

```
In [1]: import pandas as pd
```

```
In [228... call_list = pd.read_excel('/Users/mehmetsaka/Desktop/Customer Call List.xlsx')
```

```
In [230... call_list.head()
```

```
Out[230]:
```

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not Contac
0	1001	James	Milter	237-767-6961	123 Dove Lane, Mordor	Yes	No
1	1002	Tom	Rooney	869/643/9895	150 Mountain Rd	No	Yes
2	1003	Steven	/Brown	9086964392	314 October Street	N	Na
3	1004	Jane	Salt	976-358-1675	1980 Denethor Avenue, Gondor, 18503	Yes	
4	1005	Michael	Rain	413 394 4450	436 November Road	Y	No

```
In [84]: call_list.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 21 entries, 0 to 20
Data columns (total 8 columns):
#   Column                Non-Null Count  Dtype
---  -
0   CustomerID            21 non-null    int64
1   First_Name            21 non-null    object
2   Last_Name             20 non-null    object
3   Phone_Number          19 non-null    object
4   Address               21 non-null    object
5   Paying Customer       21 non-null    object
6   Do_Not_Contact        17 non-null    object
7   Not_Useful_Column     21 non-null    bool
dtypes: bool(1), int64(1), object(6)
memory usage: 1.3+ KB
```

```
In [236... df = call_list
```

```
In [238... df.tail()
```

Out[238]:

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not Contac
16	1017	Didier	Tominay	489/145/1612	121 Paper Avenue, Rivendell	Yes	No
17	1018	Victor	Watercity	3015150122	23 Fence Lane	Y	Na
18	1019	Gareth	Thorn	N/a	N/a	N/a	Ye
19	1020	Leo	Finch	760 318 3109	34 Durin Road, Moria	Yes	f
20	1020	Roman	Eagle	156 651 7016	52 Galadriel Road, Lorien	Yes	f

```
In [240]: df.value_counts("Paying Customer")
```

```
Out[240]: Paying Customer
Yes      11
N         3
No        3
Y         3
N/a       1
dtype: int64
```

```
In [242]: df = df.drop_duplicates()
df
```

Out [242]:

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Conta
0	1001	James	Milter	237-767-6961	123 Dove Lane, Mordor	Yes	I
1	1002	Tom	Rooney	869/643/9895	150 Mountain Rd	No	Y
2	1003	Steven	/Brown	9086964392	314 October Street	N	Na
3	1004	Jane	Salt	976-358-1675	1980 Denethor Avenue, Gondor, 18503	Yes	
4	1005	Michael	Rain	413 394 4450	436 November Road	Y	I
5	1006	Timothy	Weasley	212-532-5368	768 Fenton Parkway	Yes	Y
6	1007	Jeff	Guardian	NaN	1720 North Ave	No	I
7	1008	Juan	White	216 171 9058	42 Industry Drive	N	I
8	1009	Jack	NaN	N/a	487 Mill Street	Yes	Na
9	1010	Kevin	Black	461-452-2853	25th Thranduil Street, Mirkwood, 17101	Yes	I
10	1011	Sam	Hill	NaN	612 Mithrill Lane, Erebor	Yes	I
11	1012	Harry	...Gerard	4377250195	2394 Arts Avenue	Y	Na
12	1013	Clark	Ferguson//	639-852-8725	68 East Street	Yes	
13	1014	Mark	Kite	619 909 1234	2466 Fast Parkway	Yes	I
14	1015	David	Fletcher_	542-532-2127	214 Balboa Lane	N	I
15	1016	Robert	Call	455-932-9176	2445 Garnett Dr	No	

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Conta
16	1017	Didier	Tominay	489/145/1612	121 Paper Avenue, Rivendell	Yes	I
17	1018	Victor	Watercity	3015150122	23 Fence Lane	Y	Ni
18	1019	Gareth	Thorn	N/a	N/a	N/a	Y
19	1020	Leo	Finch	760 318 3109	34 Durin Road, Moria	Yes	
20	1020	Roman	Eagle	156 651 7016	52 Galadriel Road, Lorien	Yes	

```
In [244...] df["Last_Name"] = df["Last_Name"].str.strip("123._/")
```

```
In [246...] df["Phone_Number"] = df["Phone_Number"].str.replace('[^a-zA-Z0-9]', '')
```

```
/var/folders/mb/cj5yd6q12_l7yv3rw7qxrq840000gn/T/ipykernel_22277/2864832742.py:1: FutureWarning: The default value of regex will change from True to False in a future version.
df["Phone_Number"] = df["Phone_Number"].str.replace('[^a-zA-Z0-9]', '')
```

```
In [248...] df["Phone_Number"] = df["Phone_Number"].apply(lambda x: str(x))
```

```
In [250...] df["Phone_Number"] = df["Phone_Number"].apply(lambda x: x[0:3] + '-' + x[3:6] + '-')
```

```
In [126...] df
```

Out[126]:

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Col
0	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Yes	
1	1002	Abed	Nadir	123-643-9775	93 West Main Street	No	
2	1003	Walter	White	706-695-0392	298 Drugs Driveway	N	
3	1004	Dwight	Schrute	123-543-2345	980 Paper Avenue, Pennsylvania, 18503	Yes	
4	1005	Jon	Snow	876-678-3469	123 Dragons Road	Y	
5	1006	Ron	Swanson	304-762-2467	768 City Parkway	Yes	
6	1007	Jeff	Winger	nan--	1209 South Street	No	
7	1008	Sherlock	Holmes	876-678-3469	98 Clue Drive	N	
8	1009	Gandalf	NaN	Na--	123 Middle Earth	Yes	
9	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Yes	
10	1011	Samwise	Gamgee	nan--	612 Shire Lane, Shire	Yes	
11	1012	Harry	Potter	706-695-0392	2394 Hogwarts Avenue	Y	
12	1013	Don	Draper	123-543-2345	2039 Main Street	Yes	
13	1014	Leslie	Knope	876-678-3469	343 City Parkway	Yes	
14	1015	Toby	Flenderson	304-762-2467	214 HR Avenue	N	
15	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	No	
16	1017	Michael	Scott	123-643-9775	121 Paper Avenue, Pennsylvania	Yes	
17	1018	Clark	Kent	706-695-0392	3498 Super Lane	Y	
18	1019	Creed	Braton	Na--	N/a	N/a	
19	1020	Anakin	Skywalker	876-678-3469	910 Tatooine Road, Tatooine	Yes	

```
In [252... # df["Phone_Number"] = df["Phone_Number"].str.replace("Na--" , "")  
df["Phone_Number"] = df["Phone_Number"].str.replace("nan--", "")
```

```
In [254... df
```

Out [254]:

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Conta
0	1001	James	Milter	237-767-6961	123 Dove Lane, Mordor	Yes	I
1	1002	Tom	Rooney	869-643-9895	150 Mountain Rd	No	Y
2	1003	Steven	Brown		314 October Street	N	Na
3	1004	Jane	Salt	976-358-1675	1980 Denethor Avenue, Gondor, 18503	Yes	
4	1005	Michael	Rain	413-394-4450	436 November Road	Y	I
5	1006	Timothy	Weasley	212-532-5368	768 Fenton Parkway	Yes	Y
6	1007	Jeff	Guardian		1720 North Ave	No	I
7	1008	Juan	White	216-171-9058	42 Industry Drive	N	I
8	1009	Jack	NaN	Na--	487 Mill Street	Yes	Na
9	1010	Kevin	Black	461-452-2853	25th Thranduil Street, Mirkwood, 17101	Yes	I
10	1011	Sam	Hill		612 Mithrill Lane, Erebor	Yes	I
11	1012	Harry	...Gerard		2394 Arts Avenue	Y	Na
12	1013	Clark	Ferguson	639-852-8725	68 East Street	Yes	
13	1014	Mark	Kite	619-909-1234	2466 Fast Parkway	Yes	I
14	1015	David	Fletcher	542-532-2127	214 Balboa Lane	N	I
15	1016	Robert	Call	455-932-9176	2445 Garnett Dr	No	

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Conta
16	1017	Didier	Tominay	489-145-1612	121 Paper Avenue, Rivendell	Yes	I
17	1018	Victor	Watercity		23 Fence Lane	Y	Na
18	1019	Gareth	Thorn	Na--	N/a	N/a	Y
19	1020	Leo	Finch	760-318-3109	34 Durin Road, Moria	Yes	
20	1020	Roman	Eagle	156-651-7016	52 Galadriel Road, Lorien	Yes	

In [256... `df[["Street","State","Zip_Code"]] = df["Address"].str.split(", " , 2 , expand =`

```

/var/folders/mb/cj5yd6q12_l7yv3rw7qxrq840000gn/T/ipykernel_22277/4265193529.py:1: FutureWarning: In a future version of pandas all arguments of StringMethods.split except for the argument 'pat' will be keyword-only.
  df[["Street","State","Zip_Code"]] = df["Address"].str.split(", " , 2 , expand = True)

```

In [258... `df.drop(columns = "Address" , inplace = True)`

In [260... `df["Paying Customer"] = df["Paying Customer"].str.replace("No" , "N")`

In [264... `df["Paying Customer"] = df["Paying Customer"].str.replace("Yes" , "Y")`

In [266... `df["Paying Customer"] = df["Paying Customer"].str.replace("N" , "No")`

In [268... `df["Paying Customer"] = df["Paying Customer"].str.replace("Y" , "Yes")`

In [158... `df`

Out[158]:

	CustomerID	First_Name	Last_Name	Phone_Number	Paying Customer	Do_Not_Contact	Street
0	1001	Frodo	Baggins	123-545-5421	Yes	No	123 Shir Lan
1	1002	Abed	Nadir	123-643-9775	No	Yes	93 Wes Mai Stree
2	1003	Walter	White	706-695-0392	No	NaN	29 Drug Drivewa
3	1004	Dwight	Schrute	123-543-2345	Yes	Y	98 Pape Avenu
4	1005	Jon	Snow	876-678-3469	Yes	No	12 Dragon Roa
5	1006	Ron	Swanson	304-762-2467	Yes	Yes	768 Cit Parkwa
6	1007	Jeff	Winger		No	No	120 Sout Stree
7	1008	Sherlock	Holmes	876-678-3469	No	No	98 Clu Driv
8	1009	Gandalf	NaN		Yes	NaN	12 Middl Eart
9	1010	Peter	Parker	123-545-5421	Yes	No	25t Mai Stree
10	1011	Samwise	Gamgee		Yes	No	612 Shir Lan
11	1012	Harry	Potter	706-695-0392	Yes	NaN	239 Hogwart Avenu
12	1013	Don	Draper	123-543-2345	Yes	N	203 Mai Stree
13	1014	Leslie	Knope	876-678-3469	Yes	No	343 Cit Parkwa
14	1015	Toby	Flenderson	304-762-2467	No	No	214 H Avenu
15	1016	Ron	Weasley	123-545-5421	No	N	239 Hogwart Avenu
16	1017	Michael	Scott	123-643-9775	Yes	No	12 Pape Avenu
17	1018	Clark	Kent	706-695-0392	Yes	NaN	349 Supe

	CustomerID	First_Name	Last_Name	Phone_Number	Paying Customer	Do_Not_Contact	Stree
							Lan
18	1019	Creed	Braton		No/a	Yes	N/
19	1020	Anakin	Skywalker	876-678-3469	Yes	N	91 Tatooine Roa

```
In [270...] df["Do_Not_Contact"] = df["Do_Not_Contact"].str.replace("No" , "N")
```

```
In [272...] df["Do_Not_Contact"] = df["Do_Not_Contact"].str.replace("Y","Yes")
```

```
In [274...] df = df.replace("N/a" , "")
```

```
In [276...] df = df.fillna("")
```

```
In [278...] df = df.replace("No/a" , "")
```

```
In [286...] df["Paying Customer"] = df["Paying Customer"].str.replace("Yes" , "Y")
df["Paying Customer"] = df["Paying Customer"].str.replace("No" , "N")
```

```
In [288...] df["Do_Not_Contact"] = df["Do_Not_Contact"].str.replace("Yes" , "Y")
df["Do_Not_Contact"] = df["Do_Not_Contact"].str.replace("No" , "N")
```

```
In [290...] df
```

Out[290]:

	CustomerID	First_Name	Last_Name	Phone_Number	Paying Customer	Do_Not_Contact	Not_Us
0	1001	James	Milter	237-767-6961	Y		N
1	1002	Tom	Rooney	869-643-9895	N		Y
2	1003	Steven	Brown		N		
3	1004	Jane	Salt	976-358-1675	Y		Y
4	1005	Michael	Rain	413-394-4450	Y		N
5	1006	Timothy	Weasley	212-532-5368	Y		Y
6	1007	Jeff	Guardian		N		N
7	1008	Juan	White	216-171-9058	N		N
8	1009	Jack		Na--	Y		
9	1010	Kevin	Black	461-452-2853	Y		N
10	1011	Sam	Hill		Y		N
11	1012	Harry	...Gerard		Y		
12	1013	Clark	Ferguson	639-852-8725	Y		N
13	1014	Mark	Kite	619-909-1234	Y		N
14	1015	David	Fletcher	542-532-2127	N		N
15	1016	Robert	Call	455-932-9176	N		N
16	1017	Didier	Tominay	489-145-1612	Y		N
17	1018	Victor	Watercity		Y		

	CustomerID	First_Name	Last_Name	Phone_Number	Paying Customer	Do_Not_Contact	Not_Us
18	1019	Gareth	Thorn	Na--		Y	
19	1020	Leo	Finch	760-318-3109	Y	N	
20	1020	Roman	Eagle	156-651-7016	Y	N	

```
In [292... for x in df.index:
            if df.loc[x, "Do_Not_Contact"] == 'Y':
                df.drop(x, inplace = True)
```

```
In [218... df
```

Out[218]:

	CustomerID	First_Name	Last_Name	Phone_Number	Paying Customer	Do_Not_Contact	Street
0	1001	Frodo	Baggins	123-545-5421	Y	N	123 Shir Lan
2	1003	Walter	White	706-695-0392	N		29 Drug Drivewa
4	1005	Jon	Snow	876-678-3469	Y	N	12 Dragon Roa
6	1007	Jeff	Winger		N	N	120 Sout Stree
7	1008	Sherlock	Holmes	876-678-3469	N	N	98 Clu Driv
8	1009	Gandalf			Y		12 Middl Eart
9	1010	Peter	Parker	123-545-5421	Y	N	25t Mai Stree
10	1011	Samwise	Gamgee		Y	N	612 Shir Lan
11	1012	Harry	Potter	706-695-0392	Y		239 Hogwart Avenu
12	1013	Don	Draper	123-543-2345	Y	N	203 Mai Stree
13	1014	Leslie	Knope	876-678-3469	Y	N	343 Cit Parkwa
14	1015	Toby	Flenderson	304-762-2467	N	N	214 H Avenu
15	1016	Ron	Weasley	123-545-5421	N	N	239 Hogwart Avenu
16	1017	Michael	Scott	123-643-9775	Y	N	12 Pape Avenu
17	1018	Clark	Kent	706-695-0392	Y		349 Supe Lan
19	1020	Anakin	Skywalker	876-678-3469	Y	N	91 Tatooi Roa

```
In [306... for x in df.index:
            if df.loc[x, "Phone_Number"] == '':
                df.drop(x, inplace = True)
```

```
In [222... df
```

```
Out[222]:
```

	CustomerID	First_Name	Last_Name	Phone_Number	Paying Customer	Do_Not_Contact	Stree
0	1001	Frodo	Baggins	123-545-5421	Y	N	123 Shir Lan
2	1003	Walter	White	706-695-0392	N		29 Drug Drivewa
4	1005	Jon	Snow	876-678-3469	Y	N	12 Dragon Roa
7	1008	Sherlock	Holmes	876-678-3469	N	N	98 Clu Driv
9	1010	Peter	Parker	123-545-5421	Y	N	25t Mai Stree
11	1012	Harry	Potter	706-695-0392	Y		239 Hogwart Avenu
12	1013	Don	Draper	123-543-2345	Y	N	203 Mai Stree
13	1014	Leslie	Knope	876-678-3469	Y	N	343 Cit Parkwa
14	1015	Toby	Flenderson	304-762-2467	N	N	214 H Avenu
15	1016	Ron	Weasley	123-545-5421	N	N	239 Hogwart Avenu
16	1017	Michael	Scott	123-643-9775	Y	N	12 Pape Avenu
17	1018	Clark	Kent	706-695-0392	Y		349 Supe Lan
19	1020	Anakin	Skywalker	876-678-3469	Y	N	91 Tatooi Roa

```
In [296... df.reset_index(drop = True)
```

Out[296]:

	CustomerID	First_Name	Last_Name	Phone_Number	Paying Customer	Do_Not_Contact	Not_Us
0	1001	James	Milter	237-767-6961	Y		N
1	1005	Michael	Rain	413-394-4450	Y		N
2	1008	Juan	White	216-171-9058	N		N
3	1009	Jack		Na--	Y		
4	1010	Kevin	Black	461-452-2853	Y		N
5	1013	Clark	Ferguson	639-852-8725	Y		N
6	1014	Mark	Kite	619-909-1234	Y		N
7	1015	David	Fletcher	542-532-2127	N		N
8	1016	Robert	Call	455-932-9176	N		N
9	1017	Didier	Tominay	489-145-1612	Y		N
10	1020	Leo	Finch	760-318-3109	Y		N
11	1020	Roman	Eagle	156-651-7016	Y		N

In [302...

```
df['Phone_Number'] = df['Phone_Number'].str.replace('Na--', '')
```

In [308...

```
df
```

Out[308]:

	CustomerID	First_Name	Last_Name	Phone_Number	Paying Customer	Do_Not_Contact	Not_Us
0	1001	James	Milter	237-767-6961	Y		N
4	1005	Michael	Rain	413-394-4450	Y		N
7	1008	Juan	White	216-171-9058	N		N
9	1010	Kevin	Black	461-452-2853	Y		N
12	1013	Clark	Ferguson	639-852-8725	Y		N
13	1014	Mark	Kite	619-909-1234	Y		N
14	1015	David	Fletcher	542-532-2127	N		N
15	1016	Robert	Call	455-932-9176	N		N
16	1017	Didier	Tominay	489-145-1612	Y		N
19	1020	Leo	Finch	760-318-3109	Y		N
20	1020	Roman	Eagle	156-651-7016	Y		N

In []: