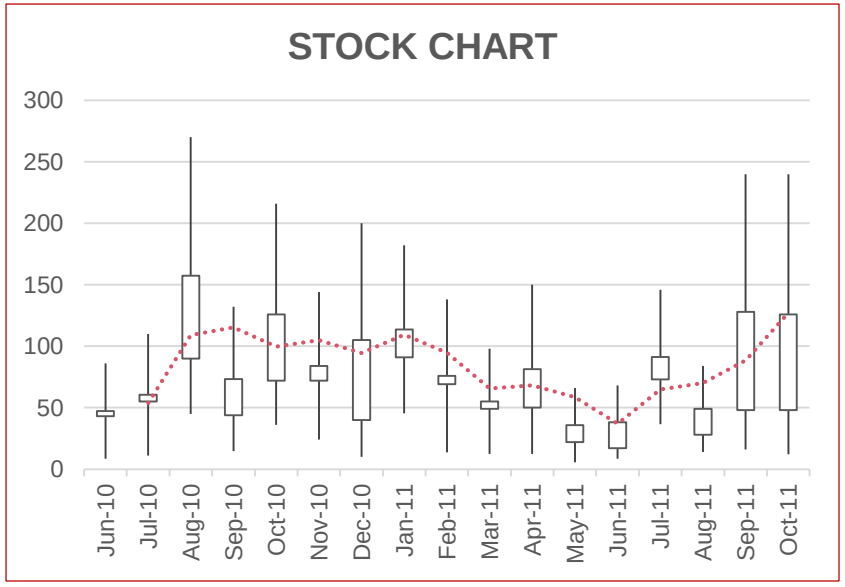
Types of visualisation by technique

Bar graphs

Are visualisations representing the relation between a categorical variable and a continuous variable.

Select Data -> Insert -> Charts -> ...





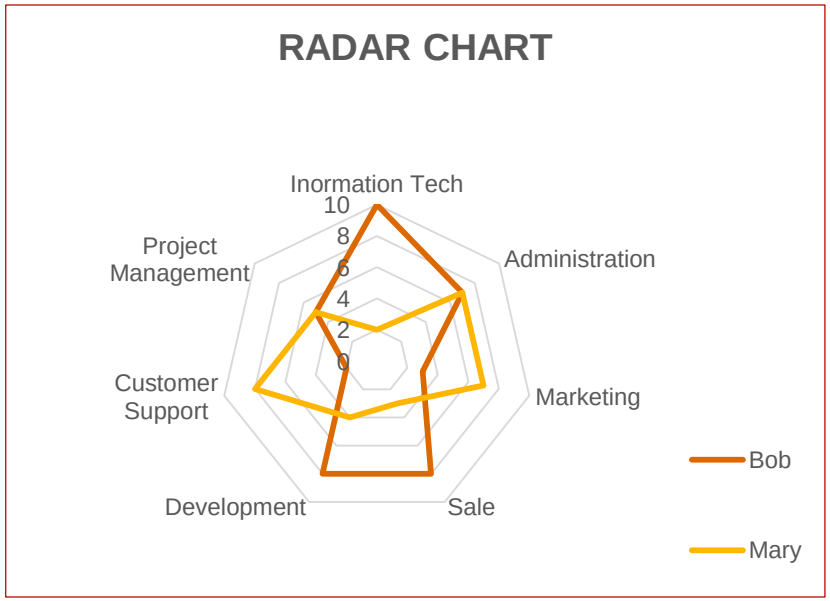
Data Visualisation

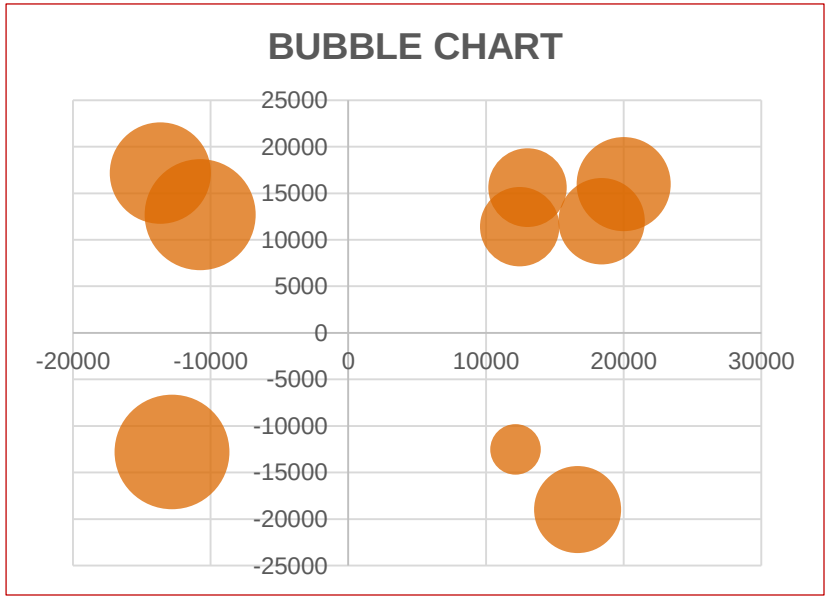
Types of visualisation by technique

Grids

Representing entities in a grid to map them according to two or more axes.

Select Data -> Insert -> Charts -> ...





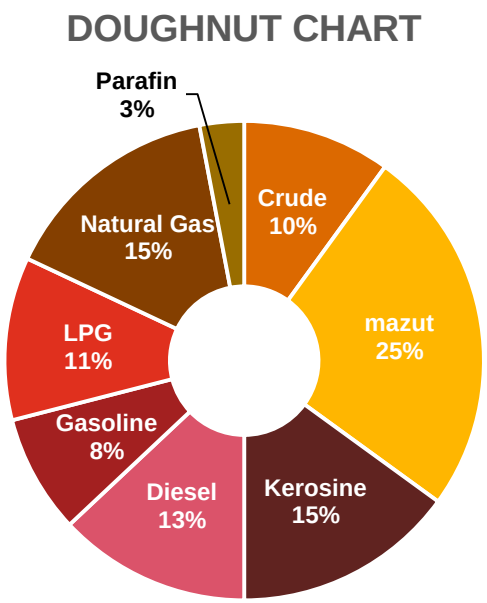
Data Visualisation

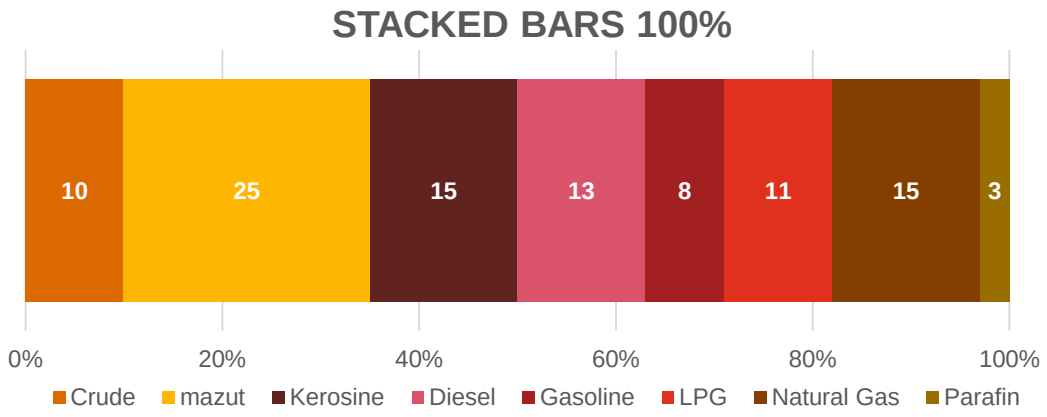
Types of visualisation by technique

Shape and proportions

Show proportions without reference to a coordinate system.

Select Data -> Insert -> Charts -> ...





Data Visualisation

Best practices

Keep these in mind

- **Identify your goal**
  Always start with a question
- **Know your audience**
  Tailor your visualisation to the intended viewer's expertise and expectations
- **Provide context**
  Provide the supporting context which makes your visualisation meaningful
- **Keep it simple**
  Don't include any unnecessary information and avoid clutter
- **Keep it engaging**
  Nobody pays attention to a boring visualisation
- **Consider the colour blind**
  Colour blindness is not uncommon and can render a visualisation useless

Introduction to Data Visualisation

- In practice

Contents

Introduction to Data Visualisation

- **Microsoft Excel hands on**

Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

Using interactively calculated data, we will create a dashboard visualising different metrics about a loan repayment.

INPUT DATA

Duration	300
Amount	€ 250.000,00
Interest	2,00%
First payment	01/11/2015

Month	Outstanding	Capital	Interest	Date line	*Total Payment*	Remaining Amount	Percent of Owners	Capital Owned	Interest Paid
Begin	€ 250.000,00					€ 250.000,00			
1/11/2015	€ 250.000,00	€ 644,54	€ 412,90		*€ 1.057,43*	€ 249.355,46	0,26%	€ 644,54	€ 412,90
1/12/2015	€ 249.355,46	€ 645,60	€ 411,83		*€ 1.057,43*	€ 248.709,86	0,52%	€ 1.290,14	€ 824,73
1/01/2016	€ 248.709,86	€ 646,67	€ 410,76		*€ 1.057,43*	€ 248.063,19	0,77%	€ 1.936,81	€ 1.235,49
1/02/2016	€ 248.063,19	€ 647,74	€ 409,70		*€ 1.057,43*	€ 247.415,45	1,03%	€ 2.584,55	€ 1.645,19
1/03/2016	€ 247.415,45	€ 648,81	€ 408,63		*€ 1.057,43*	€ 246.766,64	1,29%	€ 3.233,36	€ 2.053,81
1/04/2016	€ 246.766,64	€ 649,88	€ 407,56		*€ 1.057,43*	€ 246.116,77	1,55%	€ 3.883,23	€ 2.461,37
1/05/2016	€ 246.116,77	€ 650,95	€ 406,48		*€ 1.057,43*	€ 245.465,81	1,81%	€ 4.534,19	€ 2.867,85
1/06/2016	€ 245.465,81	€ 652,03	€ 405,41		*€ 1.057,43*	€ 244.813,79	2,07%	€ 5.186,21	€ 3.273,26
1/07/2016	€ 244.813,79	€ 653,10	€ 404,33		*€ 1.057,43*	€ 244.160,68	2,34%	€ 5.839,32	€ 3.677,59
1/08/2016	€ 244.160,68	€ 654,18	€ 403,25		*€ 1.057,43*	€ 243.506,50	2,60%	€ 6.493,50	€ 4.080,84
1/09/2016	€ 243.506,50	€ 655,26	€ 402,17		*€ 1.057,43*	€ 242.851,24	2,86%	€ 7.148,76	€ 4.483,01
1/10/2016	€ 242.851,24	€ 656,35	€ 401,09		*€ 1.057,43*	€ 242.194,89	3,12%	€ 7.805,11	€ 4.884,10
1/11/2016	€ 242.194,89	€ 657,43	€ 400,00		*€ 1.057,43*	€ 241.537,46	3,39%	€ 8.462,54	€ 5.284,10
1/12/2016	€ 241.537,46	€ 658,52	€ 398,92	€ 1.057,43	*€ 1.057,43*	€ 240.878,95	3,65%	€ 9.121,05	€ 5.683,02
1/01/2017	€ 240.878,95	€ 659,60	€ 397,83		*€ 1.057,43*	€ 240.219,34	3,91%	€ 9.780,66	€ 6.080,85
1/02/2017	€ 240.219,34	€ 660,69	€ 396,74		*€ 1.057,43*	€ 239.558,65	4,18%	€ 10.441,35	€ 6.477,59

Numberof Months	300
Borrowed Amount	250000
Annual Interest	0,02
Monthly interest	0,1651581302%
Monthly installment	€ 1.057,43

Month of calculation	01/12/2016
Capital Owned	€ 9.121,05
Interest Paid	€ 5.683,02

Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

On the "Payments" sheet we have all the data about the repayment of the loan, including:

- "Month";
- "Outstanding Amount";
- "Monthly Instalment" and its breakdown into "Capital" and "Interest";
- "Percent of Ownership";
- and the amount of "Capital Owned" and "Interest Paid".

Month	Outstanding	Capital	Interest	Date line	*Total Payment*	Remaining Amount	Percent of Owners	Capital Owned	Interest Paid
Begin	€ 250.000,00					€ 250.000,00			
1/11/2015	€ 250.000,00	€ 644,54	€ 412,90		*€ 1.057,43*	€ 249.355,46	0,26%	€ 644,54	€ 412,90
1/12/2015	€ 249.355,46	€ 645,60	€ 411,83		*€ 1.057,43*	€ 248.709,86	0,52%	€ 1.290,14	€ 824,73
1/01/2016	€ 248.709,86	€ 646,67	€ 410,76		*€ 1.057,43*	€ 248.063,19	0,77%	€ 1.936,81	€ 1.235,49
1/02/2016	€ 248.063,19	€ 647,74	€ 409,70		*€ 1.057,43*	€ 247.415,45	1,03%	€ 2.584,55	€ 1.645,19
1/03/2016	€ 247.415,45	€ 648,81	€ 408,63		*€ 1.057,43*	€ 246.766,64	1,29%	€ 3.233,36	€ 2.053,81
1/04/2016	€ 246.766,64	€ 649,88	€ 407,56		*€ 1.057,43*	€ 246.116,77	1,55%	€ 3.883,23	€ 2.461,37
1/05/2016	€ 246.116,77	€ 650,95	€ 406,48		*€ 1.057,43*	€ 245.465,81	1,81%	€ 4.534,19	€ 2.867,85
1/06/2016	€ 245.465,81	€ 652,03	€ 405,41		*€ 1.057,43*	€ 244.813,79	2,07%	€ 5.186,21	€ 3.273,26
1/07/2016	€ 244.813,79	€ 653,10	€ 404,33		*€ 1.057,43*	€ 244.160,68	2,34%	€ 5.839,32	€ 3.677,59
1/08/2016	€ 244.160,68	€ 654,18	€ 403,25		*€ 1.057,43*	€ 243.506,50	2,60%	€ 6.493,50	€ 4.080,84
1/09/2016	€ 243.506,50	€ 655,26	€ 402,17		*€ 1.057,43*	€ 242.851,24	2,86%	€ 7.148,76	€ 4.483,01
1/10/2016	€ 242.851,24	€ 656,35	€ 401,09		*€ 1.057,43*	€ 242.194,89	3,12%	€ 7.805,11	€ 4.884,10
1/11/2016	€ 242.194,89	€ 657,43	€ 400,00		*€ 1.057,43*	€ 241.537,46	3,39%	€ 8.462,54	€ 5.284,10
1/12/2016	€ 241.537,46	€ 658,52	€ 398,92	€ 1.057,43	*€ 1.057,43*	€ 240.878,95	3,65%	€ 9.121,05	€ 5.683,02
1/01/2017	€ 240.878,95	€ 659,60	€ 397,83		*€ 1.057,43*	€ 240.219,34	3,91%	€ 9.780,66	€ 6.080,85
1/02/2017	€ 240.219,34	€ 660,69	€ 396,74		*€ 1.057,43*	€ 239.558,65	4,18%	€ 10.441,35	€ 6.477,59

Numberof Months	300
Borrowed Amount	250000
Annual Interest	0,02
Monthly interest	0,1651581302%
Monthly installment	€ 1.057,43

Month of calculation	01/12/2016
Capital Owned	€ 9.121,05
Interest Paid	€ 5.683,02

Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

Furthermore on the "Payments" sheet we have an assortment of cells that we will need to create our dashboard, these include:

- "Number of Months";
- "Borrowed Amount";
- "Annual interest";
- "Monthly Installment";
- "Capital Owned";
- "Interest Paid";
- "Number of Payments";
- "Remaining Payments";
- "Amount Paid";
- "Amount remaining";
- and "Percentage Paid".

Numberof Months	300
Borrowed Amount	250000
Annual Interest	0,02
Monthly interest	0,1651581302%
Monthly installment	€ 1.057,43

Month of calculation	01/12/2016
Capital Owned	€ 9.121,05
Interest Paid	€ 5.683,02
Number of Payments	16
Remaining Payments	284

Amount Paid	Amount Remaining	Total Amount	Percentage Paid
€ 9.121,05	€ 240.878,95	€ 250.000,00	3,65%

Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

On the “Dashboard” sheet, we have a convenient input box, where we submit:

- the “Duration”, in months, of the loan;
- the “Amount” of the loan;
- the “Interest” rate;
- the date of the “First Payment”;
- and the current “DATE”.

INPUT DATA	
Duration	300
Amount	€ 250.000,00
Interest	2,00%
First payment	01/11/2015
DATE	07/12/2016

Any changes done to the “Input Data” **automatically update the data** on the “Payments” sheet.

Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

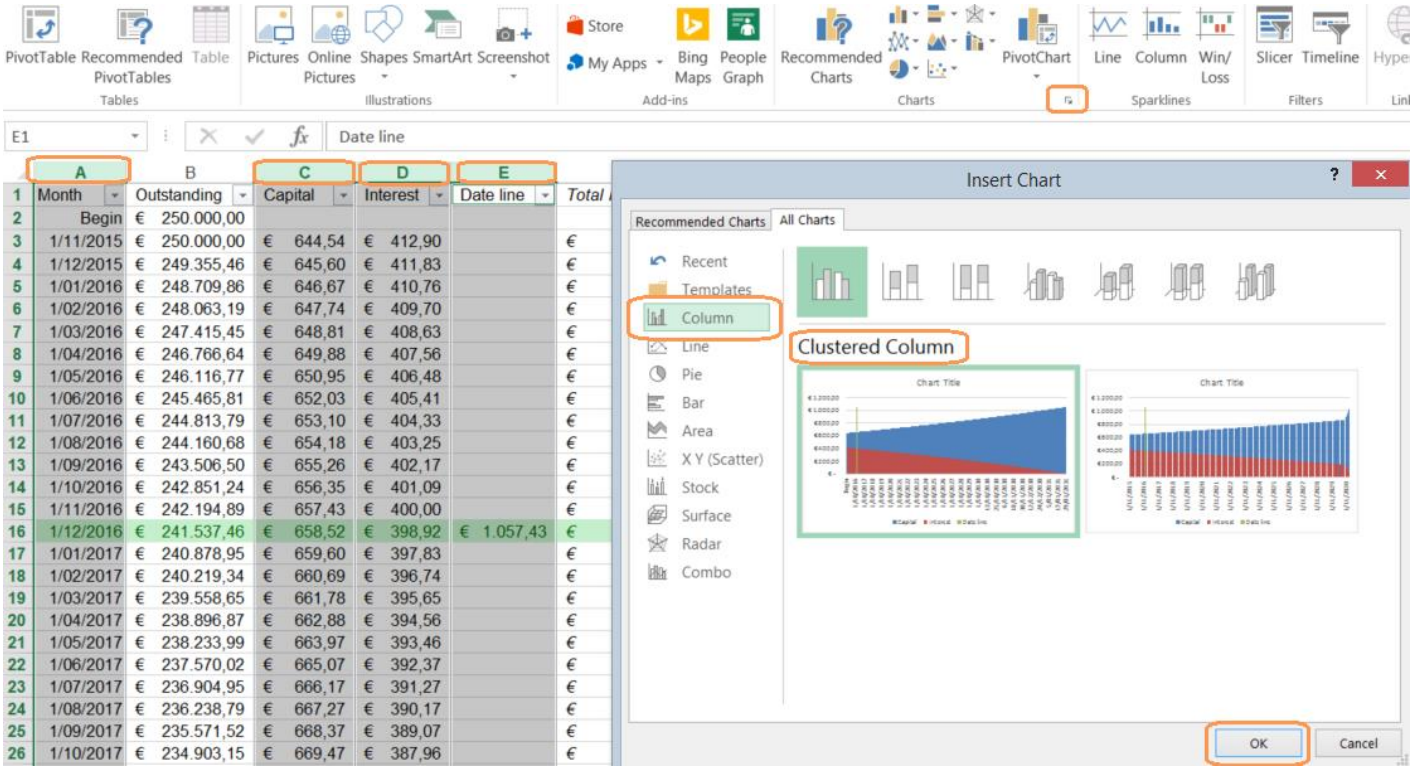
Let's create a chart showing the amount of Capital we acquire relative to the amount of interest we pay back with each instalment through time. Furthermore let's add a marker to show us where we are on that time series.

Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

Select columns A, C, D and E. Open up the “all charts” menu, and select “Clustered Column” from the “Columns” section.



	A	B	C	D	E	
1	Month	Outstanding	Capital	Interest	Date line	Total
2	Begin	€ 250.000,00				
3	1/11/2015	€ 250.000,00	€ 644,54	€ 412,90		€
4	1/12/2015	€ 249.355,46	€ 645,60	€ 411,83		€
5	1/01/2016	€ 248.709,86	€ 646,67	€ 410,76		€
6	1/02/2016	€ 248.063,19	€ 647,74	€ 409,70		€
7	1/03/2016	€ 247.415,45	€ 648,81	€ 408,63		€
8	1/04/2016	€ 246.766,64	€ 649,88	€ 407,56		€
9	1/05/2016	€ 246.116,77	€ 650,95	€ 406,48		€
10	1/06/2016	€ 245.465,81	€ 652,03	€ 405,41		€
11	1/07/2016	€ 244.813,79	€ 653,10	€ 404,33		€
12	1/08/2016	€ 244.160,68	€ 654,18	€ 403,25		€
13	1/09/2016	€ 243.506,50	€ 655,26	€ 402,17		€
14	1/10/2016	€ 242.851,24	€ 656,35	€ 401,09		€
15	1/11/2016	€ 242.194,89	€ 657,43	€ 400,00		€
16	1/12/2016	€ 241.537,46	€ 658,52	€ 398,92	€ 1.057,43	€
17	1/01/2017	€ 240.878,95	€ 659,60	€ 397,83		€
18	1/02/2017	€ 240.219,34	€ 660,69	€ 396,74		€
19	1/03/2017	€ 239.558,65	€ 661,78	€ 395,65		€
20	1/04/2017	€ 238.896,87	€ 662,88	€ 394,56		€
21	1/05/2017	€ 238.233,99	€ 663,97	€ 393,46		€
22	1/06/2017	€ 237.570,02	€ 665,07	€ 392,37		€
23	1/07/2017	€ 236.904,95	€ 666,17	€ 391,27		€
24	1/08/2017	€ 236.238,79	€ 667,27	€ 390,17		€
25	1/09/2017	€ 235.571,52	€ 668,37	€ 389,07		€
26	1/10/2017	€ 234.903,15	€ 669,47	€ 387,96		€

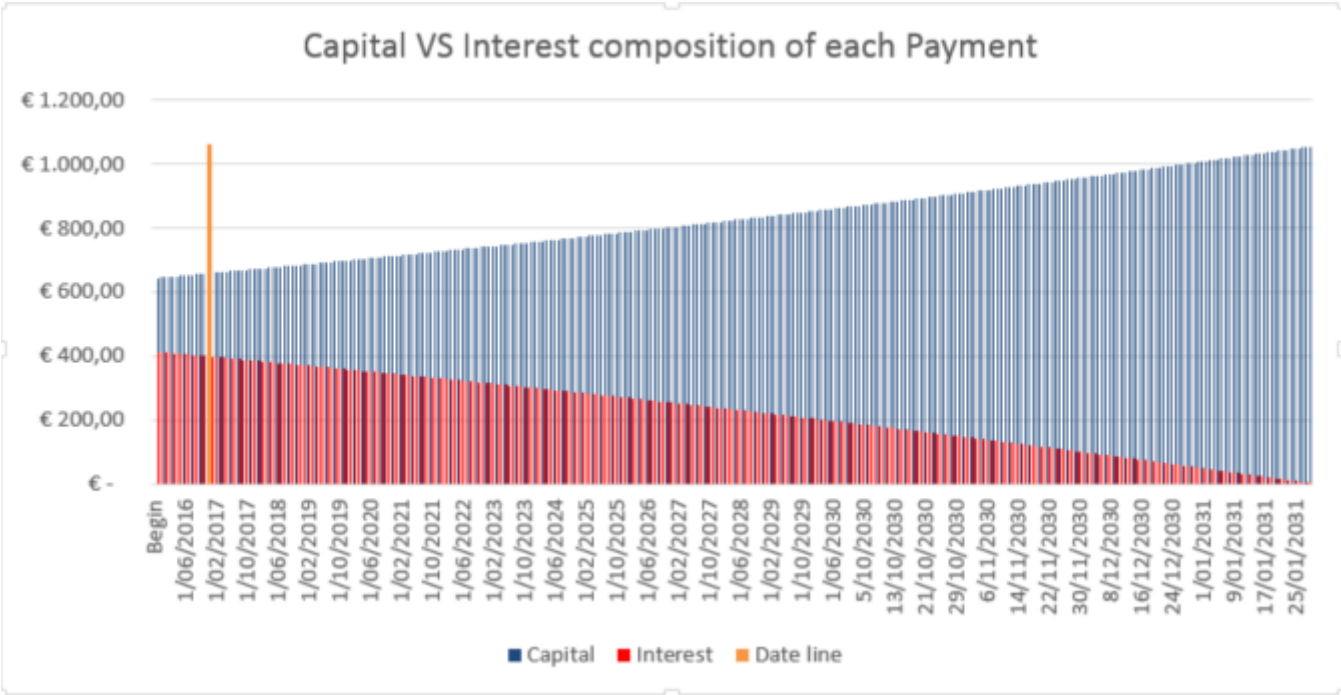
Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

The result is a chart that very conveniently displays the amount of Capital compared to the amount of Interest covered with each instalment.

Can you think of a different way this chart could be formatted?

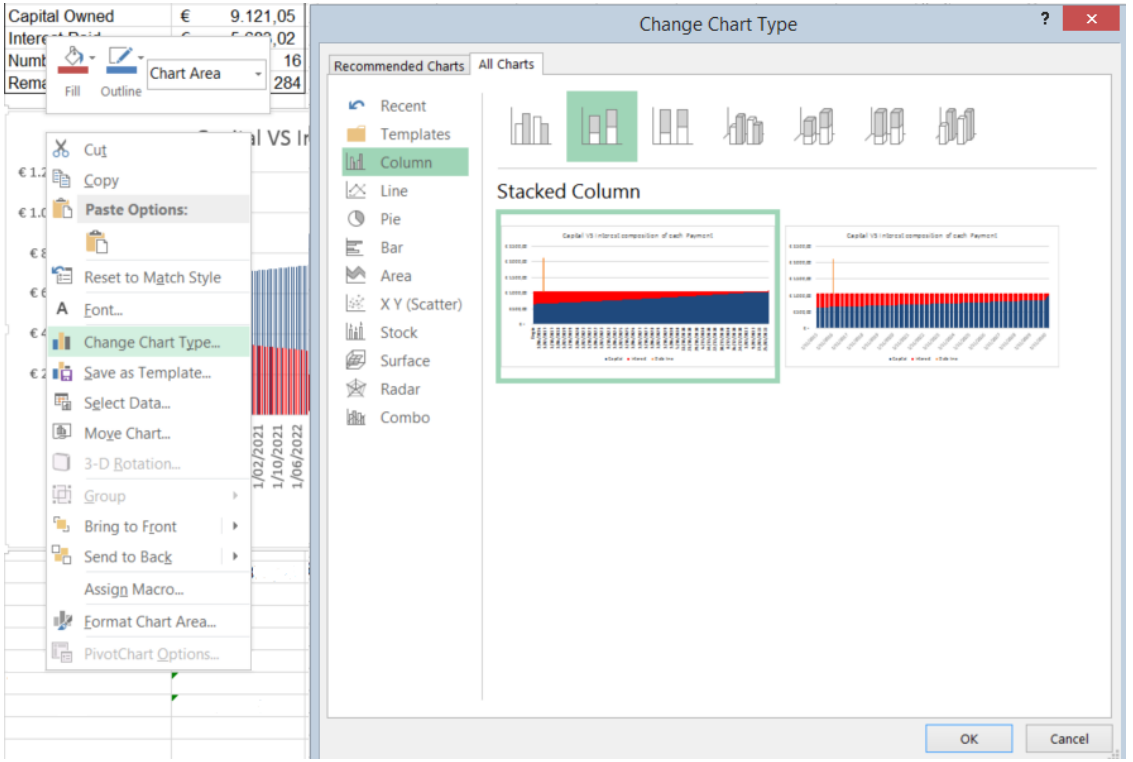


Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

Right click on the chart and click on "Change Chart Type", from the menu that opens select "Stacked Column" or "Stacked Area".

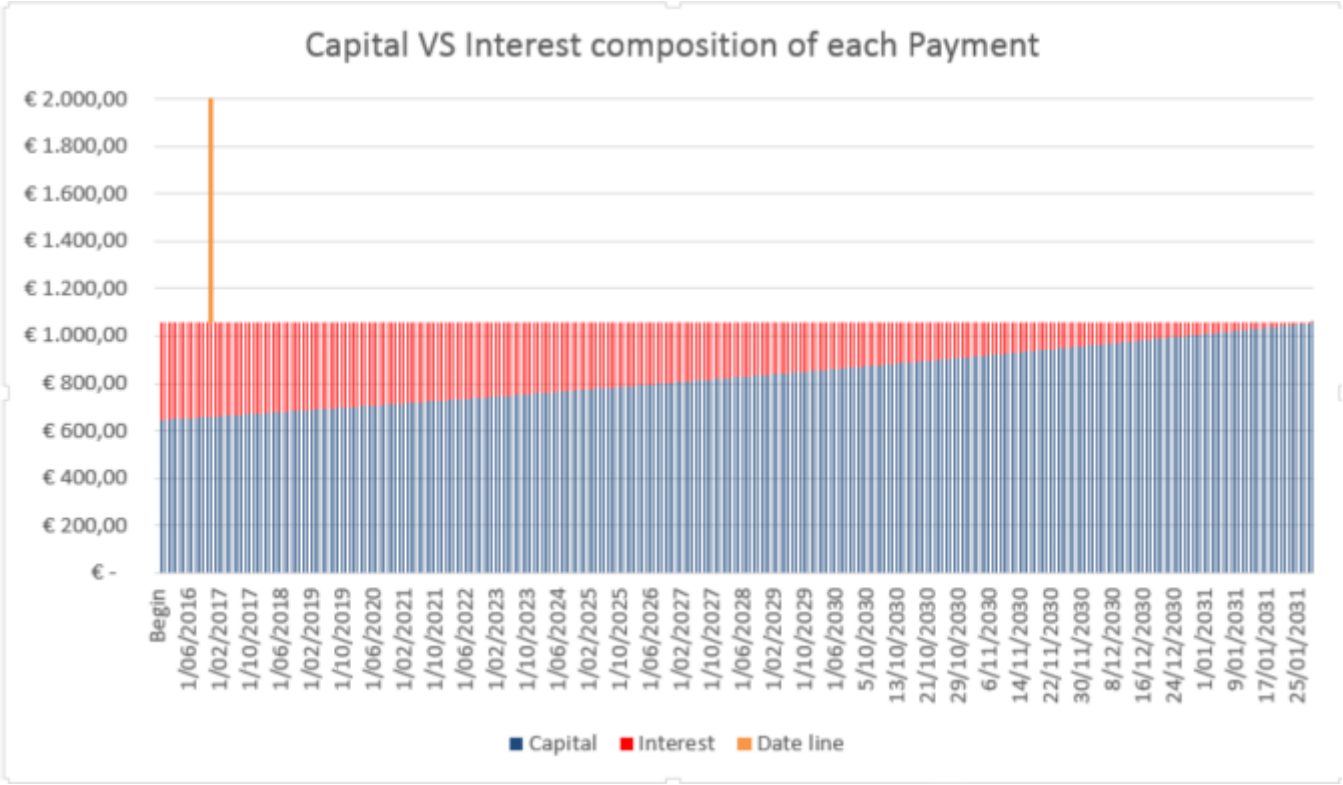


Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

The chart now changes to show how capital and interest compose the total of each instalment. Discuss which way could be best and why.

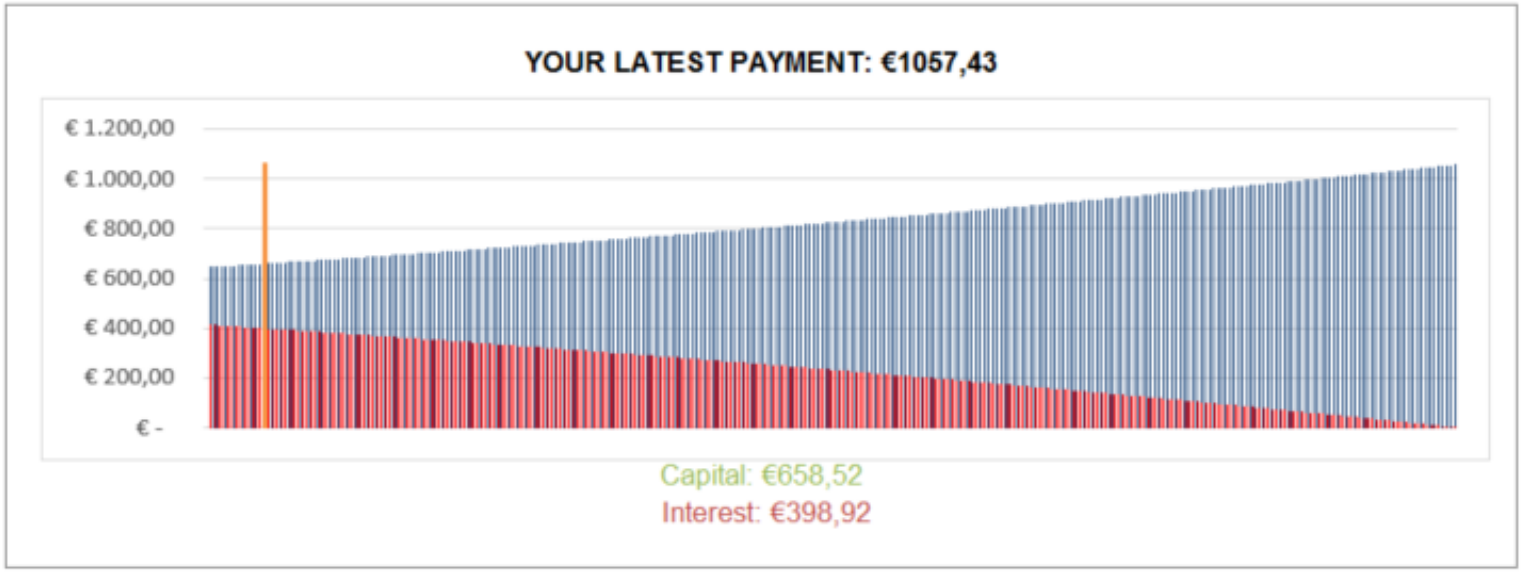


Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

Now cut and paste you chart into the “Dashboard” sheet. Remove the legends and titles, add in a cell above the amount of each instalment by referencing cell “M6” from the “Payments” sheet and add below two cells displaying the current composition of Capital and Interest in each payment by referencing the necessary cells.

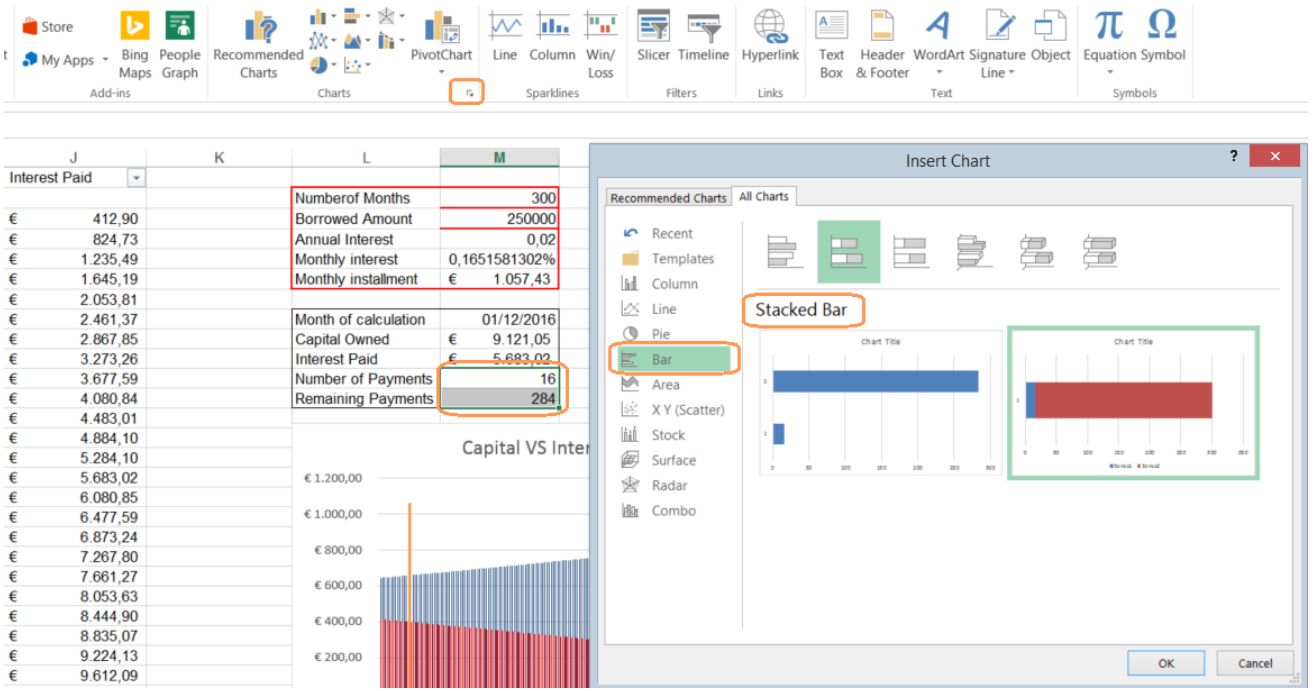


Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

Now select the cells containing the "Number of Payments" and "Remaining Payments". Open the charts menu and select "Stacked Bar" from "Bars" . Cut and paste the resulting chart into the "Dashboard" sheet.

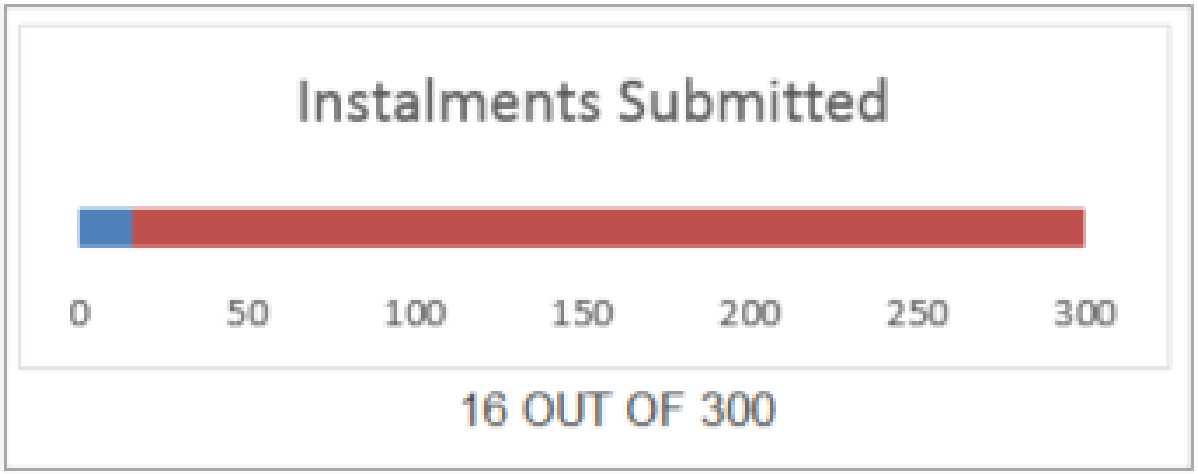


Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

Format the chart accordingly to look like the one below. Keep in mind that the label "16 out of 300" is actually a cell, referencing the "Number of Payments" and "Remaining Payments" values from the "Payments" sheet.

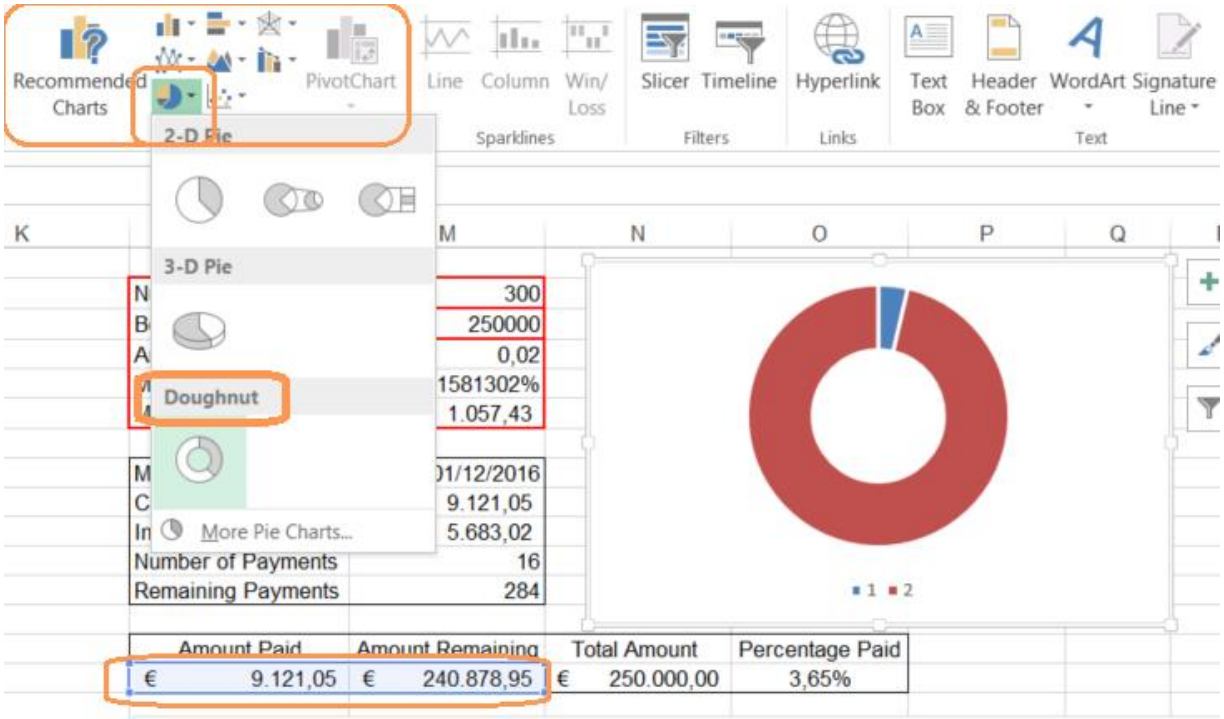


Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

Select the cells containing the "Amount Paid" and "Amount Remaining". Open the charts menu and select "Doughnut Chart". Cut and paste the chart onto the "Dashboard" sheet.

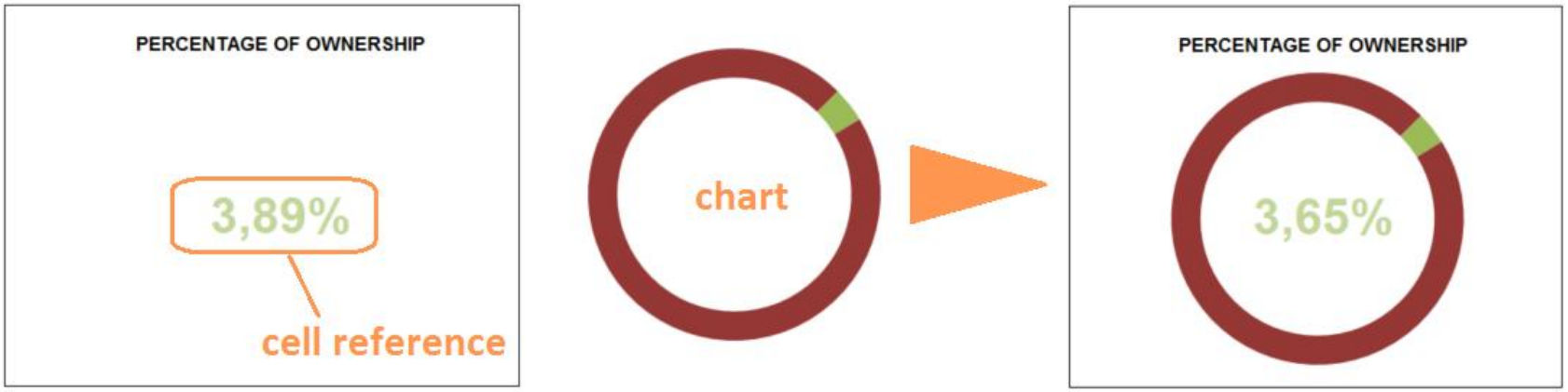


Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

Sometimes visualising requires imagination. In your dashboard create a cell with a reference to the "Percentage Paid" from your "Payments" sheet and drag your doughnut graph ON TOP of that. Now doesn't that look good?

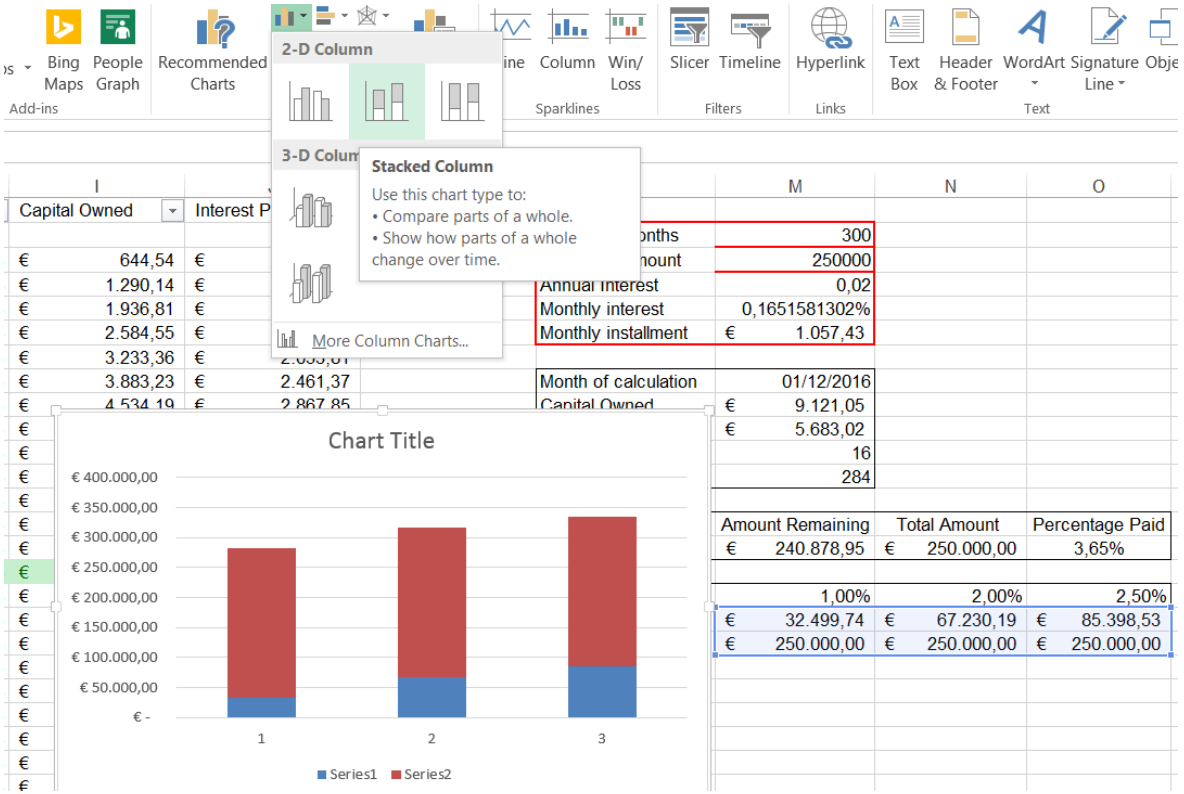


Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

Select the two rows containing the “Total Interest” and “Total Capital” from the Interest variance table and select **“Stacked Column Chart”**

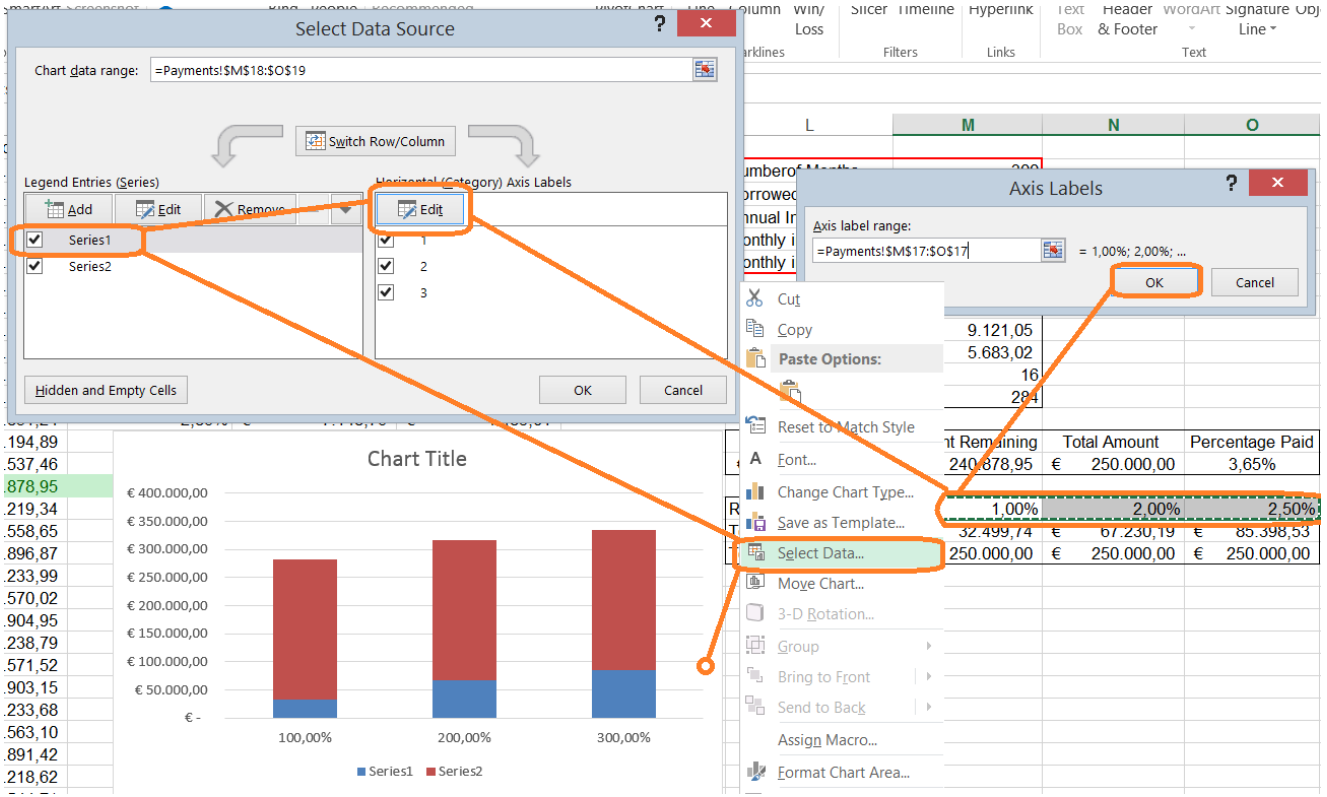


Visualising in MS Excel

Creating a dashboard in Excel

Loan repayment dashboard

Right click in the chart, “Select Data…”, and “Edit” “Series1”. Select the “Rates” and click “OK”. Copy paste the chart on your “Dashboards” sheet.

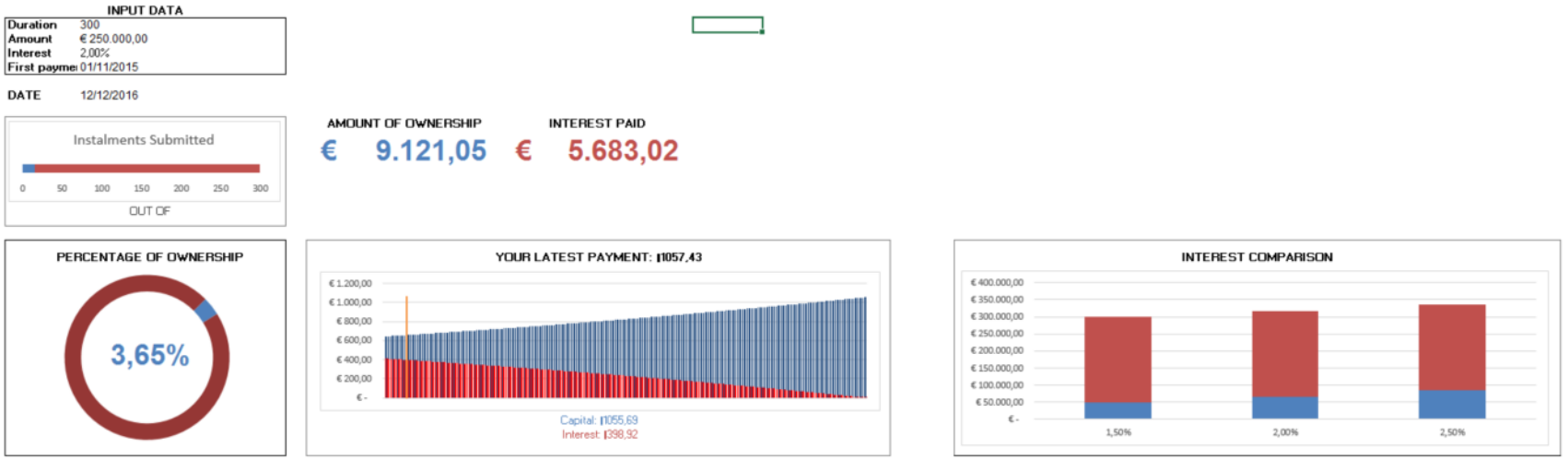


Visualising in MS Excel

Getting the data

Behold your creation

Arrange your graphs as it follows. Now isn't that a great looking dashboard? The best part is the interactivity. Whenever you change a value in the input, the whole dashboard updates live. Great for exploring your options!



Visualising in MS Excel
Getting the data

Loan repayment dashboard

Let's add one last chart that will:

- visualise three scenarios for the future value of the house;
- and based on that calculate the profit from selling the house at different moments in time.

Visualising in MS Excel

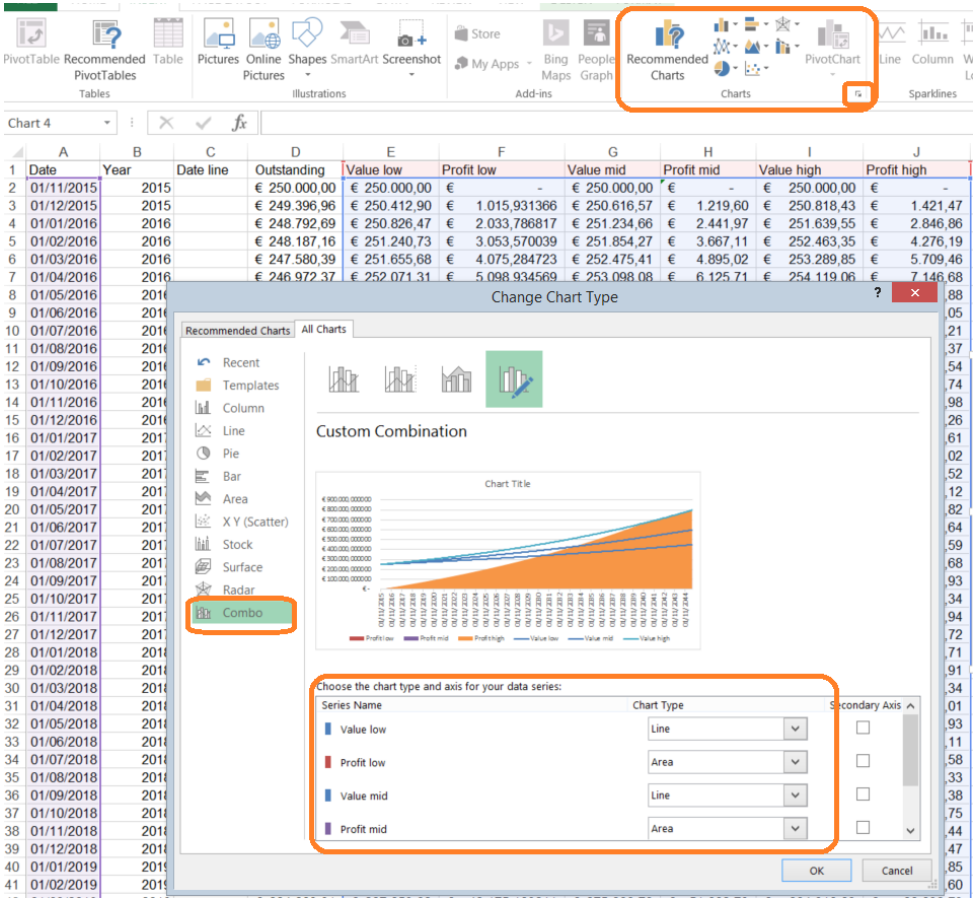
Getting the data

Loan repayment dashboard

Select the “Date” , “Value” and “Profit” columns from the “Sales forecast detail” sheet.

Open the charts menu and select “Combo” chart.

From the customisation box configure the “Value” series as “Line” and the “Profit” series as “Area”.

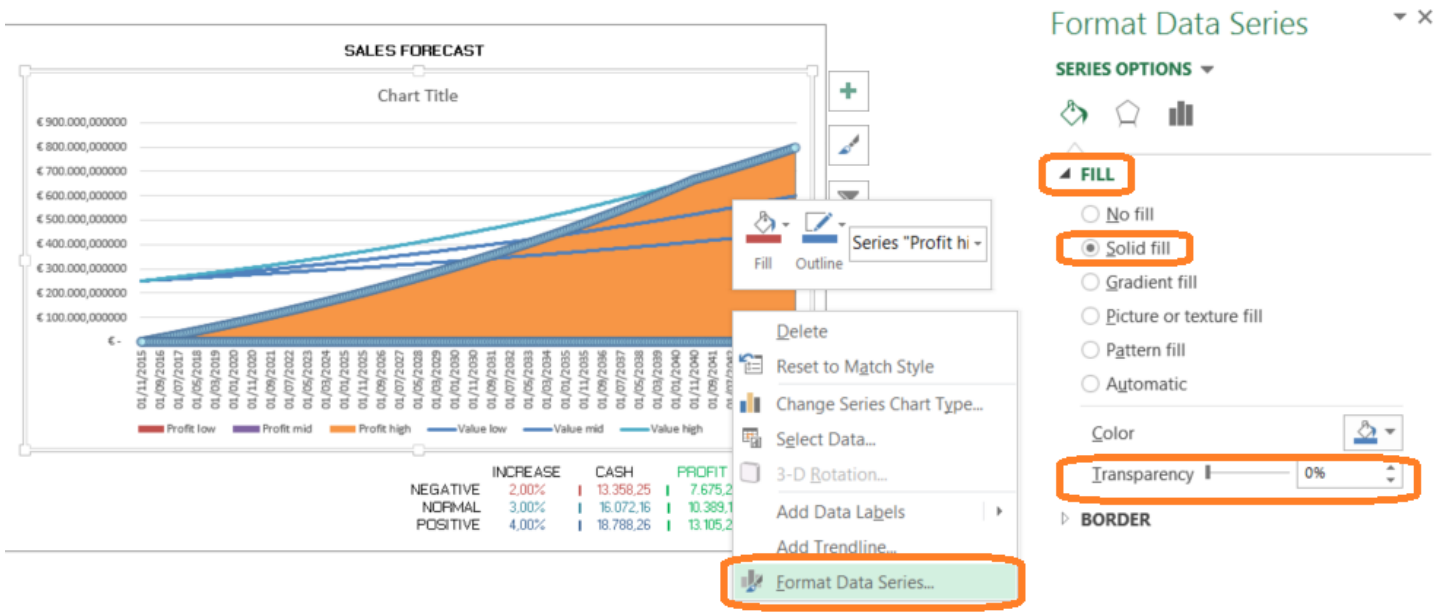


Visualising in MS Excel

Getting the data

Loan repayment dashboard

Right click on any of the "Area" series and select "Format Data Series". From the menu change the fill to "Solid" and add "Transparency". Do this for all three of the "Area" series until you can see all of them.

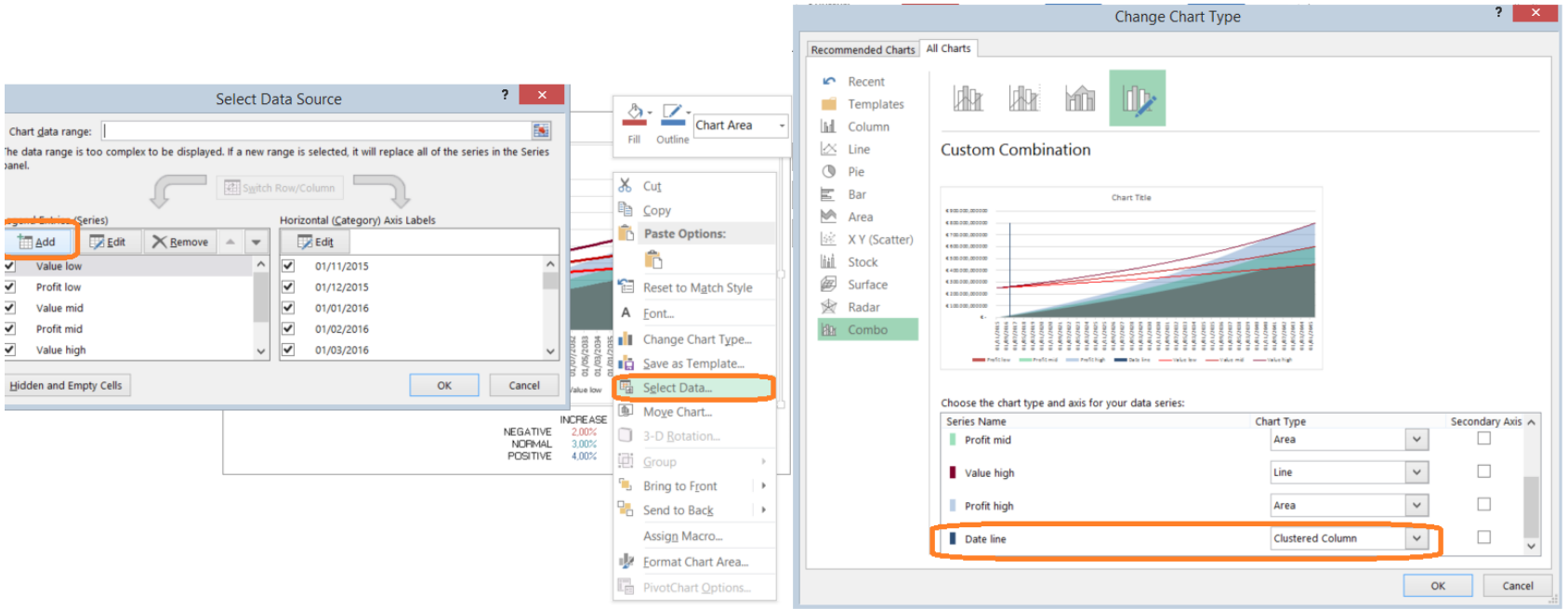


Visualising in MS Excel

Getting the data

Loan repayment dashboard

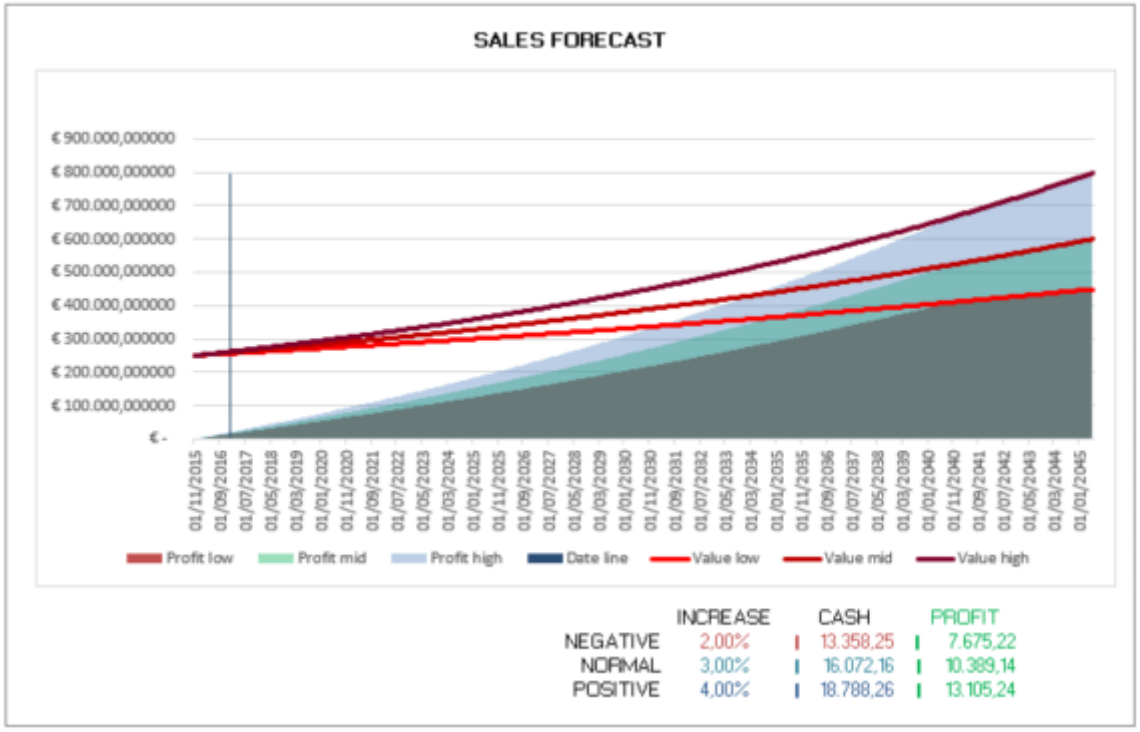
Now right click on the graph select “Select data”, click on “Add”, select the “Date line” column from the “Forecast” sheet, click “Ok”. Right click again on the graph, select “Change Chart Type” and configure the “Date line” as a “Column”.



Visualising in MS Excel
Getting the data

Loan repayment dashboard

Cut and paste your chart into the "Dashboard" sheet and format the colours until all the series are clear.

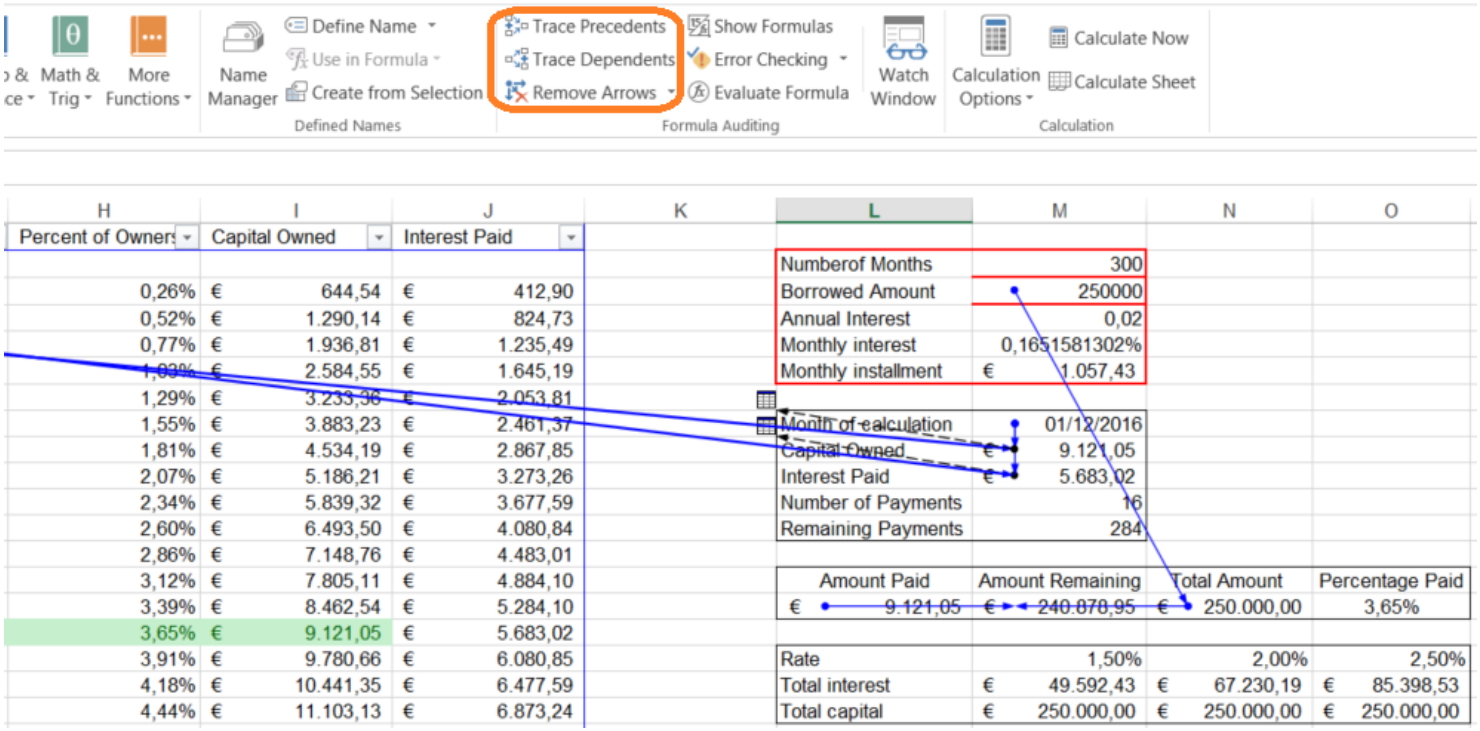


Visualising in MS Excel

Getting the data

Trace Precedent and Dependent cells in Excel

Things can get pretty complicated in such projects so don't forget to use the "Trace Precedents and Dependents" buttons.



Percent of Owners	Capital Owned	Interest Paid
0,26%	€ 644,54	€ 412,90
0,52%	€ 1.290,14	€ 824,73
0,77%	€ 1.936,81	€ 1.235,49
1,03%	€ 2.584,55	€ 1.645,19
1,29%	€ 3.233,36	€ 2.053,81
1,55%	€ 3.883,23	€ 2.461,37
1,81%	€ 4.534,19	€ 2.867,85
2,07%	€ 5.186,21	€ 3.273,26
2,34%	€ 5.839,32	€ 3.677,59
2,60%	€ 6.493,50	€ 4.080,84
2,86%	€ 7.148,76	€ 4.483,01
3,12%	€ 7.805,11	€ 4.884,10
3,39%	€ 8.462,54	€ 5.284,10
3,65%	€ 9.121,05	€ 5.683,02
3,91%	€ 9.780,66	€ 6.080,85
4,18%	€ 10.441,35	€ 6.477,59
4,44%	€ 11.103,13	€ 6.873,24

Numberof Months	300
Borrowed Amount	250000
Annual Interest	0,02
Monthly interest	0,1651581302%
Monthly installment	€ 1.057,43
Month of calculation	01/12/2016
Capital Owned	€ 9.121,05
Interest Paid	€ 5.683,02
Number of Payments	16
Remaining Payments	284

Amount Paid	Amount Remaining	Total Amount	Percentage Paid
€ 9.121,05	€ 240.878,95	€ 250.000,00	3,65%

Rate	1,50%	2,00%	2,50%
Total interest	€ 49.592,43	€ 67.230,19	€ 85.398,53
Total capital	€ 250.000,00	€ 250.000,00	€ 250.000,00

Visualising in MS Excel

Creating a dashboard in Excel

Importance of data preparation

The exercise we just completed truly illustrated the importance of proper data preparation.

Imagine how much longer it would have taken if you didn't have the variables we needed pre-thought out and conveniently calculated in prepared cells.

Introduction to Data Visualisation

- Effective Visualisation

Effective visualisation

Effective Communication

Sending a message VS Displaying information

Effective communication is getting messages across. This means getting the audience to understand something.

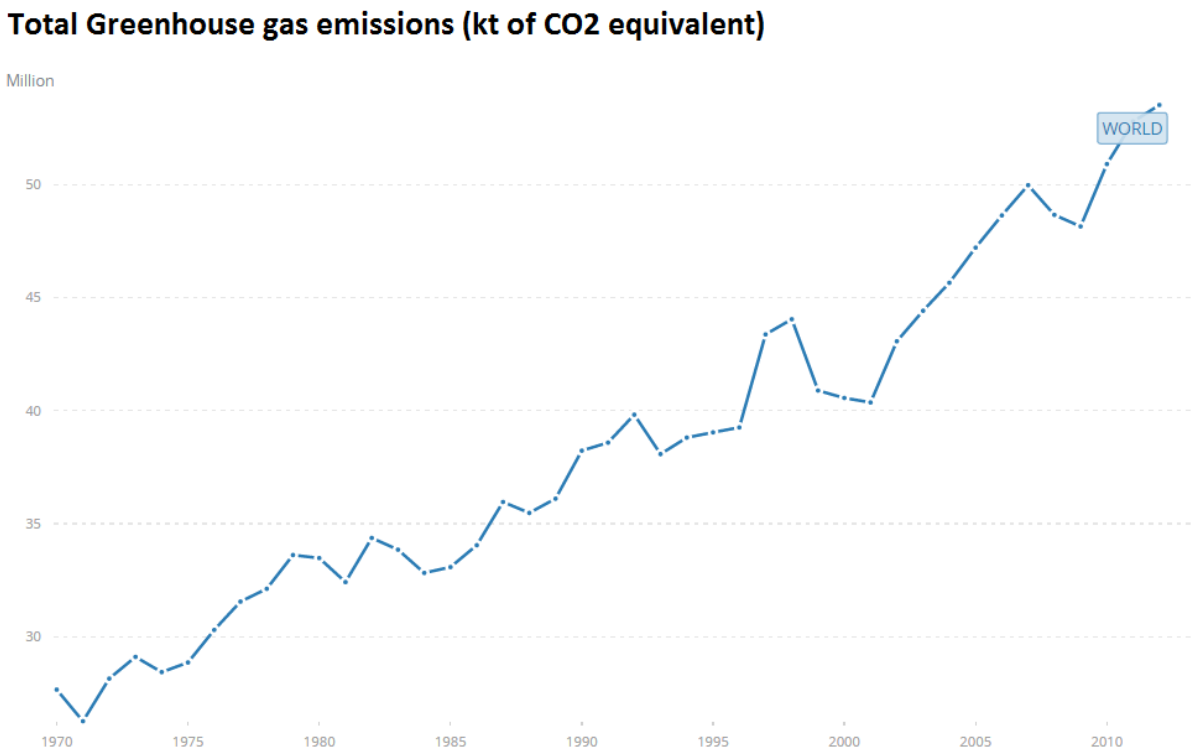
A message differs from raw information in that it presents **“intelligent added value”,** that is, something to understand about the information.

A message interprets the information for a specific audience and for a specific purpose. It conveys the **so what**, whereas information merely conveys the what. Because it makes a statement, it requires a complete sentence.

Effective visualisation

Effective Communication

"Total greenhouse gas emissions are calculated at 53,526,302 kT"

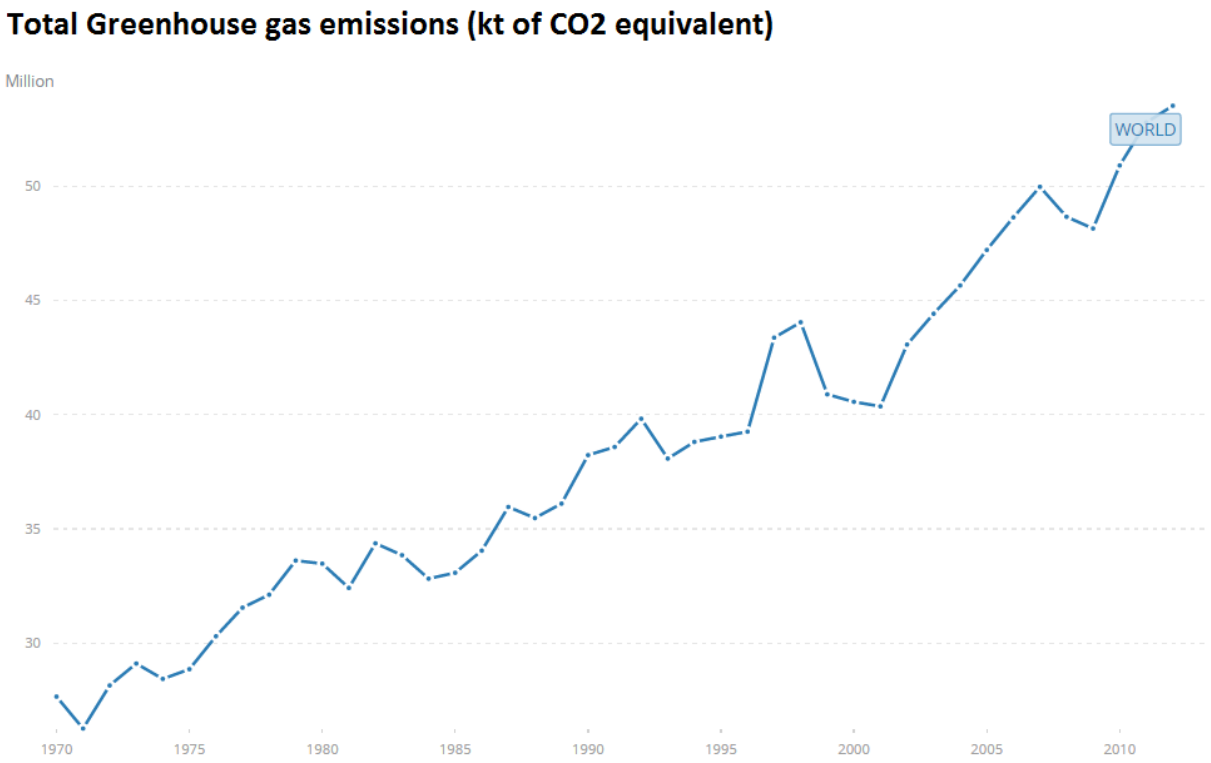


Select Data -> Line Chart with Markers…

Effective visualisation

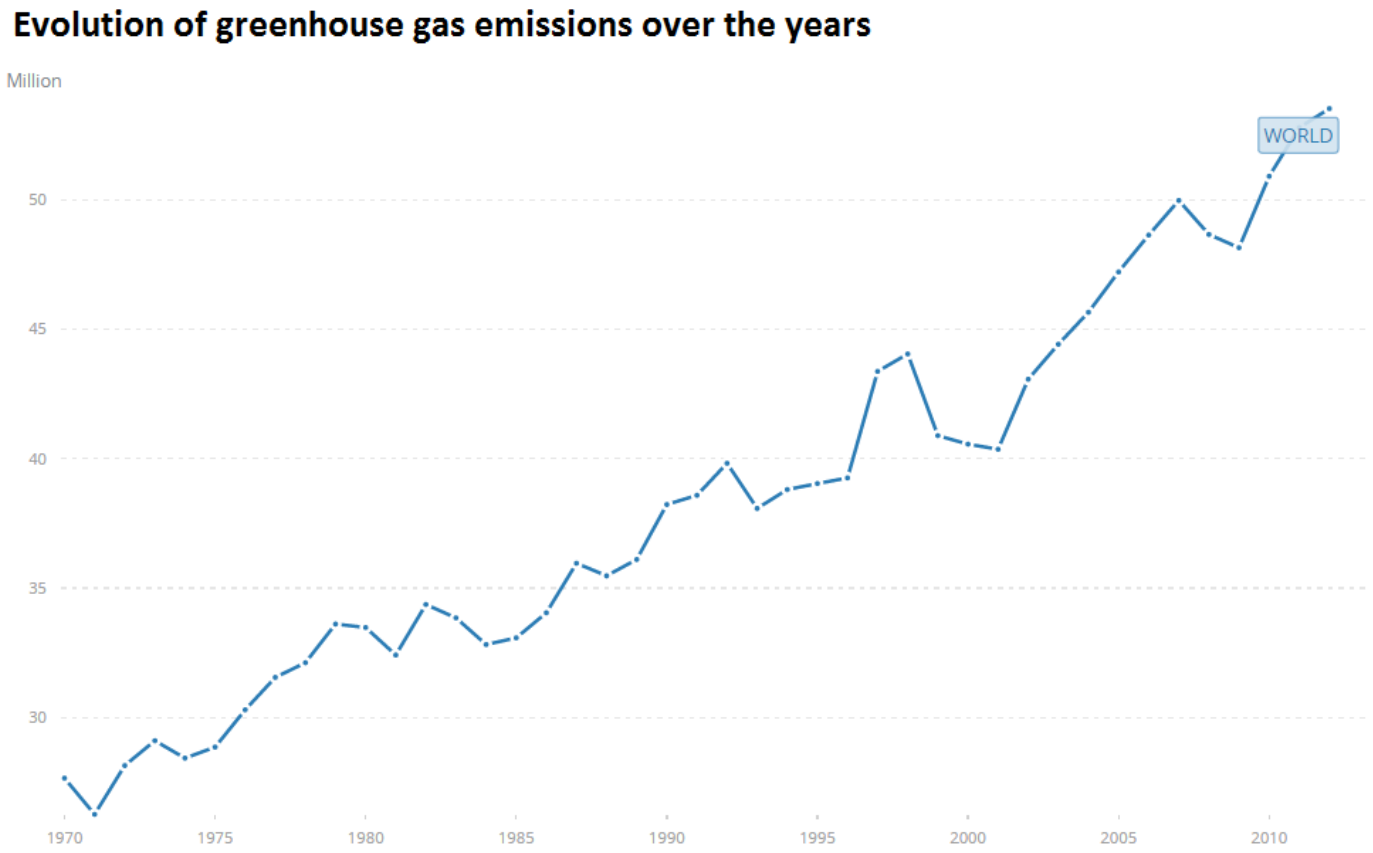
Effective Communication

"Total greenhouse gas emissions (53,526 Mt) are dangerously high!"



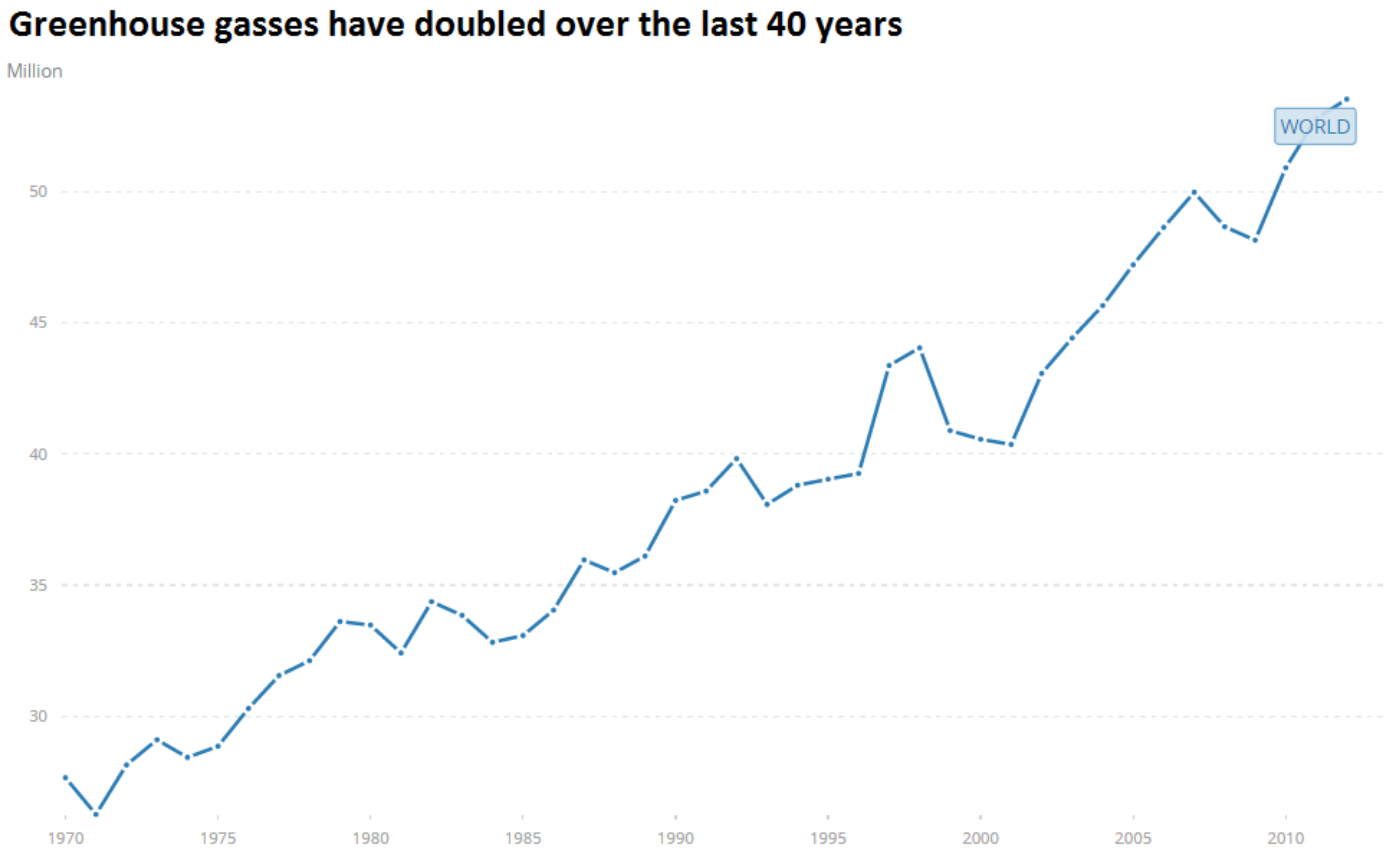
Effective visualisation
Effective Communication

"What"



Effective visualisation
Effective Communication

"So what"



Effective visualisation

Effective Communication

Know your audience

Information "A backup of the IT system is taken once per week"

Message

To a banking employee:
"When your system fails, you risk losing one week of work."

To the Legal Director:
"The backup schedule does not comply with National Bank regulations"

To the CIO:
"The backup interval should be increased to once per day."

Effective visualisation

Effective Communication

Adapt to your message!

System	Viruses	%
Windows	5	8%
Linux	25	37%
Mainframe	20	30%
HP	17	25%

Effective visualisation

Effective Communication

Adapt to your message!

Right click on Column -> Column width…

System	Viruses	%
Windows	5	8%
Linux	25	37%
Mainframe	20	30%
HP	17	25%

Effective visualisation

Effective Communication

Adapt to your message!

Select Columns -> Align Right…

System	Viruses	%
Windows	5	8%
Linux	25	37%
Mainframe	20	30%
HP	17	25%

Effective visualisation

Effective Communication

Adapt to your message!

Select Cells -> Right click -> Format Cells -> Remove Borders & Configure Font

System	Viruses	%
Windows	5	8%
Linux	25	37%
Mainframe	20	30%
HP	17	25%

Effective visualisation

Effective Communication

Adapt to your message!

Select Cells -> Select cells -> Conditional Formatting -> More Rules…

System	Viruses	%
Windows	5	8%
Linux	25	**37%**
Mainframe	20	30%
HP	17	25%

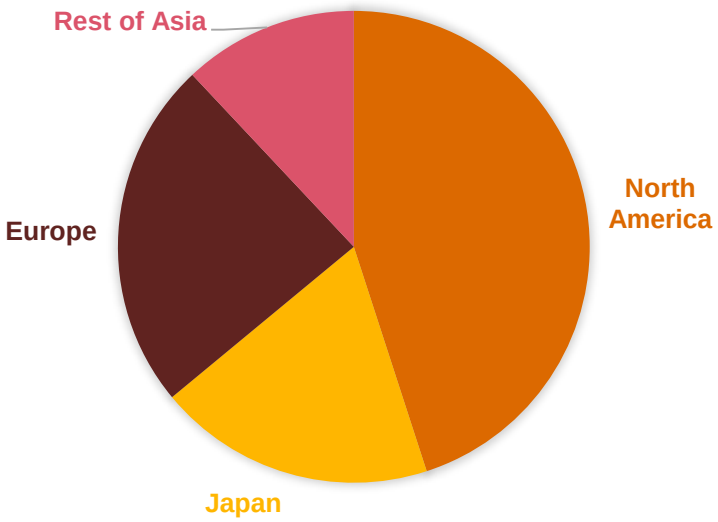
Effective visualisation

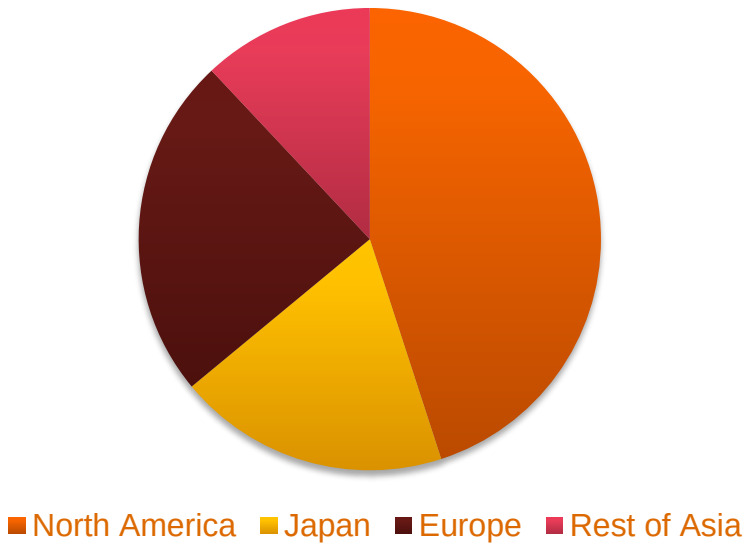
Effective Communication

Proper Labelling

Avoid legends. Label all your charts by placing the necessary words next to the items they describe. This makes reading the graph much more intuitive.

Select Data -> Pie Chart -> Right click -> Format Plot Area…





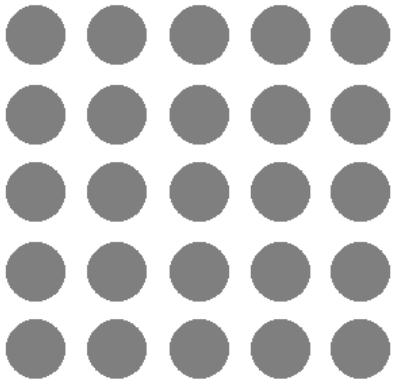
Effective visualisation

Effective Communication

Matching human intuition

Truly visual representations are in essence intuitive: they require no new interpretation rules, no verbal steps. Instead they are based on intuitive rules interpreting **proximity**, **similarity**, **prominence**, and **sequence**.

What do you see? Rows or Columns?



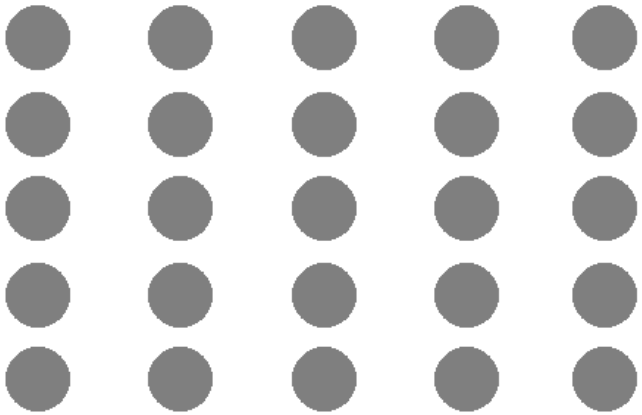
Effective visualisation

Effective Communication

Matching human intuition

Truly visual representations are in essence intuitive: they require no new interpretation rules, no verbal steps. Instead they are based on intuitive rules interpreting **proximity**, **similarity**, **prominence**, and **sequence**.

How about now?



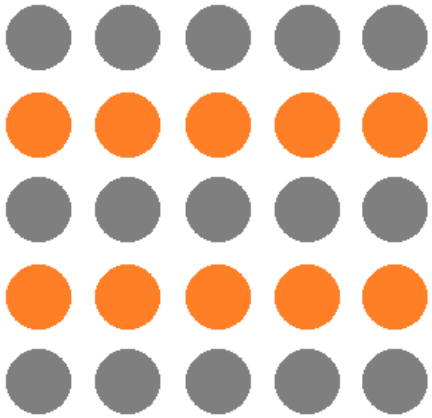
Effective visualisation

Effective Communication

Matching human intuition

Truly visual representations are in essence intuitive: they require no new interpretation rules, no verbal steps. Instead they are based on intuitive rules interpreting **proximity**, **similarity**, **prominence**, and **sequence**.

Or maybe now?



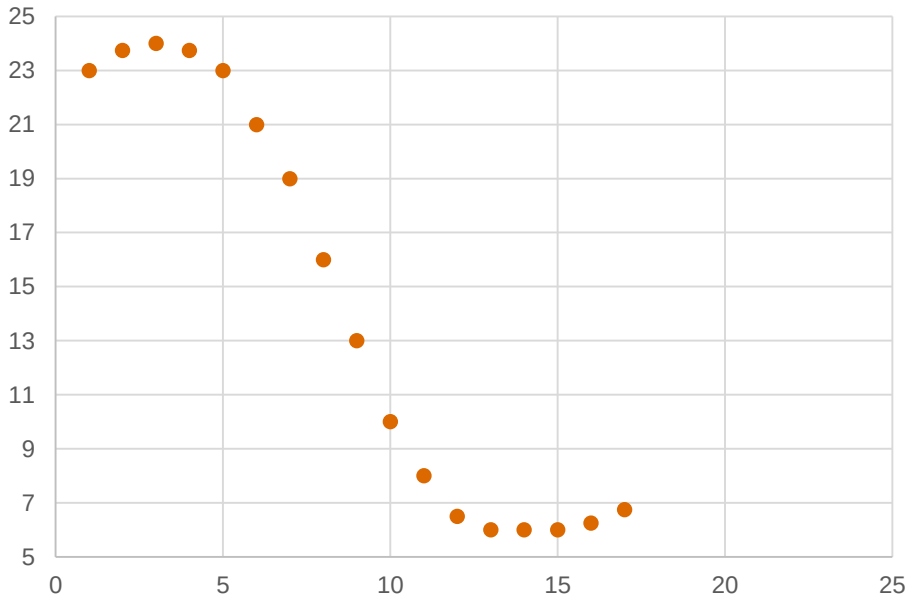
Effective visualisation

Effective Communication

Matching human intuition

A position representation need not start from zero, but one starting close to zero can mislead viewers.

Select Data -> Insert Chart -> X Y (Scatter)…

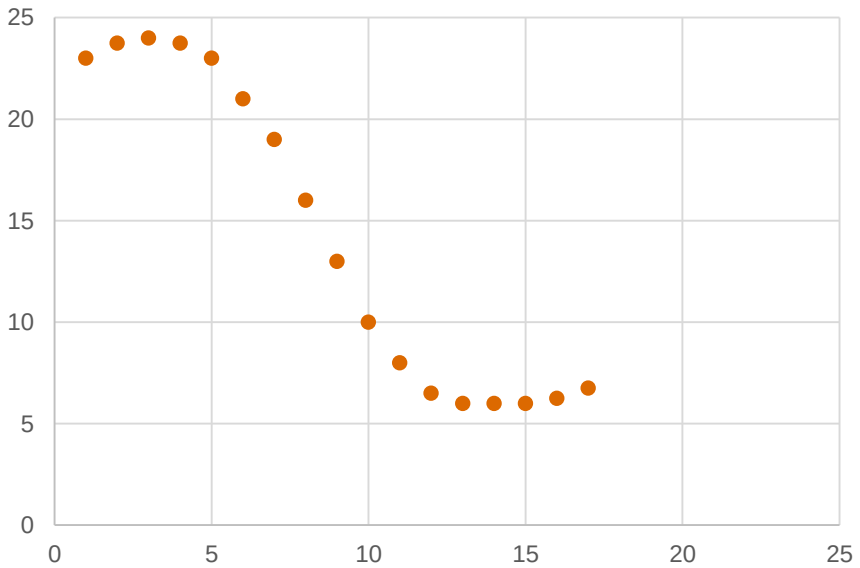


Effective visualisation
Effective Communication

Matching human intuition

It is best to extend the axis to zero for a more intuitive display.

Right click on legend -> Format Plot Area-> Axis options…

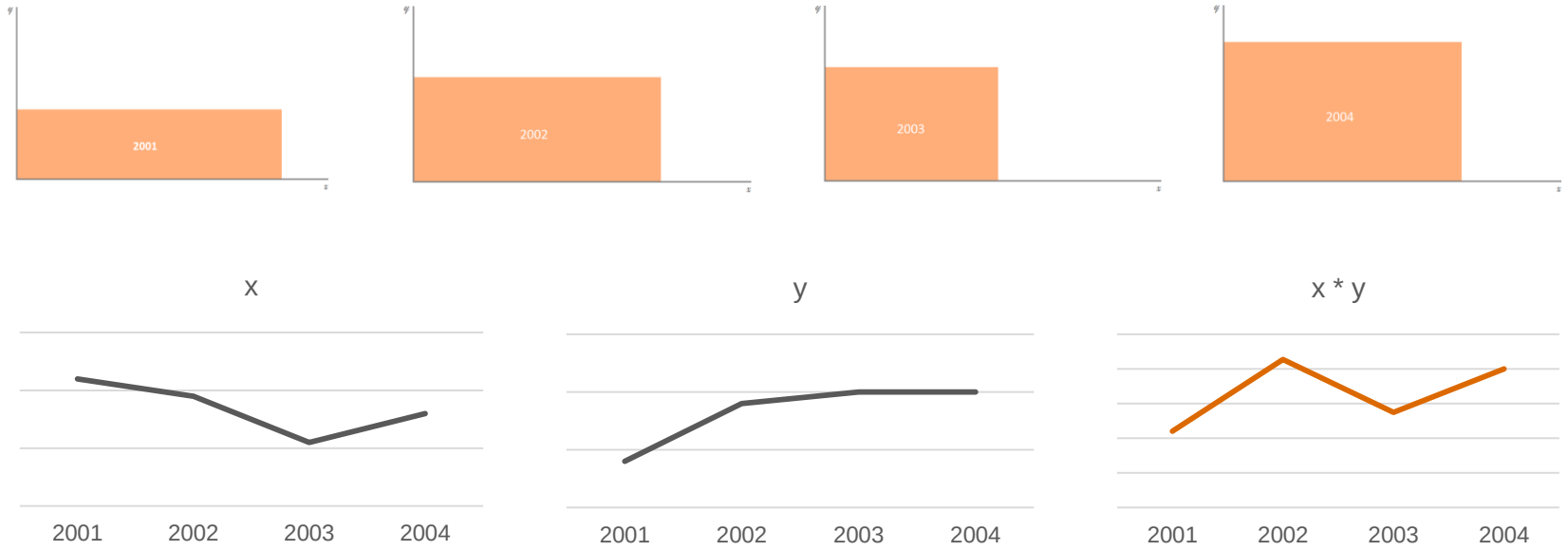


Effective visualisation
Effective Communication

Matching human intuition

A concurrent variation in two (or more) directions results in a hard-to-compare area representation.

Select Data -> Area Chart & Line Chart…



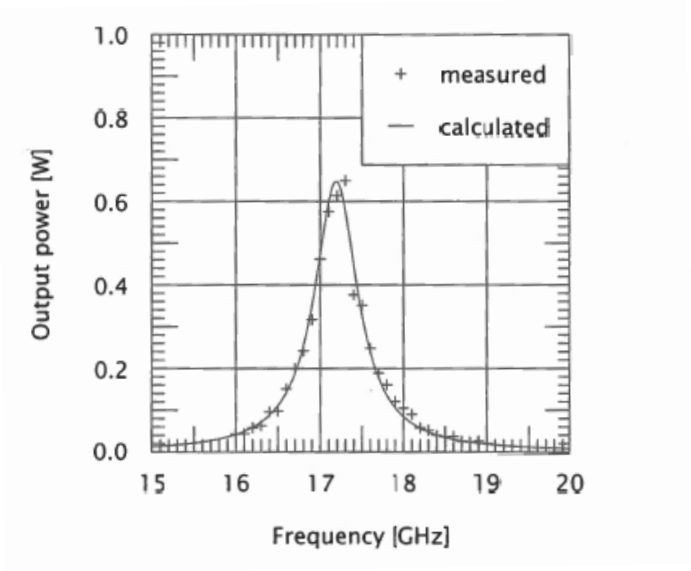
Effective visualisation
Effective Communication

Maximise the Signal-to-Noise ratio

A poor graph

The graph exhibits a very low signal-to-noise ratio, with excessive tick marks and uncalled-for grid lines, and very little emphasis on the data.

Select Data -> Line with Markers…



Effective visualisation
Effective Communication

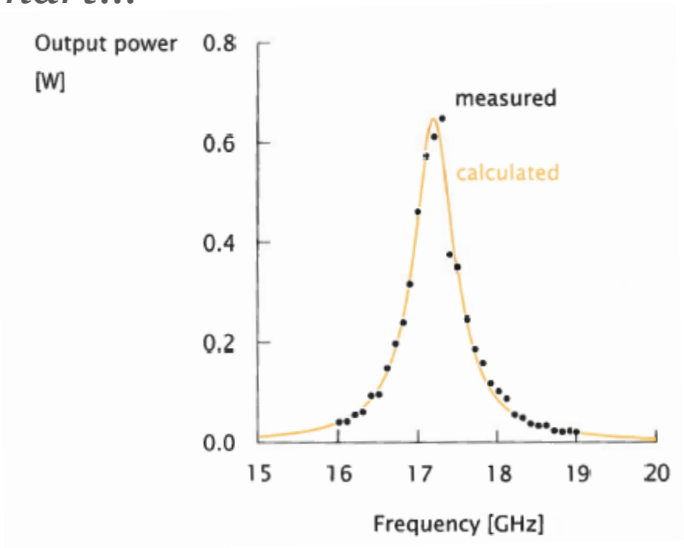
Maximise the Signal-to-Noise ratio

A good graph

The graph is plainer and better contrasted. The background no longer interferes with the data, yet it provides sufficient information about them.

The labels are intuitive by being placed where they are needed, next to the data.

Right click-> Format Chart…



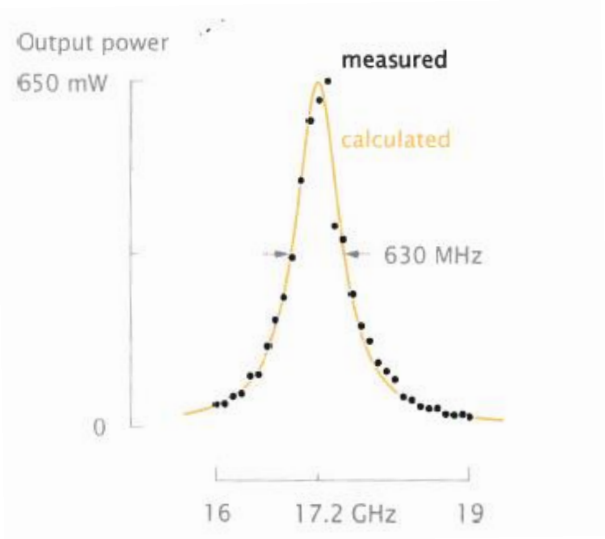
Effective visualisation
Effective Communication

Maximise the Signal-to-Noise ratio

A better graph

The graph shows the data and nothing but the data.

- Tick marks are relevant, not arbitrarily equidistant;
- and non-data lines are grey, to make the data prominent.



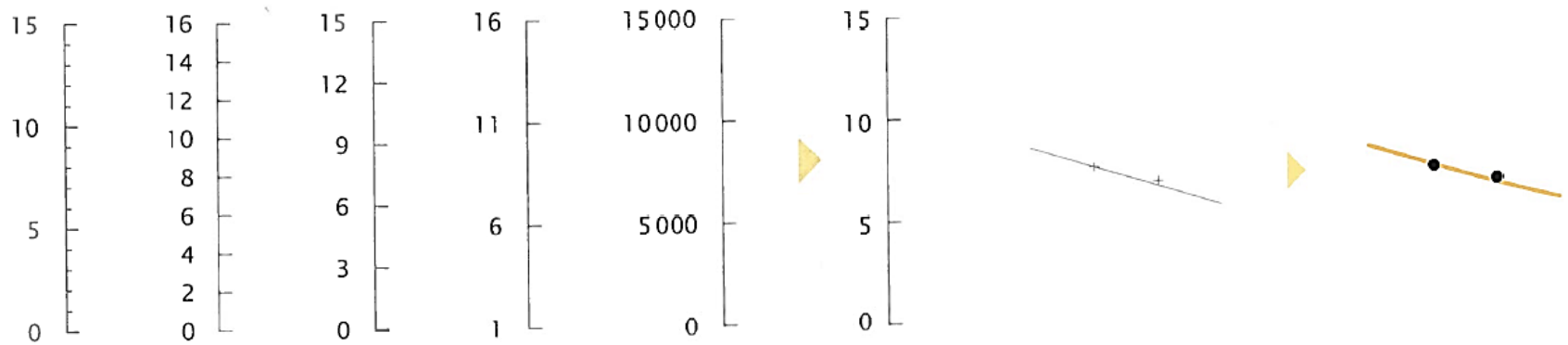
Effective visualisation
Effective Communication

Maximise the Signal-to-Noise ratio

Always avoid clutter. Keep your scales as simple as possible. Any scale is fully defined with just two tick marks. Any other tick mark should indicate a point of interest.

After reducing or eliminating the noise in the display, increase the signal by making the data more prominent.

Right click on legend -> Format Chart -> Axis Options…



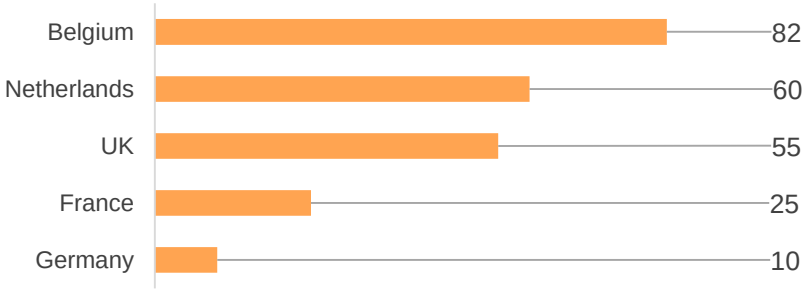
Effective visualisation

Choosing the correct graph

Comparing data

A straightforward way to compare numerical data is to represent them by lines or bars of proportional length aligned at one end. To respect the proportion, bars must start from zero.

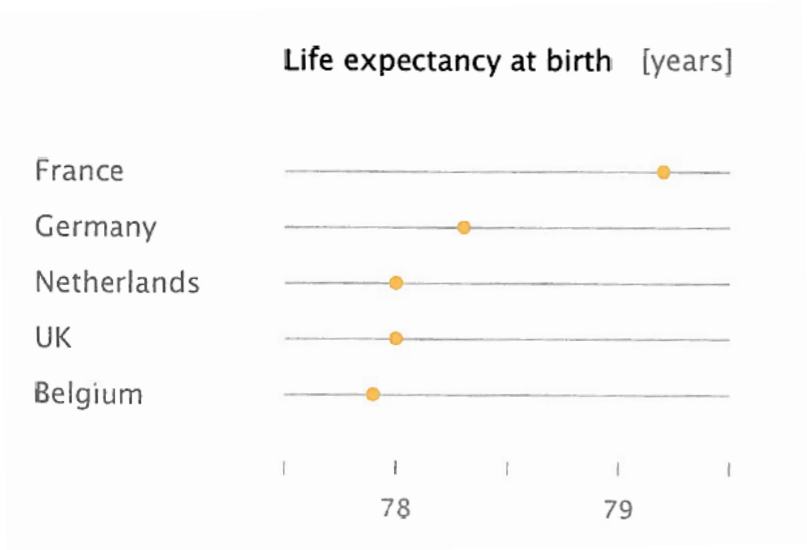
Select Data -> Bar Chart + Right click -> Format Chart



Effective visualisation
Choosing the correct graph

Comparing data

Close data values, poorly resolved by a length representation, are best encoded as positions along a scale , marked by dots. These do not need to run from zero to be meaningful

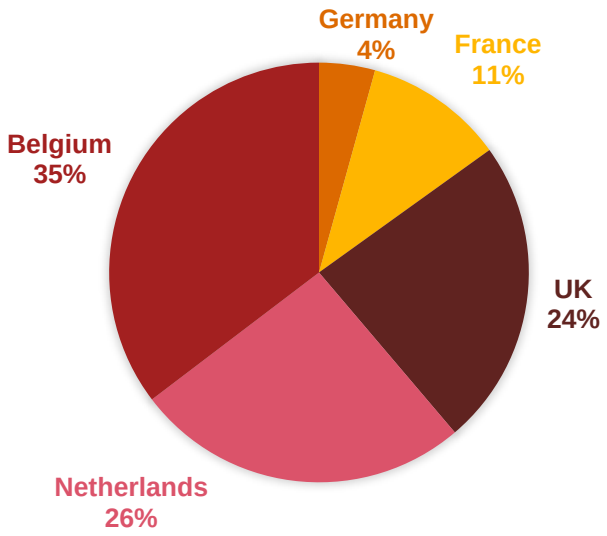


Effective visualisation
Choosing the correct graph

Avoid pie charts

Pie charts, a common way to represent fractions, are intuitive but are not very accurate, they fail to reveal small differences. They are best replaced by bar or dot charts.

Select Data -> Pie Chart…



Effective visualisation

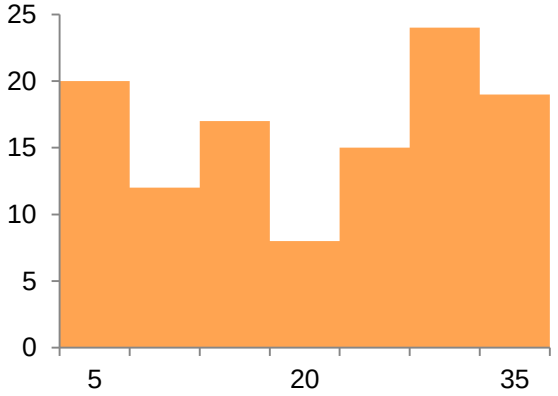
Choosing the correct graph

Displaying distribution

Showing the entire dataset as points along a scale is probably the most accurate way to convey its distribution. The resulting display is simple and truthful to individual data. For large datasets, however, it quickly becomes impractical.

Histograms reduce the dataset somewhat by grouping data n equivalent intervals. For an easy and intuitive interpretation of the fraction of total data in each interval, the bars must touch.

Select Data -> X Y (Scatter) + Format Chart & Column Chart + Format Chart

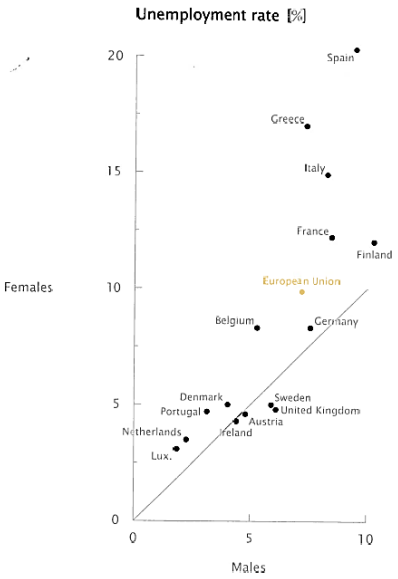


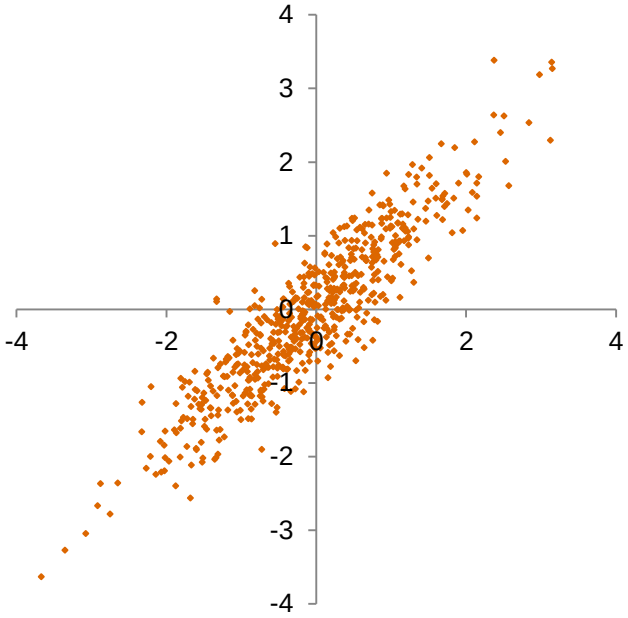
Effective visualisation
Choosing the correct graph

Revealing correlations

Correlation between two or more variables, especially when the variables are not sequenced in time, is revealed clearly by a scatter plot.

Select Data -> X Y (Scatter)…





Effective visualisation

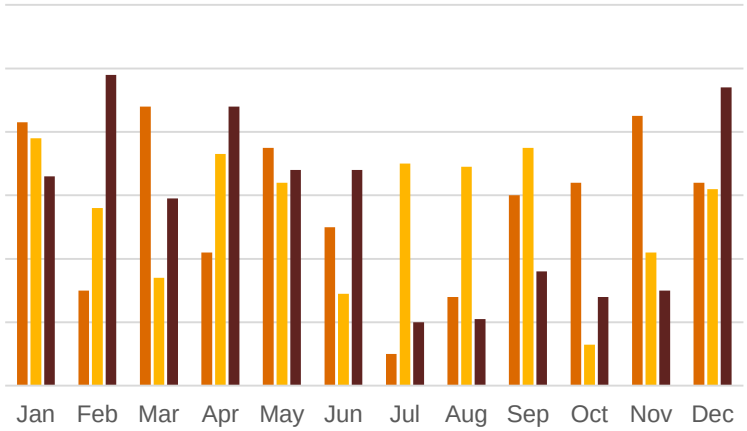
Choosing the correct graph

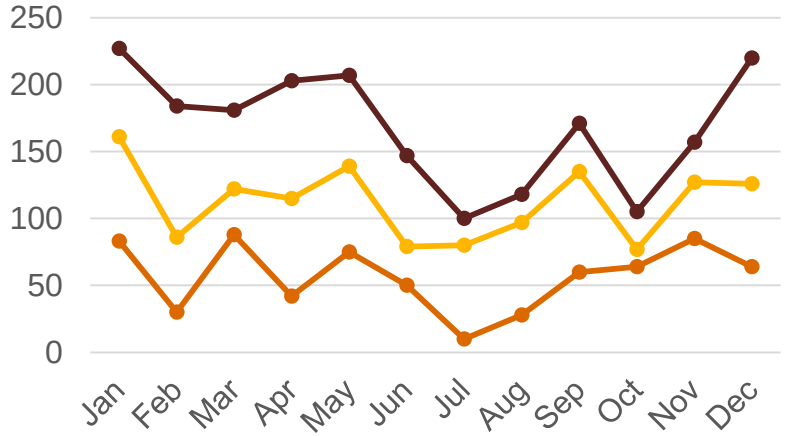
Displaying evolution

Bar charts although great for displaying univariate data, are a suboptimal display for multivariate data.

Switching from bars to dots connected by lines is the best way to display relation between two related variables, such as in evolutions through time.

Select Data -> Column Chart & Line Chart with Markers…





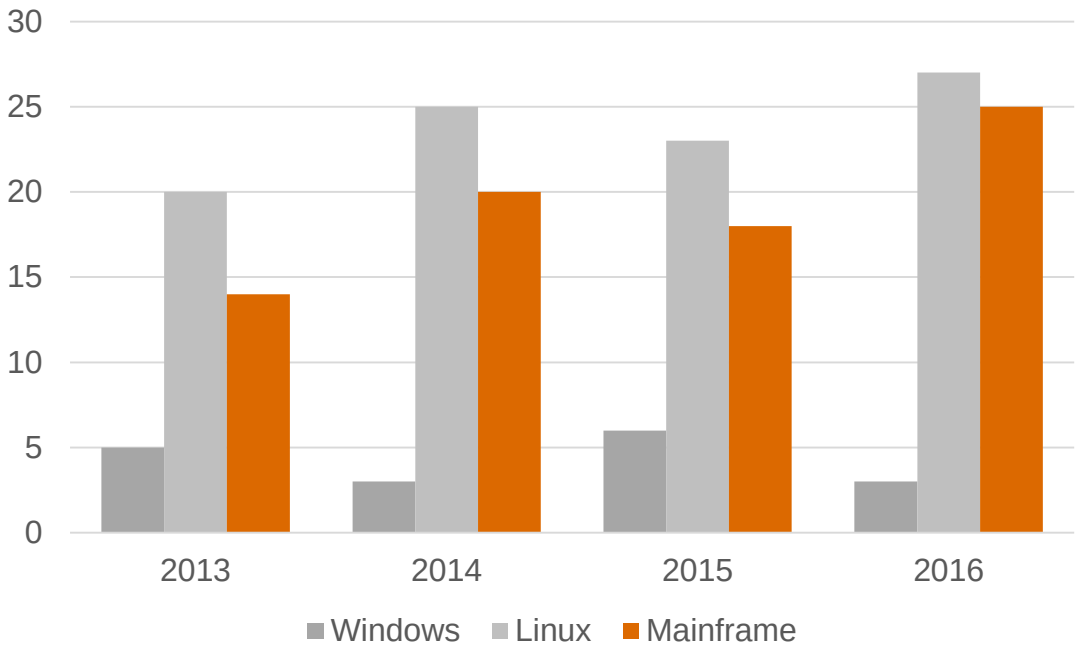
Effective visualisation

Choosing the correct graph

Displaying Evolution

- Column charts for univariate data

Select Data -> Column Chart



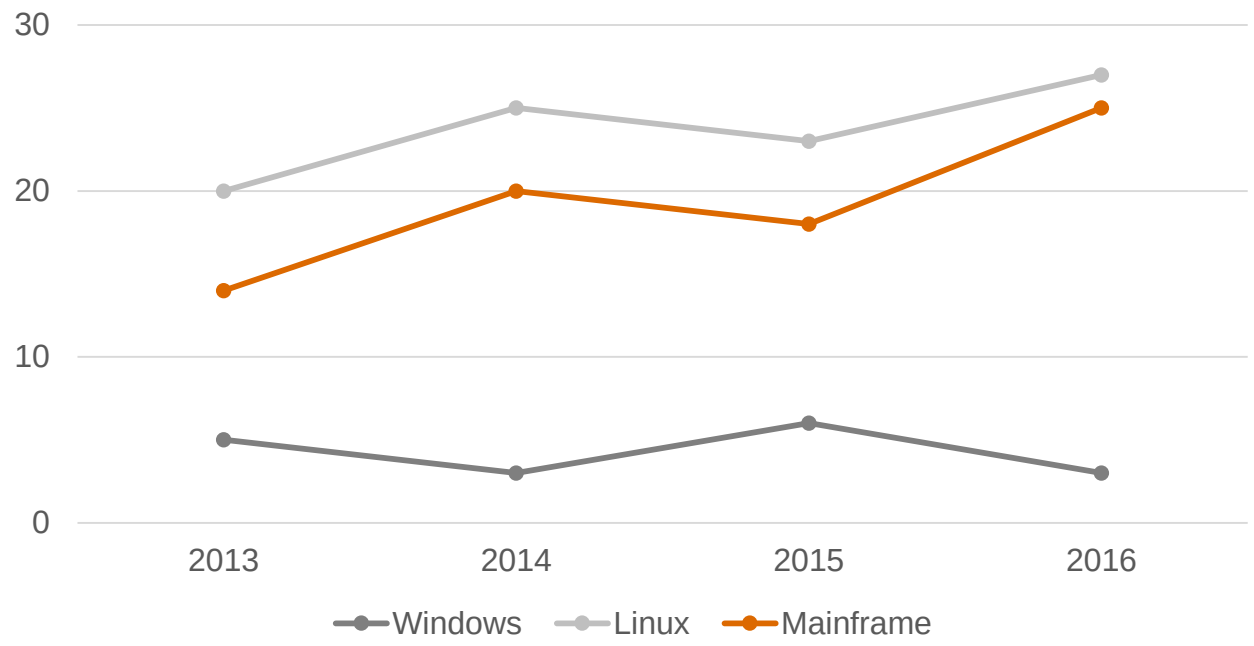
Effective visualisation

Choosing the correct graph

Displaying Evolution

- Line charts with Markers to indicate an evolution

Select Data -> Column Chart & Line Chart with Markers…



Introduction to Data Visualisation

- Pivot tables and other visualisation features

Visualising in MS Excel

Pivot tables in Excel

Massive datasets

An often overlooked form of visualisation is tables. **Yes tables!**

When dealing with massive datasets the so called, **Pivot Tables,** can seriously assist in aggregating and filtering the dataset in appropriate ways to facilitate understanding of what is hidden within.

Furthermore the table itself can be used to display the data in different ways, or to create **interactive charts** relative to the table.

Visualising in MS Excel

Obtaining the data

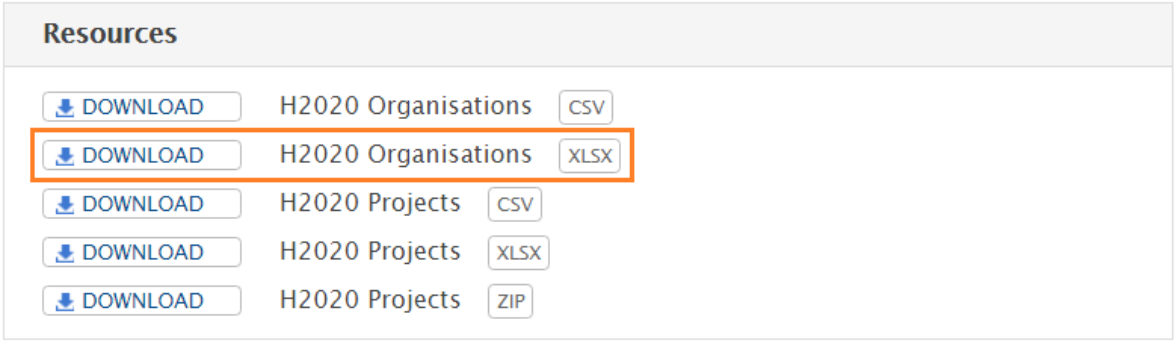
European Union Open Data Portal

H2020 funding data from the EU ODP

From the EU ODP we will obtain data containing the organisations funded by the European Union under the Horizon 2020 framework programme for research and innovation "H2020" from 2014 to 2020. The data includes institutions, countries, projects and amounts among other things.

Head to this link and download the file indicated below:

https://data.europa.eu/euodp/data/dataset/cordisH2020projects

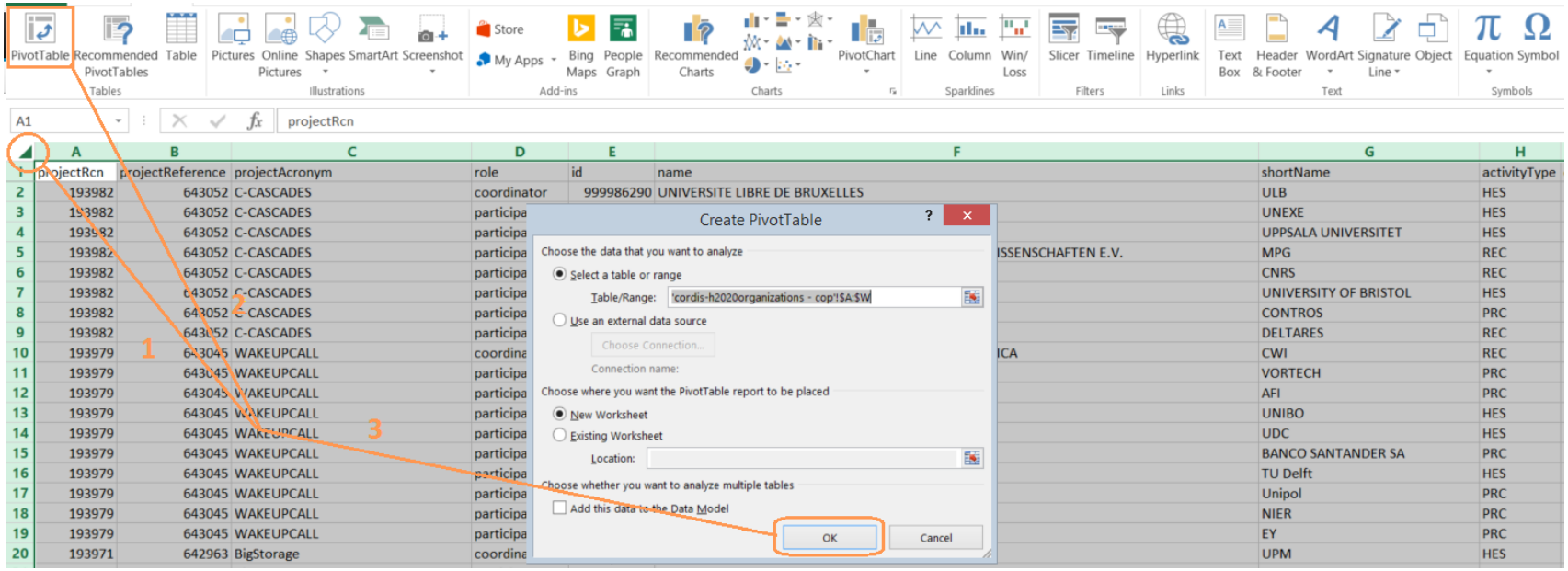


Visualising in MS Excel

Pivot tables in Excel

Creating the pivot table

1. Select the whole dataset by clicking on the top-left corner
2. Click “Pivot Table”
3. Leave default options and click “OK”



	A	B	C	D	E	F	G	H
1	projectRcn	projectReference	projectAcronym	role	id	name	shortName	activityType
2	193982	643052	C-CASCADES	coordinator	999986290	UNIVERSITE LIBRE DE BRUXELLES	ULB	HES
3	193982	643052	C-CASCADES	participa			UNEXE	HES
4	193982	643052	C-CASCADES	participa			UPPSALA UNIVERSITET	HES
5	193982	643052	C-CASCADES	participa		ISSENSCHAFTEN E.V.	MPG	REC
6	193982	643052	C-CASCADES	participa			CNRS	REC
7	193982	643052	C-CASCADES	participa			UNIVERSITY OF BRISTOL	HES
8	193982	643052	C-CASCADES	participa			CONTROS	PRC
9	193982	643052	C-CASCADES	participa			DELTARES	REC
10	193979	643045	WAKEUPCALL	coordina		ICA	CWI	REC
11	193979	643045	WAKEUPCALL	participa			VORTECH	PRC
12	193979	643045	WAKEUPCALL	participa			AFI	PRC
13	193979	643045	WAKEUPCALL	participa			UNIBO	HES
14	193979	643045	WAKEUPCALL	participa			UDC	HES
15	193979	643045	WAKEUPCALL	participa			BANCO SANTANDER SA	PRC
16	193979	643045	WAKEUPCALL	participa			TU Delft	HES
17	193979	643045	WAKEUPCALL	participa			Unipol	PRC
18	193979	643045	WAKEUPCALL	participa			NIER	PRC
19	193979	643045	WAKEUPCALL	participa			EY	PRC
20	193971	642963	BigStorage	coordina			UPM	HES

Visualising in MS Excel

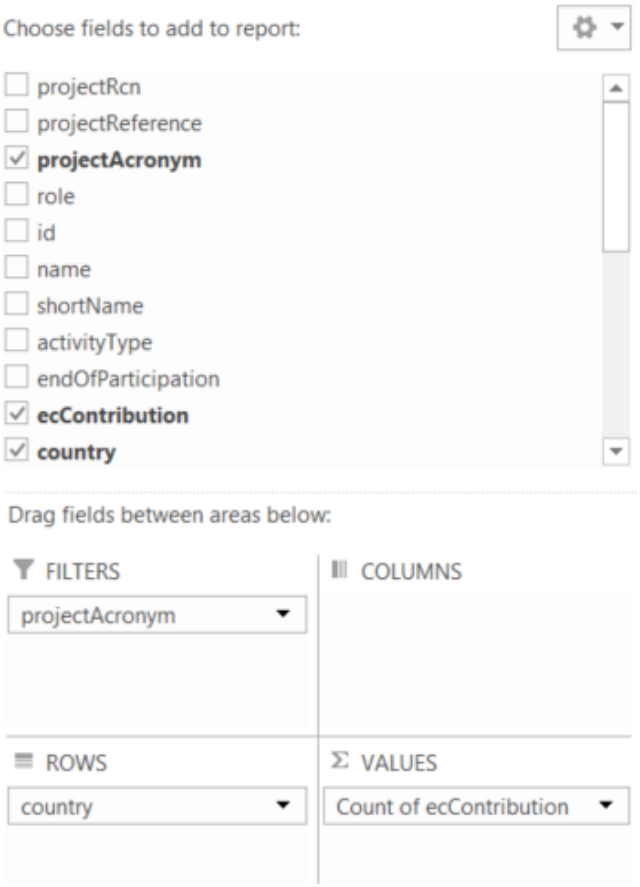
Pivot tables in Excel

Creating the pivot table

The Pivot Table field list appears. Here we select and drag the fields we want included into the different parts of the table.

Go ahead and select:

1. **"country"** for the Row
2. **"Count of ecContribution"** for Values
3. and **"projectAcronym"** for Filter



Visualising in MS Excel

Pivot tables in Excel

Refining the pivot table

The resulting table should look like this. The problem here is that the table is displaying the "Count" of contributions for each country and not the total amount.

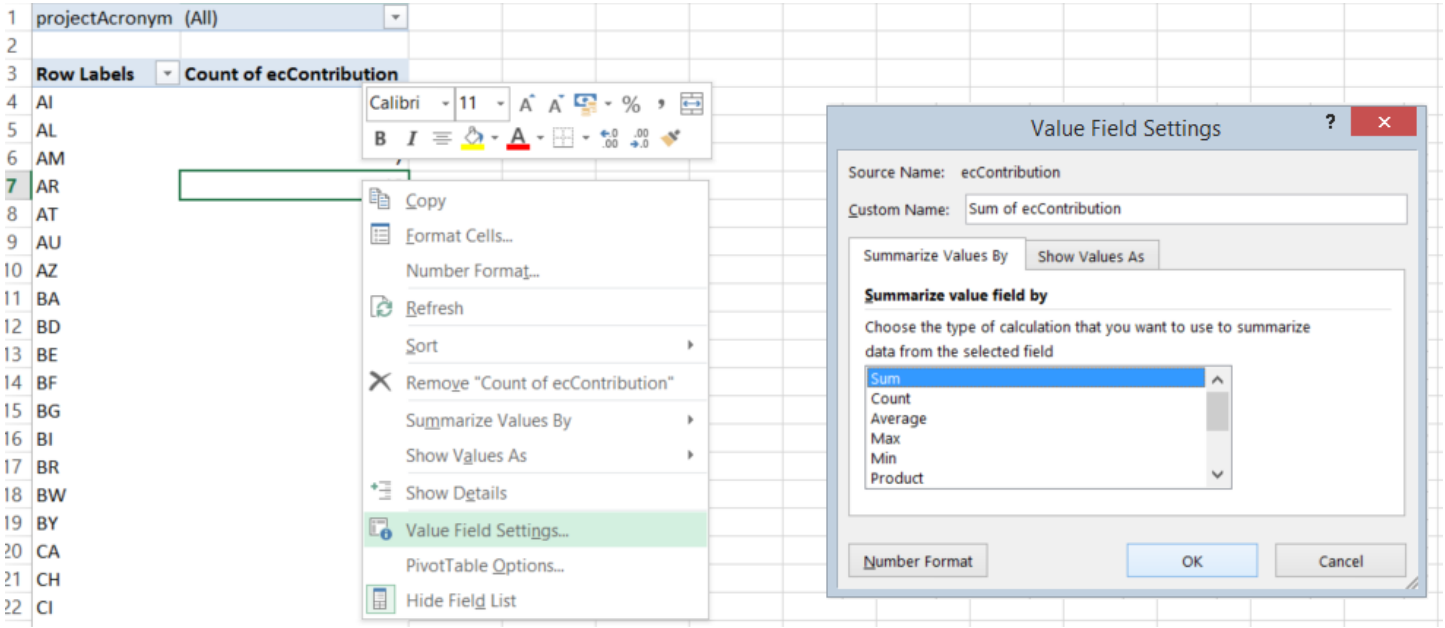
	A	B
1	projectAcronym	(All)
2		
3	**Row Labels**	**Count of ecContribution**
4	AI	3
5	AL	8
6	AM	7
7	AR	12
8	AT	1044
9	AU	32
10	AZ	3
11	BA	18
12	BD	2
13	BE	1580
14	BF	4
15	BG	199
16	BI	1
17	BR	28
18	BW	2
19	BY	10

Visualising in MS Excel

Pivot tables in Excel

Refining the pivot table

Right click on any of the cells under **“Count of ecContribution”** and then click **“Value Field Settings”** from the drop down menu. In the window that appears select **“Sum”** and click **“ok”**.

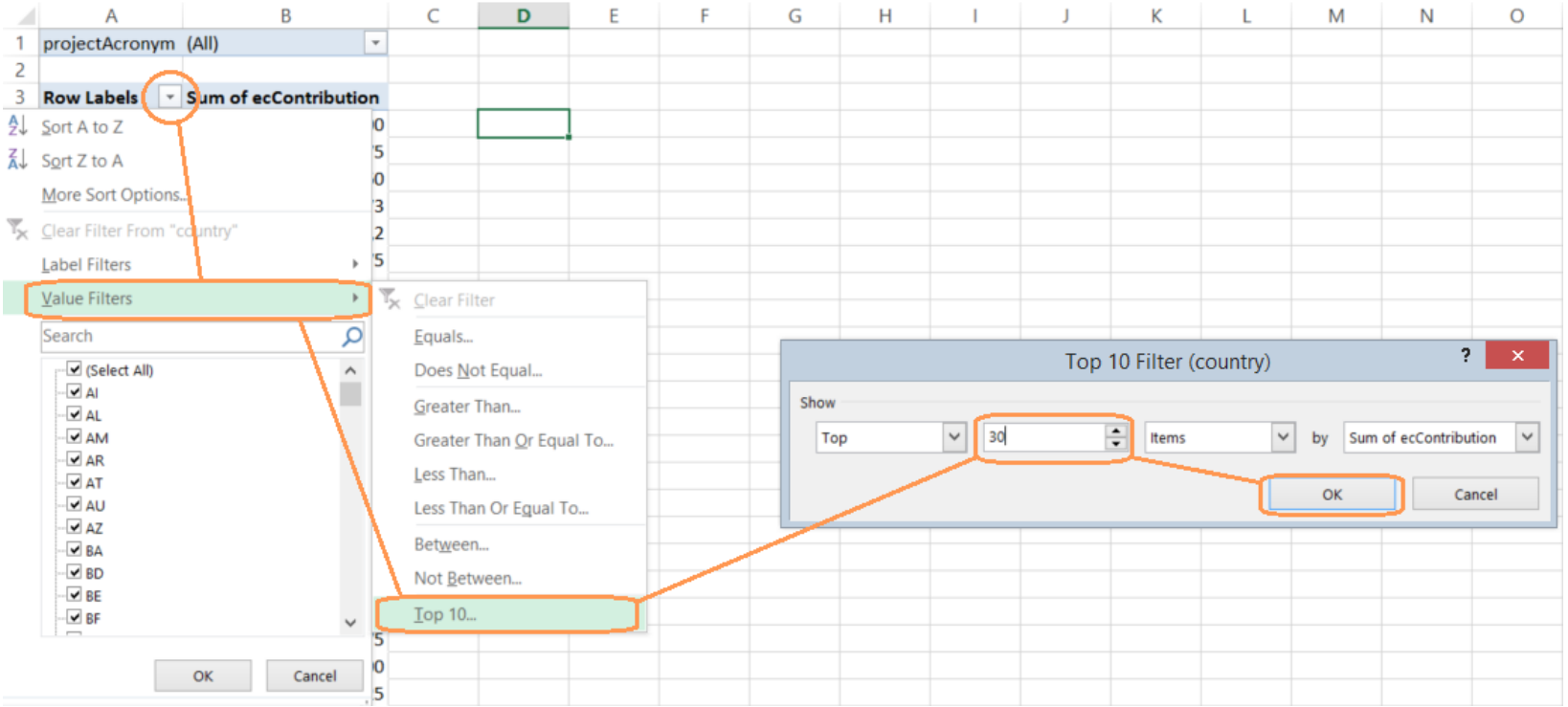


Visualising in MS Excel

Pivot tables in Excel

Refining the pivot table

Because the countries displayed are too many, lets limit our table to the 30 most funded countries. Click on the arrow next to **“Row Labels”**, from the dropdown select **“Value Filters”** and then click on **“Top 10”**. In the appearing menu select 30 and click **“OK”**.

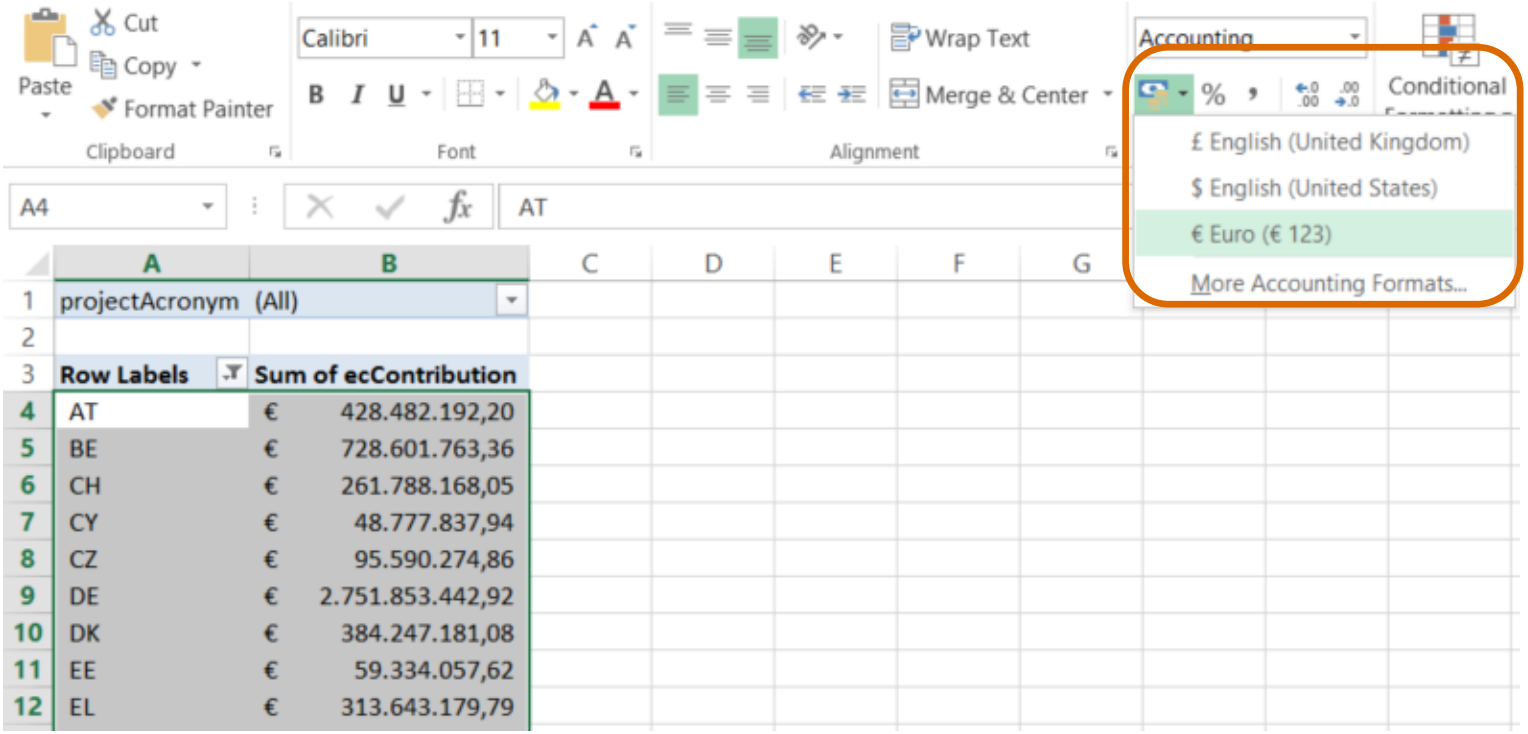


Visualising in MS Excel

Pivot tables in Excel

Refining the pivot table

That's more like it! Now select all the data and format it as "Euro" to align it and enhance readability.



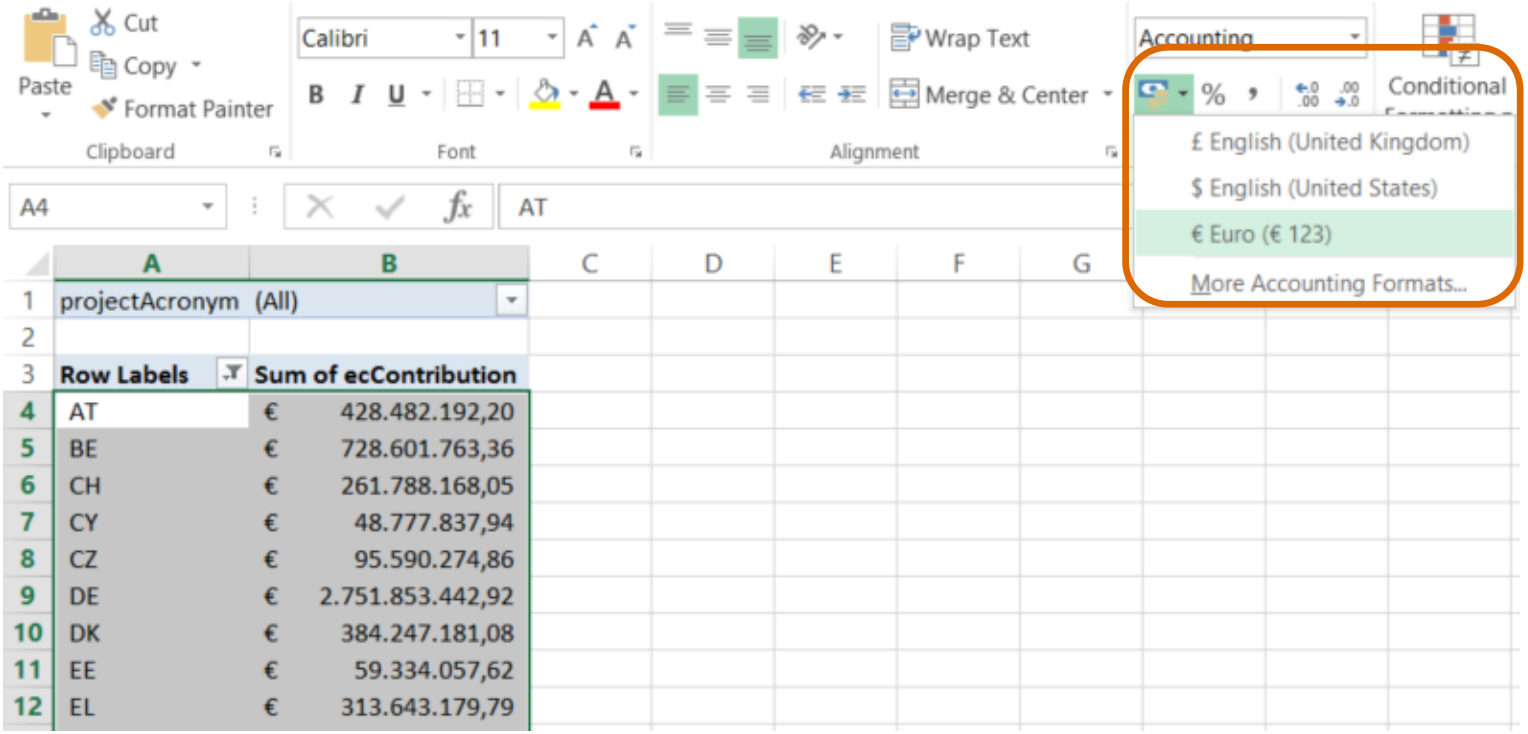
	A	B	C	D	E	F	G
1	projectAcronym	(All)					
2							
3	**Row Labels**	**Sum of ecContribution**					
4	AT	€ 428.482.192,20					
5	BE	€ 728.601.763,36					
6	CH	€ 261.788.168,05					
7	CY	€ 48.777.837,94					
8	CZ	€ 95.590.274,86					
9	DE	€ 2.751.853.442,92					
10	DK	€ 384.247.181,08					
11	EE	€ 59.334.057,62					
12	EL	€ 313.643.179,79					

Visualising in MS Excel

Pivot tables in Excel

Refining the pivot table

That's more like it! Now select all the data and format it as "Euro" to align it and enhance readability.



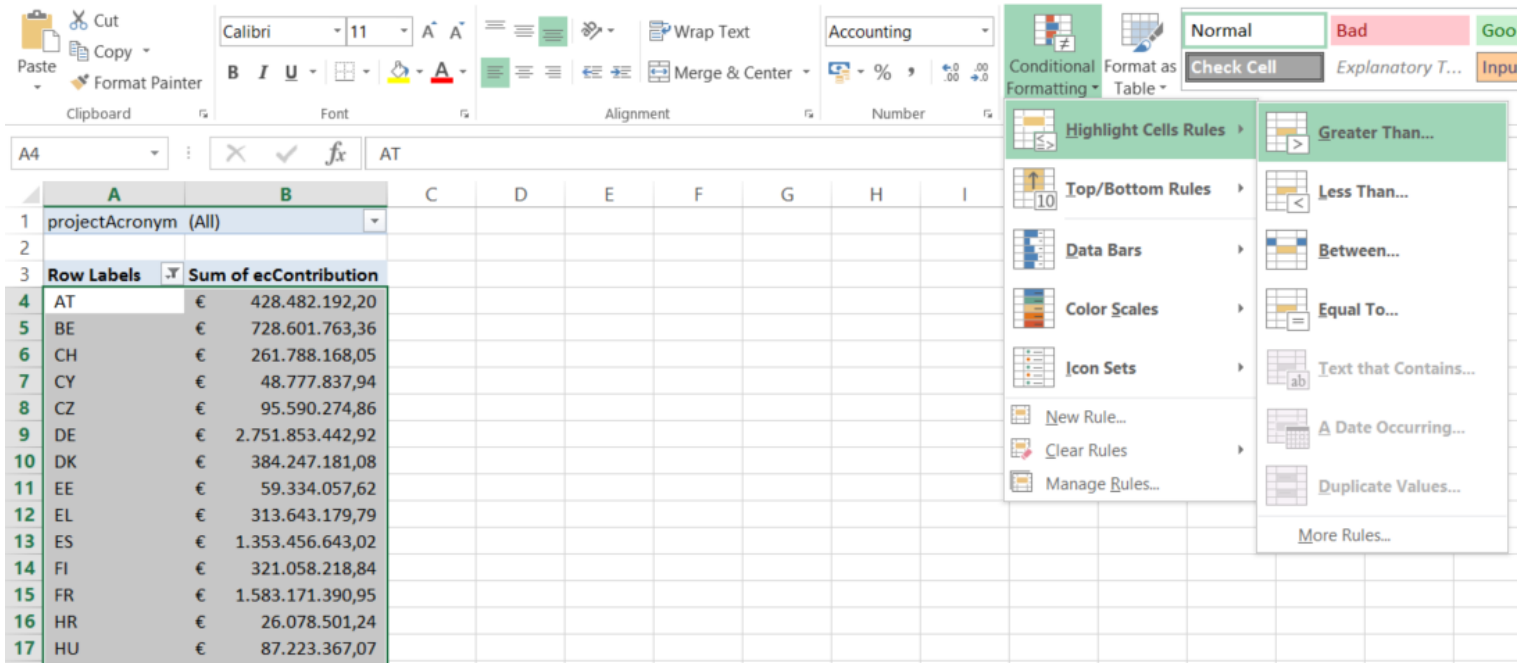
	A	B	C	D	E	F	G
1	projectAcronym	(All)					
2							
3	**Row Labels**	**Sum of ecContribution**					
4	AT	€ 428.482.192,20					
5	BE	€ 728.601.763,36					
6	CH	€ 261.788.168,05					
7	CY	€ 48.777.837,94					
8	CZ	€ 95.590.274,86					
9	DE	€ 2.751.853.442,92					
10	DK	€ 384.247.181,08					
11	EE	€ 59.334.057,62					
12	EL	€ 313.643.179,79					

Visualising in MS Excel

Pivot tables in Excel

Refining the pivot table

Now let's find out which of these 30 countries are funded more than the average! Select all the data, click on **"Conditional Formatting"** , then **"Highlight Cells Rules"** and click on **"Greater Than"**.



	A	B	C	D	E	F	G	H	I
1	projectAcronym	(All)							
2									
3	Row Labels	Sum of ecContribution							
4	AT	€ 428.482.192,20							
5	BE	€ 728.601.763,36							
6	CH	€ 261.788.168,05							
7	CY	€ 48.777.837,94							
8	CZ	€ 95.590.274,86							
9	DE	€ 2.751.853.442,92							
10	DK	€ 384.247.181,08							
11	EE	€ 59.334.057,62							
12	EL	€ 313.643.179,79							
13	ES	€ 1.353.456.643,02							
14	FI	€ 321.058.218,84							
15	FR	€ 1.583.171.390,95							
16	HR	€ 26.078.501,24							
17	HU	€ 87.223.367,07							

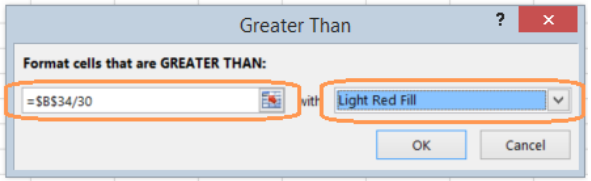
Visualising in MS Excel

Pivot tables in Excel

Refining the pivot table

In the menu that appears activate the first field and click on the cell containing the value of the **"Grand Total"** and divide by 30 by typing **"/30"** after it. In the **"with"** field select **"Light Red Fill"** and click **"OK"**.

3	Row Labels	Sum of ecContribution
4	AT	€ 428.482.192,20
5	BE	€ 728.601.763,36
6	CH	€ 261.788.168,05
7	CY	€ 48.777.837,94
8	CZ	€ 95.590.274,86
9	DE	€ 2.751.853.442,92
10	DK	€ 384.247.181,08
11	EE	€ 59.334.057,62
12	EL	€ 313.643.179,79
13	ES	€ 1.353.456.643,02
14	FI	€ 321.058.218,84
15	FR	€ 1.583.171.390,95
16	HR	€ 26.078.501,24
17	HU	€ 87.223.367,07
18	IE	€ 277.983.823,02
19	IL	€ 294.974.947,49
20	IS	€ 29.859.788,28
21	IT	€ 1.243.837.042,17
22	LU	€ 39.973.754,54
23	NL	€ 1.218.256.578,68
24	NO	€ 275.228.929,19
25	PL	€ 139.554.910,11
26	PT	€ 265.798.783,60
27	RO	€ 61.602.032,22
28	RS	€ 23.756.702,36
29	SE	€ 492.081.122,20
30	SI	€ 84.863.829,25
31	SK	€ 40.392.618,44
32	TR	€ 65.373.053,40
33	UK	€ 2.359.569.619,29
34	Grand Total	€ 15.356.413.753,18

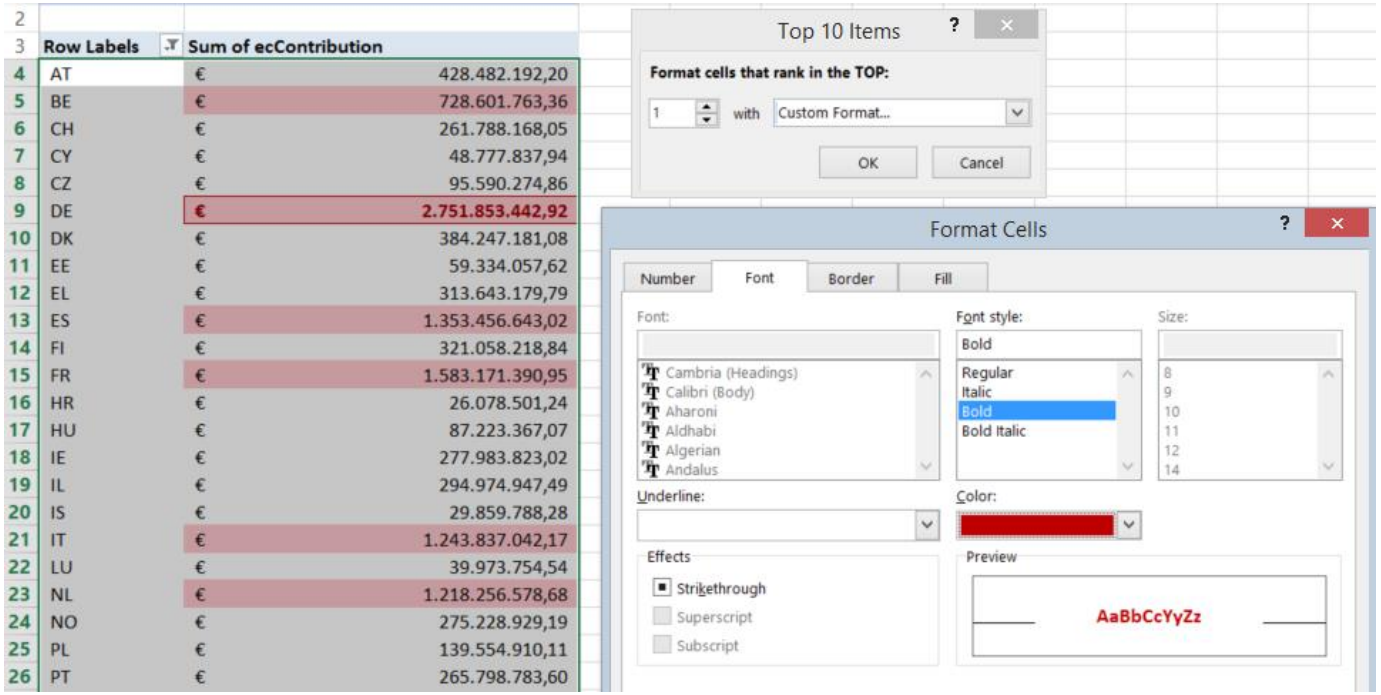


Visualising in MS Excel

Pivot tables in Excel

Refining the pivot table

Now do the same again but click on **“Conditional Formatting”** , then **“Top/Bottom Cells Rules”** and click on **“Top 10 Items”**. Customise the fields so that the Top 1 item has a red border and a bold red typeface.



Visualising in MS Excel

Pivot tables in Excel

Refining the pivot table

Hmm… It looks like a handful of countries are getting funded above average.

Lets take a closer look on the difference with the others!

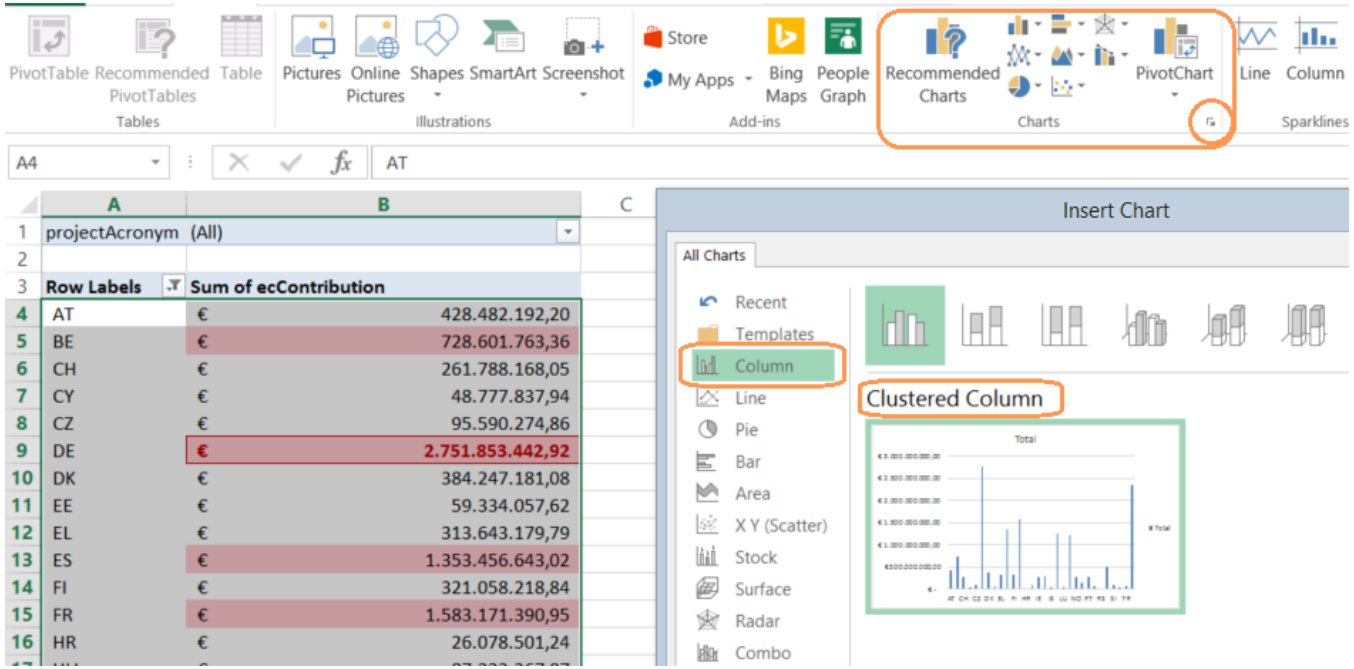
Row Labels	Sum of ecContribution
AT	€ 428.482.192,20
BE	€ 728.601.763,36
CH	€ 261.788.168,05
CY	€ 48.777.837,94
CZ	€ 95.590.274,86
DE	€ 2.751.853.442,92
DK	€ 384.247.181,08
EE	€ 59.334.057,62
EL	€ 313.643.179,79
ES	€ 1.353.456.643,02
FI	€ 321.058.218,84
FR	€ 1.583.171.390,95
HR	€ 26.078.501,24
HU	€ 87.223.367,07
IE	€ 277.983.823,02
IL	€ 294.974.947,49
IS	€ 29.859.788,28
IT	€ 1.243.837.042,17
LU	€ 39.973.754,54
NL	€ 1.218.256.578,68
NO	€ 275.228.929,19
PL	€ 139.554.910,11
PT	€ 265.798.783,60
RO	€ 61.602.032,22
RS	€ 23.756.702,36
SE	€ 492.081.122,20
SI	€ 84.863.829,25
SK	€ 40.392.618,44
TR	€ 65.373.053,40
UK	€ 2.359.569.619,29
Grand Total	**€ 15.356.413.753,18**

Visualising in MS Excel

Pivot tables in Excel

Making a chart

Let's make a chart to investigate. Click on the **"See all charts"** button at the bottom right of **"Charts"** tab. Select **"Column"** and from there **"Clustered Column".**

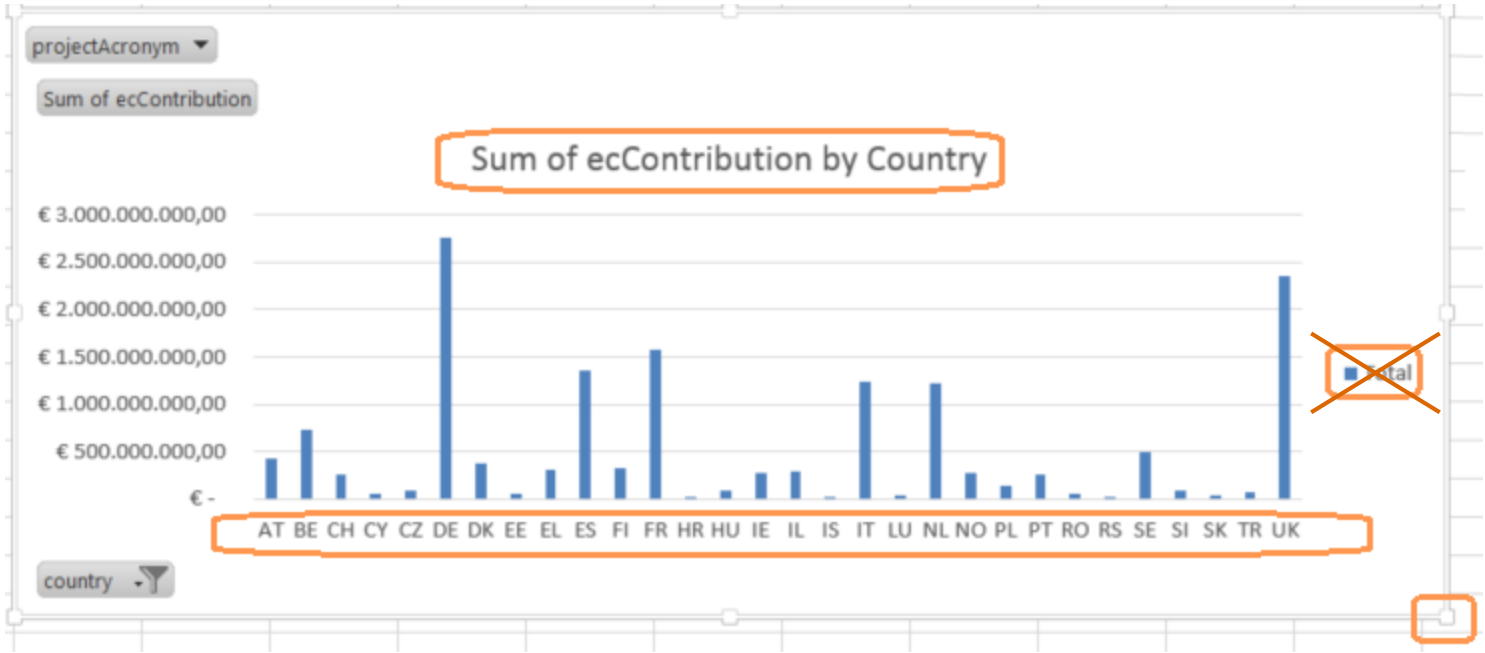


Visualising in MS Excel

Pivot tables in Excel

Making a chart

Ok nice. Now lets give the graph a meaningful label, let's widen it so the names of all the countries fit in the horizontal axis and let's remove the legend as all it does is clutter the graph.

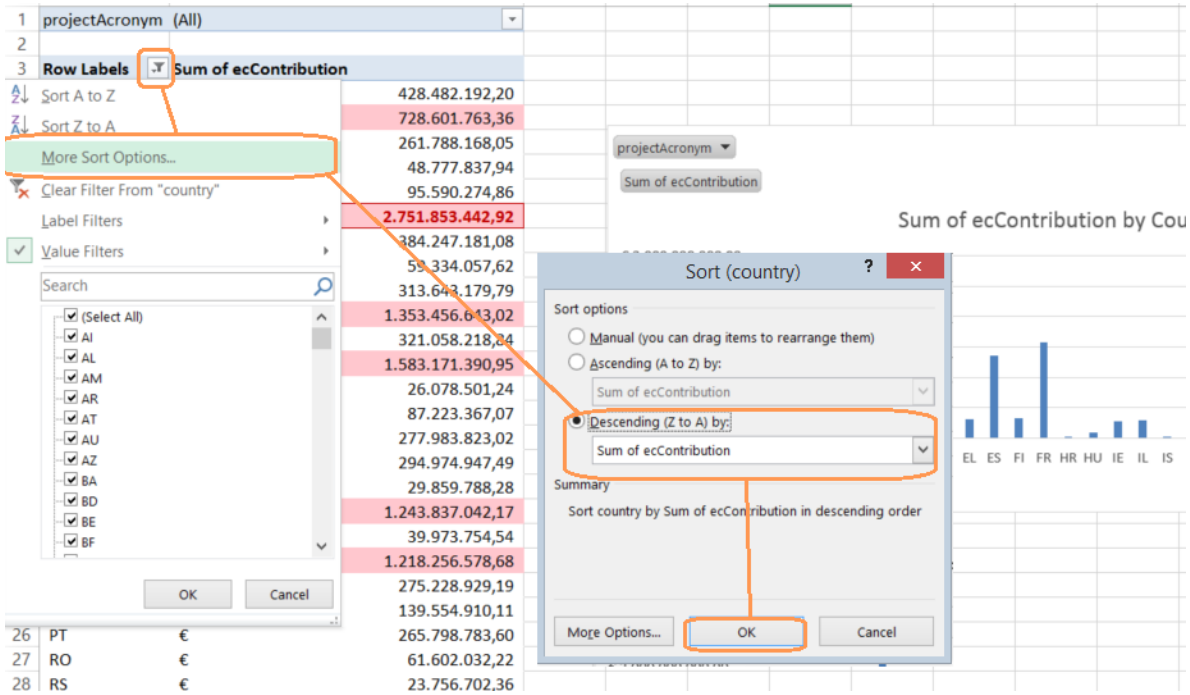


Visualising in MS Excel

Pivot tables in Excel

Making a chart

Now let's sort our chart in descending order to get clearer picture. Click on the filter symbol next to **"Row Labels"** in the pivot table, select **"More Sort Options"**, configure the field to sort descending by **"Sum of ecContribution"** and click **"OK"**.



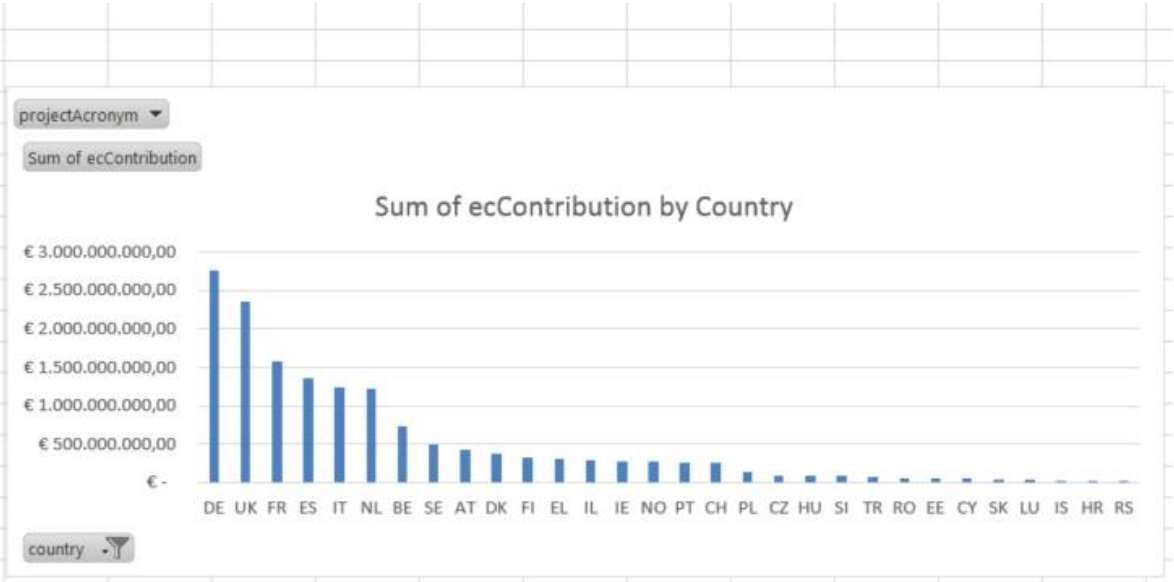
Visualising in MS Excel

Pivot tables in Excel

Making a chart

Whoah! By sorting the pivot table our chart also got sorted. When a chart is based on a pivot table any changes done to the table are transferred to the chart as well.

Row Labels	Sum of ecContribution
DE	€ 2.751.853.442,92
UK	€ 2.359.569.619,29
FR	€ 1.583.171.390,95
ES	€ 1.353.456.643,02
IT	€ 1.243.837.042,17
NL	€ 1.218.256.578,68
BE	€ 728.601.763,36
SE	€ 492.081.122,20
AT	€ 428.482.192,20
DK	€ 384.247.181,08
FI	€ 321.058.218,84
EL	€ 313.643.179,79
IL	€ 294.974.947,49
IE	€ 277.983.823,02
NO	€ 275.228.929,19
PT	€ 265.798.783,60
CH	€ 261.788.168,05
PL	€ 139.554.910,11

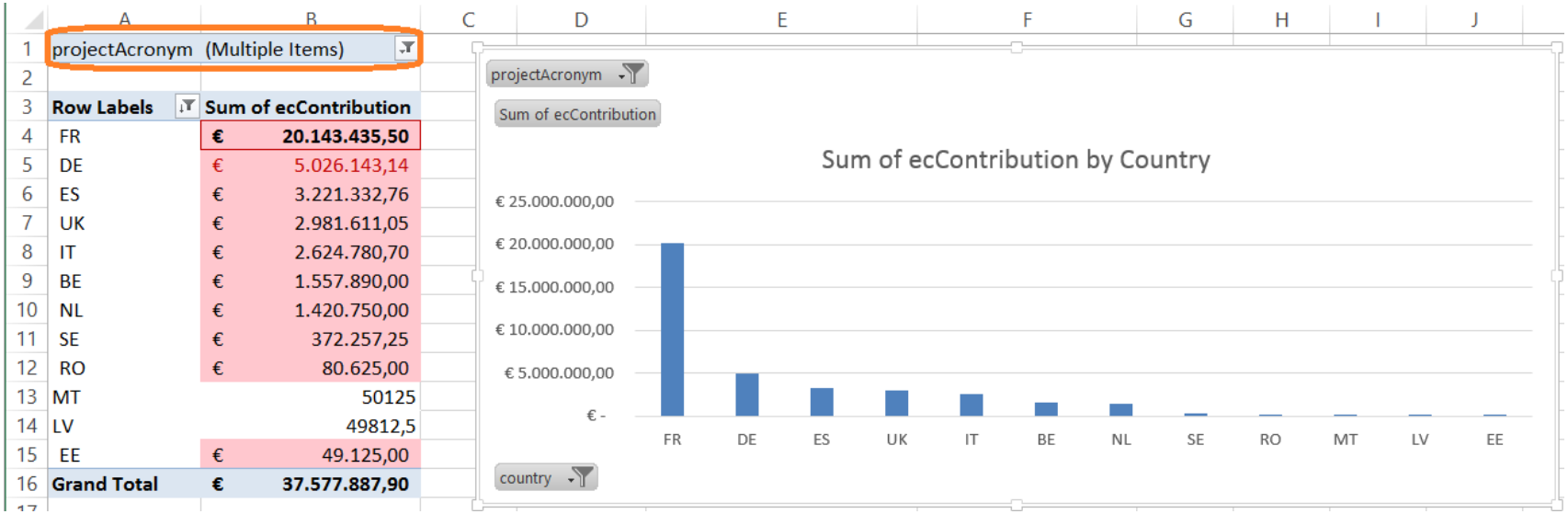


Visualising in MS Excel

Pivot tables in Excel

Making a chart

For example, try using the filter we have added. Select a few **“projectAcronyms”** and see how the chart updates to reflect the funding among the countries, for those projects.



projectAcronym	(Multiple Items)
Row Labels	Sum of ecContribution
FR	€ 20.143.435,50
DE	€ 5.026.143,14
ES	€ 3.221.332,76
UK	€ 2.981.611,05
IT	€ 2.624.780,70
BE	€ 1.557.890,00
NL	€ 1.420.750,00
SE	€ 372.257,25
RO	€ 80.625,00
MT	50125
LV	49812,5
EE	€ 49.125,00
Grand Total	€ 37.577.887,90

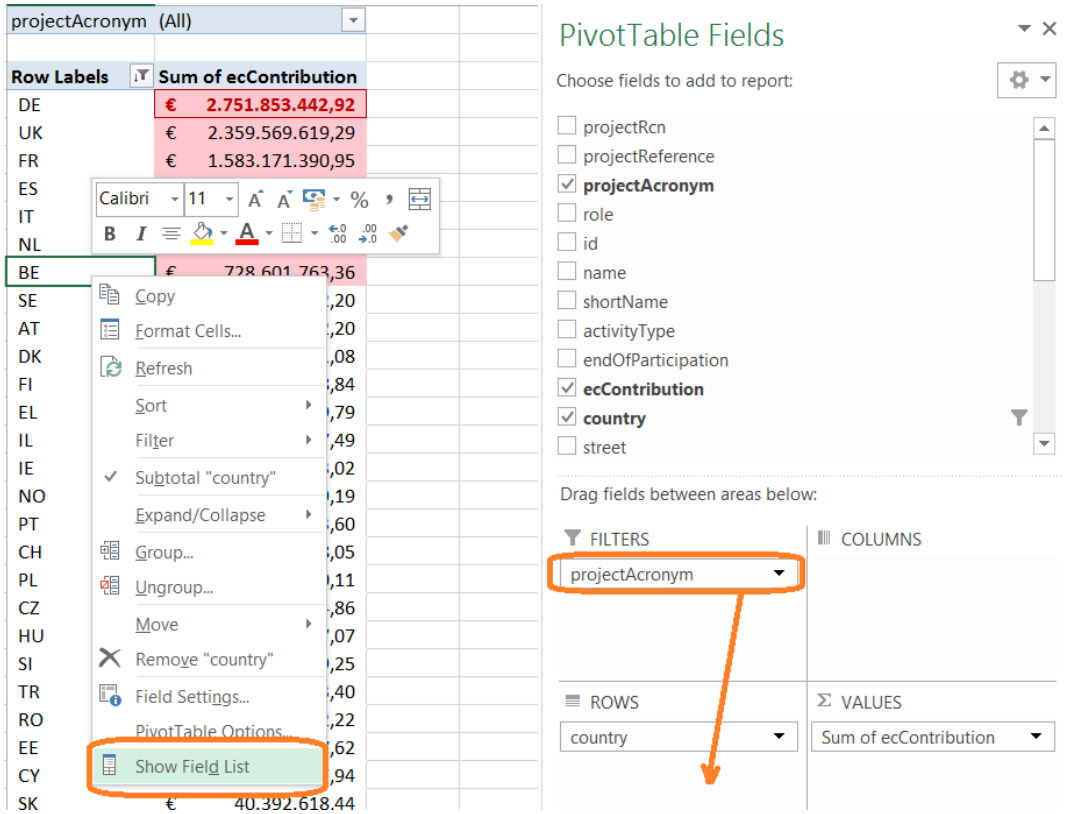
Visualising in MS Excel

Pivot tables in Excel

Refining the pivot table

Now lets say we want to see the Top 3 projects for which each country gets the most funding. Any ideas how we can do this?

Right click in the table and click **"Show Field List"**. From the menu drag **"projectAcronym"** from **"Filter"** to **"Rows"**.



Visualising in MS Excel

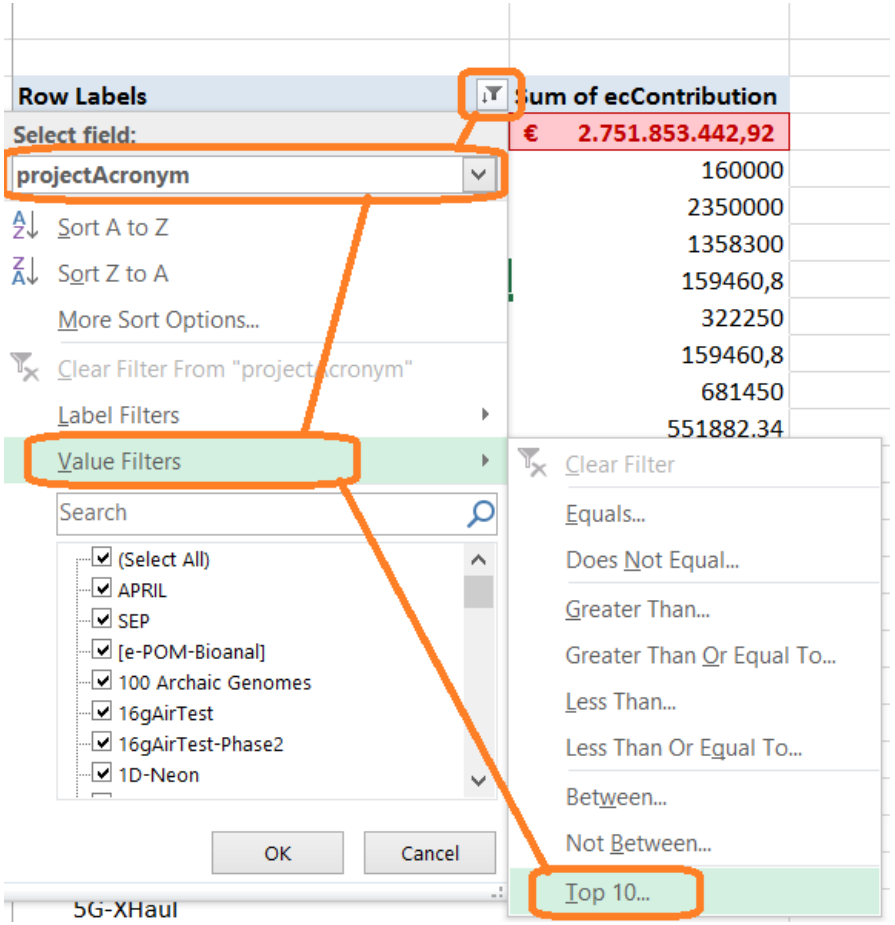
Pivot tables in Excel

Refining the pivot table

Now our pivot table is populated with ALL the projects for the countries displayed.

Click on the filter icon next to **“Row Labels”** , go to **“Value Filters”** and select **“Top 10”**. From the menu that appears select the top 3, like we did before.

Make sure you opened the **“Value Filters”** menu for **“projectAcronym”** and not for **“Country”**

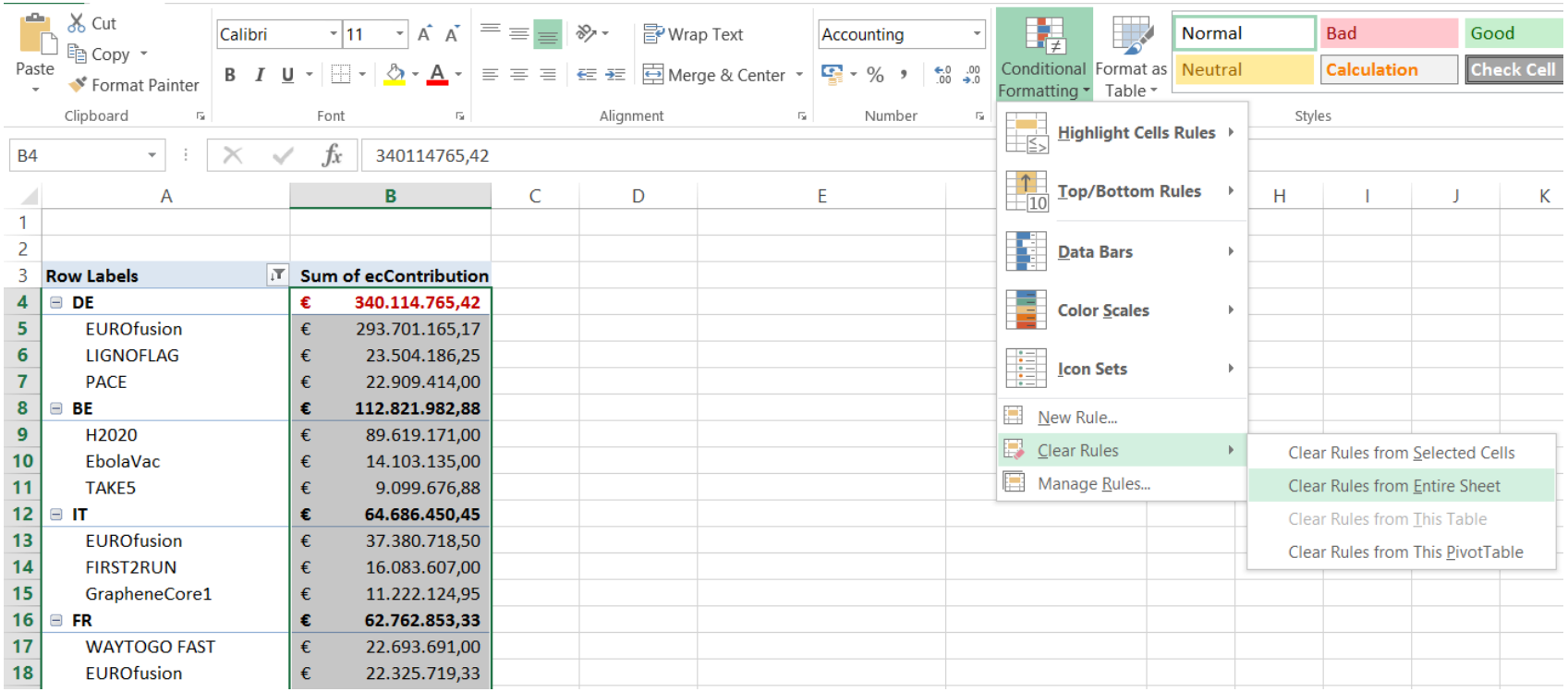


Visualising in MS Excel

Pivot tables in Excel

Refining the pivot table

Don’t forget to remove the **“Conditional formatting”** as it makes no sense anymore.



Visualising in MS Excel

Pivot tables in Excel

Refining the pivot table

Exercise:

The table below displays the “Countries” that take up 50% of the funding for each of the Top 10 funded “Projects”. Furthermore it highlights each row according to the “percentage” of each project’s funding each country receives.

Sum of ecContribution	Column Labels				
Row Labels	DE	UK	FR	ES	Grand Total
EUROfusion	€ 293.701.165,17	€ 27.479.945,00	€ 22.325.719,33	€ 9.242.254,01	€ 352.749.083,51
GrapheneCore1	€ 9.727.808,69	€ 15.941.815,87	€ 10.713.462,95	€ 10.865.078,75	€ 47.248.166,26
H2ME	€ 16.013.877,00	€ 3.969.456,00	€ 9.321.841,00		€ 29.305.174,00
WAYTOGO FAST	€ 5.535.377,10	€ 168.675,00	€ 22.693.691,00	€ 48.750,00	€ 28.446.493,10
PACE	€ 22.909.414,00	€ 870.964,51	€ 1.314.378,00	€ 247.872,96	€ 25.342.629,47
LIGNOFLAG	€ 23.504.186,25				€ 23.504.186,25
PROMOTION	€ 15.922.585,50	€ 3.946.696,25	€ 1.097.775,00	€ 970.095,00	€ 21.937.151,75
PERFORM	€ 4.493.095,00	€ 12.824.246,75	€ 865.725,00	€ 1.703.610,00	€ 19.886.676,75
CONCERT	€ 15.218.811,50	€ 450.812,00	€ 1.597.157,00	€ 2.067.879,50	€ 19.334.660,00
REPLICATE	€ 915.512,50	€ 8.265.645,14		€ 9.602.112,10	€ 18.783.269,74
Grand Total	**€ 407.941.832,71**	**€ 73.918.256,52**	**€ 69.929.749,28**	**€ 34.747.652,32**	**€ 586.537.490,83**

Visualising in MS Excel

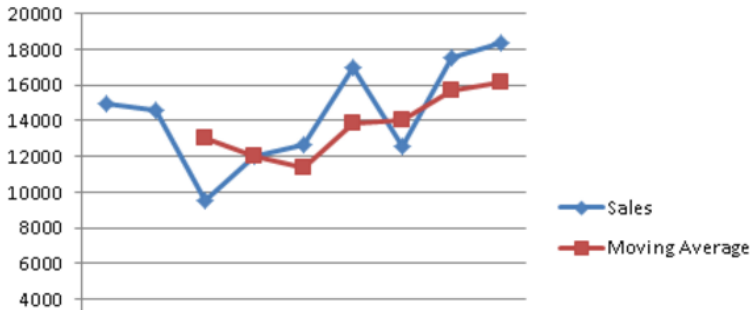
Statistical package and visualisation

Excel Data Analysis Toolpak

The Data Analysis toolpak allows the user to easily conduct descriptive and exploratory statistics jobs on their data. It also offers a number of relevant visualisations such as histograms, scatter plots and regression charts.

Examples:

Moving average



Correlation Matrix

		1	2	3	4	5	6
1	Population in thousands						
2	Number of people / sq. kilometer	-.02					
3	People living in cities (%)	-.18	.22				
4	Average female life expectancy	-.07	.13	.76			
5	Average male life expectancy	-.03	.15	.74	.98		
6	People who read (%)	-.06	.03	.65	.87	.81	

Visualising in MS Excel

Visualising on maps

Excel Power Map

Power Map is a very easy to use add-on for Excel that offers integration with Bing maps. It is usually pre-installed and only needs to be activated from the options.

Power map takes data, that includes an area variable (Countries, Cities, Regions, etc.), from a spreadsheet and automatically detects the geolocation details of the places and applies the data on a map.

Example:
https://www.youtube.com/watch?v=_NPpISageUU

Exporting your Visualisations

Remember that if you want to add your graph to another office document (Word, PowerPoint) you can simply copy paste it and you retain all formatting options in the new file.

Introduction to Data Visualisation

- Individual Projects

Visualising in MS Excel

Individual Projects

Project Instructions

For the next part we will be using the datasets you all brought to create individual visualisations.

- Each participant will have to find an interesting insight hidden in the data and visualise it accordingly to reveal it.
- We will be going around helping everyone with their project.
- In the end each participant will present their findings to the rest of us.
- Those of you who haven't brought your own dataset, can head to the EU ODP and search for one now.

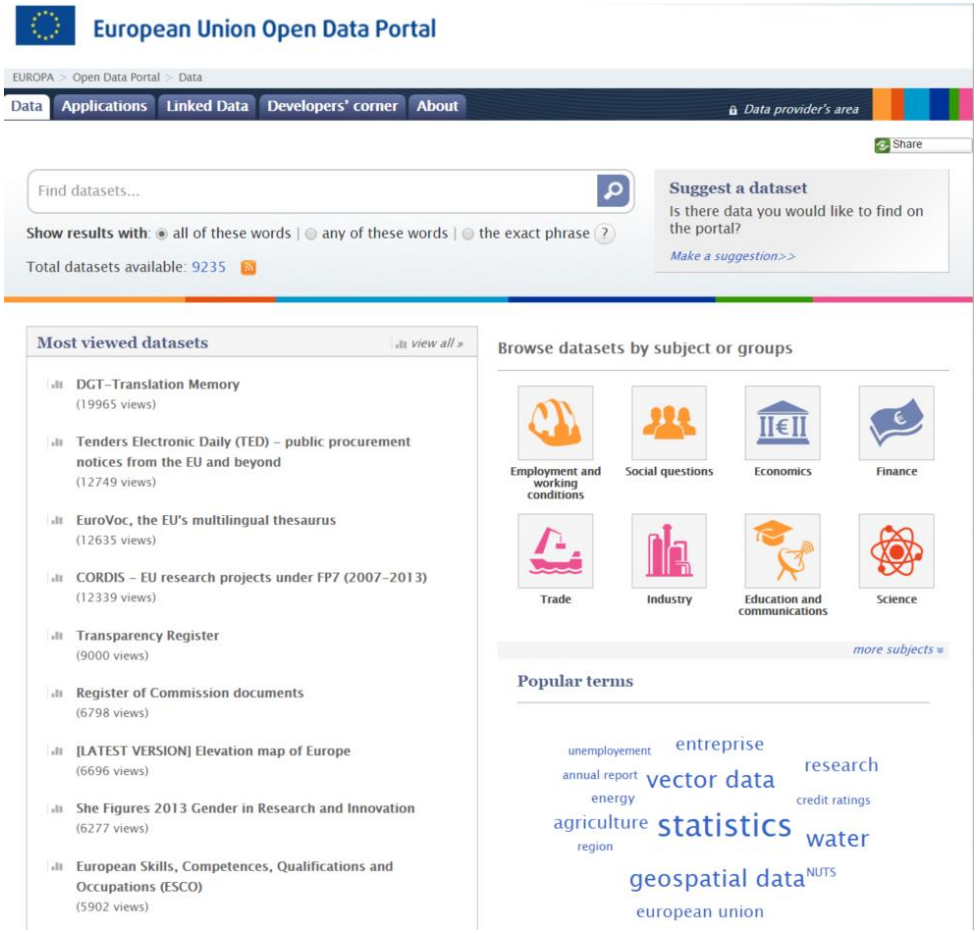
Visualising in MS Excel
Choosing Datasets

EU ODP

Head through the ODP and select a dataset.

https://data.europa.eu/euodp/en/data/

- Available as .xls or CSV
- Contains tabular data
- Contains numerical data
- Contains more than 2 variables
- Contains various granularity levels
- Contains breakdowns (e.g. country)
- Can show evolution (e.g. years)
- Can be combined with other data (e.g. employment vs. population)



Disclaimers

This presentation has been carefully compiled by PwC, but no representation is made or warranty given (either express or implied) as to the completeness or accuracy of the information it contains. PwC is not liable for the information in this presentation or any decision or consequence based on the use of it. PwC will not be liable for any damages arising from the use of the information contained in this presentation. The information contained in this presentation is of a general nature and is solely for guidance on matters of general interest. This presentation is not a substitute for professional advice on any particular matter. No reader should act on the basis of any matter contained in this publication without considering appropriate professional advice.