In [1]:

```
import pandas as pd
```

In [10]:

```
wk1=pd.read_csv("Restaurant - Week 1 Sales.csv")

wk2=pd.read_csv("Restaurant - Week 2 Sales.csv")

cust=pd.read_csv("Restaurant - Customers.csv")

food=pd.read_csv("Restaurant - Foods.csv")
```

The pd.concat() mtd part-1

In [11]:

```
wk1=pd.read_csv("Restaurant - Week 1 Sales.csv")

wk2=pd.read_csv("Restaurant - Week 2 Sales.csv")

cust=pd.read_csv("Restaurant - Customers.csv")

food=pd.read_csv("Restaurant - Foods.csv")
```

In [13]:

```
wk1.head()
```

Out[13]:

	Customer ID	Food ID
0	537	9
1	97	4
2	658	1
3	202	2
4	155	9

In [14]:

```
wk2.head()
```

Out[14]:

	Customer ID	Food ID
0	688	10
1	813	7
2	495	10
3	189	5
4	267	3

In [20]:

```
pd.concat(objs=[wk1,wk2]).tail()#0-249 +0-249      #1
```

Out[20]:

	Customer ID	Food ID
245	783	10
246	556	10
247	547	9
248	252	9
249	249	6

In [23]:

```
pd.concat(objs=[wk1,wk2],ignore_index=True).tail()
```

Out[23]:

	Customer ID	Food ID
495	783	10
496	556	10
497	547	9
498	252	9
499	249	6

In [24]:

```
wk1.append(wk2,ignore_index=True).tail()            #2
```

Out[24]:

	Customer ID	Food ID
495	783	10
496	556	10
497	547	9
498	252	9
499	249	6

The pd.concat() mtd part-

In [25]:

```
wk1=pd.read_csv("Restaurant - Week 1 Sales.csv")

wk2=pd.read_csv("Restaurant - Week 2 Sales.csv")

cust=pd.read_csv("Restaurant - Customers.csv")

food=pd.read_csv("Restaurant - Foods.csv")
```

In [28]:

```
pd.concat(objs=[wk1,wk2],ignore_index=True).head()
```

Out[28]:

	Customer ID	Food ID
0	537	9
1	97	4
2	658	1
3	202	2
4	155	9

In [30]:

```
sales=pd.concat(objs=[wk1,wk2],keys=["Week 1","Week 2"])#To make it different identifier
```

In [31]:

```
sales.head()
```

Out[31]:

		Customer ID	Food ID
Week 1	**0**	537	9
	1	97	4
	2	658	1
	3	202	2
	4	155	9

In [38]:

```
sales.loc[("Week 1",),("Customer ID")].head()
```

Out[38]:

```
0    537
1     97
2    658
3    202
4    155
Name: Customer ID, dtype: int64
```

In [39]:

```
sales.loc[("Week 1",)].head()
```

Out[39]:

	Customer ID	Food ID
0	537	9
1	97	4
2	658	1
3	202	2
4	155	9

In [42]:

```
sales.loc[("Week 1",2)].head()#####particular row
```

Out[42]:

```
Customer ID    658
Food ID          1
Name: (Week 1, 2), dtype: int64
```

In [43]:

```
sales.loc[("Week 2",248)].head()#####particular row
```

Out[43]:

```
Customer ID    252
Food ID          9
Name: (Week 2, 248), dtype: int64
```

In [46]:

```
sales.loc[("Week 1",2),["Customer ID"]].head()#####particular row
```

Out[46]:

```
Customer ID    658
Name: (Week 1, 2), dtype: int64
```

In [47]:

```
sales.loc[("Week 1",2),["Customer ID","Food ID"]].head()#####particular row
```

Out[47]:

```
Customer ID    658
Food ID          1
Name: (Week 1, 2), dtype: int64
```

The inner join part-1

In [49]:

```
wk1=pd.read_csv("Restaurant - Week 1 Sales.csv")

wk2=pd.read_csv("Restaurant - Week 2 Sales.csv")

cust=pd.read_csv("Restaurant - Customers.csv")

food=pd.read_csv("Restaurant - Foods.csv")
```

In [50]:

```
wk1.head()
```

Out[50]:

	Customer ID	Food ID
0	537	9
1	97	4
2	658	1
3	202	2
4	155	9

In [51]:

```
wk2.head()
```

Out[51]:

	Customer ID	Food ID
0	688	10
1	813	7
2	495	10
3	189	5
4	267	3

In [68]:

```
wk1.merge(wk2,how="inner",on="Customer ID").head()###customers coe on both weeks
```

Out[68]:

	Customer ID	Food ID_x	Food ID_y
0	537	9	5
1	155	9	3
2	155	1	3
3	503	5	8
4	503	5	9

In [69]:

```
wk1.merge(wk2,how="inner",on="Customer ID",suffixes=["item id from week 1","item id from we
```

Out[69]:

	Customer ID	Food IDitem id from week 1	Food IDitem id from week 2
0	537	9	5
1	155	9	3
2	155	1	3
3	503	5	8
4	503	5	9

In [70]:

```
wk1.merge(wk2,how="inner",on="Customer ID",suffixes=["- from week 1","- from week 2"]).head
```

Out[70]:

	Customer ID	Food ID- from week 1	Food ID- from week 2
0	537	9	5
1	155	9	3
2	155	1	3
3	503	5	8
4	503	5	9

In [64]:

```
wk1[wk1["Customer ID"]==155]
```

Out[64]:

	Customer ID	Food ID
4	155	9
17	155	1

In [66]:

```
wk2[wk2["Customer ID"]==155]
```

Out[66]:

	Customer ID	Food ID
208	155	3

The inner join part-2

In [71]:

```
wk1=pd.read_csv("Restaurant - Week 1 Sales.csv")

wk2=pd.read_csv("Restaurant - Week 2 Sales.csv")

cust=pd.read_csv("Restaurant - Customers.csv")

food=pd.read_csv("Restaurant - Foods.csv")
```

In [72]:

```
wk1.head()
```

Out[72]:

	Customer ID	Food ID
0	537	9
1	97	4
2	658	1
3	202	2
4	155	9

In [73]:

```
wk2.head()
```

Out[73]:

	Customer ID	Food ID
0	688	10
1	813	7
2	495	10
3	189	5
4	267	3

In [77]:

```
wk1.merge(wk2,how="inner",on=["Customer ID","Food ID"])#1
```

Out[77]:

	Customer ID	Food ID
0	304	3
1	540	3
2	937	10
3	233	3
4	21	4
5	21	4
6	922	1
7	578	5
8	578	5

In [80]:

```
wk1.merge(wk2)#2
```

Out[80]:

	Customer ID	Food ID
0	304	3
1	540	3
2	937	10
3	233	3
4	21	4
5	21	4
6	922	1
7	578	5
8	578	5

In [81]:

```
wk1[wk1["Customer ID"]==21]
```

Out[81]:

	Customer ID	Food ID
101	21	4
212	21	4

In [82]:

```
wk2[wk2["Customer ID"]==21]
```

Out[82]:

	Customer ID	Food ID
30	21	4

outer join part-1

In [83]:

```
wk1=pd.read_csv("Restaurant - Week 1 Sales.csv")

wk2=pd.read_csv("Restaurant - Week 2 Sales.csv")

cust=pd.read_csv("Restaurant - Customers.csv")

food=pd.read_csv("Restaurant - Foods.csv")
```

In [87]:

```
wk1.merge(wk2,how="outer",on="Customer ID")
```

	Customer ID	Food ID_x	Food ID_y
0	537	9.0	5.0
1	97	4.0	NaN
2	658	1.0	NaN
3	202	2.0	NaN
4	155	9.0	3.0
...	...	...	...
449	855	NaN	4.0
450	559	NaN	10.0
451	276	NaN	4.0
452	556	NaN	10.0
453	252	NaN	9.0

In [96]:

```
wk1.merge(wk2,how="outer",on="Customer ID",indicator=True)#indicator
```

Out[96]:

	Customer ID	Food ID_x	Food ID_y	_merge
0	537	9.0	5.0	both
1	97	4.0	NaN	left_only
2	658	1.0	NaN	left_only
3	202	2.0	NaN	left_only
4	155	9.0	3.0	both
...	...	...	...	...
449	855	NaN	4.0	right_only
450	559	NaN	10.0	right_only
451	276	NaN	4.0	right_only
452	556	NaN	10.0	right_only
453	252	NaN	9.0	right_only

454 rows × 4 columns

In [88]:

```
mergeOut=wk1.merge(wk2,how="outer",on="Customer ID",indicator=True)#indicator
```

In [95]:

```
mergeOut[mergeOut["_merge"]=="both"]#1          ###inner join only rows in both from outer
```

Out[95]:

	Customer ID	Food ID_x	Food ID_y	_merge
0	537	9.0	5.0	both
4	155	9.0	3.0	both
5	155	1.0	3.0	both
8	503	5.0	8.0	both
9	503	5.0	9.0	both
...	...	...	...	...
246	945	5.0	4.0	both
247	343	3.0	5.0	both
248	343	3.0	2.0	both
249	343	3.0	7.0	both
251	621	9.0	6.0	both

62 rows × 4 columns

In [94]:

```
wk1.merge(wk2,how="inner",on="Customer ID")      ###same as above
```

Out[94]:

	Customer ID	Food ID_x	Food ID_y
0	537	9	5
1	155	9	3
2	155	1	3
3	503	5	8
4	503	5	9
...	...	...	...
57	945	5	4
58	343	3	5
59	343	3	2
60	343	3	7
61	621	9	6

62 rows × 3 columns

In [99]:

```
mergeOut[mergeOut["_merge"]!="both"].head()   ### only in left and only in right
```

Out[99]:

	Customer ID	Food ID_x	Food ID_y	_merge
1	97	4.0	NaN	left_only
2	658	1.0	NaN	left_only
3	202	2.0	NaN	left_only
6	213	8.0	NaN	left_only
7	600	1.0	NaN	left_only

In [100]:

```
mergeOut[mergeOut["_merge"]!="both"].tail()   ### only in left and only in right
```

Out[100]:

	Customer ID	Food ID_x	Food ID_y	_merge
449	855	NaN	4.0	right_only
450	559	NaN	10.0	right_only
451	276	NaN	4.0	right_only
452	556	NaN	10.0	right_only
453	252	NaN	9.0	right_only

left join part-1

In [106]:

```
wk1=pd.read_csv("Restaurant - Week 1 Sales.csv")

wk2=pd.read_csv("Restaurant - Week 2 Sales.csv")

cust=pd.read_csv("Restaurant - Customers.csv")

food=pd.read_csv("Restaurant - Foods.csv")
```

In [110]:

```
LEFT=wk1.merge(food,how="left",on ="Food ID",sort=True)
LEFT.head()
```

Out[110]:

	Customer ID	Food ID	Food Item	Price
0	658	1	Sushi	3.99
1	600	1	Sushi	3.99
2	155	1	Sushi	3.99
3	341	1	Sushi	3.99
4	20	1	Sushi	3.99

In [111]:

```
LEFT.tail()
```

Out[111]:

	Customer ID	Food ID	Food Item	Price
245	809	10	Drink	1.75
246	584	10	Drink	1.75
247	274	10	Drink	1.75
248	151	10	Drink	1.75
249	535	10	Drink	1.75

left join part-2

In [112]:

```
wk1=pd.read_csv("Restaurant - Week 1 Sales.csv")

wk2=pd.read_csv("Restaurant - Week 2 Sales.csv")

cust=pd.read_csv("Restaurant - Customers.csv")

food=pd.read_csv("Restaurant - Foods.csv")
```

In [113]:

```
cust.sort_values("").head()
```

Out[113]:

	ID	First Name	Last Name	Gender	Company	Occupation
0	1	Joseph	Perkins	Male	Dynazzy	Community Outreach Specialist
1	2	Jennifer	Alvarez	Female	DabZ	Senior Quality Engineer
2	3	Roger	Black	Male	Tagfeed	Account Executive
3	4	Steven	Evans	Male	Fatz	Registered Nurse
4	5	Judy	Morrison	Female	Demivee	Legal Assistant

In [114]:

```
wk1.head()
```

Out[114]:

	Customer ID	Food ID
0	537	9
1	97	4
2	658	1
3	202	2
4	155	9

In [125]:

```
wk1.merge(cust,how ="left",left_on="Customer ID",right_on="ID",sort=True).head()###implicit
```

Out[125]:

	Customer ID	Food ID	ID	First Name	Last Name	Gender	Company	Occupation
0	3	2	3	Roger	Black	Male	Tagfeed	Account Executive
1	10	2	10	Steven	Ryan	Male	Twinder	Community Outreach Specialist
2	20	1	20	Lisa	Rice	Female	Oloo	Programmer IV
3	21	4	21	Albert	Burns	Male	Rhynoodle	Junior Executive
4	21	4	21	Albert	Burns	Male	Rhynoodle	Junior Executive

In [121]:

```
##explicit sort based on any filed
wk1.merge(cust,how ="left",left_on="Customer ID",right_on="ID").sort_values(by="Customer ID
```

Out[121]:

	Customer ID	Food ID	ID	First Name	Last Name	Gender	Company	Occupation
96	3	2	3	Roger	Black	Male	Tagfeed	Account Executive
143	10	2	10	Steven	Ryan	Male	Twinder	Community Outreach Specialist
31	20	1	20	Lisa	Rice	Female	Oloo	Programmer IV
101	21	4	21	Albert	Burns	Male	Rhynoodle	Junior Executive
212	21	4	21	Albert	Burns	Male	Rhynoodle	Junior Executive

In [122]:

```
wk1.merge(cust,how ="left",left_on="Customer ID",right_on="ID").sort_values(by="Customer ID
```

Out[122]:

	Customer ID	Food ID	ID	First Name	Last Name	Gender	Company	Occupation
11	966	5	966	Robert	Ford	Male	Jabbertype	Account Representative IV
175	968	1	968	Teresa	Reynolds	Female	Flashdog	Budget/Accounting Analyst IV
209	985	5	985	Julia	Ortiz	Female	Kwideo	Structural Analysis Engineer
194	991	2	991	Melissa	Wells	Female	Lazzy	Senior Sales Associate
146	1000	2	1000	Brian	Daniels	Male	Tazzy	Physical Therapy Assistant

Merging based on left and right indices

In [128]:

```
wk1=pd.read_csv("Restaurant - Week 1 Sales.csv")

wk2=pd.read_csv("Restaurant - Week 2 Sales.csv")

cust=pd.read_csv("Restaurant - Customers.csv",index_col="ID")

food=pd.read_csv("Restaurant - Foods.csv",index_col="Food ID")
```

In [131]:

```
cust.head()##id as index
```

Out[131]:

ID	First Name	Last Name	Gender	Company	Occupation
1	Joseph	Perkins	Male	Dynazzy	Community Outreach Specialist
2	Jennifer	Alvarez	Female	DabZ	Senior Quality Engineer
3	Roger	Black	Male	Tagfeed	Account Executive
4	Steven	Evans	Male	Fatz	Registered Nurse
5	Judy	Morrison	Female	Demivee	Legal Assistant

In [130]:

```
wk1.head()
```

Out[130]:

	Customer ID	Food ID
0	537	9
1	97	4
2	658	1
3	202	2
4	155	9

In [134]:

```
sales=wk1.merge(cust,how="left",left_on="Customer ID",right_index=True)

###to merge left colm and right index which causes no addtnal id colm
```

In [136]:

```
sales.head()
```

Out[136]:

	Customer ID	Food ID	First Name	Last Name	Gender	Company	Occupation
0	537	9	Cheryl	Carroll	Female	Zoombeat	Registered Nurse
1	97	4	Amanda	Watkins	Female	Ozu	Account Coordinator
2	658	1	Patrick	Webb	Male	Browsebug	Community Outreach Specialist
3	202	2	Louis	Campbell	Male	Rhynoodle	Account Representative III
4	155	9	Carolyn	Diaz	Female	Gigazoom	Database Administrator III

In [137]:

```
food.head()
```

Out[137]:

Food ID	Food Item	Price
1	Sushi	3.99
2	Burrito	9.99
3	Taco	2.99
4	Quesadilla	4.25
5	Pizza	2.49

In [141]:

```
sales=sales.merge(food,how="left",left_on="Food ID",right_index=True)
```

In [142]:

```
sales.head()
```

Out[142]:

	Customer ID	Food ID	First Name	Last Name	Gender	Company	Occupation	Food Item	Price
0	537	9	Cheryl	Carroll	Female	Zoombeat	Registered Nurse	Donut	0.99
1	97	4	Amanda	Watkins	Female	Ozu	Account Coordinator	Quesadilla	4.25
2	658	1	Patrick	Webb	Male	Browsebug	Community Outreach Specialist	Sushi	3.99
3	202	2	Louis	Campbell	Male	Rhynoodle	Account Representative III	Burrito	9.99
4	155	9	Carolyn	Diaz	Female	Gigazoom	Database Administrator III	Donut	0.99

In [139]:

```
food.index
```

Out[139]:

```
Int64Index([1, 2, 3, 4, 5, 6, 7, 8, 9, 10], dtype='int64', name='Food ID')
```

In [143]:

```
wk1.head()
```

Out[143]:

	Customer ID	Food ID
0	537	9
1	97	4
2	658	1
3	202	2
4	155	9

In [144]:

```
wk2.head()
```

Out[144]:

	Customer ID	Food ID
0	688	10
1	813	7
2	495	10
3	189	5
4	267	3

In [146]:

```
wk1.merge(wk2,how="left",left_index=True,right_index=True)##merge based on both indices
```

	Customer ID_x	Food ID_x	Customer ID_y	Food ID_y
0	537	9	688	10
1	97	4	813	7
2	658	1	495	10
3	202	2	189	5
4	155	9	267	3
...	...	...	...	...
245	413	9	783	10
246	926	6	556	10
247	134	3	547	9
248	396	6	252	9

The .join() method

In [149]:

```
wk1=pd.read_csv("Restaurant - Week 1 Sales.csv")

wk2=pd.read_csv("Restaurant - Week 2 Sales.csv")

cust=pd.read_csv("Restaurant - Customers.csv",index_col="ID")

food=pd.read_csv("Restaurant - Foods.csv",index_col="Food ID")

satis=pd.read_csv("Restaurant - Week 1 Satisfaction.csv")
```

In [150]:

```
satis.head()
```

Out[150]:

	Satisfaction Rating
0	2
1	7
2	3
3	7
4	10

In [151]:

```
wk1.head()
```

Out[151]:

	Customer ID	Food ID
0	537	9
1	97	4
2	658	1
3	202	2
4	155	9

In [154]:

```
wk1.merge(satis,how="left",left_index=True,right_index=True).head()
```

Out[154]:

	Customer ID	Food ID	Satisfaction Rating
0	537	9	2
1	97	4	7
2	658	1	3
3	202	2	7
4	155	9	10

In [156]:

```
wk1.join(satis).head()#######neeed to use when we have same idices and same rows and to app
```

Out[156]:

	Customer ID	Food ID	Satisfaction Rating
0	537	9	2
1	97	4	7
2	658	1	3
3	202	2	7
4	155	9	10

tHE pd.merge() method

In [157]:

```
wk1=pd.read_csv("Restaurant - Week 1 Sales.csv")

wk2=pd.read_csv("Restaurant - Week 2 Sales.csv")

cust=pd.read_csv("Restaurant - Customers.csv",index_col="ID")

food=pd.read_csv("Restaurant - Foods.csv",index_col="Food ID")

satis=pd.read_csv("Restaurant - Week 1 Satisfaction.csv")
```

In [159]:

```
pd.merge(wk1,cust,how ="left",left_on="Customer ID",right_on="ID").head()####1 latest
```

Out[159]:

	Customer ID	Food ID	First Name	Last Name	Gender	Company	Occupation
0	537	9	Cheryl	Carroll	Female	Zoombeat	Registered Nurse
1	97	4	Amanda	Watkins	Female	Ozu	Account Coordinator
2	658	1	Patrick	Webb	Male	Browsebug	Community Outreach Specialist
3	202	2	Louis	Campbell	Male	Rhynoodle	Account Representative III
4	155	9	Carolyn	Diaz	Female	Gigazoom	Database Administrator III

In [161]:

```
wk1.merge(cust,how="left",left_on="Customer ID",right_on="ID").head()####2 olddd
```

Out[161]:

	Customer ID	Food ID	First Name	Last Name	Gender	Company	Occupation
0	537	9	Cheryl	Carroll	Female	Zoombeat	Registered Nurse
1	97	4	Amanda	Watkins	Female	Ozu	Account Coordinator
2	658	1	Patrick	Webb	Male	Browsebug	Community Outreach Specialist
3	202	2	Louis	Campbell	Male	Rhynoodle	Account Representative III
4	155	9	Carolyn	Diaz	Female	Gigazoom	Database Administrator III