In [8]:
```
import pandas as pd
df = pd.read_csv("flavors.csv")
pd.set_option('display.width', 1000)
#pd.set_option('display.max_columns', None)
print(df)
```

```
                  Flavor Base_Flavor Liked  Flavor_Rating  TextureRating  TotalRating
0     Mint Chocolate Chip     Vanilla   Yes           10.0            8.0         18.0
1              Chocolate   Chocolate   Yes            8.8            7.6         16.6
2                Vanilla     Vanilla    No            4.7            5.0          9.7
3           Cookie Dough     Vanilla   Yes            6.9            6.5         13.4
4             Rocky Road   Chocolate   Yes            8.2            7.0         15.2
5              Pistachio     Vanilla    No            2.3            3.4          5.7
6            Cake Batter     Vanilla   Yes            6.5            6.0         12.5
7              Neapolitan     Vanilla    No            3.8            5.0          8.8
8  Chocolte Fudge Brownie   Chocolate   Yes            8.2            7.1         15.3
```

In [14]:
```
df.groupby('Base_Flavor').mean(numeric_only = True)
```

Out[14]:

Base_Flavor	Flavor_Rating	TextureRating	TotalRating
Chocolate	8.4	7.233333	15.70
Vanilla	5.7	5.650000	11.35

In [15]:
```
df.groupby('Base_Flavor').count()
```

Out[15]:

Base_Flavor	Flavor	Liked	Flavor_Rating	TextureRating	TotalRating
Chocolate	3	3	3	3	3
Vanilla	6	6	6	6	6

In [18]:
```
#df.groupby('Base_Flavor').sum(numberic_only = True)
df.groupby('Base_Flavor').sum(['Flavor Rating','Texture Rating','Total Rating'])
```

Out[18]:

Base_Flavor	Flavor_Rating	TextureRating	TotalRating
Chocolate	25.2	21.7	47.1
Vanilla	34.2	33.9	68.1

In [19]:
```
df.groupby('Base_Flavor').agg({'Flavor_Rating': ['mean', 'max', 'count', 'sum']})
```

Out[19]:

	Flavor_Rating			
	mean	max	count	sum
Base_Flavor				
Chocolate	8.4	8.8	3	25.2
Vanilla	5.7	10.0	6	34.2

In [20]:
```
df.groupby('Base_Flavor').agg({'Flavor_Rating': ['mean', 'max', 'count', 'sum'], 'TextureRating': ['count', 'sum']})
```

Out[20]:

	Flavor_Rating				TextureRating	
	mean	max	count	sum	count	sum
Base_Flavor						
Chocolate	8.4	8.8	3	25.2	3	21.7
Vanilla	5.7	10.0	6	34.2	6	33.9

In [23]:
```
df.groupby(['Base_Flavor', 'Liked']).mean(numeric_only = True)
```

Out[23]:

Base_Flavor	Liked	Flavor_Rating	TextureRating	TotalRating
Chocolate	Yes	8.4	7.233333	15.700000
Vanilla	No	3.6	4.466667	8.066667
	Yes	7.8	6.833333	14.633333

In [24]:
```
df.groupby('Base_Flavor').describe()
```

Out[24]:

	Flavor_Rating								TextureRating				TotalRating							
	count	mean	std	min	25%	50%	75%	max	count	mean	...	75%	max	count	mean	std	min	25%	50%	75%
Base_Flavor																				
Chocolate	3.0	8.4	0.346410	8.2	8.200	8.2	8.5	8.8	3.0	7.233333	...	7.350	7.6	3.0	15.70	0.781025	15.2	15.250	15.3	15.950
Vanilla	6.0	5.7	2.710719	2.3	4.025	5.6	6.8	10.0	6.0	5.650000	...	6.375	8.0	6.0	11.35	4.263684	5.7	9.025	11.1	13.175

2 rows × 24 columns

In [ ]:
```
#MERGE AND JOINS
```

In [25]:
```
import pandas as pd
df = pd.read_csv('LOTR.csv')
print(df)
```

```
   FellowshipID FirstName     Skills
0          1001     Frodo     Hiding
1          1002   Samwise  Gardening
2          1003   Gandalf     Spells
3          1004    Pippin  Fireworks
```

In [26]:
```
df2 = pd.read_csv('LOTR 2.csv')
print(df2)
```

```
   FellowshipID FirstName   Age
0          1001     Frodo    50
1          1002   Samwise    39
2          1006   Legolas  2931
3          1007    Elrond  6520
4          1008  Barromir    51
```

In [27]:
```
#INNER JOIN
df.merge(df2)
```

Out[27]:

	FellowshipID	FirstName	Skills	Age
0	1001	Frodo	Hiding	50
1	1002	Samwise	Gardening	39

In [28]:
```
#INNER JOIN--SAME OUTPUT (a Testing...)
df.merge(df2, how = 'inner')
```

Out[28]:

	FellowshipID	FirstName	Skills	Age
0	1001	Frodo	Hiding	50
1	1002	Samwise	Gardening	39

In [29]:
```
#INNER JOIN--SAME OUTPUT (Joining ON FellowshipID only....)
df.merge(df2, how = 'inner', on = 'FellowshipID')
```

Out[29]:

	FellowshipID	FirstName_x	Skills	FirstName_y	Age
0	1001	Frodo	Hiding	Frodo	50
1	1002	Samwise	Gardening	Samwise	39

In [30]:
```
#INNER JOIN--SAME OUTPUT (Joined ON FellowshipID AND FirstName == original merge ouput)
df.merge(df2, how = 'inner', on = ['FellowshipID', 'FirstName'])
```

Out[30]:

	FellowshipID	FirstName	Skills	Age
0	1001	Frodo	Hiding	50
1	1002	Samwise	Gardening	39

In [ ]:
```
#SHIFT.....TAB
```

In [31]:
```
#OUTER JOIN--(a Testing...)
df.merge(df2, how = 'outer')
```

Out[31]:

	FellowshipID	FirstName	Skills	Age
0	1001	Frodo	Hiding	50.0
1	1002	Samwise	Gardening	39.0
2	1003	Gandalf	Spells	NaN
3	1004	Pippin	Fireworks	NaN
4	1006	Legolas	NaN	2931.0
5	1007	Elrond	NaN	6520.0
6	1008	Barromir	NaN	51.0

In [32]:
```
#LEFT JOIN--(a Testing...)
df.merge(df2, how = 'left')
```

Out[32]:

	FellowshipID	FirstName	Skills	Age
0	1001	Frodo	Hiding	50.0
1	1002	Samwise	Gardening	39.0
2	1003	Gandalf	Spells	NaN
3	1004	Pippin	Fireworks	NaN

In [33]:
```
#RIGHT JOIN--(a Testing...)
df.merge(df2, how = 'right')
```

Out[33]:

	FellowshipID	FirstName	Skills	Age
0	1001	Frodo	Hiding	50
1	1002	Samwise	Gardening	39
2	1006	Legolas	NaN	2931
3	1007	Elrond	NaN	6520
4	1008	Barromir	NaN	51

In [36]:
```
#CONCAT
pd.concat([df,df2])
```

Out[36]:

	FellowshipID	FirstName	Skills	Age
0	1001	Frodo	Hiding	NaN
1	1002	Samwise	Gardening	NaN
2	1003	Gandalf	Spells	NaN
3	1004	Pippin	Fireworks	NaN
0	1001	Frodo	NaN	50.0
1	1002	Samwise	NaN	39.0
2	1006	Legolas	NaN	2931.0
3	1007	Elrond	NaN	6520.0
4	1008	Barromir	NaN	51.0

In [38]:
```
#CONCAT
pd.concat([df,df2], join = 'outer', axis = 1)
```

Out[38]:

	FellowshipID	FirstName	Skills	FellowshipID	FirstName	Age
0	1001.0	Frodo	Hiding	1001	Frodo	50
1	1002.0	Samwise	Gardening	1002	Samwise	39
2	1003.0	Gandalf	Spells	1006	Legolas	2931
3	1004.0	Pippin	Fireworks	1007	Elrond	6520
4	NaN	NaN	NaN	1008	Barromir	51

In [39]:
```
#CONCAT
pd.concat([df,df2], join = 'inner', axis = 1)
```

Out[39]:

	FellowshipID	FirstName	Skills	FellowshipID	FirstName	Age
0	1001	Frodo	Hiding	1001	Frodo	50
1	1002	Samwise	Gardening	1002	Samwise	39
2	1003	Gandalf	Spells	1006	Legolas	2931
3	1004	Pippin	Fireworks	1007	Elrond	6520

In [40]:
```
#CONCAT
pd.concat([df,df2], join = 'outer', axis = 0)
```

Out[40]:

	FellowshipID	FirstName	Skills	Age
0	1001	Frodo	Hiding	NaN
1	1002	Samwise	Gardening	NaN
2	1003	Gandalf	Spells	NaN
3	1004	Pippin	Fireworks	NaN
0	1001	Frodo	NaN	50.0
1	1002	Samwise	NaN	39.0
2	1006	Legolas	NaN	2931.0
3	1007	Elrond	NaN	6520.0
4	1008	Barromir	NaN	51.0

In [41]:
```
#CONCAT
pd.concat([df,df2], join = 'inner', axis = 0)
```

Out[41]:

	FellowshipID	FirstName
0	1001	Frodo
1	1002	Samwise
2	1003	Gandalf
3	1004	Pippin
0	1001	Frodo
1	1002	Samwise
2	1006	Legolas
3	1007	Elrond
4	1008	Barromir

In [ ]:
```

```