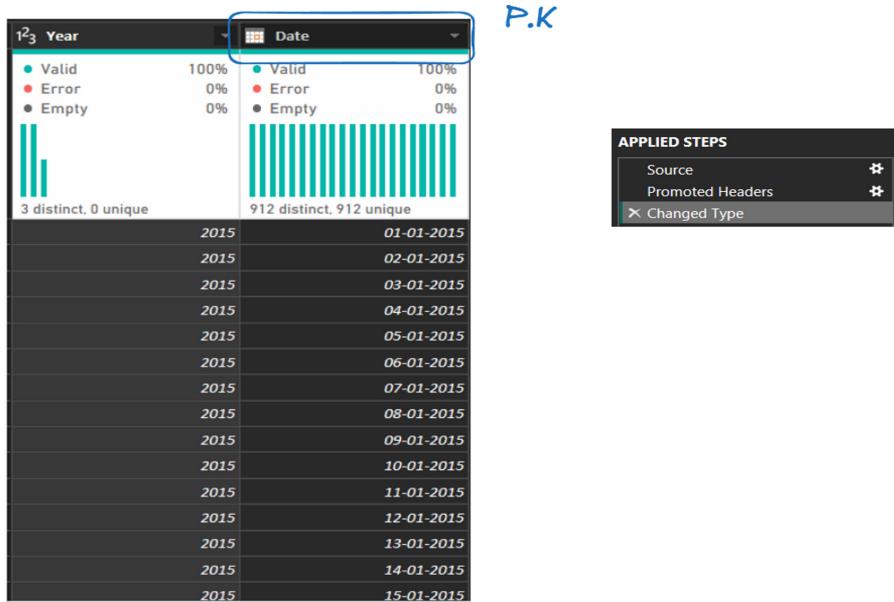
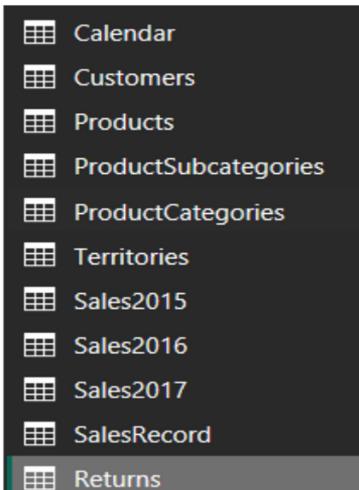


Power BI - End To End Project - 1

- Data Transformation
- Data Modelling

- Bikes Shop - 2015 till 2017 [Historical Data]

- Categories > sub-categories > Products [Hierarchy]
- Sales[Appended] Vs Returns [Fact Table]
- Calendar [Time Intelligence]
- Territory Table [Continent > Country > Region].
- Categories[Bikes, Clothing, Accessories, Component].



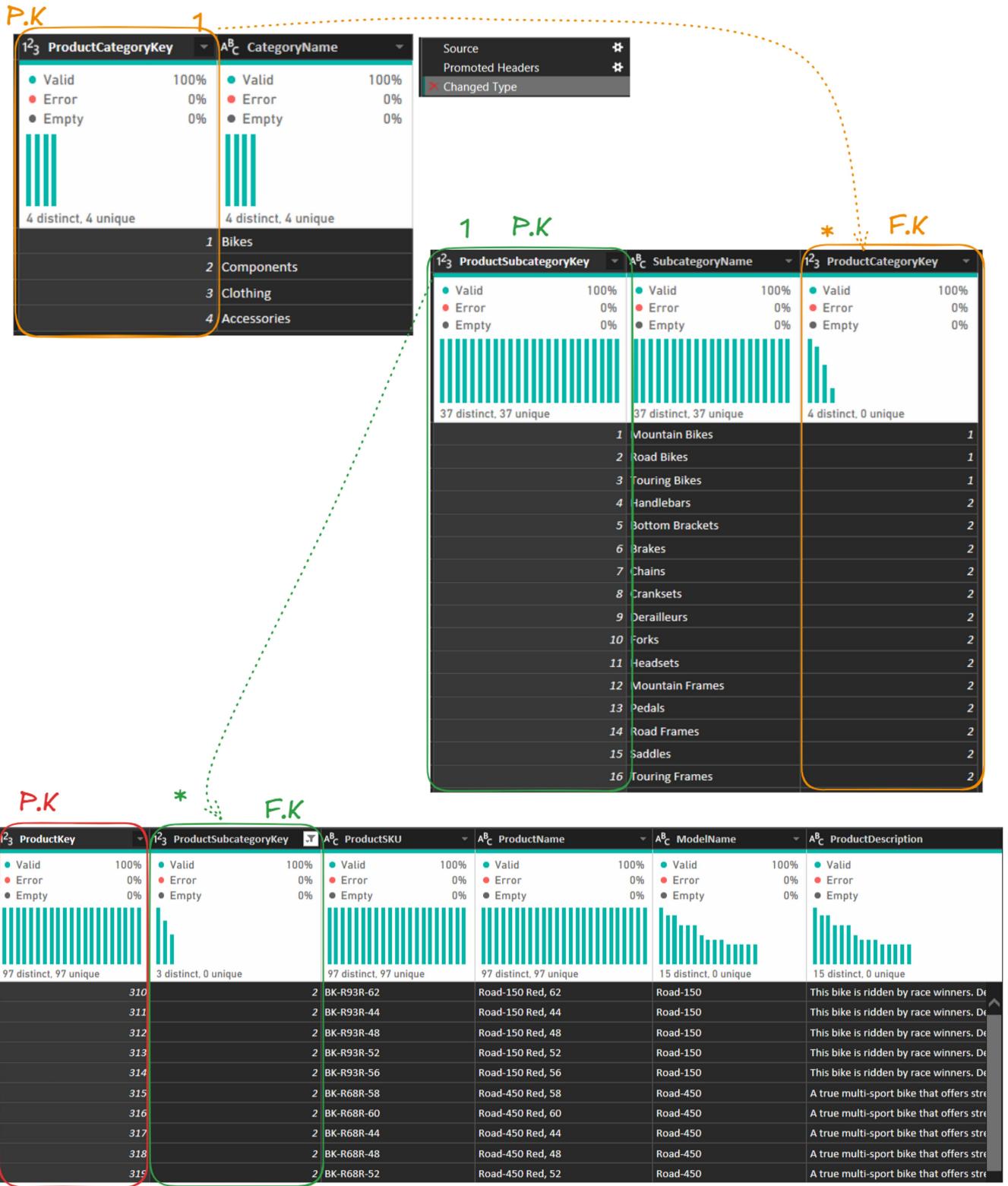
	$\#$ CustomerKey	A^B_C Customer Name	A^B_C BirthDate	A^B_C MaritalStatus	A^B_C Gender	A^B_C EmailAddress	\$ AnnualIncome
1	11000	Mr. Jon Yang	04/08/1966	M	M	jon24@learnsector.com	
2	11001	Mr. Eugene Huang	14/05/1965	S	M	eugene10@learnsector.com	
3	11002	Mr. Ruben Torres	08/12/1965	M	M	ruben35@learnsector.com	
4	11003	Ms. Christy Zhu	15/02/1968	S	F	christy12@learnsector.com	
5	11004	Mrs. Elizabeth Johnson	08/08/1968	S	F	elizabeth5@learnsector.com	
6	11005	Mr. Julio Ruiz	08/05/1965	S	M	julio1@learnsector.com	
7	11006	Mr. Marco Mehta	05/09/1964	M	M	marco14@learnsector.com	
8	11007	Mrs. Robin Verhoff	07/07/1964	S	F	rob4@learnsector.com	
9	11008	Mr. Shannon Carlson	04/01/1964	S	M	shannon38@learnsector.com	
10	11009	Ms. Jacquelyn Suarez	02/06/1964	S	F	jacquelyn20@learnsector.com	
11	11010	Mr. Curtis Lu	11/04/1963	M	M	curtis9@learnsector.com	
12	11011	Mrs. Lauren Walker	18/01/1968	M	F	lauren41@learnsector.com	
13	11012	Mr. Ian Jenkins	08/06/1968	M	M	ian47@learnsector.com	
14	11013	Mrs. Sydney Bennett	05/09/1968	S	F	sydney23@learnsector.com	
15	11014	Ms. Chloe Young	27/02/1979	S	F	chloe23@learnsector.com	
16	11015	Mr. Wyatt Hill	28/04/1979	M	M	wyatt32@learnsector.com	
17	11016	Mrs. Shannon Wang	26/06/1944	S	F	shannon1@learnsector.com	
18	11017	Mr. Clarence Rai	10/09/1944	S	M	clarence32@learnsector.com	
19	11018	Mr. Luke Lal	03/07/1978	S	M	luke18@learnsector.com	

Column From Example

\$ AnnualIncome	$\#$ TotalChildren	A^B_C EducationLevel	A^B_C Occupation	A^B_C HomeOwner	A^B_C Phone_number
90,000.00	2	Bachelors	Professional	Y	6697921661
60,000.00	3	Bachelors	Professional	N	8586376955
60,000.00	3	Bachelors	Professional	Y	6528852745
60,000.00	0	Bachelors	Professional	N	3646568427
80,000.00	5	Bachelors	Professional	Y	7132265883
70,000.00	0	Bachelors	Professional	Y	1902716743
60,000.00	3	Bachelors	Professional	Y	4203325209
60,000.00	4	Bachelors	Professional	Y	2866694333
70,000.00	0	Bachelors	Professional	N	3417265787
70,000.00	0	Bachelors	Professional	N	3166486176
60,000.00	4	Bachelors	Professional	Y	8338877898
1,00,000.00	2	Bachelors	Management	Y	8043834080
1,00,000.00	2	Bachelors	Management	Y	211712173
1,00,000.00	3	Bachelors	Management	N	435758409
1,00,000.00	0	Partial College	Skilled Manual	N	7907467471
30,000.00	0	Partial College	Skilled Manual	Y	6493845387
20,000.00	4	High School	Skilled Manual	Y	3314308824

APPLIED STEPS	
Source	☒
Promoted Headers	☒
Removed Columns	☒
Changed Type	☒
Replaced Value	☒
Replaced Value1	☒
Changed Type1	☒
Filled Down	☒
Changed Type2	☒
Inserted Merged Column	☒
Removed Columns1	☒
Reordered Columns	☒

Replace Value - "\$" " " → fixed Decimal Places



A ^B _C ProductColor	A ^B _C ProductSize	A ^B _C ProductStyle	\$ ProductCost	\$ ProductPrice
● Valid 100%	● Valid 100%	● Valid 100%	● Valid 100%	● Valid 100%
● Error 0%	● Error 0%	● Error 0%	● Error 0%	● Error 0%
● Empty 0%	● Empty 0%	● Empty 0%	● Empty 0%	● Empty 0%
5 distinct, 0 unique	13 distinct, 1 unique	2 distinct, 0 unique	19 distinct, 0 unique	18 distinct, 0 unique
Red	62	U	2,171.29	3,578.27
Red	44	U	2,171.29	3,578.27
Red	48	U	2,171.29	3,578.27
Red	52	U	2,171.29	3,578.27
Red	56	U	2,171.29	3,578.27
Red	58	U	884.71	1,457.99
Red	60	U	884.71	1,457.99

Hierarchy

Product Category -> Product SubCategory -> Products

P.K

I ² ₃ SalesTerritoryKey	A ^B _C Region	A ^B _C Country	A ^B _C Continent
● Valid 100%	● Valid 100%	● Valid 100%	● Valid 100%
● Error 0%	● Error 0%	● Error 0%	● Error 0%
● Empty 0%	● Empty 0%	● Empty 0%	● Empty 0%
10 distinct, 10 unique	10 distinct, 10 unique	6 distinct, 5 unique	3 distinct, 1 unique
1 Northwest	United States	North America	
2 Northeast	United States	North America	
3 Central	United States	North America	
4 Southwest	United States	North America	
5 Southeast	United States	North America	
6 Canada	Canada	North America	
7 France	France	Europe	
8 Germany	Germany	Europe	
9 Australia	Australia	Pacific	
10 United Kingdom	United Kingdom	Europe	

Text??

- Map Visuals
- Categories according to the map visuals.

Append

Concatenate rows from three or more tables into a single table.

Two tables Three or more tables

Available tables

- Calendar
- Customers
- Products
- ProductSubcategories
- ProductCategories
- Territories
- Sales2015
- Sales2016
- Sales2017

Tables to append

- Sales2015
- Sales2016
- Sales2017

F.K

Invoice Number

F.K

OrderDate	StockDate	OrderNumber	ProductKey	CustomerKey	TerritoryKey	OrderLineItem	OrderQuantity
Valid Error Empty	Valid Error Empty	Valid Error Empty	Valid Error Empty	Valid Error Empty	Valid Error Empty	Valid Error Empty	Valid Error Empty
158 distinct, 2 unique	231 distinct, 33 unique	1000 distinct, 1000 unique	25 distinct, 0 unique	1000 distinct, 1000 unique	7 distinct, 0 unique	1 distinct, 0 unique	1 distinct, 0 unique
01-01-2015	21-09-2001	SO45080	332	14657	1	1	1
01-01-2015	05-12-2001	SO45079	312	29255	4	1	1
01-01-2015	29-10-2001	SO45082	350	11455	9	1	1
01-01-2015	16-11-2001	SO45081	338	26782	6	1	1
02-01-2015	15-12-2001	SO45083	312	14947	10	1	1
02-01-2015	12-10-2001	SO45084	310	29143	4	1	1
02-01-2015	18-12-2001	SO45086	314	18747	9	1	1
02-01-2015	09-10-2001	SO45085	312	18746	9	1	1
03-01-2015	03-10-2001	SO45093	312	18906	9	1	1
03-01-2015	29-09-2001	SO45090	310	29170	4	1	1
03-01-2015	11-12-2001	SO45088	345	11398	10	1	1
03-01-2015	24-10-2001	SO45092	313	18899	9	1	1
03-01-2015	16-12-2001	SO45089	351	25977	4	1	1
03-01-2015	26-10-2001	SO45091	314	18909	9	1	1
03-01-2015	11-09-2001	SO45087	350	11388	10	1	1
03-01-2015	11-09-2001	SO45094	310	22785	6	1	1
04-01-2015	30-10-2001	SO45096	312	12483	7	1	1
04-01-2015	30-10-2001	SO45097	313	29151	4	1	1

F.K

ReturnDate	T ₃ TerritoryKey	T ₃ ProductKey	T ₃ ReturnQuantity
Valid Error Empty	Valid Error Empty	Valid Error Empty	Valid Error Empty
325 distinct, 128 unique	7 distinct, 0 unique	114 distinct, 23 unique	2 distinct, 0 unique
18-01-2015	9	312	1
18-01-2015	10	310	1
21-01-2015	8	346	1
22-01-2015	4	311	1
02-02-2015	6	312	1
15-02-2015	1	312	1
19-02-2015	9	311	1
24-02-2015	8	314	1
08-03-2015	8	350	1
13-03-2015	9	350	1
14-03-2015	4	346	1
15-03-2015	9	340	1
22-03-2015	4	311	1
26-03-2015	10	312	1
28-03-2015	7	312	1
28-03-2015	9	314	1
29-03-2015	9	311	1
01-04-2015	8	311	1
07-04-2015	9	311	1
07-04-2015	9	351	1

Fix all the dates using Locale.

Change Type with Locale

Change the data type and select the locale of origin.

Data Type: Date

Locale: English (United States)

Sample input values:
 3/29/2016
 Tuesday, March 29, 2016
 March 29
 March 2016

OK **Cancel**

Is a Data Model?

There is no "Relationship exist".
So, its not a data model.

CategoryName	Sum of OrderQuantity
Accessories	84174
Bikes	84174
Clothing	84174
Components	84174
Total	84174

Screenshot of a data catalog interface showing various tables:

- Calendar**: Date, Year
- Customers**: AnnualIncome, BirthDate, CustomerName, CustomerKey, EducationLevel, EmailAddress, Gender, HomeOwner, MaritalStatus
- Territories**: Continent, Country, Region, SalesTerritoryKey
- Products**: ModelName, ProductColor, ProductCost, ProductDescription, ProductKey, ProductName, ProductPrice, ProductSize, ProductSKU
- ProductSubcategories**: ProductCategoryKey, ProductSubCategoryKey, SubcategoryName
- ProductCategories**: CategoryName
- SalesRecord**: CustomerKey, OrderDate, OrderLineItem, OrderNumber, OrderQuantity, ProductKey, StockDate, TerritoryKey
- Returns**: ProductKey, ReturnDate, ReturnQuantity, TerritoryKey
- Sales2015**, **Sales2016**, **Sales2017**: CustomerKey, OrderDate, OrderLineItem, OrderNumber, OrderQuantity, ProductKey, StockDate, TerritoryKey

Report View (Search: productKey)

- Products: \sum ProductKey
- Returns: \sum ProductKey
- Sales2015: \sum ProductKey
- Sales2016: \sum ProductKey
- Sales2017: \sum ProductKey
- SalesRecord: \sum ProductKey

unnecessary records coming multiple time, make the user confuse at some point.
So, its better to hide the referenced table.

Report View (Search: productKey)

- Products: \sum ProductKey
- Returns: \sum ProductKey
- SalesRecord: \sum ProductKey

Sales2015

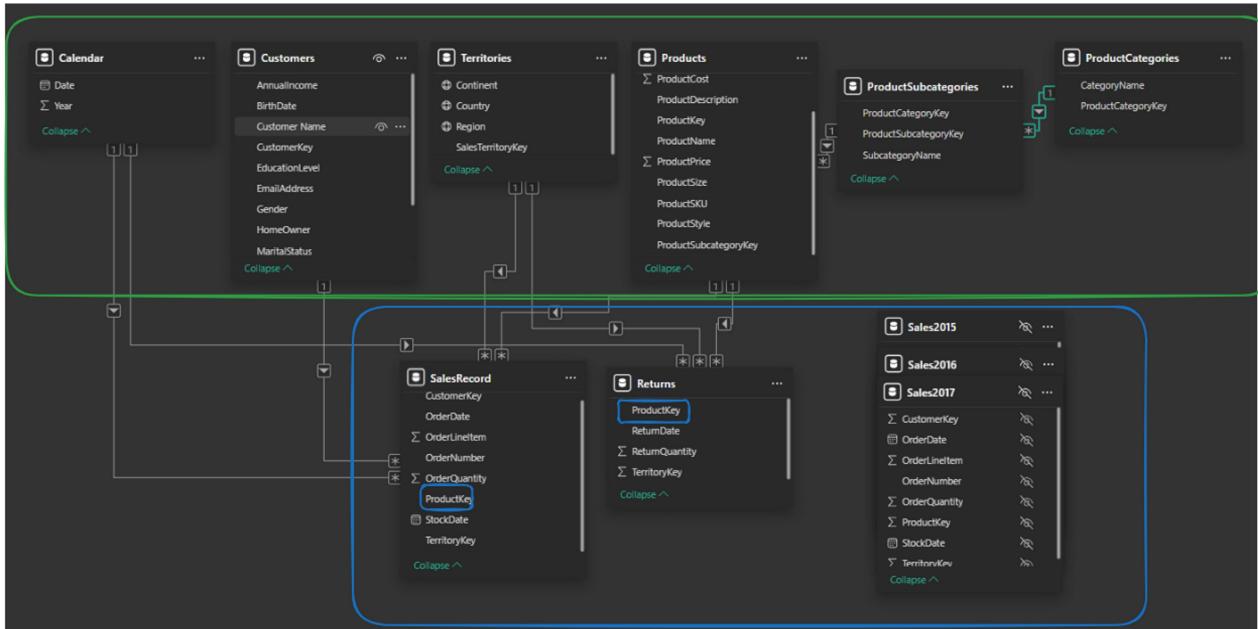
Sales2016

Sales2017

\sum CustomerKey
OrderDate
\sum OrderLineItem
OrderNumber
\sum OrderQuantity
ProductKey
StockDate
\sum TerritoryKey

Snow Flake Schema

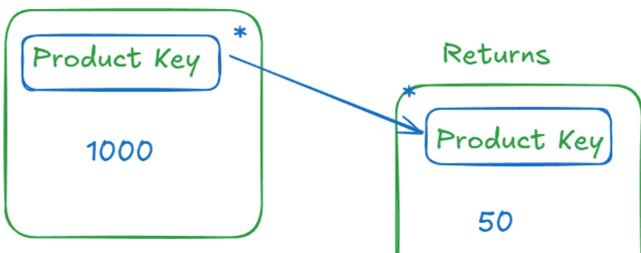
Dimension Table



Fact Table - Transaction Record \rightarrow Foreign Key

1. Star Schema \rightarrow Atleast one Fact Table with multi Dimension Table.
2. Snowflake Schema \rightarrow Extended part of Star Schema. [Star + sub-dimension Table]

Sales



Many To Many Cardinality

Final Table after joining



50K Records - $1000 * 50$