

```
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.drop_duplicates()
```

	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

```
df.drop_duplicates(subset=['Customer_ID'], keep='first',inplace=True)
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
9998	8287	Shawn Palmer	20	Male	Equatorial Guinea	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
10000 rows × 25 columns										
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.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	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

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

```
df['Email'] = df['Email'].fillna('noemail@example.com')
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
3	3377	Cindy Shaw	100	Female	Chile	Ohio	Staceyberg	carpenterlinda@yahoo.com	2024-08-19	Active
1	8287	Barber Palmer	20	Male	Uganda	Texas	Smithhaven	davidmendoza@hotmail.com	2025-05-18	NaN
4	5107	William	110	Mle	Timor-	Hawaii	South	blairivan@lyons-	2020-05-04	NaN
2	6574	Daniel Cortez	105	Mle	Tajikistan	Montana	West Jennifermouth	hoffmancole@yahoo.com	2025-05-19	Active
9991	2919	Tyler Hunter	44	Female	Guam	Indiana	Tylerville	ashley87@yahoo.com	2021-10-23	NaN
3	3377	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
9994	1905	Deborah Young	128	f	Tuvalu	Tennessee	East Anthonvtown	james28@yahoo.com	2020-05-04	NaN
9991	2919	Tyler Hunter	44	Female	Guam	Indiana	Tylerville	ashley87@yahoo.com	2021-10-23	NaN
					Guinea			gilbert.com		
9993	8538	Harry Carter	108	Female	Tokelau	Texas	West Annaberg	uhayes@clark-webb.com	2023-09-30	Active
9999	2851	Kathleen Leblanc	67	Mle	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
9994	1905	Deborah Young	128	f	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	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['Product_Name'] = df['Product_Name'].fillna('Misc')
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

```
df['Category'] = df['Category'].fillna('Misc')
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
9993	8538	Harry Carter	108	Female	Tokelau	Texas	West Annaberg	uhayes@clark-webb.com	2023-09-30	Active
1	8287	Shawn Palmer	20	Male	Uganda	Texas	Smithhaven	davidmendoza@hotmail.com	2025-05-18	NaN
9994	1905	Deborah Young	128	f	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
2	6574	Daniel Cortez	105	Mle	Tajikistan	Montana	West Jennifermouth	hoffmancole@yahoo.com	2025-05-19	Active
9999	2851	Kathleen Leblanc	67	Mle	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
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	Unknown	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['Feedback_Score'] = df.groupby('Product_ID')['Feedback_Score'].transform(lambda x: x.fillna(x.median()))
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

```
df['Quantity'] = df['Quantity'].fillna(1)
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	Unknown	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['Price'] >= 0]
```

	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	Unknown	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

5875 rows × 25 columns

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

	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

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

	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	8338	Cindy Barber	108	Female	Tokelau	Texas	Staceyberg Annaberg	carpenter@clark-webb.com	2024-09-30	Active
4	5107	William Patel	110	Mle	Timor-Leste	Hawaii	South Hollyfort	blairivan@lyons-christensen.org	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9991	2919	Tyler Hunter	44	Female	Guam	Indiana	Tylerville	ashley87@yahoo.com	2021-10-23	NaN
9999	2851	Kathleen Leblanc	67	Mle	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
9993	8538	Harry Carter	108	Female	Tokelau	Texas	Staceyberg Annaberg	uhayes@clark-webb.com	2023-09-30	Active
5875 rows × 25 columns										
9994	1905	Deborah Young	128	f	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	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
4755 rows × 25 columns										

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

	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
...

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

9990	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
9993	8538	Harry Carter	108	Female	Tokelau	Texas	West Annaberg	uhayes@clark-webb.com	2023-09-30	Active
1	8287	Shawn Palmer	20	Male	Uganda	Texas	Smithhaven	davidmendoza@hotmail.com	2025-05-18	NaN
					Equatorial			courtnevhradlev@da...		
2	6574	Daniel Cortez	105	Mle	Tajikistan	Montana	West Jennifermouth	hoffmancole@yahoo.com	2025-05-19	Active
9999	2851	Kathleen LeGlad	67	Mle	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactiv
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
...
9987	2035	Michael Lee	0	Other	Korea	NaN	East Mary	alexanderryan@gmail.com	2023-05-20	Bannec
9990	4038	Cassandra Collins	80	f	Chile	Mississippi	West Mark	stanleysarah@harrington.org	2020-02-24	Bannec
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
4041 rows × 25 columns										

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

	Customer_ID	Name	Age	Gender	Country	State	City	Email	Signup_Date	Account_St:
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
...
9983	329	Fernando Taylor	106	Female	Congo	Tennessee	Lauraport	maria37@gmail.com	2020-08-18	
9985	2694	Paul Escobar	-4	f	India	Texas	Donaldberg	millerkayla@robinson.net	2025-06-16	A
9987	2035	Michael Lee	0	Other	Korea	NaN	East Mary	alexanderryan@gmail.com	2023-05-20	Bar
9993	8538	Harry Carter	108	Female	Tokelau	Texas	West Annaberg	uhayes@clark-webb.com	2023-09-30	A

Kathleen Kristincook@thompson-

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

2979 rows × 25 columns

/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_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	Male	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	Male	Timor-Leste	Hawaii	South Hollyfort	blairivan@lyons-christensen.org	2020-05-04	NaN
...

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()
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
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
1	8287	Shawn Palmer	20	Male	Uganda	Texas	Smithhaven	davidmendoza@hotmail.com	2025-05-18	NaN
5875 rows × 25 columns										
2	6574	Daniel Cortez	105	Male	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	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
5875 rows × 25 columns										

```
df['Account_Status'] = df['Account_Status'].str.capitalize().replace({'Active':'Active','Inactiv':'Inactive','Baned':'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_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	Male	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	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive

5875 rows × 25 columns

```
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_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	Male	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

df[df['Price'] > 500] 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
9993	8538	Harry Carter	108	Female	Tokelau	Texas	West Annaberg	uhayes@clark-webb.com	2023-09-30	Active
1	8287	Shawn Palmer	20	Male	Uganda	Texas	Smithhaven	davidmendoza@hotmail.com	2025-05-18	NaN
9994	1905	Deborah Young	128	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
2	6574	Daniel Cortez	105	Male	Tajikistan	Montana	West Jennifermouth	hoffmancole@yahoo.com	2025-05-19	Active
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
3	3377	Cindy Barber	100	Female	Chile	Ohio	Staceyberg	carpenterlinda@yahoo.com	2024-08-19	Active
4	5107	William Patel	110	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
5875 rows × 25 columns										

df[df['Feedback_Score'] < 3] 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	Male	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	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN

df[(df['Customer_Loyalty']=='VIP') & (df['Total_Purchase']>1000)]
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
5875 rows × 25 columns										
1	8287	Shawn Palmer	20	Male	Uganda	Texas	Smithhaven	davidmendoza@hotmail.com	2025-05-18	NaN
2	6574	Daniel Cortez	105	Male	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	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
5875 rows × 25 columns										

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

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	Male	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	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive

5875 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_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

```
df[df['Category']=='Electronics']
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
4	5107	Patel	110	Male	Leste	Hawaii	Hollyfort	christensen.org	2020-05-04	NaN
1	8287	Shawn Palmer	20	Male	Uganda	Texas	Smithhaven	davidmendoza@hotmail.com	2025-05-18	NaN
9991	2919	Tyler Hunter	44	Female	Guam	Indiana	Tylerville	ashley87@yahoo.com	2021-10-23	NaN
2	6574	Daniel Cortez	105	Male	Tajikistan	Montana	West Jennifermouth	hoffmancole@yahoo.com	2025-05-19	Active
9994	1905	Deborah Young	128	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
4	5107	William Patel	110	Male	Timor-Leste	Hawaii	South Hollyfort	blairivan@lyons-christensen.org	2020-05-04	NaN
...	...	Kathleen	South	kristincook@thompson-
9991	2919	Tyler Hunter	44	Female	Guam	Indiana	Tylerville	ashley87@yahoo.com	2021-10-23	NaN
5875 rows x 25 columns										
9993	8538	Harry Carter	108	Female	Tokelau	Texas	West Annaberg	uhayes@clark-webb.com	2023-09-30	Active
9994	1905	Deborah Young	128	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
5875 rows x 25 columns										

```
df[df['Country']=='India']
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	Male	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	Male	Timor-Leste	Hawaii	South Hollyfort	blairivan@lyons-christensen.org	2020-05-04	NaN
...

9991 2019 Tyler 44 Female Guam Indiana Tylerville ashley87@yahoo.com 2021-10-23 NaN

df[df['Quantity'] > 10]
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
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
1	8287	Shawn Palmer	20	Male	Uganda	Texas	Smithhaven	davidmendoza@hotmail.com	2025-05-18	NaN
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
2	6574	Daniel Cortez	105	Male	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	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive

5875 rows × 25 columns

df[df['Discount'] > 50]

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

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

	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
3	3377	Cindy Barber	100	Female	Chile	Ohio	Staceyberg	carpenterlinda@yahoo.com	2024-08-19	Active
4	5107	William Patel	110	Male	Timor-Leste	Hawaii	South Hollyfort	blairivan@lyons-christensen.org	2020-05-04	NaN
5	8213	Kristen Harvey	62	Other	Greenland	NaN	South Sheilaborough	jonessusan@gmail.com	2024-12-15	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive

3704 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 deprecated. Use 'groupby(...).apply(...)' instead.
df.groupby('Customer_ID').apply(lambda x:(x['Price']*x['Quantity']).sum())

Customer_ID	
1	3495.80
2	-421.42
3	5907.22
4	2144.52
8	-574.60
...	...
9483	1407.56
9487	1453.60
9493	-932.16
9494	6815.40
9496	7894.88

5875 rows × 1 columns

dtype: float64

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

Feedback_Score	
Product_ID	
1	3.300000
2	2.953333
3	3.113333
4	2.027778
5	3.950000
...	...
496	2.694118
497	3.041667
498	2.181818
499	3.885714
500	3.000000

500 rows × 1 columns

dtype: float64

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

Category	
Customer_ID	
1	False
2	False
3	False
4	False
8	False
...	...
9483	False
9487	True
9493	False
9494	True
9496	False

5875 rows × 1 columns

dtype: bool

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

Total_Purchase	
Customer_ID	
4376	19896.80
2184	19848.60
596	19807.00
7382	19784.80
5505	19403.00
6455	19316.80
1680	19295.60
7643	18901.60
1444	18867.57
5633	18315.00

```
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[df['Product_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_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	Male	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	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
5875 rows × 25 columns										

df['Discounted_Price'] = df['Price'] - df['Discount']
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	Male	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	Male	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

```
df['Delivery_Duration'] = (df['Delivery_Date'] - df['Purchase_Date']).dt.days
df
```

	Customer_ID	Deborah	Age	Gender	Country	State	City	Email	Signup_Date	Account_Status
9994	1905	Deborah Young	128	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
0	1191	Larry Wallace	31	Other	NaN	Arkansas	Janetfort	gonzalezbrandi@shaffer-lambert.com	2021-05-23	Inactive
9999	2887	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
5875 rows x 26 columns										
2	6574	Daniel Cortez	105	Male	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	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive
5875 rows x 27 columns										

```
df['Age_Group'] = pd.cut(df['Age'], [0,12,19,59,120], labels=['Child','Teen','Adult','Senior'])
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	Male	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	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive

```
import numpy as np
df['Loyalty_Level'] = pd.cut(df['Total_Purchase'], [0,500,1000,5000,np.inf], labels=['Bronze','Silver','Gold','Platinum'])
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

```
df['Gender_Code'] = df['Gender'].map({'Male':1,'Female':2,'Other':3})
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
3	3377	Cindy Barber	100	Female	Chile	Ohio	Staceyberg	carpenterlinda@yahoo.com	2024-08-19	Active
2	6574	Daniel Cortez	105	Male	Tajikistan	Montana	West Jennifermouth	hoffmancole@yahoo.com	2025-05-19	Active
4	5107	William Patel	110	Male	Timor-Leste	Hawaii	South Hollyfort	blairivan@lyons-christensen.org	2020-05-04	NaN
3	3377	Cindy Barber	100	Female	Chile	Ohio	Staceyberg	carpenterlinda@yahoo.com	2024-08-19	Active
4	5107	William Patel	110	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive

5875 rows × 31 columns

5875 rows × 29 columns

```
df['Status_Code'] = df['Account_Status'].map({'Active':1,'Inactive':2,'Banned':3})
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	Male	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	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive

5875 rows × 31 columns

```
df['Country'] = df['Country'].str.title().str.strip()
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	Male	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	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive

5875 rows × 31 columns

```
df['Payment_Type'] = df['Payment_Type'].str.title().str.strip()
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	Male	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	Male	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	Female	Tuvalu	Tennessee	East Anthonytown	james28@yahoo.com	2020-05-04	NaN
9998	8121	Unknown	110	Male	Equatorial Guinea	Maine	Jacquelinefort	courtneybradley@day-gilbert.com	2022-11-20	Inactive
9999	2851	Kathleen Leblanc	67	Male	Albania	Virginia	South Michelle	kristincook@thompson-campbell.info	2024-04-01	Inactive

5875 rows × 31 columns

```
df = df.applymap(lambda x: x.strip().lower() if isinstance(x,str) else x)
```

```
/tmp/ipython-input-119833090.py:1: FutureWarning: DataFrame.applymap has been deprecated. Use DataFrame.map instead.
df = df.applymap(lambda x: x.strip().lower() if isinstance(x,str) else x)
```

```
df.groupby('Product_ID').apply(lambda x:(x['Price']*x['Quantity']).sum()).nlargest(5)
```



```
/tmp/ipython-input-4270780365.py:1: DeprecationWarning: DataFrameGroupBy.apply operated on the grouping columns. This behavior is deprecated. Use 'apply()' instead.  
df.groupby('Product_ID').apply(lambda x:(x['Price']*x['Quantity']).sum()).nlargest(5)
```

0

Product_ID

68	130535.02
44	118327.06
33	110736.73
316	108112.65
461	98284.10

dtype: float64

```
df.groupby('Customer_ID')['Total_Purchase'].sum().nlargest(5)
```

Total_Purchase

Customer_ID

7924	19920.0
4376	19896.8
2184	19848.6
5624	19823.0
596	19807.0

dtype: float64

```
df.groupby('Category')['Discount'].mean()
```

Discount

Category

clothing	19.068800
electronics	20.244790
misc	19.758666

dtype: float64

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

Feedback_Score

Category

clothing	3.079373
electronics	3.019122
misc	2.937315

dtype: float64

```
df['Payment_Type'].value_counts().idxmax()
```

'debit'

```
df.loc[(df['Price']*df['Quantity']).idxmax()]
```

6971	
Customer_ID	7924
Name	sydney perkins
Age	44
Gender	female
Country	grenada
State	north dakota
City	ethanberg
Email	aguilarthomas@garcia-calderon.org
Signup_Date	2020-02-18 00:00:00
Account_Status	banned
Product_ID	29
Product_Name	staff
Category	electronics
Price	996.0
Discount	NaN
Quantity	20
Purchase_Date	2025-08-02 00:00:00
Delivery_Date	2025-05-20 00:00:00
Payment_Type	debit
Feedback_Score	1.2
Comments	NaN
Customer_Loyalty	new
Last_Login	2025-10-10 00:00:00
Signup_Channel	app
Total_Purchase	19920.0

df['Customer_ID'].value_counts()[lambda x:x>5]

Delivery_Duration	Count	-74
Age_Group		adult
Customer_ID		
Loyalty_Level		platinum
dtype: int64		
Gender_Code		2

df.sort_values(by='Delivery_Duration', ascending=False)

dtype: object

	Customer_ID	Name	Age	Gender	Country	State	City	Email	Signup_Date	Account_Status
1364	8989	kevin walsh	61	female	burundi	massachusetts	lake allison	xwheeler@yahoo.com	2022-07-30	inactive

```
df.groupby('Product_ID')['Feedback_Score'].agg(['max','min'])
```

		max	min								
	Product_ID	6896	6897	schmidt	democratic republic	bradleychester					
1	6.0	0.2	st	44	male	liberia	washington	brownmouth	kimberly40@yahoo.com	2023-12-15	
2	5.9	0.1	unknown	117	male	ethiopia	georgia	west paul	aliciajohnson@yahoo.com	2023-01-18	
3	5.7	0.8	ackenzie hanson	113	male	switzerland	arkansas	rodriguezville	brenda48@james-clarke.com	2020-10-15	in
4	5.3	0.1									
5	5.8	0.3	erin lewis	56	male	south africa	NaN	north josephchester	leslie26@cook.info	2022-10-01	in
496	5.1	0.3	michael newman	11	male	bahrain	kentucky	south anthonytown	walkeramanda@brown.com	2023-07-01	
497	5.2	0.2	unknown	-1	female	poland	illinois	new keithmouth	jennyferguson@reed.org	2021-10-12	ba
498	5.3	0.3									
42 499	5.32	1.5	sandra evans	10	other	vanuatu	montana	raymondstad	nicholas79@finley.com	2024-06-02	
500	6.0	0.8	sara watkins	62	male	georgia	utah	terryburgh	gstephens@kim.info	2020-02-20	ba
500 rows × 2 columns											
500 rows × 31 columns											

```
df.groupby(['Country','Category'])['Total_Purchase'].sum()
```

Total_Purchase		
Country	Category	
afghanistan	clothing	2381.19
	electronics	16484.31
	misc	69351.28
albania	clothing	33307.00
	electronics	19355.40
...
zambia	electronics	23990.72
	misc	59923.39
zimbabwe	clothing	44268.91
	electronics	51967.59
	misc	39135.66
729 rows × 1 columns		
dtype: float64		