

```
# Import pandas as pd
df = pd.read_excel(r"C:\Users\Ttram Nguyen\Documents\00_ Project\Python - cleaning_data\Customer Call List.xlsx")
df

In [5]: df.drop_duplicates()
df

Out[5]: 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 Piper Avenue, Pennsylvania, 18503 Yes Y True
4 1005 Jon Snow 8766783469 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 8766783469 98 Clue Drive N No False
8 1009 Gandalf NaN Na 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 8766783469 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 Hogswarts 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 Bratton Na NaN Yes True
19 1020 Anakin Skywalker 8766783469 910 Tatooine Road, Tatooine Yes N True

In [5]: df = df.drop_duplicates()
df

Out[5]: 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 Piper Avenue, Pennsylvania, 18503 Yes Y True
4 1005 Jon Snow 8766783469 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 8766783469 98 Clue Drive N No False
8 1009 Gandalf NaN Na 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 Hogswarts Avenue Y NaN True
12 1013 Don Draper 123-543-2345 2039 Main Street Yes N False
13 1014 Leslie Knope 8766783469 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 Hogswarts 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 Bratton Na NaN Yes True
19 1020 Anakin Skywalker 8766783469 910 Tatooine Road, Tatooine Yes N True

In [5]: df["Last_Name"] = df["Last_Name"].str.strip("123./'")
df

Out[5]: 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 Piper Avenue, Pennsylvania, 18503 Yes Y
4 1005 Jon Snow 8766783469 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 8766783469 98 Clue Drive N No
8 1009 Gandalf NaN Na 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 Hogswarts Avenue Y NaN
12 1013 Don Draper 123-543-2345 2039 Main Street Yes N
13 1014 Leslie Knope 8766783469 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 Hogswarts 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 Bratton Na NaN Yes
19 1020 Anakin Skywalker 8766783469 910 Tatooine Road, Tatooine Yes N

In [25]: df["Last_Name"] = df["Last_Name"].str.strip("123./'")
df

Out[25]: 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 Piper Avenue, Pennsylvania, 18503 Yes Y
4 1005 Jon Snow 8766783469 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 8766783469 98 Clue Drive N No
8 1009 Gandalf NaN Na 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 Hogswarts Avenue Y NaN
12 1013 Don Draper 123-543-2345 2039 Main Street Yes N
13 1014 Leslie Knope 8766783469 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 Hogswarts 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 Bratton Na NaN Yes
19 1020 Anakin Skywalker 8766783469 910 Tatooine Road, Tatooine Yes N

In [35]: df["Phone_Number"] = df["Phone_Number"].str.replace("[^a-zA-Z0-9]", "", regex=True)
df

Out[35]: CustomerID First_Name Last_Name Phone_Number Address Paying_Customer Do_Not_Contact Cleaned_Phone_Number
0 1001 Frodo Baggins 123-545-5421 123 Shire Lane, Shire Yes No 1235455421
1 1002 Abed Nadir 123-643-9775 93 West Main Street No Yes 1236439775
2 1003 Walter White nan 298 Drugs Driveway N NaN NaN
3 1004 Dwight Schrute 123-543-2345 980 Piper Avenue, Pennsylvania, 18503 Yes Y 1235432345
4 1005 Jon Snow 876-678-3469 123 Dragons Road Y No 8766783469
5 1006 Ron Swanson 304-762-2467 768 City Parkway Yes Yes 3047622467
6 1007 Jeff Winger nan 1209 South Street No No NaN
7 1008 Sherlock Holmes 876-678-3469 98 Clue Drive N No 8766783469
8 1009 Gandalf NaN Na 123 Middle Earth Yes NaN Na
9 1010 Peter Parker 123-545-5421 25th Main Street, New York Yes No 1235455421
10 1011 Samwise Gamgee nan 612 Shire Lane, Shire Yes No NaN
11 1012 Harry Potter nan 2394 Hogswarts Avenue Y NaN NaN
12 1013 Don Draper 123-543-2345 2039 Main Street Yes N 1235432345
13 1014 Leslie Knope 8766783469 343 City Parkway Yes No 8766783469
14 1015 Toby Flenderson_ 304-762-2467 214 HR Avenue N No 3047622467
15 1016 Ron Weasley 123-545-5421 2395 Hogswarts Avenue No N 1235455421
16 1017 Michael Scott 123-643-9775 121 Paper Avenue, Pennsylvania Yes No 1236439775
17 1018 Clark Kent nan- 3498 Super Lane Y NaN NaN
18 1019 Creed Bratton Na- NaN Yes Na
19 1020 Anakin Skywalker 876-678-3469 910 Tatooine Road, Tatooine Yes N 8766783469

In [40]: df["Phone_Number"] = df["Phone_Number"].apply(lambda x: str(x))
df

Out[40]: CustomerID First_Name Last_Name Phone_Number Address Paying_Customer Do_Not_Contact Cleaned_Phone_Number
0 1001 Frodo Baggins 123-545-5421 123 Shire Lane, Shire Yes No 1235455421
1 1002 Abed Nadir 123-643-9775 93 West Main Street No Yes 1236439775
2 1003 Walter White nan 298 Drugs Driveway N NaN NaN
3 1004 Dwight Schrute 123-543-2345 980 Piper Avenue, Pennsylvania, 18503 Yes Y 1235432345
4 1005 Jon Snow 876-678-3469 123 Dragons Road Y No 8766783469
5 1006 Ron Swanson 304-762-2467 768 City Parkway Yes Yes 3047622467
6 1007 Jeff Winger nan- 1209 South Street No No NaN
7 1008 Sherlock Holmes 876-678-3469 98 Clue Drive N No 8766783469
8 1009 Gandalf NaN Na 123 Middle Earth Yes NaN Na
9 1010 Peter Parker 123-545-5421 25th Main Street, New York Yes No 1235455421
10 1011 Samwise Gamgee nan 612 Shire Lane, Shire Yes No NaN
11 1012 Harry Potter nan 2394 Hogswarts Avenue Y NaN NaN
12 1013 Don Draper 123-543-2345 2039 Main Street Yes N 1235432345
13 1014 Leslie Knope 8766783469 343 City Parkway Yes No 8766783469
14 1015 Toby Flenderson_ 304-762-2467 214 HR Avenue N No 3047622467
15 1016 Ron Weasley 123-545-5
```