

What is Market Basket Analysis

Market Basket Analysis is a data mining technique used to discover relationships between items purchased together in transactions.

It answers this question:

‘When customers buy product A, what else are they likely to buy?’

It is based on association rule learning.

Why is it called ‘Market Basket ‘?

Imagine a customer’s shopping basket: Bread, milk, eggs. This single basket is a transaction; now imagine thousands of baskets. Market Basket Analysis (MBA) looks across all baskets and finds patterns like:

- Bread → Butter
- Beer → Chips
- Cereal → Milk

These patterns are called association rules.

The data required for an MBA should be transaction-level data.

With each row = one item in a transaction.

Understanding the Core Metrics of MBA

Support

Support measures how frequently an itemset appears in the dataset,

For a rule: $A \rightarrow B$

Support measures how often A and B occur together.

Confidence

Confidence measures how often an item appears in a transaction that already contains another item.

It answers the question ‘if a customer buys A, how likely are they to also buy B’.

Lift

Lift measures how much more often item A and item B occur together than expected if they were independent.

Summary

Support → how common

Confidence → How reliable

Lift → How meaningful.

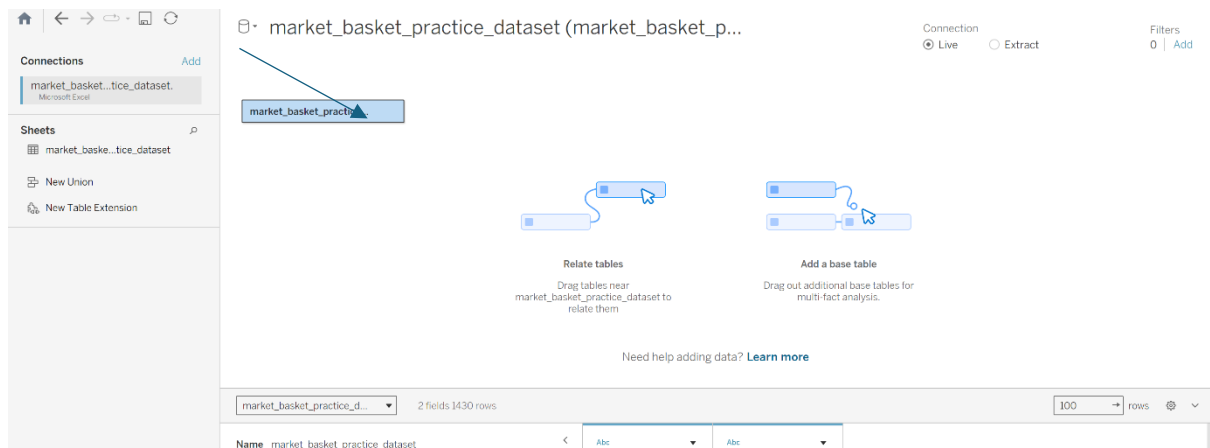
Building Market Basket Analysis in Tableau.

Step 1 – Connect Data in Tableau

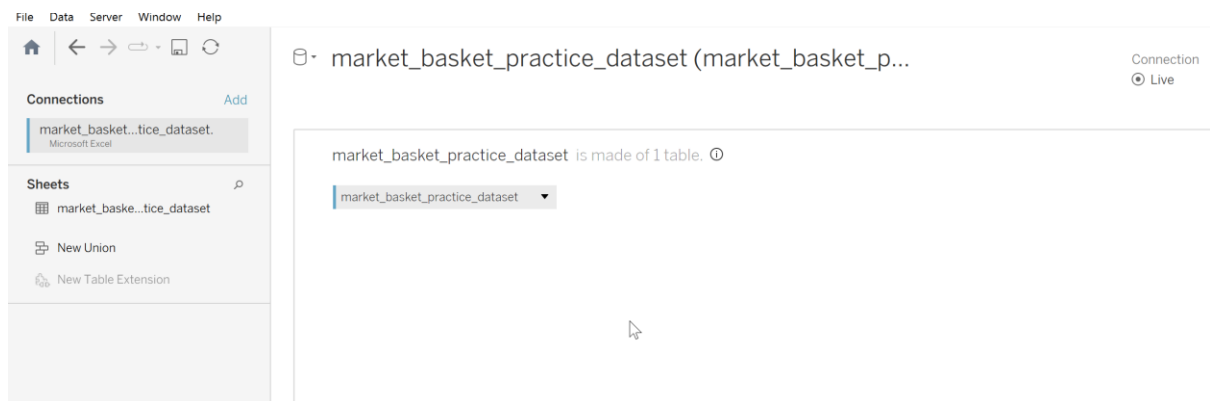
- Open Tableau
- Click Connect to connect to the file (Microsoft Excel)
- Select your data (market_basket_practice_dataset)
- It takes you to the data source tab,
You should see:
transactions ID
Item.

Create a self-join.

Double click on the blue containing the market_basket_practice_dataset



A new canvas will open



Double-click on the same data set under sheets.

Click the Inner join under data source, add Transaction ID, leave this sign (=) under Market_basket_practice enter transaction id.

Add a new join clause and enter the word item on both sides, but change the sign to <.

Close it to return to the data source page.

The screenshot shows the Tableau Desktop interface. On the left, the 'Connections' pane lists 'market_basket_practice_dataset' and 'market_basket_practice_dataset1'. The 'Sheets' pane shows a new sheet named 'market_basket_practice_dataset'. The main view displays a join configuration window for 'market_basket_practice_dataset' and 'market_basket_practice_dataset1'. The join is configured as an 'Inner' join on the 'Transaction ID' field. Below the join configuration, a data preview table is shown with 4 fields and 3209 rows. The table has columns for 'Transaction ID', 'Item', 'TransactionID (marketbasketpracticedataset1)', and 'Item (marketbasketpracticedataset1)'. The data shows transactions with items like Beer, Chips, Salsa, Bread, Butter, Yogurt, Milk, and Salsa.

Transaction ID	Item	TransactionID (marketbasketpracticedataset1)	Item (marketbasketpracticedataset1)
T001	Beer	T001	Chips
T001	Beer	T001	Salsa
T001	Beer	T001	Bread
T001	Beer	T001	Butter
T001	Beer	T001	Yogurt
T001	Beer	T001	Milk
T001	Chips	T001	Salsa

Now you will have

- Table 1
- Table 2

Rename

Item → Item A

Item (table 2) → Item B

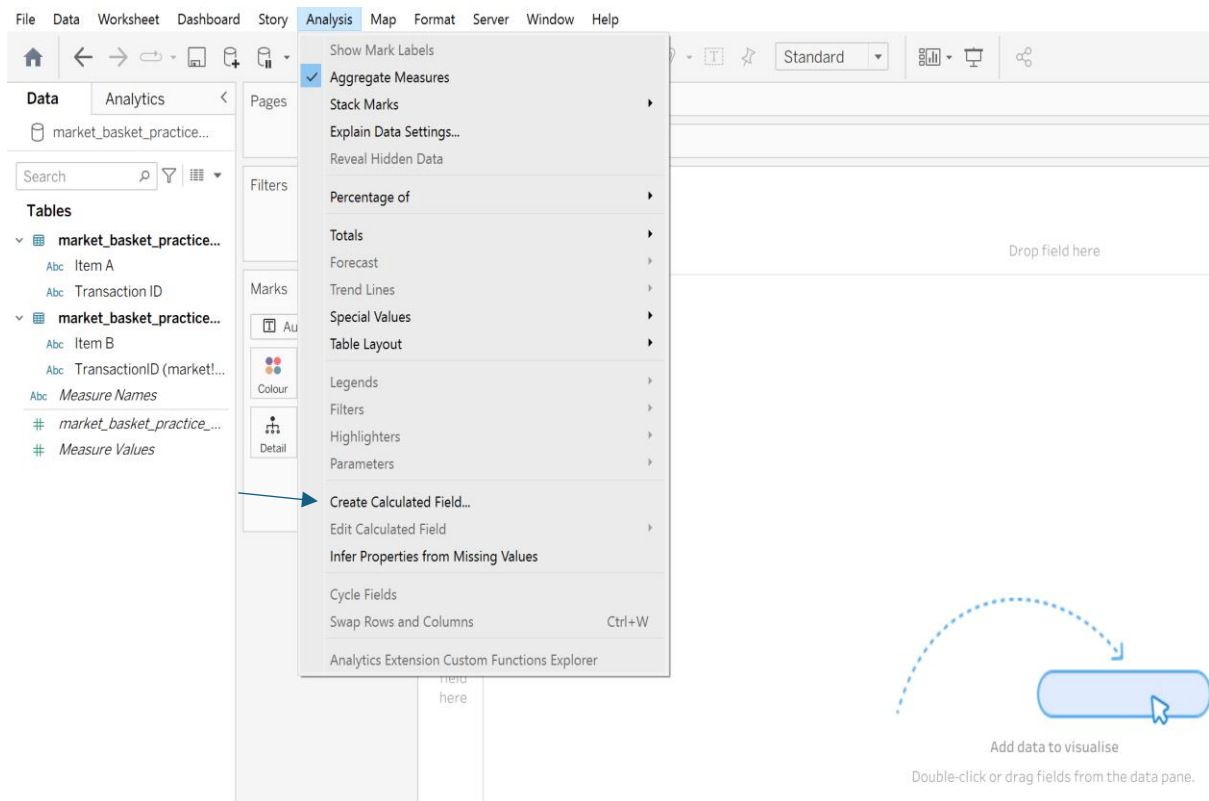
This avoids confusion later

Click on Sheet One to start exploring the data.

First, we will need to create all the calculated fields.

Total Transactions

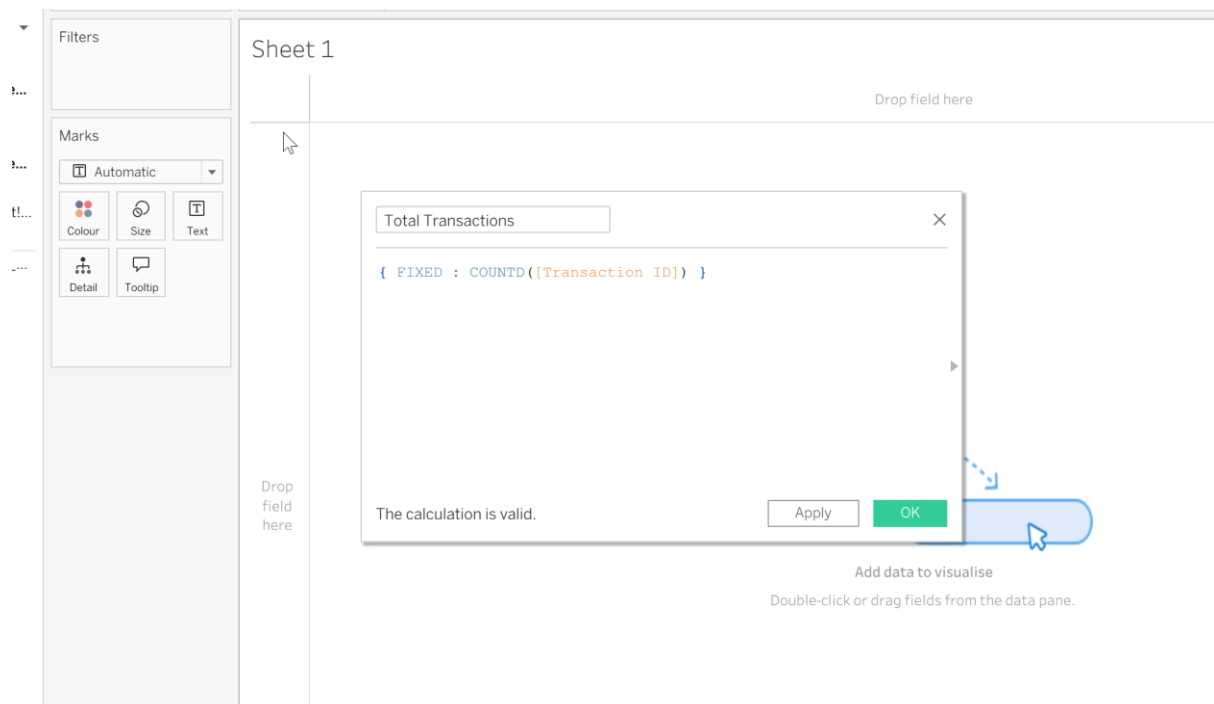
Click on the analysis tab and click on create calculated field



Change calculations to total transactions, then enter this

formula

{ FIXED : COUNTD([Transaction ID]) }

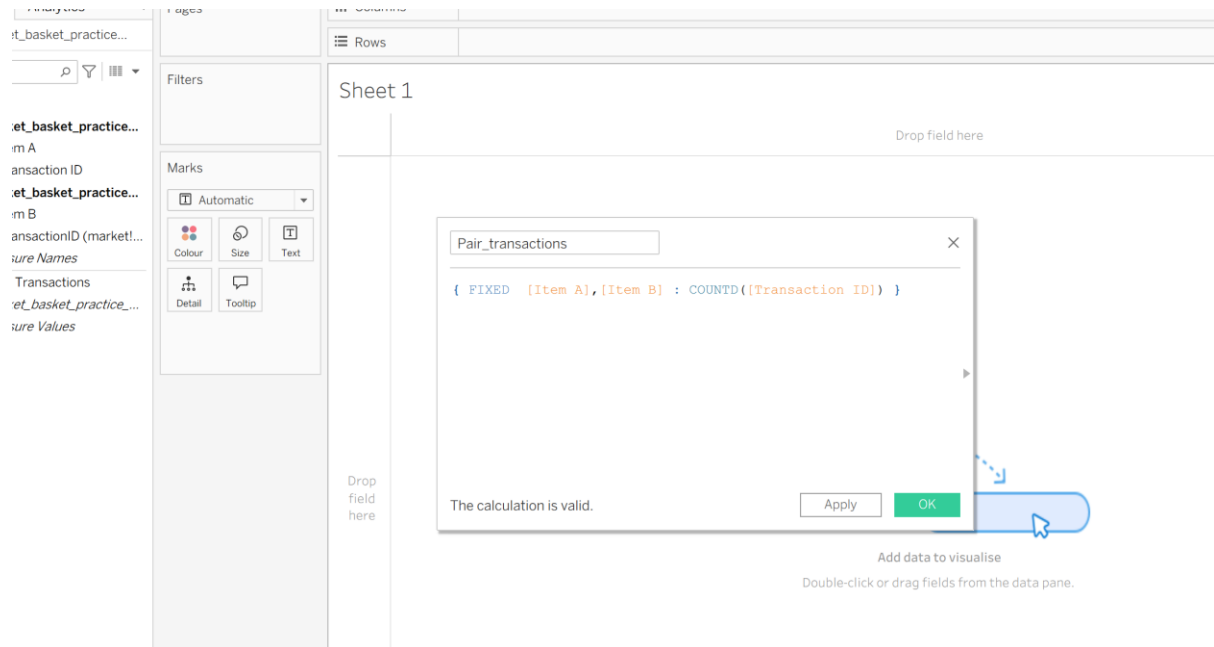


click ok to finish.

Create a calculation field for Pair Transactions

Formula

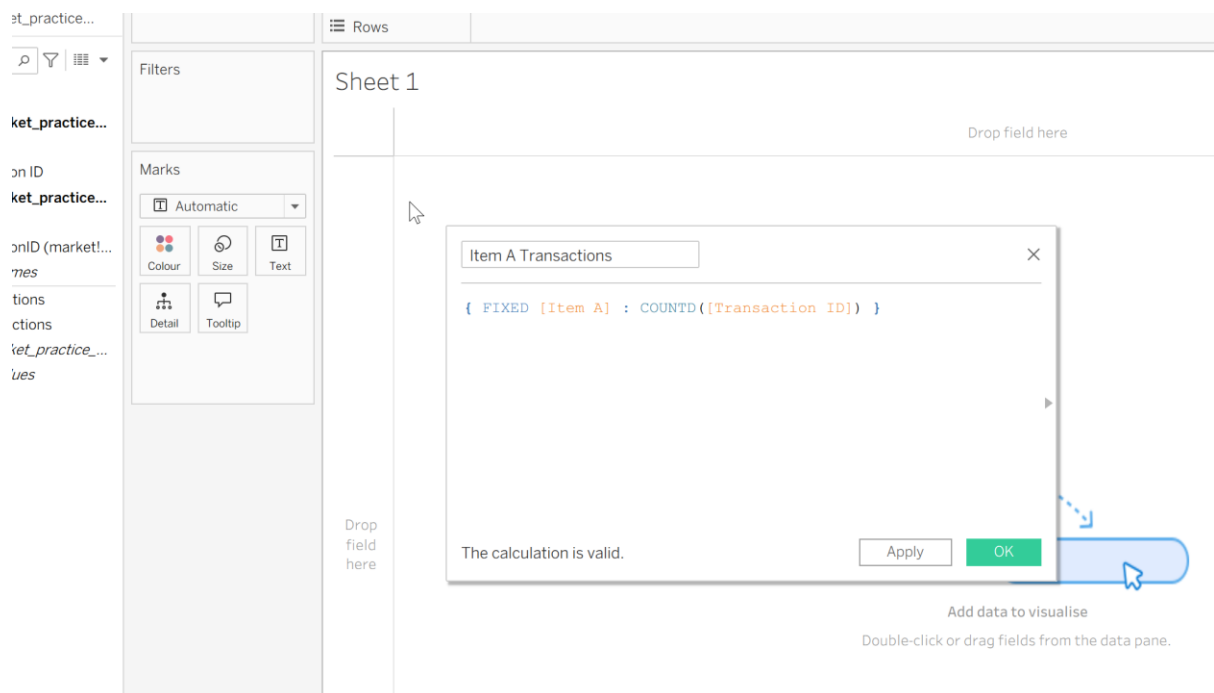
{ FIXED [Item A], [Item B] : COUNTD([Transaction ID]) }



Item A Transactions

Formula

{ FIXED [Item A] : COUNTD([Transactions ID]) }



Item B Transactions

Formula

{ FIXED [Item B] : COUNTD([Transaction ID]) }

ers

arks

Automatic

Color

Size

Text

Detail

Tooltip

Sheet 1

Drop field here

Item B Transactions

{ FIXED [Item B] : COUNTD([Transaction ID]) }

The calculation is valid.

ApplyOK

Add data to visualise

Double-click or drag fields from the data pane

Support

Formula

[Pair Transactions] / [total Transactions]

Drop field here

Support

[Pair_transactions] / [Total Transactions]

The calculation is valid.

ApplyOK

Add data to visualise

Double-click or drag fields from the data pane.

Confidence

Formula

[Pair Transactions] / [Item A Transactions]

Sheet 1

Drop field here

Drop field here

Confidence

[Pair_transactions] / [Item A Transactions]

The calculation is valid.

Apply OK

Add data to visualise
Double-click or drag fields from the data pane.

Lift

([pair Transactions] * [Total Transactions]) / ([Item A Transactions] * [Item B Transactions])

Lift

(([Pair_transactions] * [Total Transactions]) / ([Item A Transactions] * [Item B Transactions]))

The calculation is valid.

Apply OK

Add data to visualise
Double-click or drag fields from the data pane.

Create Rule

[Item A] + “ → “ + [Item B]

Build Lift Analysis Sheet

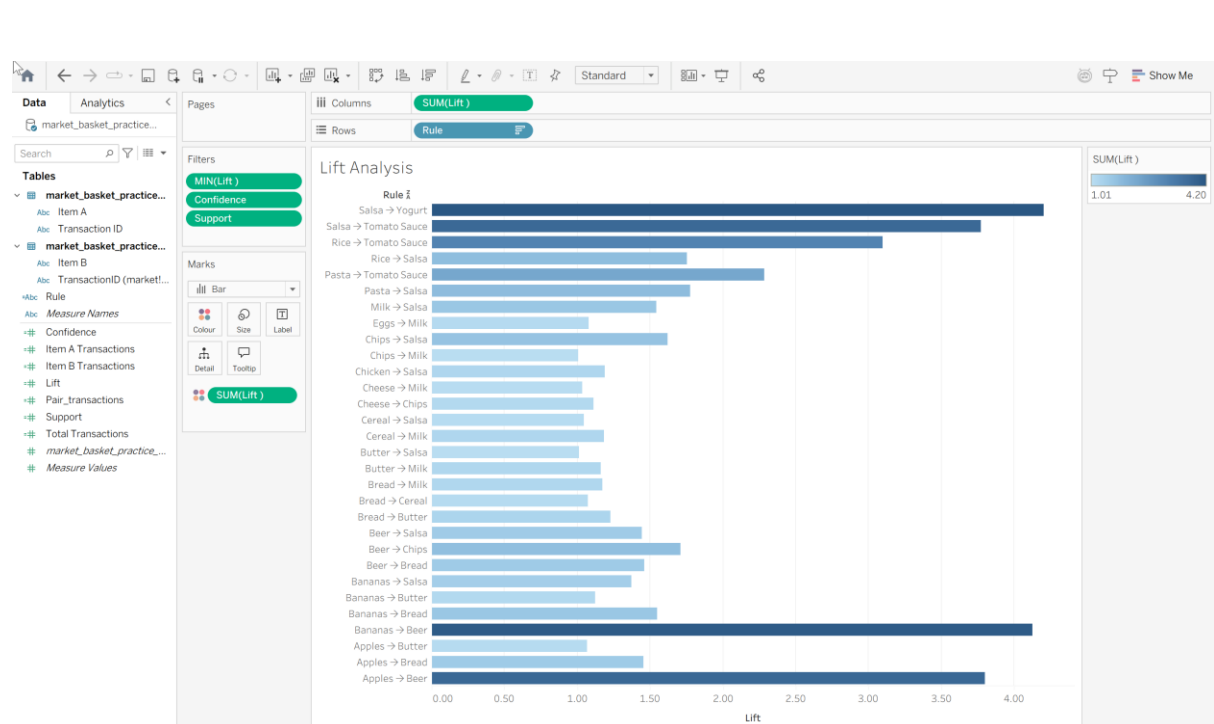
Rename Sheet 1 as Lift Analysis Sheet by double-clicking on the name Sheet 1.

Drag Rule → Rows

Drag Lift → columns

Change Marks → Bar

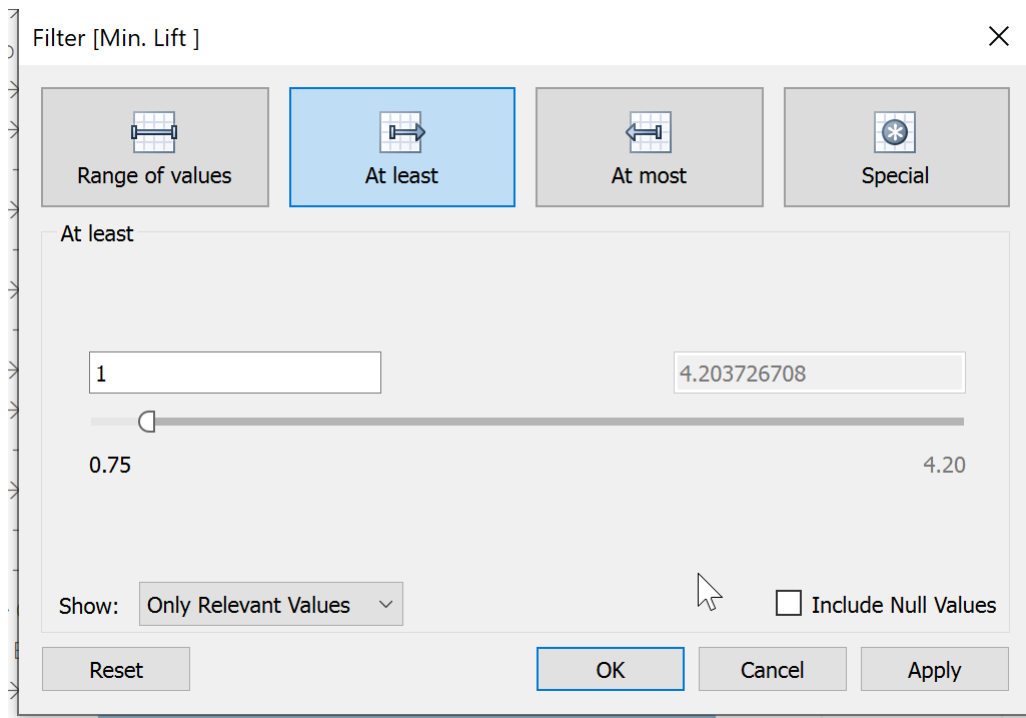
Drag Lift → colour



Add proper Filters

Drag to filters Lift, Support, confidence

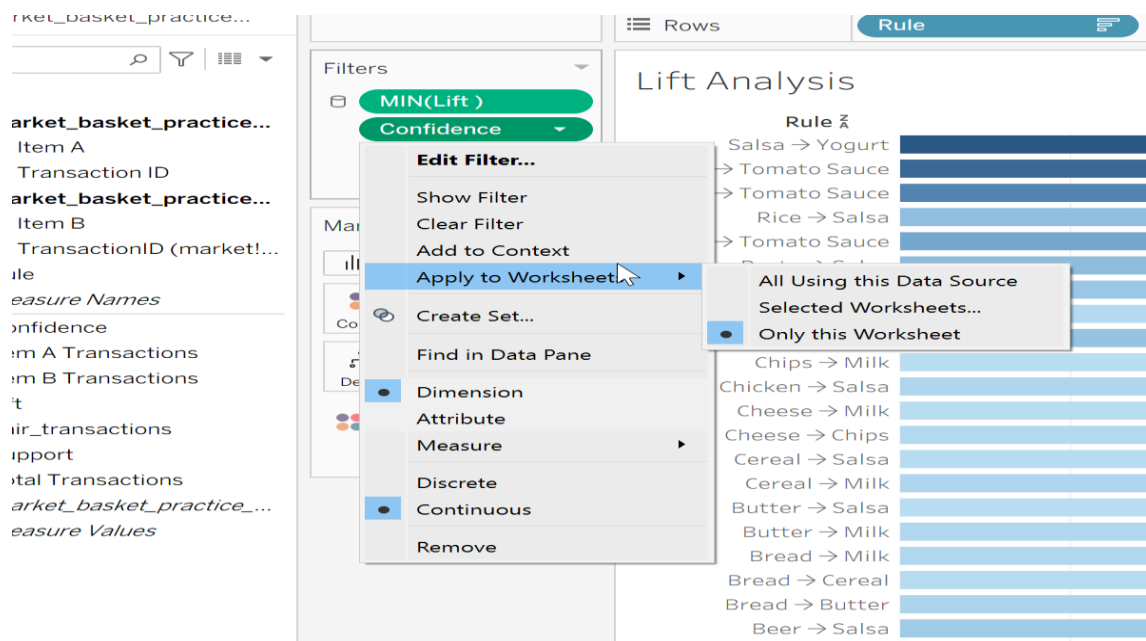
Lift → change (At least) to 1



Confidence → At least 0.3

Support → At Least 0.01

Apply this filter to all the worksheets



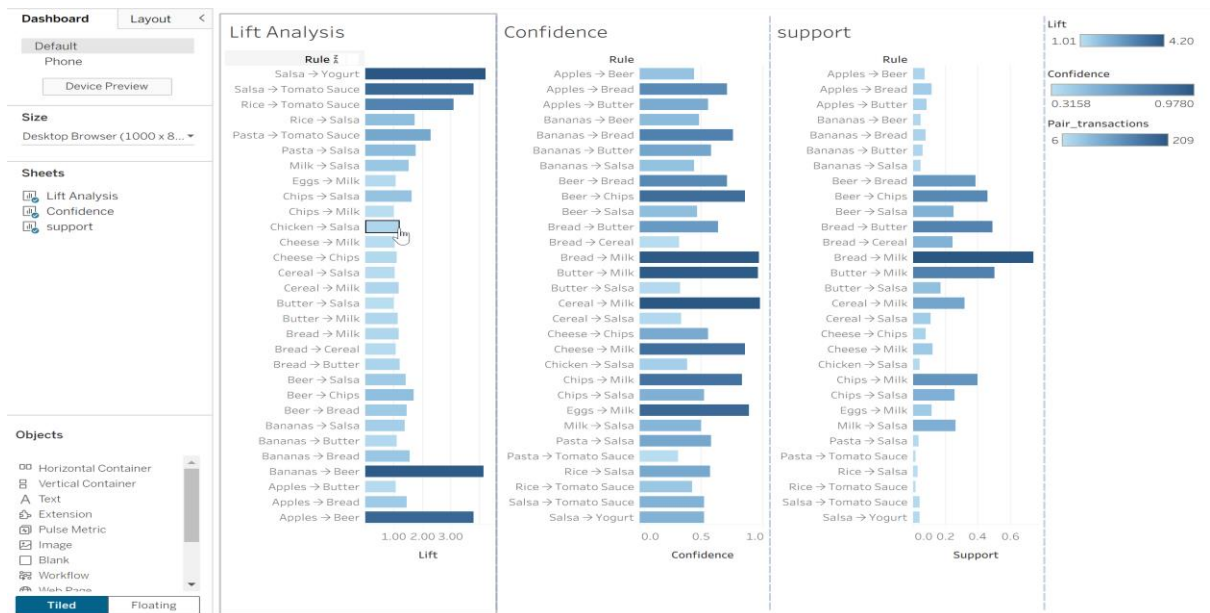
Click all using this data source

Exercise

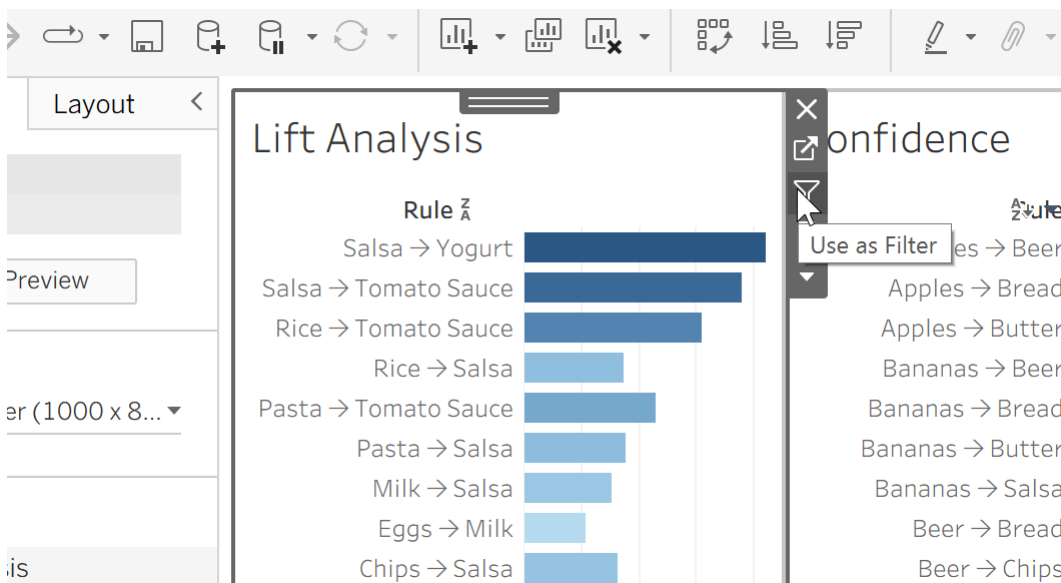
Build the confidence sheet and the support sheet.

Create a Dashboard

- Click Dashboard at the top, then click New Dashboard.
- Drag the Lift sheet to the top
- Add horizontal container below
- Drag Confidence and Support inside

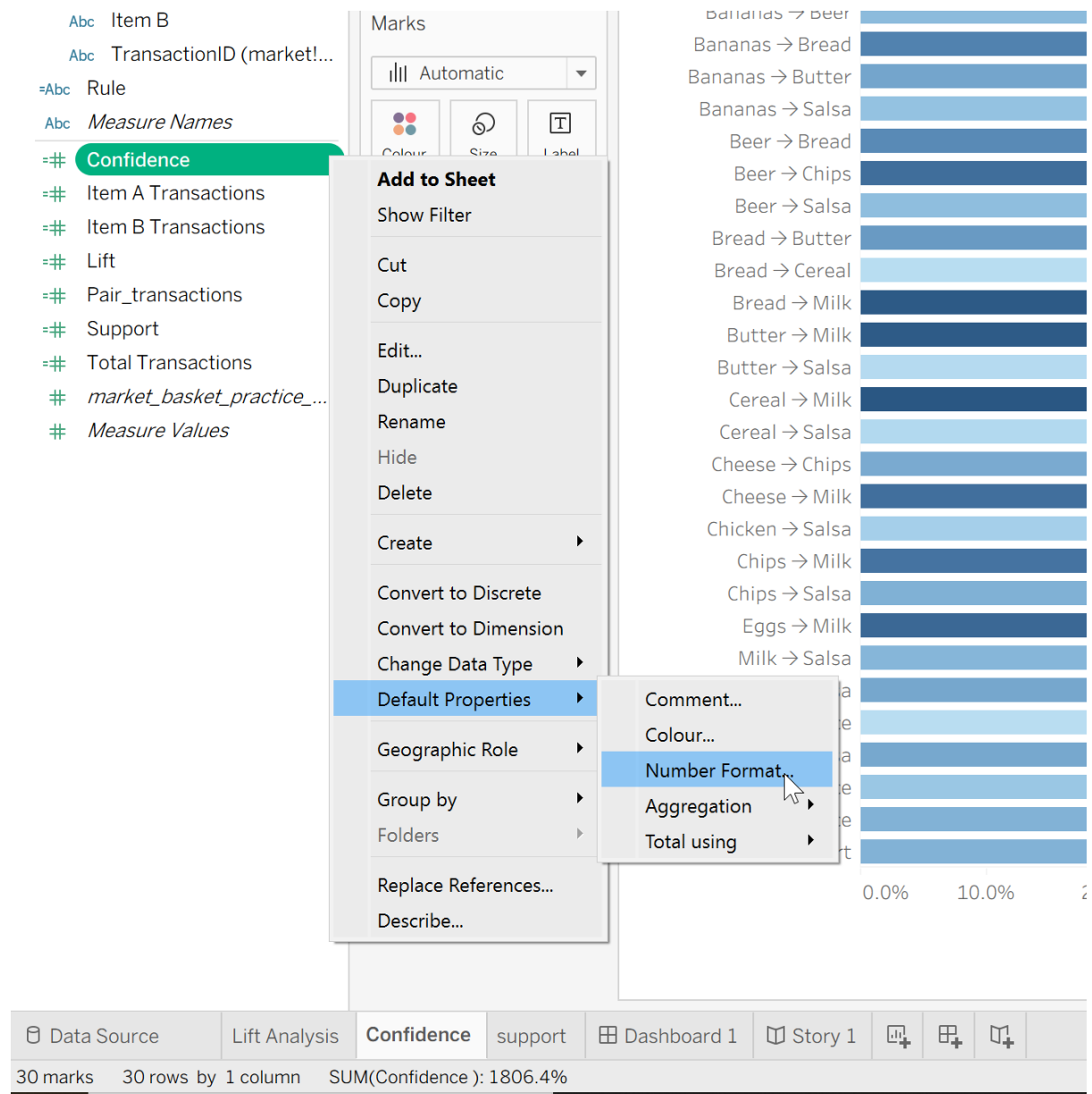


Click the Lift sheet to use as a filter



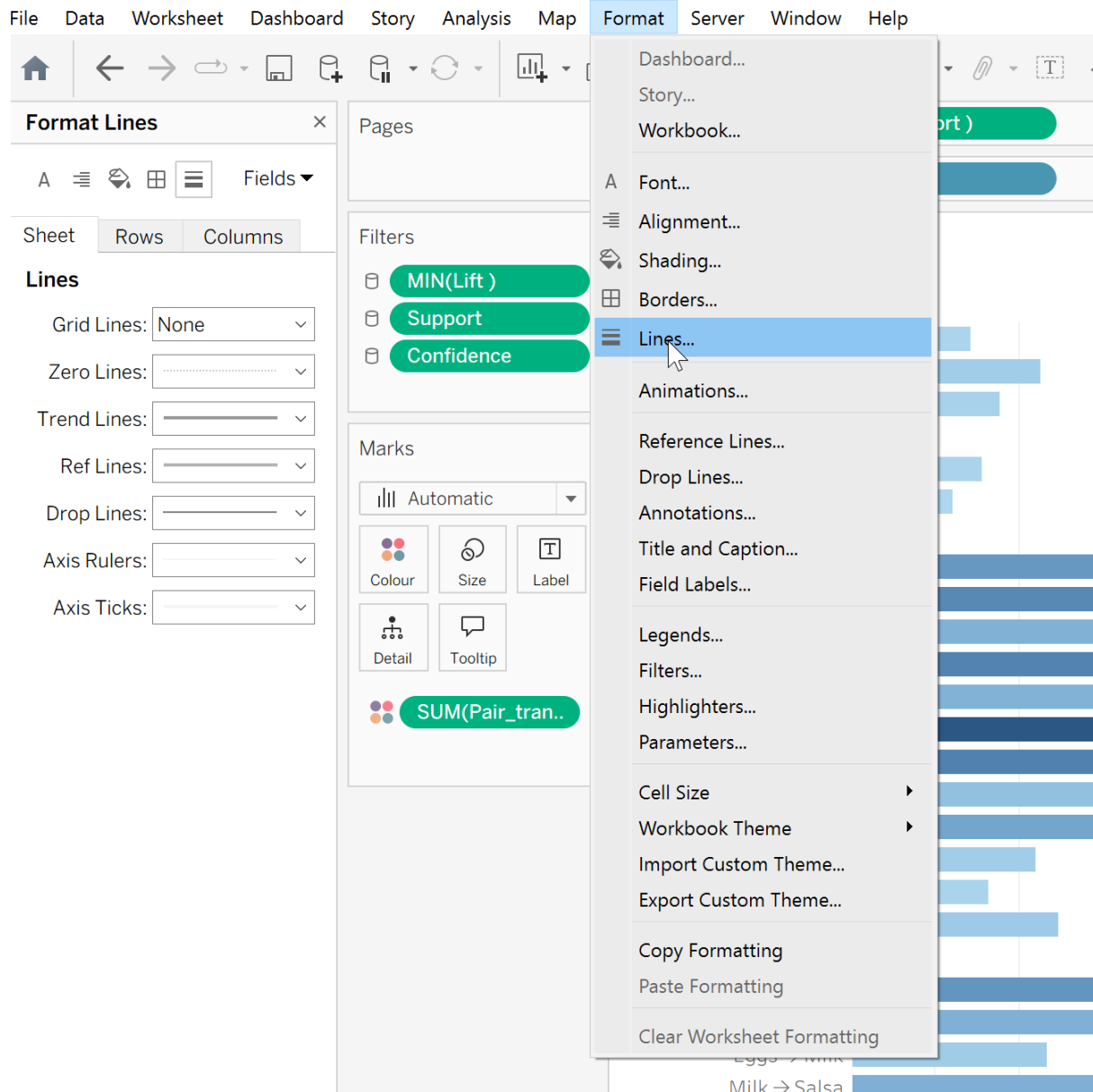
Change Confidence to a percentage

Go back to the data pane, right-click confidence, click Default Properties, then Number format. Select Percentage, set decimal place to 1.



Remove Gridlines

Open the worksheet, not the dashboard. From the top menu, click Format, then click Lines. A formatting pane will open on the left. Click None to remove the gridlines.



Exercise

- Add a title and a subtitle.

