Cont. Pandas - II

🎯 Session Objectives:

- ✅ Understand the key attributes of a DataFrame
- ✅ Use various DataFrame methods to explore and manipulate data
- ✅ Perform row and column operations
- ✅ Use joining, merging, and concatenation techniques across DataFrames

```python
import numpy as np
import pandas as pd
# Import the DataSet....
try :
    Customers = pd.read_excel('customers.xlsx')
    Products = pd.read_excel('products.xlsx')
    Purchases = pd.read_excel('purchases.xlsx')
except FileNotFoundError:
    print("File Not Found!")
```

```python
# DataFrame Attribute [.index] -> range
Customers.index # range(0,1000,step=1)
```

```
RangeIndex(start=0, stop=1000, step=1)
```

```python
Products.index
```

```
RangeIndex(start=0, stop=60, step=1)
```

```python
Purchases.index
```

```
RangeIndex(start=0, stop=6000, step=1)
```

```python
# .shape(rows , cols) # 2D Matrix [DataFrame]
Customers.shape
```

```
(1000, 11)
```

```python
Products.shape
```

```
(60, 5)
```

```python
Purchases.shape
```

```
(6000, 7)
```

```python
Purchases.shape[0] # rows
```

```
6000
```

```python
Purchases.shape[1] # cols
```

```
7
```

```python
# .columns [providing column headers [label]]
Customers.columns
```

```
Index(['id', 'first_name', 'last_name', 'email', 'gender', 'street_num',
       'street_name', 'street_suffix', 'city', 'state', 'postcode'],
      dtype='object')
```

```python
Products.columns
```

```
Index(['Unnamed: 0', 'id', 'product', 'cost', 'company'], dtype='object')
```

```python
Purchases.columns
```

```
Index(['Unnamed: 0', 'id', 'purch_date', 'customer_num', 'product_num',
       'amount', 'paid'],
      dtype='object')
```

```python
# .dtypes -> Returns the data types of each columns
Customers.dtypes
```

```
id                int64
first_name       object
last_name        object
email            object
gender           object
street_num      float64
street_name      object
street_suffix    object
city             object
state            object
postcode        float64
dtype: object
```

```python
Products.dtypes
```

```
Unnamed: 0     int64
id             int64
product       object
cost          object
company       object
dtype: object
```

```
Purchases.dtypes

Unnamed: 0      int64
id              int64
purch_date      object
customer_num    int64
product_num     int64
amount          int64
paid            object
dtype: object
```

```
# Describe <by default numeric columns>
# Objects columns -> include = "O"
Customers.describe()
```

	id	street_num	postcode
count	1000.000000	738.000000	843.000000
mean	500.500000	10536.439024	52669.548043
std	288.819436	23050.537603	28140.041026
min	1.000000	0.000000	214.000000
25%	250.750000	21.000000	29279.500000
50%	500.500000	445.500000	48232.000000
75%	750.250000	6918.500000	78337.500000
max	1000.000000	99918.000000	99812.000000

```
Customers.describe(include = 'object')
```

	first_name	last_name	email	gender	street_name	street_suffix	city	state
count	1000	1000	878	957	963	963	921	920
unique	932	993	878	2	427	21	296	48
top	Berty	Sedworth	rsouthcott0@clickbank.net	Male	Arizona	Place	Washington	Texas
freq	4	3	1	485	6	57	29	107

```
Customers.describe(include = 'O')
```

	first_name	last_name	email	gender	street_name	street_suffix	city	state
count	1000	1000	878	957	963	963	921	920
unique	932	993	878	2	427	21	296	48
top	Berty	Sedworth	rsouthcott0@clickbank.net	Male	Arizona	Place	Washington	Texas
freq	4	3	1	485	6	57	29	107

```python
Customers.describe(include = 'all')
```

	id	first_name	last_name		email	gender	street_num	street_name	street_suffix	city	stat
count	1000.000000	1000	1000		878	957	738.000000	963	963	921	92
unique	NaN	932	993		878	2	NaN	427	21	296	4
top	NaN	Berty	Sedworth	rsouthcott0@clickbank.net	Male	NaN	Arizona	Place	Washington	Texa
freq	NaN	4	3		1	485	NaN	6	57	29	10
mean	500.500000	NaN	NaN		NaN	NaN	10536.439024	NaN	NaN	NaN	Nal
std	288.819436	NaN	NaN		NaN	NaN	23050.537603	NaN	NaN	NaN	Nal
min	1.000000	NaN	NaN		NaN	NaN	0.000000	NaN	NaN	NaN	Nal
25%	250.750000	NaN	NaN		NaN	NaN	21.000000	NaN	NaN	NaN	Nal
50%	500.500000	NaN	NaN		NaN	NaN	445.500000	NaN	NaN	NaN	Nal
75%	750.250000	NaN	NaN		NaN	NaN	6918.500000	NaN	NaN	NaN	Nal
max	1000.000000	NaN	NaN		NaN	NaN	99918.000000	NaN	NaN	NaN	Nal

```python
# .values -> Reads the data of a DataFrame [2D Matrix]
Customers.values
```

```
array([[1, 'Romain', 'Southcott', ..., 'San Diego', 'California',
        92127.0],
       [2, 'Cosimo', 'Molyneaux', ..., 'El Paso', 'Texas', nan],
       [3, 'Bambi', 'Westrip', ..., 'San Antonio', 'Texas', 78220.0],
       ...,
       [998, 'Tobit', 'Birt', ..., 'Waterbury', 'Connecticut', 6721.0],
       [999, 'Issiah', 'Standbrooke', ..., 'Savannah', 'Georgia',
        31405.0],
       [1000, 'Elmore', 'Malpas', ..., 'Atlanta', 'Georgia', 30386.0]],
      dtype=object)
```

```python
Products.values
```

```
array([[0, 1, 'Liners - Baking Cups', '$6.36', 'Skipfire'],
       [1, 2, 'Nori Sea Weed - Gold Label', '$85.74', 'Dynazzy'],
       [2, 3, 'Bar Bran Honey Nut', '$65.40', 'Ntag'],
       [3, 4, 'Soup - Campbells Beef Stew', '$68.16', 'Photojam'],
       [4, 5, 'Wine - Shiraz Wolf Blass Premium', '$87.39', 'Eare'],
       [5, 6, 'Wine - White, Riesling, Semi - Dry', '$99.22', 'Livepath'],
       [6, 7, 'Brandy - Bar', '$13.83', 'Oloo'],
       [7, 8, 'Onions - White', '$42.19', 'Oozz'],
       [8, 9, 'Lettuce - Baby Salad Greens', '$30.01', 'Meevee'],
       [9, 10, 'Sambuca - Ramazzotti', '$88.99', 'Livepath'],
       [10, 11, 'Coffee - Decafenated', '$21.47', 'Meezzy'],
       [11, 12, 'Lamb Leg - Bone - In Nz', '$42.56', 'Photofeed'],
       [12, 13, 'Lettuce - Spring Mix', '$19.90', 'Twitterworks'],
       [13, 14, 'Cookies Oatmeal Raisin', '$99.40', 'Dynabox'],
       [14, 15, 'Cookies Oatmeal Raisin', '$14.13', 'Vitz'],
       [15, 16, 'Sausage - Chorizo', '$55.45', nan],
```

```
Purchases.values
```

```
array([[0, 1, datetime.datetime(2019, 1, 3, 0, 0), ..., 27, 12,
        '$568.92'],
       [1, 2, datetime.datetime(2019, 1, 3, 0, 0), ..., 28, 14,
        '$395.36'],
       [2, 3, datetime.datetime(2019, 1, 3, 0, 0), ..., 9, 17, '$510.17'],
       ...,
       [5997, 5998, '06/20/2019', ..., 19, 9, '$205.47'],
       [5998, 5999, '06/20/2019', ..., 11, 20, '$429.40'],
       [5999, 6000, '06/20/2019', ..., 57, 4, '$274.52']], dtype=object)
```

```python
# .size -> Customers.shape (1000,11) = 11000
Customers.size
```

```
11000
```

```python
Products.size
```

```
300
```

```python
Purchases.shape
```

```
(6000, 7)
```

```python
Purchases.size
```

```
42000
```

```python
# .empty -> Checks if the DataFrame is empty or not
Customers.empty
```

```
False
```

```python
Products.empty
```

```
False
```

```python
empty_df = pd.DataFrame()
print(empty_df)
print(empty_df.empty) # True
```

```
Empty DataFrame
Columns: []
Index: []
True
```

```python
# .ndim -> N Dimensional -> 2D Matrix[DataFrame] & 1D Matrix[Series]
Customers.ndim
```

```
2
```

```python
Products.ndim
```

```
2
```

```
# Top 10 Records
Customers.head(10)
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city	state
0	1	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	California
1	2	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	Texas
2	3	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	Texas
3	4	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Tennessee
4	5	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	NaN
5	6	Magdalena	Cullip	mcullip5@tiny.cc	Female	9190.0	Packers	Drive	Baltimore	Maryland
6	7	Marietta	Heball	mheball6@blog.com	Female	12.0	Mallory	Center	Carol Stream	Illinois
7	8	Tine	McSperrin	tmcsperrin7@statcounter.com	NaN	530.0	Erie	Plaza	Kansas City	Missouri
8	9	Enrichetta	de Villier	edevillier8@ox.ac.uk	Female	745.0	Annamark	Street	Fairbanks	Alaska
9	10	Sari	Poulden	spoulden9@xing.com	Female	2.0	Pond	Hill	New Orleans	Louisiana

```
# .tail(10) -> Bottom 10
Products.tail(10)
```

	Unnamed: 0	id	product	cost	company
50	50	51	Ecolab - Power Fusion	$75.32	Fliptune
51	51	52	Soup - Knorr, Country Bean	$98.74	Eimbee
52	52	53	Chicken - Thigh, Bone In	$74.06	Topiclounge
53	53	54	Plate - Foam, Bread And Butter	$37.79	Flipstorm
54	54	55	Beef - Bresaola	$8.55	Digitube
55	55	56	Peach - Fresh	$20.53	Zazio
56	56	57	Muffin Hinge - 211n	$68.63	Fanoodle
57	57	58	Bar Bran Honey Nut	$73.05	Yakijo
58	58	59	Nut - Almond, Blanched, Whole	$74.28	Eazzy
59	59	60	Table Cloth 90x90 Colour	$41.22	Quinu

```
# Transpose -> .T or .Transpose
Products.T
```

	0	1	2	3	4	5	6	7	8	9	...	50	51	5
Unnamed: 0	0	1	2	3	4	5	6	7	8	9	...	50	51	5
id	1	2	3	4	5	6	7	8	9	10	...	51	52	5
product	Liners - Baking Cups	Nori Sea Weed - Gold Label	Bar Bran Honey Nut	Soup - Campbells Beef Stew	Wine - Shiraz Wolf Blass Premium	Wine - White, Riesling, Semi - Dry	Brandy - Bar	Onions - White	Lettuce - Baby Salad Greens	Sambuca - Ramazzotti	...	Ecolab - Power Fusion	Soup - Knorr, Country Bean	Chicken Thigh, Bon I
cost	$6.36	$85.74	$65.40	$68.16	$87.39	$99.22	$13.83	$42.19	$30.01	$88.99	...	$75.32	$98.74	$74.0
company	Skipfire	Dynazzy	Ntag	Photojam	Eare	Livepath	Oloo	Oozz	Meevee	Livepath	...	Fliptune	Eimbee	Topiclounge

5 rows × 60 columns

```
Products.T.T
```

	Unnamed: 0	id	product	cost	company
0	0	1	Liners - Baking Cups	$6.36	Skipfire
1	1	2	Nori Sea Weed - Gold Label	$85.74	Dynazzy
2	2	3	Bar Bran Honey Nut	$65.40	Ntag
3	3	4	Soup - Campbells Beef Stew	$68.16	Photojam
4	4	5	Wine - Shiraz Wolf Blass Premium	$87.39	Eare
5	5	6	Wine - White, Riesling, Semi - Dry	$99.22	Livepath
6	6	7	Brandy - Bar	$13.83	Oloo
7	7	8	Onions - White	$42.19	Oozz
8	8	9	Lettuce - Baby Salad Greens	$30.01	Meevee

```
Products.transpose().transpose()
```

	Unnamed: 0	id	product	cost	company
0	0	1	Liners - Baking Cups	$6.36	Skipfire
1	1	2	Nori Sea Weed - Gold Label	$85.74	Dynazzy
2	2	3	Bar Bran Honey Nut	$65.40	Ntag
3	3	4	Soup - Campbells Beef Stew	$68.16	Photojam
4	4	5	Wine - Shiraz Wolf Blass Premium	$87.39	Eare
5	5	6	Wine - White, Riesling, Semi - Dry	$99.22	Livepath
6	6	7	Brandy - Bar	$13.83	Oloo
7	7	8	Onions - White	$42.19	Oozz
8	8	9	Lettuce - Baby Salad Greens	$30.01	Meevee

```python
# Operations on Rows & Columns in DataFrames
# Create a New Columns -> Adding a Column [Featuring Engineering]
Customers['Country'] = 'United States'
Customers
```

last_name	email	gender	street_num	street_name	street_suffix	city	state	postcode	Country
Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	California	92127.0	United States
Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	Texas	NaN	United States
Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	Texas	78220.0	United States
ankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Tennessee	38150.0	United States
Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	NaN	83732.0	United States
...	...	...	...	...	...	...	...	...	...
Alman	malmanrn@cornell.edu	Female	0.0	Thompson	Way	Reading	Pennsylvania	19610.0	United States
Heckle	whecklero@fc2.com	Female	701.0	Rowland	Hill	Indianapolis	Indiana	46278.0	United States
Birt	NaN	Male	2.0	Basil	Road	Waterbury	Connecticut	6721.0	United States
andbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive	Savannah	Georgia	31405.0	United States
Malpas	NaN	Male	205.0	Farwell	Park	Atlanta	Georgia	30386.0	United States

```python
# Membership Status -> ['Diamond','Platinum','Gold','Silver','Bronze']
Customers['Membership Status'] =  ['Diamond','Platinum','Gold','Silver','Bronze']
Customers
# ValueError: Length of values (5) does not match length of index (1000)
```

```python
Customers['Membership Status'] = 'Gold'
```

```python
membership_status =  ['Diamond','Platinum','Gold','Silver','Bronze']
Customers.loc[:4, 'Membership_status'] = membership_status
Customers
```

email	gender	street_num	street_name	street_suffix	city	state	postcode	Country	Membership_status
outhcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	California	92127.0	United States	Diamond
cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	Texas	NaN	United States	Platinum
westrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	Texas	78220.0	United States	Gold
pankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Tennessee	38150.0	United States	Silver
malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	NaN	83732.0	United States	Bronze
...	...	...	...	...	...	...	...	...	...
malmanrn@cornell.edu	Female	0.0	Thompson	Way	Reading	Pennsylvania	19610.0	United States	nan
whecklero@fc2.com	Female	701.0	Rowland	Hill	Indianapolis	Indiana	46278.0	United States	nan
NaN	Male	2.0	Basil	Road	Waterbury	Connecticut	6721.0	United States	nan

```python
membership_status = ['Diamond','Platinum','Gold','Silver','Bronze']
Customers['Membership Status'] = np.tile(membership_status,
                                 len(Customers) // len(membership_status)
                                 )
Customers
```

email	gender	street_num	street_name	street_suffix	city	state	postcode	Country	Membership_status	Membership Status
:bank.net	Male	1.0	Trailsway	Road	San Diego	California	92127.0	United States	Diamond	Diamond
viley.com	Male	NaN	NaN	NaN	El Paso	Texas	NaN	United States	Platinum	Platinum
ntec.com	Female	4057.0	Arkansas	Circle	San Antonio	Texas	78220.0	United States	Gold	Gold
viley.com	Male	74.0	Debs	Point	Memphis	Tennessee	38150.0	United States	Silver	Silver
e4@51.la	NaN	NaN	2nd	Drive	NaN	NaN	83732.0	United States	Bronze	Bronze
...	...	...	...	...	...	...	...	...	...	...
rnell.edu	Female	0.0	Thompson	Way	Reading	Pennsylvania	19610.0	United States	nan	Diamond
Əfc2.com	Female	701.0	Rowland	Hill	Indianapolis	Indiana	46278.0	United States	nan	Platinum
NaN	Male	2.0	Basil	Road	Waterbury	Connecticut	6721.0	United States	nan	Gold
ages.com	Male	NaN	Kenwood	Drive	Savannah	Georgia	31405.0	United States	nan	Silver
NaN	Male	205.0	Farwell	Park	Atlanta	Georgia	30386.0	United States	nan	Bronze

```python
# Assign based on Conditions
Customers['membershipStatus'] = Customers['state'].apply(
    lambda x : 'Gold' if x == 'Texas' else 'Silver'
)
Customers
```

et_num	street_name	street_suffix	city	state	postcode	Country	Membership_status	Membership Status	membershipStatus
1.0	Trailsway	Road	San Diego	California	92127.0	United States	Diamond	Diamond	Silver
NaN	NaN	NaN	El Paso	Texas	NaN	United States	Platinum	Platinum	Gold
4057.0	Arkansas	Circle	San Antonio	Texas	78220.0	United States	Gold	Gold	Gold
74.0	Debs	Point	Memphis	Tennessee	38150.0	United States	Silver	Silver	Silver
NaN	2nd	Drive	NaN	NaN	83732.0	United States	Bronze	Bronze	Silver
...	...	...	...	...	...	...	...	...	...

```python
# Assign based on Conditions
Customers['membershipStatus'] = Customers['state'].apply(
    lambda x : 'Gold' if (x == 'Texas') or (x == 'California') or (x == 'Georgia') else 'Silver'
)
Customers
```

treet_num	street_name	street_suffix	city	state	postcode	Country	Membership_status	Membership Status	membershipStatus
1.0	Trailsway	Road	San Diego	California	92127.0	United States	Diamond	Diamond	Gold
NaN	NaN	NaN	El Paso	Texas	NaN	United States	Platinum	Platinum	Gold
4057.0	Arkansas	Circle	San Antonio	Texas	78220.0	United States	Gold	Gold	Gold
74.0	Debs	Point	Memphis	Tennessee	38150.0	United States	Silver	Silver	Silver
NaN	2nd	Drive	NaN	NaN	83732.0	United States	Bronze	Bronze	Silver
...	...	...	...	...	...	...	...	...	...
0.0	Thompson	Way	Reading	Pennsylvania	19610.0	United States	nan	Diamond	Silver
701.0	Rowland	Hill	Indianapolis	Indiana	46278.0	United States	nan	Platinum	Silver
2.0	Basil	Road	Waterbury	Connecticut	6721.0	United States	nan	Gold	Silver
NaN	Kenwood	Drive	Savannah	Georgia	31405.0	United States	nan	Silver	Gold
205.0	Farwell	Park	Atlanta	Georgia	30386.0	United States	nan	Bronze	Gold

```python
# Adding a new columns 'CustomerName'
Customers['CustomerName'] = Customers['first_name'] + ' ' + Customers['last_name']
Customers
```

_name	street_suffix	city	state	postcode	Country	Membership_status	Membership Status	membershipStatus	CustomerName
ailsway	Road	San Diego	California	92127.0	United States	Diamond	Diamond	Gold	Romain Southcott
NaN	NaN	El Paso	Texas	NaN	United States	Platinum	Platinum	Gold	Cosimo Molyneaux
rkansas	Circle	San Antonio	Texas	78220.0	United States	Gold	Gold	Gold	Bambi Westrip
Debs	Point	Memphis	Tennessee	38150.0	United States	Silver	Silver	Silver	Roarke Pankettman
2nd	Drive	NaN	NaN	83732.0	United States	Bronze	Bronze	Silver	Mikaela Althorpe
...	...	...	...	...	...	...	...	...	...
mpson	Way	Reading	Pennsylvania	19610.0	United States	nan	Diamond	Silver	Merrili Alman
owland	Hill	Indianapolis	Indiana	46278.0	United States	nan	Platinum	Silver	Winonah Heckle
Basil	Road	Waterbury	Connecticut	6721.0	United States	nan	Gold	Silver	Tobit Birt
nwood	Drive	Savannah	Georgia	31405.0	United States	nan	Silver	Gold	Issiah Standbrooke
Farwell	Park	Atlanta	Georgia	30386.0	United States	nan	Bronze	Gold	Elmore Malpas

```
Customers.columns
```

```
Index(['id', 'first_name', 'last_name', 'email', 'gender', 'street_num',
       'street_name', 'street_suffix', 'city', 'state', 'postcode', 'Country',
       'Membership_status', 'Membership Status', 'membershipStatus',
       'CustomerName'],
      dtype='object')
```

```
Customers.dtypes
```

```
id                   int64
first_name          object
last_name           object
email               object
gender              object
street_num         float64
street_name         object
street_suffix       object
city                object
state               object
postcode           float64
Country             object
Membership_status   object
Membership Status   object
membershipStatus    object
CustomerName        object
dtype: object
```

```python
# Let's do the insert command in SQL [Row Manipulation]
# <NA>['Int64'] , None , NaN , NaT
Customers.loc['1000'] = [1001, 'Palash' , 'Mudgal', 'palash.mudgal@gmail.com', 'M', 11, 'NH-8', 'Park',
                         'San Diego', 'California' , 92127.0 , 'United States', None , None , None , None]
Customers
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city	
0	1	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	Cal
1	2	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	
2	3	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	
3	4	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Tenr
4	5	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	
...	...	...	...	...	...	...	...	...	...	
996	997	Winonah	Heckle	whecklero@fc2.com	Female	701.0	Rowland	Hill	Indianapolis	lı
997	998	Tobit	Birt	NaN	Male	2.0	Basil	Road	Waterbury	Conn
998	999	Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive	Savannah	G
999	1000	Elmore	Malpas	NaN	Male	205.0	Farwell	Park	Atlanta	G
1000	1001	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	Cal

```
Customers.loc['1001'] = [1002, 'Palash' , 'Mudgal', 'palash.mudgal@gmail.com', 'M', 11, 'NH-8', 'Park',
                         'San Diego', 'California' , 92127.0 , 'United States', np.nan , np.nan , np.nan , np.n
Customers
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city	s
0	1	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	Calif
1	2	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	T
2	3	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	T
3	4	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Tenne
4	5	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	
...	...	...	...	...	...	...	...	...	...	
997	998	Tobit	Birt	NaN	Male	2.0	Basil	Road	Waterbury	Conne
998	999	Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive	Savannah	Ge
999	1000	Elmore	Malpas	NaN	Male	205.0	Farwell	Park	Atlanta	Ge
1000	1001	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	Calif
1001	1002	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	Calif

```
Customers.loc['1002'] = None
Customers
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city	s
0	1.0	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	Califo
1	2.0	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	Te
2	3.0	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	Te
3	4.0	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Tennes
4	5.0	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	N
...	...	...	...	...	...	...	...	...	...	
998	999.0	Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive	Savannah	Geo
999	1000.0	Elmore	Malpas	NaN	Male	205.0	Farwell	Park	Atlanta	Geo
1000	1001.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	Califo
1001	1002.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	Califo
1002	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N

1003 rows × 16 columns

```
# Drop a row() -> axis = 0 [Horizontal] [Rows] [Default]
Customers.drop(index = '1002' , axis = 0 , inplace = True)
Customers
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city	
0	1.0	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	Cal
1	2.0	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	
2	3.0	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	
3	4.0	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Tenr
4	5.0	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	
...	...	...	...	...	...	...	...	...	...	
997	998.0	Tobit	Birt	NaN	Male	2.0	Basil	Road	Waterbury	Conn
998	999.0	Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive	Savannah	G
999	1000.0	Elmore	Malpas	NaN	Male	205.0	Farwell	Park	Atlanta	G
1000	1001.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	Cal
1001	1002.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	Cal

1002 rows × 16 columns

```
# Dropping multiple rows
Customers.drop(index = ['1000','1001'] , axis = 0 , inplace = True)
Customers
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city	
0	1.0	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	Ca
1	2.0	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	
2	3.0	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	
3	4.0	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Ter
4	5.0	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	
...	...	...	...	...	...	...	...	...	...	
995	996.0	Merrili	Alman	malmanrn@cornell.edu	Female	0.0	Thompson	Way	Reading	Penn:
996	997.0	Winonah	Heckle	whecklero@fc2.com	Female	701.0	Rowland	Hill	Indianapolis	
997	998.0	Tobit	Birt	NaN	Male	2.0	Basil	Road	Waterbury	Conn
998	999.0	Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive	Savannah	
999	1000.0	Elmore	Malpas	NaN	Male	205.0	Farwell	Park	Atlanta	

1000 rows × 16 columns

```
Customers.loc[1000] = None
Customers
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city	
0	1.0	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	C
1	2.0	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	
2	3.0	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	
3	4.0	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Te
4	5.0	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	
...	...	...	...	...	...	...	...	...	...	
996	997.0	Winonah	Heckle	whecklero@fc2.com	Female	701.0	Rowland	Hill	Indianapolis	
997	998.0	Tobit	Birt	NaN	Male	2.0	Basil	Road	Waterbury	Con
998	999.0	Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive	Savannah	
999	1000.0	Elmore	Malpas	NaN	Male	205.0	Farwell	Park	Atlanta	
1000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	

1001 rows × 16 columns

```
for x in range(1001,1011):
    Customers.loc[x] = None

Customers
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city	state	...	1
0	1.0	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	California	...	N
1	2.0	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	Texas	...	N
2	3.0	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	Texas	...	N
3	4.0	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Tennessee	...	N
4	5.0	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	NaN	...	N
...	...	...	...	...	...	...	...	...	...	...	...	
1006	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	N
1007	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	N
1008	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	N
1009	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	N
1010	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	

1011 rows × 26 columns

```python
idx_slicing = Customers.iloc[1000:].index
idx_slicing
```

```
Index([1000, 1001, 1002, 1003, 1004, 1005, 1006, 1007, 1008, 1009, 1010], dtype='int64')
```

```python
# Dropping multiple rows
Customers.drop(index = idx_slicing , axis = 0 , inplace = True)
Customers
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city	
0	1.0	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	Ca
1	2.0	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	
2	3.0	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	
3	4.0	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Ter
4	5.0	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	
...	...	...	...	...	...	...	...	...	...	
995	996.0	Merrili	Alman	malmanrn@cornell.edu	Female	0.0	Thompson	Way	Reading	Penns
996	997.0	Winonah	Heckle	whecklero@fc2.com	Female	701.0	Rowland	Hill	Indianapolis	
997	998.0	Tobit	Birt	NaN	Male	2.0	Basil	Road	Waterbury	Coni
998	999.0	Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive	Savannah	(
999	1000.0	Elmore	Malpas	NaN	Male	205.0	Farwell	Park	Atlanta	(

1000 rows × 26 columns

```python
Customers.columns
```

```
Index([                       'id',          'first_name',          'last_name',
                          'email',              'gender',          'street_num',
                    'street_name',       'street_suffix',                'city',
                          'state',            'postcode',             'Country',
              'Membership_status', 'Membership Status',   'membershipStatus',
                   'CustomerName',                  1001,                  1002,
                             1003,                  1004,                  1005,
                             1006,                  1007,                  1008,
                             1009,                  1010],
      dtype='object')
```

```python
mapper = [1001,1002,1003,1004,1005,1006,1007,1008,1009,1010]
Customers.drop(mapper , axis = 1 , inplace = True)
Customers
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city
0	1.0	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego
1	2.0	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso
2	3.0	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio
3	4.0	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis
4	5.0	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN
...	...	...	...	...	...	...	...	...	...
995	996.0	Merrili	Alman	malmanrn@cornell.edu	Female	0.0	Thompson	Way	Reading
996	997.0	Winonah	Heckle	whecklero@fc2.com	Female	701.0	Rowland	Hill	Indianapolis
997	998.0	Tobit	Birt	NaN	Male	2.0	Basil	Road	Waterbury
998	999.0	Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive	Savannah

```python
Customers.loc['1000'] = [1001, 'Palash' , 'Mudgal', 'palash.mudgal@gmail.com', 'M', 11, 'NH-8', 'Park',
                         'San Diego', 'California' , 92127.0 , 'United States', np.nan , np.nan , np.nan , np.
Customers.loc['1001'] = [1002, 'Palash' , 'Mudgal', 'palash.mudgal@gmail.com', 'M', 11, 'NH-8', 'Park',
                         'San Diego', 'California' , 92127.0 , 'United States', np.nan , np.nan , np.nan , np.
Customers.loc['1002'] = [1003, 'Palash' , 'Mudgal', 'palash.mudgal@gmail.com', 'M', 11, 'NH-8', 'Park',
                         'San Diego', 'California' , 92127.0 , 'United States', np.nan , np.nan , np.nan , np.
Customers.loc['1003'] = [1004, 'Palash' , 'Mudgal', 'palash.mudgal@gmail.com', 'M', 11, 'NH-8', 'Park',
                         'San Diego', 'California' , 92127.0 , 'United States', np.nan , np.nan , np.nan , np.
Customers.loc['1004'] = [1005, 'Palash' , 'Mudgal', 'palash.mudgal@gmail.com', 'M', 11, 'NH-8', 'Park',
                         'San Diego', 'California' , 92127.0 , 'United States', np.nan , np.nan , np.nan , np.
Customers
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city	state	p
0	1.0	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	California	
1	2.0	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	Texas	
2	3.0	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	Texas	
3	4.0	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Tennessee	
4	5.0	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	NaN	
...	...	...	...	...	...	...	...	...	...	...	
1000	1001.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	California	
1001	1002.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	California	
1002	1003.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	California	
1003	1004.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	California	
1004	1005.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	California	

1005 rows × 16 columns

```python
Customers[Customers['id'].between(1001,1005)]
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city	state	postcod
1000	1001.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	California	92127.
1001	1002.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	California	92127.
1002	1003.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	California	92127.
1003	1004.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	California	92127.
1004	1005.0	Palash	Mudgal	palash.mudgal@gmail.com	M	11.0	NH-8	Park	San Diego	California	92127.

```python
idx_range = Customers[Customers['id'].between(1001,1005)].index
idx_range
```
```
Index(['1000', '1001', '1002', '1003', '1004'], dtype='object')
```

```python
Customers = Customers.drop(idx_range)
Customers
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix	city	
0	1.0	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	Ca
1	2.0	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	
2	3.0	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	
3	4.0	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Ten
4	5.0	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	
...	...	...	...	...	...	...	...	...	...	
995	996.0	Merrili	Alman	malmanrn@cornell.edu	Female	0.0	Thompson	Way	Reading	Penns
996	997.0	Winonah	Heckle	whecklero@fc2.com	Female	701.0	Rowland	Hill	Indianapolis	I
997	998.0	Tobit	Birt	NaN	Male	2.0	Basil	Road	Waterbury	Conr
998	999.0	Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive	Savannah	C

```
Customers.info()
```
```
<class 'pandas.core.frame.DataFrame'>
Index: 1000 entries, 0 to 999
Data columns (total 16 columns):
 #   Column             Non-Null Count   Dtype
---  ------             --------------   -----
 0   id                 1000 non-null    float64
 1   first_name         1000 non-null    object
 2   last_name          1000 non-null    object
 3   email              878 non-null     object
 4   gender             957 non-null     object
 5   street_num         738 non-null     float64
 6   street_name        963 non-null     object
 7   street_suffix      963 non-null     object
 8   city               921 non-null     object
 9   state              920 non-null     object
 10  postcode           843 non-null     float64
 11  Country            1000 non-null    object
 12  Membership_status  1000 non-null    object
 13  Membership Status  1000 non-null    object
 14  membershipStatus   1000 non-null    object
 15  CustomerName       1000 non-null    object
dtypes: float64(3), object(13)
memory usage: 165.1+ KB
```

```python
Customers[Customers['postcode'].isna()]
```

first_name	last_name	email	gender	street_num	street_name	street_suffix	city	state	postcode
Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	Texas	NaN
Sari	Poulden	spoulden9@xing.com	Female	2.0	Pond	Hill	New Orleans	Louisiana	NaN
Tobi	Dewicke	tdewickeg@hostgator.com	Female	6817.0	Orin	Junction	San Rafael	NaN	NaN
Karim	Woosnam	kwoosnamk@ifeng.com	Male	NaN	Grayhawk	Place	Houston	Texas	NaN
Woodrow	Fidoe	wfidoel@creativecommons.org	Male	1.0	Artisan	Junction	Gainesville	NaN	NaN
...	...		...	...	...	...	...	...	...
Ruddie	Eckhard	reckhardqk@free.fr	Male	853.0	Little Fleur	Trail	Abilene	Texas	NaN
Vanya	Headey	vheadeyqo@networkadvertising.org	Female	NaN	Starling	Point	Hicksville	New York	NaN
Corena	Pelz	NaN	Female	62987.0	Grayhawk	Plaza	Dallas	Texas	NaN
Lidia	Glendza	lglendzard@techcrunch.com	Female	90.0	Golf Course	Circle	Clearwater	Florida	NaN
Angelica	Ludlamme	aludlammerf@unc.edu	Female	5849.0	Steensland	Park	Saint Petersburg	Florida	NaN

olumns

```python
Customers[Customers['postcode'].isna()].index
```

```
Index([   1,    9,   16,   20,   21,   29,   31,   36,   42,   58,
       ...
        905, 907, 944, 950, 954, 956, 960, 976, 985, 987],
      dtype='object', length=157)
```

```python
# Individual Column Drop [Pass the column header]
Customers.drop(columns = 'CustomerName' , axis = 1 , inplace = True)
Customers
```

treet_num	street_name	street_suffix	city	state	postcode	Country	Membership_status	Membership Status	membershipStatus
1.0	Trailsway	Road	San Diego	California	92127.0	United States	Diamond	Diamond	Gold
NaN	NaN	NaN	El Paso	Texas	NaN	United States	Platinum	Platinum	Gold
4057.0	Arkansas	Circle	San Antonio	Texas	78220.0	United States	Gold	Gold	Gold
74.0	Debs	Point	Memphis	Tennessee	38150.0	United States	Silver	Silver	Silver
NaN	2nd	Drive	NaN	NaN	83732.0	United States	Bronze	Bronze	Silver
...	...	...	...	...	...	...	...	...	...
0.0	Thompson	Way	Reading	Pennsylvania	19610.0	United States	nan	Diamond	Silver
701.0	Rowland	Hill	Indianapolis	Indiana	46278.0	United States	nan	Platinum	Silver
2.0	Basil	Road	Waterbury	Connecticut	6721.0	United States	nan	Gold	Silver
NaN	Kenwood	Drive	Savannah	Georgia	31405.0	United States	nan	Silver	Gold
205.0	Farwell	Park	Atlanta	Georgia	30386.0	United States	nan	Bronze	Gold

```python
dropped_columns = ['Country','Membership_status','Membership Status','membershipStatus']
Customers.drop(columns = dropped_columns, axis = 1, inplace = True)
Customers
```

name	last_name	email	gender	street_num	street_name	street_suffix	city	state	postcode
omain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road	San Diego	California	92127.0
osimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN	El Paso	Texas	NaN
Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle	San Antonio	Texas	78220.0
oarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point	Memphis	Tennessee	38150.0
ikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive	NaN	NaN	83732.0
...	...	...	...	...	...	...	...	...	...
Merrili	Alman	malmanrn@cornell.edu	Female	0.0	Thompson	Way	Reading	Pennsylvania	19610.0
nonah	Heckle	whecklero@fc2.com	Female	701.0	Rowland	Hill	Indianapolis	Indiana	46278.0
Tobit	Birt	NaN	Male	2.0	Basil	Road	Waterbury	Connecticut	6721.0
Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive	Savannah	Georgia	31405.0
lmore	Malpas	NaN	Male	205.0	Farwell	Park	Atlanta	Georgia	30386.0

```python
# Renaming the rows index
index_rename = {
    0:'FirstCustomer',
    1:'SecondCustomer',
    2:'ThirdCustomer',
    3:'FourthCustomer',
    4:'FifthCustomer',
}
# Customers.rename(index = index_rename , inplace = True)
Customers = Customers.rename(index = index_rename)
Customers
```

	id	first_name	last_name	email	gender	street_num	street_name	street_suffix
FirstCustomer	1.0	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road
SecondCustomer	2.0	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN
ThirdCustomer	3.0	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle
FourthCustomer	4.0	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point
FifthCustomer	5.0	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive
...	...	...	...	...	...	...	...	...
995	996.0	Merrili	Alman	malmanrn@cornell.edu	Female	0.0	Thompson	Way
996	997.0	Winonah	Heckle	whecklero@fc2.com	Female	701.0	Rowland	Hill
997	998.0	Tobit	Birt	NaN	Male	2.0	Basil	Road
998	999.0	Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive
999	1000.0	Elmore	Malpas	NaN	Male	205.0	Farwell	Park

1000 rows × 11 columns

```python
# Renaming the Columns headers
column_rename = {
    'street_num':'street_number',
    'street_name':'street_address'
}
# Customers.rename(columns = column_rename , inplace = True)
Customers = Customers.rename(columns = column_rename)
Customers
```

	id	first_name	last_name	email	gender	street_number	street_address	street_suffix
FirstCustomer	1.0	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road
SecondCustomer	2.0	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN
ThirdCustomer	3.0	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle
FourthCustomer	4.0	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point
FifthCustomer	5.0	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive
...	...	...	...	...	...	...	...	...
995	996.0	Merrili	Alman	malmanrn@cornell.edu	Female	0.0	Thompson	Way
996	997.0	Winonah	Heckle	whecklero@fc2.com	Female	701.0	Rowland	Hill
997	998.0	Tobit	Birt	NaN	Male	2.0	Basil	Road
998	999.0	Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive
999	1000.0	Elmore	Malpas	NaN	Male	205.0	Farwell	Park

1000 rows × 11 columns

```python
# SELECT DISTINCT STATE FROM Customers -> Unique Values
Customers['state'].shape # Series
```

```
(1000,)
```

```python
Customers['state'].ndim
```

```
1
```

```python
Customers['state'].unique()
```

```
array(['California', 'Texas', 'Tennessee', nan, 'Maryland', 'Illinois',
       'Missouri', 'Alaska', 'Louisiana', 'Florida', 'New York',
       'District of Columbia', 'South Carolina', 'Washington', 'Nevada',
       'Indiana', 'Pennsylvania', 'Georgia', 'Virginia', 'Michigan',
       'Iowa', 'Kentucky', 'Oregon', 'North Carolina', 'Ohio', 'Kansas',
       'West Virginia', 'Oklahoma', 'North Dakota', 'Arizona', 'Utah',
       'Minnesota', 'Wisconsin', 'New Jersey', 'Montana', 'New Hampshire',
       'Colorado', 'Massachusetts', 'Alabama', 'Delaware', 'Mississippi',
       'Idaho', 'Arkansas', 'Hawaii', 'Nebraska', 'Connecticut',
       'Rhode Island', 'New Mexico', 'Maine'], dtype=object)
```

```python
Customers['state'].nunique() # Count -> Excluding NaN
```

```
48
```

```python
Customers['gender'].unique()
```

```
array(['Male', 'Female', nan], dtype=object)
```

```python
Customers['gender'].nunique() # Count -> Excluding NaN
```

```
2
```

```
# .value_counts
Products['company'].value_counts().head()

company
Vitz          2
Ntag          2
Livepath      2
Cogibox       2
Browsezoom    2
Name: count, dtype: int64
```

```
# TypeCasting .astype()
Customers.info()

<class 'pandas.core.frame.DataFrame'>
Index: 1000 entries, FirstCustomer to 999
Data columns (total 11 columns):
 #   Column          Non-Null Count   Dtype
---  ------          --------------   -----
 0   id              1000 non-null    float64
 1   first_name      1000 non-null    object
 2   last_name       1000 non-null    object
 3   email           878 non-null     object
 4   gender          957 non-null     object
 5   street_number   738 non-null     float64
 6   street_address  963 non-null     object
 7   street_suffix   963 non-null     object
 8   city            921 non-null     object
 9   state           920 non-null     object
 10  postcode        1000 non-null    int32
dtypes: float64(2), int32(1), object(8)
memory usage: 122.1+ KB
```

```
Customers['postcode'] = Customers['postcode'].fillna(0).astype(int)
Customers
```

	id	first_name	last_name	email	gender	street_number	street_address	street_suffix
FirstCustomer	1.0	Romain	Southcott	rsouthcott0@clickbank.net	Male	1.0	Trailsway	Road
SecondCustomer	2.0	Cosimo	Molyneaux	cmolyneaux1@wiley.com	Male	NaN	NaN	NaN
ThirdCustomer	3.0	Bambi	Westrip	bwestrip2@symantec.com	Female	4057.0	Arkansas	Circle
FourthCustomer	4.0	Roarke	Pankettman	rpankettman3@wiley.com	Male	74.0	Debs	Point
FifthCustomer	5.0	Mikaela	Althorpe	malthorpe4@51.la	NaN	NaN	2nd	Drive
...	...	...	...	...	...	...	...	...
995	996.0	Merrili	Alman	malmanrn@cornell.edu	Female	0.0	Thompson	Way
996	997.0	Winonah	Heckle	whecklero@fc2.com	Female	701.0	Rowland	Hill
997	998.0	Tobit	Birt	NaN	Male	2.0	Basil	Road
998	999.0	Issiah	Standbrooke	istandbrookerq@yellowpages.com	Male	NaN	Kenwood	Drive
999	1000.0	Elmore	Malpas	NaN	Male	205.0	Farwell	Park

1000 rows × 11 columns