

LESSON HANDOUT

Correlation Plots and Scatter Charts

Why:

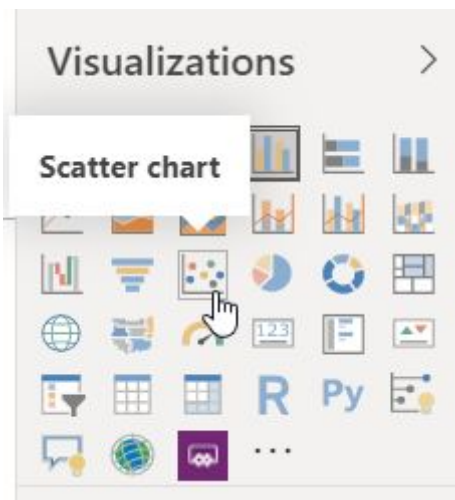
Correlation indicates the direction and strength to which two variables are related to one another and can set the condition for further exploratory analysis.

Create a Scatter Plot in Power BI:

Why: A scatter chart is used to plot the relationship between two variables.

How:

- **Method 1:** Select two numerical fields and then click on 'Scatter Chart'



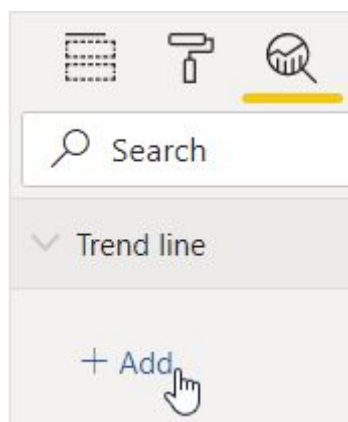
- **Method 2:** Select the Scatter chart, then click and drag from the fields list into your X and Y axes

Add a trend line:

Why: Trend lines make it easy to visualise a positive or negative relationship between the two values.

How:

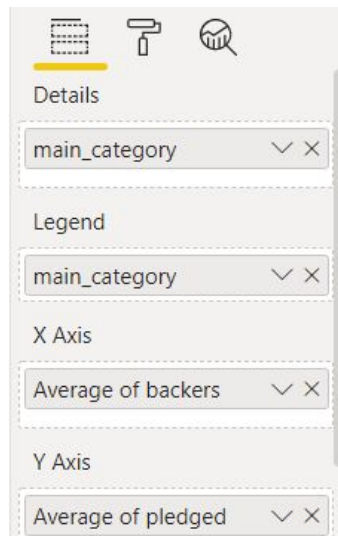
- Analytics > Trend Line > Add



Aggregate:

How:

- You can aggregate this chart by bring a third field in the 'Details' section, and then summarising your X and Y axes by Average. When doing this it's often helpful to add a Legend, so choose the same field and drag it into the Legend field as well. This will make it easier to read.

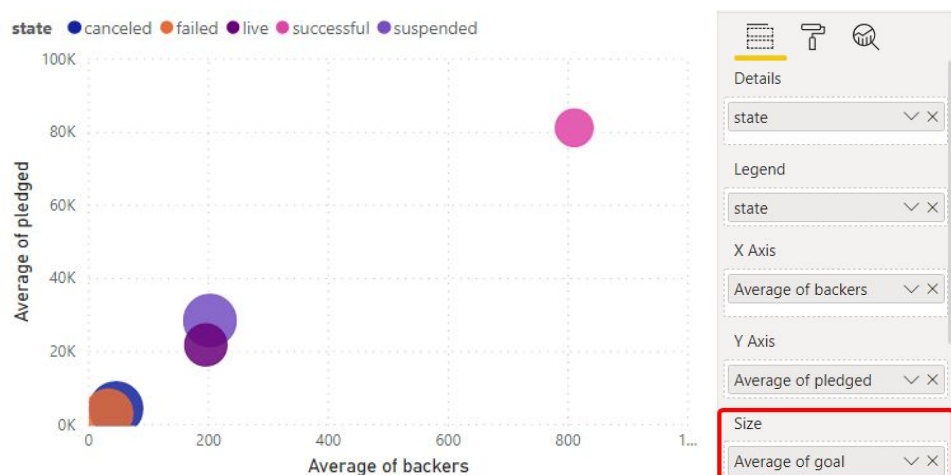


Bubble Plot:

Why: A bubble plot is a way of showing a relationship between 3 variables, by adding the size of data points.

How:

- Create a scatter chart as normal, then click and drag your 3rd variable from the fields list into 'Size', which will give you a chart like the example below.



Adding a Play Axis:


Why: Visualise changes over time

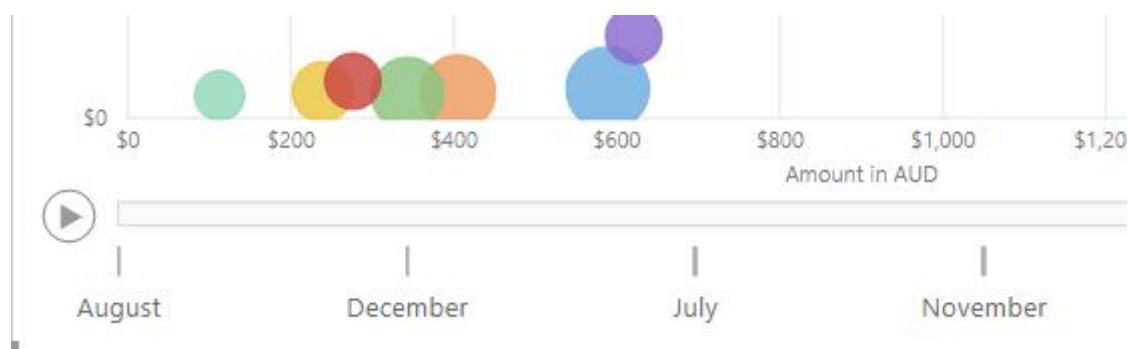
How:

- Use the 'Play Axis' field to bring a date field into your visualisation (note: depending on your data, it's usually better to bring in Month or Year hierarchy).

The configuration panel shows the following settings:

- Details:** Add data fields here
- Legend:** Category
- X Axis:** Amount in AUD
- Y Axis:** Average of Amount in A
- Size:** Count of Amount in AU
- Play Axis:** MonthName (highlighted with a red box)

- Click the  button to see your chart dynamically update, or click on various points along the date line to update.

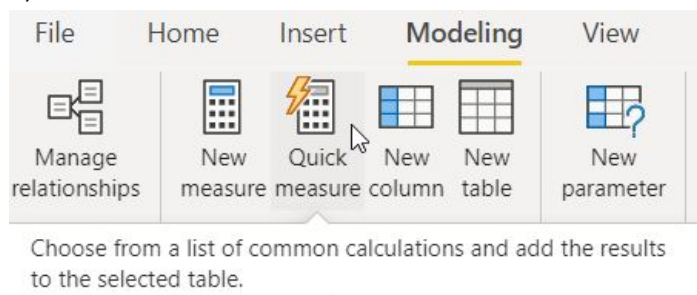


Create the Correlation Coefficient:

Why: Measures the strength of the relationship between data

How:

- Create new quick measure - In the ribbon, on the Modeling tab under Calculations, click 'Quick Measure'



- Select the calculation as 'Correlation Coefficient'
- Click and drag from the Fields list to choose your Category, Measure X and Measure Y

Calculation

Correlation coefficient ▼

Calculate the correlation coefficient between two values over a category. Originally suggested by Daniil Maslyuk in the quick measures gallery. [Learn more](#)

Category ⓘ

name

Measure X ⓘ

Sum of backers

Measure Y ⓘ

Sum of pledged

Fields

Search

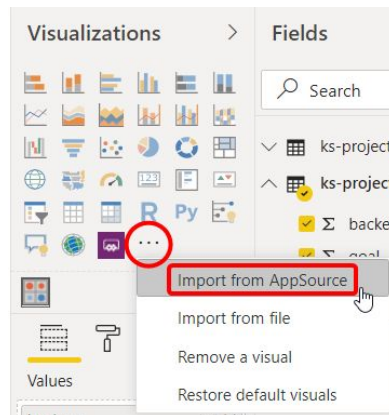
- ks-projects-info
 - category
 - deadline
 - ID
 - launched
 - main_category
 - name
- ks-projects_financials
 - backers
 - goal
 - ID
 - pledged
 - state

Create a Correlation Plot (Custom Visual):

Why: Allows quick visualisation of the correlation between values

How:

- In the Visualisation tab, click on the 3 dots (...) and choose 'Import from AppSource' (note: You'll need to be signed into Power BI for access)



- Within the marketplace, search for 'Correlation', then click 'Add' on the Correlation Plot visual.

Power BI Visuals

MARKETPLACE | MY ORGANIZATION

Add-ins may access personal and document information. By using an add-in, you agree to its Permissions, License Terms and Privacy Policy.

correlation

Suggested for you

Category

All



Correlation plot

Highlight the most correlated variables in a data table

★★★★☆

Add

- Your new visualisation will be added to your list for use. Click on this visualisation to create it, then add your required fields. (Note: you'll need to have R setup in Power BI, which was covered in the previous lesson). You'll end up with something like this:

