Efficient Data Cleaning for Customer Sales Analysis using Pandas

Importing the library

In [288…

```python
import pandas as pd
```

Reading the dataset

In [289…

```python
df = pd.read_excel("C:\\Users\\HP\\Desktop\\EDA\\mosh\\Customer Call List.xlsx")
```

In [290…

```python
df
```

Out[290…

	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

To drop the duplicates in the entire column

In [291…

```python
df=df.drop_duplicates()
df
```

Out[291…

	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

To remove the unwanted columns

In [292…

```python
df=df.drop(columns = "Not_Useful_Column")
df
```

Out[292…

	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

Using 'Strip' command to remove the carbages in a particular column

In [293…

```python
#df["Last_Name"]=df["Last_Name"].str.lstrip()       # left strip is used here
#df["Last_Name"]=df["Last_Name"].str.lstrip("...")  # left strip is used here
#df["Last_Name"]=df["Last_Name"].str.lstrip("/")    # left strip is used here
#df["Last_Name"]=df["Last_Name"].str.rstrip("_")    # right strip is used here
df["Last_Name"]=df["Last_Name"].str.strip("123._/")  # Removing all carbages in one commnad using 'strip' command
df
```

Out[293…

	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

In [294…

```python
df
```

Out[294…

	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

Cleaning the phone number

In [295…

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

#df["Phone_Number"]=df["Phone_Number"].apply(lambda x: x[0:3]+ '-' + x[3:6] + '-' + x[6:10])

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

df["Phone_Number"] = df["Phone_Number"].apply(lambda x: x[0:3]+ '-' + x[3:6] + '-' + x[6:10])

df["Phone_Number"]=df["Phone_Number"].str.replace('nan--','')

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

df
```

```
C:\Users\HP\AppData\Local\Temp/ipykernel_9892/4225034453.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]','')
```

Out[295…

	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		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		1209 South Street	No	No
7	1008	Sherlock	Holmes	876-678-3469	98 Clue Drive	N	No
8	1009	Gandalf	NaN		123 Middle Earth	Yes	NaN
9	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Yes	No
10	1011	Samwise	Gamgee		612 Shire Lane, Shire	Yes	No
11	1012	Harry	Potter		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		3498 Super Lane	Y	NaN
18	1019	Creed	Braton		N/a	N/a	Yes
19	1020	Anakin	Skywalker	876-678-3469	910 Tatooine Road, Tatooine	Yes	N

To expand the address column

In [296…

```python
#df["Address"].str.split(',',1)              # didnt make any change
#df["Address"].str.split(',',1,expand=True) # for creating one column
df["Address"].str.split(',',2,expand=True)  # for creating two columns
```

Out[296…

	0	1	2
0	123 Shire Lane	Shire	None
1	93 West Main Street	None	None
2	298 Drugs Driveway	None	None
3	980 Paper Avenue	Pennsylvania	18503
4	123 Dragons Road	None	None
5	768 City Parkway	None	None
6	1209 South Street	None	None
7	98 Clue Drive	None	None
8	123 Middle Earth	None	None
9	25th Main Street	New York	None
10	612 Shire Lane	Shire	None
11	2394 Hogwarts Avenue	None	None
12	2039 Main Street	None	None
13	343 City Parkway	None	None
14	214 HR Avenue	None	None
15	2395 Hogwarts Avenue	None	None
16	121 Paper Avenue	Pennsylvania	None
17	3498 Super Lane	None	None
18	N/a	None	None
19	910 Tatooine Road	Tatooine	None

Assigning a column name for the expanded address

In [297…

```python
df[["Street_Address","State","Zip_Code"]] = df["Address"].str.split(',',2,expand=True)
df
```

Out[297…

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Street_Address	State	Zip_Code
0	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Yes	No	123 Shire Lane	Shire	None
1	1002	Abed	Nadir	123-643-9775	93 West Main Street	No	Yes	93 West Main Street	None	None
2	1003	Walter	White		298 Drugs Driveway	N	NaN	298 Drugs Driveway	None	None
3	1004	Dwight	Schrute	123-543-2345	980 Paper Avenue, Pennsylvania, 18503	Yes	Y	980 Paper Avenue	Pennsylvania	18503
4	1005	Jon	Snow	876-678-3469	123 Dragons Road	Y	No	123 Dragons Road	None	None
5	1006	Ron	Swanson	304-762-2467	768 City Parkway	Yes	Yes	768 City Parkway	None	None
6	1007	Jeff	Winger		1209 South Street	No	No	1209 South Street	None	None
7	1008	Sherlock	Holmes	876-678-3469	98 Clue Drive	N	No	98 Clue Drive	None	None
8	1009	Gandalf	NaN		123 Middle Earth	Yes	NaN	123 Middle Earth	None	None
9	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Yes	No	25th Main Street	New York	None
10	1011	Samwise	Gamgee		612 Shire Lane, Shire	Yes	No	612 Shire Lane	Shire	None
11	1012	Harry	Potter		2394 Hogwarts Avenue	Y	NaN	2394 Hogwarts Avenue	None	None
12	1013	Don	Draper	123-543-2345	2039 Main Street	Yes	N	2039 Main Street	None	None
13	1014	Leslie	Knope	876-678-3469	343 City Parkway	Yes	No	343 City Parkway	None	None
14	1015	Toby	Flenderson	304-762-2467	214 HR Avenue	N	No	214 HR Avenue	None	None
15	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	No	N	2395 Hogwarts Avenue	None	None
16	1017	Michael	Scott	123-643-9775	121 Paper Avenue, Pennsylvania	Yes	No	121 Paper Avenue	Pennsylvania	None
17	1018	Clark	Kent		3498 Super Lane	Y	NaN	3498 Super Lane	None	None
18	1019	Creed	Braton		N/a	N/a	Yes	N/a	None	None
19	1020	Anakin	Skywalker	876-678-3469	910 Tatooine Road, Tatooine	Yes	N	910 Tatooine Road	Tatooine	None

working on paying customer

In [298…

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


df["Paying Customer"] = df["Paying Customer"].str.replace('No','N')

df
```

Out[298…

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Street_Address	State	Zip_Code
0	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Y	No	123 Shire Lane	Shire	None
1	1002	Abed	Nadir	123-643-9775	93 West Main Street	N	Yes	93 West Main Street	None	None
2	1003	Walter	White		298 Drugs Driveway	N	NaN	298 Drugs Driveway	None	None
3	1004	Dwight	Schrute	123-543-2345	980 Paper Avenue, Pennsylvania, 18503	Y	Y	980 Paper Avenue	Pennsylvania	18503
4	1005	Jon	Snow	876-678-3469	123 Dragons Road	Y	No	123 Dragons Road	None	None
5	1006	Ron	Swanson	304-762-2467	768 City Parkway	Y	Yes	768 City Parkway	None	None
6	1007	Jeff	Winger		1209 South Street	N	No	1209 South Street	None	None
7	1008	Sherlock	Holmes	876-678-3469	98 Clue Drive	N	No	98 Clue Drive	None	None
8	1009	Gandalf	NaN		123 Middle Earth	Y	NaN	123 Middle Earth	None	None
9	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Y	No	25th Main Street	New York	None
10	1011	Samwise	Gamgee		612 Shire Lane, Shire	Y	No	612 Shire Lane	Shire	None
11	1012	Harry	Potter		2394 Hogwarts Avenue	Y	NaN	2394 Hogwarts Avenue	None	None
12	1013	Don	Draper	123-543-2345	2039 Main Street	Y	N	2039 Main Street	None	None
13	1014	Leslie	Knope	876-678-3469	343 City Parkway	Y	No	343 City Parkway	None	None
14	1015	Toby	Flenderson	304-762-2467	214 HR Avenue	N	No	214 HR Avenue	None	None
15	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	N	N	2395 Hogwarts Avenue	None	None
16	1017	Michael	Scott	123-643-9775	121 Paper Avenue, Pennsylvania	Y	No	121 Paper Avenue	Pennsylvania	None
17	1018	Clark	Kent		3498 Super Lane	Y	NaN	3498 Super Lane	None	None
18	1019	Creed	Braton		N/a	N/a	Yes	N/a	None	None
19	1020	Anakin	Skywalker	876-678-3469	910 Tatooine Road, Tatooine	Y	N	910 Tatooine Road	Tatooine	None

Working on do_not_contact

In [299…

```python
df["Do_Not_Contact"] = df["Do_Not_Contact"].str.replace('Yes','Y')


df["Do_Not_Contact"] = df["Do_Not_Contact"].str.replace('No','N')

df
```

Out[299…

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Street_Address	State	Zip_Code
0	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Y	N	123 Shire Lane	Shire	None
1	1002	Abed	Nadir	123-643-9775	93 West Main Street	N	Y	93 West Main Street	None	None
2	1003	Walter	White		298 Drugs Driveway	N	NaN	298 Drugs Driveway	None	None
3	1004	Dwight	Schrute	123-543-2345	980 Paper Avenue, Pennsylvania, 18503	Y	Y	980 Paper Avenue	Pennsylvania	18503
4	1005	Jon	Snow	876-678-3469	123 Dragons Road	Y	N	123 Dragons Road	None	None
5	1006	Ron	Swanson	304-762-2467	768 City Parkway	Y	Y	768 City Parkway	None	None
6	1007	Jeff	Winger		1209 South Street	N	N	1209 South Street	None	None
7	1008	Sherlock	Holmes	876-678-3469	98 Clue Drive	N	N	98 Clue Drive	None	None
8	1009	Gandalf	NaN		123 Middle Earth	Y	NaN	123 Middle Earth	None	None
9	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Y	N	25th Main Street	New York	None
10	1011	Samwise	Gamgee		612 Shire Lane, Shire	Y	N	612 Shire Lane	Shire	None
11	1012	Harry	Potter		2394 Hogwarts Avenue	Y	NaN	2394 Hogwarts Avenue	None	None
12	1013	Don	Draper	123-543-2345	2039 Main Street	Y	N	2039 Main Street	None	None
13	1014	Leslie	Knope	876-678-3469	343 City Parkway	Y	N	343 City Parkway	None	None
14	1015	Toby	Flenderson	304-762-2467	214 HR Avenue	N	N	214 HR Avenue	None	None
15	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	N	N	2395 Hogwarts Avenue	None	None
16	1017	Michael	Scott	123-643-9775	121 Paper Avenue, Pennsylvania	Y	N	121 Paper Avenue	Pennsylvania	None
17	1018	Clark	Kent		3498 Super Lane	Y	NaN	3498 Super Lane	None	None
18	1019	Creed	Braton		N/a	N/a	Y	N/a	None	None
19	1020	Anakin	Skywalker	876-678-3469	910 Tatooine Road, Tatooine	Y	N	910 Tatooine Road	Tatooine	None

To clear the 'N/a' and 'NaN' from all the columns

df.fillna('') # removes 'NaN' and 'None' in all coulumns

Looping in "Do_Not_Contact" column

In [300…

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

Out[300…

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Street_Address	State	Zip_Code
0	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Y	N	123 Shire Lane	Shire	None
2	1003	Walter	White		298 Drugs Driveway	N	NaN	298 Drugs Driveway	None	None
4	1005	Jon	Snow	876-678-3469	123 Dragons Road	Y	N	123 Dragons Road	None	None
6	1007	Jeff	Winger		1209 South Street	N	N	1209 South Street	None	None
7	1008	Sherlock	Holmes	876-678-3469	98 Clue Drive	N	N	98 Clue Drive	None	None
8	1009	Gandalf	NaN		123 Middle Earth	Y	NaN	123 Middle Earth	None	None
9	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Y	N	25th Main Street	New York	None
10	1011	Samwise	Gamgee		612 Shire Lane, Shire	Y	N	612 Shire Lane	Shire	None
11	1012	Harry	Potter		2394 Hogwarts Avenue	Y	NaN	2394 Hogwarts Avenue	None	None
12	1013	Don	Draper	123-543-2345	2039 Main Street	Y	N	2039 Main Street	None	None
13	1014	Leslie	Knope	876-678-3469	343 City Parkway	Y	N	343 City Parkway	None	None
14	1015	Toby	Flenderson	304-762-2467	214 HR Avenue	N	N	214 HR Avenue	None	None
15	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	N	N	2395 Hogwarts Avenue	None	None
16	1017	Michael	Scott	123-643-9775	121 Paper Avenue, Pennsylvania	Y	N	121 Paper Avenue	Pennsylvania	None
17	1018	Clark	Kent		3498 Super Lane	Y	NaN	3498 Super Lane	None	None
19	1020	Anakin	Skywalker	876-678-3469	910 Tatooine Road, Tatooine	Y	N	910 Tatooine Road	Tatooine	None

To remove the empty space in "phone_Number" column

In [301…

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

Out[301…

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Street_Address	State	Zip_Code
0	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Y	N	123 Shire Lane	Shire	None
4	1005	Jon	Snow	876-678-3469	123 Dragons Road	Y	N	123 Dragons Road	None	None
7	1008	Sherlock	Holmes	876-678-3469	98 Clue Drive	N	N	98 Clue Drive	None	None
9	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Y	N	25th Main Street	New York	None
12	1013	Don	Draper	123-543-2345	2039 Main Street	Y	N	2039 Main Street	None	None
13	1014	Leslie	Knope	876-678-3469	343 City Parkway	Y	N	343 City Parkway	None	None
14	1015	Toby	Flenderson	304-762-2467	214 HR Avenue	N	N	214 HR Avenue	None	None
15	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	N	N	2395 Hogwarts Avenue	None	None
16	1017	Michael	Scott	123-643-9775	121 Paper Avenue, Pennsylvania	Y	N	121 Paper Avenue	Pennsylvania	None
19	1020	Anakin	Skywalker	876-678-3469	910 Tatooine Road, Tatooine	Y	N	910 Tatooine Road	Tatooine	None

To reset the index

In [302…

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

Out[302…

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Street_Address	State	Zip_Code
0	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Y	N	123 Shire Lane	Shire	None
1	1005	Jon	Snow	876-678-3469	123 Dragons Road	Y	N	123 Dragons Road	None	None
2	1008	Sherlock	Holmes	876-678-3469	98 Clue Drive	N	N	98 Clue Drive	None	None
3	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Y	N	25th Main Street	New York	None
4	1013	Don	Draper	123-543-2345	2039 Main Street	Y	N	2039 Main Street	None	None
5	1014	Leslie	Knope	876-678-3469	343 City Parkway	Y	N	343 City Parkway	None	None
6	1015	Toby	Flenderson	304-762-2467	214 HR Avenue	N	N	214 HR Avenue	None	None
7	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	N	N	2395 Hogwarts Avenue	None	None
8	1017	Michael	Scott	123-643-9775	121 Paper Avenue, Pennsylvania	Y	N	121 Paper Avenue	Pennsylvania	None
9	1020	Anakin	Skywalker	876-678-3469	910 Tatooine Road, Tatooine	Y	N	910 Tatooine Road	Tatooine	None

To remove "None" in all the column

In [303…

```python
df.fillna('')
```

Out[303…

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Street_Address	State	Zip_Code
0	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Y	N	123 Shire Lane	Shire	
1	1005	Jon	Snow	876-678-3469	123 Dragons Road	Y	N	123 Dragons Road		
2	1008	Sherlock	Holmes	876-678-3469	98 Clue Drive	N	N	98 Clue Drive		
3	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Y	N	25th Main Street	New York	
4	1013	Don	Draper	123-543-2345	2039 Main Street	Y	N	2039 Main Street		
5	1014	Leslie	Knope	876-678-3469	343 City Parkway	Y	N	343 City Parkway		
6	1015	Toby	Flenderson	304-762-2467	214 HR Avenue	N	N	214 HR Avenue		
7	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	N	N	2395 Hogwarts Avenue		
8	1017	Michael	Scott	123-643-9775	121 Paper Avenue, Pennsylvania	Y	N	121 Paper Avenue	Pennsylvania	
9	1020	Anakin	Skywalker	876-678-3469	910 Tatooine Road, Tatooine	Y	N	910 Tatooine Road	Tatooine	

Summary:

This project focuses on efficiently cleaning a customer sales dataset using the pandas library.

Various functions from the pandas library are utilized to effectively clean the data, ensuring it is suitable for analysis purposes.