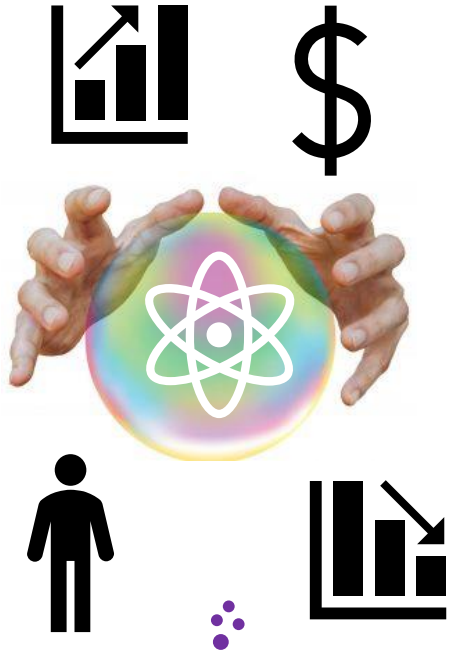
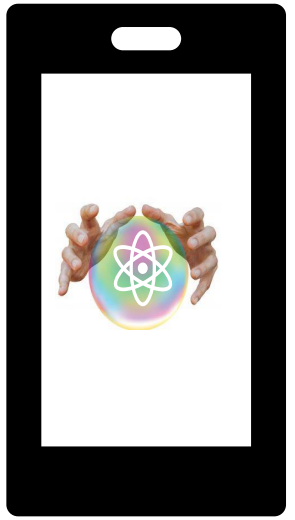
PowerBi

Visualizations and Reporting

Dash

By: Kristoffer Kofoed Rødvei





Data Engineering

Reporting

Data Analytics

Production

Engineer+

Data Engineer

ETL
point & click

DBA+:
-View,
-Indexes

Server-Tweaking/ Data-Governance

Reporting
custom:
- DAX

Presenting

Deep Learning:
- MLP
- LSTM
- CNN

- RL

ML-Models:
- Decision tree
- Monte Carlo

- Train, test, val analytics

Stat-Models:
- Linear reg
- Logistic reg
- Clustering
- H_0, CI, ROC

Cloud:

Server:

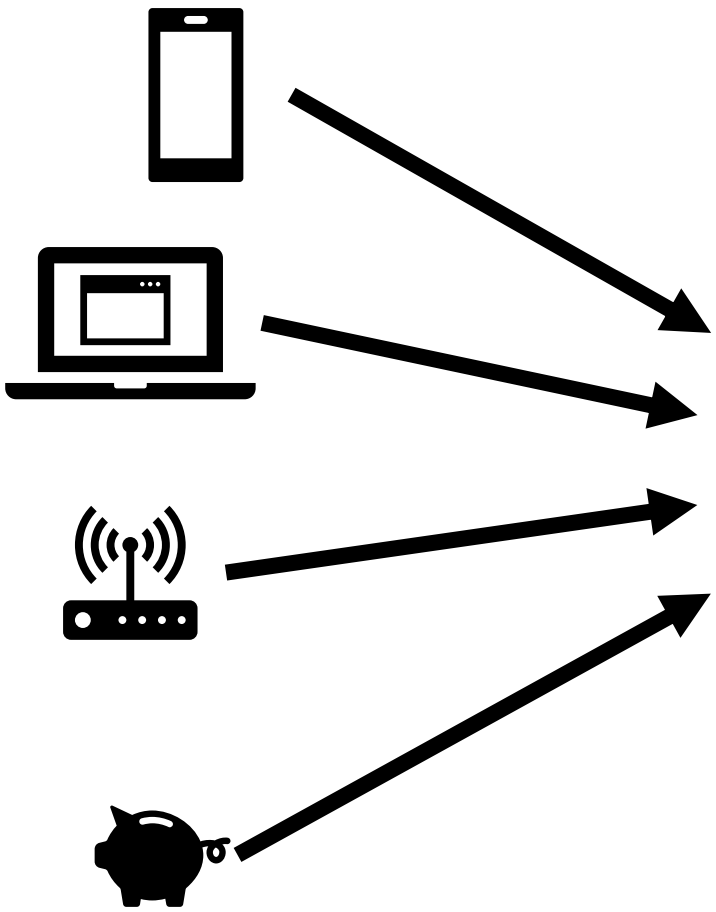
Data Management

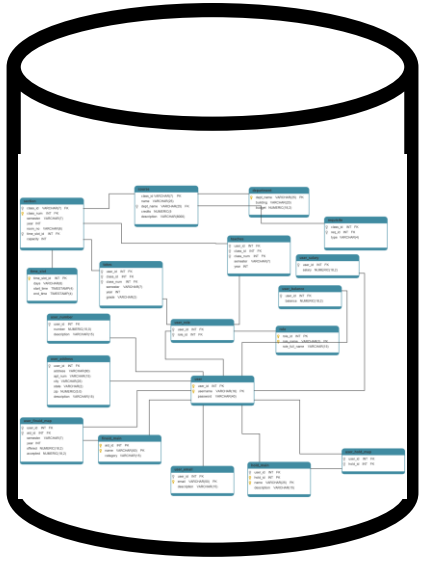
Reporting

Data Analytics

Production

Summary week 1, 2, 4 and 5



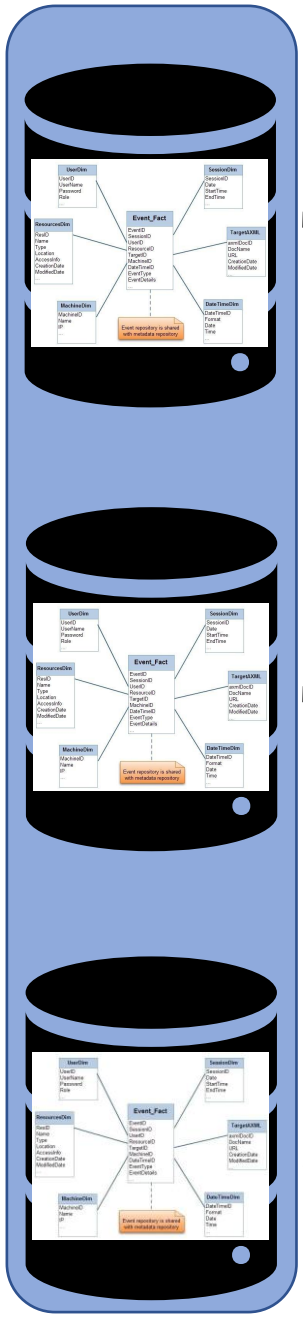


Database/Datalake

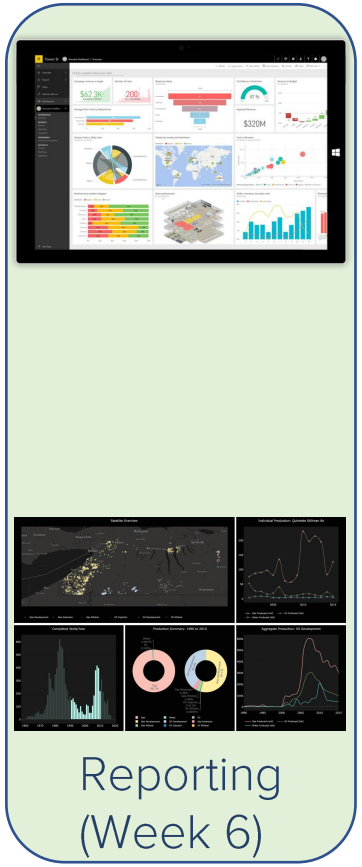
ETL

ETL

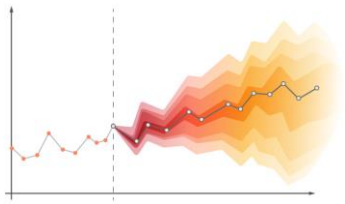
ETL



Data Warehouse

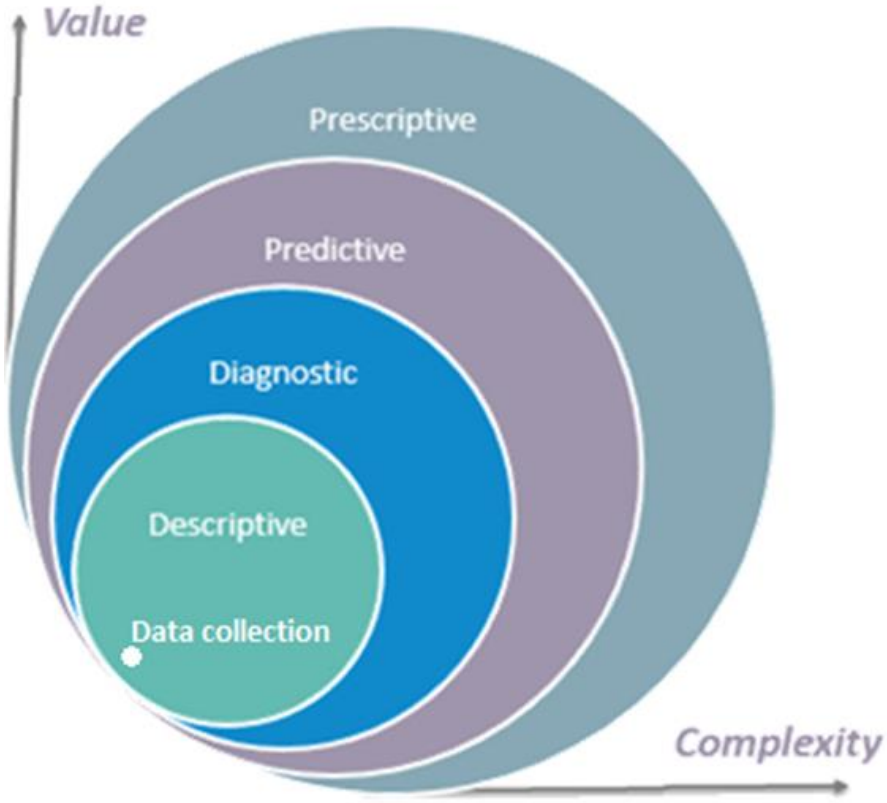


Reporting (Week 6)



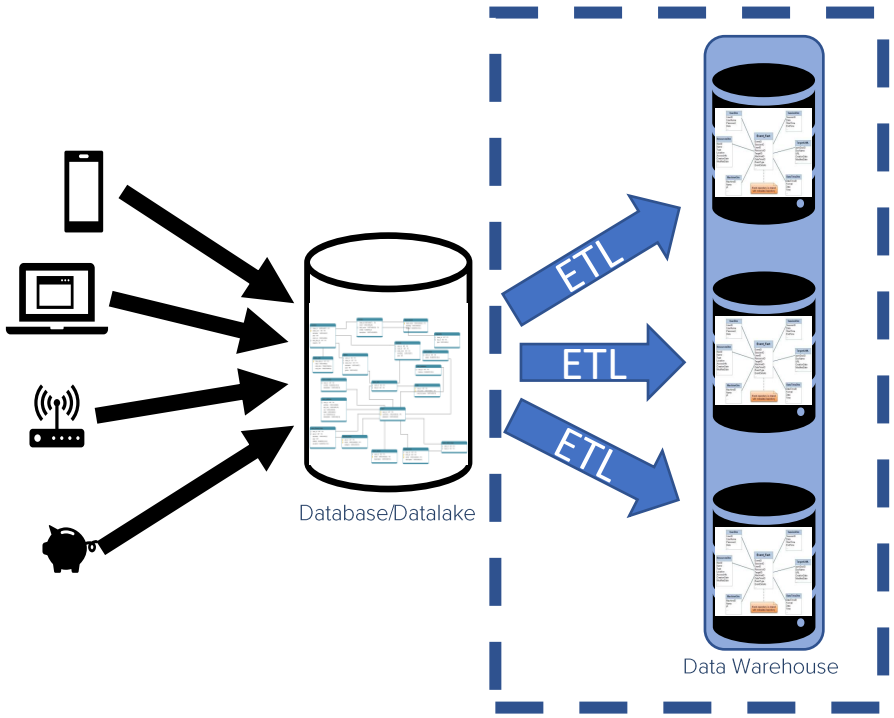
Prediction

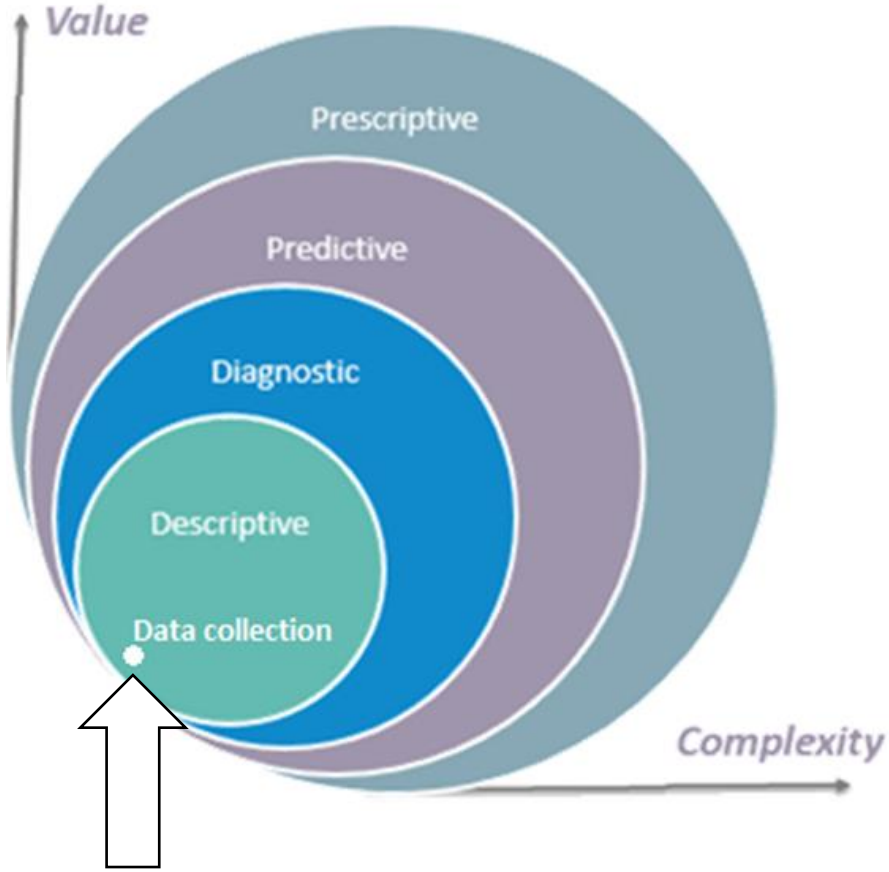
Data Science Levels



Data Science Levels

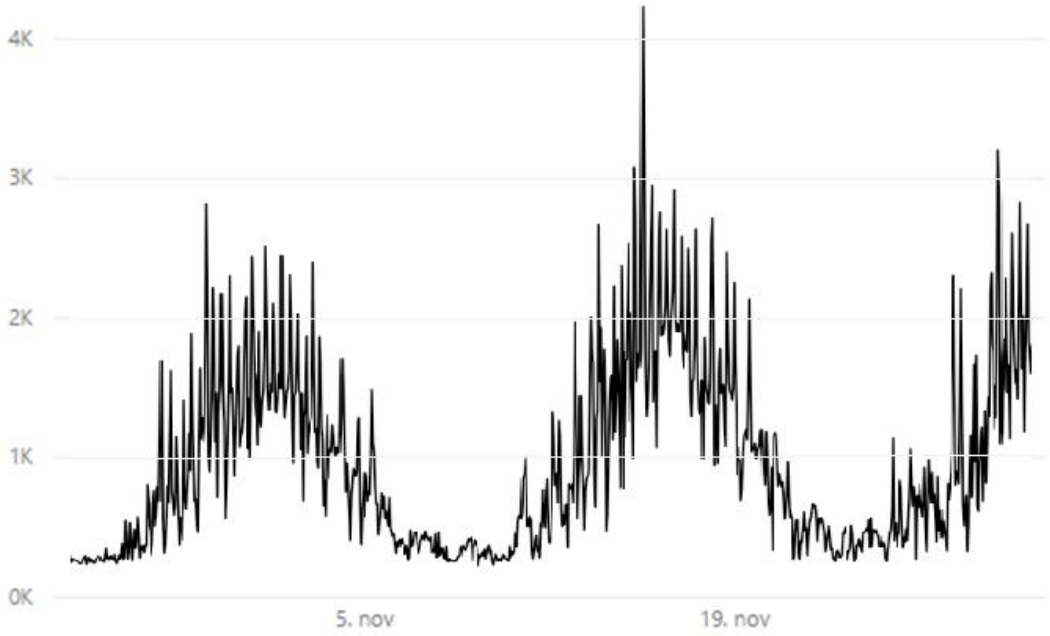
- Data collection

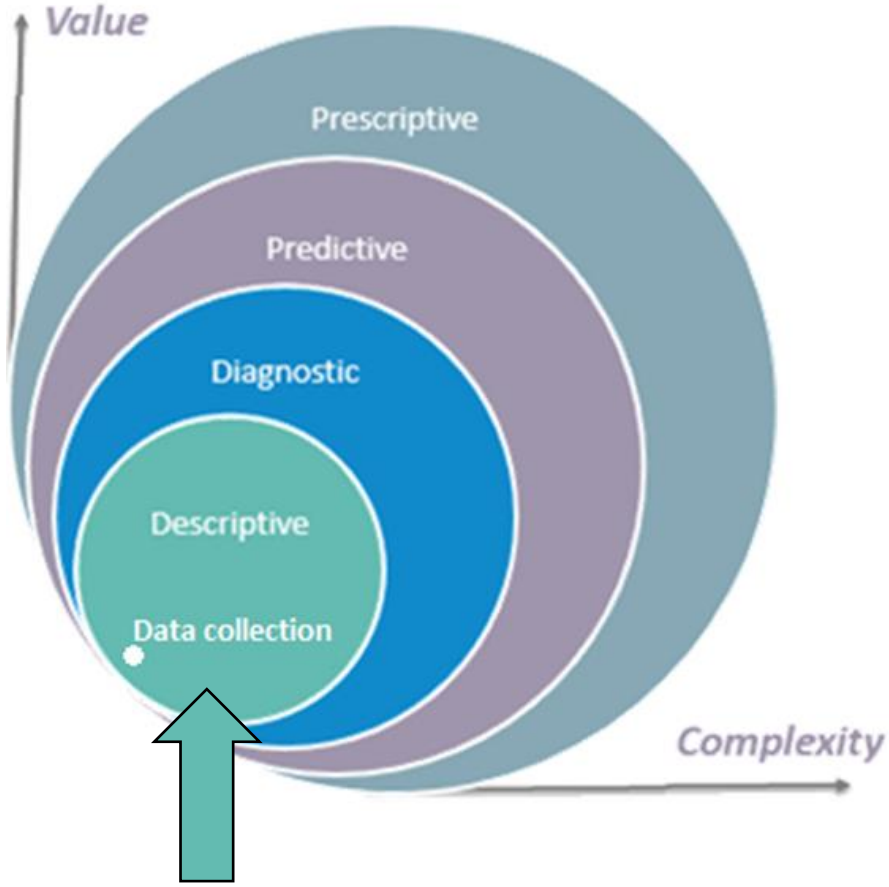




Data Science Levels

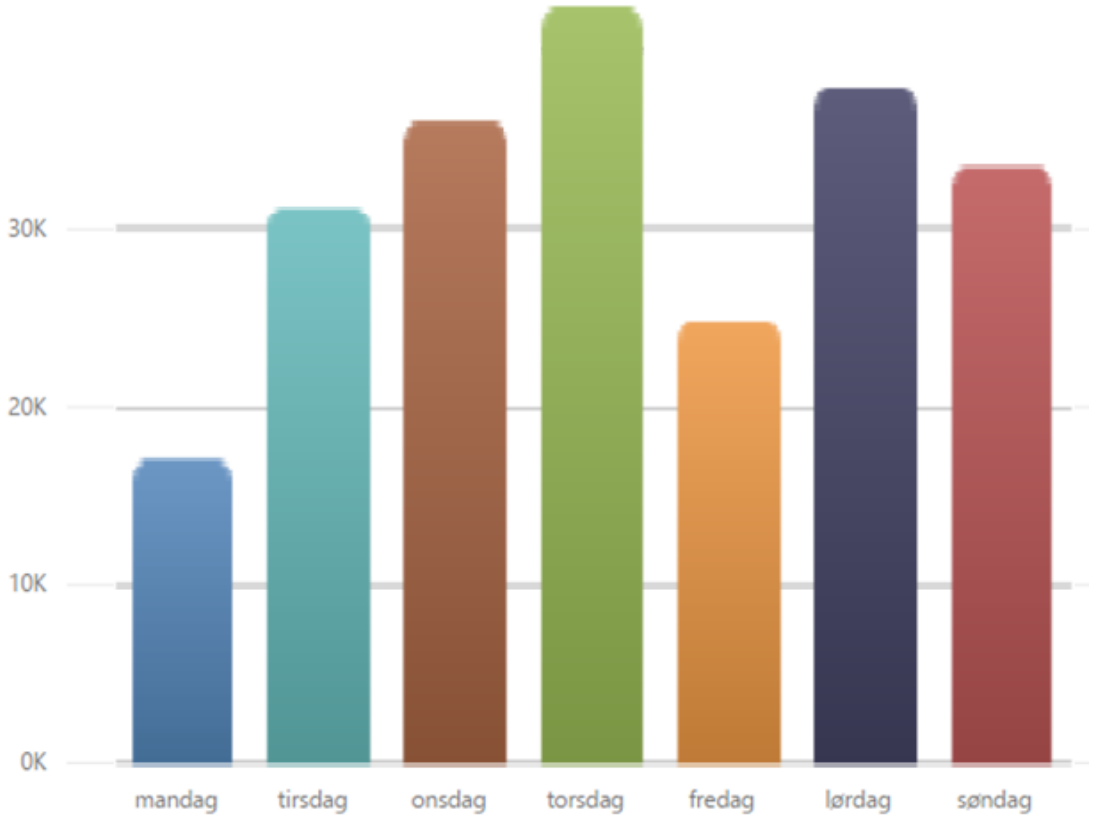
- Data collection
- Descriptive

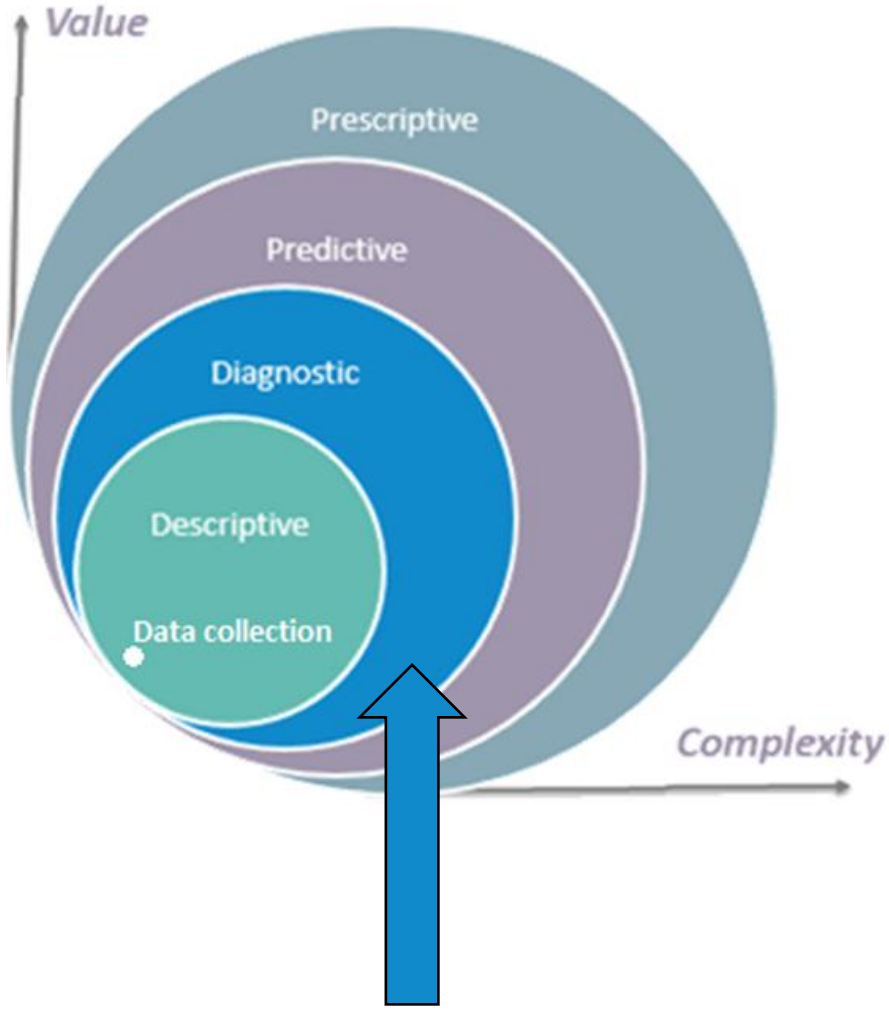




Data Science Levels

- Data collection
- Descriptive
- Diagnostic





Business Intelligence

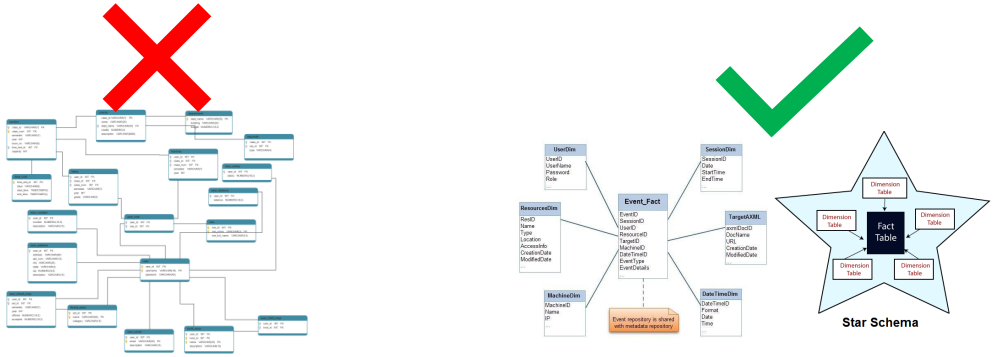
“Give the right information to the right people at the right time”

Hench provide the information to:

- Understand the business status/situation
- Make the business more **Data Driven**
- Help focus where Return On Investment (**ROI**) is highest
- Early warning if deviating from plans/budgets
- Understanding unexpected events
- Measure performance

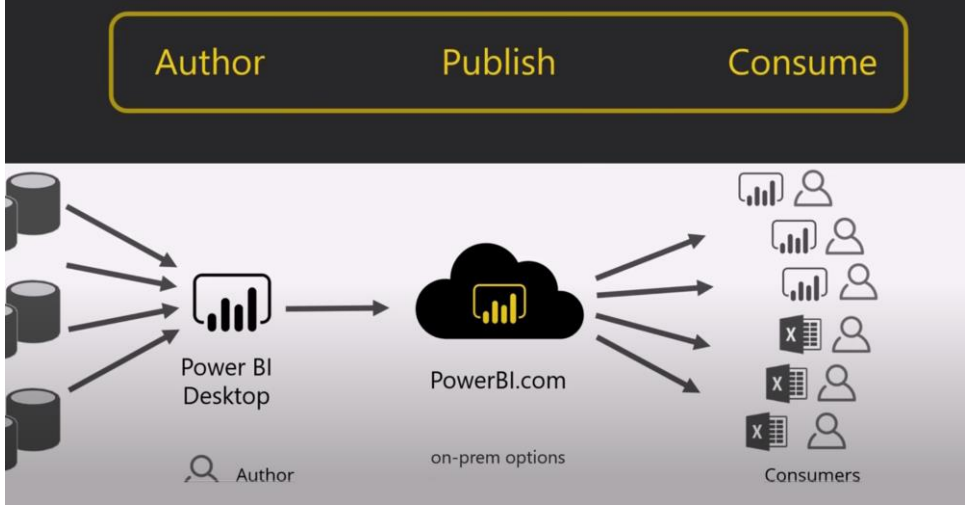
Power Bi

- Download: https://www.microsoft.com/en-us/download/details.aspx?id=58494
- Microsoft PowerBi popular in Norway because of Azure
- No data editing (unlike Excel)
- Star structure (fact and dimension tables)



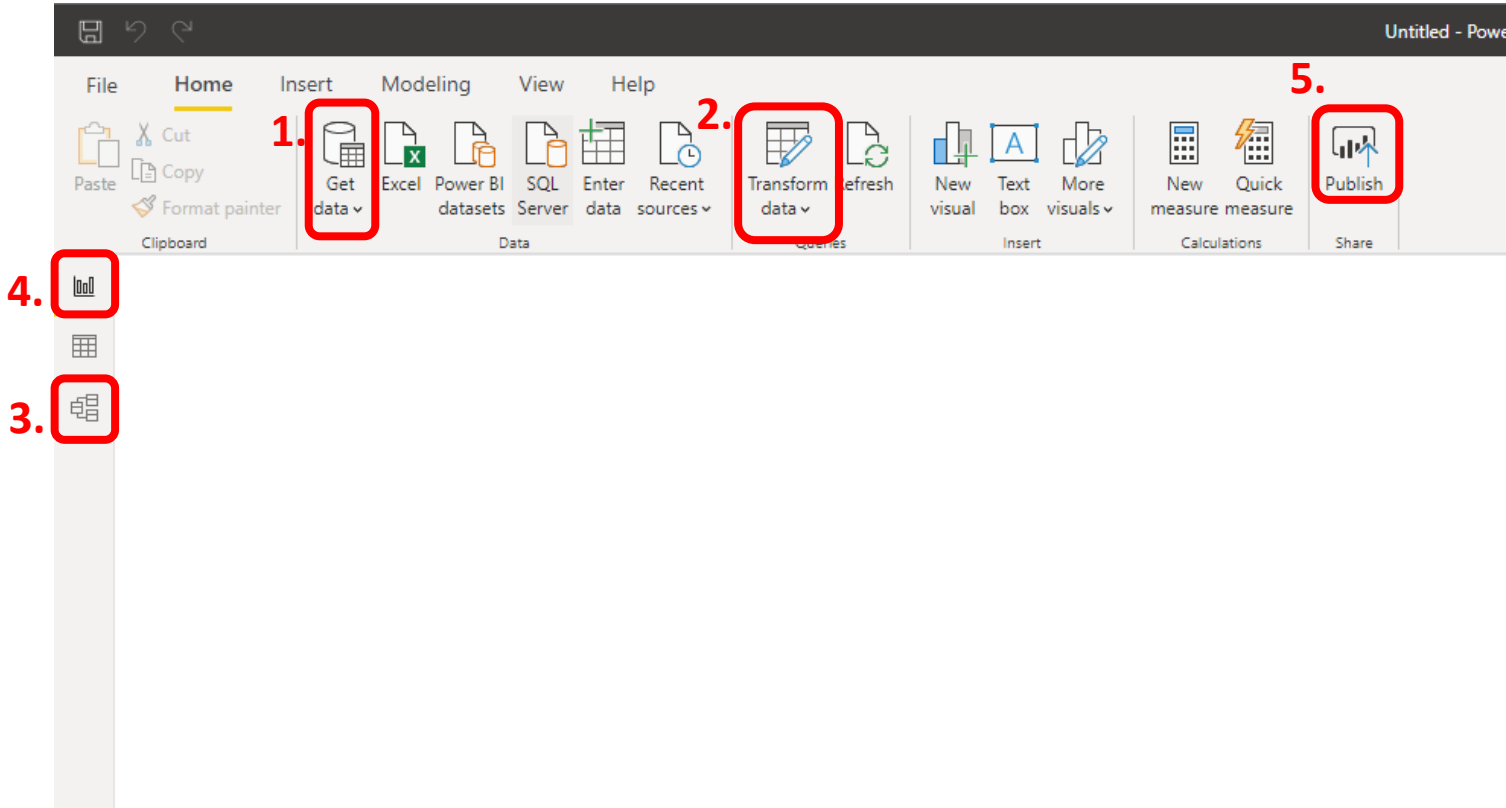
Power Bi

- App.powerbi.com (free trail?)



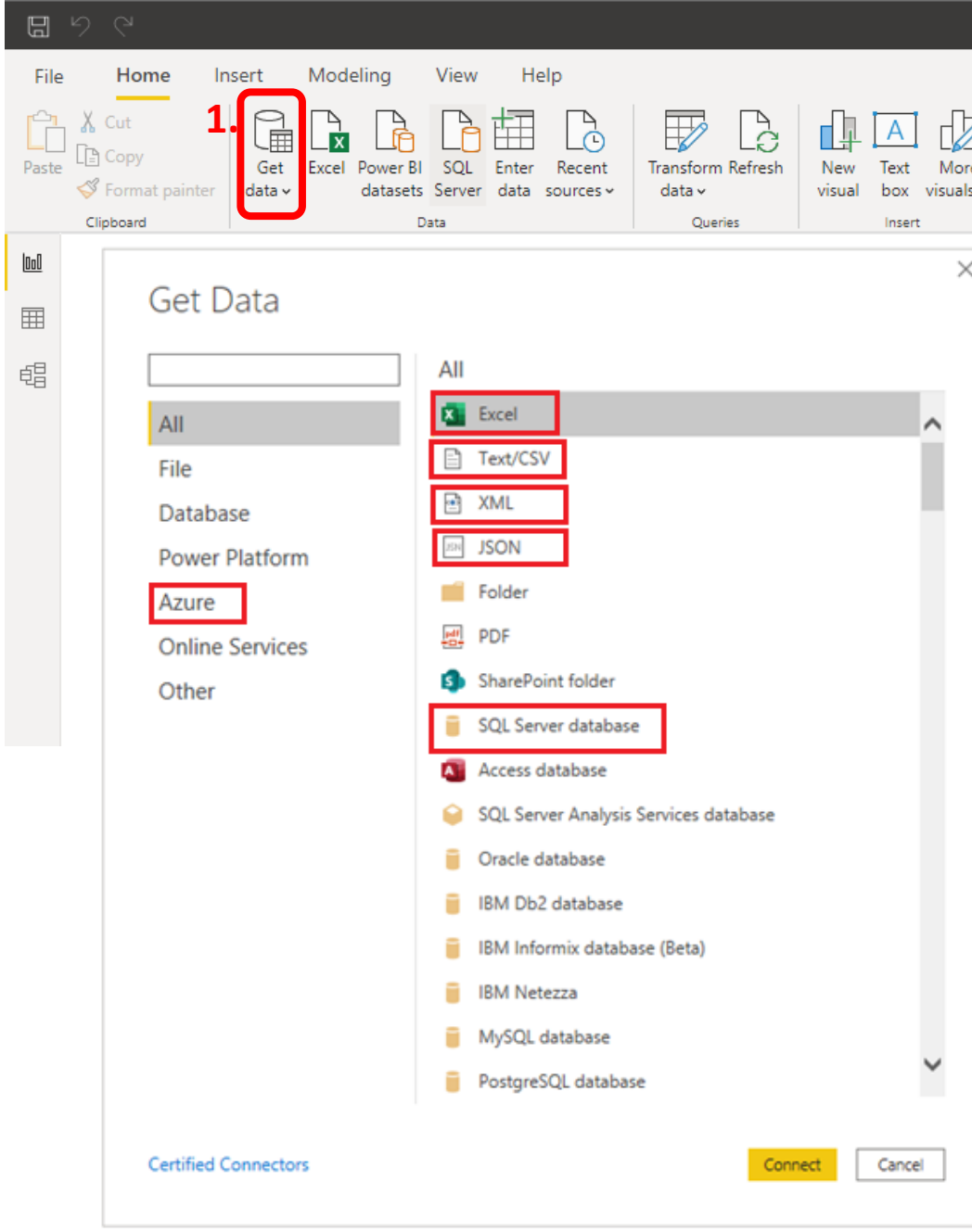
- Power BI desktop
  1. Get data
  2. Transform data (query editor)
  3. Model data, connect data relation
  4. Report, where the report is made
  5. Publish, push the report to app.powerbi.com

Power BI desktop



Power BI desktop - Get data

- Support large amount of data sources
- Can combined multiple sources
- Not there? might still work using ODBC
- For data base import there are two options
  - Import:
    - Copy all data to Power BI (hence space limited)
    - Every Power Bi features available (allow more logic to be placed in Power BI)
    - Max 8 refresh/24h, using scheduled refresh
  - Direct Query:
    - Leaves the data in the SQL database
    - “Live” connection to the database
    - Limits the available Power BI features (no data tab, no new column, etc…, requires the data on the server to be more ready for reporting)



Power BI desktop - Transform data (query editor)

- One of two Power BI engines
  - Query Editor: Filter where all data goes through
  - Model: Logic for the report (more on this later)
- Query Editor combines two “languages”
  - “SQL”- selects for data (when connected to database)
  - “Power Query M language” maybe familiar to Excel pro users?
  - Query Editor can shift from SQL to Power Query M, but not back!
- “Dataflow” in App.powerbi.com can now separate out query editor from the report

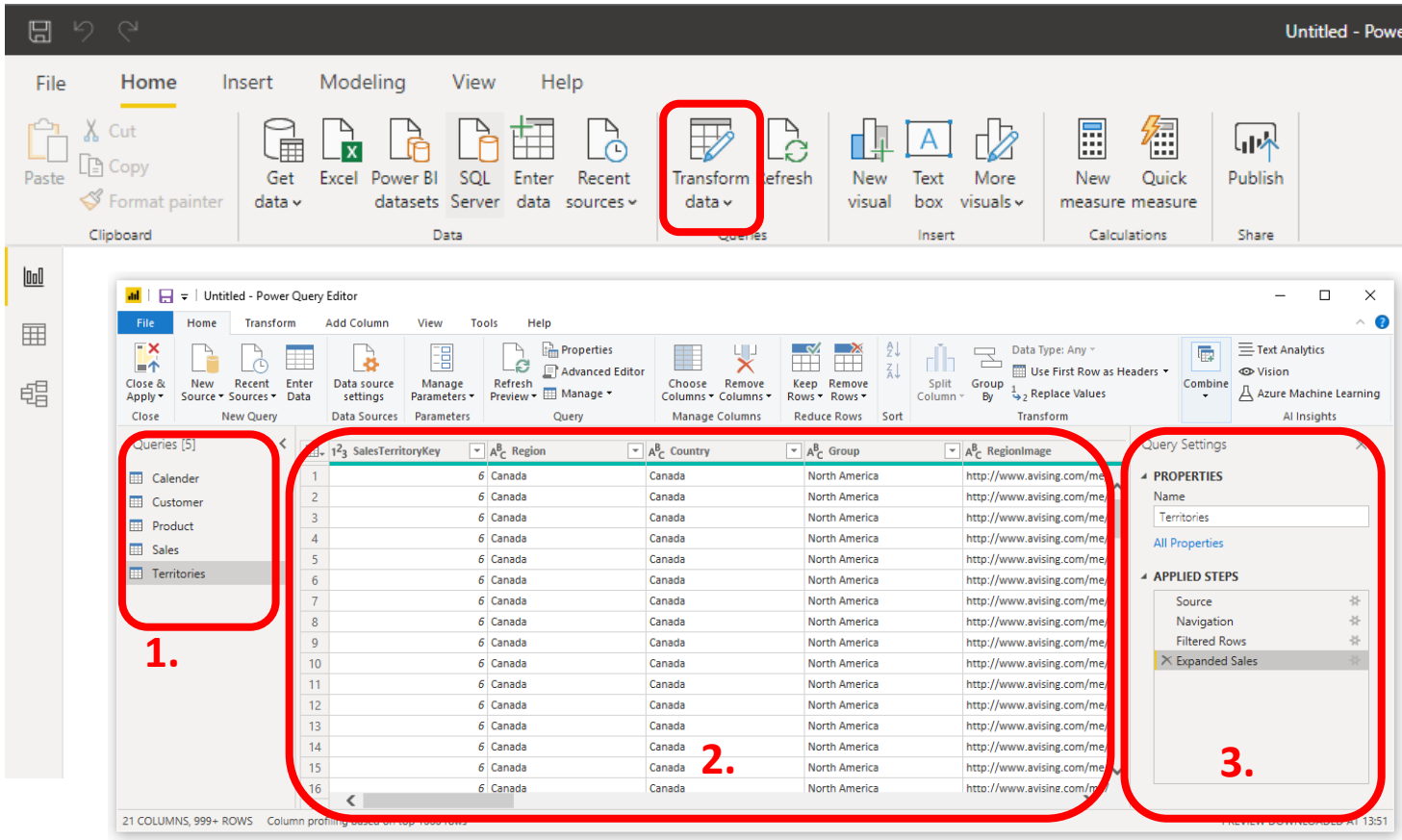
Query Editor

Model

Relationships + DAX

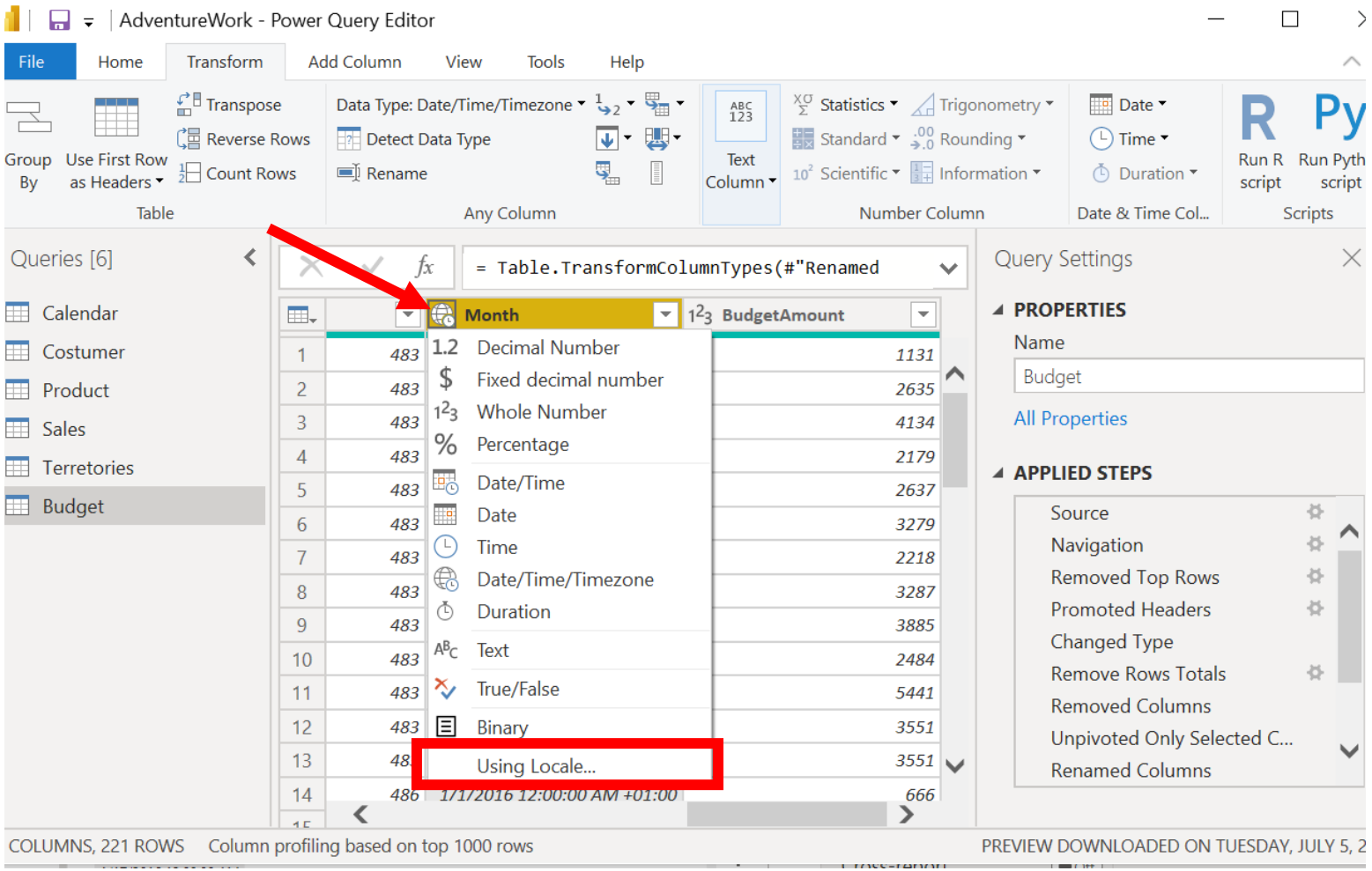
Power BI desktop - Transform data (query editor)

1. Available tables
2. Top 1000 rows of table columns
3. Applied data transformation
   - History of transformations
   - Order Dependent
   - Right klick step, View Native Query to see SQL call
   - If grayed out, using Power Query M language logic.



Change Column Interpretation Region

- How to handle multiple different datetime, currency ++ format e.g:
  - Norway
  - Sweden
  - US



Get data, SQL server

- server: My_IpAdress,49170
- username: Academy
- password: Academy123

Database:

- Academy_AdventureWorks

Show Demo

- Direct Query:
  - Show missing features
- Import (Pretend the SQL database is huge):
  - Transform data (query editor)
    - Remove columns
    - Restrict data amount by filtering for Canada
      (Decrease Fact-Table size > Decrease Dimension-Table size)
    - How SQL vs Power Query M language (order depentecy)

Demo, and try it out!

Get data, SQL server

- server: My_IpAdress,49170
- username: Academy
- password: Academy123

Database:

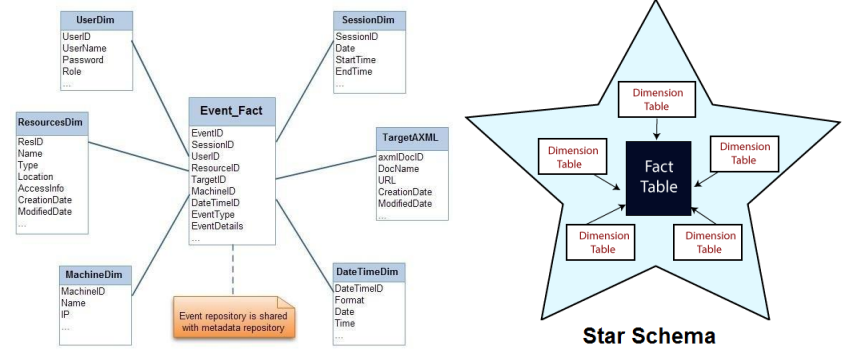
- Academy_AdventureWorks

1. Import data from SQL-server
2. Transform data
   a) Add a column for the sales table called “Customer Pay Price” which is the sum of SalesAmount and TaxAmt
   b) The company is considering buying a new more efficient sewing machine and we expect the cost of each product of category “Clothing” to fall by 2$.
   Add the new column called “SewingMachineExpectedCost” (with corresponding logic) in the Product table.
   c) Using the new SewingMachineExpectedCost column and the ListPrice, create a new column called SewingMachineExpectedNetIncome (where netincome=revenue - cost)
   d) The company is considering using cheaper steel for its bicycles, and we expect the cost of each product of category “Bikes” to fall by 1%.
   Add the new column called “BikeMetalExpectedCost” (with corresponding logic) in the Product table.
   e) Using the new BikeMetalExpectedCost column and the ListPrice, create a new column called BikeMetalExpectedNetIncome (where netincome=revenue - cost)
   f) Close & Apply

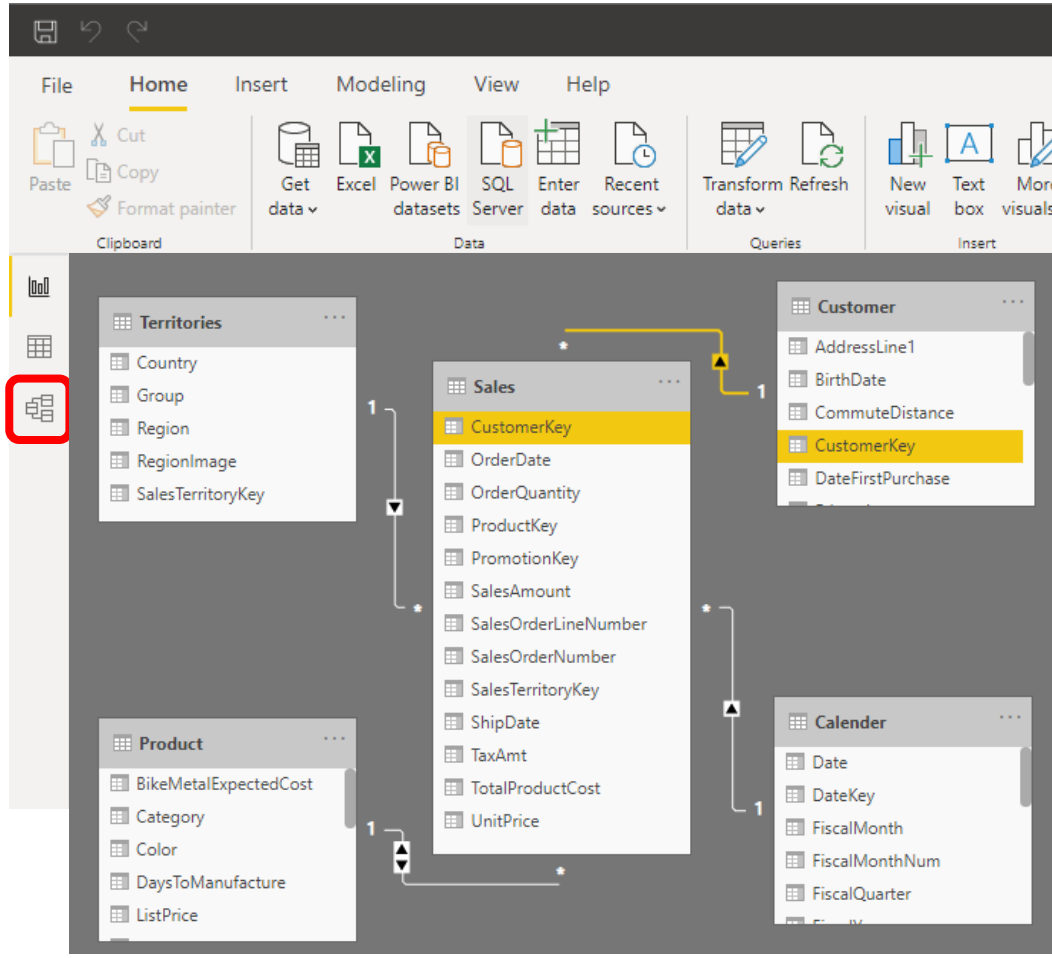
Demo, and try it out!

1. Import data from Budget.xlsx (sent to you) to the same report. Budgets are normally manually updated by people working in operation, hence messier for computers to understand.
2. Transform and clean the Budget data so its ready for use in Power BI
   a) Remove all automatically applied steps (except, Source and Navigation)
   b) Remove Top 3 rows
   c) Now the top row look like the column names. Make that change.
   d) There are multiple sub-totals in the budget, which is redundant to the data.
      Remove each row where Category contains the word “Total”. Click the down arrow next to category column and select Text filters (text filter is case sensitive).
   e) Also remove the column called “Grand Total”
   f) We don’t want one column for each month of data. Combine them all using the “Unpivot Columns” tool.
   g) Rename the new column with date information to “Month” and change Data Type to date
   h) Rename the new column with budget values to “BudgetAmount”
   i) Close & Apply

Power BI desktop - Model data, connect data relation

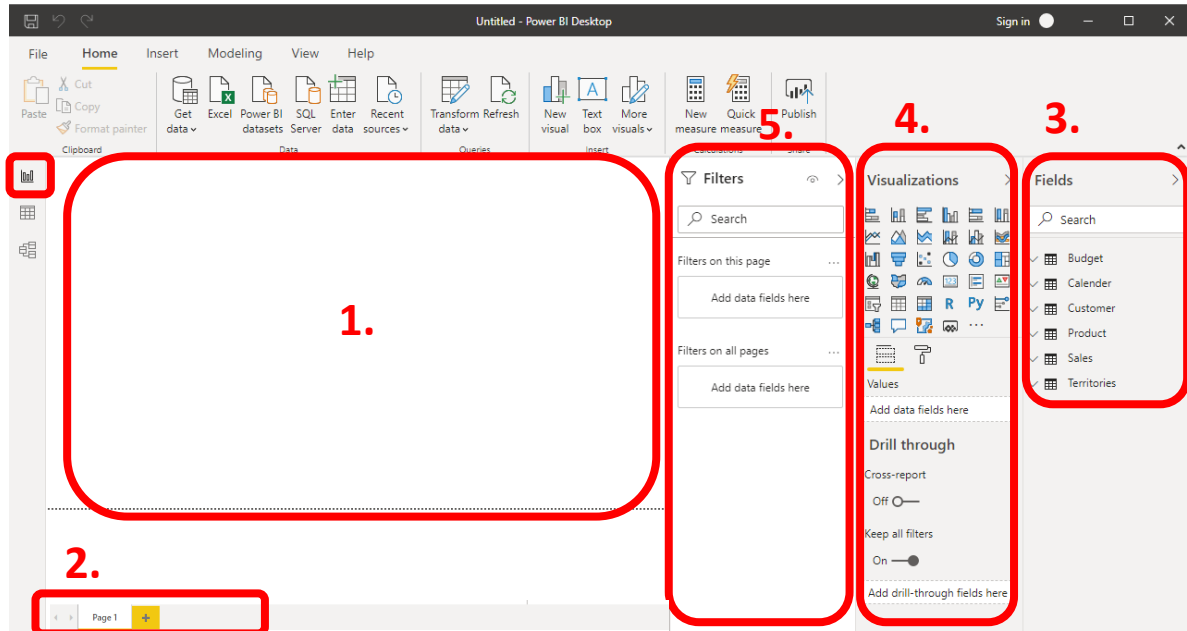


- Make relationships like SQL left-join without making flat tables which increasing data size.
- Connect keys by dragging
- Organize data in Star structures.
  - Fact-table (or Data tables) in middle of the star. Often large number records and rows.
  - Dimension-tables (or Lookup tables) edge of the star. Normally smaller tables with unique values of who, what where when, how.
- “1 > *” indicates one to many single direction relationship.
  “<>” Indicates both directional.
  (more on this later)



Power BI desktop – Report

1. The report page. This is where you build the report
2. Select and create new pages for that report
3. Table information with all columns for each table (by expanding the table)
4. Where you select and adjust the visualizations
5. With filters, you can select visualizations, page or report filters



Get data, SQL server

- server: 192.168.15.195,49170
- username: Academy
- password: Academy123

Database:

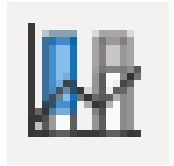
- Academy_AdventureWorks

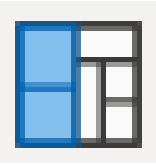
Show Small Demo

- Show clean empty page setup
- Show “Field”, “Table”, “column” image types (for measures, numbers, etc)
- Report:
  - Slicer for Date
  - Slicer Country
  - Table with target information
  - Line diagram X-axises = Time (MonthYear), Y=Revenue and legend = Country
  - The line plot x-axis order is wrong, order MonthYear by MonthYearNum
  - Show the tree plot tools

Demo, and try it out!

1. In the Model-tab,
   a) Delete all data relations (key connections)
   b) Re-combine the connections. Using Sales as Fact-table (also called data-table), and Customer, Territories, Calender (to Sales[OrderDate]) and Product as Dimention-table (also called lookup-table).
2. Report:
   a) Add slicer for Customer FullName
   b) Add slicer for Product ProductName
   c) Add table with single column, Customer FullName
   d) Add table with single column, Product ProductName
   e) Add a new page and call it graphs (the rest of the tasks should be done on this page)
   f) Add a Year slicer
   g) Add a Country slicer, and remover Country NA
   h) Add a table which shows multiple columns of some self picked information (copy the table and place it on the first page too)
   i) Create a “Line and stacked column chart” with Month as x-axis (written month name, so remember to order it), Sales amount y-axis country as legend, TaxAmt as tooltips and number of sales as line values (number of sales can be a little tricky, can you find a way?)
   j) Make a Treemap, using columns:
   Category, SubCategory, OrderQuantity, SalesAmount
   for the visualizations fields:
   Group, Details, Values and Tooltips
   correspondingly.
   k) If you have more time, go through the plot tools, and change the color of the plot classes, line, turn on data labels, create a better title, etc…

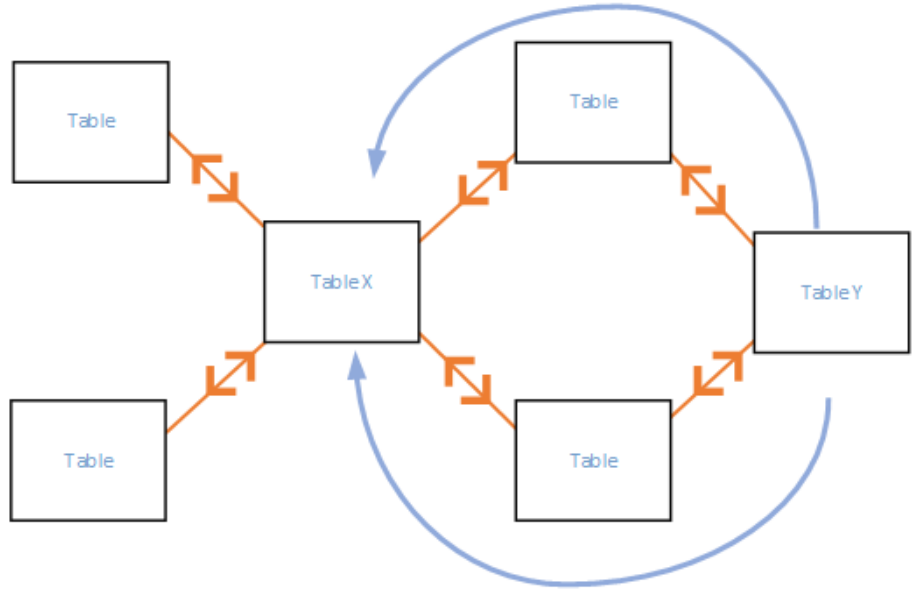


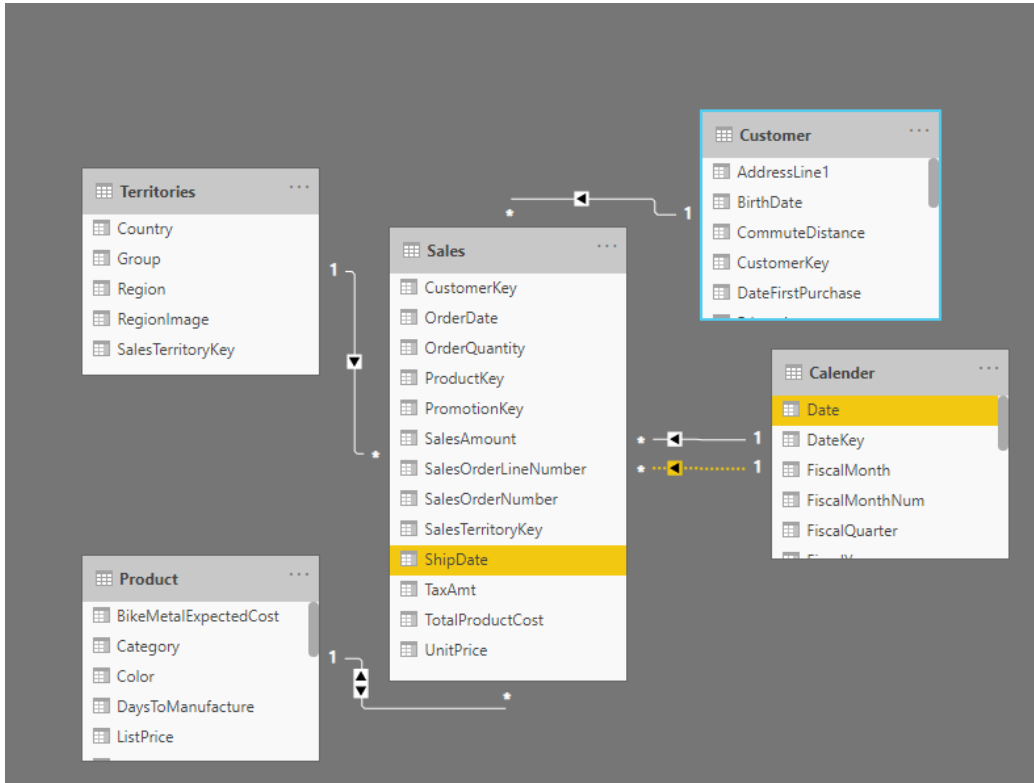


https://docs.microsoft.com/en-us/power-bi/transform-model/desktop-create-and-manage-relationships

Power BI desktop - Model data, connect data relation

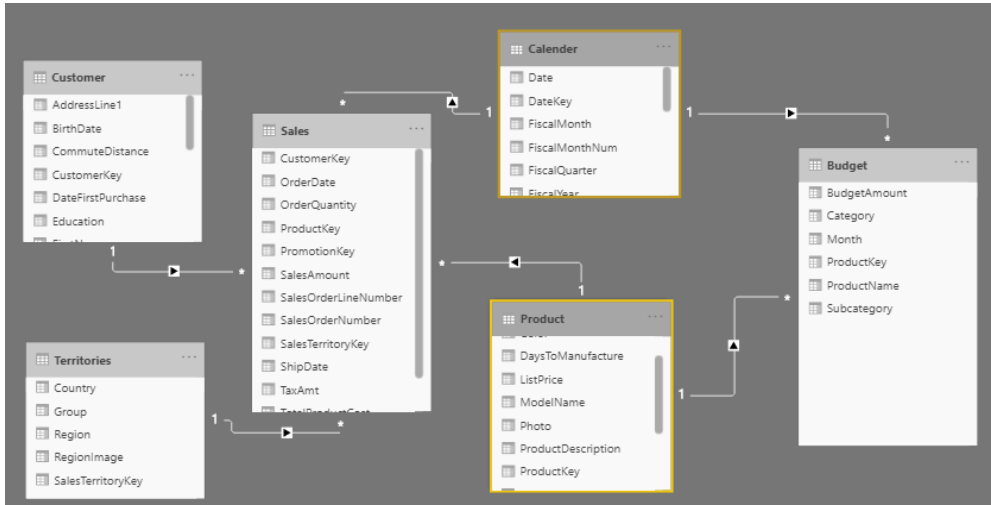
- From the first page of your demo, select a name, or select a product. Why is not the other tables filtered by this selections?
- Fixed by changing Customer and Product table connection both direction (1<>*)
- Both direction are acceptable for star structure
- Can cause problem when connecting extra tables (as Table Y in figure1, no longer star structure), because connection loops create ambiguous set of relationships
- Inactive relationships (dashed) can happen when multiple relations are establishd between two tables (see Calender Sales relationship).
  Only solid line used for filtering, while dashed can be used with DAX expression.





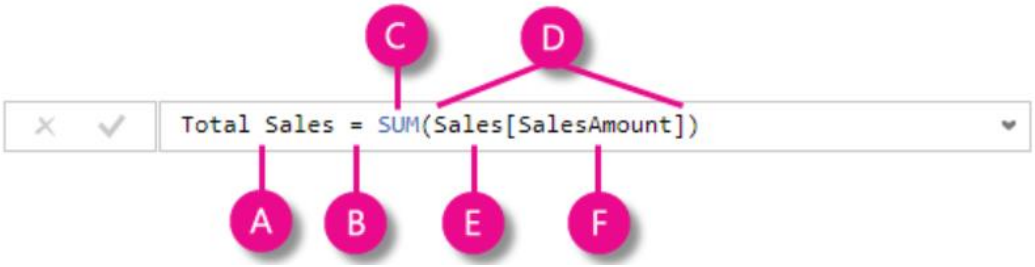
Demo, and try it out!

- On you Power BI first page, try selecting some customer and products and see how it affects the tables
- Add a connection between Sales[ShipDate] and Calender[Date], what effect does this make?
- Make all connections both directional and do the same kind of selections and see the effect.
- Add connections between Budget and Product (using ProductKey) and Budget, and Calender using Budget[Month] and Calender[Date]
  Get any problems?
- Any ideas how we can achieve both directions behaviours between Budget and Sales? If you have time try implementing your workaround.
- Finally set all connections to single directions (as seen in the photo, alternatively leave Territories and Customers both directions)
- Notice, only Calender and Product can filter Sales and Budget
  This is no longer Star Structure



Power BI – Data Analysis Expressions (DAX)

- DAX is a formula and query language used for extracting additional information from data already in your model
- Like formulas in Excel
- Used to create Measures and Calculated columns
- Syntax (use “” if table name is unusual characters)
- Example:
  Create plot with automatic SUM, and a Measure that gives the same result



More advance:

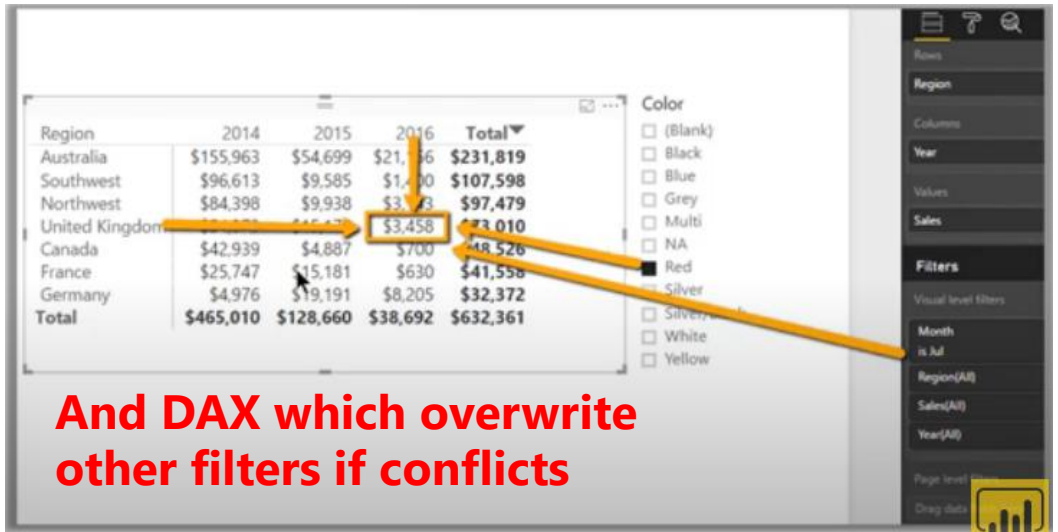
```
Previous Quarter Sales = CALCULATE(SUM(Sales[SalesAmount]), PREVIOUSQUARTER(Calender[Date]))
Where,
CALCULATE(*calculation*, *filter*) (one of the most used DAX functions)
PREVIOUSQUARTER(*range of dates*)
```

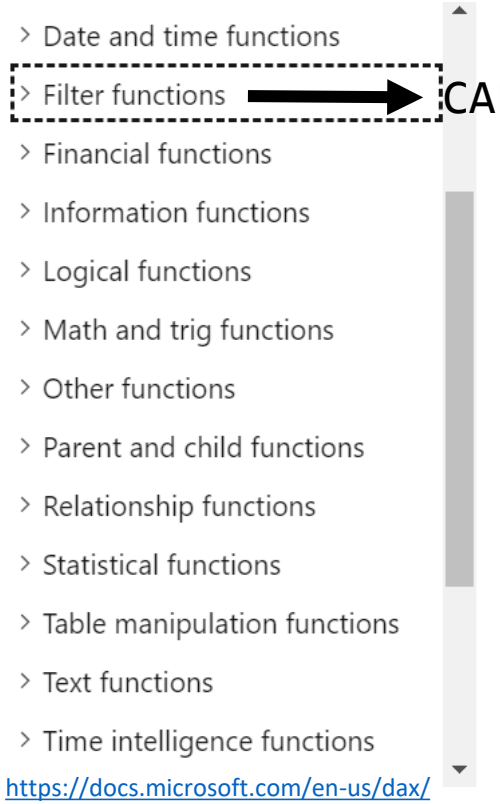
https://docs.microsoft.com/en-us/power-bi/transform-model/desktop-quickstart-learn-dax-basics

Power BI – Data Analysis Expressions (DAX)

- Functions
  - List of all built in DAX functions can be found in the “DAX-BIBLE”: https://docs.microsoft.com/en-us/dax/
  - Function references a complete column, a table or another measure (can be specific rows if filters are used)
  - DAX includes functions that return a table rather than a value (used as input to other functions returning a value)
  - DAX includes time intelligence functions (like PREVIOUSQUARTER)

- Contex, two types:
  1. Row context, X-function like SUMX, row by row (other DAX functions whole columns)
     Total Sales SUMX = SUMX(Sales, Sales[Qty] * Sales[Price Per Unit])
  2. Filter context, see photo





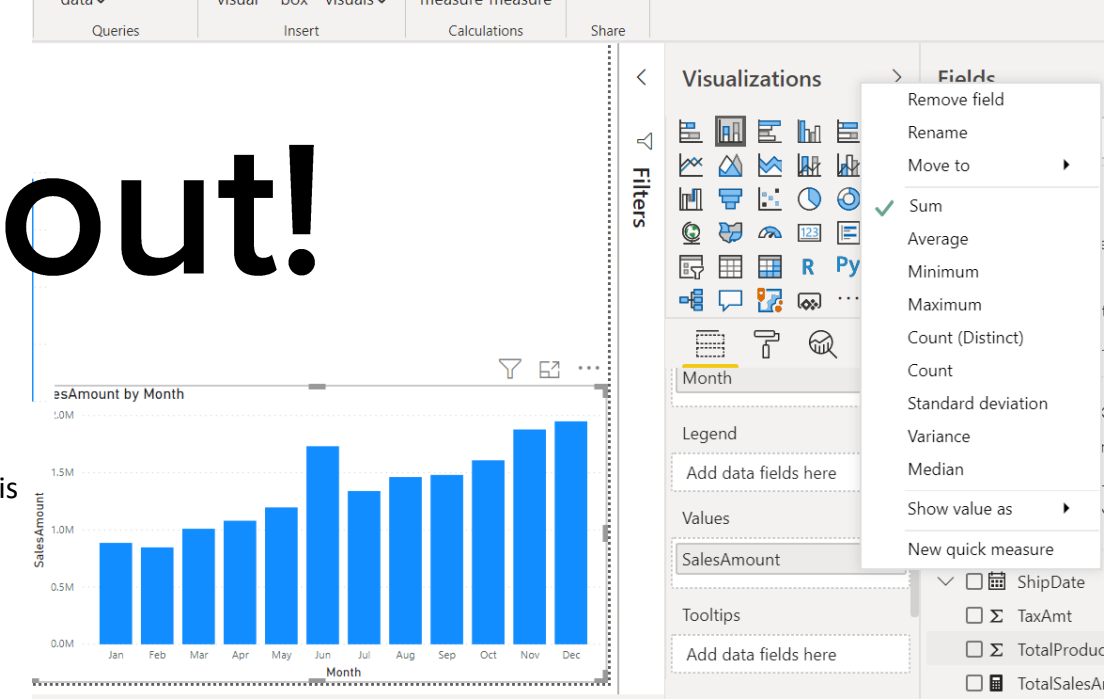
CALCULATE

https://docs.microsoft.com/en-us/dax/

Query Editor

Model
Relationships + DAX

Demo, and try it out!



Look up the DAX function documentation online when solving this
Questions:

1. Create a new page and set a year slicer (type list). Make a Column plot with Month on X-axis and SalseAmount on Y-axis. Make sure SalesAmount is set to “SUM”
2. Create a new Column plot at the same page, also using Month on X-axis, and try to make a measure that gives the same plot as above hint: SUM DAX function.
3. Change the first plot from SUM to AVERAGE and make a new measure for the second plot such that it replicate the first plot.
4. Repeat question 3 for min, max, count and count (Distinct)
5. Using the CALCULATE function, make two measure and calculate the Sum of SalesAmount for Males and Females using the Customer table. Self test that it works and gives the correct result (can test this using Card and slicers)
6. Do the same as in 5 but use the REMOVEFILTERS function to let the measure ignore the year selection.
7. Create a function that calculate % of sales of a chosen Product[category] (by slicer or graph) vs all categories using the functions DIVIDE, SUM, CALCULATE and REMOVEFILTERS.
   For example for the Product[Category] Bieks the measure should output the number of [Total Sales of Bikes]/[Total Sales]
8. Make a measure that uses the IF function to check if the SUM of SalesAmount is larger than 5million. If true, reurn “High”, else return “Low”
9. Meake a measure called ColSalseAmount that calculate the SUM of Sales[OrderQuantity] and multiply it (by using *) by the SUM of Sales[UnitPrice].
   Create a new measure called RowSalesAmount that calculate OrderQuantity multiply by UnitPrice rowbased using SUMX. Which of this two methods will be close to equal to the sum of SalesAmount, and why is the other measure wrong??
10. Add a country slicer.
    Make a measure called CanadaSalesAmount using the CALCULATE function to get the SUM of SalesAmount for Canada (using filter Territories[Country]=“Canada”)
    Make another measure called FilterCanadaSalesAmount using the CALCULATE function to get the SUM of SalesAmount for Canada, using the FILTER function on the CALCULATE filter cell.
    Change each sliser and see how the two measure behave, any differences?

Learn these first:

SUM, AVERAGE, MIN, MAX

COUNT, COUNTROWS

CALCULATE

FILTER

IF

Demo, and try it out!

Get data, SQL server

- server: My_IpAdress,49170
- username: Academy
- password: Academy123

Database:

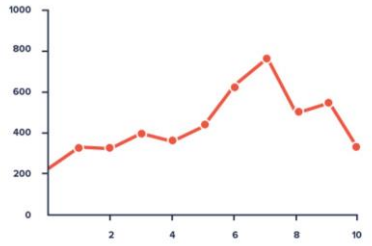
- Academy_AdventureWorks

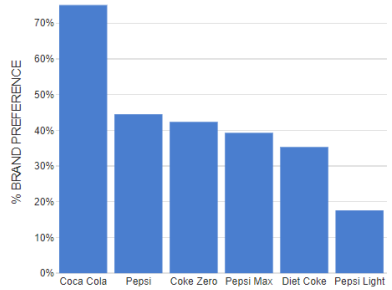
Questions:

1. Two and two together
2. Create self define reports but must include one of the following subjects:
   a) Budget analysis, how well did we follow the budgets
   b) Sewing machine vs new bike metal comparison
   look back in time and see how the revenue/cost/net-income would have been affected if this deals was there from the start. How lucrative is each of the deals? At what price would you invest in a new Sewing machine?
   c) Customer profiles (who is our customers, and which of those groups have been the most important until now)
   d) Territory analysis (What is the territory trends, any upcoming countries, where should future recourses be focused?
   e) Product analysis.
   What product makes most money, which product should we focus on improving, should we cancel any? What product Categories are important and how should that affect the company?
3. Try to make business relevant plot. For this company what is the most valuable information you can give them? And how can you present it? Value is often related to Cost, Revenue, efficiency, customers satisfaction,
4. Google to get inspiration
5. Must use some measures with DAX syntax
6. On Friday, or at the end of the day, Each group show their reports.

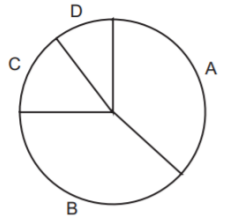
Practical Uses of Data visualization

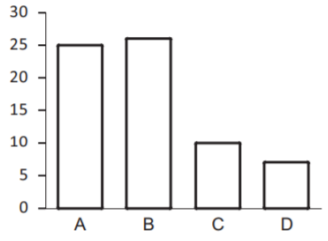
- Stephen Few – “The King” of practical data visualization
- Lots of example but one:
  Space Shuttle Challenger disaster was predicted, bu not taken seriously
- Important that data and plots show correct data!

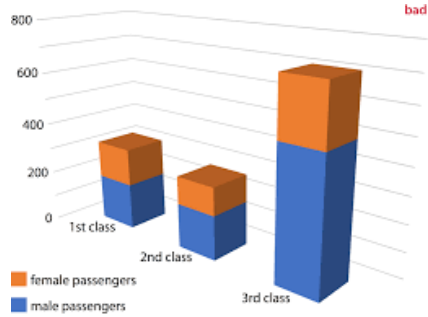




- Line plot – for Continues x-axis (like time)
- Column plot – for Discreet x-axis (categories)
- Popular Saying:
  Don’t use pie chart except when showing how mucl of a pie has been eaten
- Too fancy (example 3d) can make it harder to read







Sources and more information:

- https://www.youtube.com/watch?v=AGrl-H87pRU
- https://docs.microsoft.com/en-us/power-bi/transform-model/desktop-create-and-manage-relationships