

```
import pandas as pd
```

```
import pandas as pd
```

```
df=pd.read_excel('/content/mega_data_cleaning_dataset.xlsx')
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_S |
|-------------------------|-------------|------------------|-----|--------|-------------------------|--------------|--------------------|------------------------------------|-------------|-----------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ir |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 9995 | 8287 | Deborah Garcia | 11 | Mle | Cape Verde | Maine | Lake Albert | timothy71@hotmail.com | 2025-02-18 | Ir |
| 9996 | 8932 | Donald Pineda | 69 | Female | NaN | Pennsylvania | Raymondbury | gsmith@hotmail.com | 2022-07-03 | |
| 9997 | 6328 | Alexandria Moore | -5 | Other | Cocos (Keeling) Islands | Pennsylvania | Port Frank | morgansamantha@hotmail.com | 2023-03-01 | B |
| 9998 | 8121 | NaN | 110 | Male | Equatorial Guinea | Maine | Jacquelinefort | courtneybradley@day-gilbert.com | 2022-11-20 | Ir |
| 9999 | 2851 | Kathleen Leblanc | 67 | Mle | Albania | Virginia | South Michelle | kristincook@thompson-campbell.info | 2024-04-01 | Ir |
| 10000 rows × 25 columns | | | | | | | | | | |

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

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_S |
|------|-------------|------------------|-----|--------|-------------------------|--------------|--------------------|------------------------------------|-------------|-----------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ir |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 9995 | 8287 | Deborah Garcia | 11 | Mle | Cape Verde | Maine | Lake Albert | timothy71@hotmail.com | 2025-02-18 | Ir |
| 9996 | 8932 | Donald Pineda | 69 | Female | NaN | Pennsylvania | Raymondbury | gsmith@hotmail.com | 2022-07-03 | |
| 9997 | 6328 | Alexandria Moore | -5 | Other | Cocos (Keeling) Islands | Pennsylvania | Port Frank | morgansamantha@hotmail.com | 2023-03-01 | B |
| 9998 | 8121 | NaN | 110 | Male | Equatorial Guinea | Maine | Jacquelinefort | courtneybradley@day-gilbert.com | 2022-11-20 | Ir |
| 9999 | 2851 | Kathleen Leblanc | 67 | Mle | Albania | Virginia | South Michelle | kristincook@thompson-campbell.info | 2024-04-01 | Ir |

10000 rows × 25 columns

```
df=df.drop_duplicates(subset='Customer_ID', keep='first')
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Status |
|------|-------------|------------------|-----|--------|-------------------|-----------|--------------------|------------------------------------|-------------|----------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inactive |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | NaN |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Active |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Active |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | NaN |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 9991 | 2919 | Tyler Hunter | 44 | Female | Guam | Indiana | Tylerville | ashley87@yahoo.com | 2021-10-23 | NaN |
| 9993 | 8538 | Harry Carter | 108 | Female | Tokelau | Texas | West Annaberg | uhayes@clark-webb.com | 2023-09-30 | Active |
| 9994 | 1905 | Deborah Young | 128 | f | Tuvalu | Tennessee | East Anthonytown | james28@yahoo.com | 2020-05-04 | NaN |
| 9998 | 8121 | NaN | 110 | Male | Equatorial Guinea | Maine | Jacquelinefort | courtneybradley@day-gilbert.com | 2022-11-20 | Inactive |
| 9999 | 2851 | Kathleen Leblanc | 67 | Mle | Albania | Virginia | South Michelle | kristincook@thompson-campbell.info | 2024-04-01 | Inactive |

6172 rows × 25 columns

```
df = df.drop_duplicates(subset='Email', keep='first')
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Status |
|------|-------------|------------------|-----|--------|-------------------|-----------|--------------------|------------------------------------|-------------|----------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inactive |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | NaN |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Active |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Active |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | NaN |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 9991 | 2919 | Tyler Hunter | 44 | Female | Guam | Indiana | Tylerville | ashley87@yahoo.com | 2021-10-23 | NaN |
| 9993 | 8538 | Harry Carter | 108 | Female | Tokelau | Texas | West Annaberg | uhayes@clark-webb.com | 2023-09-30 | Active |
| 9994 | 1905 | Deborah Young | 128 | f | Tuvalu | Tennessee | East Anthonytown | james28@yahoo.com | 2020-05-04 | NaN |
| 9998 | 8121 | NaN | 110 | Male | Equatorial Guinea | Maine | Jacquelinefort | courtneybradley@day-gilbert.com | 2022-11-20 | Inactive |
| 9999 | 2851 | Kathleen Leblanc | 67 | Mle | Albania | Virginia | South Michelle | kristincook@thompson-campbell.info | 2024-04-01 | Inactive |

5824 rows × 25 columns

```
df = df.drop_duplicates(subset='Product_ID', keep='first')
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Status |
|------|-------------|------------------|-----|--------|----------------------|----------------|--------------------|------------------------------------|-------------|----------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inactiv |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | Na |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Activ |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Activ |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | Na |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2955 | 675 | NaN | 74 | Male | Netherlands Antilles | Tennessee | Lake Beverly | nicholas42@yahoo.com | 2025-09-17 | Activ |
| 3032 | 1437 | Paula Wheeler | 111 | Mle | Greenland | Connecticut | Port Emilyton | hmunoz@kent.com | 2024-09-14 | Na |
| 3034 | 3786 | Karen Miller | 22 | f | France | NaN | West Brenda | zsmall@gmail.com | 2023-05-20 | Na |
| 3039 | 201 | Kirk Clayton PhD | 80 | f | Russian Federation | South Carolina | Andrewton | cbennett@wright.com | 2022-01-24 | Na |
| 3298 | 5796 | Angel Hood | 53 | Mle | NaN | Arkansas | Lake Derek | omar32@yahoo.com | 2021-08-19 | Banne |

500 rows × 25 columns

```
df['Name'] = df['Name'].fillna('Unknown')
df
```

/tmp/ipython-input-2059556363.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
df['Name'] = df['Name'].fillna('Unknown')

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Stat |
|------|-------------|------------------|-----|--------|----------------------|----------------|--------------------|------------------------------------|-------------|--------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inact |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | N |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Act |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Act |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | N |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2955 | 675 | Unknown | 74 | Male | Netherlands Antilles | Tennessee | Lake Beverly | nicholas42@yahoo.com | 2025-09-17 | Act |
| 3032 | 1437 | Paula Wheeler | 111 | Mle | Greenland | Connecticut | Port Emilyton | hmunoz@kent.com | 2024-09-14 | N |
| 3034 | 3786 | Karen Miller | 22 | f | France | NaN | West Brenda | zsmall@gmail.com | 2023-05-20 | N |
| 3039 | 201 | Kirk Clayton PhD | 80 | f | Russian Federation | South Carolina | Andrewton | cbennett@wright.com | 2022-01-24 | N |
| 3298 | 5796 | Angel Hood | 53 | Mle | NaN | Arkansas | Lake Derek | omar32@yahoo.com | 2021-08-19 | Banr |

```
df['Email'].fillna('noemail@example.com')
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Stat |
|------|-------------|------------------|-----|--------|----------------------|----------------|--------------------|------------------------------------|-------------|--------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inact |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | N |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Act |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Act |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | N |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2955 | 675 | Unknown | 74 | Male | Netherlands Antilles | Tennessee | Lake Beverly | nicholas42@yahoo.com | 2025-09-17 | Act |
| 3032 | 1437 | Paula Wheeler | 111 | Mle | Greenland | Connecticut | Port Emilyton | hmunoz@kent.com | 2024-09-14 | N |
| 3034 | 3786 | Karen Miller | 22 | f | France | NaN | West Brenda | zsmall@gmail.com | 2023-05-20 | N |
| 3039 | 201 | Kirk Clayton PhD | 80 | f | Russian Federation | South Carolina | Andrewton | cbennett@wright.com | 2022-01-24 | N |
| 3298 | 5796 | Angel Hood | 53 | Mle | NaN | Arkansas | Lake Derek | omar32@yahoo.com | 2021-08-19 | Banr |

500 rows × 25 columns

```
df['Product_Name'] = df['Product_Name'].fillna('Misc')
df
```

/tmp/ipython-input-1685545772.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
df['Product_Name'] = df['Product_Name'].fillna('Misc')

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Stat |
|------|-------------|------------------|-----|--------|----------------------|----------------|--------------------|------------------------------------|-------------|--------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inact |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | N |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Act |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Act |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | N |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2955 | 675 | Unknown | 74 | Male | Netherlands Antilles | Tennessee | Lake Beverly | nicholas42@yahoo.com | 2025-09-17 | Act |
| 3032 | 1437 | Paula Wheeler | 111 | Mle | Greenland | Connecticut | Port Emilyton | hmunoz@kent.com | 2024-09-14 | N |
| 3034 | 3786 | Karen Miller | 22 | f | France | NaN | West Brenda | zsmall@gmail.com | 2023-05-20 | N |
| 3039 | 201 | Kirk Clayton PhD | 80 | f | Russian Federation | South Carolina | Andrewton | cbennett@wright.com | 2022-01-24 | N |
| 3298 | 5796 | Angel Hood | 53 | Mle | NaN | Arkansas | Lake Derek | omar32@yahoo.com | 2021-08-19 | Banr |

df['Category'] = df['Category'].fillna('Misc')
df

/tmp/ipython-input-1577046602.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
df['Category'] = df['Category'].fillna('Misc')

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Stat |
|------|-------------|------------------|-----|--------|----------------------|----------------|--------------------|------------------------------------|-------------|--------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inact |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | N |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Act |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Act |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | N |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2955 | 675 | Unknown | 74 | Male | Netherlands Antilles | Tennessee | Lake Beverly | nicholas42@yahoo.com | 2025-09-17 | Act |
| 3032 | 1437 | Paula Wheeler | 111 | Mle | Greenland | Connecticut | Port Emilyton | hmunoz@kent.com | 2024-09-14 | N |
| 3034 | 3786 | Karen Miller | 22 | f | France | NaN | West Brenda | zsmall@gmail.com | 2023-05-20 | N |
| 3039 | 201 | Kirk Clayton PhD | 80 | f | Russian Federation | South Carolina | Andrewton | cbennett@wright.com | 2022-01-24 | N |
| 3298 | 5796 | Angel Hood | 53 | Mle | NaN | Arkansas | Lake Derek | omar32@yahoo.com | 2021-08-19 | Banr |

df['Feedback_Score'] = df.groupby('Product_ID')['Feedback_Score'].transform(lambda x: x.fillna(x.median()))
df

```
/tmp/ipython-input-2618572678.py:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy  
df['Feedback_Score'] = df.groupby('Product_ID')['Feedback_Score'].transform(lambda x: x.fillna(x.median()))
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Stat |
|------|-------------|------------------|-----|--------|----------------------|----------------|--------------------|------------------------------------|-------------|--------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inact |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | N |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Act |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Act |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | N |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2955 | 675 | Unknown | 74 | Male | Netherlands Antilles | Tennessee | Lake Beverly | nicholas42@yahoo.com | 2025-09-17 | Act |
| 3032 | 1437 | Paula Wheeler | 111 | Mle | Greenland | Connecticut | Port Emilyton | hmunoz@kent.com | 2024-09-14 | N |
| 3034 | 3786 | Karen Miller | 22 | f | France | NaN | West Brenda | zsmall@gmail.com | 2023-05-20 | N |
| 3039 | 201 | Kirk Clayton PhD | 80 | f | Russian Federation | South Carolina | Andrewton | cbennett@wright.com | 2022-01-24 | N |
| 3298 | 5796 | Angel Hood | 53 | Mle | NaN | Arkansas | Lake Derek | omar32@yahoo.com | 2021-08-19 | Banr |

```
df['Quantity'] = df['Quantity'].fillna(1)  
df
```

/tmp/ipython-input-54306166.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
df['Quantity'] = df['Quantity'].fillna(1)

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Stat |
|------|-------------|------------------|-----|--------|----------------------|----------------|--------------------|------------------------------------|-------------|--------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inact |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | N |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Act |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Act |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | N |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2955 | 675 | Unknown | 74 | Male | Netherlands Antilles | Tennessee | Lake Beverly | nicholas42@yahoo.com | 2025-09-17 | Act |
| 3032 | 1437 | Paula Wheeler | 111 | Mle | Greenland | Connecticut | Port Emilyton | hmunoz@kent.com | 2024-09-14 | N |
| 3034 | 3786 | Karen Miller | 22 | f | France | NaN | West Brenda | zsmall@gmail.com | 2023-05-20 | N |
| 3039 | 201 | Kirk Clayton PhD | 80 | f | Russian Federation | South Carolina | Andrewton | cbennett@wright.com | 2022-01-24 | N |
| 3298 | 5796 | Angel Hood | 53 | Mle | NaN | Arkansas | Lake Derek | omar32@yahoo.com | 2021-08-19 | Banr |

```
df = df[df['Price'] >= 0]  
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Stat |
|------|-------------|------------------|-----|--------|----------------------|----------------|--------------------|------------------------------------|-------------|--------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inact |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | N |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Act |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Act |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | N |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2955 | 675 | Unknown | 74 | Male | Netherlands Antilles | Tennessee | Lake Beverly | nicholas42@yahoo.com | 2025-09-17 | Act |
| 3032 | 1437 | Paula Wheeler | 111 | Mle | Greenland | Connecticut | Port Emilyton | hmunoz@kent.com | 2024-09-14 | N |
| 3034 | 3786 | Karen Miller | 22 | f | France | NaN | West Brenda | zsmall@gmail.com | 2023-05-20 | N |
| 3039 | 201 | Kirk Clayton PhD | 80 | f | Russian Federation | South Carolina | Andrewton | cbennett@wright.com | 2022-01-24 | N |
| 3298 | 5796 | Angel Hood | 53 | Mle | NaN | Arkansas | Lake Derek | omar32@yahoo.com | 2021-08-19 | Banr |

475 rows × 25 columns

```
df = df[df['Price'] > 0]
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Stat |
|------|-------------|------------------|-----|--------|----------------------|----------------|--------------------|------------------------------------|-------------|--------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inact |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | N |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Act |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Act |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | N |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2955 | 675 | Unknown | 74 | Male | Netherlands Antilles | Tennessee | Lake Beverly | nicholas42@yahoo.com | 2025-09-17 | Act |
| 3032 | 1437 | Paula Wheeler | 111 | Mle | Greenland | Connecticut | Port Emilyton | hmunoz@kent.com | 2024-09-14 | N |
| 3034 | 3786 | Karen Miller | 22 | f | France | NaN | West Brenda | zsmall@gmail.com | 2023-05-20 | N |
| 3039 | 201 | Kirk Clayton PhD | 80 | f | Russian Federation | South Carolina | Andrewton | cbennett@wright.com | 2022-01-24 | N |
| 3298 | 5796 | Angel Hood | 53 | Mle | NaN | Arkansas | Lake Derek | omar32@yahoo.com | 2021-08-19 | Banr |

475 rows × 25 columns

```
df = df[df['Quantity'] >= 0]
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Stat |
|------|-------------|------------------|-----|--------|----------------------|----------------|--------------------|------------------------------------|-------------|--------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inact |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | N |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Act |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Act |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | N |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2955 | 675 | Unknown | 74 | Male | Netherlands Antilles | Tennessee | Lake Beverly | nicholas42@yahoo.com | 2025-09-17 | Act |
| 3032 | 1437 | Paula Wheeler | 111 | Mle | Greenland | Connecticut | Port Emilyton | hmunoz@kent.com | 2024-09-14 | N |
| 3034 | 3786 | Karen Miller | 22 | f | France | NaN | West Brenda | zsmall@gmail.com | 2023-05-20 | N |
| 3039 | 201 | Kirk Clayton PhD | 80 | f | Russian Federation | South Carolina | Andrewton | cbennett@wright.com | 2022-01-24 | N |
| 3298 | 5796 | Angel Hood | 53 | Mle | NaN | Arkansas | Lake Derek | omar32@yahoo.com | 2021-08-19 | Banr |

391 rows × 25 columns

```
df = df[(df['Age'] >= 0) & (df['Age'] <= 120)]
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Stat |
|------|-------------|------------------|-----|--------|----------------------|----------------|--------------------|------------------------------------|-------------|--------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inact |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | N |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Act |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Act |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | N |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2955 | 675 | Unknown | 74 | Male | Netherlands Antilles | Tennessee | Lake Beverly | nicholas42@yahoo.com | 2025-09-17 | Act |
| 3032 | 1437 | Paula Wheeler | 111 | Mle | Greenland | Connecticut | Port Emilyton | hmunoz@kent.com | 2024-09-14 | N |
| 3034 | 3786 | Karen Miller | 22 | f | France | NaN | West Brenda | zsmall@gmail.com | 2023-05-20 | N |
| 3039 | 201 | Kirk Clayton PhD | 80 | f | Russian Federation | South Carolina | Andrewton | cbennett@wright.com | 2022-01-24 | N |
| 3298 | 5796 | Angel Hood | 53 | Mle | NaN | Arkansas | Lake Derek | omar32@yahoo.com | 2021-08-19 | Banr |

333 rows × 25 columns

```
df = df[(df['Feedback_Score'] >= 1) & (df['Feedback_Score'] <= 5)]
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Stat |
|-----------------------|-------------|------------------|-----|--------|----------------------|----------------|--------------------|------------------------------------|-------------|--------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inact |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | N |
| 2 | 6574 | Daniel Cortez | 105 | Mle | Tajikistan | Montana | West Jennifermouth | hoffmancole@yahoo.com | 2025-05-19 | Act |
| 3 | 3377 | Cindy Barber | 100 | Female | Chile | Ohio | Staceyberg | carpenterlinda@yahoo.com | 2024-08-19 | Act |
| 4 | 5107 | William Patel | 110 | Mle | Timor-Leste | Hawaii | South Hollyfort | blairivan@lyons-christensen.org | 2020-05-04 | N |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2657 | 1802 | Brian Simon | 114 | Mle | Latvia | Tennessee | New Cynthia | uwalker@graham.org | 2021-04-27 | Banr |
| 2710 | 1332 | Courtney Jenkins | 99 | f | Mauritania | Alabama | Gonzalezton | jenniferreid@hotmail.com | 2025-10-13 | Act |
| 2955 | 675 | Unknown | 74 | Male | Netherlands Antilles | Tennessee | Lake Beverly | nicholas42@yahoo.com | 2025-09-17 | Act |
| 3032 | 1437 | Paula Wheeler | 111 | Mle | Greenland | Connecticut | Port Emilyton | hmunoz@kent.com | 2024-09-14 | N |
| 3039 | 201 | Kirk Clayton PhD | 80 | f | Russian Federation | South Carolina | Andrewton | cbennett@wright.com | 2022-01-24 | N |
| 233 rows × 25 columns | | | | | | | | | | |

```
df = df[df['Delivery_Date'] >= df['Purchase_Date']]
df
```


| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Mle | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Mle | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Mle | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Bar |

111 rows × 25 columns

```
df['Gender'] = df['Gender'].replace({'Mle': 'Male', 'm': 'Male', 'f': 'Female', 'Fmale': 'Female'})
df
```

```
/tmp/ipython-input-1943499802.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy
df['Gender'] = df['Gender'].replace({'Mle': 'Male', 'm': 'Male', 'f': 'Female', 'Fmale': 'Female'})
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |

```
df['Payment_Type'] = df['Payment_Type'].str.strip().str.title()
df
```

```
/tmp/ipython-input-359501598.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
df['Payment_Type'] = df['Payment_Type'].str.strip().str.title()
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Status |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|----------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inactive |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | Active |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthacheater | yvonne52@hall.com | 2020-08-02 | Active |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | Active |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | Active |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Inactive |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Banned |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | Active |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Inactive |

```
df['Account_Status'] = df['Account_Status'].str.capitalize().replace({'Active': 'Active', 'Inactive': 'Inactive', 'Banned': 'Banned'})
df
```

```
/tmp/ipython-input-2896010401.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy
df['Account_Status'] = df['Account_Status'].str.capitalize().replace({'Active':'Active', 'Inactiv':'Inactive', 'Baned':'Banned'})
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |

```
df = df[df['Email'].str.contains(r'^[w\.-]+@[w\.-]+\.\w+$', na=False)]
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Bar |

111 rows × 25 columns

```
df[df['Price'] > 500]
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Bar |

111 rows × 25 columns

```
df[df['Feedback_Score'] < 3]
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Bar |

111 rows × 25 columns

```
df[(df['Customer_Loyalty']=='VIP') & (df['Total_Purchase']>1000)]
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Bar |

111 rows × 25 columns

```
df[df['Account_Status']=='Banned']
df
```


| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Bar |

111 rows × 25 columns

```
df[(df['Delivery_Date'] - df['Purchase_Date']).dt.days > 7]
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Bar |

111 rows × 25 columns

```
df[df['Category']=='Electronics']
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Bar |

111 rows × 25 columns

```
df[df['Country']=='India']
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Bar |

111 rows × 25 columns

```
df[df['Quantity'] > 10]
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Bar |

111 rows × 25 columns

```
df[df['Discount'] > 50]
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Bar |

111 rows × 25 columns

```
df[df['Price']*df['Quantity'] > 1000]
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | A |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Ina |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Bar |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Ina |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Bar |

111 rows × 25 columns

```
df.groupby('Customer_ID').apply(lambda x:(x['Price']*x['Quantity']).sum())
```

```
/tmp/ipython-input-47298081.py:1: DeprecationWarning: DataFrameGroupBy.apply operated on the grouping columns. This behavior is
df.groupby('Customer_ID').apply(lambda x:(x['Price']*x['Quantity']).sum())
```

0

| Customer_ID | |
|-------------|----------|
| 116 | 3578.55 |
| 211 | 4000.36 |
| 334 | 10235.68 |
| 415 | 5209.60 |
| 454 | 7222.65 |
| ... | ... |
| 9025 | 6468.63 |
| 9044 | 1001.52 |
| 9347 | 1797.08 |
| 9425 | 4421.12 |
| 9437 | 9585.50 |

111 rows × 1 columns

dtype: float64

```
df.groupby('Product_ID')['Feedback_Score'].mean()
```

| Feedback_Score | |
|----------------|-----|
| Product_ID | |
| 3 | 2.2 |
| 4 | 4.8 |
| 8 | 3.1 |
| 9 | 1.7 |
| 10 | 1.4 |
| ... | ... |
| 481 | 1.4 |
| 483 | 1.0 |
| 485 | 1.0 |
| 486 | 3.7 |
| 493 | 1.1 |

111 rows × 1 columns

dtype: float64

```
df.groupby('Customer_ID')['Category'].apply(lambda x:set(x)=='Electronics'))
```


| Category | |
|-------------|-------|
| Customer_ID | |
| 116 | False |
| 211 | False |
| 334 | True |
| 415 | True |
| 454 | False |
| ... | ... |
| 9025 | True |
| 9044 | True |
| 9347 | True |
| 9425 | True |
| 9437 | False |

111 rows × 1 columns

dtype: bool

```
df[df['Customer_Loyalty']=='VIP'].groupby('Customer_ID').agg({'Total_Purchase':'sum'}).nlargest(10,'Total_Purchase')
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Status |
|------|-------------|----------------|-----|--------|----------------------------------------------|-------------|-------------------|------------------------------------|-------------|----------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Inactive |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | Active |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | Active |
| 11 | 7739 | Tracy Reed | 109 | Male | Hong Kong | Mississippi | New Jamieland | erica26@yahoo.com | 2021-09-03 | Active |
| 13 | 211 | Lynn Garcia | 15 | Male | Rwanda | Nebraska | Lake Michael | ocruz@yahoo.com | 2020-08-19 | Active |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 1756 | 5178 | Robert Price | 113 | Other | Isle of Man | Colorado | Robertsberg | englishdiane@guzman-schneider.com | 2021-04-30 | Inactive |
| 1950 | 2268 | Derek Martinez | 116 | Male | Antarctica (the territory South of 60 deg S) | Missouri | South Douglasbury | timothypeters@yahoo.com | 2020-07-26 | Barred |
| 2175 | 3253 | Erin Goodwin | 58 | Male | Antigua and Barbuda | New Mexico | Kimstad | brose@chavez.com | 2023-12-15 | Active |
| 2240 | 8324 | Michael Lee | 110 | Female | Afghanistan | California | Robinberg | mariamoyer@yahoo.com | 2020-01-08 | Inactive |
| 2417 | 3364 | Timothy Miller | 9 | Male | Germany | Arkansas | Morganhaven | bthomas@evans.com | 2023-05-13 | Barred |

111 rows × 25 columns

```
df[df['Customer_ID'].isna()]
```

| Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Status | ... | Quantity | Purchase_Date | Delivery |
|---------------------|------|-----|--------|---------|-------|------|-------|-------------|----------------|-----|----------|---------------|----------|
| 0 rows × 25 columns | | | | | | | | | | | | | |

```
df['Total_Purchase'] = df['Price'] * df['Quantity']
df
```

| | Customer_ID | Name | Age | Gender | Country | State | City | Email | Signup_Date | Account_Sta |
|----|-------------|---------------|-----|--------|---------|----------|-----------------|------------------------------------|-------------|-------------|
| 0 | 1191 | Larry Wallace | 31 | Other | NaN | Arkansas | Janetfort | gonzalezbrandi@shaffer-lambert.com | 2021-05-23 | Ina |
| 1 | 8287 | Shawn Palmer | 20 | Male | Uganda | Texas | Smithhaven | davidmendoza@hotmail.com | 2025-05-18 | |
| 10 | 1379 | Louis Cuevas | 83 | Male | Germany | Idaho | Samanthachester | yvonne52@hall.com | 2020-08-02 | |
| | | Tracy | | | | | | | | |