Efficient Data Cleaning for Customer Sales Analysis using Pandas Importing the library In [288.. import pandas as pd Reading the dataset df = pd.read_excel("C:\\Users\\HP\\Desktop\\EDA\\mosh\\Customer Call List.xlsx") In [290.. CustomerID First Name Address Paying Customer Do_Not_Contact Not_Useful_Column Out [290.. Last_Name Phone_Number 0 1001 123-545-5421 123 Shire Lane, Shire True Frodo **Baggins** Yes No 93 West Main Street 1 1002 Abed 123/643/9775 No Yes False Nadir 2 1003 Walter /White 7066950392 Ν NaN 298 Drugs Driveway True 3 1004 Dwight Schrute 123-543-2345 980 Paper Avenue, Pennsylvania, 18503 Yes True 4 Υ 1005 876|678|3469 123 Dragons Road No True Jon Snow 5 1006 304-762-2467 768 City Parkway Yes Yes Ron Swanson True 6 No 1007 1209 South Street False Jeff NaN No Winger 98 Clue Drive 1008 Sherlock 876|678|3469 Ν No False Holmes 8 1009 123 Middle Earth NaN False Gandalf NaN N/a Yes 25th Main Street, New York 9 1010 Peter Parker 123-545-5421 Yes No True 10 1011 NaN 612 Shire Lane, Shire No Samwise Yes True Gamgee 7066950392 11 1012 ...Potter 2394 Hogwarts Avenue Υ NaN True Harry 12 1013 123-543-2345 Ν Don Draper 2039 Main Street Yes False 13 1014 Leslie 876|678|3469 343 City Parkway Yes No False Knope No 14 1015 Flenderson_ 304-762-2467 214 HR Avenue Ν False Toby 15 1016 123-545-5421 2395 Hogwarts Avenue No False Ron Weasley 121 Paper Avenue, Pennsylvania 1017 Michael 123/643/9775 No False 16 Scott Yes 17 1018 Clark 7066950392 3498 Super Lane Υ NaN True Kent 18 1019 N/a N/a Yes Creed Braton N/a True 19 1020 Anakin Skywalker 876|678|3469 910 Tatooine Road, Tatooine Yes Ν True 20 1020 876|678|3469 Ν Anakin Skywalker 910 Tatooine Road, Tatooine Yes True To drop the duplicates in the entire column In [291. df=df.drop_duplicates() df Out [291. CustomerID First_Name Last_Name Phone_Number Address Paying Customer Do_Not_Contact Not_Useful_Column 0 123-545-5421 123 Shire Lane, Shire No 1001 Frodo **Baggins** Yes True 1 1002 Abed Nadir 123/643/9775 93 West Main Street No Yes False 2 7066950392 NaN 1003 Walter /White 298 Drugs Driveway Ν True 3 123-543-2345 980 Paper Avenue, Pennsylvania, 18503 Υ 1004 Dwight Schrute Yes True 123 Dragons Road Υ No 1005 Jon Snow 876|678|3469 True 5 1006 304-762-2467 Ron Swanson 768 City Parkway Yes Yes True 1007 Jeff Winger NaN 1209 South Street No False 1008 Sherlock 876|678|3469 98 Clue Drive Ν No False Holmes 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 612 Shire Lane, Shire 10 1011 Samwise Gamgee NaN Yes No True NaN 11 1012 ...Potter 7066950392 2394 Hogwarts Avenue Υ True Harry 12 123-543-2345 Ν 1013 Don Draper 2039 Main Street Yes False No 13 1014 Leslie Knope 876|678|3469 343 City Parkway Yes False 304-762-2467 214 HR Avenue 14 1015 Toby Flenderson_ Ν No False 15 1016 Weasley 123-545-5421 2395 Hogwarts Avenue No Ν False Ron 123/643/9775 No 16 1017 Michael Scott 121 Paper Avenue, Pennsylvania Yes False 17 1018 Clark 7066950392 3498 Super Lane Υ NaN True Kent 18 1019 Creed Braton N/a Yes True 19 1020 Anakin Skywalker 876|678|3469 910 Tatooine Road, Tatooine Yes Ν True To remove the unwanted columns In [292.. df=df.drop(columns = "Not_Useful_Column") CustomerID First_Name Last_Name Phone_Number Out [292.. Address Paying Customer Do_Not_Contact 0 1001 Frodo 123-545-5421 123 Shire Lane, Shire Yes No Baggins 1 123/643/9775 1002 Abed Nadir 93 West Main Street No Yes 2 1003 Walter /White 7066950392 298 Drugs Driveway Ν NaN 3 123-543-2345 980 Paper Avenue, Pennsylvania, 18503 Υ 1004 Dwight Schrute Yes 4 1005 Jon Snow 876|678|3469 123 Dragons Road Υ No 5 1006 Ron Swanson 304-762-2467 768 City Parkway Yes Yes 6 1007 Jeff NaN 1209 South Street No No Winger 7 1008 Sherlock Holmes 876|678|3469 98 Clue Drive Ν 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 NaN 612 Shire Lane, Shire Yes No Samwise Gamgee 11 7066950392 Υ 1012 Harry ...Potter 2394 Hogwarts Avenue NaN 12 1013 123-543-2345 2039 Main Street Yes Ν Don Draper 13 1014 Leslie Knope 876|678|3469 343 City Parkway Yes No 14 1015 Toby Flenderson_ 304-762-2467 214 HR Avenue Ν No 15 1016 Ron Weasley 123-545-5421 2395 Hogwarts Avenue No Ν 16 1017 Michael 123/643/9775 121 Paper Avenue, Pennsylvania Yes No Scott 17 Υ 1018 Clark Kent 7066950392 3498 Super Lane NaN 18 1019 Creed Braton N/a N/a Yes N/a 19 1020 Anakin Skywalker 876|678|3469 910 Tatooine Road, Tatooine Yes Ν Using 'Strip' command to remove the carbages in a particular column In [293... #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 Out[293.. CustomerID First_Name Last_Name Phone_Number Address Paying Customer Do_Not_Contact 0 1001 123-545-5421 123 Shire Lane, Shire No Frodo Baggins Yes 1002 123/643/9775 93 West Main Street 1 Abed Nadir No Yes 2 1003 Walter White 7066950392 298 Drugs Driveway Ν NaN 3 980 Paper Avenue, Pennsylvania, 18503 Υ 1004 123-543-2345 Dwight Schrute Yes 4 Snow 876|678|3469 Υ 1005 Jon 123 Dragons Road No 5 1006 Ron Swanson 304-762-2467 768 City Parkway Yes Yes 6 1007 Jeff Winger NaN 1209 South Street No No 7 876|678|3469 Ν 1008 Sherlock Holmes 98 Clue Drive 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 NaN 612 Shire Lane, Shire No Samwise Gamgee Yes 11 1012 7066950392 2394 Hogwarts Avenue NaN Harry Potter 12 Draper 123-543-2345 2039 Main Street 1013 Don Yes Ν 1014 13 Leslie Knope 876|678|3469 343 City Parkway Yes No 304-762-2467 14 1015 Toby Flenderson 214 HR Avenue Ν No 123-545-5421 Ν 15 1016 Ron Weasley 2395 Hogwarts Avenue No 16 1017 Michael Scott 123/643/9775 121 Paper Avenue, Pennsylvania Yes No 1018 7066950392 Υ NaN 17 Clark Kent 3498 Super Lane 18 1019 Creed Braton N/a N/a N/a Yes Ν 19 1020 Anakin Skywalker 876|678|3469 910 Tatooine Road, Tatooine Yes In [294.. CustomerID First_Name Last_Name Phone_Number Address Paying Customer Do_Not_Contact Out[294.. 0 1001 123-545-5421 123 Shire Lane, Shire Frodo **Baggins** Yes No 1 1002 Abed Nadir 123/643/9775 93 West Main Street No Yes 2 Walter 7066950392 NaN 1003 White 298 Drugs Driveway Ν 3 Υ 1004 Dwight Schrute 123-543-2345 980 Paper Avenue, Pennsylvania, 18503 Yes 4 1005 Υ Jon Snow 876|678|3469 123 Dragons Road 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 Ν No 8 1009 N/a 123 Middle Earth NaN Gandalf NaN Yes 123-545-5421 9 1010 Peter Parker 25th Main Street, New York No Yes 10 No 1011 Samwise Gamgee NaN 612 Shire Lane, Shire Yes 11 1012 Harry Potter 7066950392 2394 Hogwarts Avenue Υ NaN 12 1013 123-543-2345 Ν Don Draper 2039 Main Street Yes 13 1014 Leslie Knope 876|678|3469 343 City Parkway Yes No 14 1015 Toby Flenderson 304-762-2467 214 HR Avenue Ν No Ν 15 1016 Ron Weasley 123-545-5421 2395 Hogwarts Avenue No 16 1017 Michael Scott 123/643/9775 121 Paper Avenue, Pennsylvania Yes No 17 1018 Clark 7066950392 Υ NaN Kent 3498 Super Lane 1019 18 Creed Braton N/a N/a Yes Ν 19 1020 Anakin Skywalker 876|678|3469 910 Tatooine Road, Tatooine Yes Cleaning the phone number In [295.. 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]','') CustomerID First_Name Last_Name Phone_Number Address Paying Customer Do_Not_Contact Out [295.. 0 1001 123-545-5421 123 Shire Lane, Shire No Frodo Baggins Yes 1 1002 Abed 123-643-9775 93 West Main Street No Yes Nadir 2 Ν 1003 Walter White NaN 298 Drugs Driveway 3 1004 123-543-2345 980 Paper Avenue, Pennsylvania, 18503 Dwight Schrute Yes 4 123 Dragons Road Υ 1005 876-678-3469 No Jon Snow 304-762-2467 768 City Parkway 5 1006 Ron Swanson Yes Yes 6 1007 1209 South Street No No Jeff Winger 98 Clue Drive 1008 Sherlock Holmes 876-678-3469 Ν No 8 123 Middle Earth NaN 1009 Gandalf NaN Yes 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 Potter 2394 Hogwarts Avenue Υ NaN Harry 123-543-2345 2039 Main Street Ν 12 1013 Don Draper Yes 13 1014 Leslie 876-678-3469 343 City Parkway Yes No Knope 214 HR Avenue 1015 304-762-2467 Ν No 14 Toby Flenderson 2395 Hogwarts Avenue 15 1016 Ron Weasley 123-545-5421 No Ν 1017 123-643-9775 No 16 Michael Scott 121 Paper Avenue, Pennsylvania Yes 17 1018 Clark Kent 3498 Super Lane NaN N/a 18 1019 N/a Yes Creed Braton 19 1020 Skywalker 876-678-3469 910 Tatooine Road, Tatooine Anakin Yes To expand the address column In [296.. #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 93 West Main Street None None 298 Drugs Driveway None None 980 Paper Avenue Pennsylvania 18503 123 Dragons Road None None 768 City Parkway None 6 1209 South Street None None 98 Clue Drive None None 8 123 Middle Earth None None 25th Main Street New York None 10 612 Shire Lane Shire None **11** 2394 Hogwarts Avenue None None 2039 Main Street 12 None None 13 343 City Parkway None None 214 HR Avenue 14 None None 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.. df[["Street_Address", "State", "Zip_Code"]] = df["Address"].str.split(',',2,expand=True) CustomerID First_Name Last_Name Phone_Number Address Paying Customer Do_Not_Contact Street_Address Out[297. State Zip_Code 123-545-5421 123 Shire Lane, Shire 123 Shire Lane Shire Frodo Baggins Yes None 1002 123-643-9775 93 West Main Street 1 Abed Nadir 93 West Main Street No Yes None None 2 1003 Walter White 298 Drugs Driveway Ν NaN 298 Drugs Driveway None None 123-543-2345 980 Paper Avenue, Pennsylvania, 18503 3 1004 Υ 18503 Dwight Schrute Yes 980 Paper Avenue Pennsylvania 876-678-3469 1005 Jon Snow 123 Dragons Road Υ No 123 Dragons Road None None 5 1006 304-762-2467 Yes 768 City Parkway Ron 768 City Parkway Yes None None Swanson 6 1007 Jeff 1209 South Street No No 1209 South Street Winger None None 876-678-3469 98 Clue Drive Ν 98 Clue Drive 7 1008 Sherlock Holmes No None None 8 1009 Gandalf NaN 123 Middle Earth NaN 123 Middle Earth Yes None None 123-545-5421 9 1010 Peter Parker 25th Main Street, New York No 25th Main Street New York None Yes 10 1011 612 Shire Lane, Shire 612 Shire Lane Shire Samwise Gamgee Yes None NaN 2394 Hogwarts Avenue 11 1012 2394 Hogwarts Avenue Harry Potter None None 123-543-2345 12 1013 2039 Main Street 2039 Main Street Don Draper Yes None None 13 1014 876-678-3469 343 City Parkway No 343 City Parkway Leslie Knope Yes None None 14 1015 Toby Flenderson 304-762-2467 214 HR Avenue Ν No 214 HR Avenue None None 1016 123-545-5421 2395 Hogwarts Avenue 2395 Hogwarts Avenue 15 Ron Weasley No None None 123-643-9775 121 Paper Avenue, Pennsylvania 16 1017 Michael Scott Yes 121 Paper Avenue Pennsylvania None 17 Clark 1018 Kent 3498 Super Lane NaN 3498 Super Lane None None 18 1019 Creed N/a Yes Braton None None 876-678-3469 19 1020 Anakin Skywalker 910 Tatooine Road, Tatooine Yes 910 Tatooine Road Tatooine None working on paying customer In [298.. df["Paying Customer"] = df["Paying Customer"].str.replace('Yes','Y') df["Paying Customer"] = df["Paying Customer"].str.replace('No','N') df CustomerID First_Name Last_Name Phone_Number Out[298.. Address Paying Customer Do_Not_Contact Street_Address State Zip_Code 1001 123-545-5421 123 Shire Lane, Shire 123 Shire Lane Shire None Frodo Baggins 123-643-9775 Ν 1 1002 Abed 93 West Main Street 93 West Main Street Nadir Yes None None 298 Drugs Driveway 2 1003 Walter White Ν NaN 298 Drugs Driveway None None 3 Υ 1004 Schrute 123-543-2345 980 Paper Avenue, Pennsylvania, 18503 980 Paper Avenue Pennsylvania 18503 Dwight 4 1005 Jon Snow 876-678-3469 123 Dragons Road Υ No 123 Dragons Road None None 5 1006 304-762-2467 Ron 768 City Parkway Yes 768 City Parkway None None Swanson 6 1007 Jeff Winger 1209 South Street Ν No 1209 South Street None None 7 Ν 1008 876-678-3469 No 98 Clue Drive Sherlock Holmes 98 Clue Drive None None 8 1009 Gandalf NaN 123 Middle Earth Υ NaN 123 Middle Earth None None 9 123-545-5421 25th Main Street, New York Υ 1010 Peter Parker No 25th Main Street New York None 10 1011 Samwise 612 Shire Lane, Shire Υ No 612 Shire Lane Shire None Gamgee 1012 Υ NaN 2394 Hogwarts Avenue 11 Harry Potter 2394 Hogwarts Avenue None None 12 1013 Don 123-543-2345 2039 Main Street Υ 2039 Main Street Draper None None Leslie 13 1014 876-678-3469 No 343 City Parkway Knope 343 City Parkway None None 14 1015 Toby Flenderson 304-762-2467 214 HR Avenue Ν No 214 HR Avenue None None 15 1016 123-545-5421 Ν N 2395 Hogwarts Avenue None Ron Weasley 2395 Hogwarts Avenue None 16 1017 Michael Scott 123-643-9775 121 Paper Avenue, Pennsylvania Υ No 121 Paper Avenue Pennsylvania None Υ 17 1018 Clark Kent 3498 Super Lane NaN 3498 Super Lane None None 18 1019 Creed **Braton** N/a N/a Yes None None 1020 Ν 19 Anakin Skywalker 876-678-3469 910 Tatooine Road, Tatooine 910 Tatooine Road **Tatooine** None Working on do_not_contact In [299.. 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 1001 Frodo **Baggins** 123-545-5421 123 Shire Lane, Shire Υ 123 Shire Lane Shire None 123-643-9775 Ν 1 1002 Abed 93 West Main Street 93 West Main Street Nadir None None 2 1003 Walter White 298 Drugs Driveway Ν NaN 298 Drugs Driveway None None 3 1004 123-543-2345 980 Paper Avenue, Pennsylvania, 18503 18503 Dwight Schrute 980 Paper Avenue Pennsylvania 4 1005 Jon Snow 876-678-3469 123 Dragons Road Υ 123 Dragons Road None None 5 1006 304-762-2467 Ron Swanson 768 City Parkway 768 City Parkway None None 6 1007 Jeff Winger 1209 South Street Ν Ν 1209 South Street None None Ν Ν 7 1008 876-678-3469 98 Clue Drive 98 Clue Drive Sherlock Holmes None None 8 1009 Gandalf NaN 123 Middle Earth Υ NaN 123 Middle Earth None None 9 1010 123-545-5421 Ν Peter Parker 25th Main Street, New York 25th Main Street New York None 10 1011 612 Shire Lane, Shire Υ Ν 612 Shire Lane Shire Samwise Gamgee None Υ NaN 2394 Hogwarts Avenue 1012 11 Harry Potter 2394 Hogwarts Avenue None None 12 1013 Don 123-543-2345 2039 Main Street Υ 2039 Main Street None Draper None 13 1014 Leslie Knope 876-678-3469 343 City Parkway 343 City Parkway None None 14 1015 Toby Flenderson 304-762-2467 214 HR Avenue Ν Ν 214 HR Avenue None Ν 15 1016 Ron Weasley 123-545-5421 2395 Hogwarts Avenue N 2395 Hogwarts Avenue None None 16 1017 Michael 123-643-9775 121 Paper Avenue, Pennsylvania Υ 121 Paper Avenue Pennsylvania Scott None 17 1018 Clark 3498 Super Lane NaN 3498 Super Lane None Kent None 18 1019 Creed **Braton** N/a Υ None None 19 1020 Anakin Skywalker 876-678-3469 910 Tatooine Road, Tatooine Ν 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.. for x in df.index: if df.loc[x,"Do_Not_Contact"] == 'Y': df.drop(x, inplace=True) df CustomerID First_Name Last_Name Phone_Number Street_Address Out[300.. Address Paying Customer Do_Not_Contact State Zip_Code 0 1001 Frodo 123-545-5421 123 Shire Lane, Shire Ν 123 Shire Lane Shire None **Baggins** 2 1003 Walter White 298 Drugs Driveway NaN 298 Drugs Driveway None None 4 1005 Jon Snow 876-678-3469 123 Dragons Road Υ Ν 123 Dragons Road None None 1209 South Street Ν Ν 1209 South Street 1007 Jeff Winger None None Holmes 7 1008 Sherlock 876-678-3469 98 Clue Drive Ν Ν 98 Clue Drive None None 8 1009 123 Middle Earth NaN 123 Middle Earth Gandalf NaN None None 9 25th Main Street, New York Υ 1010 Peter Parker 123-545-5421 Ν 25th Main Street New York None 10 1011 Samwise Gamgee 612 Shire Lane, Shire Ν 612 Shire Lane Shire None 11 1012 Harry Potter 2394 Hogwarts Avenue Υ NaN 2394 Hogwarts Avenue None None 2039 Main Street 12 1013 123-543-2345 Don Draper 2039 Main Street None None Υ None 13 1014 Leslie 876-678-3469 343 City Parkway Ν 343 City Parkway Knope None 304-762-2467 214 HR Avenue Ν Ν 214 HR Avenue 14 1015 Toby Flenderson None None None 15 1016 Ron Weasley 123-545-5421 2395 Hogwarts Avenue Ν 2395 Hogwarts Avenue None 1017 123-643-9775 121 Paper Avenue, Pennsylvania 16 Michael Scott 121 Paper Avenue Pennsylvania None 17 Υ 1018 Clark Kent 3498 Super Lane NaN 3498 Super Lane None None 19 1020 Skywalker 876-678-3469 Ν Anakin 910 Tatooine Road, Tatooine 910 Tatooine Road **Tatooine** None To remove the empty space in "phone_Number" column In [301.. 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 123-545-5421 123 Shire Lane, Shire Ν 123 Shire Lane Frodo Baggins Shire None 123 Dragons Road 4 1005 Snow 876-678-3469 Ν 123 Dragons Road Jon None None 98 Clue Drive 7 1008 Sherlock Holmes 876-678-3469 98 Clue Drive Ν Ν None None 9 1010 123-545-5421 Ν 25th Main Street Peter Parker 25th Main Street, New York New York None 2039 Main Street 12 1013 Don Draper 123-543-2345 Υ Ν 2039 Main Street None None 13 876-678-3469 Ν 343 City Parkway 1014 Leslie 343 City Parkway Knope None None 214 HR Avenue 14 1015 304-762-2467 Ν Ν 214 HR Avenue Toby Flenderson None None 15 1016 Weasley Ν Ron 123-545-5421 2395 Hogwarts Avenue N 2395 Hogwarts Avenue None None 16 1017 Michael 123-643-9775 121 Paper Avenue, Pennsylvania 121 Paper Avenue Pennsylvania None 19 1020 Anakin Skywalker 876-678-3469 910 Tatooine Road, Tatooine 910 Tatooine Road **Tatooine** None To reset the index In [302.. df = df.reset_index(drop=True) CustomerID First Name Last Name Phone Number Out[302.. Address Paying Customer Do_Not_Contact Street_Address State Zip_Code 0 1001 123-545-5421 123 Shire Lane, Shire 123 Shire Lane Frodo **Baggins** Shire None 1005 Ν 123 Dragons Road 1 Jon 876-678-3469 123 Dragons Road Snow None None 2 Ν 1008 876-678-3469 98 Clue Drive Ν 98 Clue Drive Sherlock Holmes None None 3 Ν 25th Main Street 1010 Peter Parker 123-545-5421 25th Main Street, New York New York None Υ 4 1013 Don Draper 123-543-2345 2039 Main Street Ν 2039 Main Street None None 5 1014 Leslie Knope 876-678-3469 343 City Parkway Ν 343 City Parkway None None 6 1015 304-762-2467 Ν Ν Toby Flenderson 214 HR Avenue 214 HR Avenue None None 7 1016 N 2395 Hogwarts Avenue Ron Weasley 123-545-5421 2395 Hogwarts Avenue Ν None None 8 121 Paper Avenue, Pennsylvania Υ 1017 Michael Scott 123-643-9775 Ν 121 Paper Avenue Pennsylvania None 1020 Anakin Skywalker 876-678-3469 910 Tatooine Road, Tatooine 910 Tatooine Road Tatooine None To remove "None" in all the column In [303.. df.fillna('') Address Paying Customer Do_Not_Contact State Zip_Code Out[303.. CustomerID First_Name Last_Name Phone_Number Street_Address 0 1001 Frodo Baggins 123-545-5421 123 Shire Lane, Shire Ν 123 Shire Lane Shire 1 1005 Jon 876-678-3469 123 Dragons Road 123 Dragons Road Snow 2 1008 876-678-3469 98 Clue Drive Ν Ν Sherlock 98 Clue Drive Holmes 3 1010 Peter Parker 123-545-5421 25th Main Street, New York Ν 25th Main Street New York 1013 2039 Main Street Υ Ν 4 123-543-2345 2039 Main Street Don Draper 5 1014 Leslie Knope 876-678-3469 343 City Parkway 343 City Parkway 214 HR Avenue 6 1015 304-762-2467 Ν Ν 214 HR Avenue Toby Flenderson 7 1016 Ron Weasley 123-545-5421 2395 Hogwarts Avenue N 2395 Hogwarts Avenue Υ 8 1017 123-643-9775 121 Paper Avenue, Pennsylvania Michael Scott Ν 121 Paper Avenue Pennsylvania 876-678-3469 1020 Anakin Skywalker 910 Tatooine Road, Tatooine 910 Tatooine Road 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.