

DATA PREPARATION TALEND

talend DATA PREPARATION

HOME

PREPARATIONS

ADD PREPARATION

ADD FOLDER

DATASETS

NAME

ecommerce\_customer\_data\_add on city Preparation

ecommerce\_customer\_data\_DIRTY Preparation

hypothetical\_customer\_data\_with\_missing\_ids\_v1 Preparation

AUTHOR

CREATED

MODIFIED

DATASET

STEPS

talend DATA PREPARATION

ecommerce\_customer\_data\_DIRTY Preparation

1 Replace the cells that match on column Payment Method

2 Clear the cells with invalid values on column Customer ID

3 Delete the rows with empty cell on column Customer ID

4 Delete the rows with empty cell on column Returns

5 Extract date parts on column Purchase Date

6 Change date format on column Purchase Date

7 Delete column on column Purchase Date\_DAY

8 Delete column on column Purchase Date\_MONTH

9 Delete column on column Purchase Date\_YEAR

Filters

19476/19476

Customer ID

fr\_postal\_code

Purchase Date

date

Product Category

text

Product Price

integer

Quantity

integer

Total Purch

1	44605	03-05-2023	Home	177	1	
2	44605	16-05-2021	Electronics	174	3	
3	44605	13-07-2020	Books	413	1	
4	44605	17-01-2023	Electronics	396	3	
5	44605	01-05-2021	Books	259	4	
6	13738	25-08-2022	Home	191	3	
8	13738	05-02-2023	Books	370	5	
10	13738	09-02-2023	Electronics	40	4	
12	33969	05-01-2023	Home	304	1	
13	33969	18-07-2023	Books	54	2	
14	33969	20-12-2021	Electronics	428	4	
15	33969	07-03-2020	Books	281	1	
16	33969	21-07-2022	Home	193	2	
17	33969	05-07-2023	Clothing	473	3	
18	42650	18-10-2020	Books	127	5	
19	42650	17-05-2020	Home	284	2	
20	42650	18-03-2022	Electronics	256	2	
21	42650	21-04-2021	Clothing	252	2	
23	42650	11-11-2020	Electronics	193	2	

Customer ID

COLUMN ROW

Find a function ...

SUGGESTIONS

Compare numbers...

Add, multiply, subtract or divide...

BOOLEAN

Negate value

COLUMNS

CHART VALUE PATTERN ADVANCED

ROW COUNT

Occurrences

20,000 30,000 40,000 50,000

Filters

19476/19476

Customer ID

fr\_postal\_code

Purchase Date

date

Product Category

text

Product Price

integer

Quantity

integer

Total Purch

1	44605	03-05-2023	Home	177	1	
2	44605	16-05-2021	Electronics	174	3	
3	44605	13-07-2020	Books	413	1	
4	44605	17-01-2023	Electronics	396	3	
5	44605	01-05-2021	Books	259	4	
6	13738	25-08-2022	Home	191	3	
8	13738	05-02-2023	Books	370	5	
10	13738	09-02-2023	Electronics	40	4	
12	33969	05-01-2023	Home	304	1	
13	33969	18-07-2023	Books	54	2	
14	33969	20-12-2021	Electronics	428	4	
15	33969	07-03-2020	Books	281	1	
16	33969	21-07-2022	Home	193	2	
17	33969	05-07-2023	Clothing	473	3	
18	42650	18-10-2020	Books	127	5	
19	42650	17-05-2020	Home	284	2	
20	42650	18-03-2022	Electronics	256	2	
21	42650	21-04-2021	Clothing	252	2	

Customer ID

COLUMN ROW

Find a function ...

SUGGESTIONS

Compare numbers...

Add, multiply, subtract or divide...

BOOLEAN

Negate value

COLUMNS

CHART VALUE PATTERN ADVANCED

99999