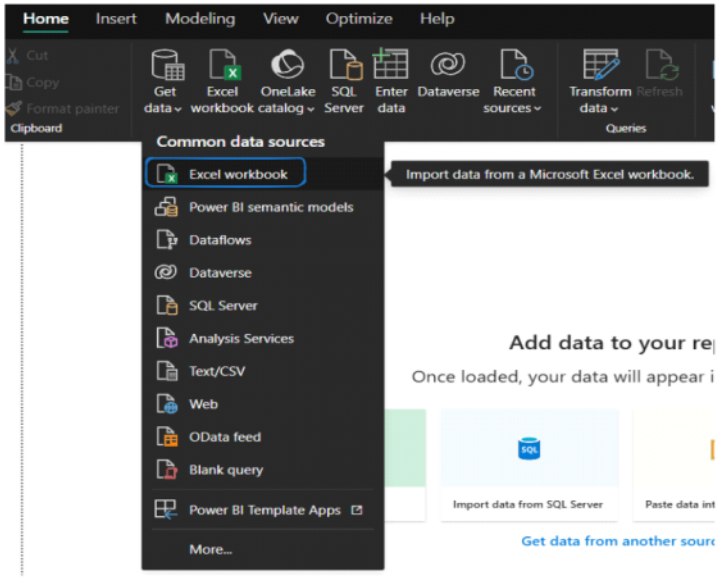
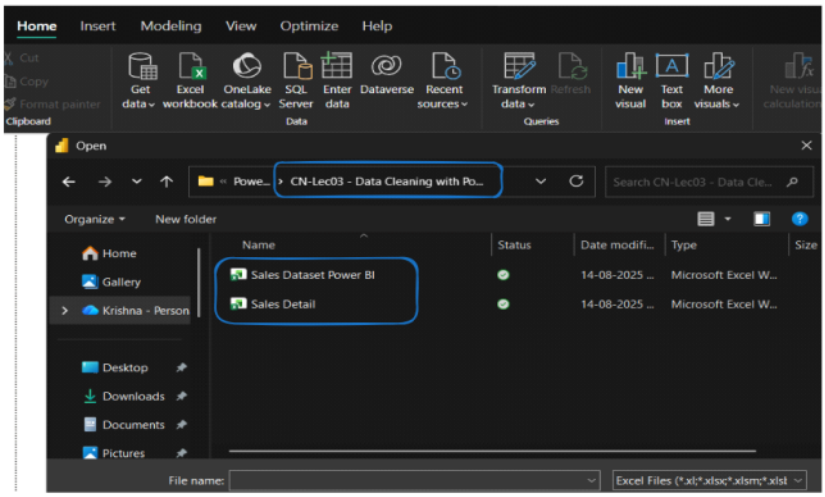
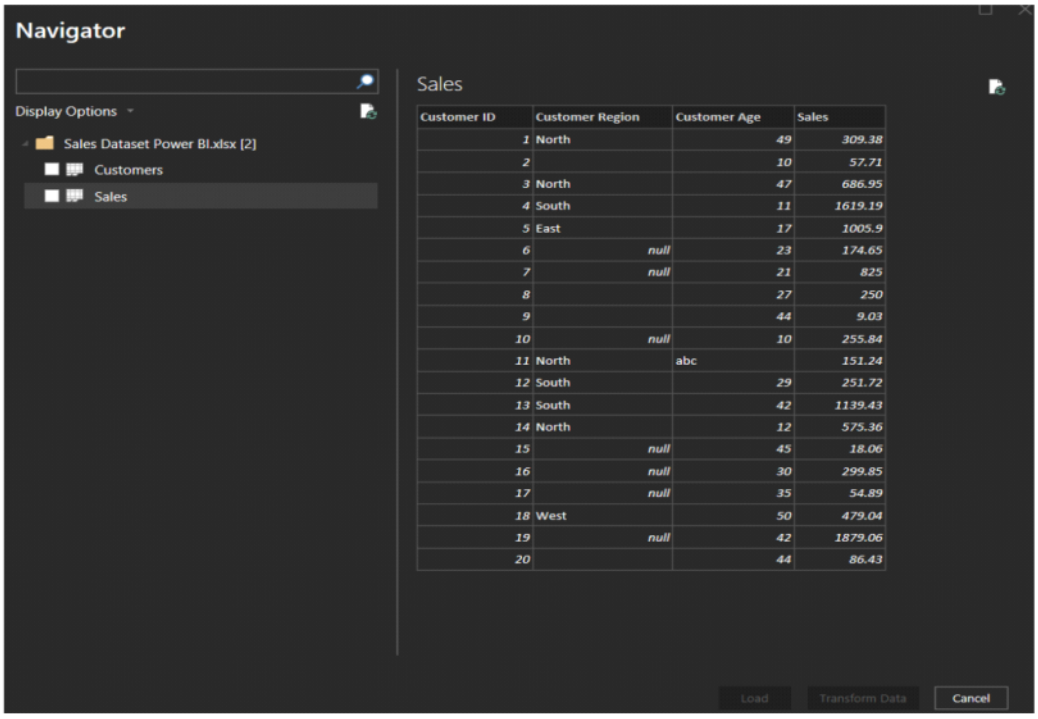
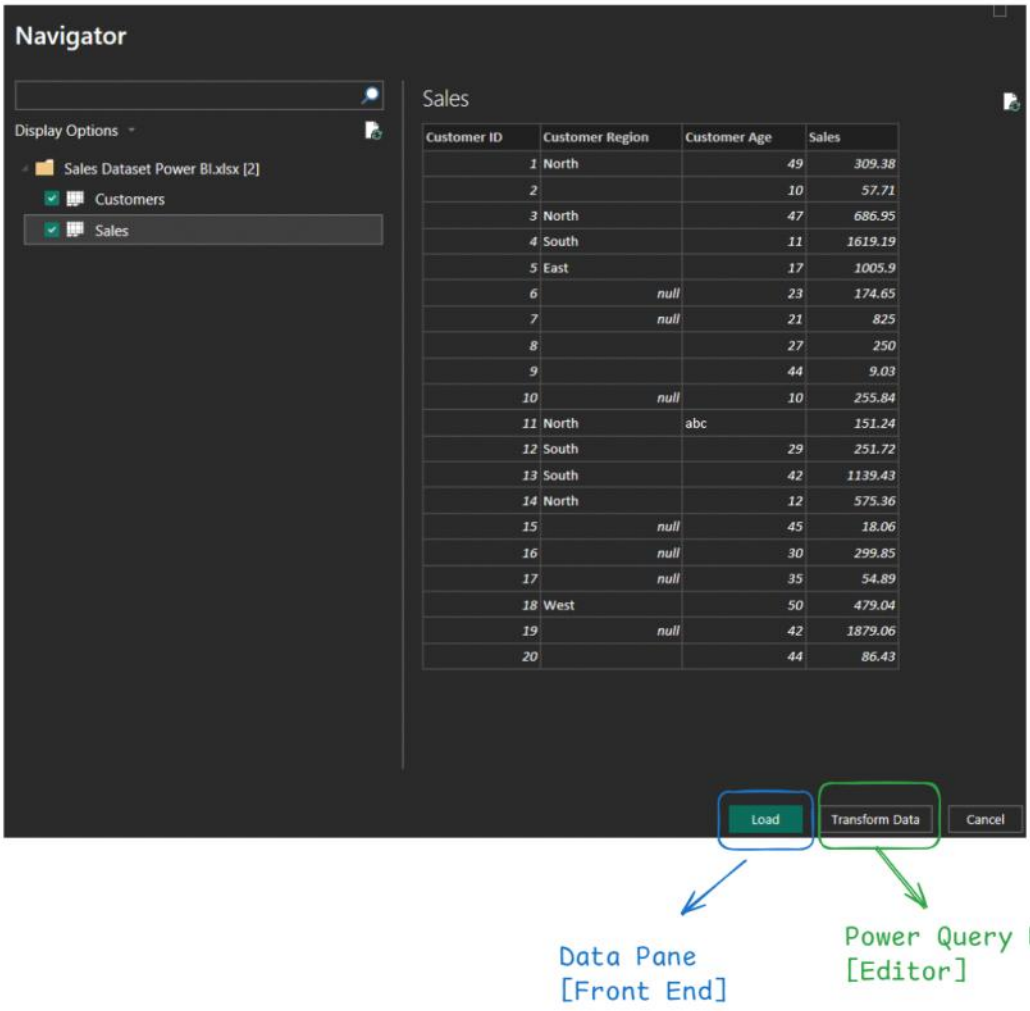
Data Cleaning with Power BI







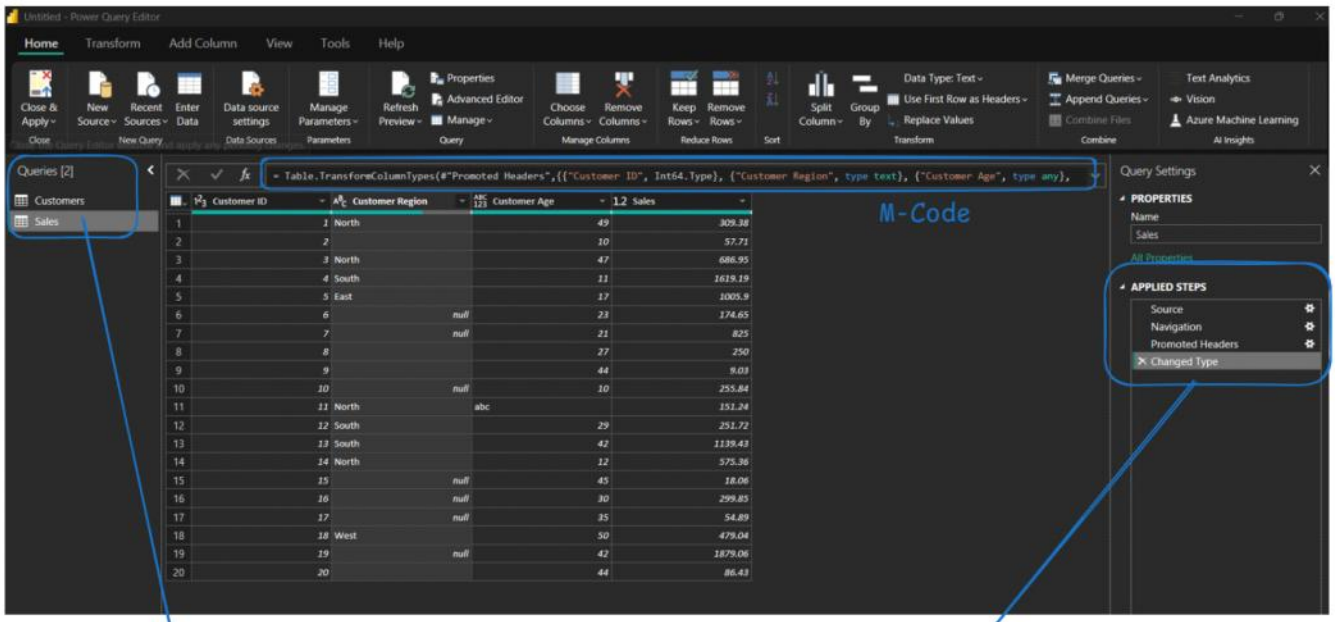
Customer ID	Customer Region	Customer Age	Sales
1	North	49	309.38
2		10	57.71
3	North	47	686.95
4	South	11	1619.19
5	East	17	1005.9
6	null	23	174.65
7	null	21	825
8		27	250
9		44	9.03
10	null	10	255.84
11	North	abc	151.24
12	South	29	251.72
13	South	42	1139.43
14	North	12	575.36
15	null	45	18.06
16	null	30	299.85
17	null	35	54.89
18	West	50	479.04
19	null	42	1879.06
20		44	86.43



Customer ID	Customer Region	Customer Age	Sales
1	North	49	309.38
2		10	57.71
3	North	47	686.95
4	South	11	1619.19
5	East	17	1005.9
6	null	23	174.65
7	null	21	825
8		27	250
9		44	9.03
10	null	10	255.84
11	North	abc	151.24
12	South	29	251.72
13	South	42	1139.43
14	North	12	575.36
15	null	45	18.06
16	null	30	299.85
17	null	35	54.89
18	West	50	479.04
19	null	42	1879.06
20		44	86.43

Data Pane
[Front End]

Power Query Editor
[Editor]



M-Code

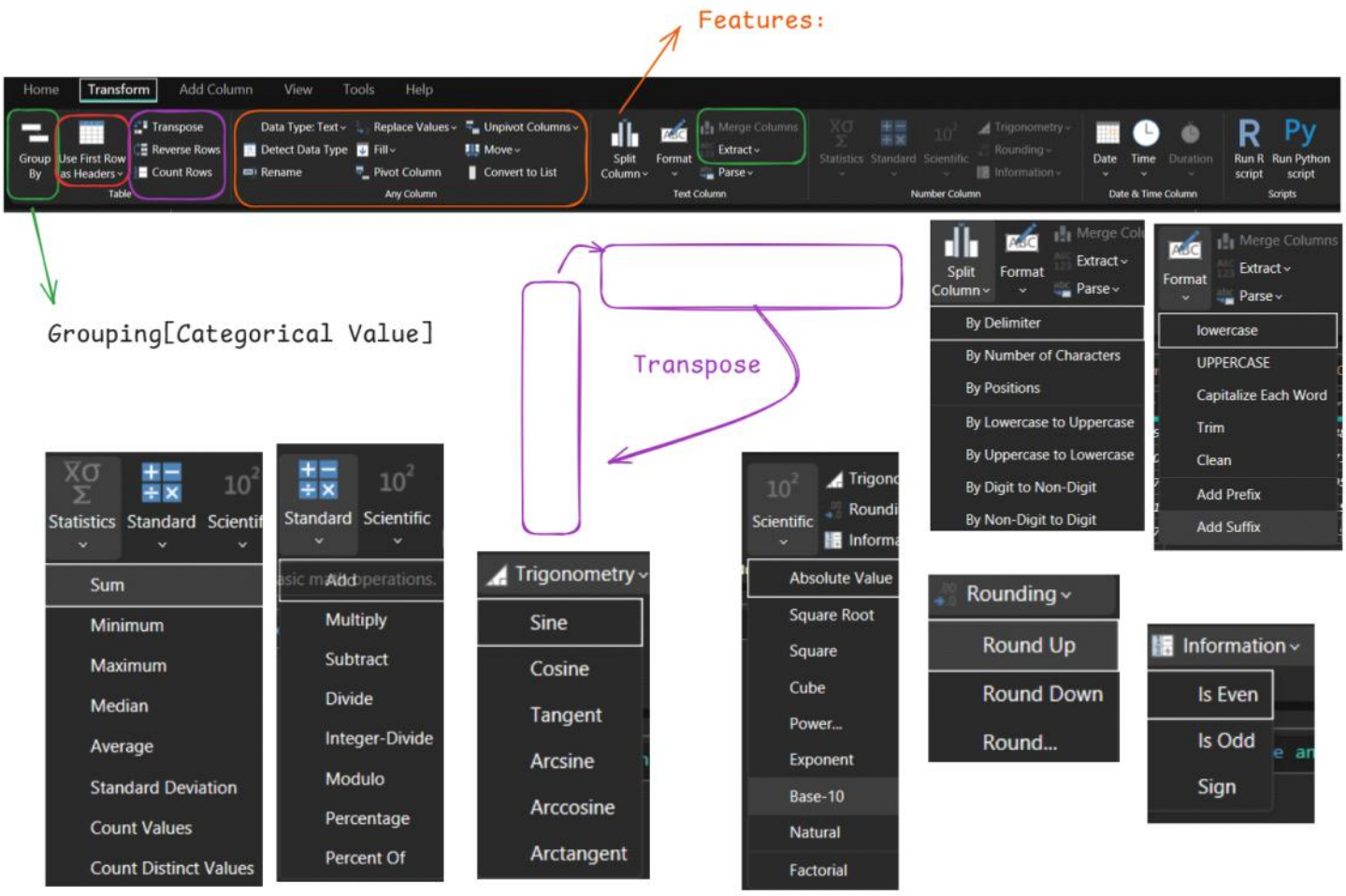
Shows Table..

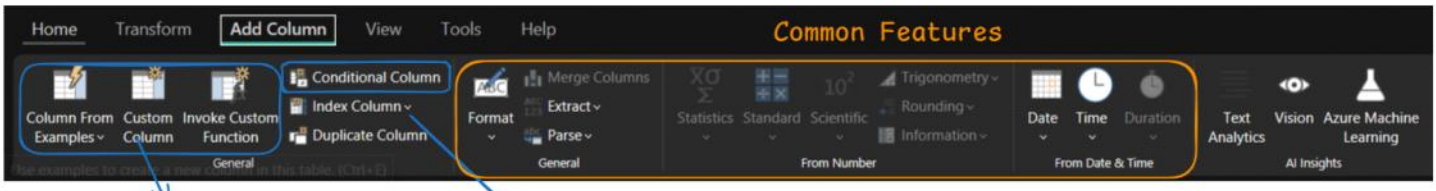
Steps performed
while cleaning
the dataset

Transform Vs Add Column

New Column

Existing Column

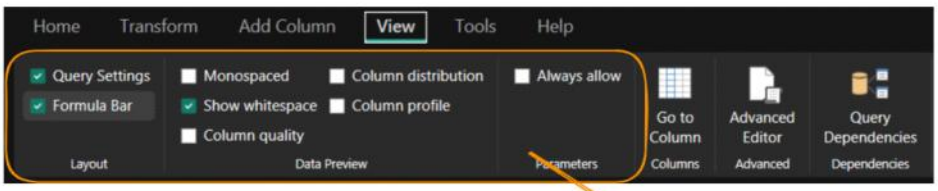


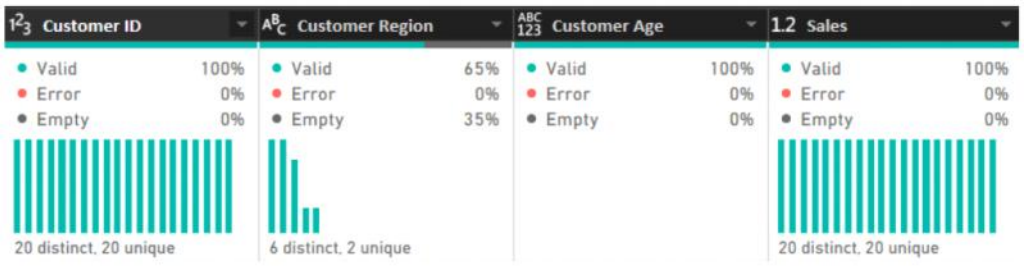


New Column Pattern

If-else

Start with 0/1/custom





at least 1

Sports

Cricket Badminton Cricket Volleyball Hockey Badminton Cricket Kabaddi Volleyball Volleyball Badminton BasketBall Cricket Football BasketBall Kabaddi Badminton Football Cricket BasketBall Golf

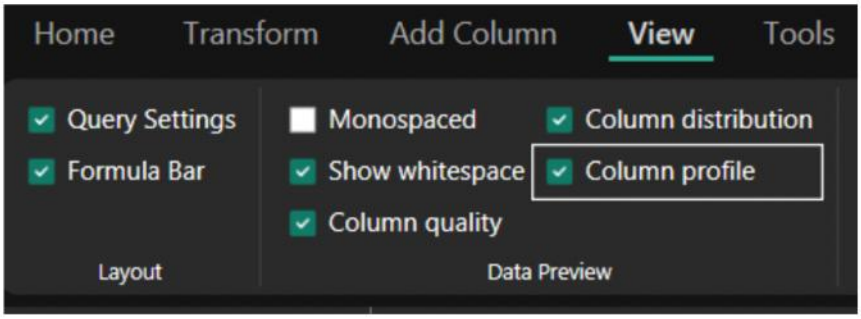
Distinct 8

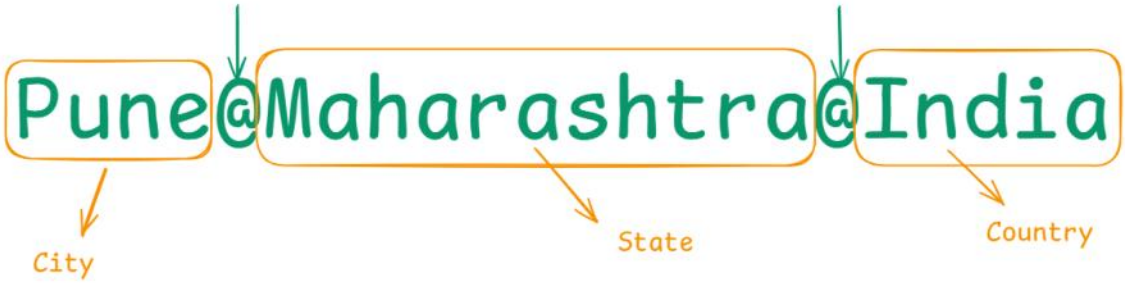
Cricket
Badminton
Kabaddi
Volleyball
Golf
Hockey
BasketBall
Football

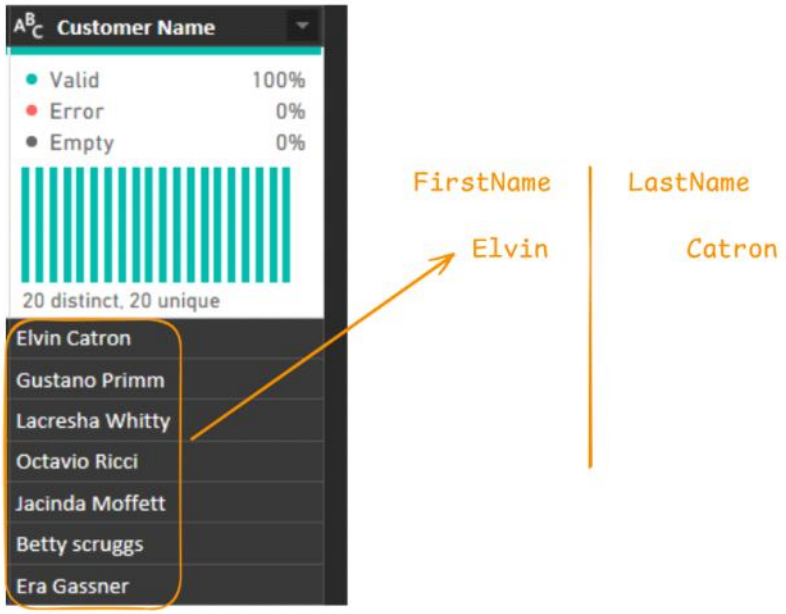
Cricket - 5
Badminton - 4
Kabaddi - 2
Volleyball - 3
Golf - 1
Hockey - 1
Basketball - 3
Football - 2

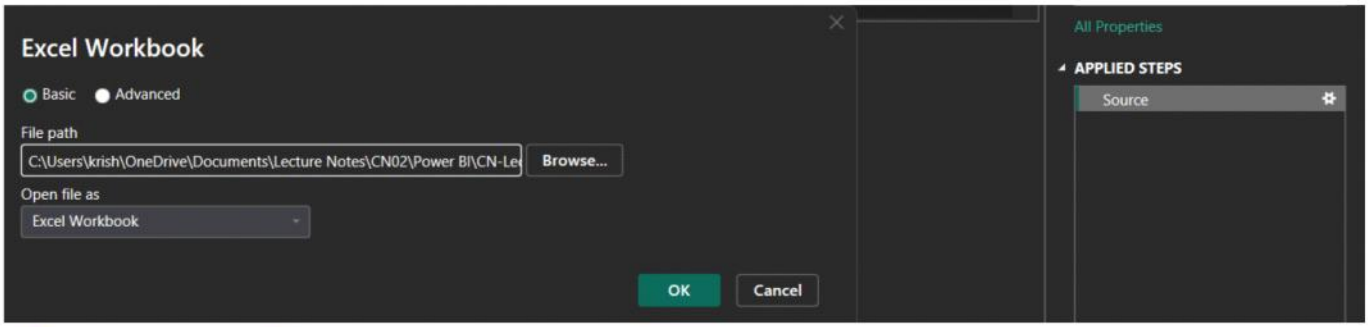
Unique 2

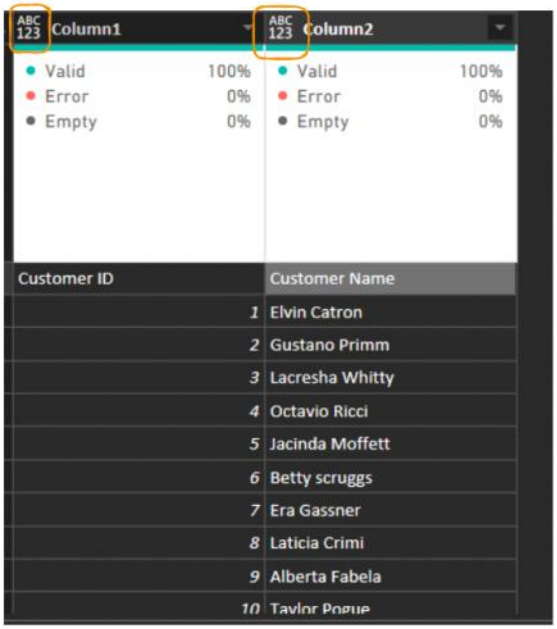
Golf exactly 1
Hockey

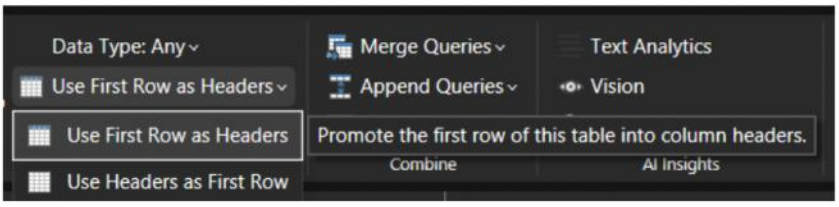


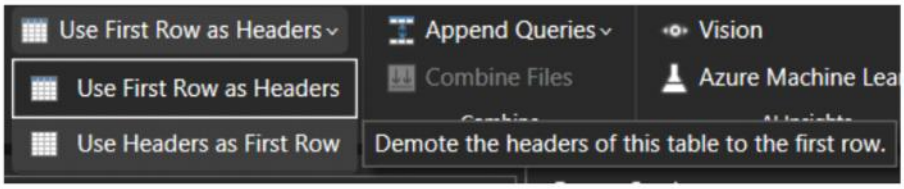


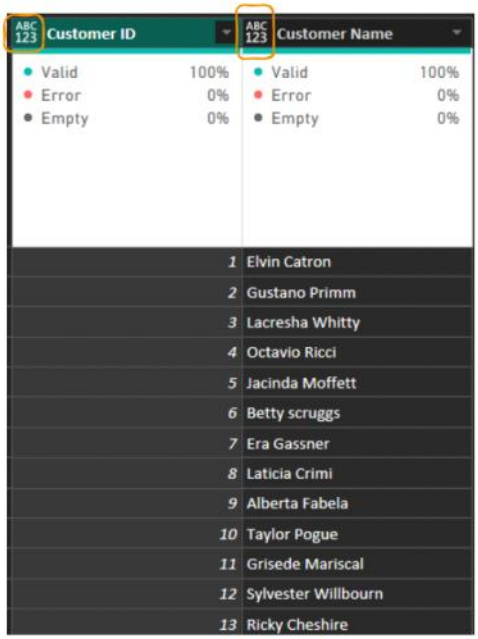

Customer Name
Valid 100%
Error 0%
Empty 0%
20 distinct, 20 unique
Elvin Catron
Gustano Primm
Lacresha Whitty
Octavio Ricci
Jacinda Moffett
Betty scruggs
Era Gassner
FirstName
LastName
Elvin
Catron

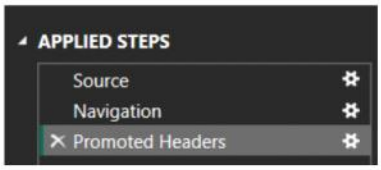

Excel Workbook
Basic
Advanced
File path
C:\Users\krish\OneDrive\Documents\Lecture Notes\CN02\Power BI\CN-Le
Browse...
Open file as
Excel Workbook
OK
Cancel
All Properties
APPLIED STEPS
Source

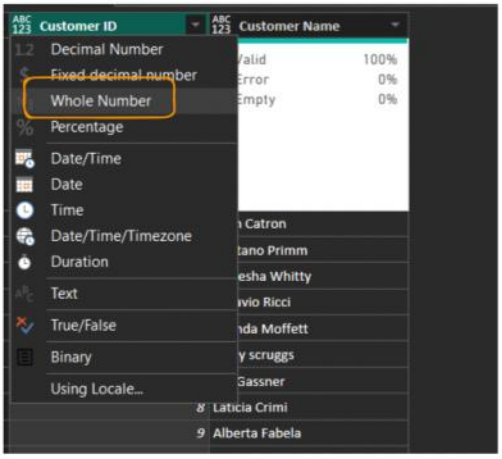

Column1
Column2
Valid 100%
Error 0%
Empty 0%
Customer ID
Customer Name
1 Elvin Catron
2 Gustano Primm
3 Lacresha Whitty
4 Octavio Ricci
5 Jacinda Moffett
6 Betty scruggs
7 Era Gassner
8 Laticia Crimi
9 Alberta Fabela
10 Taylor Pogue

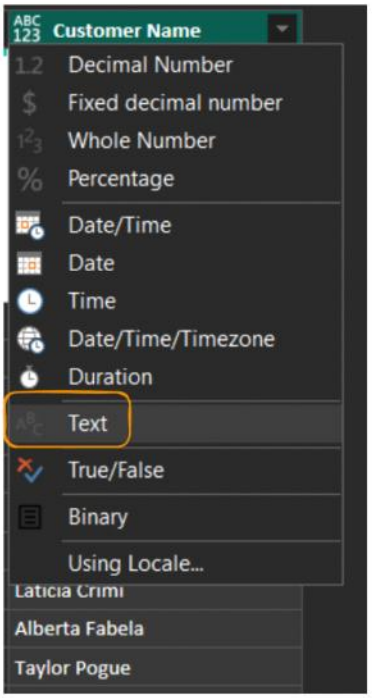

Data Type: Any
Use First Row as Headers
Merge Queries
Append Queries
Text Analytics
Vision
Use First Row as Headers
Use Headers as First Row
Promote the first row of this table into column headers.
Combine
AI Insights

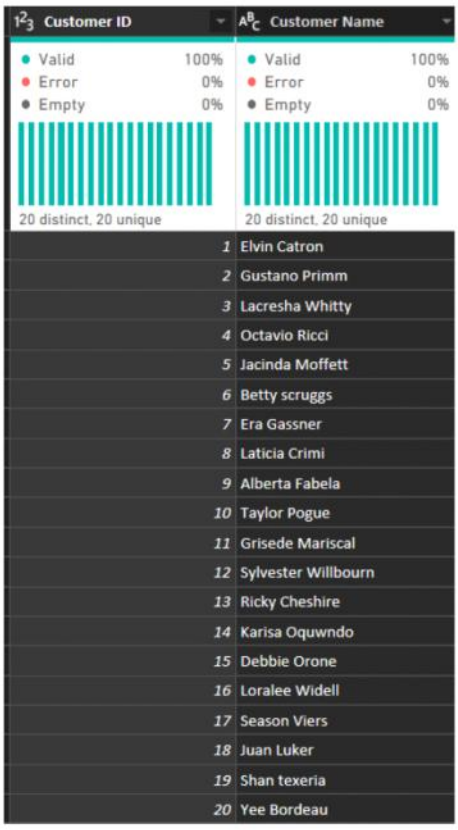

Use First Row as Headers
Append Queries
Vision
Use First Row as Headers
Combine Files
Azure Machine Lea
Use Headers as First Row
Demote the headers of this table to the first row.

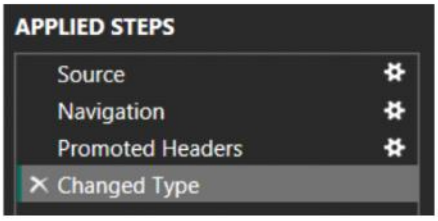

Customer ID
Customer Name
Valid 100%
Error 0%
Empty 0%
1 Elvin Catron
2 Gustano Primm
3 Lacresha Whitty
4 Octavio Ricci
5 Jacinda Moffett
6 Betty scruggs
7 Era Gassner
8 Laticia Crimi
9 Alberta Fabela
10 Taylor Pogue
11 Grisede Mariscal
12 Sylvester Willbourn
13 Ricky Cheshire


APPLIED STEPS
Source
Navigation
Promoted Headers


Customer ID
Customer Name
Decimal Number
Fixed decimal number
Whole Number
Percentage
Date/Time
Date
Time
Date/Time/Timezone
Duration
Text
True/False
Binary
Using Locale...
8 Laticia Crimi
9 Alberta Fabela


Customer Name
Decimal Number
Fixed decimal number
Whole Number
Percentage
Date/Time
Date
Time
Date/Time/Timezone
Duration
Text
True/False
Binary
Using Locale...
Laticia Crimi
Alberta Fabela
Taylor Pogue


Customer ID
Customer Name
Valid 100%
Error 0%
Empty 0%
20 distinct, 20 unique
20 distinct, 20 unique
1 Elvin Catron
2 Gustano Primm
3 Lacresha Whitty
4 Octavio Ricci
5 Jacinda Moffett
6 Betty scruggs
7 Era Gassner
8 Laticia Crimi
9 Alberta Fabela
10 Taylor Pogue
11 Grisede Mariscal
12 Sylvester Willbourn
13 Ricky Cheshire
14 Karisa Oquwndo
15 Debbie Orone
16 Loralee Widell
17 Season Viers
18 Juan Luker
19 Shan texeria
20 Yee Bordeau


APPLIED STEPS
Source
Navigation
Promoted Headers
Changed Type

Column1	Column2	Column3	Column4
Valid 100%	Valid 67%	Valid 100%	Valid 100%
Error 0%	Error 0%	Error 0%	Error 0%
Empty 0%	Empty 33%	Empty 0%	Empty 0%

Customer ID	Customer Region	Customer Age	Sales
1	North	49	309.38
2		10	57.71
3	North	47	686.95
4	South	11	1619.19
5	East	17	1005.9
6	null	23	174.65
7	null	21	825
8		27	250
9		44	9.03
10	null	10	255.84
11	North	abc	151.24
12	South	29	251.72
13	South	42	1139.43
14	North	12	575.36
15	null	45	18.06
16	null	30	299.85
17	null	35	54.89
18	West	50	479.04
19	null	42	1879.06

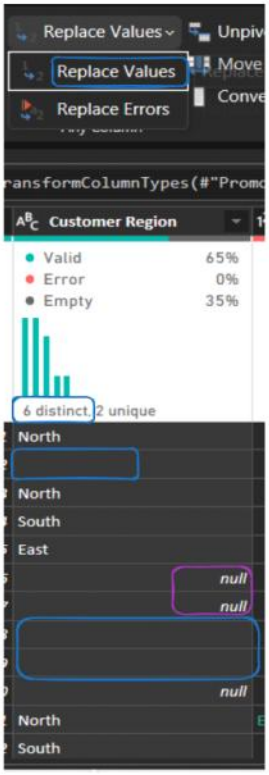
Promoting Headers

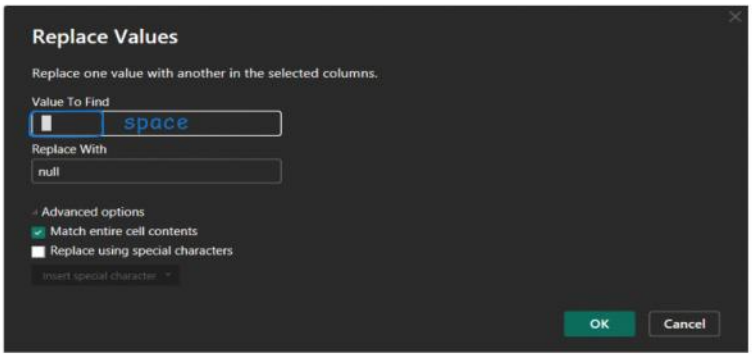
Customer ID	Customer Region	Customer Age	Sales
Valid 100% Error 0% Empty 0% 20 distinct, 20 unique	Valid 65% Error 0% Empty 35% 6 distinct, 2 unique	Valid - % Error 5% Empty - %	Valid 100% Error 0% Empty 0% 20 distinct, 20 unique
1	North	49	309.38
2		10	57.71
3	North	47	686.95
4	South	11	1,619.19
5	East	17	1,005.90
6	null	23	174.65
7	null	21	825.00
8		27	250.00
9		44	9.03
10	null	10	255.84
11	North	Error	151.24
12	South	29	251.72
13	South	42	1,139.43
14	North	12	575.36
15	null	45	18.06
16	null	30	299.85
17	null	35	54.89
18	West	50	479.04
19	null	42	1,879.06
20		44	86.43

P.K

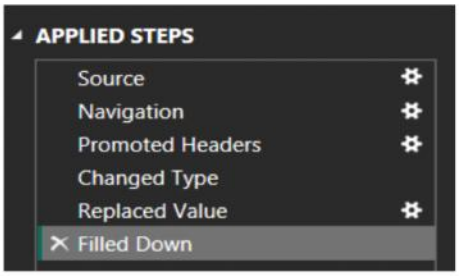
Fill_Down

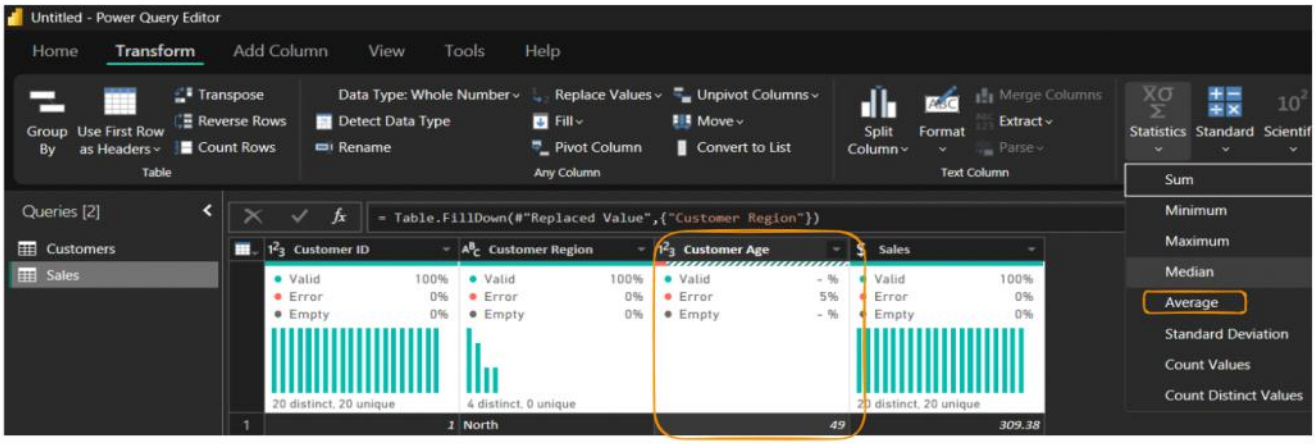
'abc' = abc[❌❌]

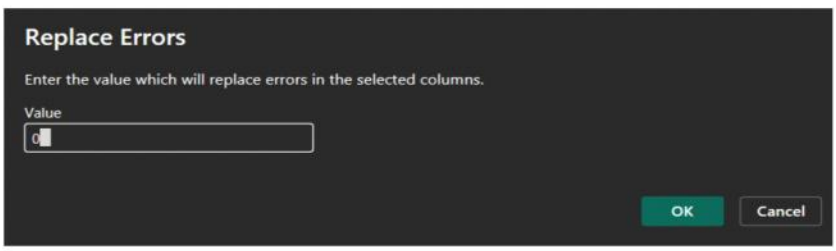


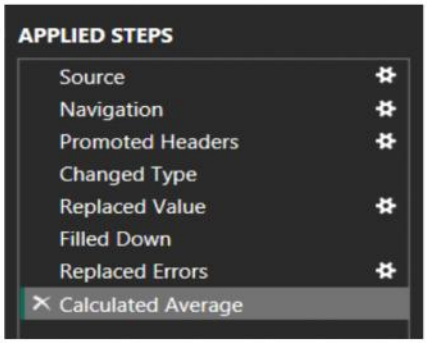


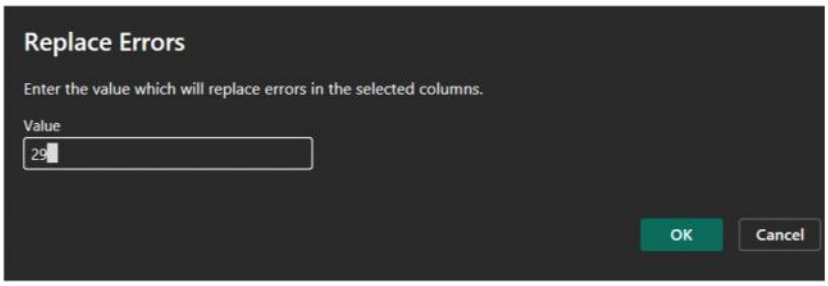
Now all cells are balanced,
Now Perform Fill Down



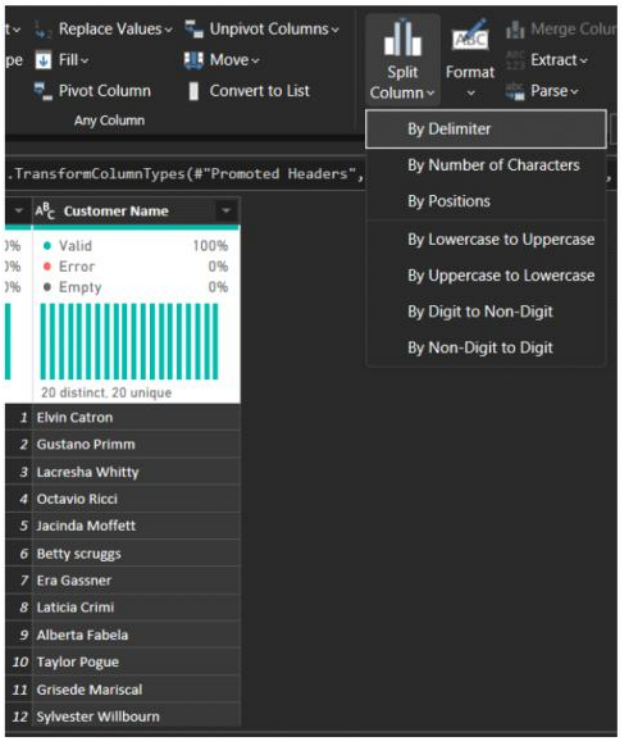


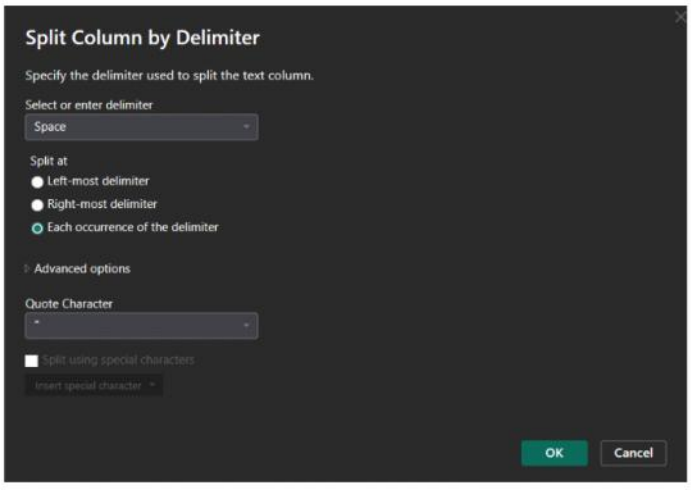




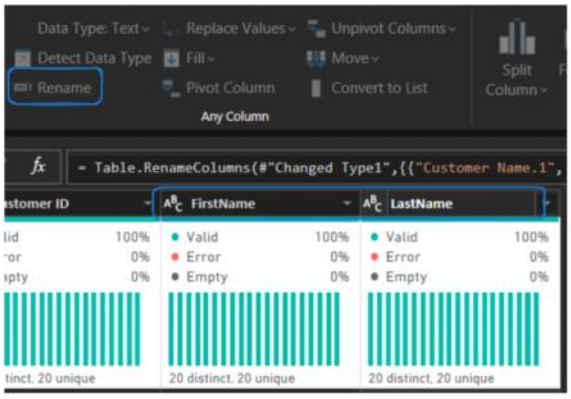


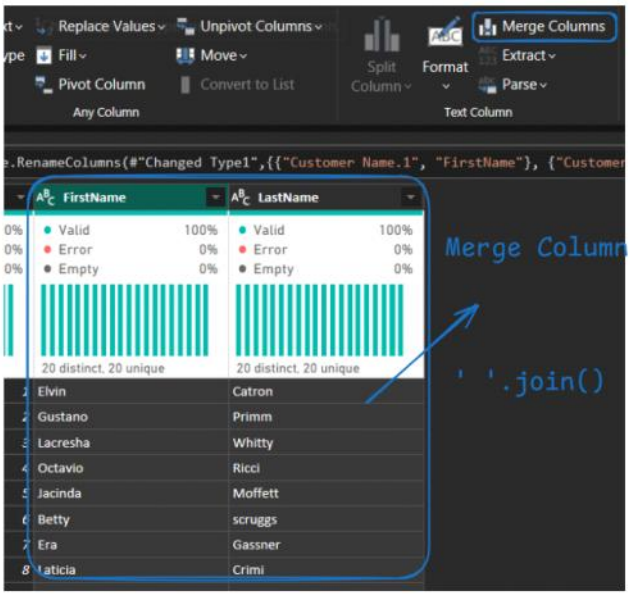
Customer ID	Customer Region	Customer Age	Sales
Valid 100% Error 0% Empty 0%	Valid 100% Error 0% Empty 0%	Valid 100% Error 0% Empty 0%	Valid 100% Error 0% Empty 0%
20 distinct, 20 unique	4 distinct, 0 unique	16 distinct, 12 unique	20 distinct, 20 unique
1	North	49	309.38
2	North	10	57.71
3	North	47	686.95
4	South	11	1,619.19
5	East	17	1,005.90
6	East	23	174.65
7	East	21	825.00
8	East	27	250.00
9	East	44	9.03
10	East	10	255.84
11	North	29	151.24
12	South	29	251.72
13	South	42	1,139.43
14	North	12	575.36
15	North	45	18.06
16	North	30	299.85
17	North	35	54.89
18	West	50	479.04
19	West	42	1,879.06
20	West	44	86.43

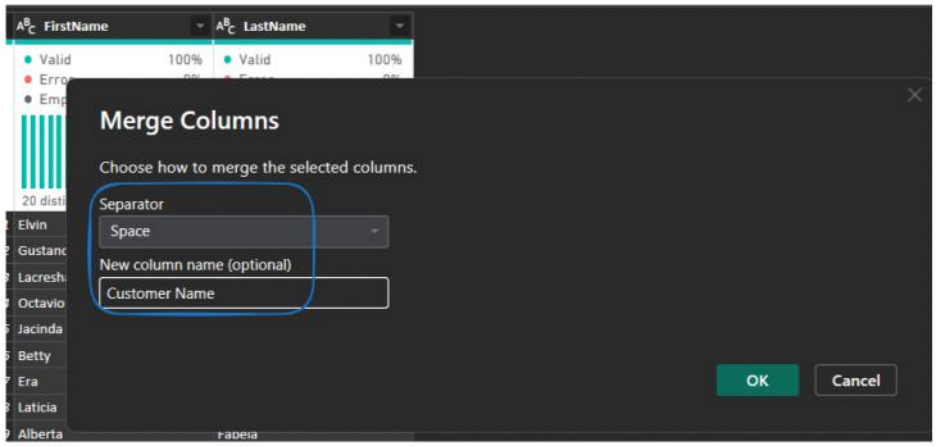


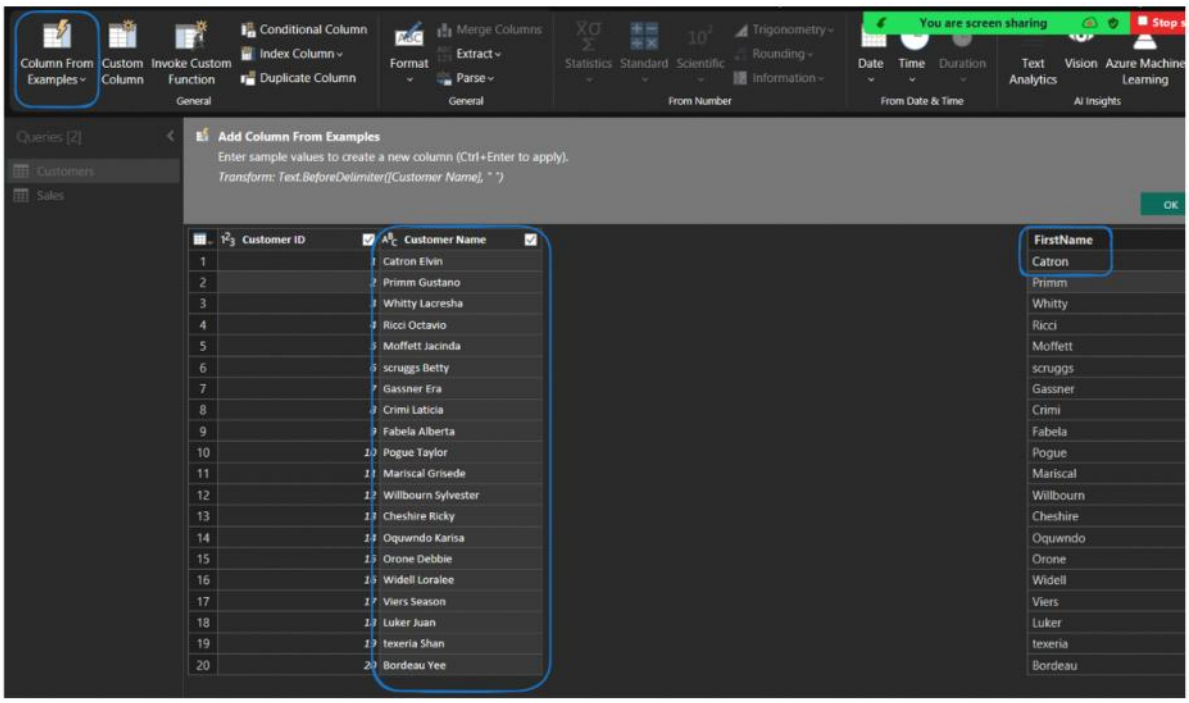


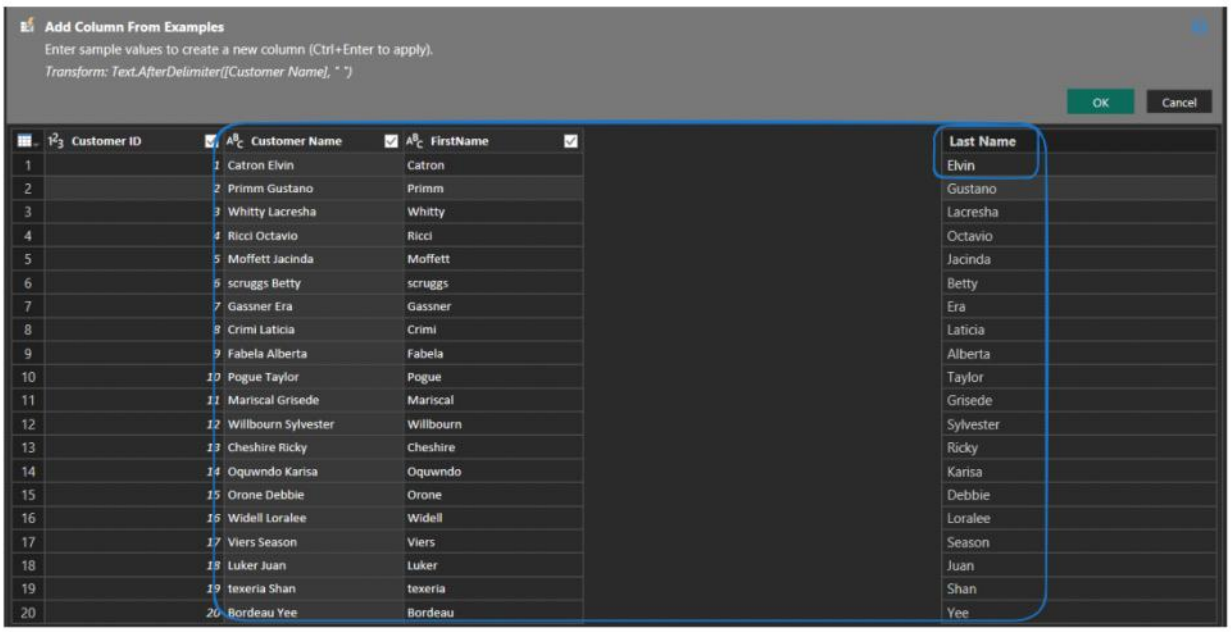
Customer Name.1
Customer Name.2
Valid 100%
Error 0%
Empty 0%
20 distinct, 20 unique
Rename the column
Elvin Catron
Gustano Primm
Lacresha Whitty
Octavio Ricci
Jacinda Moffett
Betty scruggs
Era Gassner
Laticia Crimi
Alberta Fabela
Taylor Pogue
Grisede Mariscal
Sylvester Willbourn
Ricky Cheshire
Karisa Oquwndo
Debbie Orone
Loralee Widell
Season Viers
Juan Luker
Shan texeria
Yee Bordeau

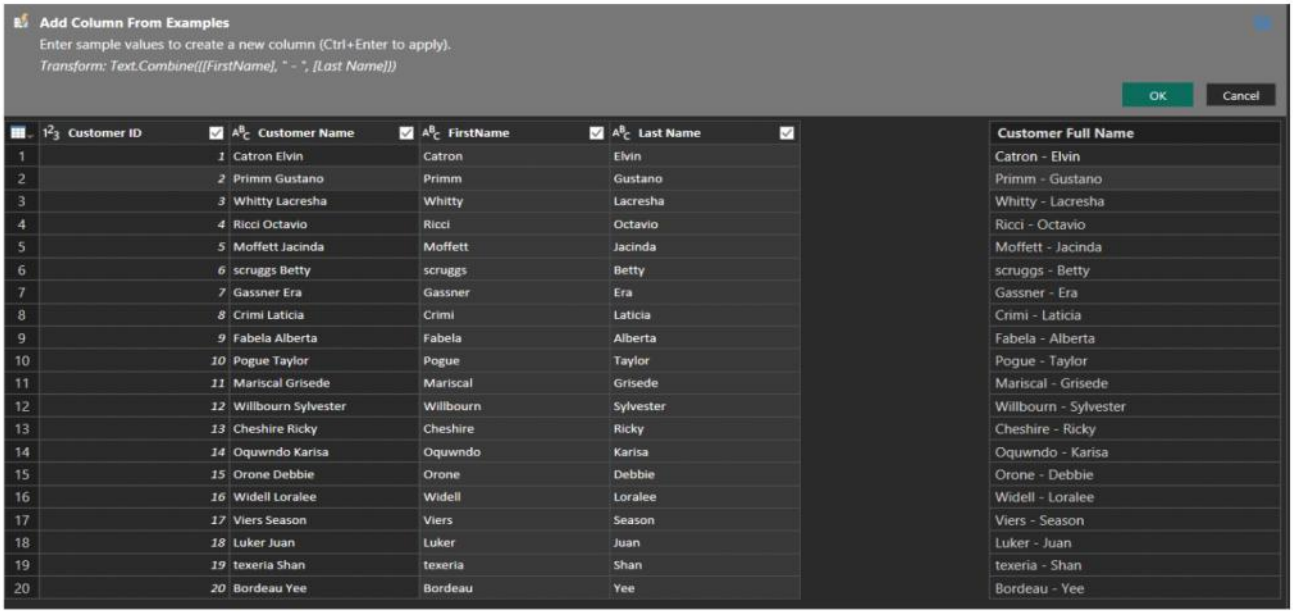

Data Type: Text
Replace Values
Unpivot Columns
Detect Data Type
Fill
Move
Rename
Pivot Column
Convert to List
Split Column
Any Column
= Table.RenameColumns(#"Changed Type1",{{"Customer Name.1",
FirstName
LastName
20 distinct, 20 unique


Replace Values
Unpivot Columns
Merge Columns
Fill
Move
Extract
Pivot Column
Convert to List
Split Column
Format
Parse
Any Column
Text Column
.RenameColumns(#"Changed Type1",{{"Customer Name.1", "FirstName"}, {"Customer
FirstName
LastName
Merge Columns
' '.join()
Elvin Catron
Gustano Primm
Lacresha Whitty
Octavio Ricci
Jacinda Moffett
Betty scruggs
Era Gassner
Laticia Crimi

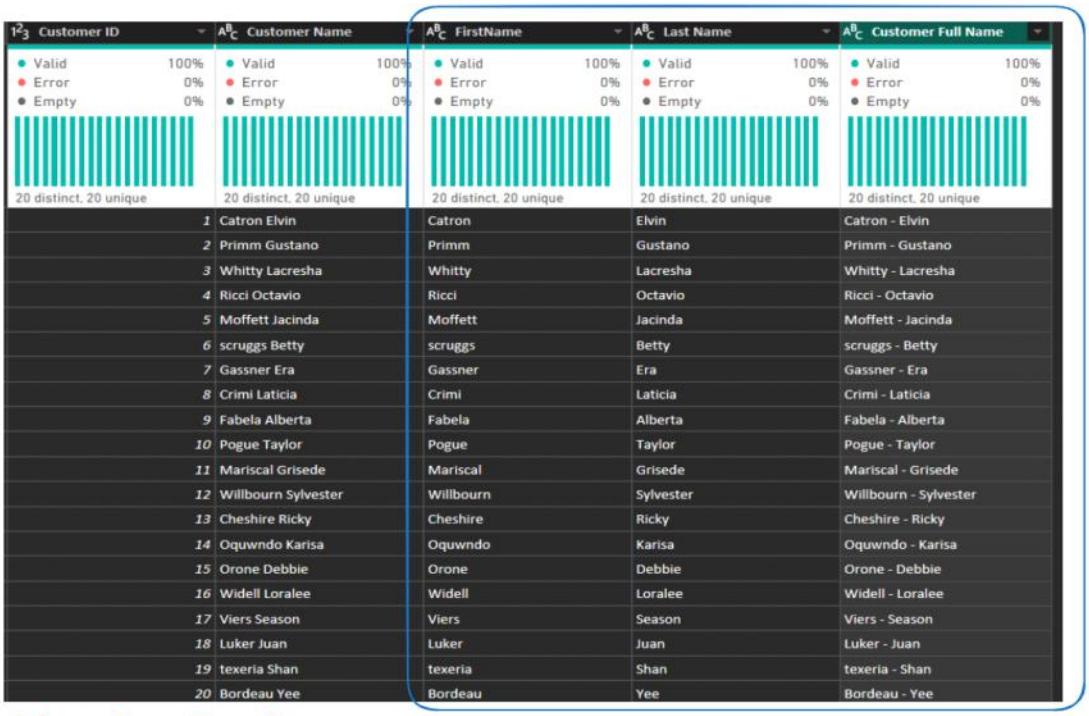

FirstName
LastName
Valid 100%
Valid 100%
Merge Columns
Choose how to merge the selected columns.
Separator
Space
New column name (optional)
Customer Name
OK
Cancel


Column From Examples
Custom Column
Invoke Custom Function
Conditional Column
Index Column
Duplicate Column
Format
Merge Columns
Extract
Parse
Statistics
Standard
Scientific
Trigonometry
Rounding
Information
You are screen sharing
Date
Time
Duration
Text Analytics
Vision
Azure Machine Learning
Add Column From Examples
Enter sample values to create a new column (Ctrl+Enter to apply).
Transform: Text.BeforeDelimiter([Customer Name], " ")
OK
Customer ID
Customer Name
FirstName
1 Catron Elvin Catron
2 Primm Gustano Primm
3 Whitty Lacresha Whitty
4 Ricci Octavio Ricci
5 Moffett Jacinda Moffett
6 scruggs Betty scruggs
7 Gassner Era Gassner
8 Crimi Laticia Crimi
9 Fabela Alberta Fabela
10 Pogue Taylor Pogue
11 Mariscal Grisede Mariscal
12 Willbourn Sylvester Willbourn
13 Cheshire Ricky Cheshire
14 Oquwndo Karisa Oquwndo
15 Orone Debbie Orone
16 Widell Loralee Widell
17 Viers Season Viers
18 Luker Juan Luker
19 texeria Shan texeria
20 Bordeau Yee Bordeau


Add Column From Examples
Enter sample values to create a new column (Ctrl+Enter to apply).
Transform: Text.AfterDelimiter([Customer Name], " ")
OK
Cancel
Customer ID
Customer Name
FirstName
Last Name
1 Catron Elvin Catron Elvin
2 Primm Gustano Primm Gustano
3 Whitty Lacresha Whitty Lacresha
4 Ricci Octavio Ricci Octavio
5 Moffett Jacinda Moffett Jacinda
6 scruggs Betty scruggs Betty
7 Gassner Era Gassner Era
8 Crimi Laticia Crimi Laticia
9 Fabela Alberta Fabela Alberta
10 Pogue Taylor Pogue Taylor
11 Mariscal Grisede Mariscal Grisede
12 Willbourn Sylvester Willbourn Sylvester
13 Cheshire Ricky Cheshire Ricky
14 Oquwndo Karisa Oquwndo Karisa
15 Orone Debbie Orone Debbie
16 Widell Loralee Widell Loralee
17 Viers Season Viers Season
18 Luker Juan Luker Juan
19 texeria Shan texeria Shan
20 Bordeau Yee Bordeau Yee



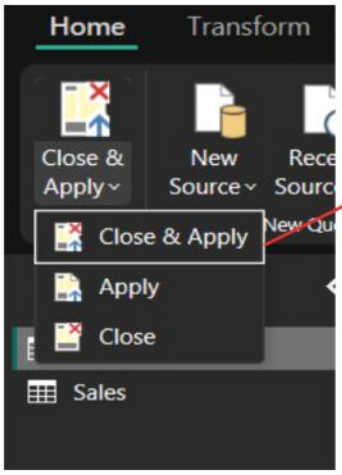
	Customer ID	Customer Name	FirstName	Last Name	Customer Full Name
1	1	Catron Elvin	Catron	Elvin	Catron - Elvin
2	2	Primm Gustano	Primm	Gustano	Primm - Gustano
3	3	Whitty Lacresha	Whitty	Lacresha	Whitty - Lacresha
4	4	Ricci Octavio	Ricci	Octavio	Ricci - Octavio
5	5	Moffett Jacinda	Moffett	Jacinda	Moffett - Jacinda
6	6	scruggs Betty	scruggs	Betty	scruggs - Betty
7	7	Gassner Era	Gassner	Era	Gassner - Era
8	8	Crimi Laticia	Crimi	Laticia	Crimi - Laticia
9	9	Fabela Alberta	Fabela	Alberta	Fabela - Alberta
10	10	Pogue Taylor	Pogue	Taylor	Pogue - Taylor
11	11	Mariscal Grisede	Mariscal	Grisede	Mariscal - Grisede
12	12	Willbourn Sylvester	Willbourn	Sylvester	Willbourn - Sylvester
13	13	Cheshire Ricky	Cheshire	Ricky	Cheshire - Ricky
14	14	Oquwndo Karisa	Oquwndo	Karisa	Oquwndo - Karisa
15	15	Orone Debbie	Orone	Debbie	Orone - Debbie
16	16	Widell Loralee	Widell	Loralee	Widell - Loralee
17	17	Viers Season	Viers	Season	Viers - Season
18	18	Luker Juan	Luker	Juan	Luker - Juan
19	19	texeria Shan	texeria	Shan	texeria - Shan
20	20	Bordeau Yee	Bordeau	Yee	Bordeau - Yee



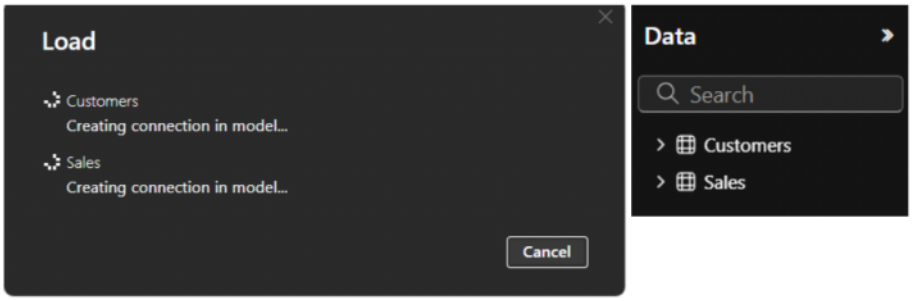
Customer ID	Customer Name	FirstName	Last Name	Customer Full Name
Valid 100% / Error 0% / Empty 0%	Valid 100% / Error 0% / Empty 0%	Valid 100% / Error 0% / Empty 0%	Valid 100% / Error 0% / Empty 0%	Valid 100% / Error 0% / Empty 0%
20 distinct, 20 unique	20 distinct, 20 unique	20 distinct, 20 unique	20 distinct, 20 unique	20 distinct, 20 unique
1	Catron Elvin	Catron	Elvin	Catron - Elvin
2	Primm Gustano	Primm	Gustano	Primm - Gustano
3	Whitty Lacresha	Whitty	Lacresha	Whitty - Lacresha
4	Ricci Octavio	Ricci	Octavio	Ricci - Octavio
5	Moffett Jacinda	Moffett	Jacinda	Moffett - Jacinda
6	scruggs Betty	scruggs	Betty	scruggs - Betty
7	Gassner Era	Gassner	Era	Gassner - Era
8	Crimi Laticia	Crimi	Laticia	Crimi - Laticia
9	Fabela Alberta	Fabela	Alberta	Fabela - Alberta
10	Pogue Taylor	Pogue	Taylor	Pogue - Taylor
11	Mariscal Grisede	Mariscal	Grisede	Mariscal - Grisede
12	Willbourn Sylvester	Willbourn	Sylvester	Willbourn - Sylvester
13	Cheshire Ricky	Cheshire	Ricky	Cheshire - Ricky
14	Oquwndo Karisa	Oquwndo	Karisa	Oquwndo - Karisa
15	Orone Debbie	Orone	Debbie	Orone - Debbie
16	Widell Loralee	Widell	Loralee	Widell - Loralee
17	Viers Season	Viers	Season	Viers - Season
18	Luker Juan	Luker	Juan	Luker - Juan
19	texeria Shan	texeria	Shan	texeria - Shan
20	Bordeau Yee	Bordeau	Yee	Bordeau - Yee

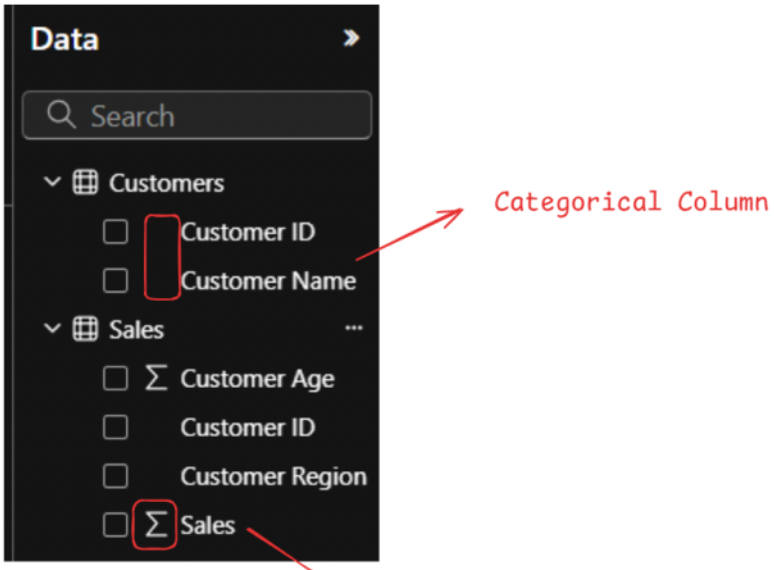
Add Column Feature

Column From Example



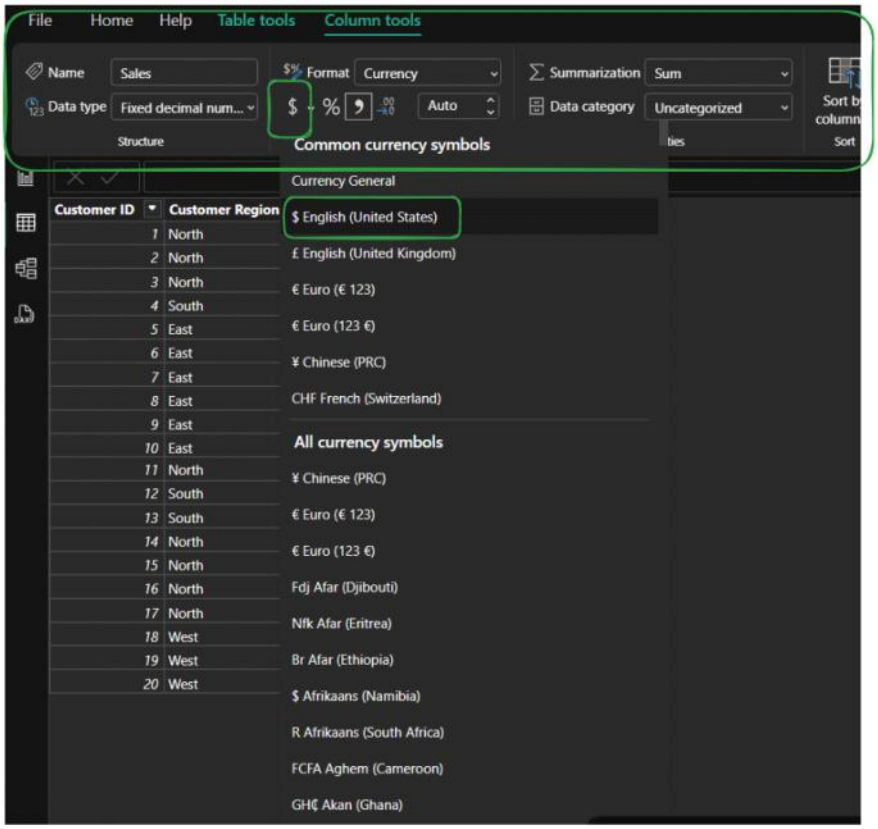
Close and Apply Changes -> Reflected on Front End



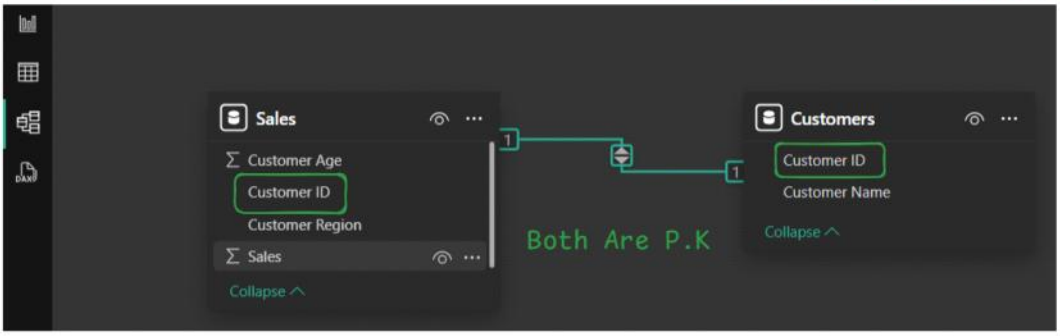


Customer ID	Customer Region	Customer Age	Sales
1	North	49	₹ 309.38
2	North	10	₹ 57.71
3	North	47	₹ 686.95
4	South	11	₹ 1,619.19
5	East	17	₹ 1,005.9
6	East	23	₹ 174.65
7	East	21	₹ 825
8	East	27	₹ 250
9	East	44	₹ 9.03
10	East	10	₹ 255.84
11	North	29	₹ 151.24
12	South	29	₹ 251.72
13	South	42	₹ 1,139.43
14	North	12	₹ 575.36
15	North	45	₹ 18.06
16	North	30	₹ 299.85
17	North	35	₹ 54.89
18	West	50	₹ 479.04
19	West	42	₹ 1,879.06
20	West	44	₹ 86.43

Regional Settings



Customer ID	Customer Region
1	North
2	North
3	North
4	South
5	East
6	East
7	East
8	East
9	East
10	East
11	North
12	South
13	South
14	North
15	North
16	North
17	North
18	West
19	West
20	West

Customer ID	Customer Region	Customer Age	Sales
1	North	49	$309.38
2	North	10	$57.71
3	North	47	$686.95
4	South	11	$1,619.19
5	East	17	$1,005.90
6	East	23	$174.65
7	East	21	$825.00
8	East	27	$250.00
9	East	44	$9.03
10	East	10	$255.84
11	North	29	$151.24
12	South	29	$251.72
13	South	42	$1,139.43
14	North	12	$575.36
15	North	45	$18.06
16	North	30	$299.85
17	North	35	$54.89
18	West	50	$479.04
19	West	42	$1,879.06
20	West	44	$86.43

1:1 Cardinality [Relationship]