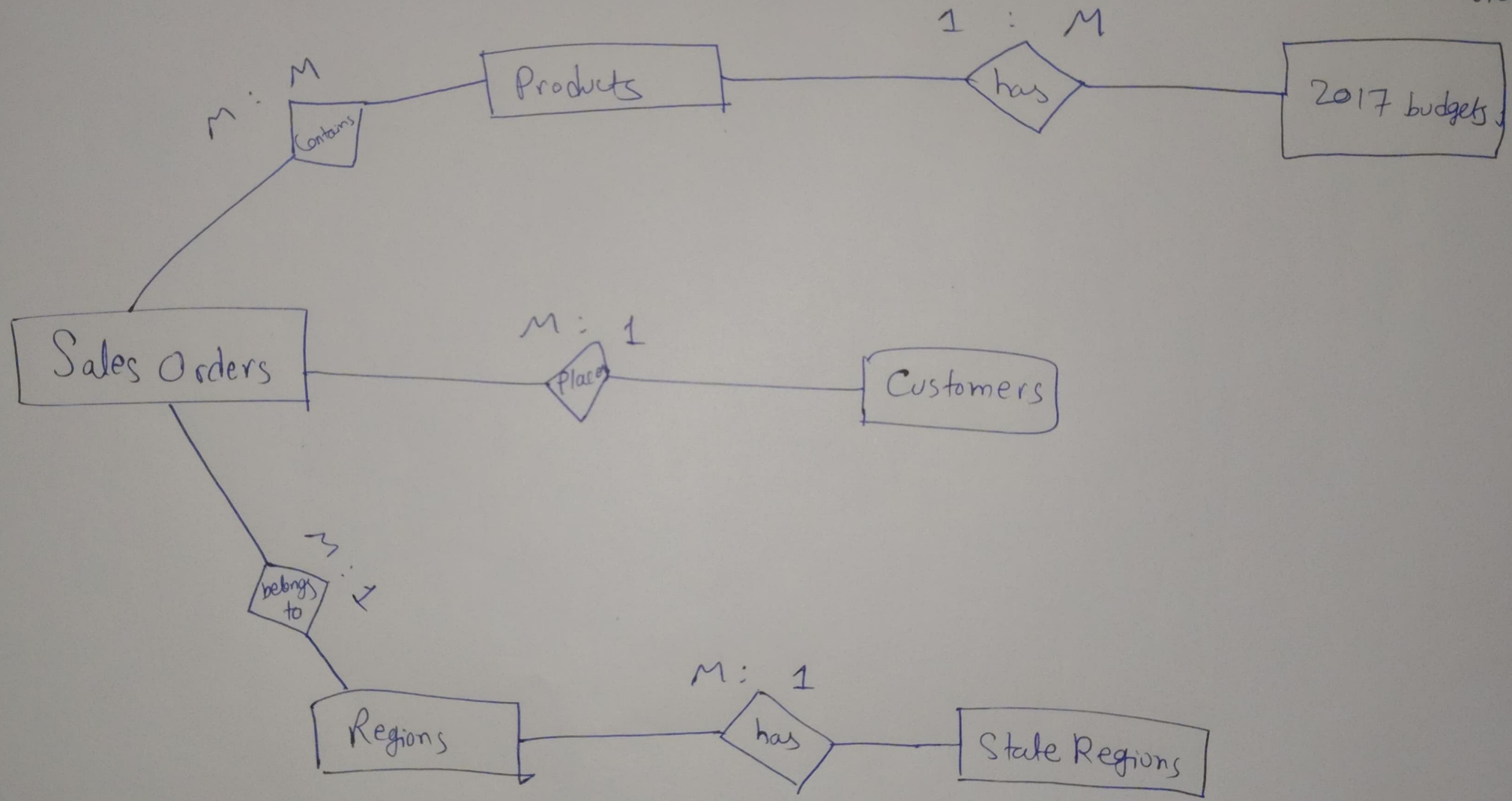
ER Diagram

Attributes are omitted for better visibility & not so messy view.



Entity Relationship Diagram.

Entity → Rows
Attributes → Columns.

Cardinality → min
→ max

Advanced ERD → Primary key
→ Foreign key
→ Bridge Tables.

Phone no. not used as PK as phone no can be changed but we need something which is never changing.

more than 1 Foreign key → possible.

Making ER Diagram of the Data

Sales Orders.

- → OrderNumber (PK)
- → OrderDate
- → Customer Name Index (FK of Table customers [Customer Index])
- → Channel
- → Currency Code.
- → Warehouse Code
- → Delivery Region Index (FK of Table Regions [id])
- → Product Description Index. (FK of Table Products [index])
- → Order Quantity
- → Unit Price
- → Line Total
- → Total Unit Cost

(a) Customers

- Customer Index (PK)
- Customer Names

Regions

- id (PK)
- name
- Country
- State_code (FK of Table State Regions [State Code])
- State (FK of " " [State]).
- type
- latitude
- longitude
- area_code
- population
- households.
- median_income
- land_area
- water_area
- time_zone

State Regions

- State Code (PK)
- State
- Region

Products

- index (pk)
- Product Name (~~fk of 2017 Budgets (Prod~~ we need.

2017 Budgets.

- Product Name (PK) ~~& fk of Pro~~
- 2017 Budgets.

Now Understanding Relationships.

Orders (O) & Customers. (C)

1 C → min 0 Orders max may Orders.

1 O → min 1C max 1C.

Regions (R) & orders. (ord)

1 R → min 0 Ord max may Ord

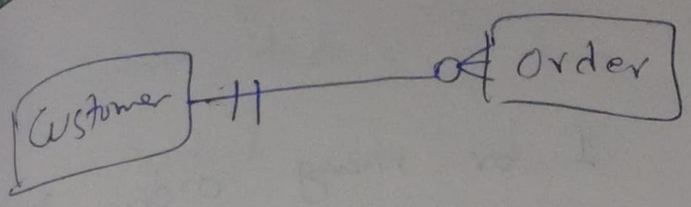
1 O → max 1R min 1R.

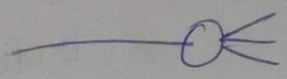
---

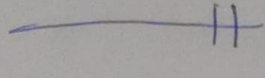
ord (Orders) & Products. (P)

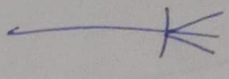
1 O → max may P min 1 Prod.

1 P → max may min 0 ord.



 zero or many

 one (and only one).

 one (or more). (>=1)

1 order can be delivered to one & only one region.

1 Region can have 0 or many orders.

---

1 Region contains 1 & only 1 state_Region

1 State Region can have 0 or many regions.

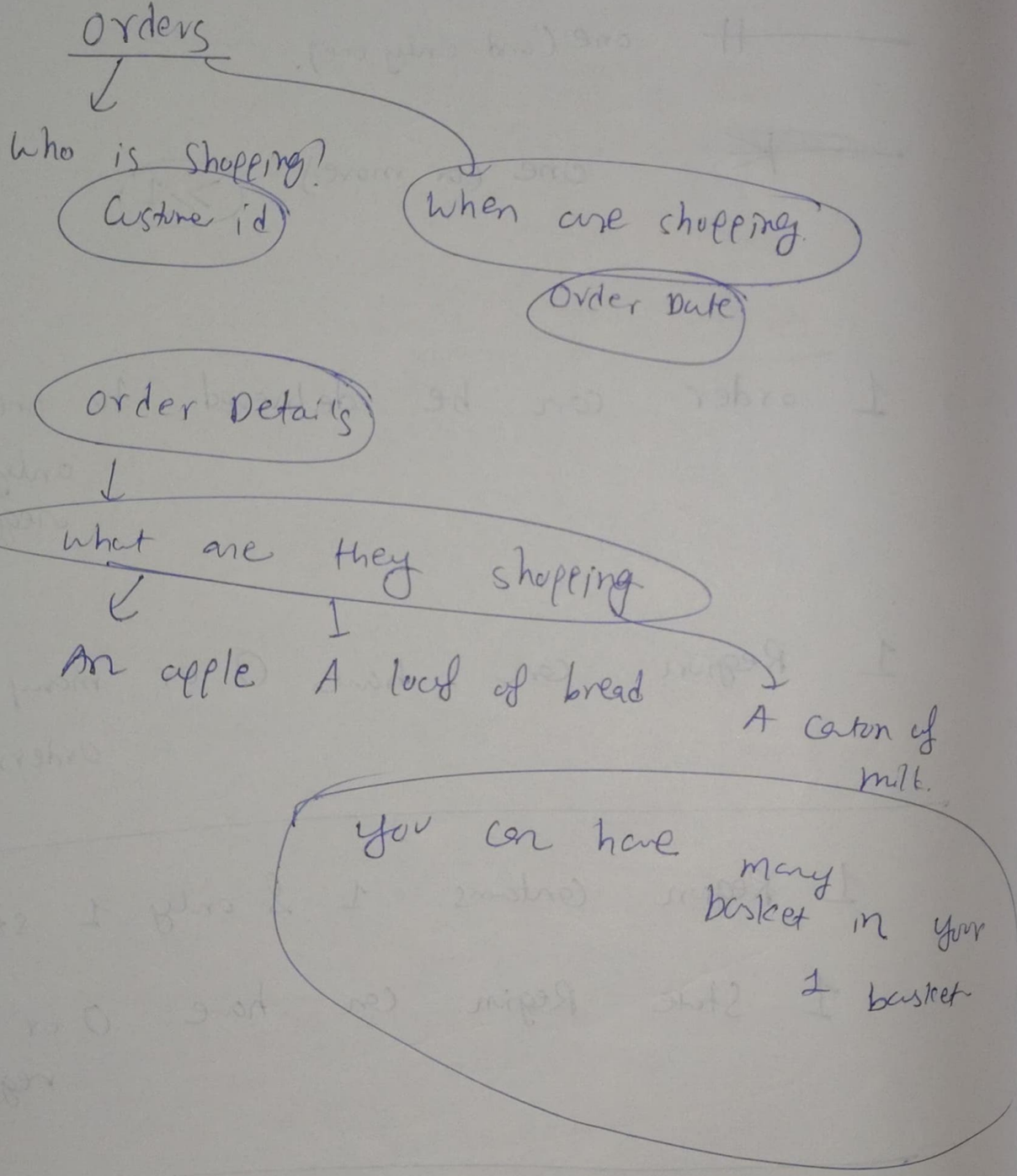
---

1 order can be placed by 1 & only 1 customer

1 Customer can place 0 or many orders.

1 order can have 1 or many order details.

1 order Details can have one & only one order.



1 order-details have 1 and only 1 product
1 product can have many order details.

Products Table is the Product Catalogue.

#1 laptop
#2 Tab
#3 Keyboard.

Budget Table is the Annual Sales.

2017 → $3000 worth of Laptop.
2017 → $7000 worth of Tab.
2018 → $70000 worth of keyboard.

∴ 1 Product has 0 or many budgets,
1 budget has one & only 1 Product.

Exploratory Data Analysis with Pandas Python

Rob Mulla

first learning Pandas from.
Learn Pandas in 30 min

Tech with Tim

```
df.drop(["Location"], axis=1)
```

↙ Column name to be dropped

Suggesting to drop a column & not a row.

Types of Analysis in EDA → Univariate, Bi, Multivariate.

Uni ↙ one feature

Bi ↙ Two feature

Multi → 3 or 4 or more features

Try to Determine output

Working in analysis_notebook_eda.ipynb

↓

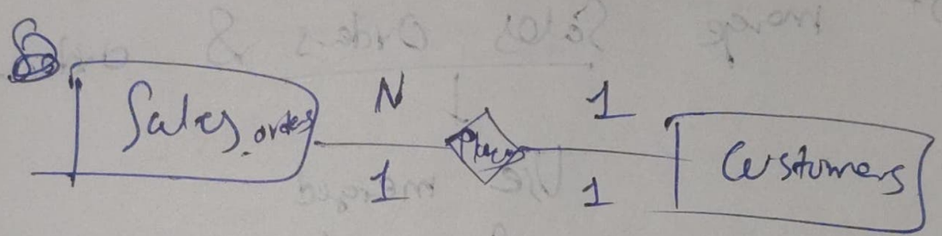
interactive python note book.

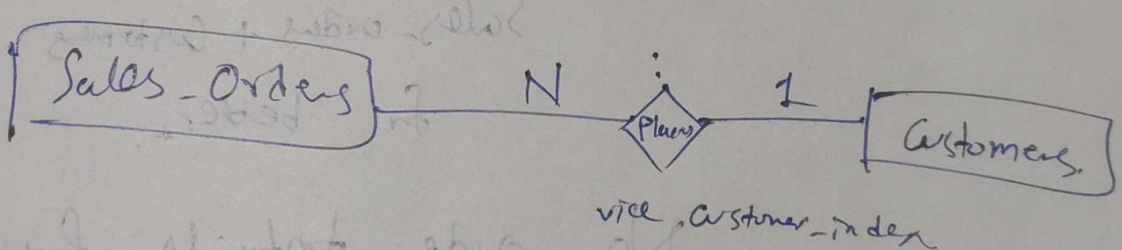
in data Cleaning & wrangling Part.

Here we will try to merge tables.

↓

1st Sales_Orders & Customers tables.



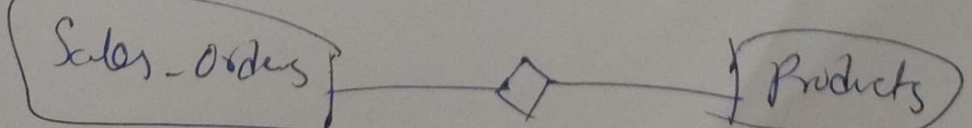


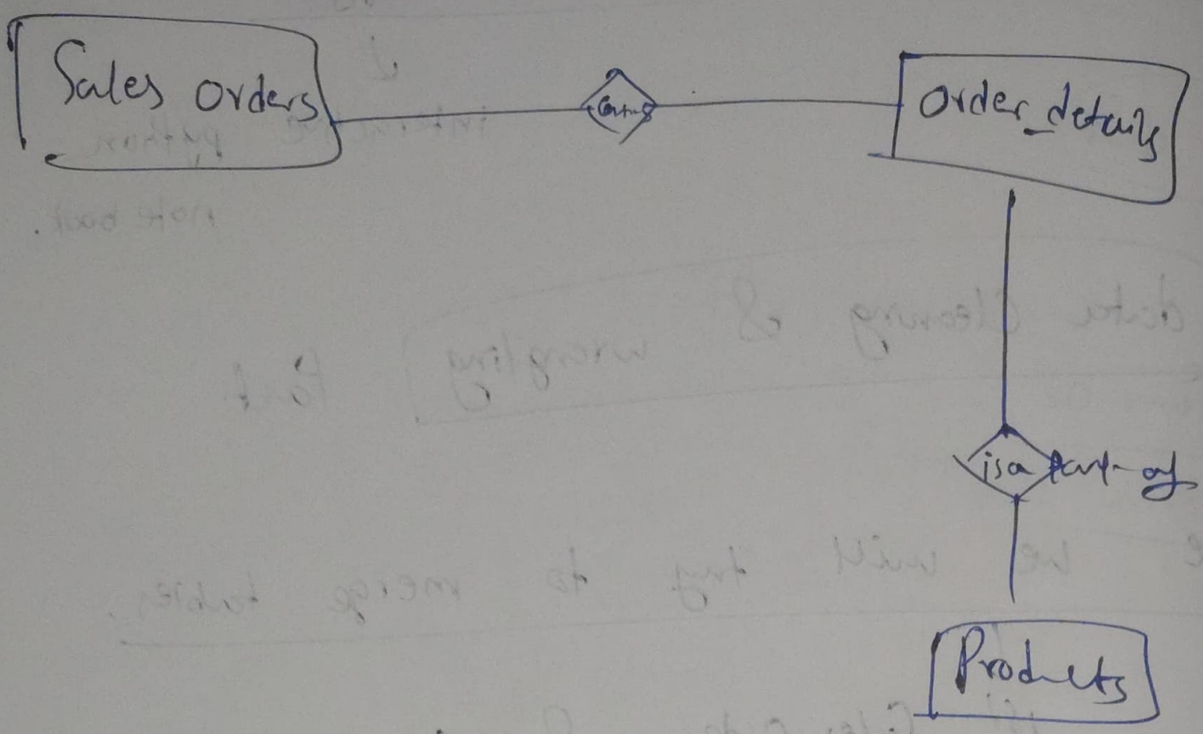
vice. customer_index

Sales_orders, Customers left join

Now we can ignore customers table.

Now Sales_Orders merged with Products





∴ 1st merge Sales orders & order_details.

↓

Use merged form of Sales_orders + Customers for better.

No order details present

Products

merge Sales_orders + Customers

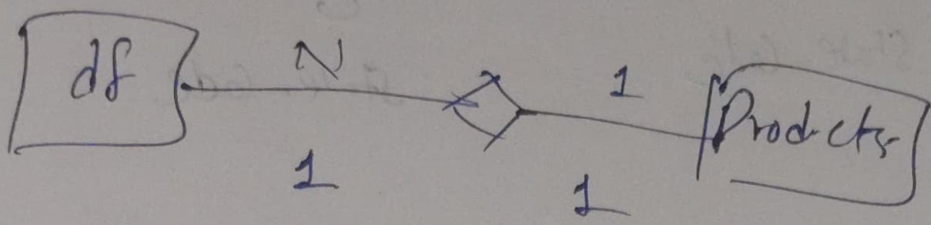
↓

df

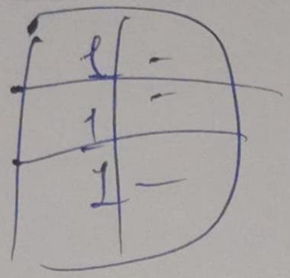
with Products.

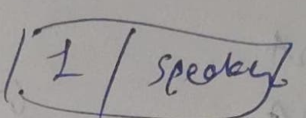
df
↓
Product description
index.

Products
↓
index.



N : 1.



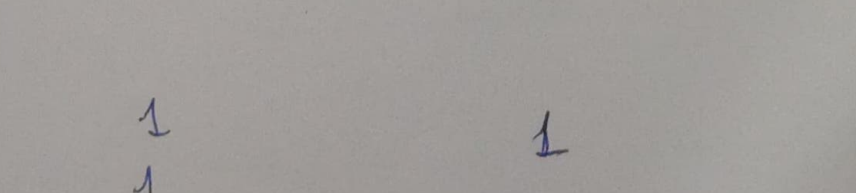


Now we don't need customers & products table.

Now merging with Regions.

df
↓
delivery
region index

Regions
↓
id.

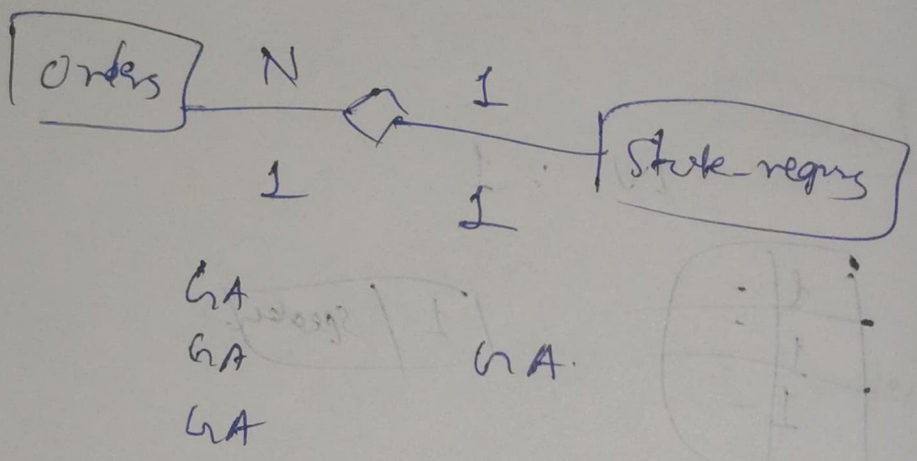


1 1
1
1

Now merging df with State Regions.

df
↓
State-Code

State-regions
↓
State-Code



---

bold underline for Primary key
dotted, dashed underline for
foreign key.