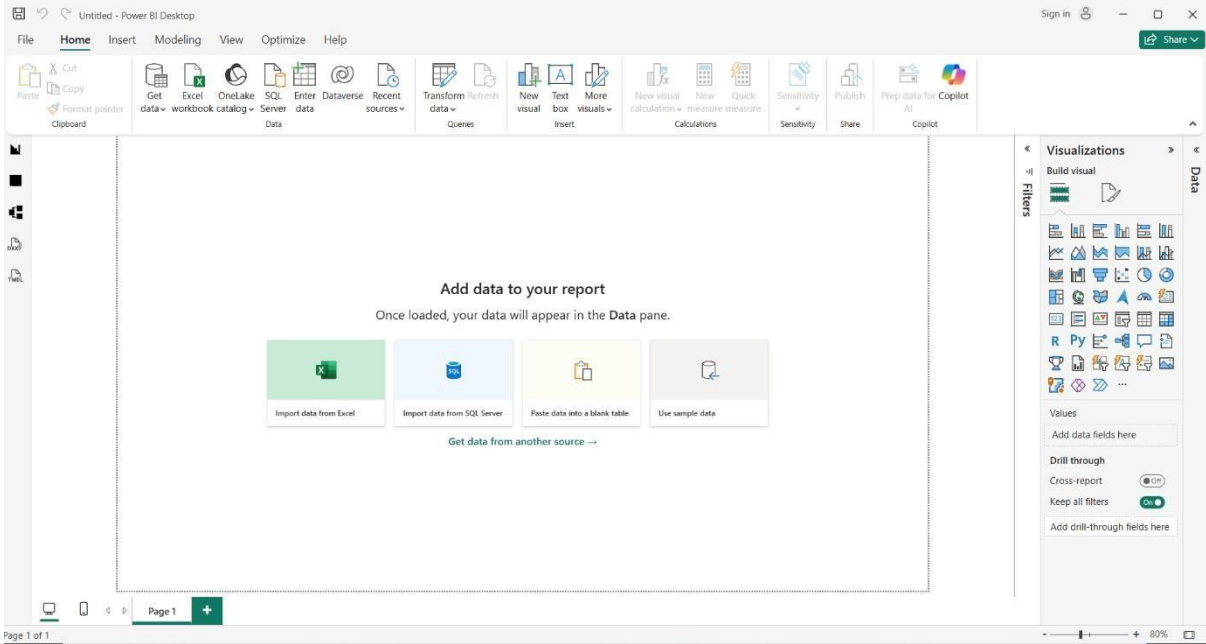


DATA CLEANING IN POWER BI [POWER QUERY]



1.Power Bi Dashboard

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R		
disasterNumber	declarationDate	disasterName	incidentBe	incidentEn	declaratio	stateCode	stateName	incidentTy	entryDate	updateDat	closeoutD	region	ihProgram	iaProgram	paProgram	hmProgar	designatetc	dec	
5243	2018-06-22T00:00:00.000Z	GRAHAM FIRE	2018-06-2	2018-06-2	Fire Manaj	OR	Oregon	Fire	2018-06-2	2025-03-1	2025-03-1	10	0	0	0	1	1	201	
5554	2025-03-07T00:00:00.000Z	COVINGTON DRIVE FIRE	2025-03-01T00:00:00.000Z	2025-03-01T00:00:00.000Z	Fire Manaj	SC	South Car	Fire	2025-03-1	2025-03-10T00:00:00.000Z	2025-03-10T00:00:00.000Z	4	0	0	0	1	1 R	202	
4859	2025-01-15T00:00:00.000Z	SEVERE STORM AND FLOODING	2024-10-2	2024-10-2	Major Dis	AK	Alaska	Severe Sto	2025-01-1	2025-01-16T00:00:00.000Z	2025-01-16T00:00:00.000Z	10	0	0	0	1	1 5,W	202	
4856	2025-01-08T00:00:00.000Z	WILDFIRES AND STRAIGHT-LINE WINDS	2025-01-0	2025-01-3	Major Dis	CA	California	Fire	2025-01-0	2025-02-18T00:00:00.000Z	2025-02-18T00:00:00.000Z	9	1	0	0	1	1 2,R	202	
5551	2025-01-08T00:00:00.000Z	HURST FIRE	2025-01-07T00:00:00.000Z	2025-01-07T00:00:00.000Z	Fire Manaj	CA	California	Fire	2025-01-0	2025-01-08T00:00:00.000Z	2025-01-08T00:00:00.000Z	9	0	0	0	1	1 R	202	
5550	2025-01-08T00:00:00.000Z	EATON FIRE	2025-01-07T00:00:00.000Z	2025-01-07T00:00:00.000Z	Fire Manaj	CA	California	Fire	2025-01-0	2025-01-08T00:00:00.000Z	2025-01-08T00:00:00.000Z	9	0	0	0	1	1 R	202	
5549	2025-01-07T00:00:00.000Z	PALISADES FIRE	2025-01-07T00:00:00.000Z	2025-01-07T00:00:00.000Z	Fire Manaj	CA	California	Fire	2025-01-0	2025-01-08T00:00:00.000Z	2025-01-08T00:00:00.000Z	9	0	0	0	1	1 R	202	
4854	2025-01-01T00:00:00.000Z	WILDFIRES	2024-07-1	2024-08-2	Major Dis	OR	Oregon	Fire	2025-01-0	2025-01-02T00:00:00.000Z	2025-01-02T00:00:00.000Z	10	0	0	0	1	1 R,W	202	
53	1956-04-05T00:00:00.000Z	TORNADO	1956-04-0	1956-04-0	Major Dis	MI	Michigan	Tornado	1993-07-2	2001-09-0	1956-04-3	5	0	0	1	1	1	195	
52	1956-03-29T00:00:00.000Z	FLOOD	1956-03-2	1956-03-2	Major Dis	NY	New York	Flood	1993-07-2	2001-09-0	1957-03-0	2	0	0	1	1	1	195	
4852	2024-12-24T00:00:00.000Z	WILDFIRES AND STRAIGHT-LINE WINDS	2024-10-0	2024-10-0	Major Dis	ND	North Dak	Fire	2024-12-2	2024-12-25T00:00:00.000Z	2024-12-25T00:00:00.000Z	8	0	0	0	1	1 2,R	202	
5548	2024-12-10T00:00:00.000Z	FRANKLIN FIRE	2024-12-1	2024-12-1	Fire Manaj	CA	California	Fire	2024-12-1	2025-02-07T00:00:00.000Z	2025-02-07T00:00:00.000Z	9	0	0	0	1	1 R	202	
4851	2024-12-09T00:00:00.000Z	POST-TROPICAL STORM HELENE	2024-09-2	2024-09-2	Major Dis	WV	West Virgi	Tropical St	2024-12-0	2024-12-09T00:00:00.000Z	2024-12-09T00:00:00.000Z	3	1	0	0	0	1	4	202
4849	2024-11-26T00:00:00.000Z	WILDFIRES	2024-07-1	2024-08-2	Major Dis	WA	Washingtc	Fire	2024-11-2	2024-11-27T00:00:00.000Z	2024-11-27T00:00:00.000Z	10	0	0	0	1	1 R	202	
4847	2024-11-14T00:00:00.000Z	SEVERE STORM AND STRAIGHT-LINE WI	2024-08-0	2024-08-0	Major Dis	MT	Montana	Severe Sto	2024-11-1	2024-11-14T00:00:00.000Z	2024-11-14T00:00:00.000Z	8	1	0	0	0	1 2,W	202	
4846	2024-11-13T00:00:00.000Z	LANDSLIDES	2024-08-2	2024-08-2	Major Dis	AK	Alaska	Mud/Land	2024-11-1	2024-11-14T00:00:00.000Z	2024-11-14T00:00:00.000Z	10	0	0	0	1	1 M	202	
4845	2024-11-13T00:00:00.000Z	WILDFIRES	2024-08-2	2024-08-3	Major Dis	WY	Wyoming	Fire	2024-11-1	2024-11-14T00:00:00.000Z	2024-11-14T00:00:00.000Z	8	0	0	0	1	1 R	202	
5546	2024-11-11T00:00:00.000Z	CALLAHAN FIRE	2024-11-1	2024-11-1	Fire Manaj	NV	Nevada	Fire	2024-11-1	2025-02-07T00:00:00.000Z	2025-02-07T00:00:00.000Z	9	0	0	0	1	1 R	202	
5545	2024-11-06T00:00:00.000Z	MOUNTAIN FIRE	2024-11-0	2024-11-1	Fire Manaj	CA	California	Fire	2024-11-1	2025-02-07T00:00:00.000Z	2025-02-07T00:00:00.000Z	9	0	0	0	1	1 R	202	
4844	2024-11-05T00:00:00.000Z	HURRICANE MILTON	2024-10-0	2024-11-0	Major Dis	FL	Florida	Hurricane	2024-11-0	2024-11-20T00:00:00.000Z	2024-11-20T00:00:00.000Z	4	1	0	0	1	1 H	202	
4843	2024-11-01T00:00:00.000Z	SEVERE STORM AND FLOODING	2024-10-1	2024-10-2	Major Dis	NM	New Mexi	Flood	2024-11-0	2024-11-01T00:00:00.000Z	2024-11-01T00:00:00.000Z	6	1	0	0	1	1 F	202	
5544	2024-10-30T00:00:00.000Z	NORTH ROAD FIRE	2024-10-29T00:00:00.000Z	2024-10-29T00:00:00.000Z	Fire Manaj	OK	Oklahoma	Fire	2024-10-3	2024-10-30T00:00:00.000Z	2024-10-30T00:00:00.000Z	6	0	0	0	1	1 R	202	
5543	2024-10-30T00:00:00.000Z	EUCHEE CREEK FIRE	2024-10-29T00:00:00.000Z	2024-10-29T00:00:00.000Z	Fire Manaj	OK	Oklahoma	Fire	2024-10-3	2024-10-30T00:00:00.000Z	2024-10-30T00:00:00.000Z	6	0	0	0	1	1 R	202	
51	1956-03-15T00:00:00.000Z	FLOOD	1956-03-1	1956-03-1	Major Dis	PA	Pennsylvai	Flood	1993-07-2	2001-09-0	1959-06-0	3	0	0	1	1	1	195	
4841	2024-10-25T00:00:00.000Z	TROPICAL STORM ERNESTO	2024-08-1	2024-08-1	Major Dis	VI	Virgin Islar	Tropical St	2024-10-2	2024-10-25T00:00:00.000Z	2024-10-25T00:00:00.000Z	2	0	0	0	1	1 4,W	202	
4840	2024-10-25T00:00:00.000Z	FLOODING	2024-08-2	2024-08-2	Major Dis	AZ	Arizona	Flood	2024-10-2	2024-10-25T00:00:00.000Z	2024-10-25T00:00:00.000Z	9	1	0	0	1	1 F	202	
5542	2024-10-24T00:00:00.000Z	HAWTHORNE FIRE	2024-10-21T00:00:00.000Z	2024-10-21T00:00:00.000Z	Fire Manaj	CT	Connectici	Fire	2024-10-2	2024-10-25T00:00:00.000Z	2024-10-25T00:00:00.000Z	1	0	0	0	1	1 R	202	
4836	2024-10-16T00:00:00.000Z	FLOODING	2024-08-0	2024-08-0	Major Dis	AK	Alaska	Flood	2024-10-1	2024-10-16T00:00:00.000Z	2024-10-16T00:00:00.000Z	10	1	0	0	1	1 F	202	
4834	2024-10-11T00:00:00.000Z	HURRICANE MILTON	2024-10-0	2024-11-0	Major Dis	FL	Florida	Hurricane	2024-10-1	2024-11-04T00:00:00.000Z	2024-11-04T00:00:00.000Z	4	1	0	0	1	1 F,H	202	
5540	2024-10-06T00:00:00.000Z	BEAR DEN FIRE	2024-10-0	2024-10-2	Fire Manaj	ND	North Dak	Fire	2024-10-1	2025-01-24T00:00:00.000Z	2025-01-24T00:00:00.000Z	8	0	0	0	1	1 R	202	
4833	2024-10-04T00:00:00.000Z	WATCH FIRE	2024-07-1	2024-07-1	Maior Dis	AZ	Arizona	Fire	2024-10-0	2024-10-04T00:00:00.000Z	2024-10-04T00:00:00.000Z	9	1	0	0	1	1 R	202	

2.Raw Data

DATA CLEANING IN POWER BI [POWER QUERY]

FemaWebDisasterDeclarations.csv

File Origin

1252: Western European (Windows)

Delimiter

Comma

Data Type Detection

Based on first 200 rows

dDate	declarationType	stateCode	stateName	incidentType	entryDate	updateDate	closeoutDate	region
05:30:00	Fire Management	OR	Oregon	Fire	22-06-2018 05:30:00	13-03-2025 05:30:00	13-03-2025 05:30:00	10
	Fire Management	SC	South Carolina	Fire	10-03-2025 05:30:00	10-03-2025 05:30:00	null	4
05:30:00	Major Disaster	AK	Alaska	Severe Storm	16-01-2025 05:30:00	16-01-2025 05:30:00	null	10
05:30:00	Major Disaster	CA	California	Fire	09-01-2025 05:30:00	18-02-2025 05:30:00	null	9
	Fire Management	CA	California	Fire	08-01-2025 05:30:00	08-01-2025 05:30:00	null	9
	Fire Management	CA	California	Fire	08-01-2025 05:30:00	08-01-2025 05:30:00	null	9
	Fire Management	CA	California	Fire	08-01-2025 05:30:00	08-01-2025 05:30:00	null	9
05:30:00	Major Disaster	OR	Oregon	Fire	02-01-2025 05:30:00	02-01-2025 05:30:00	null	10
05:30:00	Major Disaster	MI	Michigan	Tornado	21-07-1993 05:30:00	09-09-2001 05:30:00	30-04-1996 05:30:00	5
05:30:00	Major Disaster	NY	New York	Flood	21-07-1993 05:30:00	09-09-2001 05:30:00	01-03-1957 05:30:00	2
05:30:00	Major Disaster	ND	North Dakota	Fire	25-12-2024 05:30:00	25-12-2024 05:30:00	null	8
05:30:00	Fire Management	CA	California	Fire	10-12-2024 05:30:00	07-02-2025 05:30:00	null	9
05:30:00	Major Disaster	WV	West Virginia	Tropical Storm	09-12-2024 05:30:00	09-12-2024 05:30:00	null	3
05:30:00	Major Disaster	WA	Washington	Fire	27-11-2024 05:30:00	27-11-2024 05:30:00	null	10
05:30:00	Major Disaster	MT	Montana	Severe Storm	14-11-2024 05:30:00	14-11-2024 05:30:00	null	8
05:30:00	Major Disaster	AK	Alaska	Mud/Landslide	14-11-2024 05:30:00	14-11-2024 05:30:00	null	10
05:30:00	Major Disaster	WY	Wyoming	Fire	14-11-2024 05:30:00	14-11-2024 05:30:00	null	8
05:30:00	Fire Management	NV	Nevada	Fire	13-11-2024 05:30:00	07-02-2025 05:30:00	null	9
05:30:00	Fire Management	CA	California	Fire	12-11-2024 05:30:00	07-02-2025 05:30:00	null	9
05:30:00	Major Disaster	FL	Florida	Hurricane	06-11-2024 05:30:00	20-11-2024 05:30:00	null	4

Extract Table Using Examples

Load

Transform Data

Cancel

Query Settings

×

PROPERTIES

Name

FemaWebDisasterDeclarations

All Properties

APPLIED STEPS

Renamed Columns8

Changed Type12

Renamed Columns9

Replaced Value

Changed Type13

Renamed Columns10

Changed Type14

Removed Columns

Renamed Columns11

Changed Type15

Renamed Columns12

Replaced Value1

Removed Columns1

Changed Type16

Filtered Rows1

Changed Type17

Filtered Rows2

Replaced Value2

Changed Type18

Replaced Value3

Uppercased Text3

3. Import/Load The Data

Home

Transform

Add Column

View

Tools

Help

Close & Apply

New Source

Recent Sources

Enter Data

Data source settings

Manage Parameters

Export query results

Refresh Preview

Properties

Advanced Editor

Manage

Choose Columns

Remove Columns

Keep Rows

Remove Rows

Sort

Split Column

Group By

Replace Values

Data Type: Text

Use First Row as Headers

Merge Queries

Append Queries

Combine Files

Combine

Close

New Query

Data Sources

Parameters

Output Data

Query

Manage Columns

Reduce Rows

Transform

×

✓

fx

= Table.TransformColumns(Table.TransformColumnTypes("#Replaced Value3", {"Disaster Number", type text}, {"Declaration Date", type text}, {"Incident Begin Date", type text},

Queries

	Disaster Number	Declaration Date	Disaster Name	Incident Begin Date	Incident End Date	Declaration Type	State Code	State Name
1	1	02-05-1953	TORNADO	02-05-1953	02-05-1953	MAJOR DISASTER	GA	GEORGIA
2	2	15-05-1953	TORNADO & HEAVY RAINFALL	15-05-1953	15-05-1953	MAJOR DISASTER	TX	TEXAS
3	3	29-05-1953	FLOOD	29-05-1953	29-05-1953	MAJOR DISASTER	LA	LOUISIANA
4	4	02-06-1953	TORNADO	02-06-1953	02-06-1953	MAJOR DISASTER	MI	MICHIGAN
5	5	06-06-1953	FLOODS	06-06-1953	06-06-1953	MAJOR DISASTER	MT	MONTANA
6	6	09-06-1953	TORNADO	09-06-1953	09-06-1953	MAJOR DISASTER	MI	MICHIGAN
7	7	11-06-1953	TORNADO	11-06-1953	11-06-1953	MAJOR DISASTER	MA	MASSACHUSETTS
8	8	11-06-1953	FLOOD	11-06-1953	11-06-1953	MAJOR DISASTER	IA	IOWA
9	9	19-06-1953	FLOOD	19-06-1953	19-06-1953	MAJOR DISASTER	TX	TEXAS
10	11	02-07-1953	FOREST FIRE	02-07-1953	02-07-1953	MAJOR DISASTER	NH	NEW HAMPSHIRE
11	12	22-10-1953	FLOOD	22-10-1953	22-10-1953	MAJOR DISASTER	FL	FLORIDA
12	13	30-10-1953	SEVERE HARDSHIP	30-10-1953	30-10-1953	MAJOR DISASTER	AK	ALASKA
13	14	06-12-1953	TORNADO	06-12-1953	06-12-1953	MAJOR DISASTER	MS	MISSISSIPPI
14	15	05-02-1954	FLOOD & EROSION	05-02-1954	05-02-1954	MAJOR DISASTER	CA	CALIFORNIA
15	16	17-03-1954	TORNADO	17-03-1954	17-03-1954	MAJOR DISASTER	GA	GEORGIA
16	17	23-06-1954	FLOOD	23-06-1954	23-06-1954	MAJOR DISASTER	IA	IOWA
17	18	01-07-1954	FLOOD	01-07-1954	01-07-1954	MAJOR DISASTER	TX	TEXAS
18	19	14-07-1954	EARTHQUAKE	14-07-1954	14-07-1954	MAJOR DISASTER	NV	NEVADA
19	20	31-07-1954	FLOOD	31-07-1954	31-07-1954	MAJOR DISASTER	SD	SOUTH DAKOTA
20	21	04-08-1954	FLOOD	04-08-1954	04-08-1954	MAJOR DISASTER	WV	WEST VIRGINIA
21	22	02-09-1954	HURRICANES	02-09-1954	02-09-1954	MAJOR DISASTER	MA	MASSACHUSETTS
22	23	02-09-1954	HURRICANES	02-09-1954	02-09-1954	MAJOR DISASTER	RI	RHODE ISLAND
23	24	13-09-1954	HURRICANES	13-09-1954	13-09-1954	MAJOR DISASTER	ME	MAINE
24	25	17-09-1954	HURRICANES	17-09-1954	17-09-1954	MAJOR DISASTER	CT	CONNECTICUT
25	26	07-10-1954	HURRICANES	07-10-1954	07-10-1954	MAJOR DISASTER	NY	NEW YORK
26	27	12-10-1954	FLOOD	12-10-1954	12-10-1954	MAJOR DISASTER	MA	MASSACHUSETTS

5.Cleaned Data

DATA CLEANING USING PYTHON [VS CODE, JUPYTER NOTEBOOK]

File Edit Selection View Go Run Terminal Help

Sample.ipynb X

C: > Users > User > OneDrive > Desktop > INTERNSHIP 6.0 > Sample.ipynb > import pandas as pd

Generate + Code + Markdown | Run All | Restart | Clear All Outputs | Jupyter Variables | Outline

```
import pandas as pd
df = pd.read_excel(r"C:\Users\User\OneDrive\Desktop\Customer Call List.xlsx")
df
```

[8] ✓ 0.8s

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Not_Useful_Column
0	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Yes	No	True
1	1002	Abed	Nadir	123/643/9775	93 West Main Street	No	Yes	False
2	1003	Walter	/White	7066950392	298 Drugs Driveway	N	NaN	True
3	1004	Dwight	Schrute	123-543-2345	980 Paper Avenue, Pennsylvania, 18503	Yes	Y	True
4	1005	Jon	Snow	876[678]3469	123 Dragons Road	Y	No	True
5	1006	Ron	Swanson	304-762-2467	768 City Parkway	Yes	Yes	True
6	1007	Jeff	Winger	NaN	1209 South Street	No	No	False
7	1008	Sherlock	Holmes	876[678]3469	98 Clue Drive	N	No	False
8	1009	Gandalf	NaN	N/a	123 Middle Earth	Yes	NaN	False
9	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Yes	No	True
10	1011	Samwise	Gamgee	NaN	612 Shire Lane, Shire	Yes	No	True
11	1012	Harry	...Potter	7066950392	2394 Hogwarts Avenue	Y	NaN	True
12	1013	Don	Draper	123-543-2345	2039 Main Street	Yes	N	False
13	1014	Leslie	Knope	876[678]3469	343 City Parkway	Yes	No	False
14	1015	Toby	Flenderson_	304-762-2467	214 HR Avenue	N	No	False
15	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	No	N	False
16	1017	Michael	Scott	123/643/9775	121 Paper Avenue, Pennsylvania	Yes	No	False
17	1018	Clark	Kent	7066950392	3498 Super Lane	Y	NaN	True
18	1019	Creed	Braton	N/a	N/a	N/a	Yes	True
19	1020	Anakin	Skywalker	876[678]3469	910 Tatooine Road, Tatooine	Yes	N	True
20	1020	Anakin	Skywalker	876[678]3469	910 Tatooine Road, Tatooine	Yes	N	True

1.Importing/Load Data

```
df.drop_duplicates()
```

[10] ✓ 0.0s

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Not_Useful_Column
0	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Yes	No	True
1	1002	Abed	Nadir	123/643/9775	93 West Main Street	No	Yes	False
2	1003	Walter	/White	7066950392	298 Drugs Driveway	N	NaN	True
3	1004	Dwight	Schrute	123-543-2345	980 Paper Avenue, Pennsylvania, 18503	Yes	Y	True
4	1005	Jon	Snow	876[678]3469	123 Dragons Road	Y	No	True
5	1006	Ron	Swanson	304-762-2467	768 City Parkway	Yes	Yes	True
6	1007	Jeff	Winger	NaN	1209 South Street	No	No	False
7	1008	Sherlock	Holmes	876[678]3469	98 Clue Drive	N	No	False
8	1009	Gandalf	NaN	N/a	123 Middle Earth	Yes	NaN	False
9	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Yes	No	True
10	1011	Samwise	Gamgee	NaN	612 Shire Lane, Shire	Yes	No	True
11	1012	Harry	...Potter	7066950392	2394 Hogwarts Avenue	Y	NaN	True
12	1013	Don	Draper	123-543-2345	2039 Main Street	Yes	N	False
13	1014	Leslie	Knope	876[678]3469	343 City Parkway	Yes	No	False
14	1015	Toby	Flenderson_	304-762-2467	214 HR Avenue	N	No	False
15	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	No	N	False
16	1017	Michael	Scott	123/643/9775	121 Paper Avenue, Pennsylvania	Yes	No	False
17	1018	Clark	Kent	7066950392	3498 Super Lane	Y	NaN	True
18	1019	Creed	Braton	N/a	N/a	N/a	Yes	True
19	1020	Anakin	Skywalker	876[678]3469	910 Tatooine Road, Tatooine	Yes	N	True

2.Removing Duplicates


```
df.columns
✓ 0.0s

Index(['CustomerID', 'First_Name', 'Last_Name', 'Phone_Number', 'Address',
      'Paying Customer', 'Do_Not_Contact', 'Not_Useful_Column'],
      dtype='object')
```

3.No.of Columns

```
df = df.drop(columns = ["Not_Useful_Column"])
✓ 0.0s

df["Last_Name"] = df["Last_Name"].str.strip("/..._")
✓ 0.0s

df
✓ 0.0s
```

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact
0	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Yes	No
1	1002	Abed	Nadir	123/643/9775	93 West Main Street	No	Yes
2	1003	Walter	White	7066950392	298 Drugs Driveway	N	NaN
3	1004	Dwight	Schrute	123-543-2345	980 Paper Avenue, Pennsylvania, 18503	Yes	Y
4	1005	Jon	Snow	876[678]3469	123 Dragons Road	Y	No
5	1006	Ron	Swanson	304-762-2467	768 City Parkway	Yes	Yes
6	1007	Jeff	Winger	NaN	1209 South Street	No	No
7	1008	Sherlock	Holmes	876[678]3469	98 Clue Drive	N	No
8	1009	Gandalf	NaN	N/a	123 Middle Earth	Yes	NaN
9	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Yes	No
10	1011	Samwise	Gamgee	NaN	612 Shire Lane, Shire	Yes	No
11	1012	Harry	Potter	7066950392	2394 Hogwarts Avenue	Y	NaN
12	1013	Don	Draper	123-543-2345	2039 Main Street	Yes	N
13	1014	Leslie	Knope	876[678]3469	343 City Parkway	Yes	No
14	1015	Toby	Flenderson	304-762-2467	214 HR Avenue	N	No
15	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	No	N
16	1017	Michael	Scott	123/643/9775	121 Paper Avenue, Pennsylvania	Yes	No
17	1018	Clark	Kent	7066950392	3498 Super Lane	Y	NaN
18	1019	Creed	Braton	N/a	N/a	N/a	Yes
19	1020	Anakin	Skywalker	876[678]3469	910 Tatooine Road, Tatooine	Yes	N

4.Removing Unwanted Columns

```
df = df.rename(columns={"Last_Name" : "Test_Name"})
df
✓ 0.0s
```

	CustomerID	First_Name	Test_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Street	City
0	1001	Frodo	Baggins	1235455421	123 Shire Lane, Shire	Y	N	123 Shire Lane	Shire
4	1005	Jon	Snow	8766783469	123 Dragons Road	Y	N	123 Dragons Road	None
7	1008	Sherlock	Holmes	8766783469	98 Clue Drive	N	N	98 Clue Drive	None
9	1010	Peter	Parker	1235455421	25th Main Street, New York	Y	N	25th Main Street	New York
12	1013	Don	Draper	1235432345	2039 Main Street	Y	N	2039 Main Street	None
13	1014	Leslie	Knope	8766783469	343 City Parkway	Y	N	343 City Parkway	None
14	1015	Toby	Flenderson	3047622467	214 HR Avenue	N	N	214 HR Avenue	None
15	1016	Ron	Weasley	1235455421	2395 Hogwarts Avenue	N	N	2395 Hogwarts Avenue	None
16	1017	Michael	Scott	1236439775	121 Paper Avenue, Pennsylvania	Y	N	121 Paper Avenue	Pennsylvania
19	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Y	N	910 Tatooine Road	Tatooine
20	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Y	N	910 Tatooine Road	Tatooine

5.Rename Columns

DATA CLEANING USING PYTHON [VS CODE, JUPYTER NOTEBOOK]

```
df["Phone_Number"] = df["Phone_Number"].str.replace("-", "")
df["Phone_Number"] = df["Phone_Number"].str.replace("/", "")
df["Phone_Number"] = df["Phone_Number"].str.replace("|", "")
df
```

[14] ✓ 0.0s

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact
0	1001	Frodo	Baggins	1235455421	123 Shire Lane, Shire	Yes	No
1	1002	Abed	Nadir	1236439775	93 West Main Street	No	Yes
2	1003	Walter	White	NaN	298 Drugs Driveway	N	NaN
3	1004	Dwight	Schrute	1235432345	980 Paper Avenue, Pennsylvania, 18503	Yes	Y
4	1005	Jon	Snow	8766783469	123 Dragons Road	Y	No
5	1006	Ron	Swanson	3047622467	768 City Parkway	Yes	Yes
6	1007	Jeff	Winger	NaN	1209 South Street	No	No
7	1008	Sherlock	Holmes	8766783469	98 Clue Drive	N	No
8	1009	Gandalf	NaN	Na	123 Middle Earth	Yes	NaN
9	1010	Peter	Parker	1235455421	25th Main Street, New York	Yes	No
10	1011	Samwise	Gamgee	NaN	612 Shire Lane, Shire	Yes	No
11	1012	Harry	Potter	NaN	2394 Hogwarts Avenue	Y	NaN
12	1013	Don	Draper	1235432345	2039 Main Street	Yes	N
13	1014	Leslie	Knope	8766783469	343 City Parkway	Yes	No
14	1015	Toby	Flenderson	3047622467	214 HR Avenue	N	No
15	1016	Ron	Weasley	1235455421	2395 Hogwarts Avenue	No	N
16	1017	Michael	Scott	1236439775	121 Paper Avenue, Pennsylvania	Yes	No
17	1018	Clark	Kent	NaN	3498 Super Lane	Y	NaN
18	1019	Creed	Braton	Na	N/a	N/a	Yes
19	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Yes	N
20	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Yes	N

```
df = df.fillna(" ")
```

[15] ✓ 0.0s

```
df = df.replace("Na", " ")
df = df.replace("N/a", " ")
df = df.replace("Yes", "Y")
df = df.replace(["No", "N"], "N")
df
```

[16] ✓ 0.0s

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact
0	1001	Frodo	Baggins	1235455421	123 Shire Lane, Shire	Y	N
1	1002	Abed	Nadir	1236439775	93 West Main Street	N	Y
2	1003	Walter	White		298 Drugs Driveway	N	
3	1004	Dwight	Schrute	1235432345	980 Paper Avenue, Pennsylvania, 18503	Y	Y
4	1005	Jon	Snow	8766783469	123 Dragons Road	Y	N
5	1006	Ron	Swanson	3047622467	768 City Parkway	Y	Y
6	1007	Jeff	Winger		1209 South Street	N	N
7	1008	Sherlock	Holmes	8766783469	98 Clue Drive	N	N
8	1009	Gandalf			123 Middle Earth	Y	
9	1010	Peter	Parker	1235455421	25th Main Street, New York	Y	N
10	1011	Samwise	Gamgee		612 Shire Lane, Shire	Y	N
11	1012	Harry	Potter		2394 Hogwarts Avenue	Y	
12	1013	Don	Draper	1235432345	2039 Main Street	Y	N
13	1014	Leslie	Knope	8766783469	343 City Parkway	Y	N
14	1015	Toby	Flenderson	3047622467	214 HR Avenue	N	N
15	1016	Ron	Weasley	1235455421	2395 Hogwarts Avenue	N	N
16	1017	Michael	Scott	1236439775	121 Paper Avenue, Pennsylvania	Y	N
17	1018	Clark	Kent		3498 Super Lane	Y	
18	1019	Creed	Braton				Y
19	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Y	N
20	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Y	N

```
df = df.replace("Y", "Yes")
df = df.replace("NO", "No")
df
```

✓ 0.0s

	CustomerID	First_Name	Test Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Street	City
0	1001	Frodo	Baggins	1235455421	123 Shire Lane, Shire	Yes	N	123 Shire Lane	Shire
4	1005	Jon	Snow	8766783469	123 Dragons Road	Yes	N	123 Dragons Road	None
7	1008	Sherlock	Holmes	8766783469	98 Clue Drive	N	N	98 Clue Drive	None
9	1010	Peter	Parker	1235455421	25th Main Street, New York	Yes	N	25th Main Street	New York
12	1013	Don	Draper	1235432345	2039 Main Street	Yes	N	2039 Main Street	None
13	1014	Leslie	Knope	8766783469	343 City Parkway	Yes	N	343 City Parkway	None
14	1015	Toby	Flenderson	3047622467	214 HR Avenue	N	N	214 HR Avenue	None
15	1016	Ron	Weasley	1235455421	2395 Hogwarts Avenue	N	N	2395 Hogwarts Avenue	None
16	1017	Michael	Scott	1236439775	121 Paper Avenue, Pennsylvania	Yes	N	121 Paper Avenue	Pennsylvania
19	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Yes	N	910 Tatooine Road	Tatooine
20	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Yes	N	910 Tatooine Road	Tatooine

6.Replacing Values

```
df.columns
```

✓ 0.0s

```
Index(['CustomerID', 'First_Name', 'Last_Name', 'Phone_Number', 'Address',  
      'Paying Customer', 'Do_Not_Contact', 'Street', 'City'],  
      dtype='object')
```

7.Reading No.of Column

```
df[["Street", "City"]] = df["Address"].str.split(",", expand=True)
df
```

✓ 0.0s

C:\Users\User\AppData\Local\Temp\ipykernel_17312\589710751.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
df[["Street", "City"]] = df["Address"].str.split(",", expand=True)

C:\Users\User\AppData\Local\Temp\ipykernel_17312\589710751.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
df[["Street", "City"]] = df["Address"].str.split(",", expand=True)

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Street	City
0	1001	Frodo	Baggins	1235455421	123 Shire Lane, Shire	Y	N	123 Shire Lane	Shire
4	1005	Jon	Snow	8766783469	123 Dragons Road	Y	N	123 Dragons Road	None
7	1008	Sherlock	Holmes	8766783469	98 Clue Drive	N	N	98 Clue Drive	None
9	1010	Peter	Parker	1235455421	25th Main Street, New York	Y	N	25th Main Street	New York
12	1013	Don	Draper	1235432345	2039 Main Street	Y	N	2039 Main Street	None
13	1014	Leslie	Knope	8766783469	343 City Parkway	Y	N	343 City Parkway	None
14	1015	Toby	Flenderson	3047622467	214 HR Avenue	N	N	214 HR Avenue	None
15	1016	Ron	Weasley	1235455421	2395 Hogwarts Avenue	N	N	2395 Hogwarts Avenue	None
16	1017	Michael	Scott	1236439775	121 Paper Avenue, Pennsylvania	Y	N	121 Paper Avenue	Pennsylvania
19	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Y	N	910 Tatooine Road	Tatooine
20	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Y	N	910 Tatooine Road	Tatooine

8. Splitting Column

DATA CLEANING USING PYTHON [VS CODE, JUPYTER NOTEBOOK]

```
df = df[df["Phone_Number"] != " "]
df
```

✓ 0.0s

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact
0	1001	Frodo	Baggins	1235455421	123 Shire Lane, Shire	Y	N
4	1005	Jon	Snow	8766783469	123 Dragons Road	Y	N
7	1008	Sherlock	Holmes	8766783469	98 Clue Drive	N	N
9	1010	Peter	Parker	1235455421	25th Main Street, New York	Y	N
12	1013	Don	Draper	1235432345	2039 Main Street	Y	N
13	1014	Leslie	Knope	8766783469	343 City Parkway	Y	N
14	1015	Toby	Flenderson	3047622467	214 HR Avenue	N	N
15	1016	Ron	Weasley	1235455421	2395 Hogwarts Avenue	N	N
16	1017	Michael	Scott	1236439775	121 Paper Avenue, Pennsylvania	Y	N
19	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Y	N
20	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Y	N

```
df.reset_index(drop=True)
```

Python

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Street	City
0	1001	Frodo	Baggins	1235455421	123 Shire Lane, Shire	Y	N	123 Shire Lane	Shire
1	1005	Jon	Snow	8766783469	123 Dragons Road	Y	N	123 Dragons Road	None
2	1008	Sherlock	Holmes	8766783469	98 Clue Drive	N	N	98 Clue Drive	None
3	1010	Peter	Parker	1235455421	25th Main Street, New York	Y	N	25th Main Street	New York
4	1013	Don	Draper	1235432345	2039 Main Street	Y	N	2039 Main Street	None
5	1014	Leslie	Knope	8766783469	343 City Parkway	Y	N	343 City Parkway	None
6	1015	Toby	Flenderson	3047622467	214 HR Avenue	N	N	214 HR Avenue	None
7	1016	Ron	Weasley	1235455421	2395 Hogwarts Avenue	N	N	2395 Hogwarts Avenue	None
8	1017	Michael	Scott	1236439775	121 Paper Avenue, Pennsylvania	Y	N	121 Paper Avenue	Pennsylvania
9	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Y	N	910 Tatooine Road	Tatooine
10	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Y	N	910 Tatooine Road	Tatooine

9.Condition Based

```
df
```

✓ 0.0s

	CustomerID	First_Name	Test_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Street	City
0	1001	Frodo	Baggins	1235455421	123 Shire Lane, Shire	Yes	N	123 Shire Lane	Shire
4	1005	Jon	Snow	8766783469	123 Dragons Road	Yes	N	123 Dragons Road	None
7	1008	Sherlock	Holmes	8766783469	98 Clue Drive	N	N	98 Clue Drive	None
9	1010	Peter	Parker	1235455421	25th Main Street, New York	Yes	N	25th Main Street	New York
12	1013	Don	Draper	1235432345	2039 Main Street	Yes	N	2039 Main Street	None
13	1014	Leslie	Knope	8766783469	343 City Parkway	Yes	N	343 City Parkway	None
14	1015	Toby	Flenderson	3047622467	214 HR Avenue	N	N	214 HR Avenue	None
15	1016	Ron	Weasley	1235455421	2395 Hogwarts Avenue	N	N	2395 Hogwarts Avenue	None
16	1017	Michael	Scott	1236439775	121 Paper Avenue, Pennsylvania	Yes	N	121 Paper Avenue	Pennsylvania
19	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Yes	N	910 Tatooine Road	Tatooine
20	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Yes	N	910 Tatooine Road	Tatooine

10.Cleaned Data

