

Power BI

Professor Ernesto Lee

What we will cover

- The Power Query Editor
- Transform basics
- Advanced data transformation c
- AI Insights
- Univariate Analysis
- Bivariate Analysis
- Correlation Analysis



Where Do I Manage Cleaning My Data in Power BI?

Power Query Editor

The screenshot displays the Power Query Editor interface. Five red circles with numbers 1 through 5 are overlaid on the interface, with red arrows pointing to specific features:

- 1** points to the **Close & Apply** button in the **File** tab.
- 2** points to the **Queries** list on the left sidebar.
- 3** points to the **Query Settings** pane on the right, specifically the **APPLIED STEPS** section.
- 4** points to the **Advanced Editor** button in the **Transform** tab.
- 5** points to the **Close & Apply** button in the **File** tab.

The main data view shows a table with columns: **series_id**, **year**, and **period**. The data is filtered to show rows where **series_id** is 'APU000074712' and **year** is '1974'. The **period** column shows values from M01 to M09.

The **Query Settings** pane on the right shows the **NAME** of the query as 'Historical Gas Prices' and the **APPLIED STEPS** list, which includes 'Source', 'Promoted Headers', and 'Changed Type'.

Data Cleaning in Power BI

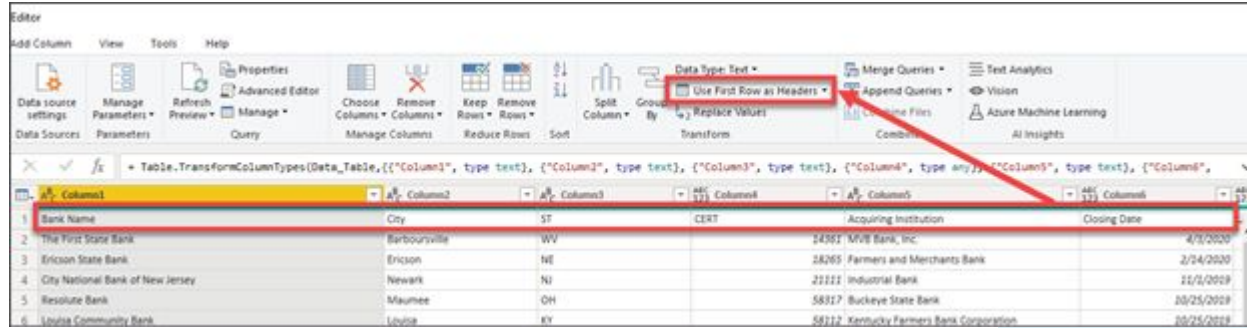
Deleting transforms: If you make a mistake and need to undo a step, you can click the Delete button next to a step.

Modifying transforms: This can be done with any step that has a gear icon next to it.

Changing the order of transforms: If you realize that it is better for one step to execute before another one, you can change the order of how the steps are executed.

Selecting previous steps: Clicking on any step prior to the current one will allow you to see how your query results would change one step earlier in the process.

Use First Row as Headers



The screenshot shows the Power Query Editor interface. The 'Data Type Text' dropdown menu is open, and the 'Use First Row as Headers' option is selected. A red box highlights the menu, and a red arrow points to the selected option. Below the menu, a table of bank data is visible with columns: Bank Name, City, ST, CERT, Acquiring Institution, and Closing Date.

Bank Name	City	ST	CERT	Acquiring Institution	Closing Date
The First State Bank	Barbourville	WV		24361 MVB Bank, Inc.	4/7/2020
Ericson State Bank	Ericson	NE		28265 Farmers and Merchants Bank	2/24/2020
City National Bank of New Jersey	Newark	NJ		21111 Industrial Bank	11/1/2019
Resolute Bank	Maumee	OH		58327 Buckeye State Bank	10/25/2019
Louisia Community Bank	Louisia	KY		58112 Kentucky Farmers Bank Corporation	10/25/2019

<https://bit.ly/2410C>

*Do Mini.md in the link above



Remove Columns

The screenshot displays the Power Query Editor interface. The ribbon at the top includes tabs for Column, View, Tools, and Help. The 'Column' tab is active, showing various options like 'Choose Columns', 'Remove Columns', 'Keep Rows', and 'Remove Rows'. A red arrow points to the 'Remove Columns' button in the ribbon. Below the ribbon, a data table is visible with columns: Bank Name, City, ST, and Acquiring Institution. A context menu is open over the 'ST' column header, with the 'Remove Columns' option highlighted by a red rectangle. The menu also includes options like 'Remove Other Columns', 'Add Column From Examples...', 'Remove Duplicates', 'Remove Errors', 'Replace Values...', 'Fill', 'Change Type', 'Merge Columns', 'Group By...', 'Unpivot Columns', 'Unpivot Other Columns', 'Unpivot Only Selected Columns', and 'Move'.

Bank Name	City	ST	Acquiring Institution
The First State Bank	Barbourville	WV	MVB Bank, Inc.
Ericson State Bank	Ericson	NE	Farmers and Merchants Bank
City National Bank of New Jersey	Newark	NJ	Industrial Bank
Resolute Bank	Maumee	OH	Buckeye State Bank
Louisa Community Bank	Louisa	KY	Kentucky Farmers Bank Corporation
The Enloe State Bank	Cooper	TX	Legend Bank, N. A.
Washington Federal Bank for Savings	Chicago	IL	Royal Savings Bank
The Farmers and Merchants State Bank of Argonia	Argonia	KS	Conway Bank
Fayette County Bank	Saint Elmo	IL	United Fidelity Bank, fsb
Guaranty Bank, (d/b/a BestBank in Georgia & Michigan)	Milwaukee	WI	First-Citizens Bank & Trust Company
First NBC Bank	New Orleans	LA	Whitney Bank
Proflco Bank	Cottonwood Heights	UT	Cache Valley Bank
Seaway Bank and Trust Company	Chicago	IL	State Bank of Texas
Harvest Community Bank	Pennsville	NI	First-Citizens Bank & Trust Company
Allied Bank	Mulberry	AR	Today's Bank
The Woodbury Banking Company	Woodbury	GA	United Bank
First CornerStone Bank	King of Prussia	PA	First-Citizens Bank & Trust Company

Quantitative or Qualitative

```
graph TD; A[Quantitative or Qualitative] --> B[Quantitative]; A --> C[Qualitative]; B --> D[Continuous]; B --> E[Discrete]; D --> F[Interval]; D --> G[Ratio]; E --> H[Count]; C --> I[Categorical]; C --> J[Binary]; C --> K[Ordinal];
```

Quantitative

Continuous

Interval

Ratio

Discrete

Count

Qualitative

Categorical

Binary

Ordinal

Changing Data Types

Editor

Add Column View Tools Help

Data source settings Manage Parameters Refresh Preview Properties Advanced Editor Manage

Choose Columns Remove Columns Keep Rows Remove Rows

Sort Split Column Group By

Data Type: Date Use First Row as Headers Replace Values

Merge Queries Append Queries Combine Files Combine

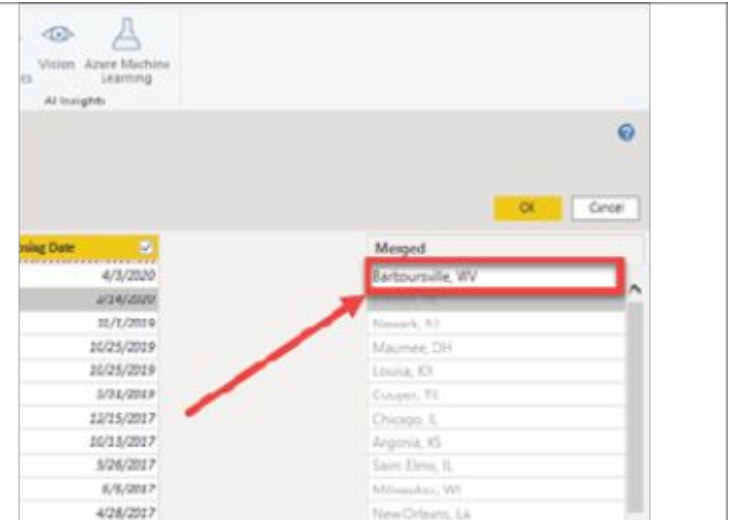
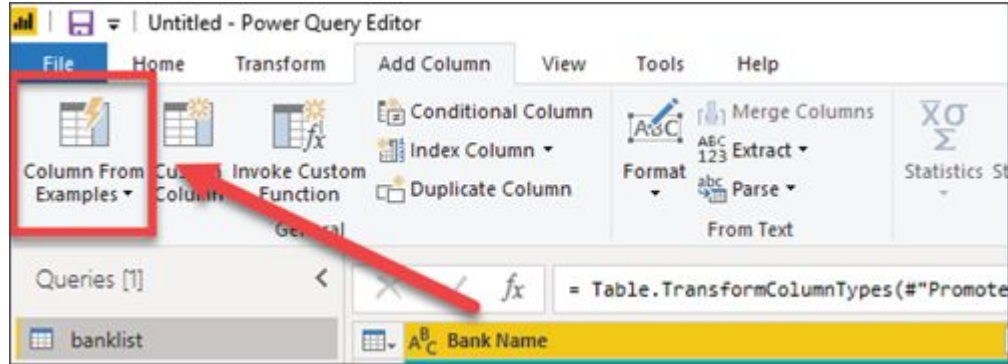
Table.SelectColumns(#"Changed Type1",{"Bank Name", "City", "ST", "Closing Date"})

	Bank Name	City	ST	Closing Date
1	The First State Bank	Barboursville	WV	
2	Ericson State Bank	Ericson	NE	
3	City National Bank of New Jersey	Newark	NJ	
4	Resolute Bank	Maumee	OH	
5	Louisa Community Bank	Louisa	KY	
6	The Enloe State Bank	Cooper	TX	
7	Washington Federal Bank for Savings	Chicago	IL	
8	The Farmers and Merchants State Bank of Argonia	Argonia	KS	
9	Fayette County Bank	Saint Elmo	IL	
10	Guaranty Bank (d/b/a BestBank in Georgia & Michigan)	Millwaukee	WI	
11	First NBC Bank	New Orleans	LA	
12	Proflco Bank	Cottonwood Heights	UT	
13	Seaway Bank and Trust Company	Chicago	IL	
14	Harvest Community Bank	Pennsville	NJ	
15	Allied Bank	Mulberry	AR	
16	The Woodbury Banking Company	Woodbury	GA	

1.2 Decimal Number
\$ Fixed decimal number
123 Whole Number
% Percentage
Date/Time
Date
Time
Date/Time/Timezone
Duration
Text
True/False
Binary
Using Locale...

8/19/2016

Create Column from Examples



Advanced Data Transformations

Conditional Columns

Split Column by Delimiter

Specify the delimiter used to split the text column.

Select or enter delimiter

--Custom--

,

- Split at
- ☐ Left-most delimiter
 - ☐ Right-most delimiter
 - ☒ Each occurrence of the delimiter

Advanced options

Add Conditional Column

Add a conditional column that is computed from the other columns or values.

New column name
State Name

Column Name	Operator	Value	Output
If State Abbreviation	equals	ABC 123 null	Then County Name

Add Clause

Else
ABC 123 null

OK Cancel

ABC 123

Enter a value

Select a column

Parameter

Queries [1]

FIPS_CountyName

FIPS_CountyCode	County Name	State Abbreviation	State Name
01000	ALABAMA		ALABAMA
01001	Autauga County	AL	
01003	Baldwin County	AL	
01005	Barbour County	AL	
01007	Bibb County	AL	
01009	Blount County	AL	
01011	Bullock County	AL	
01013	Butler County	AL	
01015	Calhoun County	AL	
01017	Chambers County	AL	
01019	Cherokee County	AL	
01021	Chilton County	AL	
01023	Choctaw County	AL	
01025	Clarke County	AL	
01027	Clay County	AL	
01029	Cleburne County	AL	
01031	Coffee County	AL	

Table.AddColumn(#"Renamed Columns", "State Name", each if [State Abbreviation] = "AL" then "ALABAMA" else null)

Query Settings

PROPERTIES

Name
FIPS_CountyName

APPLIED STEPS

- Source
- Changed Type
- Filtered Rows
- Split Column by Delimiter
- Changed Type1
- Renamed Columns
- Added Conditional Column

Fill Down

Table.SelectRows("#Filled Down", each ([State Abbreviation] <> null))

	A ^B _C County Code	A ^B _C County Name	A ^B _C State Abbreviation	A ^B _C State Name
1	01001	Autauga County	AL	Alabama
2	01003	Baldwin County	AL	Alabama
3	01005	Barbour County	AL	Alabama
4	01007	Bibb County	AL	Alabama
5	01009	Blount County	AL	Alabama
6	01011	Bullock County	AL	Alabama
7	01013	Butler County	AL	Alabama
8	01015	Calhoun County	AL	Alabama
9	01017	Chambers County	AL	Alabama
10	01019	Cherokee County	AL	Alabama
11	01021	Chilton County	AL	Alabama
12	01023	Choctaw County	AL	Alabama
13	01025	Clarke County	AL	Alabama
14	01027	Clay County	AL	Alabama
15	01029	Cleburne County	AL	Alabama
16	01031	Coffee County	AL	Alabama
17	01033	Colbert County	AL	Alabama

Query Settings

PROPERTIES

Name
FIPS_CountyName
[All Properties](#)

APPLIED STEPS

- Source
- Changed Type
- Filtered Rows
- Split Column by Delimiter
- Changed Type1
- Renamed Columns
- Added Conditional Column
- Capitalized Each Word
- Filled Down
- Filtered Rows1**

Unipivot

¹² ₃ Year	¹² ₃ Florida	¹² ₃ Pennsylvania	¹² ₃ New York	¹² ₃ California	¹² ₃ Texas
2018	21312211	12823989	19862512	39776830	28704330
2017	20984400	12805537	19849399	39536653	28304596
2016	20656589	12787085	19836286	39296576	27904862

Unipivot

¹⁻² ₃ Year	^A _C State	¹⁻² ₃ Population
2018	Florida	21312211
2018	Pennsylvania	12823989
2018	New York	19862512
2018	California	39776830
2018	Texas	28704330
2017	Florida	20984400
2017	Pennsylvania	12905537
2017	New York	19849399
2017	California	39536653
2017	Texas	28304596
2016	Florida	20556589
2016	Pennsylvania	12787085
2016	New York	19836286
2016	California	39296576
2016	Texas	27904862

Unipivot

	1.2 Column1	1.2 Column2	1.2 Column3	1.2 Column4	1.2 Column5	1.2 Column6
1	GDP per capita PPP, with projections	1764	1765	1766	1767	1768
2	Abkhazia	null	null	null	null	null
3	Afghanistan	null	null	null	null	null
4	Akrotiri and Dhekelia	null	null	null	null	null
5	Albania	null	null	null	null	null
6	Algeria	null	null	null	null	null
7	American Samoa	null	null	null	null	null
8	Andorra	null	null	null	null	null
9	Angola	null	null	null	null	null
10	Anguilla	null	null	null	null	null
11	Antigua and Barbuda	null	null	null	null	null
12	Argentina	null	null	null	null	null
13	Armenia	null	null	null	null	null
14	Aruba	null	null	null	null	null
15	Australia	null	null	null	null	null
16	Austria	1403.59853	1404.448052	1405.298088	1406.148639	1406.999704
17	Azerbaijan	null	null			
18	Bahamas	null	null			
19	Bahrain	null	null			
20	Bangladesh	614.52257	614.3627258			
21	Barbados	null	null			
22	Belarus	null	null			
23	Belgium	1532.258272	1532.450647			

Queries [1]

Income

Table.RenameColumns(#"Unpivoted Other Columns",{"Attribute",

Country

Year

Income

1	Afghanistan	1800	472.0534996
2	Afghanistan	1820	472.0534996
3	Afghanistan	1913	638.3786419
4	Afghanistan	1950	757.3187954
5	Afghanistan	1951	766.7521974
6	Afghanistan	1952	779.4453145
7	Afghanistan	1953	812.8562795
8	Afghanistan	1954	815.3595213
9	Afghanistan	1955	816.414838
10	Afghanistan	1956	837.0669354
11	Afghanistan	1957	820.8530296
12	Afghanistan	1958	849.7400695
13	Afghanistan	1959	856.2288434
14	Afghanistan	1960	868.4982226
15	Afghanistan	1961	857.3586549
16	Afghanistan	1962	853.10071
17	Afghanistan	1963	849.4447177

Query Settings

PROPERTIES

Name

Income

All Properties

APPLIED STEPS

Source

Navigation

Changed Type

Promoted Headers

Changed Type1

Renamed Columns

Unpivoted Other Columns

Renamed Columns1

1	2	3	4	5	6	7
Column1	Column2	Column3	Column4	Column5	Column6	Column7
Total population		1086	1100	1290	1300	1341
Abkhazia		null	null		null	null
Afghanistan		null	null	null	null	null
Azotiri and Dhekelia		null		null	null	
Albania		null	null	null	null	null

The screenshot shows the Power BI Desktop interface. At the top, the formula bar contains the DAX expression: `= Table.RenameColumns(#"Changed Type1",{{"Total population",`. The main area displays a data table with three columns: **2013**, **2050**, and **2100**. The table contains 21 rows of data. A right-click context menu is open over the table, listing various actions. The option **Unpivot Other Columns** is highlighted with a red rectangle, and a red arrow points to it from the bottom left.

	2013	2050	2100
1	null	null	null
2	33397058	34499915	76249991
3	null	null	null
4	3227373	3238316	2989747
5	36485828	36983924	46521560
6	70680	71834	120458
7	87518	88909	136617
8	20162517	20714494	42334143
9	15911	16158	18300
10	90510	91404	111577
11	41118986	41473982	50559806
12	3108972	3117722	2930743
13	108587	108883	103749
14	22918688	23213944	31385363
15	8428915	8441263	8427060
16	9421233	9533445	11578484
17	351275	355233	445096
18	1359485	1377273	1801018
19	152408774	154393847	194352619
20	274530	275141	264477
21			

Context Menu Options:

- Copy
- Remove Columns
- Remove Other Columns
- Add Column From Examples...
- Remove Duplicates
- Remove Errors
- Replace Values...
- Fill
- Change Type
- Transform
- Merge Columns
- Sum
- Product
- Group By...
- Unpivot Columns
- Unpivot Other Columns**
- Unpivot Only Selected Columns
- Move

Merge Queries

- Left Outer (all rows from the first table, only matching rows from the second)
- Right Outer (all rows from the second table, only matching rows from the first)
- Full Outer (all rows from both tables)
- Inner (only matching rows from both tables)
- Left Anti (rows only in the first table)
- Right Anti (rows only in the second table)

Merge

Select tables and matching columns to create a merged table.

Population

Country	1	Year	2	Population
Afghanistan		1800		3280000
Afghanistan		1820		3280000
Afghanistan		1870		4207000
Afghanistan		1913		5730000
Afghanistan		1950		8151455

Income

Country	1	Year	2	Income
Afghanistan		1800		472.0534996
Afghanistan		1820		472.0534996
Afghanistan		1913		638.3786419
Afghanistan		1950		757.3187954
Afghanistan		1951		766.7521974

Join Kind

Inner (only matching rows)

☐ Use fuzzy matching to perform the merge

► Fuzzy matching options

✓ The selection matches 17221 of 22127 rows from the first table, and 1722...

OK

Cancel

	A ^B Country	A ^B Year	1.2 Population	1.2 Income
1	Afghanistan	1800	3280000	472.0534996
2	Afghanistan	1820	3280000	472.0534996
3	Afghanistan	1913	5730000	638.3786419
4	Afghanistan	1950	8151455	757.3187954
5	Afghanistan	1951	8276820	766.7521974
6	Afghanistan	1952	8407148	779.4453145
7	Afghanistan	1953	8542906	812.8562795
8	Afghanistan	1954	8684494	815.3595213
9	Afghanistan	1955	8832253	816.414838
10	Afghanistan	1956	8986449	837.0669354
11	Afghanistan	1957	9147286	820.8530296
12	Afghanistan	1958	9314915	849.7400695
13	Afghanistan	1959	9489453	856.2288434
14	Afghanistan	1960	9671046	868.4982226
15	Afghanistan	1961	9859928	857.3586549
16	Afghanistan	1962	10056480	853.10071
17	Afghanistan	1963	10261254	849.4447177
18	Afghanistan	1964	10474903	846.2672818

Queries [3]

Income

Population

Country Stat

- Copy
- Paste
- Delete
- Rename
- ☒ Enable load
- ☒ Include in report refresh
- Duplicate
- Reference
- Move To Group
- Move Up
- Move Down
- Create Function...
- Convert To Parameter
- Advanced Editor
- Properties...

Append Query

Append

Concatenate rows from two tables into a single table.

☒ Two tables ☐ Three or more tables

Primary table

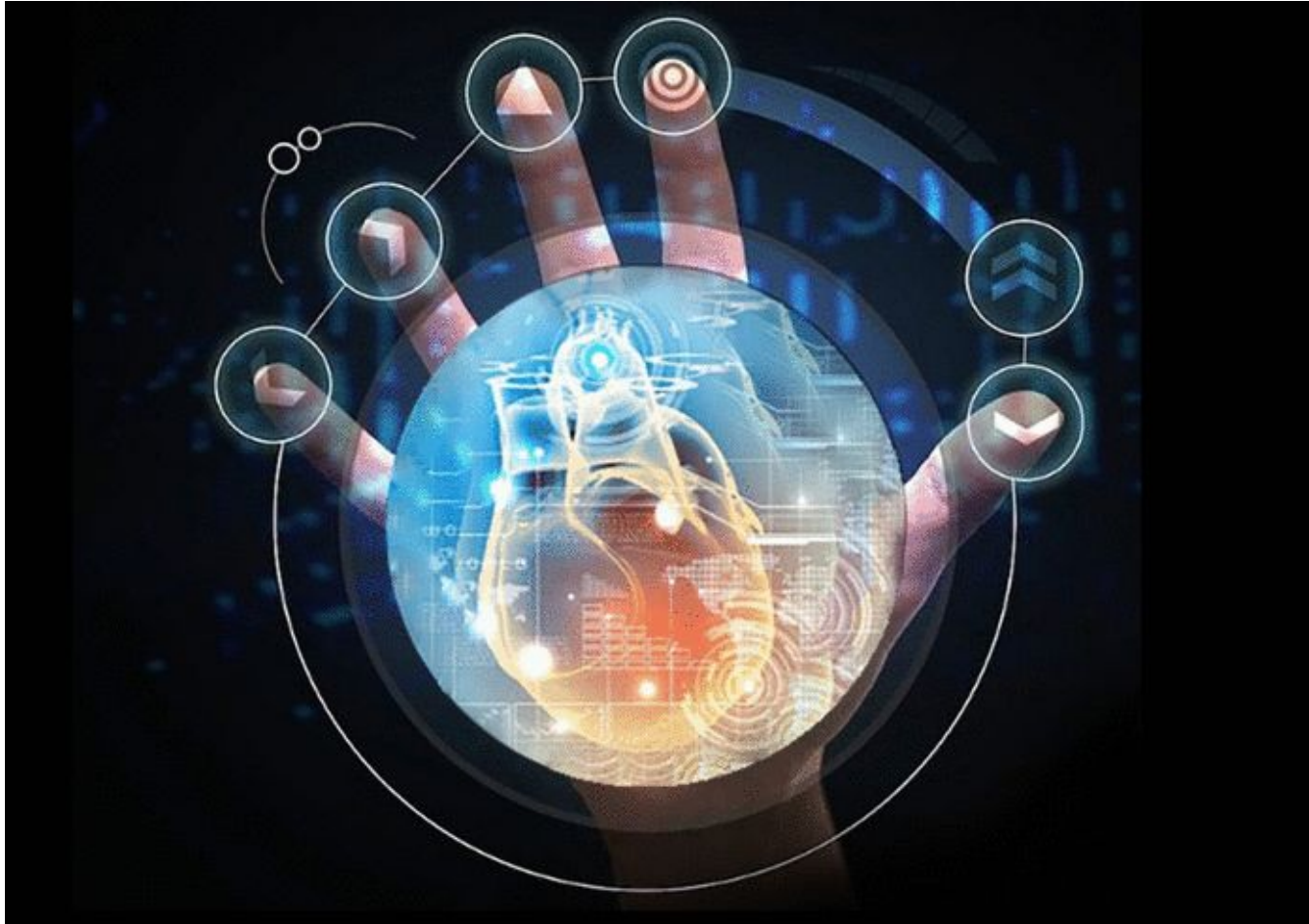
Credit Card Complaints

Table to append to the primary table

Student Loan Complaints

	Complaint ID	Product	Sub-product	Issue	Sub-issue	State
1	1429682	Credit card		Credit line increase/decrease		NC
2	1430686	Credit card		Billing statement		PA
3	1429869	Credit card		Advertising and marketing		IL
4	1427544	Credit card		Billing disputes		WI
5	1427873	Credit card		APR or interest rate		NJ
6	1428112	Credit card		Credit determination		NY
7	1430459	Credit card		Billing disputes		NJ
8	1425175	Credit card		Privacy		NC
9	1425203	Credit card		Other		FL
10	1424483	Credit card		Advertising and marketing		NY
11	1424843	Credit card		Billing disputes		FL
12	1422868	Credit card		Advertising and marketing		MN
13	1423026	Credit card		Late fee		VA
14	1423176	Credit card		Customer service / Customer relations		CA
15	1423531	Credit card		Closing/Cancelling account		WA
16	1424194	Credit card		Other		CA
17	1424522	Credit card		Late fee		MD
18	1421071	Credit card		Identity theft / Fraud / Embezzlement		CA
19	1420918	Credit card		Billing disputes		MD
20	1420893	Credit card		Payoff process		OH
21	1421027	Credit card		Closing/Cancelling account		FL
22	1421773	Credit card		Transaction issue		AZ
23	1421915	Credit card		Customer service / Customer relations		MO

Power BI and AI



Text Analytics

Text Analytics [3]

- Detect language
- Extract key phrases
- Score sentiment

Score sentiment

Measure the positive or negative sentiment of words and phrases. Sentiment is rated on a scale of zero to one, with one being the most positive.

[Learn more](#)

Text

Language ISO code (optional)

Premium capacity used for AI Insights

ReviewText	Title	Score sentiment
Everything that has been written by other reviewers was spot on. It is ...	Very accommodating	0.762831211
This place needs a complete overhaul from top to bottom. It is run do...	OLD, RUN-DOWN SMELLS MUSTY - AV...	0.030658334
I can't say enough good things about the Fort Conde Inn!! The place it...	A True Gem	0.929421484
We've stayed in several of the hotels in Griffin, I feel this is one of the ...	One of the best in Griffin	0.770861864
I travel a lot and see a lot hotels. However this was the worst bathroo...	Not good at all	0.206676841
My family and I recently moved to Jacksonville and experienced some ...	Great Experience!	0.880086839
Beds were had as rocks, light came in through the wind since there we...	Felt more like a boarding house	0.821575588
I knew when I booked this hotel, a few several months in advance, I kn...	Buyer Beware	0.134426296
This was a terrific place to stay. Just minutes from Spearfish Canyon an...	great place to stay	0.9487046
If you just need a room in an modestly accessible area, this property d...	Amenities Lacking	0.677353084
Love how I can count on La Quinta for a great nights sleep. Best beds!	Great Room	0.987293124
We enjoyed a one night stay while passing through. Very convenient o...	Very nice property	0.941030383
Comfortable accommodations and friendly, excellent staff. Stayed for ...	Enjoyable	0.985230923
Hampton Inn and suites were a vet pleasant surprise for us. The king s...	Best place in town	0.947005153
The staff was very friendly.	Comfortable Suite Hotel	0.981755257
A very pleasant stay, convenient to everything in Albany.	My six night stay at Days Inn	0.950045943
My husband spent 6 days at Hyatt Place Northpoint. The staff was extr...	Great place to stay	0.933157802
I was visiting family at a nearby Army base and wanted to stay at a hot...	A great find!	0.702835798
Underwhelmed. This hotel smells bad, has some loud, partying guests.	Don't Bother Staying Here....	0.020978987
This is our 3rd year returning to Gatlinburg and I must say this is the b...	Love!!!!!!	0.24581942

Univariate Analysis

Bivariate Analysis

Correlation