Excel and Power BI

BATH SPA UNIVERSITY

Data Analytics and Machine Learning

Lesson Objectives

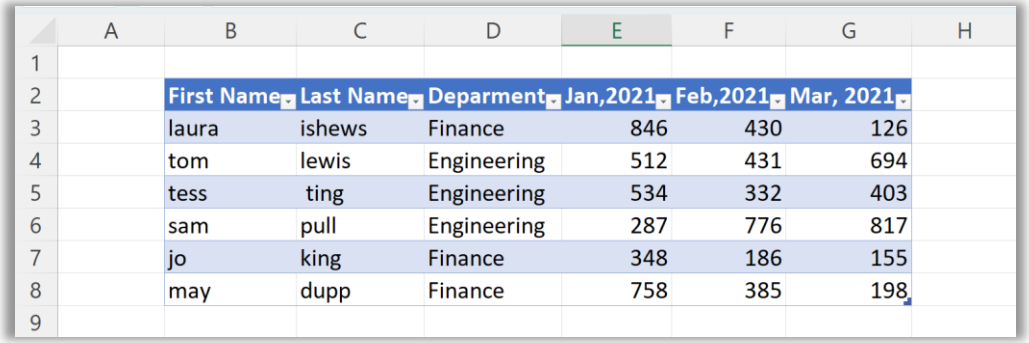
- Understand how to prepare data using Excel functions
- Explore the Power BI interface
- Import a range of data sources into Power BI
- Create visualisations in Power BI

Keyword	Description
Functions	A function in Excel is **a preset formula, that helps perform mathematical, statistical and logical operations**. Once you are familiar with the function you want to use, all you must do is enter an equal sign (=) in the cell, followed by the name of the function and the cell range it applies to.
Data	**Information**, especially facts or numbers, **collected to be examined** and considered and **used to help decision-making**, or information in an electronic form that can be stored and used by a computer: The data was collected by various researchers.
Syntax	Syntax is **the set of rules that define what the various combinations of symbols mean**. This tells the computer how to read the code. Syntax refers to a concept in writing code dealing with a very specific set of words and a very specific order to those words when we give the computer instructions.
Continuous	Continuous data is data that **can take any value**. Height, weight, temperature and length are all examples of continuous data. Some continuous data will change over time; the weight of a baby in its first year or the temperature in a room throughout the day.
Discrete	Discrete data is a count that involves integers **— only a limited number of values is possible**. This type of data cannot be subdivided into different parts. Discrete data includes discrete variables that are finite, numeric, countable, and non-negative integers.
Nesting	Nesting is **a common technique in computer programming, whereby one piece of code is placed inside another**. This can be used to create more complex code structures and can make code more readable and easier to understand.
Pivot Table	A **dynamic summary report** generated from a database
Pivot Chart	A **visualisation** option to display a Pivot Table in multiple **different chart types**.

BATH SPA UNIVERSITY

Excel in Data Analysis

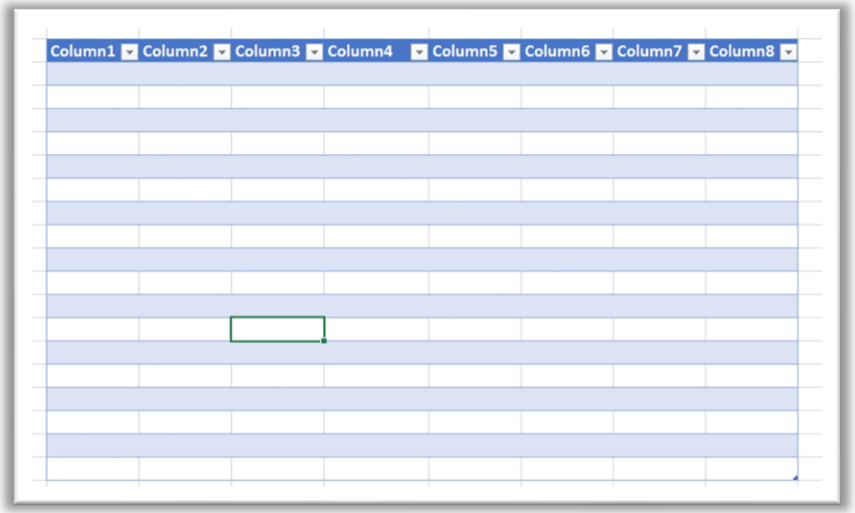
- Import data from various sources
- Cleaning the data
- Perform Calculations

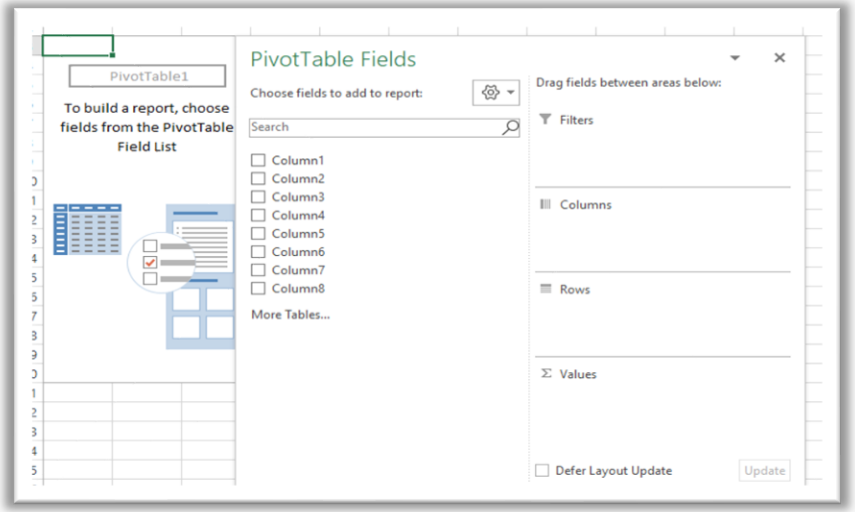


	A	B	C	D	E	F	G	H
1								
2		First Name	Last Name	Deparment	Jan,2021	Feb,2021	Mar, 2021	
3		laura	ishews	Finance	846	430	126	
4		tom	lewis	Engineering	512	431	694	
5		tess	ting	Engineering	534	332	403	
6		sam	pull	Engineering	287	776	817	
7		jo	king	Finance	348	186	155	
8		may	dupp	Finance	758	385	198	
9								

Excel Tools

- Data cleaning
- Tables
- Functions
- Pivot Tables





~~How to use Excel~~

When to use Excel

Why to use Excel

	A	B	C	D	E	F	G	H
1								
2		First Name	Last Name	Deparment	Jan,2021	Feb,2021	Mar, 2021	
3		laura	ishews	Finance	846	430	126	
4		tom	lewis	Engineering	512	431	694	
5		tess	ting	Engineering	534	332	403	
6		sam	pull	Engineering	287	776	817	
7		jo	king	Finance	348	186	155	
8		may	dupp	Finance	758	385	198	
9								

BATH SPA UNIVERSITY

~~How to use Excel~~

When to use Excel

Why to use Excel

- When the file breaks your code
- When the file fails to load or returns nonsense
- When you suspect the file will be poor quality
- When you're not sure what the data looks like
- When you're not sure the data contains the fields you need

	A	B	C	D	E	F	G	H
1								
2		First Name	Last Name	Deparment	Jan,2021	Feb,2021	Mar, 2021	
3		laura	ishews	Finance	846	430	126	
4		tom	lewis	Engineering	512	431	694	
5		tess	ting	Engineering	534	332	403	
6		sam	pull	Engineering	287	776	817	
7		jo	king	Finance	348	186	155	
8		may	dupp	Finance	758	385	198	
9								

BATH SPA UNIVERSITY

~~How to use Excel~~

When to use Excel

Why to use Excel

- It can open almost any tabular data format
- It's easy to edit, clean and validate data using Excel functions
- It has pivot tables (perfect for exploratory analysis)
- Any major data analysis and visualisation platform in use today will take Excel files as inputs

	A	B	C	D	E	F	G	H
1								
2		First Name	Last Name	Deparment	Jan,2021	Feb,2021	Mar, 2021	
3		laura	ishews	Finance	846	430	126	
4		tom	lewis	Engineering	512	431	694	
5		tess	ting	Engineering	534	332	403	
6		sam	pull	Engineering	287	776	817	
7		jo	king	Finance	348	186	155	
8		may	dupp	Finance	758	385	198	
9								

BATH SPA UNIVERSITY

Excel Functions

- Preset Formula
  - Mathematical
  - Statistical
  - Logical operations

- Function Structure
  - Always start with an ‘equals’ sign
    - =
  - Function name
    - =SUM
  - Cell range to apply to
    - =SUM(E3+E4)

- AVERAGE()
- AVERAGEIF()
- AVERAGEIFS()
- CONCATENATE()
- COUNT()
- COUNTA()
- COUNTBLANK()
- COUNTIF()
- EXACT()
- FILTER()
- IF()
- LEFT()
- LEN()
- LOWER()
- MAX()
- MEDIAN()
- MIN()
- MONTH()
- NOW()
- OR()
- PROPER()
- RIGHT()
- ROUND()
- SORT()
- SUBSTITUTE()
- SUM()
- SUMIF()
- TEXT()
- TODAY()
- TRIM()
- UNIQUE()
- UPPER()
- VALUE()
- VLOOKUP()
- WEEKDAY()
- YEAR()

Excel Functions

- Preset Formula
  - Mathematical
  - Statistical
  - Logical operations

- Function Structure
  - Always start with an ‘equals’ sign
    - =
  - Function name
    - =SUM
  - Cell range to apply to
    - =SUM(E3+E4)

- AVERAGE()
- AVERAGEIF()
- AVERAGEIFS()
- CONCATENATE()
- COUNT()
- COUNTA()
- COUNTBLANK()
- COUNTIF()
- EXACT()
- FILTER()
- IF()
- LEFT()
- LEN()
- LOWER()
- MAX()
- MEDIAN()
- MIN()
- MONTH()
- NOW()
- OR()
- PROPER()
- RIGHT()
- ROUND()
- SORT()
- SUBSTITUTE()
- SUM()
- SUMIF()
- TEXT()
- TODAY()
- TRIM()
- UNIQUE()
- UPPER()
- VALUE()
- VLOOKUP()
- WEEKDAY()
- YEAR()

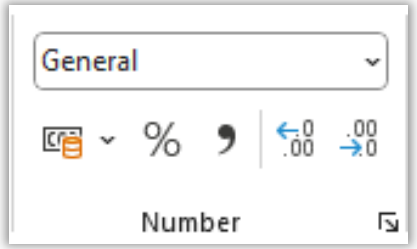
Excel Functions
Data Types

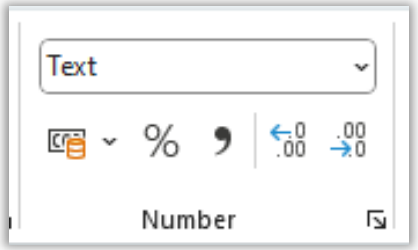
Revisiting Excel Data Types

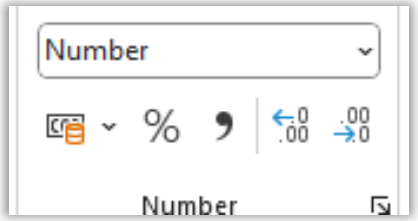
- Classification of data
- Allows the computer to determine how to process the data

- Programming languages have various data types
  - Integer, real, float, character, string, boolean, etc..

- Excel Data Types
  - Number
  - Text
  - Logical
  - Error

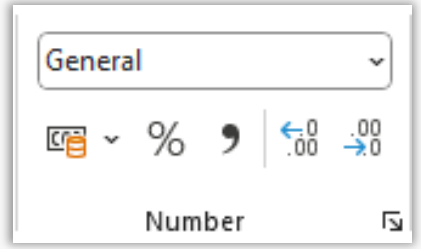


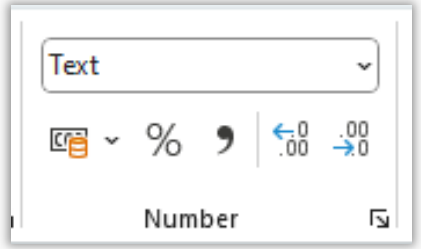


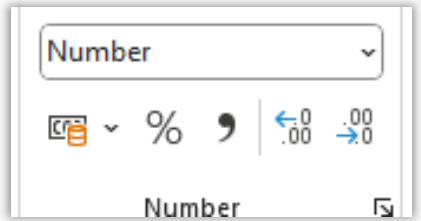


Data Types

- Numeric
  - Any kind of number not set to Text type
- Text
  - Characters, special symbols and numeric set to Text type
    - IDs tend to be text type
    - No math functions needed to be used
  - Also known as a ‘string’ of characters
- Logical
  - AND, OR, XOR, NOT
  - Returns boolean values: TRUE or FALSE
- Error
  - Indication of error present in a formula
  - #VALUE!, #REF!





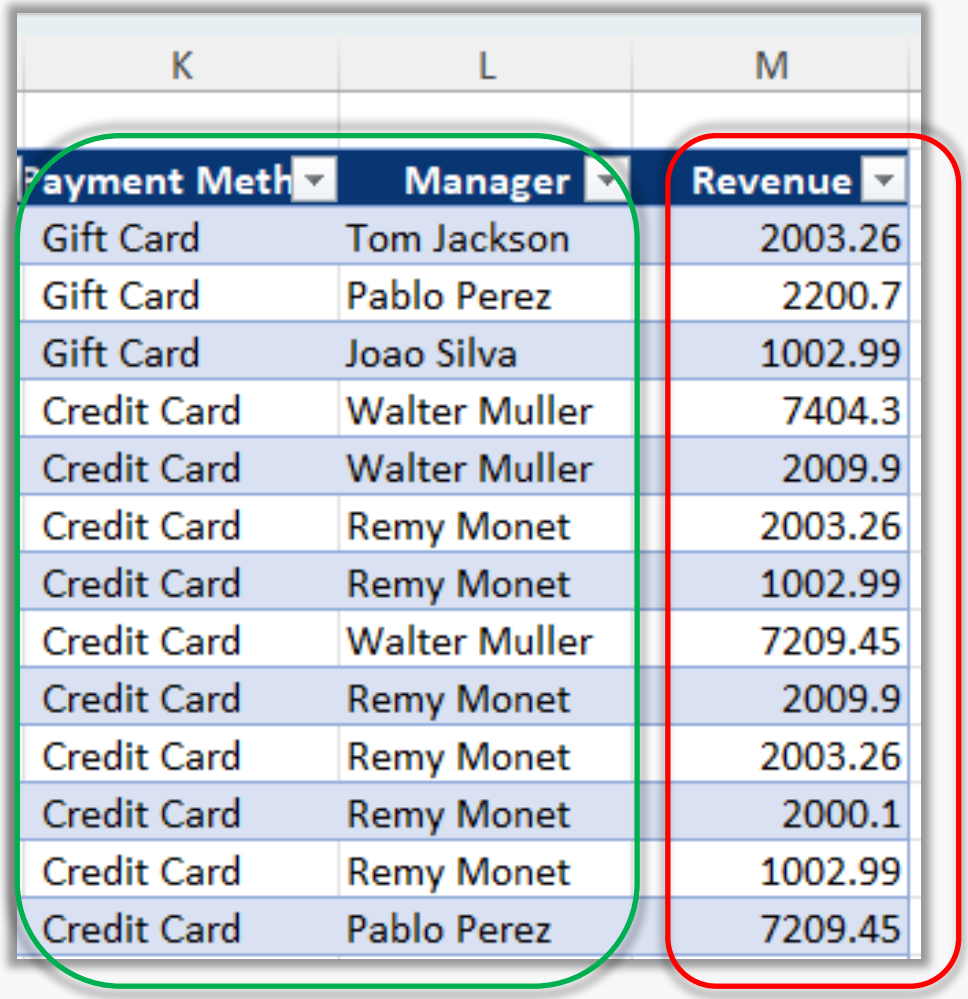


Data Types Alignment

- Excel uses a ‘General’ type as default
- Attempts to guess type

- You can manually set a column type

- Formatting hints at what data type is set
  - Text is Left aligned
  - Number is Right aligned

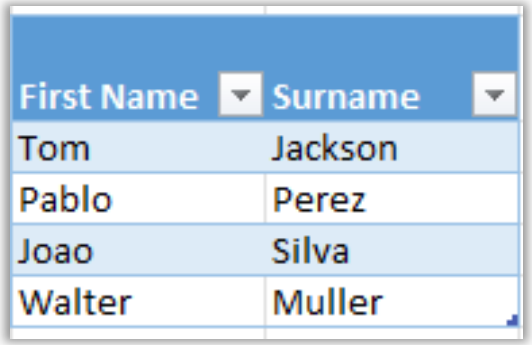


K	L	M
Payment Meth	Manager	Revenue
Gift Card	Tom Jackson	2003.26
Gift Card	Pablo Perez	2200.7
Gift Card	Joao Silva	1002.99
Credit Card	Walter Muller	7404.3
Credit Card	Walter Muller	2009.9
Credit Card	Remy Monet	2003.26
Credit Card	Remy Monet	1002.99
Credit Card	Walter Muller	7209.45
Credit Card	Remy Monet	2009.9
Credit Card	Remy Monet	2003.26
Credit Card	Remy Monet	2000.1
Credit Card	Remy Monet	1002.99
Credit Card	Pablo Perez	7209.45

CONCAT

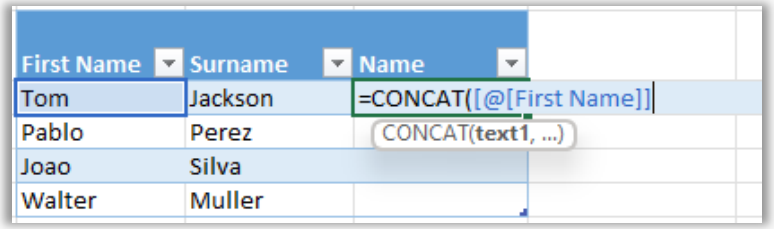
- Join text strings together in a new cell
  - Can also use the & symbol in newer versions of excel

1. First Name and Surname



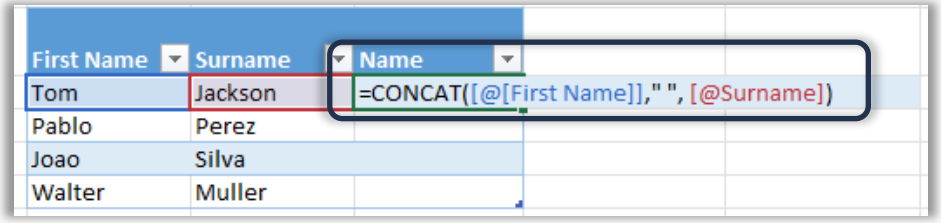
First Name	Surname
Tom	Jackson
Pablo	Perez
Joao	Silva
Walter	Muller

2. Start =CONCAT() Function
   - Select ‘First Name’ column



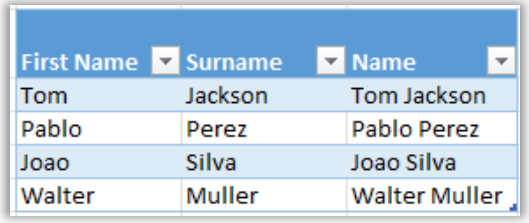
First Name	Surname	Name
Tom	Jackson	=CONCAT([@[First Name]]
Pablo	Perez	CONCAT(text1, ...)
Joao	Silva	
Walter	Muller	

3. Add a space between “ “ and select ‘Surname’ column



First Name	Surname	Name
Tom	Jackson	=CONCAT([@[First Name]]," ", [@Surname])
Pablo	Perez	
Joao	Silva	
Walter	Muller	

4. First Name and Surname are combined in the new cell



First Name	Surname	Name
Tom	Jackson	Tom Jackson
Pablo	Perez	Pablo Perez
Joao	Silva	Joao Silva
Walter	Muller	Walter Muller

‘&’ Symbol version:

Formula: `="The"&" "&"sun"&" "&"will"&" "&"come"&" "&"up"&" "&"tomorrow."`

Result: `"The sun will come up tomorrow."`

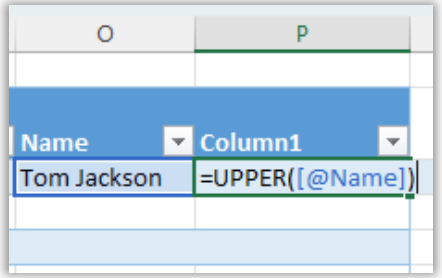
LOWER, UPPER, PROPER

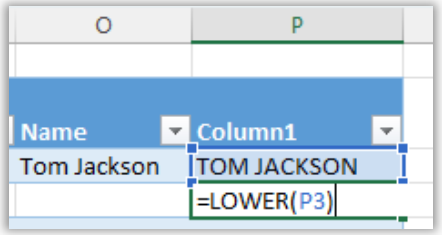
- These functions change the capitalisation in text fields

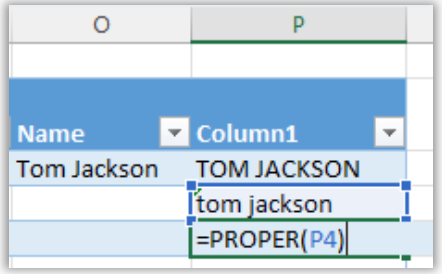
- =UPPER()
  - Change all text to upper-case

- =LOWER()
  - Change all text to lower-case

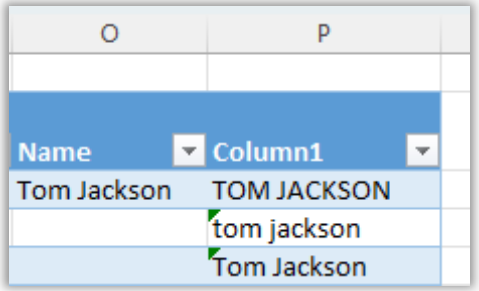
- =PROPER()
  - Change all text to have correct capitalisation







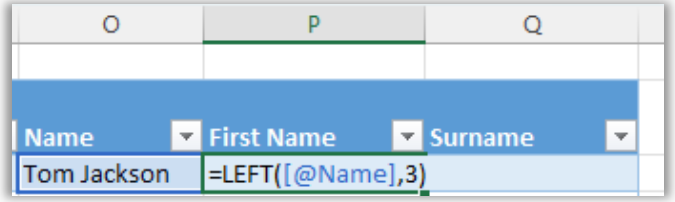
Result:



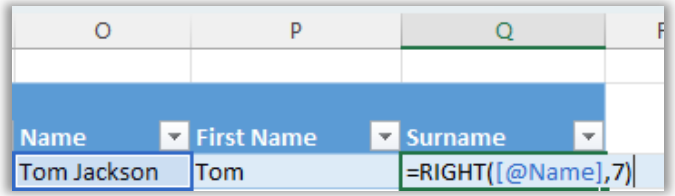
LEFT, RIGHT

- Pull out a number of characters from a cell

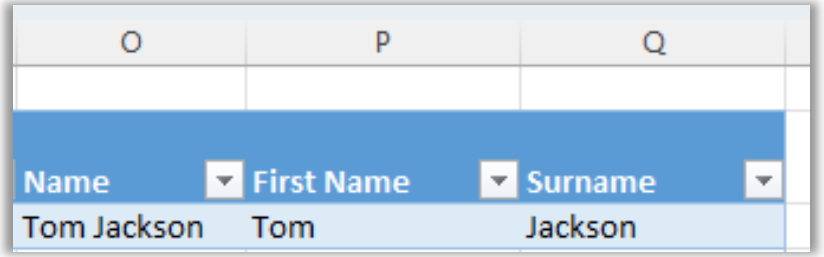
- =LEFT()
  - A number of character from the left of a cell



- =RIGHT()
  - A number of character from the right of a cell



Result:



NOW and TODAY

- NOW()
  - Returns exact date and time

- TODAY()
  - Returns only the date
  - Better for dates with no time element

NOW function

Result	Formula	
31-May-21 17:39	=NOW()	// current date and time
24-May-21 17:39	=NOW()-7	// last week same time
7-Jun-21 17:39	=NOW()+7	// next week same time
29-Aug-21 17:39	=NOW()+90	// 90 days from now
29-Aug-21 18:00	=MROUND(NOW()+90,"1:00")	// 90 days from now, rounded to nearest hour
31-Aug-21 0:00	=EDATE(NOW(),3)	// 3 months from now, time removed
31-May-22 0:00	=EDATE(NOW(),12)	// 12 months from now, time removed
1-May-21 0:00	=EOMONTH(NOW(),-1)+1	// first day of current month
1-Dec-21 17:39	=EDATE(NOW(),6)+MOD(NOW(),6)	// 6 months from now, time preserved

TODAY function

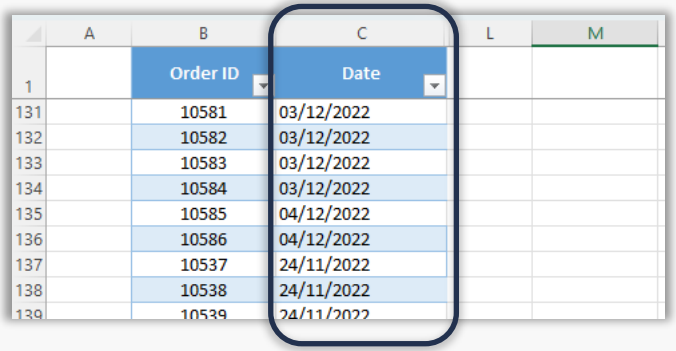
Result	Formula	
31-May-21	=TODAY()	// current date
24-May-21	=TODAY()-7	// one week in past
7-Jun-21	=TODAY()+7	// one week in future
29-Aug-21	=TODAY()+90	// 90 days from today
31-Aug-21	=EDATE(TODAY(),3)	// 3 months from today
31-May-22	=EDATE(TODAY(),12)	// 1 year from today
31-May-20	=EDATE(TODAY(),-12)	// 1 year in the past
1-May-21	=EOMONTH(TODAY(),-1)+1	// first day of current month
31-May-21 18:00	=TODAY()+TIME(18,0,0)	// today at 6:00 PM
1-Jun-21 12:00	=TODAY()+1+TIME(12,0,0)	// tomorrow at noon

BATH SPA UNIVERSITY

MONTH

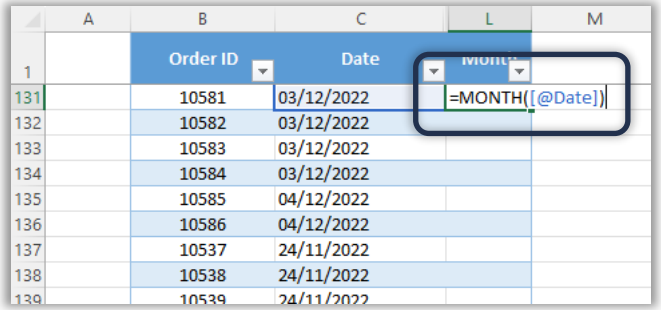
- Converting continuous dates into discrete dates
- Extracting the relevant part of the date to a new column

1. Date column



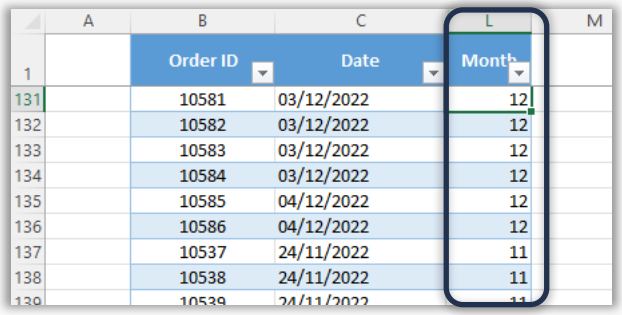
	Order ID	Date
131	10581	03/12/2022
132	10582	03/12/2022
133	10583	03/12/2022
134	10584	03/12/2022
135	10585	04/12/2022
136	10586	04/12/2022
137	10537	24/11/2022
138	10538	24/11/2022
139	10539	24/11/2022

2. Extract month using =MONTH()



	Order ID	Date	Month
131	10581	03/12/2022	=MONTH([@Date])
132	10582	03/12/2022	
133	10583	03/12/2022	
134	10584	03/12/2022	
135	10585	04/12/2022	
136	10586	04/12/2022	
137	10537	24/11/2022	
138	10538	24/11/2022	
139	10539	24/11/2022	

3. Shows the Month as a number



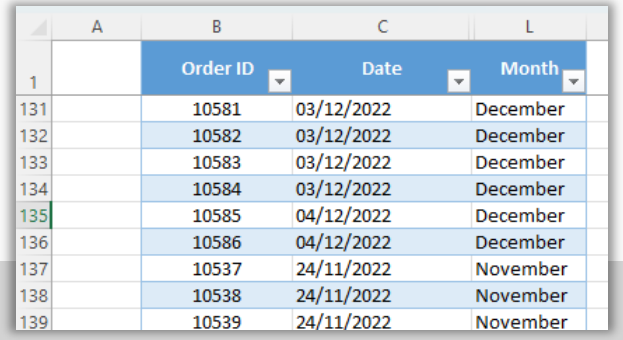
	Order ID	Date	Month
131	10581	03/12/2022	12
132	10582	03/12/2022	12
133	10583	03/12/2022	12
134	10584	03/12/2022	12
135	10585	04/12/2022	12
136	10586	04/12/2022	12
137	10537	24/11/2022	11
138	10538	24/11/2022	11
139	10539	24/11/2022	11

4. Extracting the month as a text value

- =TEXT(cell, “mmmm”)

```
=TEXT([@Date], "mmmm")
```

```
=TEXT([@Date], "mmm")
```



	Order ID	Date	Month
131	10581	03/12/2022	December
132	10582	03/12/2022	December
133	10583	03/12/2022	December
134	10584	03/12/2022	December
135	10585	04/12/2022	December
136	10586	04/12/2022	December
137	10537	24/11/2022	November
138	10538	24/11/2022	November
139	10539	24/11/2022	November

Note: “mmmm”: Displays the full date word

“mmm”: Displays the 3-letter date abbreviation

WEEKDAY

- Extracting the weekday into a new column

1. Date column

	A	B	C	L	M
1		Order ID	Date		
131		10581	03/12/2022		
132		10582	03/12/2022		
133		10583	03/12/2022		
134		10584	03/12/2022		
135		10585	04/12/2022		
136		10586	04/12/2022		
137		10537	24/11/2022		
138		10538	24/11/2022		
139		10539	24/11/2022		

2. Extract full weekday name =TEXT(cell, “mmmm”)

	A	B	C	L
1		Order ID	Date	Weekday
2		10452	07/11/2022	Monday
3		10453	07/11/2022	Monday
4		10454	07/11/2022	Monday
5		10455	08/11/2022	Tuesday
6		10456	08/11/2022	Tuesday
7		10457	08/11/2022	Tuesday
8		10457	08/11/2022	Tuesday
9		10459	08/11/2022	Tuesday

3. Extract abbreviated weekday name =TEXT(cell, “mmm”)

	A	B	C	L
1		Order ID	Date	Weekday
2		10452	07/11/2022	Mon
3		10453	07/11/2022	Mon
4		10454	07/11/2022	Mon
5		10455	08/11/2022	Tue
6		10456	08/11/2022	Tue
7		10457	08/11/2022	Tue
8		10457	08/11/2022	Tue
9		10459	08/11/2022	Tue

4. Weekday as a number =WEEKDAY(cell, 2)
   - Optional argument “2” is the week start day
     - Excel week starts on Sunday “1”, “2” = Monday

	A	B	C	L
1		Order ID	Date	Weekday
2		10452	07/11/2022	1
3		10453	07/11/2022	1
4		10454	07/11/2022	1
5		10455	08/11/2022	2
6		10456	08/11/2022	2
7		10457	08/11/2022	2
8		10457	08/11/2022	2
9		10459	08/11/2022	2

BATH SPA UNIVERSITY

Note: If no optional argument was entered, Monday would show as the number 2 in the last example

Excel Functions
Math and Logic

BATH SPA UNIVERSITY

Data Analytics and Machine Learning

Excel Calculations

- Performing calculations in Excel
- Simple calculations still work
  - E.g., E2 – D2
- Functions
  - Multiple rows to be calculated
  - SUM(), MIN(), MAX(), etc.

=SUM(E3,F3)

=MIN(E3,F3)

=MAX(E3,F3)

=AVERAGE(E3,F3)

=MEDIAN(E3,F3)

SUM

- Addition
- SUM() adds values in the selected range
  - Excel skips cells containing text or dates data
  - Count() will only count the number of cells that contain a numeric

=SUM(E3,F3)

MIN and MAX

- MAX() and MIN() functions often go hand in hand
- **MAX** returns the **largest** value in the range and **MIN** returns the **smallest**.
- Note
  - if the range doesn't contain any numbers both functions will return zero

=MIN(E3,F3)

=MAX(E3,F3)

AVERAGE

- Also called the ‘Mean’
- Provide a data range
  - Excel automatically calculate the average

- Note
  - Blank cells within the range will not be included

=AVERAGE(E3,F3)

MEDIAN

- Middle number of a group of numbers
  - Listed smallest to largest first

- Half are greater than the median
- Half are less than the median

- Often used as a comparison against the Mean
  - "The 2018 median income per capita was $33,706. The mean income per capita was $50,431."
  - The closer Median and Mean are, the more evenly distributed the values
    - If far apart then the data likely contains outliers
    - Study ‘Descriptive Statistics’ if you would like to know more in this area

=MEDIAN(E3,F3)

BATH SPA UNIVERSITY

Logic Functions

- Logic values are either TRUE or FALSE

- Ask complex questions and analyse the data
  - Determine if certain conditions are met
    - Returning TRUE or FALSE

- Logic is used often
  - If your hand gets too close to a fire, then you move it away
    - If you don't you might get hurt

  - If it's raining outside, you will use an umbrella on your walk
    - If you don't, you will get wet

=	equal	A1 = B1
>	greater than	A1 > B1
>=	greater than or equal to	A1 >= B1
<	less than	A1 < B1
<=	less than or equal to	A1 <= B1
<>	not equal to	A1 <> B1

- In data analysis, you can ask Excel to make similar True-False determinations.
  - Identify all rows in which a project had more than 50 backers
  - Check if a project is from the United Kingdom

IF

- Perform logic checks
- Arguments
  - Check criteria
  - Return value
    - On TRUE
  - Return value
    - On FALSE

```
=IF(L2>5, "Weekend", "Week Day")
```

- IF *Funding* (J2) is greater than *Goal* (K2)
  - On True
    - Label as ‘Successful Project’
  - On False
    - Label ‘Failed Project’

```
=IF(J2>K2, "Successful Project", "Failed Project")
```

Excel Functions
Nested Functions

BATH SPA UNIVERSITY

Data Analytics and Machine Learning

Nested Functions

- Multiple functions used on the same calculation
  - Test multiple criteria
  - **Executed from most inner function first**

- Can be done with various combination
  - E.g., SORT and FILTER
- “” quotations are not needed for numeric values

- Example:
- Find all In-store purchases using a Gift Card
  - If true = “x”
  - If false = “” (blank)

- Note: Ensure white space and cleaning has been performed

```
=IF(AND(H3="In-store",I3="Gift Card"),"x","")
```

Purchase Type	Payment Method	In-store Gift card
Online	Gift Card	
Online	Gift Card	
In-store	Gift Card	x
In-store	Credit Card	
In-store	Credit Card	
In-store	Credit Card	
In-store	Credit Card	
In-store	Credit Card	

Task 1 - Exploring data in Excel (15 minutes)

1. Download and open bad_data.xlsx. Start on the ‘Marketing’ tab.

2. A manager needs to use this data to work out how effective the marketing team has been. They have the following questions:
   - How many sign-ups have there been this month?
   - How does that compare to the last 6 months?
   - Which member of the team is getting the most sign-ups?
   - How has the total potential value changed over the last 5 years?

3. In your groups, discuss:
   - Can this data answer the manager’s questions?
   - If not, why not?
   - What issues can you see with the data?

4. Each group to feed back – **don’t forget to elect a spokesperson**!

Task 2 - Cleaning data in Excel (30 minutes)

1. Go back to the Marketing tab in bad_data.xlsx.

2. Use Excel functions to clean and structure the data so it can answer the manager's questions:
   - How many sign-ups have there been this month?
   - How does that compare to the last 6 months?
   - Which member of the team is getting the most sign-ups?
   - How has the total potential value changed over the last 5 years?

Business Intelligence

“Business intelligence is the delivery of accurate, useful information to the appropriate decision-makers within the necessary time frame to support effective decision making.”

Keywords

Keyword	Description
Relationship	Database relationships are associations between tables that are created using join statements to retrieve data.
Row	A row is a data record within a table. Each row, which represents a complete record of specific item data, holds different data within the same structure.
Column	A column is a set of data values of a particular type, one value for each row of the database. A column may contain text values, numbers, or even pointers to files in the operating system.
Visualisation	A visual display of information to achieve one or more objective. It alerts users on issues or problems Operational, Performance, Personal, etc.
Datasets	A dataset is something which you import or connect to. Datasets can be renamed, refreshed, removed, and explored.
Dashboard	The dashboard is a collection which contains zero or more tiles and widgets. It is used to represent a customized view of some subset of the underlying datasets.
Reports	A Power BI report is one or multiple pages of visualizations. It can be created from scratch, imported to a dashboard, and created using datasets.
Tile	Is a single visualization found in a report or on a rectangular dashboard box which contains each visual.

Business Intelligence

- Business Intelligence (BI)
- Optimise business performance
- Inform business decisions
  - Long-term strategic planning
  - Shorter-term tactical choices
- Improve performance
- Gain a competitive advantage
- Collect, analyse, visualise and present company data

- **Example 1**
  - A retail store analysing sales data to identify which products are selling well and which are not.
- **Example 2**
  - A manufacturing company optimising its production processes and reducing costs.
- **Example 3**
  - A hospital analysing patient data to improve patient outcomes.

Collecting Data

BI involves collecting data from various sources

1. Identify sources
    - Databases, spreadsheets, market research
2. Collect
    - Importing files, connecting databases, APIs
    - Collate in a single location
3. Data quality
    - Ensuring data is high quality
    - Validate and clean the data
4. Store and organise
    - Store in table format
        - Database or Power BI

- **Example 1**
    - A marketing team uses BI to collect data
        - Social media
        - Website analytics
        - Customer surveys
- **Example 2**
    - A logistics company uses BI to collect data
        - GPS tracking devices
        - Warehouse management systems
        - Shipping providers

Data Analysis

- Analyse data and identify patterns, trends, and relationships
- Explore data
  - Understand structure and content
  - Summary statistics, visualisations and data profiling
- Clean and transform
  - Suitable format for analysis
  - Remove outliers, fill in missing values, convert data types
- Create models
  - Show relationships between different data variables
  - Regression, clustering and decision trees
- Generate insights
  - Identifying patterns and trends
  - Predictions or recommendations

Example

A supermarket chain wants to understand the shopping habits of its customers to improve its marketing strategies.

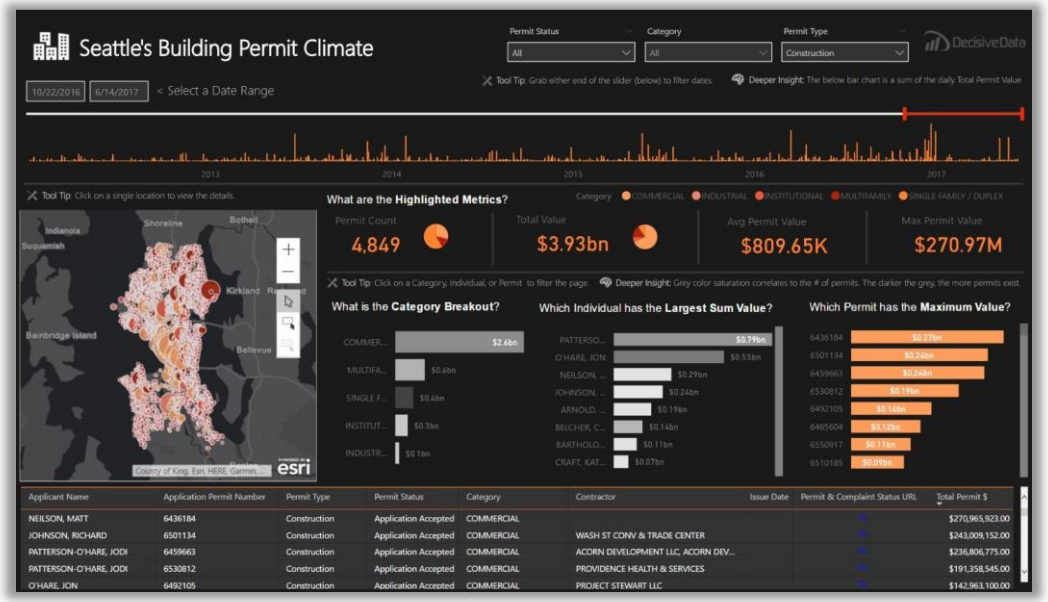
1. Collects data from its loyalty program - customer demographics, purchase history, and store visits.
2. Analyses this data to identify patterns and trends
3. Clustering techniques to segment customers into different groups based on their shopping habits.
4. Create targeted marketing campaigns for each customer group.

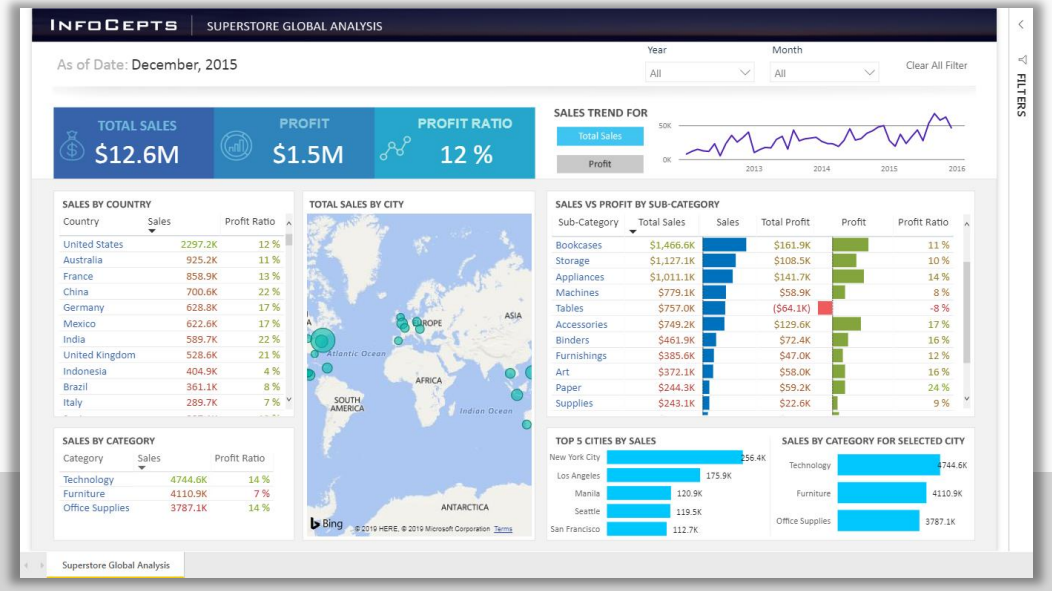
Goal: Increase sales and customer loyalty.

Data Visualisation

Create visualisations such as charts, graphs, and dashboards to help users understand the data

- Type of visualisation
  - Different types of data
  - Better to use certain visualisations
- Keep it simple
  - Easy to understand
  - Avoid cluttering with too much information
  - Avoid unnecessary details
- Provide context
  - Captions, titles and explanatory text
- Interactivity
  - Allow users to explore the data
- Test visualisations
  - Ask for feedback

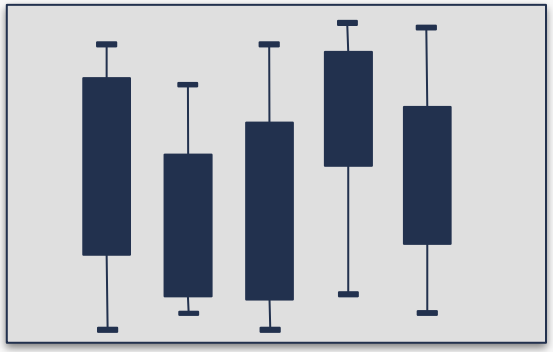




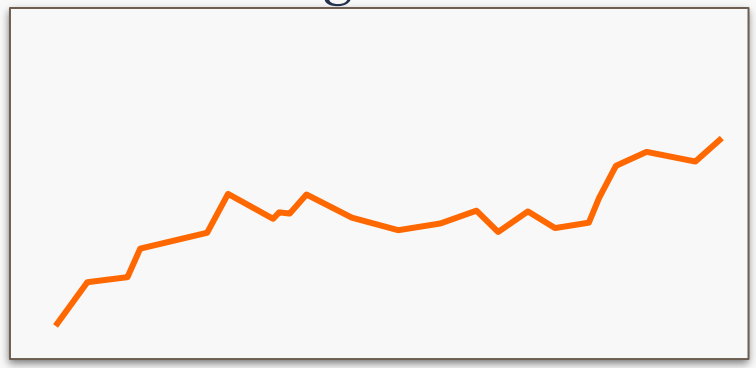
The language of dashboards

"What is it right now?"

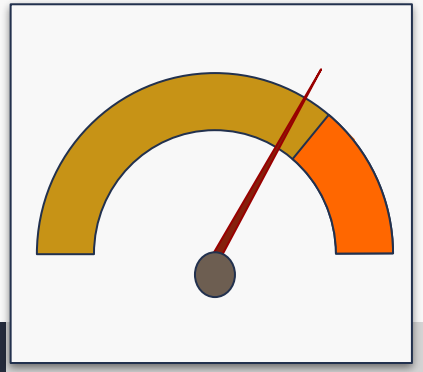
"Is it the same in all cases?"



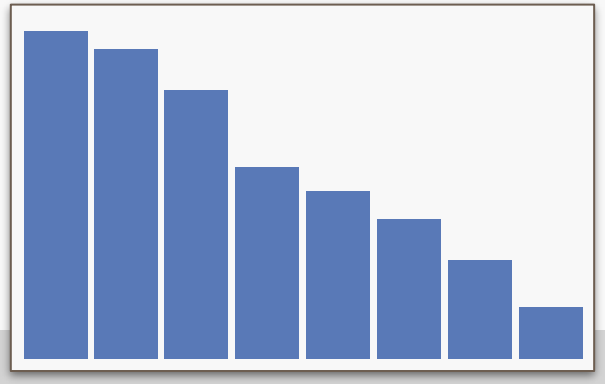
"Does it change over time?"



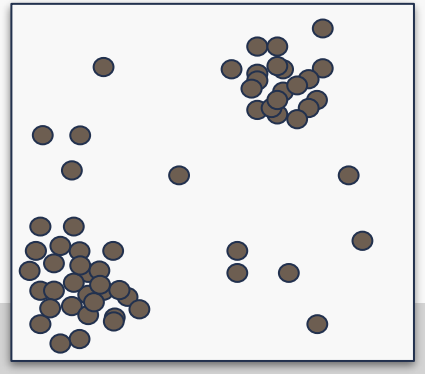
"How far away is the goal?"



"How much / how many?"



"Is there a pattern?"



Introduction to Power BI

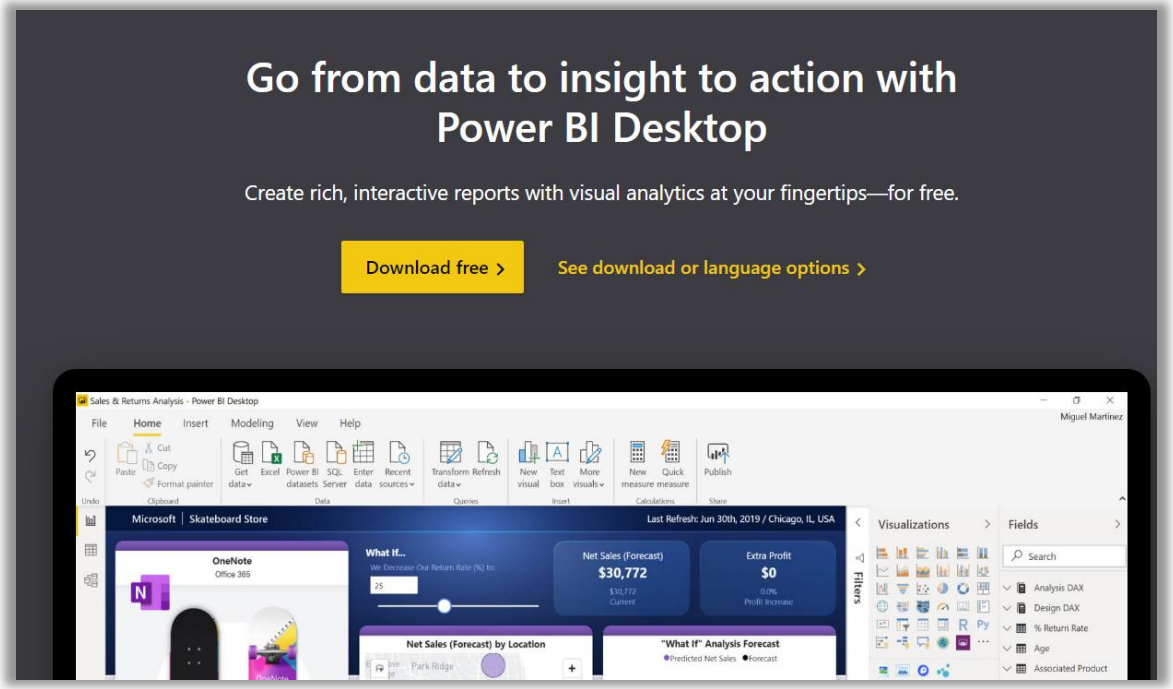
Power BI

BATH SPA UNIVERSITY

Data Analytics and Machine Learning

Installing Power BI

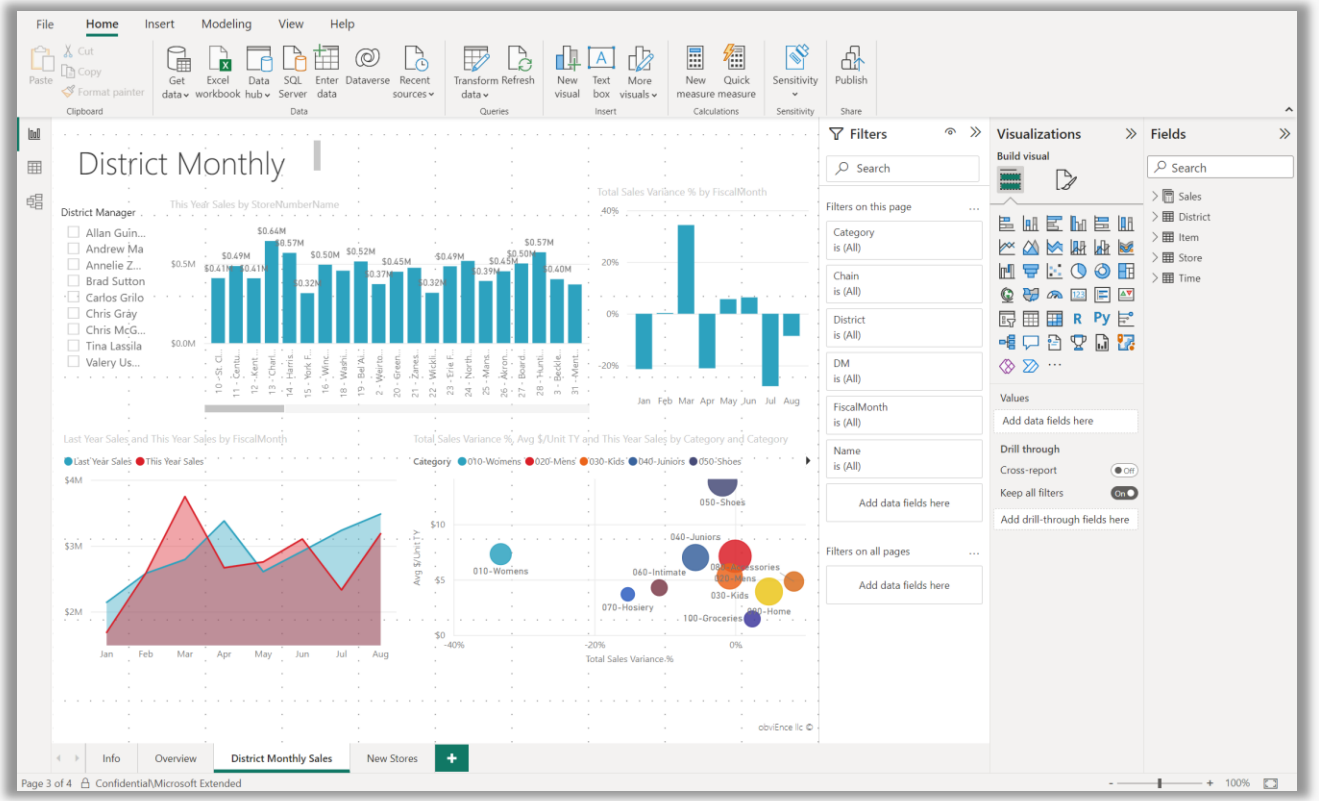
- Power BI Desktop (Windows users only): https://powerbi.microsoft.com/en-us/desktop/
  - Use Windows where possible
- Mac users: Downloads | Microsoft Power BI
  - Very limited options
- Tasks will be set as group work
  - At least one windows user per group to share screen



What is Power BI?

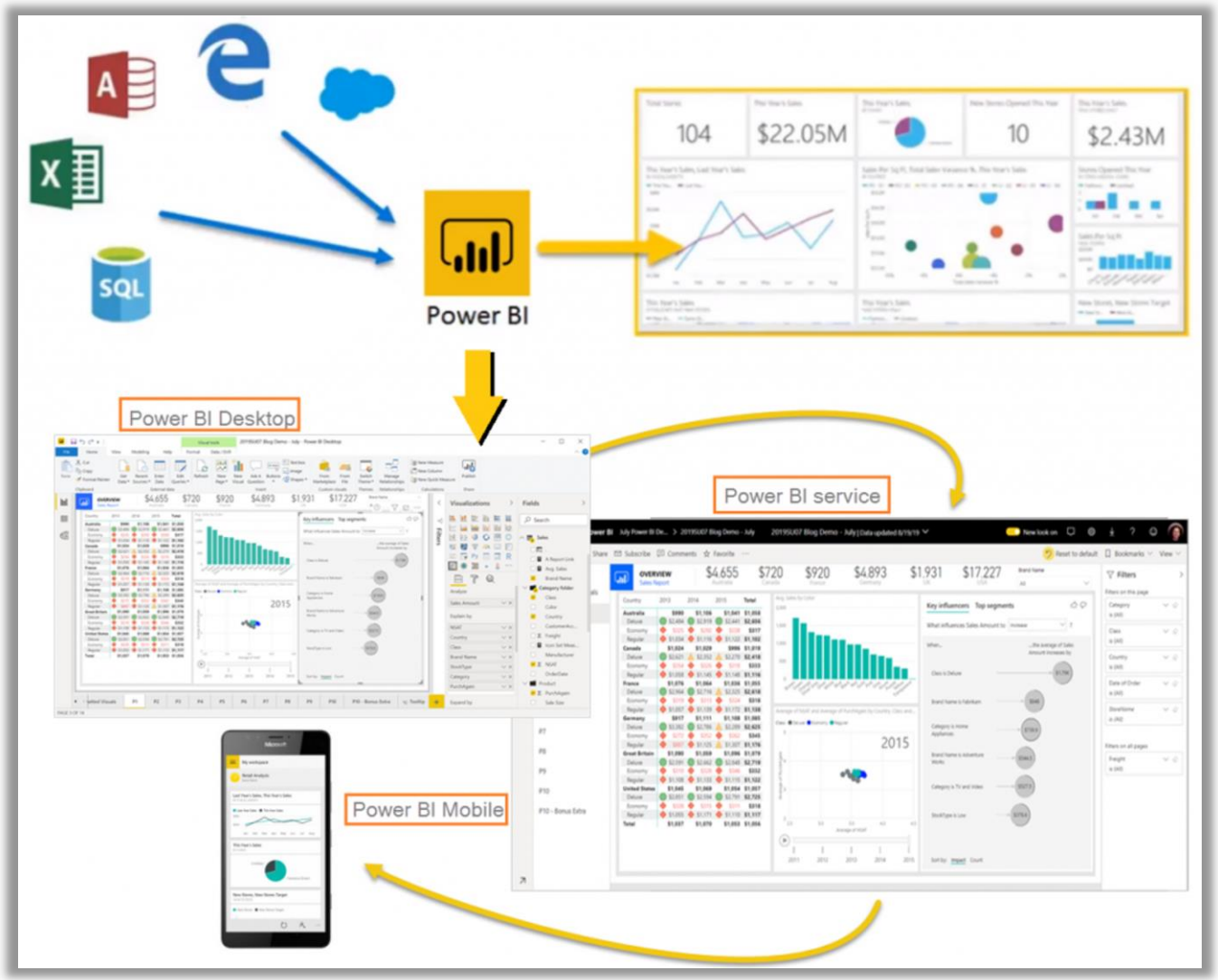
- Power BI is a business analytics service by Microsoft
- Power query included
- Interactive visualisations, reporting and apps
- Business intelligence capabilities
- Input data from various sources
  - Databases, webpages, PDFs, structured files
- Several services
  - Power BI Desktop
  - Power BI service
  - Power BI Mobile apps
- Connect to multiple data sources
  - On-premises or in the cloud sources

Power BI



What is Power BI?

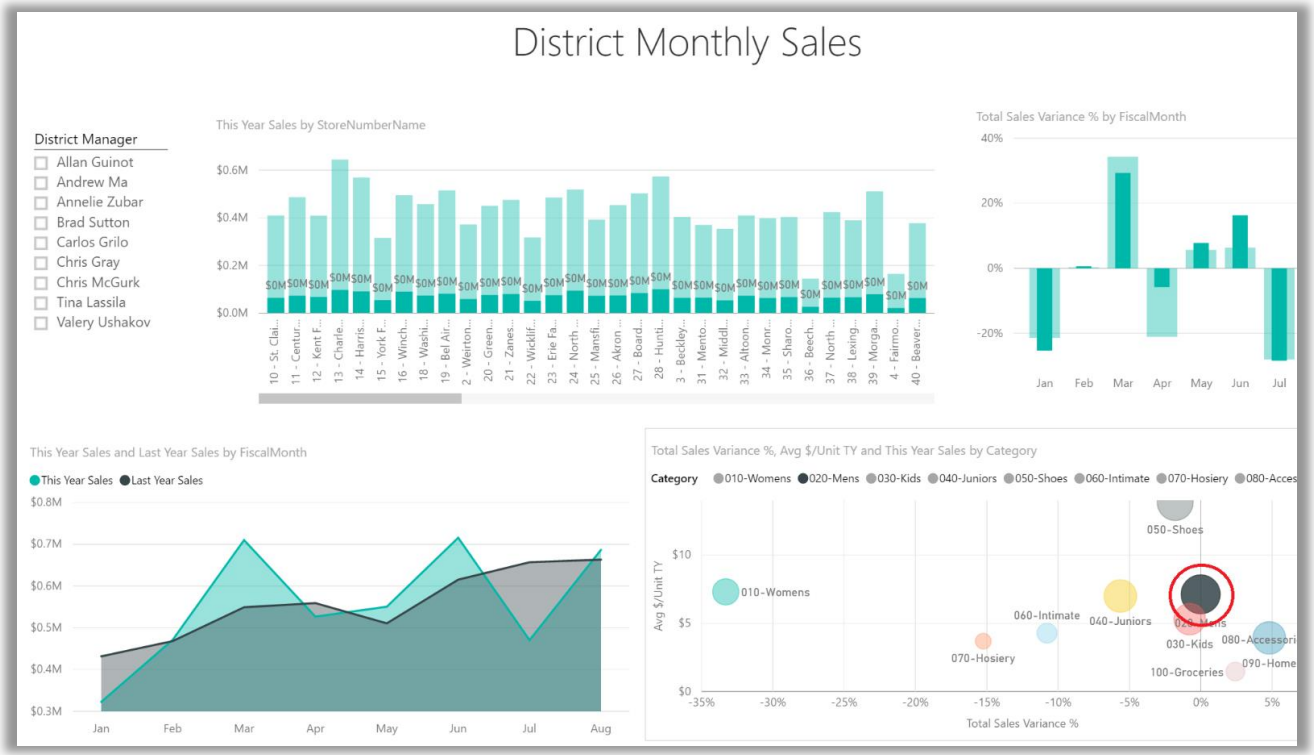
- Power BI Desktop
  - Free data analysis and report authoring tool
  - Windows operating systems
  - Connect to data
  - Reports
  - Visualisations
  - Share reports
- Power BI service
  - Secure cloud service
  - Share and collaborate
    - View dashboards
    - Reports
- Power BI Mobile apps
  - Windows
  - iOS
  - Android



Why use Power BI?

- Power BI is the leading BI tool (according to Gartner).
- Over 97% of Fortune 500 companies use Power BI.
- Power BI has over 6 million customers.
- Ease of Use
  - Simple and easy-to-use interface
  - No programming experience required
  - Simple UI to connect to data sources
  - Drag & drop selection of attributes for reports
- Allows for collaboration within an organisation
- Multiple data sources
- Data shaping
- Report Builder

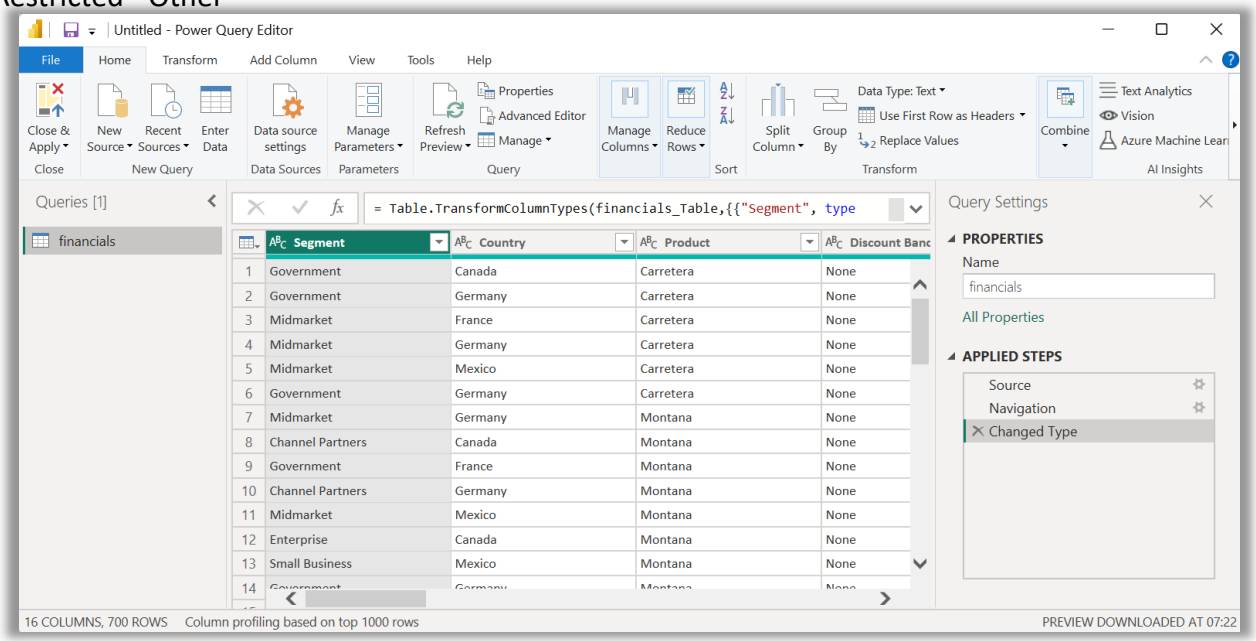
Power BI

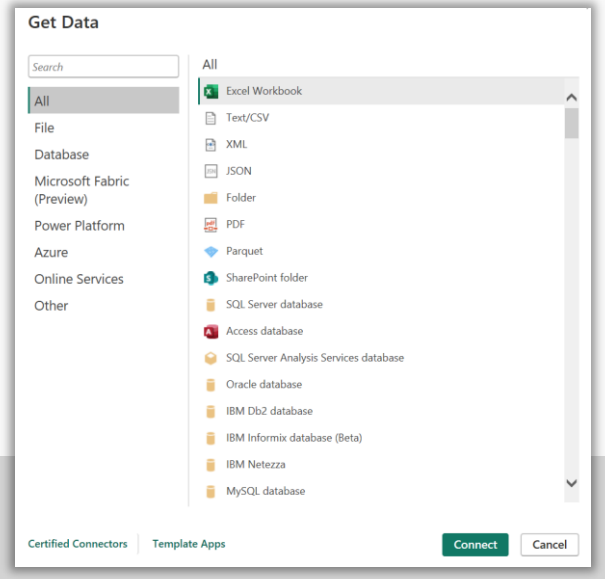


Power Query Editor

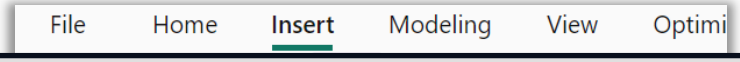
- Tool for data gathering and transformation
- Extract data from a variety of sources
- Manipulate data
  - More suitable for analysis

- Data Sources
  - Text and Excel files
  - Databases
  - Internet sites
- Data Manipulation
  - Splitting data into multiple columns
  - Pivoting and unpivoting
  - Replacing values
  - Modifying data types
  - Merging data
    - Multiple data sources
    - A single table





Power View



- V
- C
- C
- N
- C

USE WITH CAUTION!

Relying on AI tools can lead to nonsense or unpredictable errors in your analysis.

Power BI Differences: Desktop and Web

Desktop

- Cannot share Report or Dashboard
- Can write DAX Queries
- More Data source types
- Able to apply Relationships between multiple data sources
- More features overall

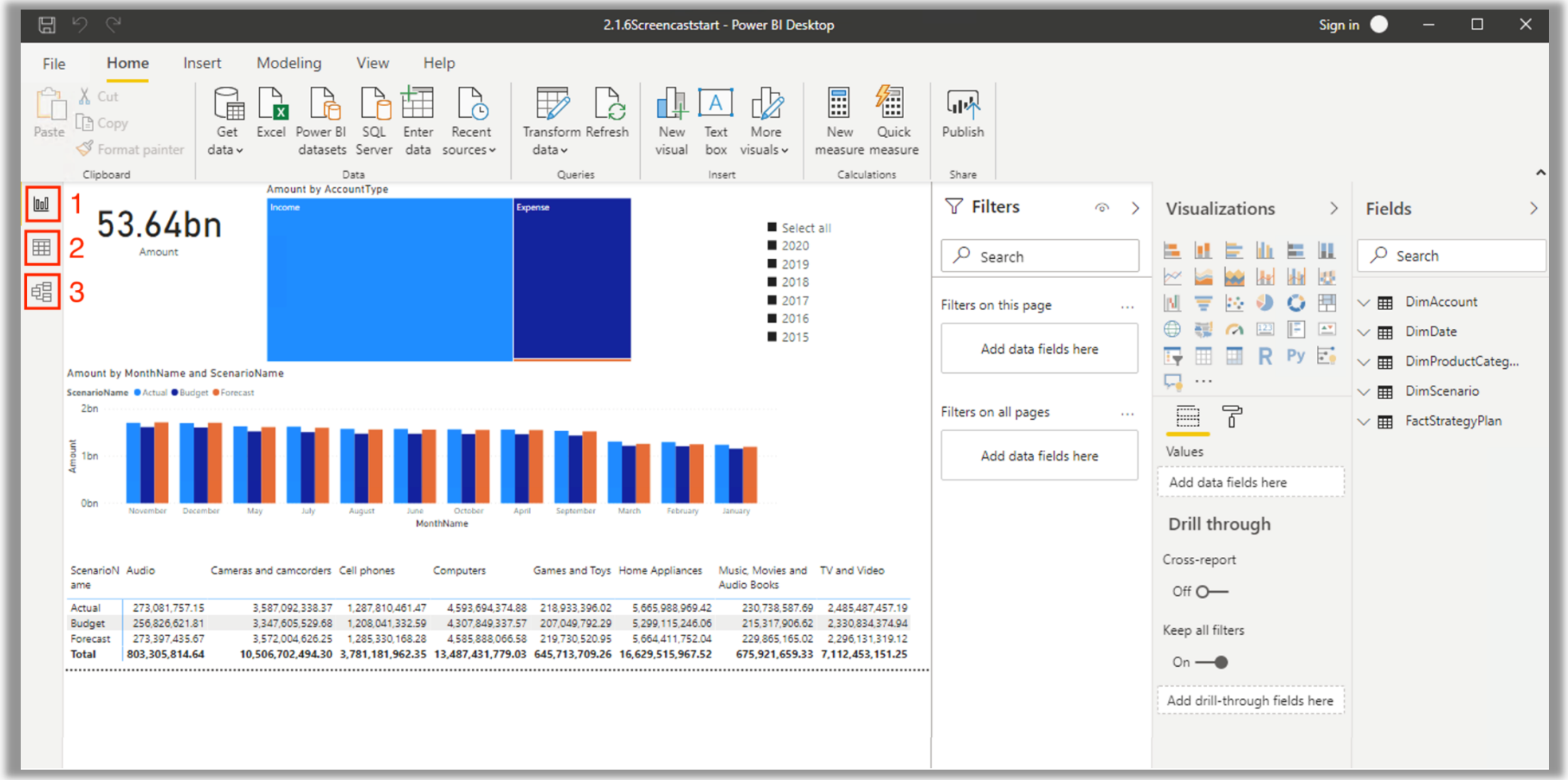
Web (and Mac)

- Can share Report or Dashboard
- Unable to write DAX Queries
- Less Data source types
- Unable to apply Relationships between multiple data sources
- Less features overall

Power BI Interface

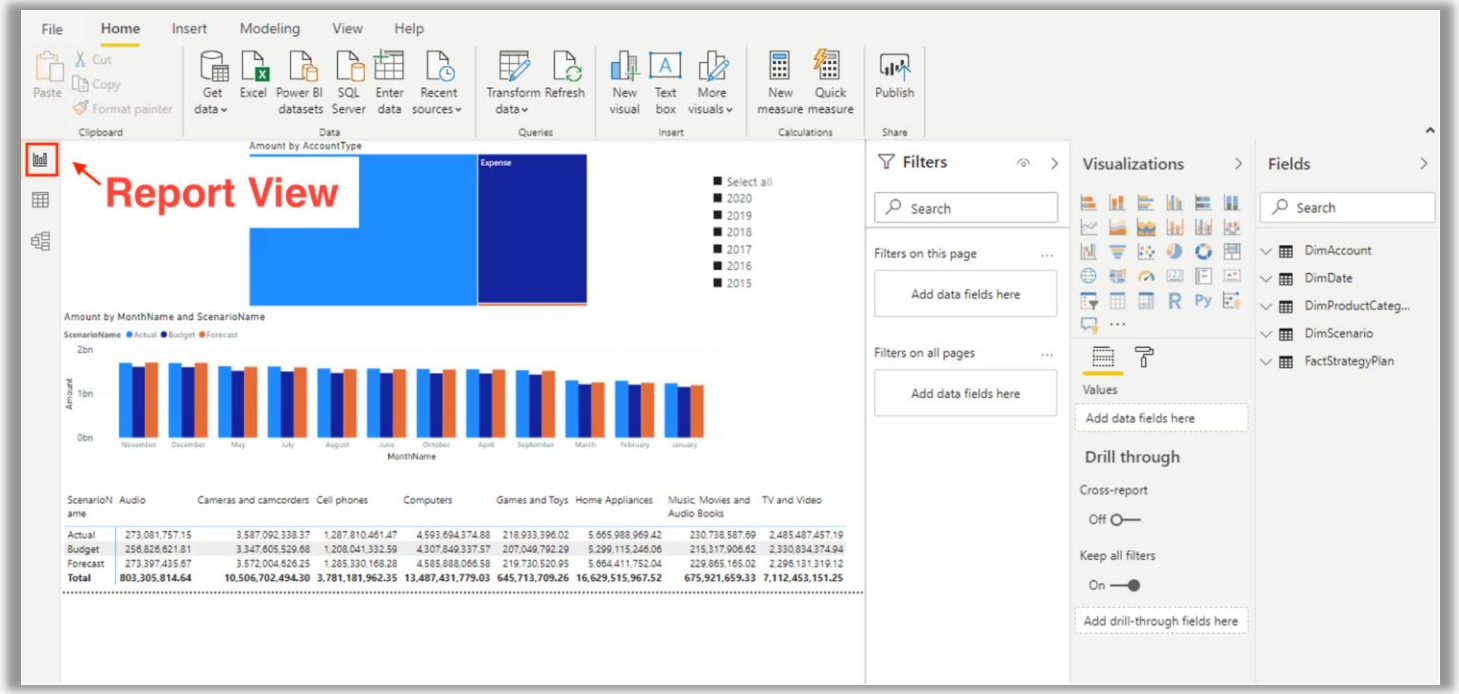
Power BI Interface

1. Report View
2. Data View
3. Model View



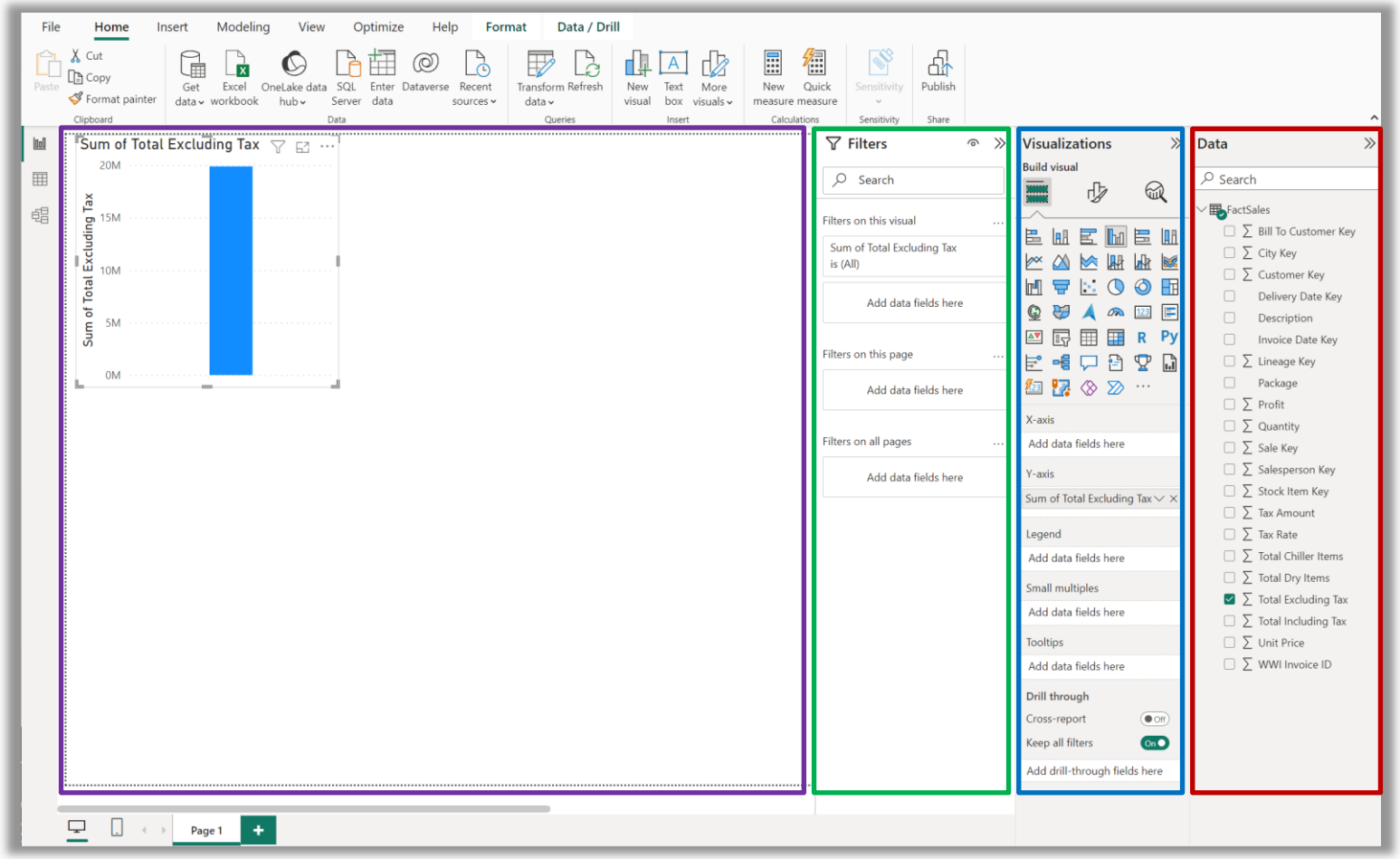
Report View

- The report view
  - Default view
- Features available in this view
- Create and customize visualizations
- Add and format text
- Add images and shapes
- Interact with data
- Apply filters and sorting
- Create custom calculations



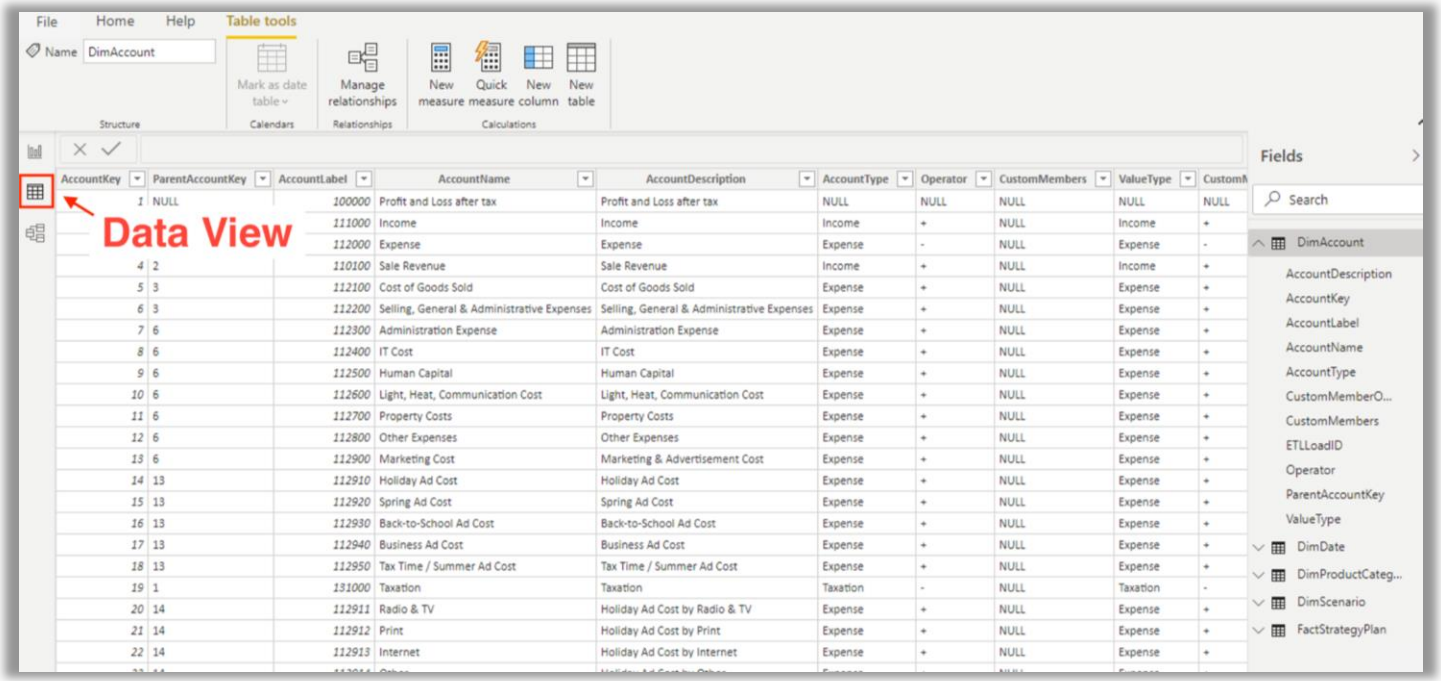
Report View

- Canvas
  - Visualisation Layout
- Filters
  - Filter data for your visualisations
  - Changes to canvas element selected
- Visualisations
  - Add, Change, Customise Visualisations
- Data
  - *'Fields'* depending on Power BI version
  - Available Data
    - Tables
    - Fields
- All panes change to match selected Canvas visualisation



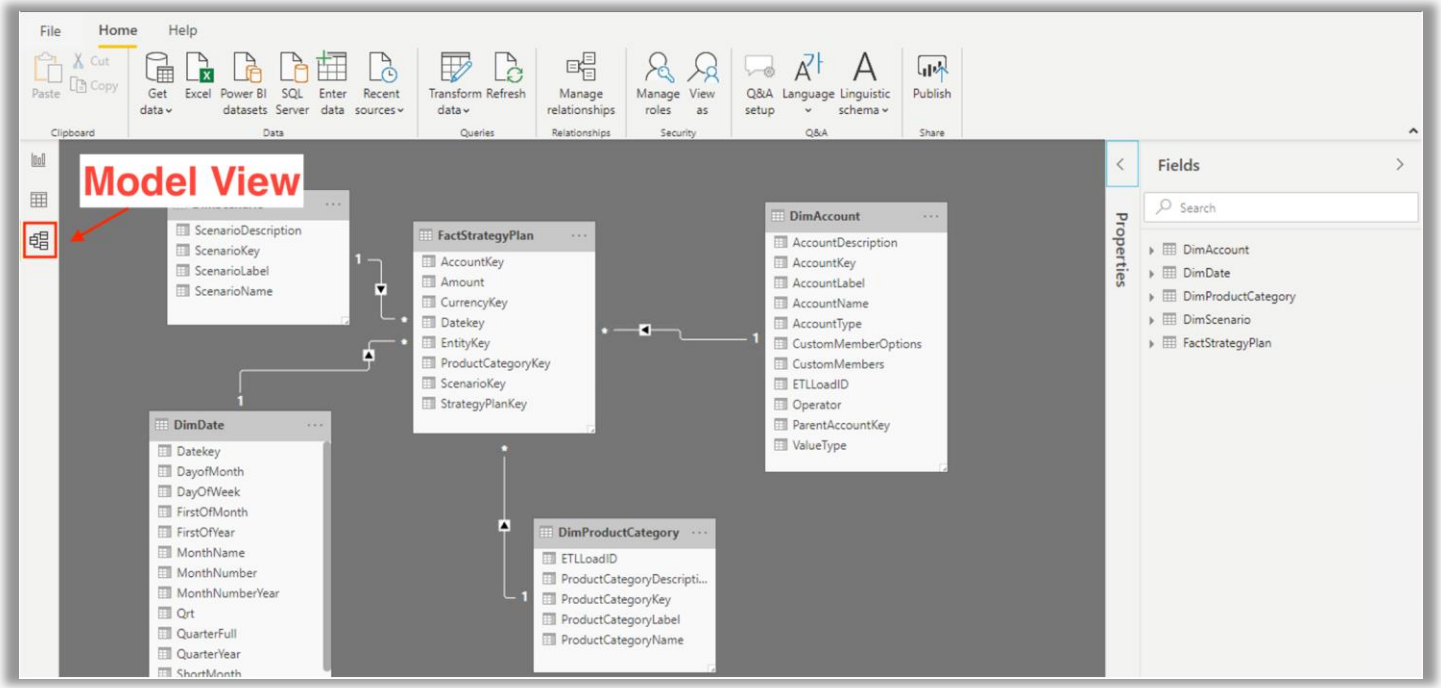
Data View

- View and edit data
  - In tables and columns
- Create relationships
  - between tables to link them together
- Create calculated columns
  - Calculate columns using DAX formulas to add new data to your tables
- Create measures
  - Measures using DAX formulas to perform calculations on your data
- Format data
  - Control how data is displayed in visualisations
- Manage columns and tables
  - Rename, delete, hide, and reorder columns and tables



Model View

- View and manage relationships
  - Between tables in your data model
- Create and edit hierarchies
  - Organize data into levels of detail
- Manage columns and tables
  - Rename, delete, hide, and reorder columns and tables
- Create calculated tables
  - Add new data to your model
- Create KPIs
  - Create Key Performance Indicators (KPIs)
  - Measure progress towards goals
- Manage perspectives
  - Provide focused views of your data model



Power BI Data

Task 3 - Importing Data (Groups) - 60 Mins

Must be at least one Power BI Desktop (Windows) user sharing screen in each group

1. Import multiple files
   - Bad_data.xlsx – Membership sheet
   - sicCodes.xlsx

2. Power Query
   - Use Power Query to transform and clean the data
     - Are the column headers descriptive?
     - Does each column contain only one data type?
     - How will you create relationships between tables?
     - What other issues can you find with the data?

3. Each group to present to class
   - Actions taken
   - Reasons why
   - In a PowerPoint slide or Word document with images

4. All members to upload **PDF version** of slides or document to Teams

This is designed as an exploratory group task. There may be differences in opinion! Work as a team to discuss which fits best

Power BI Data

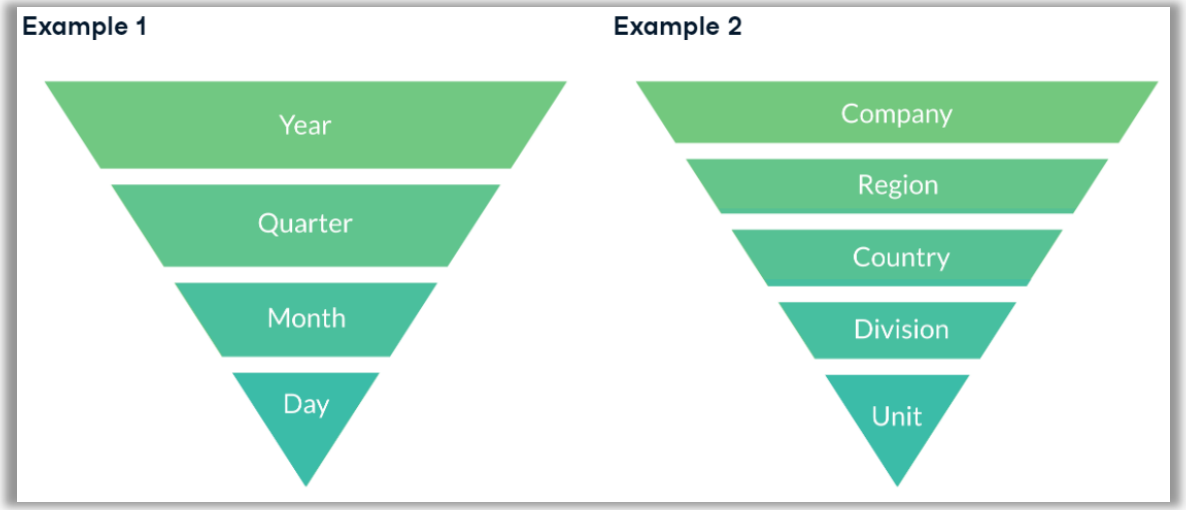
Hierarchies

BATH SPA UNIVERSITY

Data Analytics and Machine Learning

What are Hierarchies in Power BI

- Hierarchies allow you to organize data into different levels of detail
- Arranging columns in a specific order
- Drilldown
  - Explore data at different levels of detail
    - Within a hierarchy
  - Drill down to see more detailed data
- Drill up
  - More aggregated data



Task 4 - Create a visualisation (Homework)

1. Using the Membership table, create a visualisation that shows the relationship between month and renewal.
2. Using the membership table and the SIC codes table, create a second visualisation that shows the proportion of each high-level industrial classification in the Membership data. (HINT – industrial classification is in the first column ‘SECTION’ of the SIC codes table).
3. Optional: do some research and see if you can find the names or descriptions of the high-level industrial classifications. How would you add them to your visualisation?

Any questions?