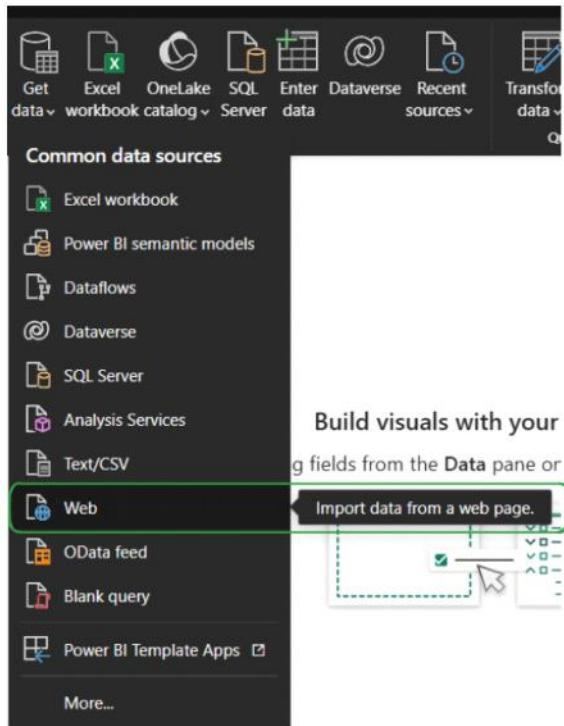


Data Cleaning with Power BI

- Web Scrapping
- Scenario Based Cleaning
- Add Columns
- Rolling Calendar [Invoke Custom Columns]



- Website**
1. Cricbuzz / ESPNCricInfo / Crex
 2. Wikipedia[Tabular]
 3. Screener [Stock Analysis]
 4. IPL-T20

Navigator

Display Options

- HTML Tables [11]
 - Table 1
 - Table 2
 - Table 3
 - Table 4
 - Table 5
 - Table 6
 - Table 7
 - Table 8
 - Table 9
 - Table 10
 - Table 11
- Suggested Tables [6]
 - Table 12
 - Table 13
 - Table 14
 - Table 15
 - Table 16
 - Table 17
- Text [2]
 -
 -

Table View

Table 11

Column1	Column2	Column3	Column4	Column5	Column6
Promoters +	27.37%	27.54%	27.74%	27.88%	27.98%
Fits +	10.07%	7.99%	6.11%	5.24%	1.12%
Dilts +	10.52%	13.46%	15.37%	17.77%	14.41%
Public +	52.04%	51.01%	50.79%	49.11%	44.47%
No. of Shareholders	38,782	46,283	61,429	63,243	50,000

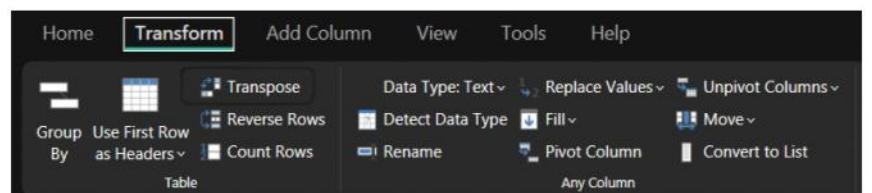
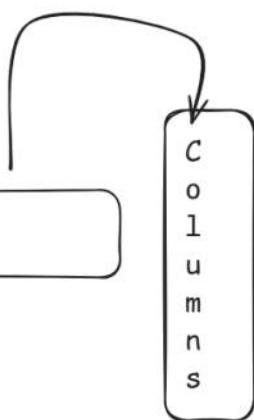
Add Table Using Examples Load Transform Data Cancel

Queries [11]

	ABC Column1	ABC Column2	ABC Column3	ABC Column4	ABC Column5	ABC Column6	ABC Column7
1	Sep 2013	Sep 2014	Mar 2016	Mar 2017	Mar 2018	Mar 2019	
2	Equity Capital	4	4	4	4	4	
3	Reserves	3,651	4,535	7,220	8,637	9,734	10,833
4	Borrowings +	1,599	1,903	2,464	2,333	2,163	2,506
5	Other Liabilities +	3,023	3,391	3,566	4,075	4,577	5,098
6	Total Liabilities	8,278	9,833	13,254	15,048	16,478	18,441
7	Fixed Assets +	2,974	3,436	4,608	5,302	6,092	6,786
8	CWIP	359	628	1,059	848	1,079	1,403
9	Investments	904	1,081	3,138	3,382	4,145	3,855
10	Other Assets +	4,041	4,688	4,448	5,316	5,162	6,397
11	Total Assets	8,278	9,833	13,254	15,048	16,478	18,441

Queries [11]

	ABC Column1	ABC Column2	ABC Column3	ABC Column4	ABC Column5	ABC Column6	ABC Column7
1	Sep 2022	Dec 2022	Mar 2023	Jun 2023	Sep 2023	Dec 2023	
2	Sales +	5,826	5,645	5,842	6,440	6,217	6,162
3	Expenses +	5,346	5,083	4,588	5,310	5,060	5,108
4	Operating Profit	481	562	854	1,130	1,157	1,055
5	OPM %	8%	10%	15%	18%	19%	17%
6	Other Income +	77	70	75	71	78	
7	Interest	75	86	92	84	86	90
8	Depreciation	309	316	330	333	351	360
9	Profit before tax	174	231	501	787	791	682
10	Tax %	25%	24%	32%	25%	26%	25%
11	Net Profit +	130	175	341	589	587	510
12	EPS in Rs	306.19	412.23	803.19	1,388.20	1,383.27	1,201.83



Rows

Transpose

Queries [11]

	ABC Column1	ABC Column2	ABC Column3	ABC Column4	ABC Column5	ABC Column6	ABC Column7
1	Sales +	Expenses +	Operating Profit	OPM %	Other Income +	Interest	
2	5,826	5,346	481	8%	77	75	
3	5,645	5,083	562	10%	71	88	
4	5,842	4,988	854	15%	70	92	
5	6,440	5,310	1,130	18%	75	84	
6	6,217	5,060	1,157	19%	71	86	
7	6,162	5,108	1,055	17%	78	90	
8	6,349	5,437	912	14%	94	93	
9	7,196	6,037	1,160	16%	84	85	
10	6,481	5,870	1,011	15%	113	84	
11	7,001	6,166	835	12%	98	94	
12	7,075	5,996	1,079	15%	113	98	
13	7,076	6,604	1,071	14%	126	98	
14	7,379	6,253	1,126	15%	108	90	

Table.PromoteHeaders("#"Transposed Table", [PromoteAllScalars=true])

ABC 123	ABC 123 Sales +	ABC 123 Expenses +	ABC 123 Operating Profit	ABC 123 OPM %	ABC 123 Other Income +	ABC 123 Interest
● Valid	100%	● Valid	100%	● Valid	100%	● Valid
● Error	0%	● Error	0%	● Error	0%	● Error
● Empty	0%	● Empty	0%	● Empty	0%	● Empty
1 Sep 2022	5,826	5,346	481	8%	77	75
2 Dec 2022	5,645	5,083	562	10%	71	86
3 Mar 2023	5,842	4,988	854	15%	70	92
4 Jun 2023	6,440	5,310	1,130	18%	75	84
5 Sep 2023	6,217	5,060	1,157	19%	71	86
6 Dec 2023	6,162	5,108	1,055	17%	78	90
7 Mar 2024	6,349	5,437	912	14%	94	93
8 Jun 2024	7,196	6,037	1,160	16%	84	85
9 Sep 2024	6,881	5,870	1,011	15%	113	84
10 Dec 2024	7,001	6,166	835	12%	98	94
11 Mar 2025	7,075	5,996	1,079	15%	113	98
12 Jun 2025	7,676	6,604	1,071	14%	126	98
13 Sep 2025	7,379	6,253	1,126	15%	108	90

Query Settings X

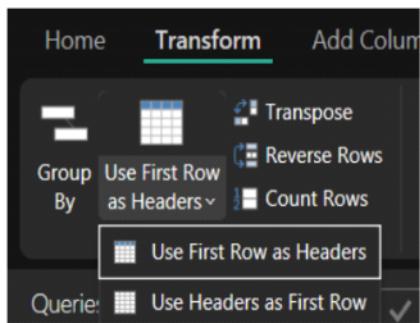
PROPERTIES

Name: Quarterly Results

All Properties

APPLIED STEPS

- Source: Extracted Table From Html
- Transposed Table
- Promoted Headers



ABC 123 Raw PDF

Properties

Copy

Remove

Remove Other Columns

Duplicate Column

Add Column From Examples...

Remove Duplicates

Remove Errors

Change Type

Transform

Replace Values...

Replace Errors...

Split Column

Group By...

Fill

Unpivot Columns

Unpivot Other Columns

Unpivot Only Selected Columns

Rename...

Move

Drill Down

Add as New Query

Replace Values

Replace one value with another in the selected columns.

Value To Find: 18m

Replace With: []

Advanced options

OK Cancel

Table.ReplaceValue(#"Promoted Headers", "18m", "", Replacer.ReplaceText, "")

	ABC	ABC Sales +	ABC 123 Expenses +	ABC 123 Operating Profit	ABC 123 OPM %	ABC 123 Other Income +
● Valid	100%	● Valid	100%	● Valid	100%	● Valid
● Error	0%	● Error	0%	● Error	0%	● Error
● Empty	0%	● Empty	0%	● Empty	0%	● Empty
	13 distinct, 13 unique					
1	Sep 2013	12,248				
2	Sep 2014	13,329				
3	Mar 2016	20,179				
4	Mar 2017	13,412				
5	Mar 2018	14,954				
6	Mar 2019	16,062	13,744	2,317	14%	416
7	Mar 2020	16,237	13,855	2,382	15%	335
8	Mar 2021	16,162	13,208	2,954	18%	205
9	Mar 2022	19,317	17,256	2,061	11%	315
10	Mar 2023	23,008	20,604	2,404	10%	245
11	Mar 2024	25,169	20,896	4,272	17%	306
12

Keep Top Rows

Specify how many rows to keep.

Number of rows: 12

OK Cancel

Queries [11]

Customer Sales Sales detail Product Information Po... sample_json_column Product Summary Exploring_Features Split_Feature Quarterly Results Profit & Loss Balance Sheet

Table.TransformColumnTypes(#"Promoted Headers", {"Date", type date}, {"Equity Capital", Int64.Type}, {"Reserves", Currency.Type}, {"Borrowings", Currency.Type}, {"Other Liabilities", Currency.Type}, {"Total Liabilities", Currency.Type}, {"Fixed Assets", Currency.Type})

	Date	Equity Capital	Reserves	Borrowings	Other Liabilities	Total Liabilities	Fixed Assets
● Valid	100%	● Valid	100%	● Valid	100%	● Valid	100%
● Error	0%	● Error	0%	● Error	0%	● Error	0%
● Empty	0%	● Empty	0%	● Empty	0%	● Empty	0%
	13 distinct, 13 unique						
1	01-09-2013	4	3,651.00	1,599.00	3,023.00	8,278.00	
2	01-09-2014	4	4,535.00	1,903.00	3,391.00	9,833.00	
3	01-03-2016	4	7,220.00	2,464.00	3,566.00	13,254.00	
4	01-03-2017	4	8,637.00	2,333.00	4,075.00	15,048.00	
5	01-03-2018	4	9,734.00	2,163.00	4,577.00	16,478.00	
6	01-03-2019	4	10,833.00	2,506.00	5,098.00	18,441.00	
7	01-03-2020	4	12,210.00	1,854.00	5,373.00	19,442.00	
8	01-03-2021	4	13,409.00	2,388.00	6,780.00	22,582.00	
9	01-03-2022	4	14,028.00	3,229.00	5,799.00	23,060.00	
10	01-03-2023	4	14,703.00	3,014.00	6,648.00	24,369.00	
11	01-03-2024	4	16,699.00	2,822.00	7,325.00	26,849.00	
12	01-03-2025	4	18,484.00	3,771.00	7,307.00	29,567.00	
13	01-09-2025	4	19,437.00	3,631.00	7,248.00	30,320.00	

Query Settings

Properties Name: Balance Sheet All Properties

Applied Steps

- Source
- Extracted Table From Html
- Transposed Table
- Promoted Headers
- Changed Type

New Source Recent Sources Enter Data Data source settings Manage Parameters Refresh Preview Manage Choose Columns Remove Columns Keep Rows Remove Rows Reduce Rows Sort Split Column Group By Replace Values Data Type: Date Use First Row as

New Query Data Sources Parameters Query Manage Columns Reduce Rows Sort Transform

From Web

Basic Advanced

URL: <https://www.espncricinfo.com/live-cricket-scorecard/108793/aus-vs-eng-2nd-test-the-ashes-2025-26>

OK Cancel

	Date	Sales	Expenses	Operating Profit	OPM %	Other Income
● Valid	100%	● Valid	100%	● Valid	100%	● Valid
● Error	0%	● Error	0%	● Error	0%	● Error
● Empty	0%	● Empty	0%	● Empty	0%	● Empty
	13 distinct, 13 unique					
1	01-09-2022					
2	01-12-2022					
3	01-03-2023					
4	01-06-2023					
5	01-09-2023					
6	01-12-2023					
7	01-03-2024					
8	01-06-2024					

A ^B _C Column1	A ^B _C Column2	A ^B _C Column3	A ^B _C Column4
Valid Error Empty	95% 0% 5%	Valid Error Empty	97% 0% 3%
38 distinct, 37 unique	30 distinct, 22 unique	32 distinct, 26 unique	26 distinct, 18 unique
Model	Release(d)	Release(d)	Discontinued
null With OS	Date		
null	null		null
iPhone	iPhone OS 1.0	June 29, 2007 (2007-06-29)	June 9, 2008 (2008-06-09)
iPhone 3G	iPhone OS 2.0	July 11, 2008 (2008-07-11)	August 9, 2010 (2010-08-09)
iPhone 3GS	iPhone OS 3.0	June 19, 2009 (2009-06-19)	September 12, 2012 (2012-09-12)
iPhone 4	iOS 4.0	June 24, 2010 (2010-06-24)	September 10, 2013 (2013-09-10)
iPhone 4s	iOS 5.0	October 14, 2011 (2011-10-14)	September 9, 2014 (2014-09-09)
iPhone 5	iOS 6.0	September 21, 2012 (2012-09-21)	September 10, 2013 (2013-09-10)
iPhone 5c	iOS 7.0	September 20, 2013 (2013-09-20)	September 9, 2015 (2015-09-09)
iPhone 5s	iOS 7.0	September 20, 2013 (2013-09-20)	March 21, 2016 (2016-03-21)
iPhone 6 / 6s Plus	iOS 8.0	September 19, 2014 (2014-09-19)	September 7, 2016 (2016-09-07)
iPhone 6s / 6s Plus	iOS 9.0	September 25, 2015 (2015-09-25)	September 12, 2018 (2018-09-12)
iPhone SE (1st)	iOS 9.3	March 31, 2016 (2016-03-31)	September 12, 2018 (2018-09-12)
iPhone 7 / 7 Plus	iOS 10.0	September 16, 2016 (2016-09-16)	September 10, 2019 (2019-09-10)
iPhone 8 / 8 Plus	iOS 11.0	September 22, 2017 (2017-09-22)	April 15, 2020 (2020-04-15)
iPhone X	iOS 11.0.1	November 3, 2017 (2017-11-03)	September 12, 2018 (2018-09-12)
iPhone XS / XS Max	iOS 12.0	September 21, 2018 (2018-09-21)	September 10, 2019 (2019-09-10)

A^B_C Release(d)_1

- Valid
- Error
- Empty

27 distinct, 21 unique

June 29, 2007 (2007-06-29)
July 11, 2008 (2008-07-11)
June 19, 2009 (2009-06-19)
June 24, 2010 (2010-06-24)
October 14, 2011 (2011-10-14)

A^B_C Discontinued

Text Between Delimiters

Enter the delimiters that mark the beginning and end of what you would like to extract.

Start delimiter: (

End delimiter:)

Advanced options

PROPERTIES

Name: iPhone_Info

APPLIED STEPS

- Source
- Extracted Table From Html
- Promoted Headers
- Removed Top Rows
- Removed Bottom Rows
- Extracted Text Between Delim...

2007-06-29

2008-07-11

2009-06-19

2010-06-24

2011-10-14

A^B_C Discontinued

- Valid
- Error
- Empty

23 distinct, 15 unique

June 9, 2008 (2008-06-09)
August 9, 2010 (2010-08-09)
September 12, 2012 (2012-09-12)
September 10, 2013 (2013-09-10)
September 9, 2014 (2014-09-09)

A^B_C Support

Text Before Delimiter

Enter the delimiter that marks the end of what you would like to extract.

Delimiter: (

Advanced options

PROPERTIES

June 9, 2008

August 9, 2010

September 12, 2012

September 10, 2013

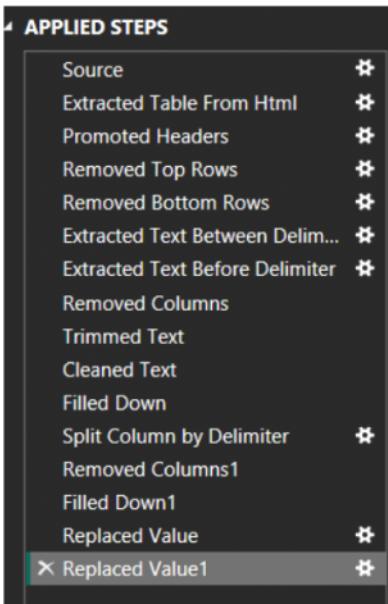
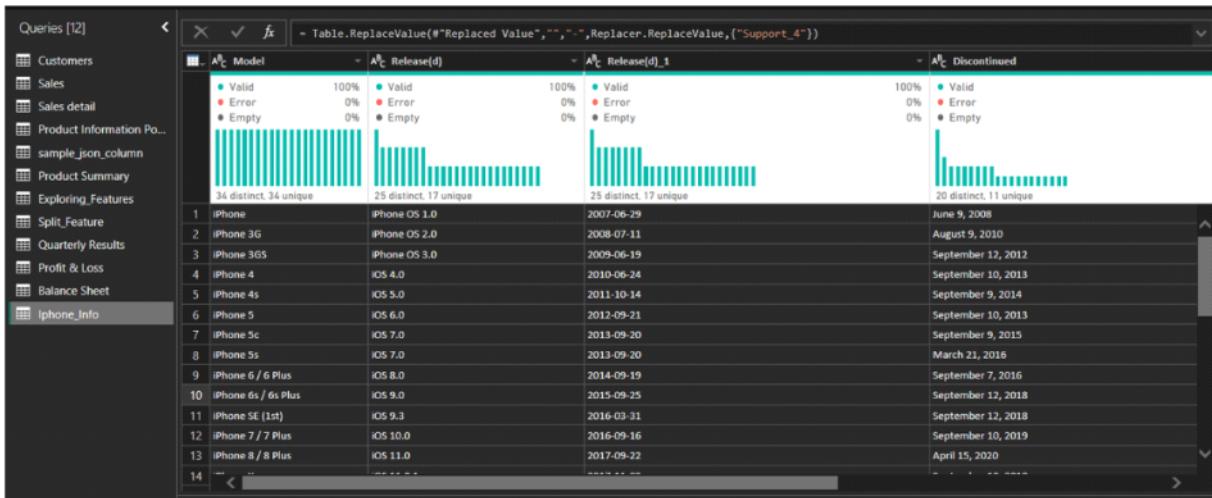
September 9, 2014

September 10, 2013

September 9, 2015

March 21, 2016

September 7, 2016



Append Queries

Sales2015

Sales2016

Sales2017

Sales2018

Appending [Top to Bottom]
[Table Vertically Increases]

Identical Table
[Have Same No.
Of Columns with
Similar Data Types]

Sales 2015 - 2018

Queries [13]

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%

1 21 North 10 6,532.00
2 22 North 23 7,958.00
3 23 South 56 1,012.00
4 24 East 38 5,816.00
5 25 East 22 5,631.00
6 26 West 22 3,715.00
7 27 West 58 8,314.00
8 28 East 11 3,575.00
9 29 East 43 7,459.00
10 30 North 45 1,607.00
11 31 North 14 6,078.00
12 32 South 27 9,065.00
13 33 East 57 4,423.00
14 34 East 17 850.00
15 35 East 53 3,701.00
16 36 East 59 5,873.00
17 37 North 49 4,385.00
18 38 North 33 9,561.00
19 39 South 38 2,322.00
20 40 East 59 4,881.00

Sales - part 2

Properties

Name: Sales - part 2

Applied Steps

Source: Navigation Promoted Headers: Changed Type: Filtered Rows

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%

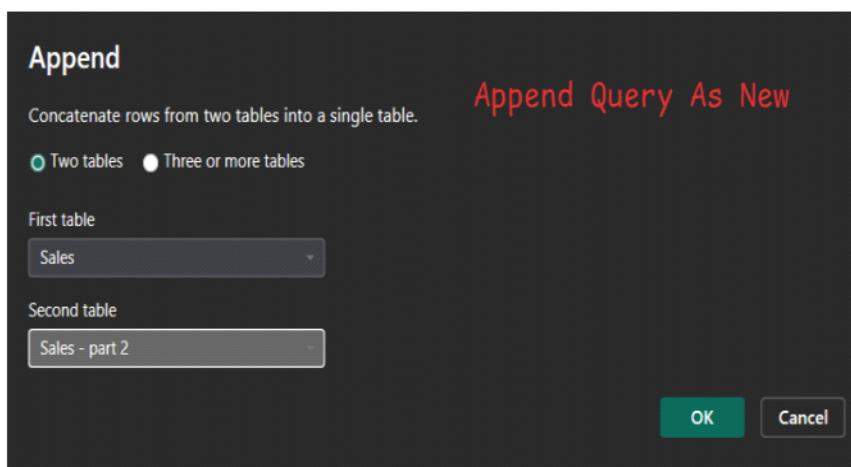
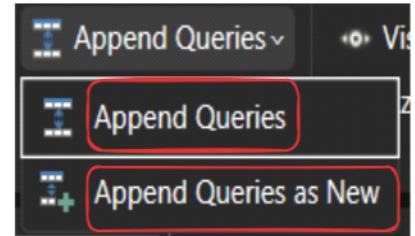
1 North 49 309.38
2 North 10 57.71
3 North 47 686.95
4 South 11 1,619.19
5 East 17 1,005.90
6 East 23 174.65
7 East 21 825.00
8 East 27 250.00
9 East 44 9.03
10 East 10 255.84
11 North 29 151.24
12 South 29 251.72
13 South 42 1,139.43
14 North 12 575.36
15 North 45 18.06
16 North 30 299.85
17 North 35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

Customer ID

100 (10%)	0 (0%)	899 (90%)
Valid	Error	Empty

Remove Empty

	Customer ID		Customer Region		Customer Age	\$ Sales	
	100%		100%		100%		100%
	0%		0%		0%		0%
	0%		0%		0%		0%
100 distinct, 100 unique		4 distinct, 0 unique		43 distinct, 15 unique		100 distinct, 100 unique	
21	North			10		6,532.00	
22	North			23		7,958.00	
23	South			56		1,012.00	
24	East			38		5,816.00	
25	East			22		5,651.00	
26	West			22		3,215.00	
27	West			58		8,314.00	
28	East			11		3,575.00	
29	East			43		7,459.00	
30	North			45		1,607.00	
31	North			14		6,078.00	
32	South			27		9,065.00	
33	East			57		4,423.00	
34	East			17		850.00	
35	East			53		3,701.00	
36	East			59		5,873.00	
37	North			49		4,385.00	
38	North			33		9,561.00	
39	South			38		2,322.00	
40	East			59		4,881.00	



Queries [14]

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

Query Settings

Properties

Name: SalesRecord

All Properties

Applied Steps

Source

Load

- Quarterly Results Evaluating...
- Profit & Loss Evaluating...
- Balance Sheet Evaluating...
- Iphone_Info Evaluating...
- Sales - part 2 Evaluating...

Data

Search

- > Balance Sheet
- > Customers
- > Exploring_Features
- > Iphone_Info
- > Product Information P...
- > Product Summary
- > Profit & Loss
- > Quarterly Results
- > Sales
- > Sales - part 2
- > Sales detail
- > SalesRecord
- > sample_json_column
- > Split_Feature