

Dodging DAX



DATA MODELING BY EXAMPLE

DAX

Alexander Arvidsson | [@arcticdba.se](https://arcticdba.se)

MANDO BIKES



Thank you to our Fabric February Friends!

twoday



bouvvet

sopra  steria



DATAmasterminds



 Profisee
Master Data Management



Tabular Editor

KURANT

 Fraktal

 CluedIn

 Dufrain
THE DATA COMPANY



Power BI **reports** on
operational data

Modeling **concepts**

Explore a **better** way of
doing Power BI **reports**





(he/him)

Alexander Arvidsson

Principal Solutions Architect @
Data Masterminds

| Data Platform MVP | MCT | Speaker |



DATAmasterminds

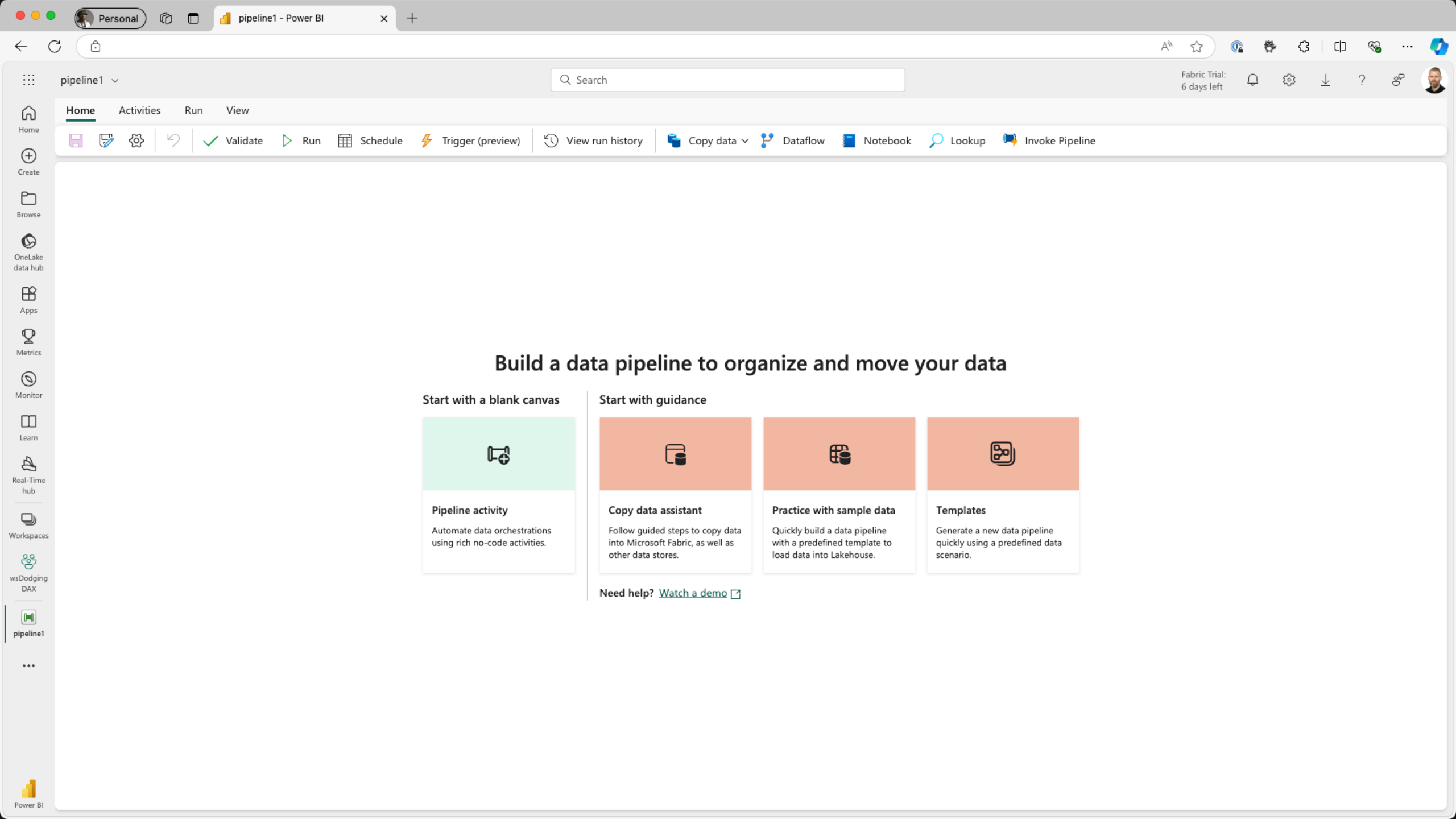




Bike
Production
System



Price lists



Personal

pipeline1 - Power BI

Search

Fabric Trial: 6 days left

Home

Activities

Run

View

Copy data

Choose data source

Select a connector. Then enter the connection information.

Connect to data source

Choose data destination

Connect to data destination

Review + save

Home

OneLake data hub

Sample data

New

Azure

Search

All

File

Database

Power Platform

Azure

Online services

Other

Folder

File

MySQL database

Database

Google BigQuery

Database

Azure Synapse Analytics (SQL ...)

Azure

Azure Data Lake Storage Gen2

Azure

Odbc

Other

Amazon RDS for SQL Server

Database

MongoDB for Pipeline

Database

Azure Files

Azure

Dynamics AX

Other

SQL Server database

Database

PostgreSQL database

Database

Amazon Redshift

Database

Azure Blobs

Azure

SharePoint Online list

Online services

Azure Database for PostgreSQL

Database

Azure Cosmos DB for MongoDB

Azure

Amazon S3

Other

Dynamics CRM

Other

Oracle database

Database

SAP HANA database

Database

Dataverse

Power Platform

Azure Tables

Azure

Salesforce objects

Online services

Oracle Cloud Storage

File

Azure SQL Managed Instance

Database

Azure Cosmos DB v2

Azure

Amazon S3 Compatible

Other

Google Cloud Storage

Other

IBM Db2 database

Database

Snowflake

Database

Azure SQL database

Azure

Azure Data Explorer (Kusto)

Azure

OData

Other

SFTP

File

MongoDB Atlas for Pipelines

Database

Azure Database for MySQL

Azure

Dynamics 365

Other

Http

Other

Power BI

Personal

pipeline1 - Power BI

Search

Fabric Trial: 6 days left

Home

Activities

Run

View

Copy data

Choose data source

Connect to data source

Choose data destination

Connect to data destination

Review + save

Select, preview, and choose the data.

Connect to data source

Select a table

(Connection: sqldodgingdax.databas...)

Tables

Query

Search

☒

production.brands

☐ production.categ...

☐ production.produ...

☐ production.stocks

☐ sales.customers

☐ sales.order_items

☐ sales.orders

☐ sales.staffs

☐ sales.stores

Preview data: production.brands

123 brand_id	abc brand_name
1	Electra
2	Haro
3	Heller
4	Pure Cycles
5	Ritchey
6	Strider
7	Sun Bicycles
8	Surly
9	Trek

Back

Next

Home

Create

Browse

OneLake data hub

Apps

Metrics

Monitor

Learn

Real-Time hub

Workspaces

wsDodging DAX

pipeline1

Power BI

Personal

Power BI

Power BI

←

↺

🔒

dfIngestPrices

Search

Fabric Trial: 23 days left

Home

Transform

Add column

View

Help

Get data

Enter data

Manage connections

Options

Manage parameters

Refresh

Advanced editor

Add data destination

Choose columns

Remove columns

Keep rows

Remove rows

Filter rows

Sort

Split column

Group by

Use first row as headers

Replace values

Merge queries

Append queries

Combine files

Map to entity

Copilot

Export template

Queries [1]

Table.TransformColumnTypes("#Promoted headers", {"product_id", Int64.Type}, {"product_name", type text}, {"model_year", Int64.Type}, {"list_price", type number})

	product_id	product_name	model_year	list_price
1	1	Trek 820 - 2016	2016	379.99
2	2	Ritchey Timberwolf Frameset - 2016	2016	749.99
3	3	Surly Wednesday Frameset - 2016	2016	999.99
4	4	Trek Fuel EX 8 29 - 2016	2016	2899.99
5	5	Heller Shagamaw Frame - 2016	2016	1320.99
6	6	Surly Ice Cream Truck Frameset - 2016	2016	469.99
7	7	Trek Slash 8 27.5 - 2016	2016	3999.99
8	8	Trek Remedy 29 Carbon Frameset - 2016	2016	1799.99
9	9	Trek Conduit+ - 2016	2016	2999.99
10	10	Surly Straggler - 2016	2016	1549
11	11	Surly Straggler 650b - 2016	2016	1680.99
12	12	Electra Townie Original 21D - 2016	2016	549.99
13	13	Electra Cruiser 1 (24-Inch) - 2016	2016	269.99
14	14	Electra Girl's Hawaii 1 (16-inch) - 2015/2016	2016	269.99
15	15	Electra Moto 1 - 2016	2016	529.99
16	16	Electra Townie Original 7D EQ - 2016	2016	599.99
17	17	Pure Cycles Vine 8-Speed - 2016	2016	429
18	18	Pure Cycles Western 3-Speed - Women's - 20...	2016	449
19	19	Pure Cycles William 3-Speed - 2016	2016	449
20	20	Electra Townie Original 7D EQ - Women's - 2016	2016	599.99
21	21	Electra Cruiser 1 (24-Inch) - 2016	2016	269.99
22	22	Electra Girl's Hawaii 1 (16-inch) - 2015/2016	2016	269.99
23	23	Electra Girl's Hawaii 1 (20-inch) - 2015/2016	2016	299.99
24	24	Electra Townie Original 21D - 2016	2016	549.99
25	25	Electra Townie Original 7D - 2015/2016	2016	499.99
26	26	Electra Townie Original 7D EQ - 2016	2016	599.99
27	27	Surly Big Dummy Frameset - 2017	2017	999.99
28	28	Surly Karate Monkey 27.5+ Frameset - 2017	2017	2499.99
29	29	Trek X-Caliber 8 - 2017	2017	999.99
30	30	Surly Ice Cream Truck Frameset - 2017	2017	999.99

Query settings

Properties

Name

sheet1

Entity type

Custom

Applied steps

Source

Navigation 1

Promoted ...

Changed c...

Data destination

Lakehouse

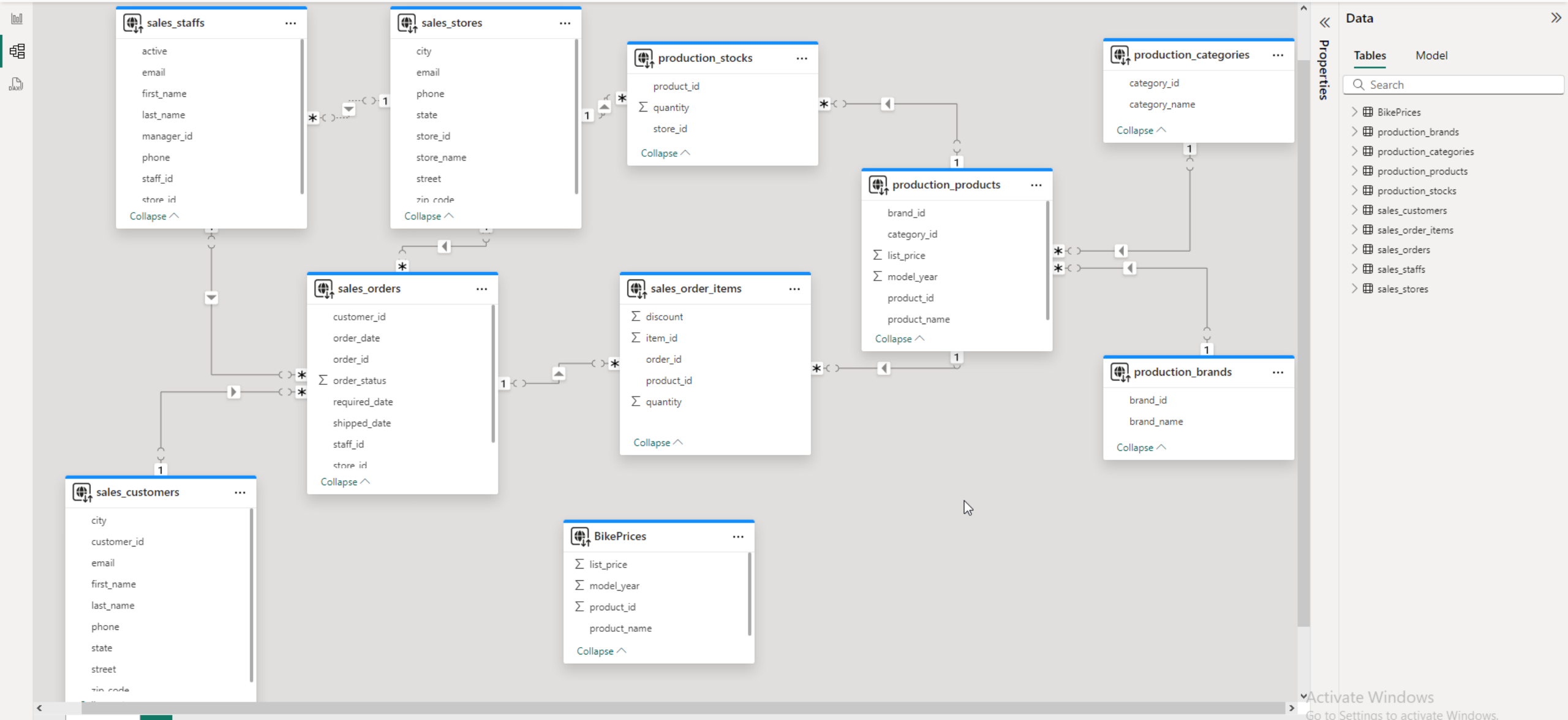
Completed (1.35 s)

Columns: 4

Rows: 99+

Step

Publish

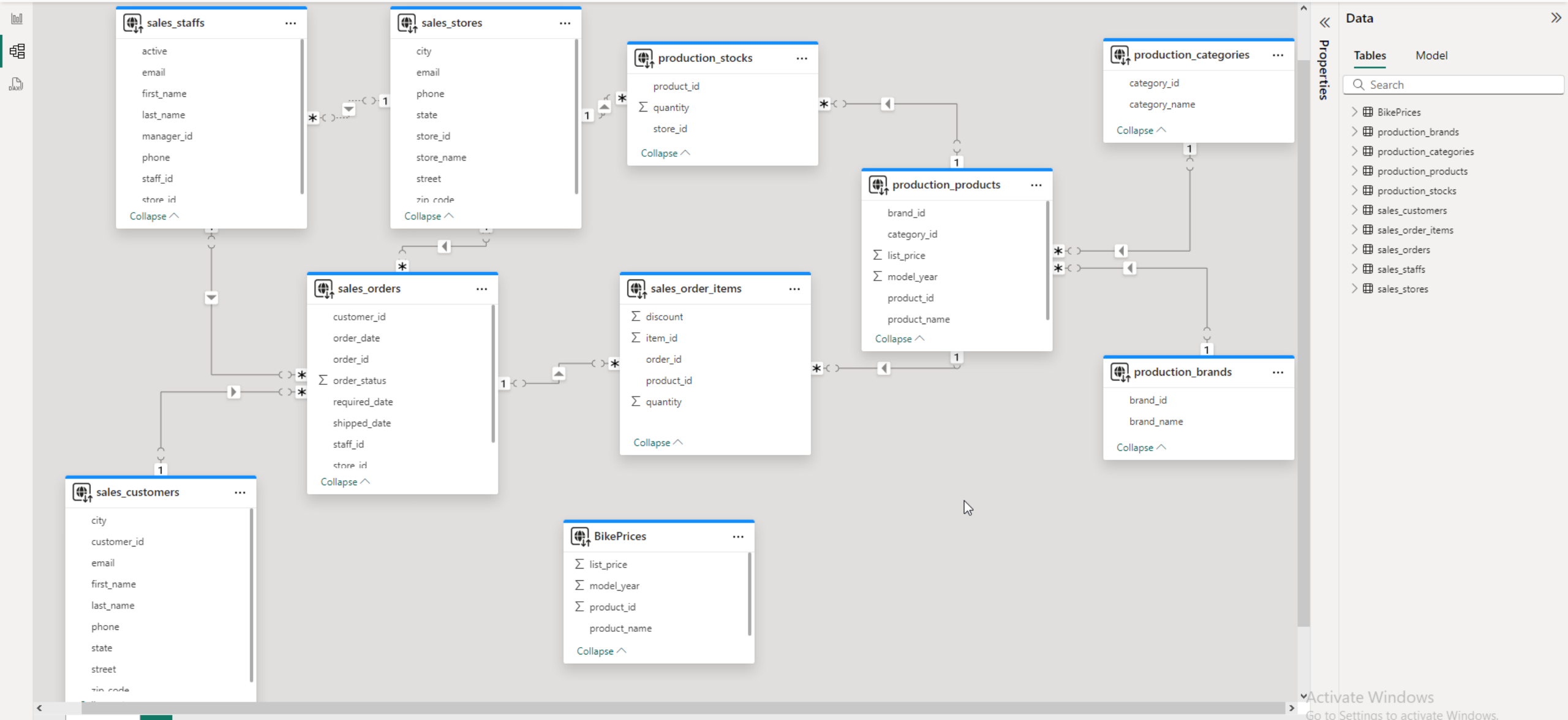


Report **requirements:**

Total Sales Revenue

Rolling 3-month Average Sale

Net Sales per Quarter



Making these reports
won't be **fun**.

In fact, it **won't** even
be **possible** with this
data model.

It's actually a
really bad idea.

Seriously.

Why is this going to
be **difficult**?

The operational data
model is **not suitable**
for **reporting**.



[Visible confusion]

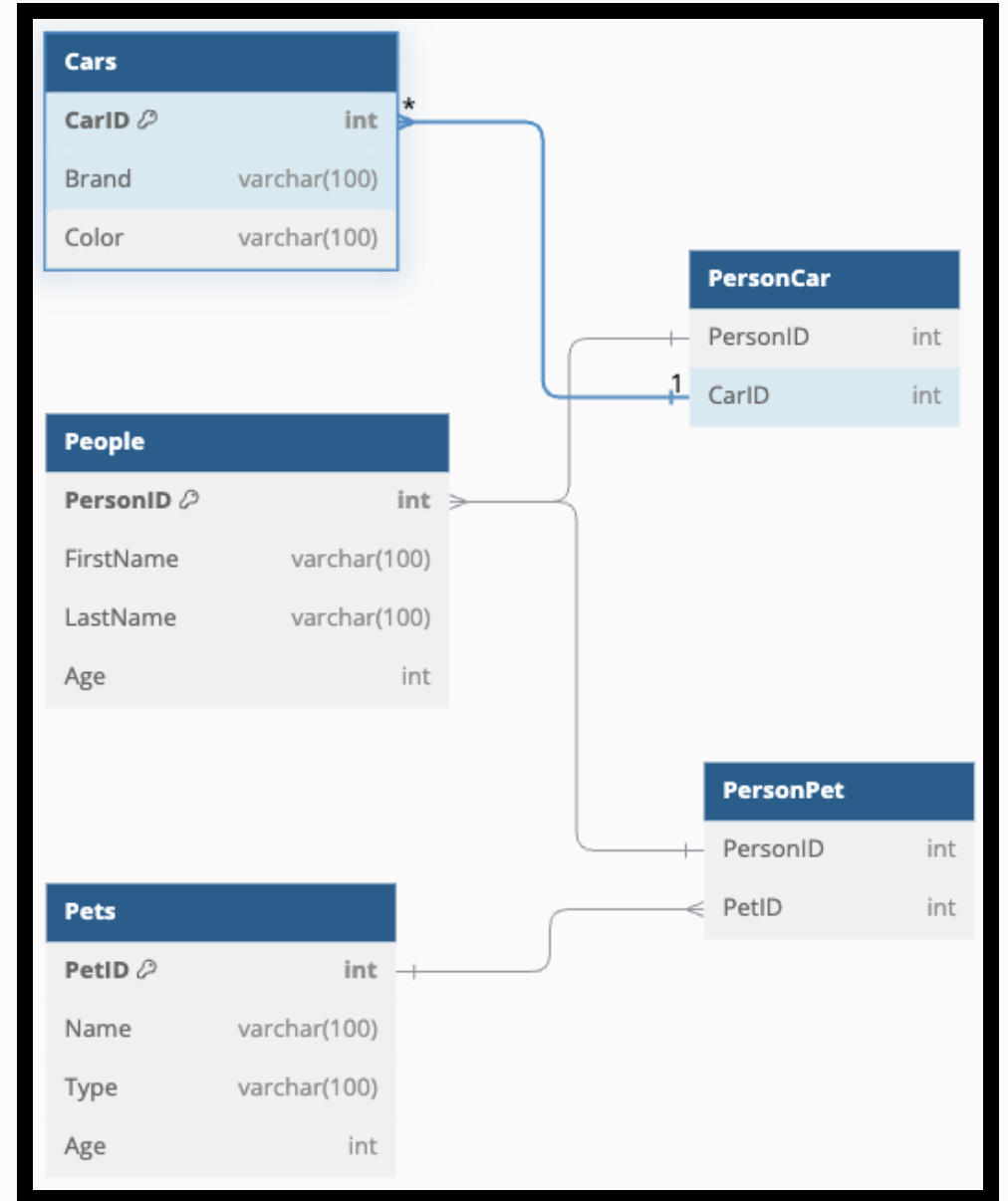
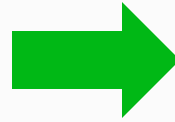
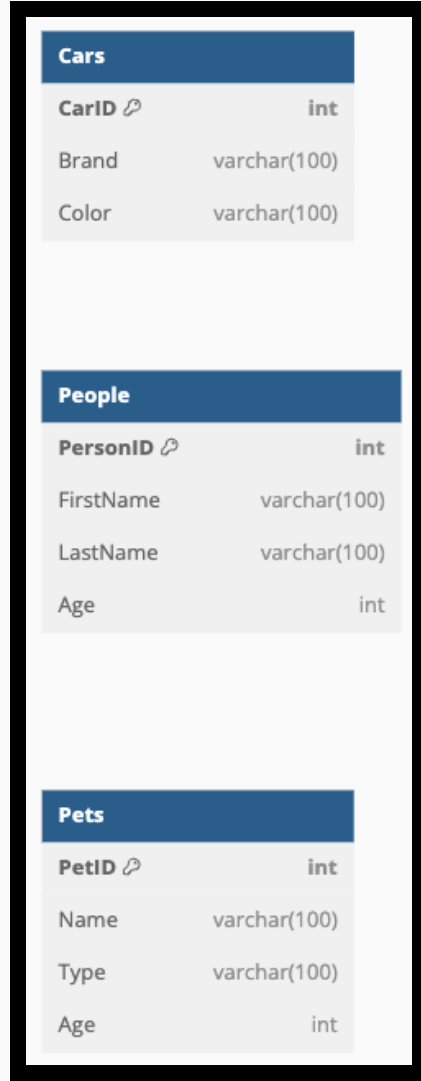
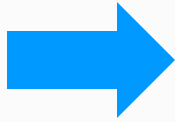
Several
reasons for **that.**

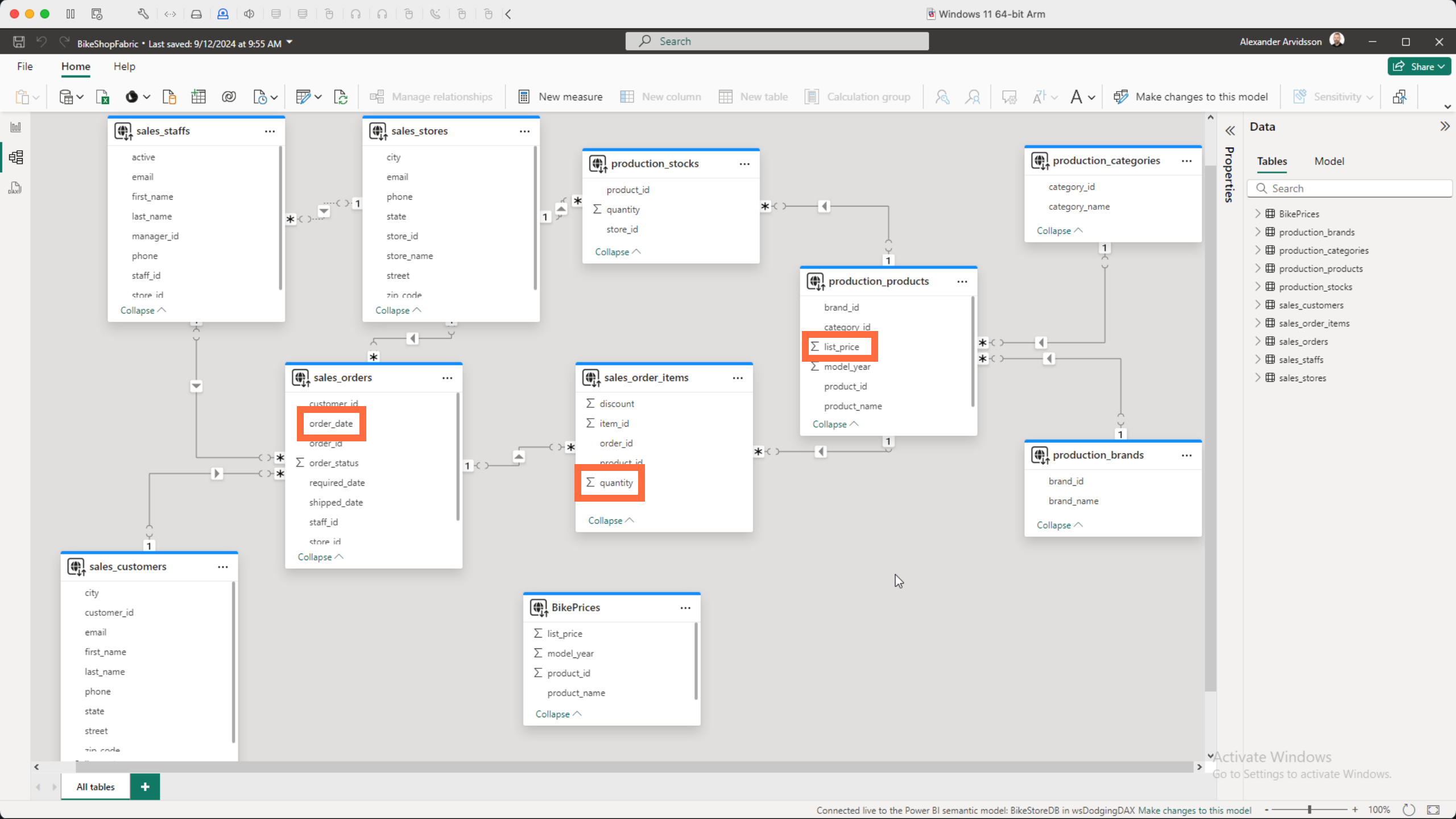
The data model is
normalized.

No.

Normalized**ED**.

Stuff	
Name	varchar(200)
Car	varchar(200)
Car color	varchar(20)
Pet name	varchar(30)



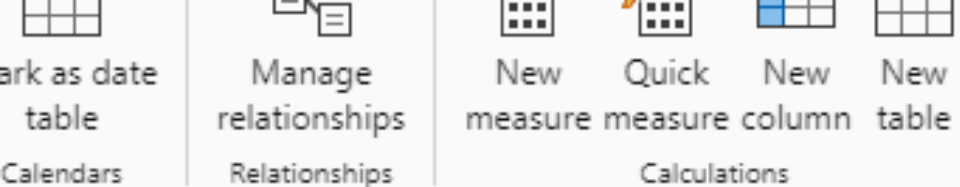


A **normalized**
data model is
not intuitive.

A **normalized**
data model has
many **uses.**

...but **reporting** is
not one of them.

The **data model** has
a **date issue**.



order_status	order_date	required_date	shipped_date	store_id	staff_id	OrderDateYear
4	Friday, 1 January 2016	Monday, 4 January 2016	Sunday, 3 January 2016	2	6	2016
4	Sunday, 3 January 2016	Wednesday, 6 January 2016	Wednesday, 6 January 2016	2	6	2016
4	Monday, 4 January 2016	Thursday, 7 January 2016	Tuesday, 5 January 2016	2	6	2016
4	Monday, 4 January 2016	Thursday, 7 January 2016	Tuesday, 5 January 2016	2	6	2016
4	Tuesday, 5 January 2016	Wednesday, 6 January 2016	Wednesday, 6 January 2016	2	6	2016
4	Friday, 8 January 2016	Monday, 11 January 2016	Monday, 11 January 2016	2	6	2016
4	Friday, 15 January 2016	Saturday, 16 January 2016	Monday, 18 January 2016	2	6	2016
4	Monday, 18 January 2016	Thursday, 21 January 2016	Thursday, 21 January 2016	2	6	2016
4	Tuesday, 19 January 2016	Wednesday, 20 January 2016	Thursday, 21 January 2016	2	6	2016
4	Wednesday, 20 January 2016	Friday, 22 January 2016	Thursday, 21 January 2016	2	6	2016
4	Wednesday, 20 January 2016	Thursday, 21 January 2016	Thursday, 21 January 2016	2	6	2016
4	Thursday, 21 January 2016	Friday, 22 January 2016	Friday, 22 January 2016	2	6	2016
4	Friday, 22 January 2016	Monday, 25 January 2016	Saturday, 23 January 2016	2	6	2016
4	Saturday, 23 January 2016	Sunday, 24 January 2016	Sunday, 24 January 2016	2	6	2016
4	Monday, 25 January 2016	Thursday, 28 January 2016	Tuesday, 26 January 2016	2	6	2016
4	Wednesday, 27 January 2016	Saturday, 30 January 2016	Friday, 29 January 2016	2	6	2016
4	Sunday, 31 January 2016	Monday, 1 February 2016	Tuesday, 2 February 2016	2	6	2016
4	Wednesday, 3 February 2016	Thursday, 4 February 2016	Friday, 5 February 2016	2	6	2016

The numbers will
seem absolutely **fine**

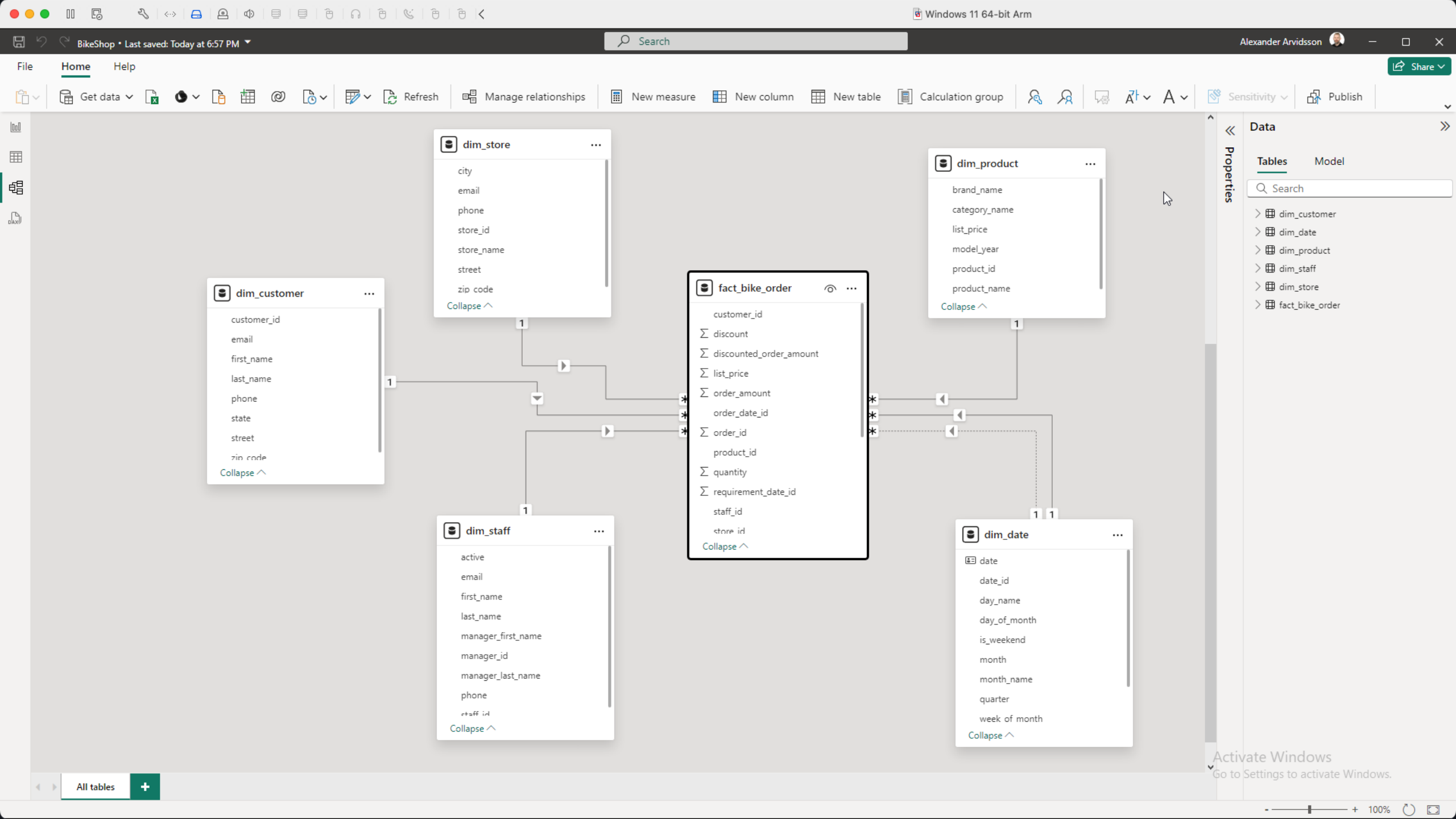
...when, **in fact,**
they're **not.**

The report will be
unbearably slow –
especially as the
dataset grows.

I **told** you it was
a **bad** idea.



We can make the
data model
much more **intuitive**.

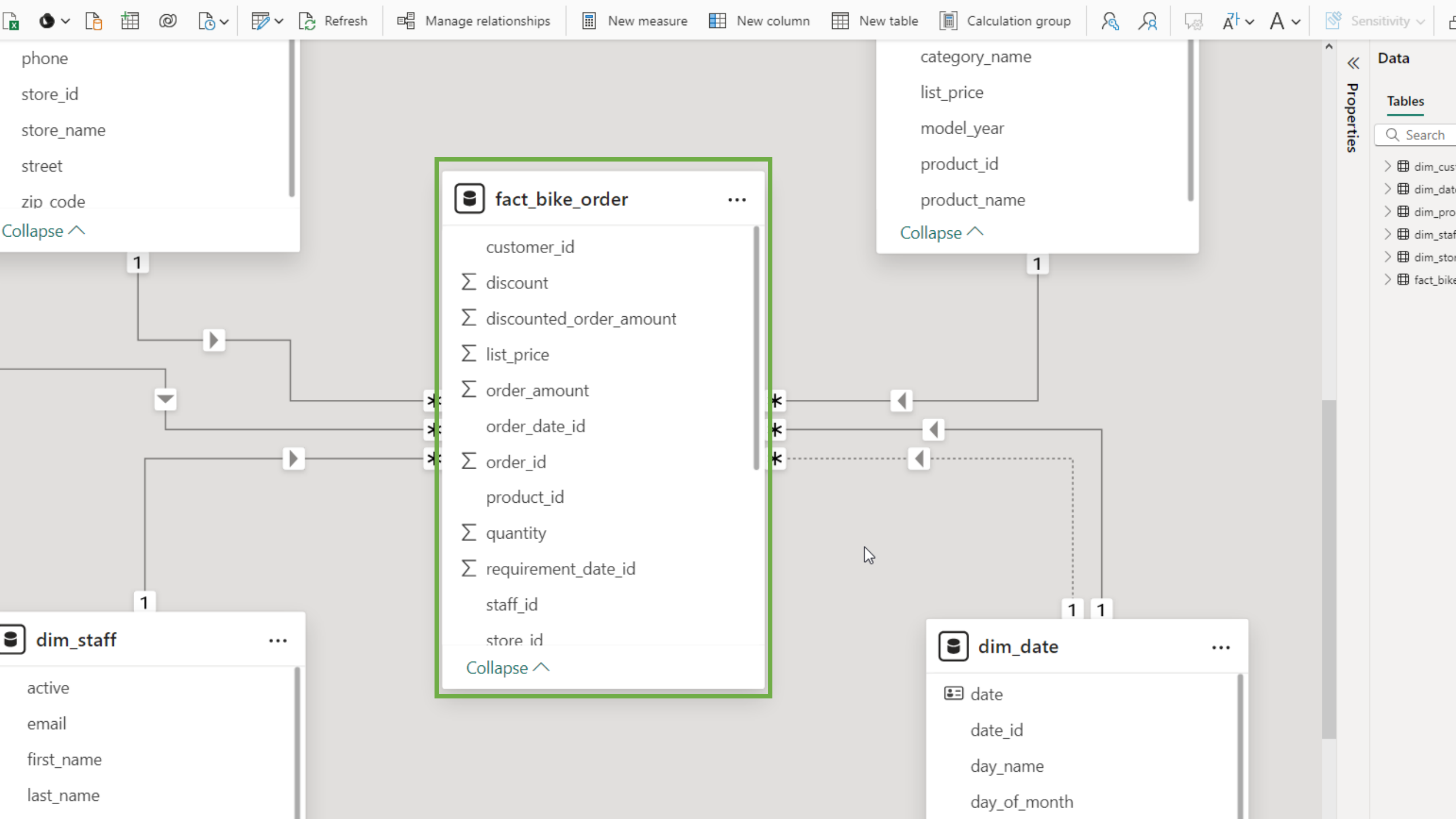


This is the **same**
information in a
dimensional structure.

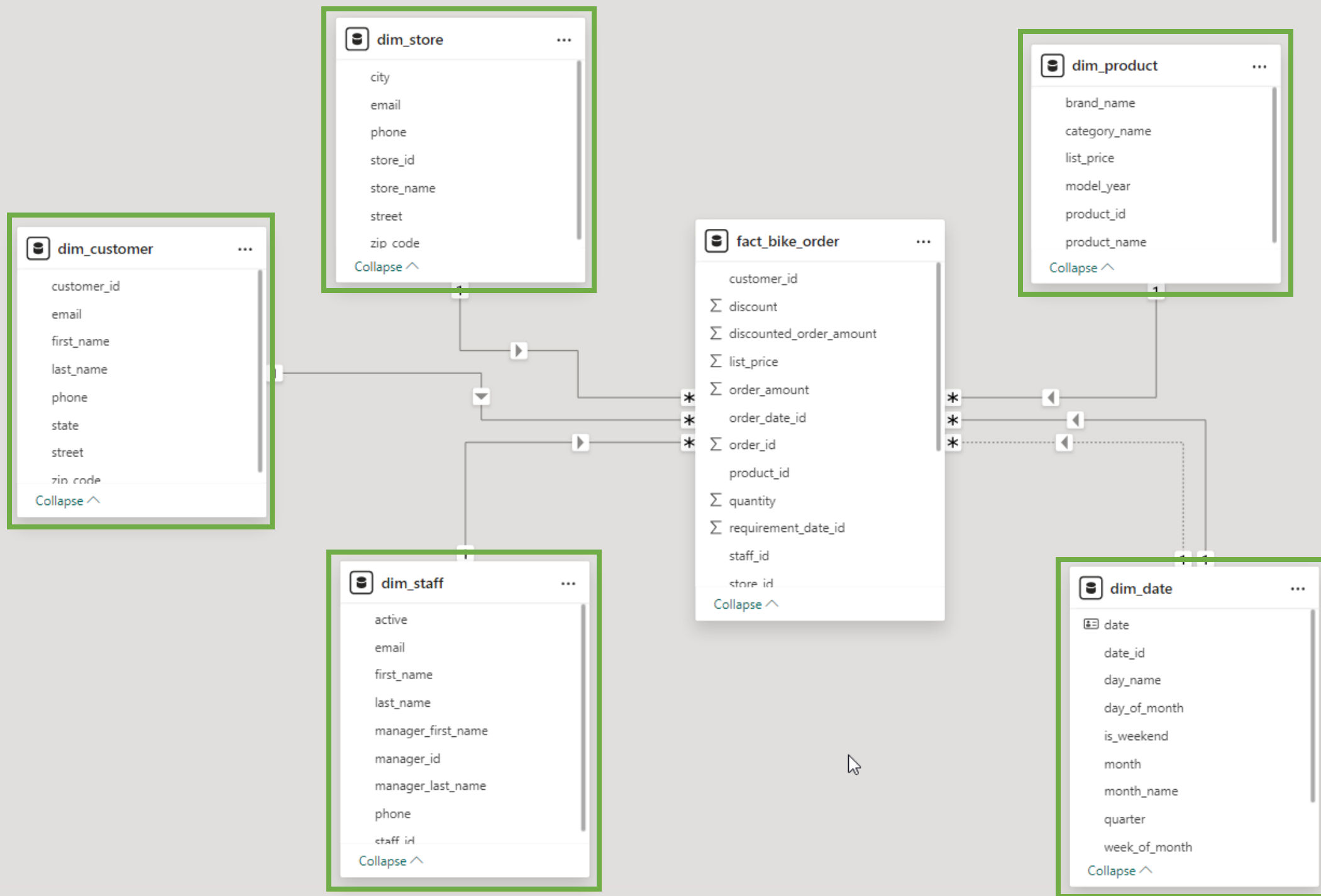
Also **known** as a
"star schema"

Two **parts** to
a **dimensional**
data model:

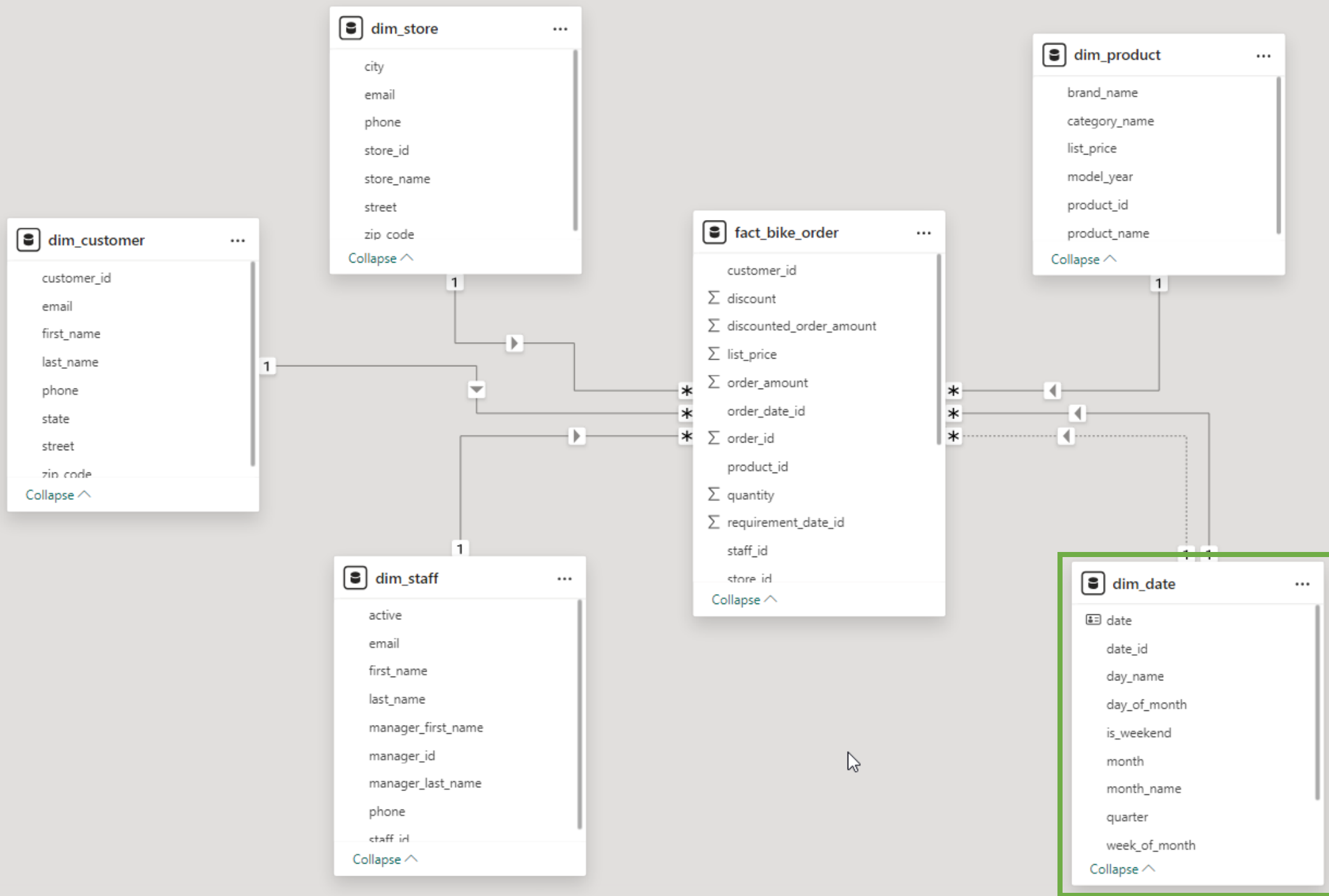
Facts



Dimensions



Remember the
date issue?



This is the way.

A **dimensional**
data model is much
more **intuitive**.

Windows 11 64-bit Arm

Alexander Arvidsson

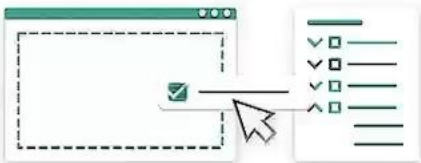
Share

FileHomeInsertModelingViewOptimizeHelp

Get dataNew visualMore visualsNew measureSensitivityPublishCopilot

Build visuals with your data

Select or drag fields from the Data pane onto the report canvas.



Visualizations

Build visual

Filters

Values

Drill through

Cross-report

Keep all filters

Add drill-through fields here

dim_customer

dim_date

dim_product

dim_staff

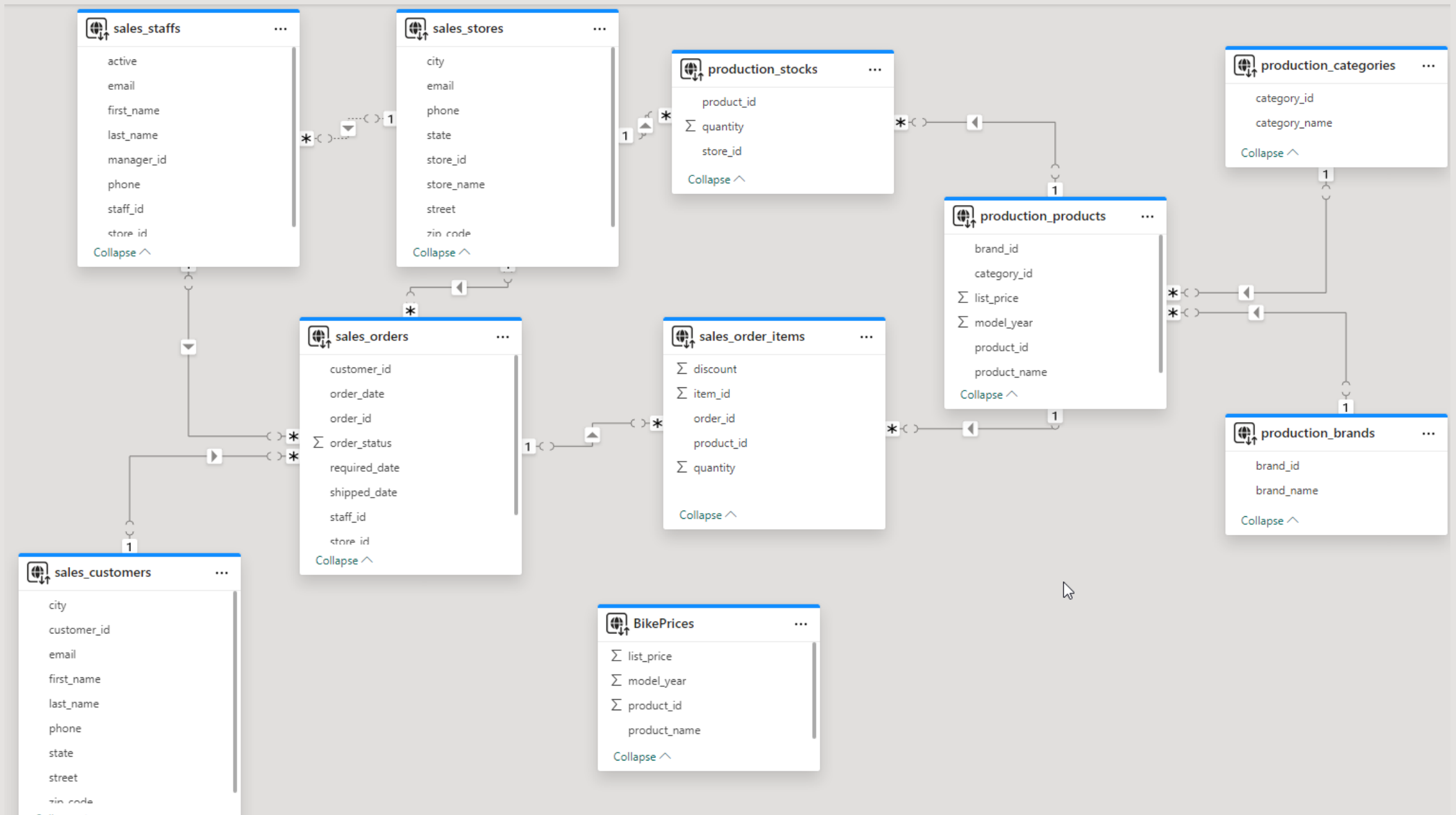
dim_store

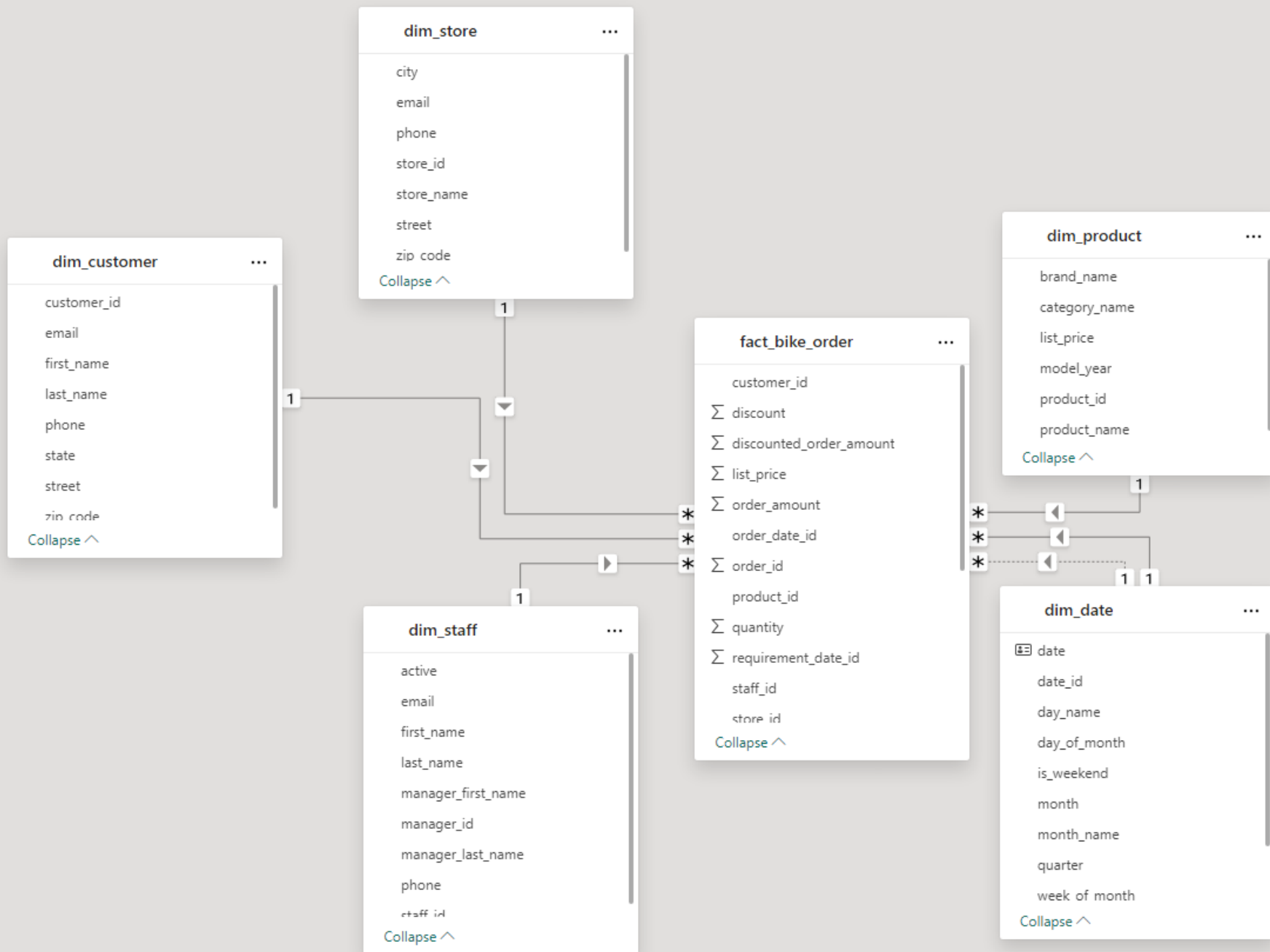
fact_bike_order

Page 1

116%

But - creating a
dimensional data
model requires **effort**.





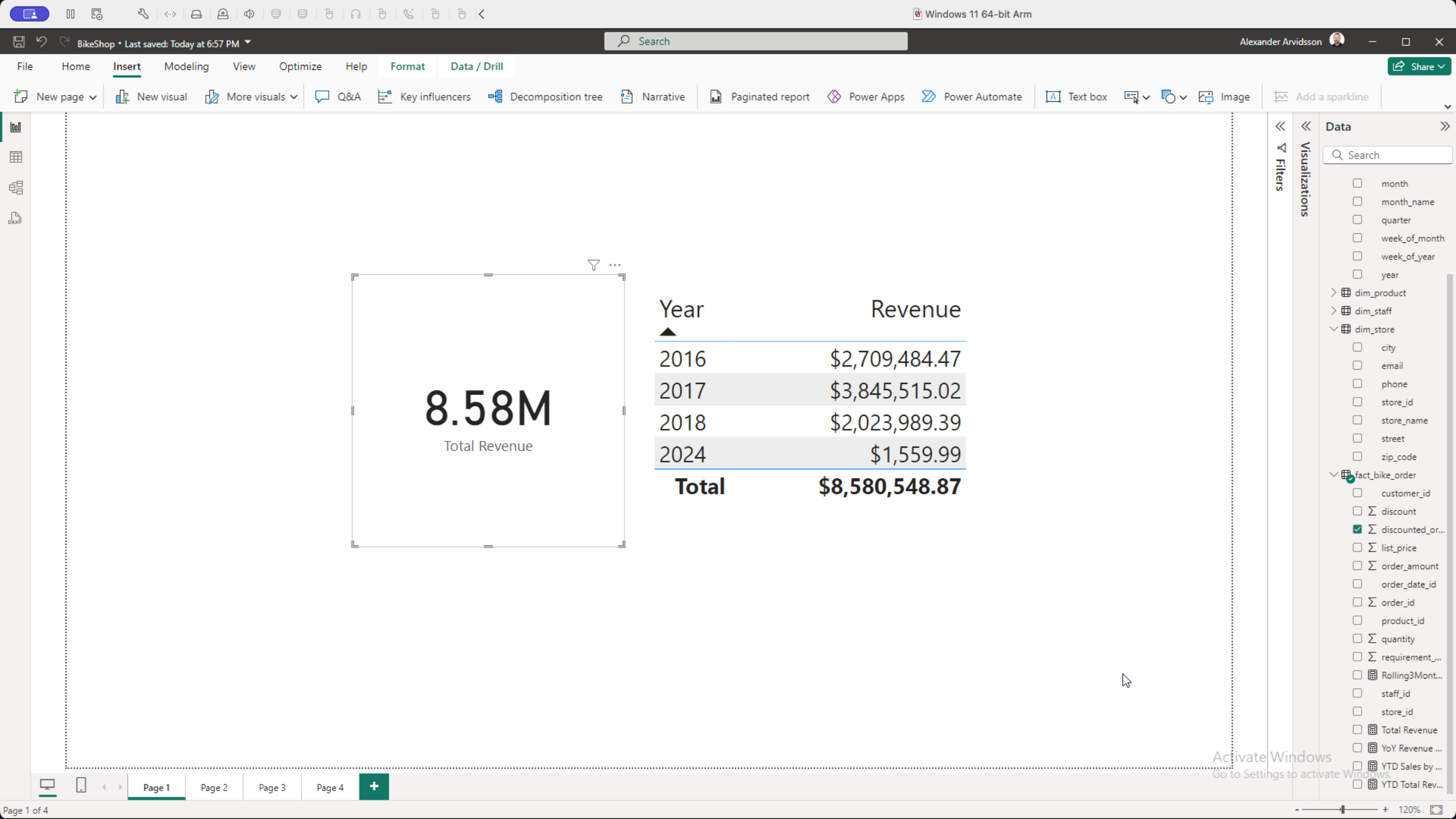
Let's put it **all together**.

Report requirements:

Total Sales Revenue

Rolling 3-month Average Sale

Net Sales per Quarter



Report requirements:

Total Sales Revenue

Rolling 3-month Average Sale

Net Sales per Quarter


```
Rolling3MonthAvgSales =  
VAR CurrentDate = MAX(dim_date[date])  
VAR ThreeMonthsAgo = EDATE(CurrentDate, -2)  
RETURN  
CALCULATE(  
    AVERAGEX(  
        VALUES(dim_date[date]),  
        CALCULATE(SUM(fact_bike_order[discounted_order_amount]))  
    ),  
    dim_date[date] >= ThreeMonthsAgo && dim_date[date] <= CurrentDate  
)
```

Date for analysis

Friday, August 05, 2016

Store Name	Rolling 3-month average sales
Baldwin Bikes	7,351.55
Rowlett Bikes	5,011.69
Santa Cruz Bikes	6,729.99
Total	9,052.87

Data

Visualizations

Filters

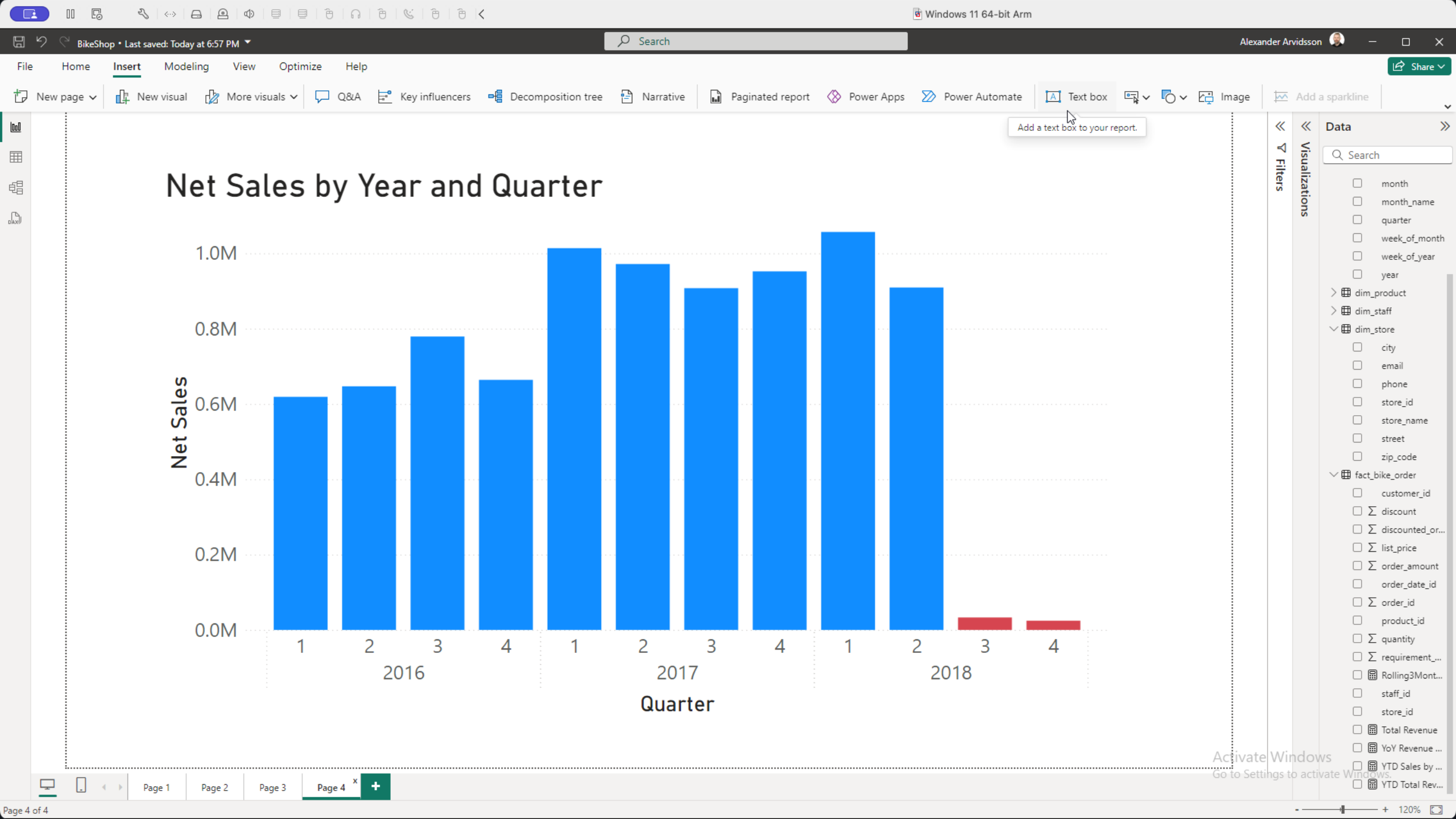
- month
- month_name
- quarter
- week_of_month
- week_of_year
- year
- dim_product
- dim_staff
- dim_store
 - city
 - email
 - phone
 - store_id
 - store_name
 - street
 - zip_code
- fact_bike_order
 - customer_id
 - discount
 - discounted_or...
 - list_price
 - order_amount
 - order_date_id
 - order_id
 - product_id
 - quantity
 - requirement_...
 - Rolling3Mont...
 - staff_id
 - store_id
 - Total Revenue
 - YoY Revenue ...
 - YTD Sales by ...
 - YTD Total Rev...

Report requirements:

Total Sales Revenue

Rolling 3-month Average Sale

Net Sales per Quarter



Measures?

Maybe. Maybe not.

“The **best** DAX is
the DAX you did
not have to write.”

- Marco Russo



“I’m not lazy,
I’m **efficient.**”

-Patrick LeBlanc



Better **Performance?**

Yes.

Bonus:

Better **Security?**

%#@\$! YES.

Just because you **can**, **doesn't**
mean you **should**.

Data modeling is a
complicated activity.

Even **more** important with
Microsoft Fabric.

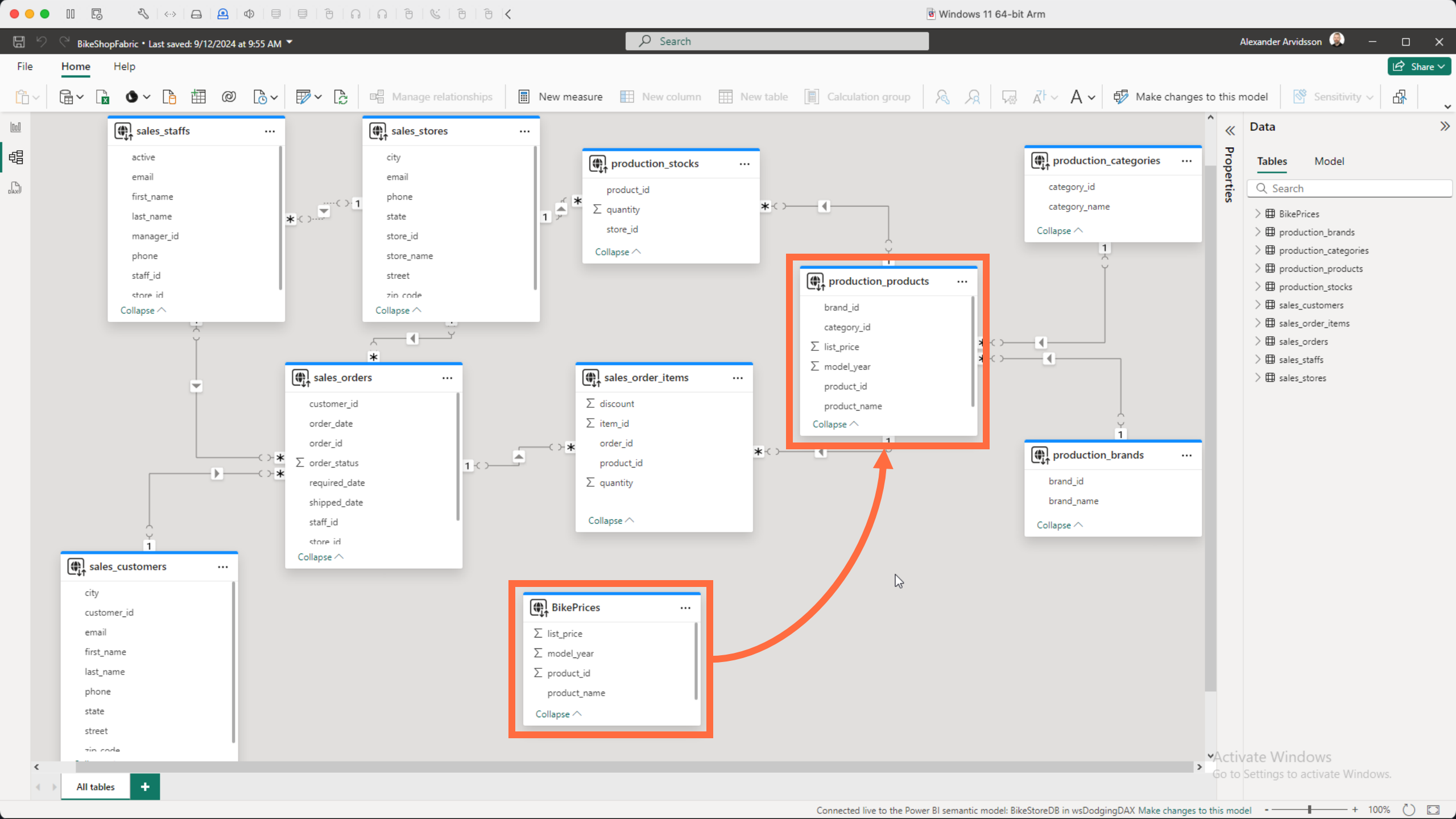
STAR SCHEMA

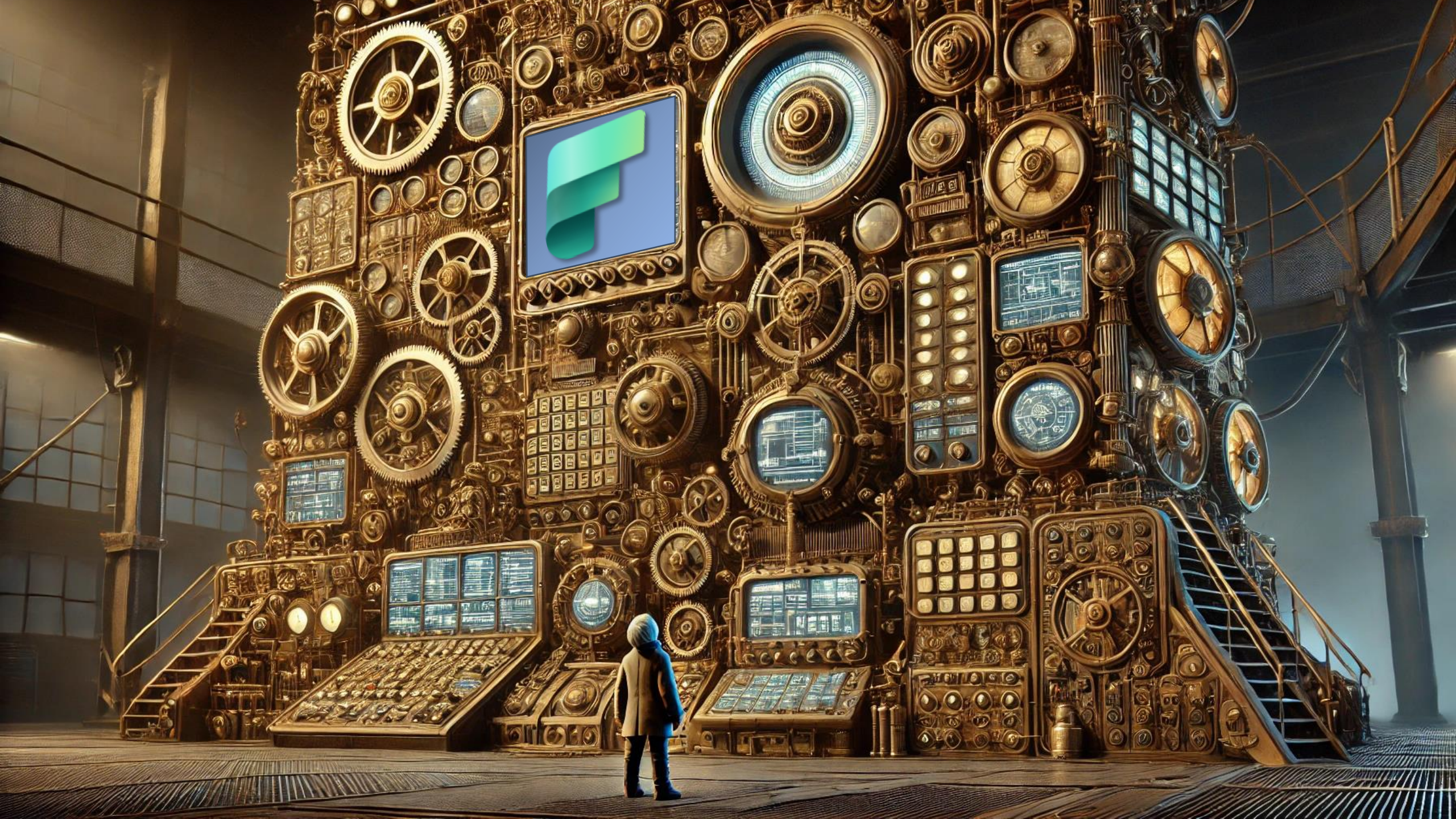


ALL THE THINGS

Dirty secret:

None of these sales
numbers are actually
correct.







Thank you!

@arcticdba.se

alexanderarvidsson@masterminds.io