```
# Perform the following operations using ‘python’ Language on the Amazon
book review Data set.
1) Create data subsets(Row wise)
2) Merge Data (Row wise)
3) Sort Data
4) Transposing Data
5) Reshape Data
```

In [35]:
```
import pandas as pd
import numpy as np
df = pd.read_csv(r"C:\Users\yasha\Desktop\Ashish\sem 6\DSBDA\DSBDA Lab Dataset
df.head()
```

Out[35]:

	reviewerID	asin	reviewername	helpful	reviewerText	overallRating	Summary	U
0	A2SUAM1J3GNN3B	23456.0	SNEHA	2	I bought this for my husband who plays the pia...	4.0	SUBWAY SURF	
1	A2SUAN1J3GNN0C	879877.0	KARTHII	3	good book to read	3.0	PILOT	
2	A2SUAN1J3GNN7B	23456.0	SHASHANK	3	high cost	2.5	DRIVE IN HEAVEN	
3	A5SUAN1J3GNN9B	67578.0	NITIN	3	good useof book	2.2	SPORTS HEAVEN	
4	A1SUAM1J3GNN3B	123.0	PRANOTI	3	I m happy with the book	3.0	HEAVEN OF PEARLS	

In [36]:
```
#creating the subset1

df1 = df[['asin','helpful','overallRating']].loc[0:15]
df1
```

Out[36]:

	asin	helpful	overallRating
0	23456.0	2	4.0
1	879877.0	3	3.0
2	23456.0	3	2.5
3	67578.0	3	2.2
4	123.0	3	3.0
5	236789.0	3	3.0
6	NaN	3	3.0
7	224442.0	3	3.0
8	98767.0	2	3.0
9	78761.0	2	5.0
10	321456.0	2	4.0
11	876231.0	2	4.0
12	987.0	2	4.0
13	234.0	2	3.0
14	124.0	2	3.0
15	156.0	2	3.0

In [37]:
```
#creating the subset2

df2 = df[['asin','helpful','overallRating']].loc[16:30]
df2
```

Out[37]:

	asin	helpful	overallRating
16	876.0	2	3.0
17	123.0	2	2.0
18	987.0	2	2.0
19	22456.0	2	2.0
20	9876.0	2	2.0
21	3451.0	2	4.0
22	879.0	2	4.0
23	321.0	2	4.0
24	2.0	3	4.0
25	1.0	3	4.0
26	23.0	3	4.0
27	11.0	3	4.0
28	11.0	3	4.0
29	23.0	3	4.0
30	21221.0	3	4.0

In [38]:
```
#creating the subset3

df3 = df[['asin','helpful','overallRating']].loc[31:44]
df3
```

Out[38]:

	asin	helpful	overallRating
31	21.0	3	4.0
32	12212.0	3	3.0
33	2121.0	3	3.0
34	21212.0	3	3.0
35	21221.0	3	3.0
36	22245.0	3	2.0
37	678.0	3	2.0
38	111.0	3	2.0
39	223.0	3	2.0
40	2312.0	3	4.0
41	1222.0	2	4.0
42	11111.0	2	4.0
43	9800.0	2	4.0
44	NaN	2	4.0

In [39]:
```
#merge
merging = pd.concat([df1,df2,df3])
merging
```

Out[39]:

	asin	helpful	overallRating
0	23456.0	2	4.0
1	879877.0	3	3.0
2	23456.0	3	2.5
3	67578.0	3	2.2
4	123.0	3	3.0
5	236789.0	3	3.0
6	NaN	3	3.0
7	224442.0	3	3.0
8	98767.0	2	3.0
9	78761.0	2	5.0
10	321456.0	2	4.0
11	876231.0	2	4.0
12	987.0	2	4.0
13	234.0	2	3.0
14	124.0	2	3.0
15	156.0	2	3.0
16	876.0	2	3.0
17	123.0	2	2.0
18	987.0	2	2.0
19	22456.0	2	2.0
20	9876.0	2	2.0
21	3451.0	2	4.0
22	879.0	2	4.0
23	321.0	2	4.0
24	2.0	3	4.0
25	1.0	3	4.0
26	23.0	3	4.0
27	11.0	3	4.0
28	11.0	3	4.0
29	23.0	3	4.0
30	21221.0	3	4.0
31	21.0	3	4.0
32	12212.0	3	3.0
33	2121.0	3	3.0
34	21212.0	3	3.0
35	21221.0	3	3.0

	asin	helpful	overallRating
36	22245.0	3	2.0
37	678.0	3	2.0
38	111.0	3	2.0
39	223.0	3	2.0
40	2312.0	3	4.0
41	1222.0	2	4.0
42	11111.0	2	4.0
43	9800.0	2	4.0
44	NaN	2	4.0

In [42]:
```
#sort
sort_values=df.sort_values('overallRating',ascending=False)
sort_values
```

Out[42]:

	reviewerID	asin	reviewername	helpful	reviewerText	overallRating	Sur
9	1234ZSDFREQWEG	78761.0	SATISH	2	many errors	5.0	DATA IS
0	A2SUAM1J3GNN3B	23456.0	SNEHA	2	I bought this for my husband who plays the pia...	4.0	SUBWAY
21	NIKI12NIKI1231	3451.0	SHREYA	2	dvdasaewb	4.0	DATA IS
43	123LKIUHNFG	9800.0	MAHI	2	DAER PLEASE READ BOOK	4.0	MC
42	QWCVFED123	11111.0	DHONI	2	HELLO DEAR	4.0	DEAR T
41	ASWER1234	1222.0	PREETI	2	LOVE U	4.0	
40	67FDASAA123	2312.0	NANAJJI	3	BYE MY LOVE	4.0	HELL OF
31	1234ZSDFREQWEG	21.0	VARSHA	3	stare the raeder	4.0	INDIA COU
30	MNHJ879WERFGH	21221.0	JOHN	3	money is dis	4.0	WINE OF
29	ASDF345671QWERT	23.0	HARRY	3	good buy for money	4.0	MC
28	CADERVG3450EREW	11.0	MORRIS	3	staff is good	4.0	DEAR T
27	BA2345WERSDFRER	11.0	CHRIS	3	helpful	4.0	
26	GHAU8768QWER	23.0	MARYLS	3	9.666666667	4.0	HELL OF
25	MIKISDF1232XAD	1.0	JACOB	3	0.888888888	4.0	DARK KI
24	NIKI12NIKI1231	2.0	PRAJWAL	3	NA	4.0	MOTIVAT
23	GHAU8768QWER	321.0	SUMEET	2	NA	4.0	FRIENDS
22	MIKISDF1232XAD	879.0	AISHWARYA	2	NA	4.0	I MA P
44	987DFHGJI657	NaN	DDE	2	BOOK	4.0	WINE OF
12	HGKIUYE345SFDG	987.0	MOHIT	2	the dealer is charged more money	4.0	MOTIVAT
11	NaN	876231.0	ROHIT	2	high vat	4.0	FRIENDS
10	9345GHIKLFLLED	321456.0	RAJU	2	no mini	4.0	I MA P
8	MNHJ879WERFGH	98767.0	AKSHAY	2	The music is at times hard to read because we ...	3.0	I MA P
4	A1SUAM1J3GNN3B	123.0	PRANOTI	3	I m happy with the book	3.0	HEAV PE

	reviewerID	asin	reviewername	helpful	reviewerText	overallRating	Sur
5	BA2345WERSDFRER	236789.0	SHALINI	3	im not happy with the delay of product	3.0	GODS
35	NIKRT2398701222	21221.0	NEIL	3	8.78E+21	3.0	good
34	HGKIUYE345SFDG	21212.0	KRISSH	3	fdghi	3.0	staff i
33	NaN	2121.0	SAIF	3	faskjas	3.0	
32	9345GHIKLFLLED	12212.0	SHRADDHA	3	reader is giood	3.0	ONE FOF
7	ASDF345671QWERT	224442.0	SHIKHA	3	Wonderful book	3.0	LIFE (
6	CADERVG3450EREW	NaN	SNEHAL	3	delay in book	3.0	GODES
13	NIKRT2398701222	234.0	NILESH	2	helpful	3.0	DARK KI
14	MAHI2348900122	124.0	KAILASH	2	staff is good	3.0	HELL OF
15	LIHJKD87634O21	156.0	SONAM	2	good buy for money	3.0	
1	A2SUAN1J3GNN0C	879877.0	KARTHII	3	good book to read	3.0	
16	SISISWER4657893	876.0	RAJESH	2	money is dis	3.0	DEAR T
2	A2SUAN1J3GNN7B	23456.0	SHASHANK	3	high cost	2.5	DR HE
3	A5SUAN1J3GNN9B	67578.0	NITIN	3	good useof book	2.2	SF HE
18	AAAJI37DFG378I	987.0	PANKAJ	2	reader is giood	2.0	WINE OF
36	MAHI2348900122	22245.0	KANGANA	3	67667	2.0	mone
37	LIHJKD87634O21	678.0	KHANN	3	3434567	2.0	stare the
38	SISISWER4657893	111.0	RIYAA	3	098O	2.0	reader is
39	8786112GHAUI	223.0	NaN	3	GOOD LUCK	2.0	
20	VABI9876ERTT1	9876.0	MANISH	2	fdghi	2.0	ONE FOF
19	FAJI234GAJI987	22456.0	DIPTI	2	faskjas	2.0	INDIA COU
17	MIKI23412341QW	123.0	SALMAN	2	stare the raeder	2.0	MC

In [43]:
```
#transpose
df.transpose()
```

Out[43]:

	0	1	2	3	
)	A2SUAM1J3GNN3B	A2SUAN1J3GNN0C	A2SUAN1J3GNN7B	A5SUAN1J3GNN9B	A1SUAM1J3GNN3
n	23456.0	879877.0	23456.0	67578.0	123
e	SNEHA	KARTHII	SHASHANK	NITIN	PRANO
il	2	3	3	3	
t	I bought this for my husband who plays the pia...	good book to read	high cost	good useof book	I m happy with tl boo
g	4.0	3.0	2.5	2.2	3
y	SUBWAY SURF	PILOT	DRIVE IN HEAVEN	SPORTS HEAVEN	HEAVEN C PEARL
e	NaN	NaN	NaN	NaN	Na
e	NaN	NaN	NaN	NaN	Na

lumns

In [13]:
```
df.shape
```

Out[13]: (45, 9)

In [44]:
```
#reshape
pivot_table = pd.pivot_table(df,index=['helpful','overallRating'], values='as
pivot_table
```

Out[44]:

		asin
helpful	**overallRating**	
2	**2.0**	8360.500
	3.0	20031.400
	4.0	124891.400
	5.0	78761.000
3	**2.0**	5814.250
	2.2	67578.000
	2.5	23456.000
	3.0	174749.625
	4.0	2625.000

In [47]:
```
# merge another way
data1 = {'id':['1','2','3','4'],
         'name':['ashish','ramesh','banty','munna'],
         'subj':['maths','chem','phy','bio']}

data2 = {'id':['1','5','3','7'],
         'address':['dange','chnichwad','pune','mumbai'],
         'color':['red','green','blue','yellow']}

#converting dictionary to dataframe
demo1=pd.DataFrame(data1)

#converting dictionaries to dataframe
demo2=pd.DataFrame(data2)

data = pd.merge(demo1,demo2, on='id')
data
```

Out[47]:

	id	name	subj	address	color
0	1	ashish	maths	dange	red
1	3	banty	phy	pune	blue

In [ ]:
```

```