

## Answer from a to d

In Power BI, in Data Tab, choose Get Data and choose SQL Server database to extract the data from Platform1, choose Direct Query and write the query to extract the necessary columns.

Example :

select billingstatus, countrycode, email, customers, tradingname from Platform1..BusinessContacts

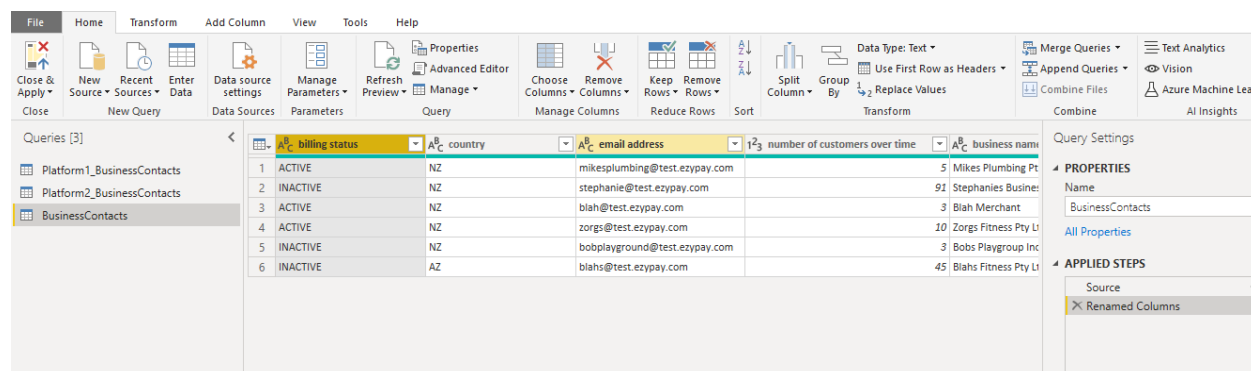
Also, choose PostgreSQL database to extract data from Platform2, choose Direct Query and write the query to extract the necessary columns and transform the data to match the data with platform1 (billingstatusid, countrycodeid) and rename the columns, so that these 2 datasets able to combine into a new dataset.

Example:

```
select case cast(billingstatusid as varchar(5))
        when '1' then 'ACTIVE'
        when '2' then 'INACTIVE'
        else 'NA'
      end as billingstatus,
       case cast(countrycodeid as varchar(5))
        when '1' then 'NZ'
        when '4' then 'AZ'
        else 'NA'
      end as countrycode,
       email, customercount as customers, tradingname
from "Platform2"."BusinessContacts"
```

In Power BI 's Power Query Editor, choose "Append Queries as New" to combine the data from these 2 platforms, make sure all the columns name are same in these 2 platforms. Rename the column name for easy understanding.

## Dataset output



	billing status	country	email address	number of customers over time	business name
1	ACTIVE	NZ	mikesplumbing@test.ezypay.com	5	Mikes Plumbing Pt
2	INACTIVE	NZ	stephanie@test.ezypay.com	91	Stephanies Busine
3	ACTIVE	NZ	blah@test.ezypay.com	3	Blah Merchant
4	ACTIVE	NZ	zorgs@test.ezypay.com	10	Zorgs Fitness Pty Li
5	INACTIVE	NZ	bobplayground@test.ezypay.com	3	Bobs Playgroup Inc
6	INACTIVE	AZ	blahs@test.ezypay.com	45	Blahs Fitness Pty Li

## Report

The screenshot displays the Microsoft Power BI Desktop interface. The top ribbon includes tabs for File, Home, Insert, Modeling, View, Help, and Table tools. The Table tools tab is active, showing options like 'Mark as date table', 'Manage relationships', and 'New measure'. The main workspace shows a data table with the following data:

business name	billing status	country	email address	number of customers over time
Blah Merchant	ACTIVE	NZ	blah@test.ezypay.com	3
Mikes Plumbing Pty Ltd	ACTIVE	NZ	mikesplumbing@test.ezypay.com	5
Zorgs Fitness Pty Ltd	ACTIVE	NZ	zorgs@test.ezypay.com	10
Blahs Fitness Pty Ltd	INACTIVE	AZ	blahs@test.ezypay.com	45
Bobs Playgroup Inc.	INACTIVE	NZ	bobplayground@test.ezypay.com	3
Stephanies Businessd	INACTIVE	NZ	stephanie@test.ezypay.com	91
<b>Total</b>				<b>157</b>

On the right side, the Fields pane is visible, showing a search bar and a list of fields under the 'BusinessContacts' table. The fields listed are: billing status, business name, country, email address, and number of customers over time. The 'number of customers over time' field is currently selected.

## e. When to use a dataflow vs a dataset (if choosing PowerBI)

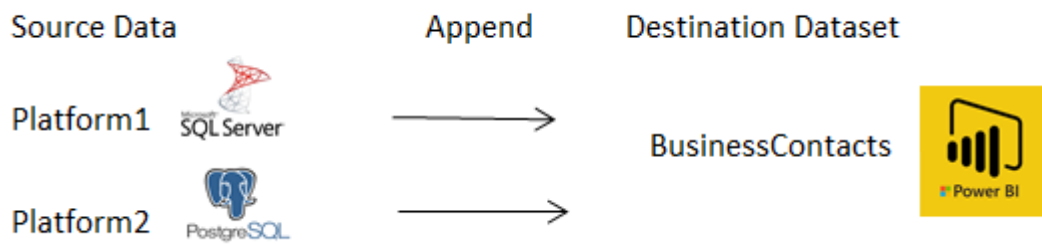
Dataflow is a data transformation component in PowerBI. It is Power Query process that runs in Cloud and store data in Common Data Model, CDM inside Azure Data Lake storage.

Dataset is the object that contains the connection to the data source, data tables, the data itself, the relationship between tables and DAX calculation.

Dataflow is the Data Transformation layer in Power BI. This will extract data from source, transform it and load into CDM. Dataset is the layer of all calculations and modelling, it will get data from dataflow, and build an in memory data model in Power BI engine. The result of dataflow will fed into a dataset for further modelling.

Although dataset can directly get data from a data source, but it is a best practice that a shared dataset gets the data from dataflow.

## f. How you will document your data model



BusinessContacts (metadata)

column name	data type
billing status	text
country	text
email address	text
business name	text
number of customers over time	number

