--- Assignment Soution By----------------------- -----

Rohit kumar pandey

Q1. 1. Create an output file "filteredCountry.csv" inside a folder named "output". This file should contain only those records where country contains the word USA.

```
In [490]:  import pandas as pd
```

```
In [491]:  data = pd.read_csv("input/main.csv")
```

```
In [492]:  data.head()
```

Out[492]:

	SKU	DESCRIPTION	YEAR	CAPACITY	URL	PRICE	SELLER_I
0	11331	Barbaresco,Angelo Gaja	1990	750ml	https://www.wine-searcher.com/find/barbaresco+...	$295.00	TopWines
1	11331	Barbaresco,Angelo Gaja	1990	750ml	https://www.wine-searcher.com/find/barbaresco+...	$324.62	Vin
2	11331	Barbaresco,Angelo Gaja	1990	750ml	https://www.wine-searcher.com/find/barbaresco+...	$329.00	The F
3	11331	Barbaresco,Angelo Gaja	1990	750ml	https://www.wine-searcher.com/find/barbaresco+...	$338.68	Wine Walk
4	11331	Barbaresco,Angelo Gaja	1990	750ml	https://www.wine-searcher.com/find/barbaresco+...	$396.83	Enc

```
In [493]:  df.columns
```

Out[493]: Index(['SKU', 'PRICE'], dtype='object')

```
In [494]:  df1 = data.loc[data['COUNTRY']=='USA (CA)']
```

In [495]: df1

Out[495]:

	SKU	DESCRIPTION	YEAR	CAPACITY	URL	PRICE	SEL
2	11331	Barbaresco,Angelo Gaja	1990	750ml	https://www.wine-searcher.com/find/barbaresco+...	$329.00	
6	11332	Barbaresco,Angelo Gaja	2010	750ml	https://www.wine-searcher.com/find/barbaresco+...	$175.00	
18	11333	Barbaresco,Costa Russi,Angelo Gaja	2009	750ml	https://www.wine-searcher.com/find/barbaresco+...	$304.00	
23	11333	Barbaresco,Costa Russi,Angelo Gaja	2009	750ml	https://www.wine-searcher.com/find/barbaresco+...	$379.00	
24	11333	Barbaresco,Costa Russi,Angelo Gaja	2009	750ml	https://www.wine-searcher.com/find/barbaresco+...	$379.98	
...	...	...	...	...	...	...	
11170	11624	Cote-Rotie,La Turque,Guigal	2009	750ml	https://www.wine-searcher.com/find/cote-rotie+...	$699.97	
11172	11624	Cote-Rotie,La Turque,Guigal	2009	750ml	https://www.wine-searcher.com/find/cote-rotie+...	$739.89	
11185	11625	Dominio De Pingus,Peter Sisseck	2005	750ml	https://www.wine-searcher.com/find/dominio+de+...	$949.00	
11186	11625	Dominio De Pingus,Peter Sisseck	2005	750ml	https://www.wine-searcher.com/find/dominio+de+...	$980.00	
11194	11625	Dominio De Pingus,Peter Sisseck	2005	750ml	https://www.wine-searcher.com/find/dominio+de+...	$5875.00	

1490 rows × 9 columns

In [496]: df1.to_csv('output/filteredCountry.csv',index=False)

In [ ]:

In [ ]:

Q2. Now consider "filteredCountry.csv" as the input file. For each group of "SKU" find 2 minimum prices and store this result in "lowestPrice.csv" inside a folder named "output"

In [ ]:

In [509]: data = pd.read_csv("output/filteredCountry.csv")

In [510]: `data.sort_values(by=['SKU']).head(50)`

Out[510]:

	SKU	DESCRIPTION	YEAR	CAPACITY	URL	PRICE	S
0	11331	Barbaresco,Angelo Gaja	1990	750ml	https://www.wine-searcher.com/find/barbaresco+...	$329.00	
1	11332	Barbaresco,Angelo Gaja	2010	750ml	https://www.wine-searcher.com/find/barbaresco+...	$175.00	
2	11333	Barbaresco,Costa Russi,Angelo Gaja	2009	750ml	https://www.wine-searcher.com/find/barbaresco+...	$304.00	
3	11333	Barbaresco,Costa Russi,Angelo Gaja	2009	750ml	https://www.wine-searcher.com/find/barbaresco+...	$379.00	
4	11333	Barbaresco,Costa Russi,Angelo Gaja	2009	750ml	https://www.wine-searcher.com/find/barbaresco+...	$379.98	
5	11334	Barbaresco,Sori San Lorenzo,Angelo Gaja	2000	750ml	https://www.wine-searcher.com/find/barbaresco+...	$339.00	
6	11334	Barbaresco,Sori San Lorenzo,Angelo Gaja	2000	750ml	https://www.wine-searcher.com/find/barbaresco+...	$348.00	
7	11334	Barbaresco,Sori San Lorenzo,Angelo Gaja	2000	750ml	https://www.wine-searcher.com/find/barbaresco+...	$348.00	
8	11334	Barbaresco,Sori San Lorenzo,Angelo Gaja	2000	750ml	https://www.wine-searcher.com/find/barbaresco+...	$700.00	
9	11334	Barbaresco,Sori San Lorenzo,Angelo Gaja	2000	750ml	https://www.wine-searcher.com/find/barbaresco+...	$4175.00	
13	11335	Barbaresco,Sori Tildin,Angelo Gaja	2009	750ml	https://www.wine-searcher.com/find/barbaresco+...	$3775.00	
12	11335	Barbaresco,Sori Tildin,Angelo Gaja	2009	750ml	https://www.wine-searcher.com/find/barbaresco+...	$449.97	
11	11335	Barbaresco,Sori Tildin,Angelo Gaja	2009	750ml	https://www.wine-searcher.com/find/barbaresco+...	$315.00	
10	11335	Barbaresco,Sori Tildin,Angelo Gaja	2009	750ml	https://www.wine-searcher.com/find/barbaresco+...	$304.00	
14	11336	Barbaresco,Sori Tildin,Angelo Gaja	2010	1.5L	https://www.wine-searcher.com/find/barbaresco+...	$395.00	E
15	11336	Barbaresco,Sori Tildin,Angelo Gaja	2010	1.5L	https://www.wine-searcher.com/find/barbaresco+...	$455.00	
16	11336	Barbaresco,Sori Tildin,Angelo Gaja	2010	1.5L	https://www.wine-searcher.com/find/barbaresco+...	$455.00	
17	11336	Barbaresco,Sori Tildin,Angelo Gaja	2010	1.5L	https://www.wine-searcher.com/find/barbaresco+...	$489.97	
18	11336	Barbaresco,Sori Tildin,Angelo Gaja	2010	1.5L	https://www.wine-searcher.com/find/barbaresco+...	$489.99	
19	11336	Barbaresco,Sori Tildin,Angelo Gaja	2010	1.5L	https://www.wine-searcher.com/find/barbaresco+...	$910.00	
20	11336	Barbaresco,Sori Tildin,Angelo Gaja	2010	1.5L	https://www.wine-searcher.com/find/barbaresco+...	$2725.00	
27	11337	Barbaresco,Sori Tildin,Angelo Gaja	2010	750ml	https://www.wine-searcher.com/find/barbaresco+...	$2725.00	
26	11337	Barbaresco,Sori Tildin,Angelo Gaja	2010	750ml	https://www.wine-searcher.com/find/barbaresco+...	$910.00	

	SKU	DESCRIPTION	YEAR	CAPACITY	URL	PRICE	S
25	11337	Barbaresco,Sori Tildin,Angelo Gaja	2010	750ml	https://www.wine-searcher.com/find/barbaresco+...	$489.99	
21	11337	Barbaresco,Sori Tildin,Angelo Gaja	2010	750ml	https://www.wine-searcher.com/find/barbaresco+...	$395.00	E
23	11337	Barbaresco,Sori Tildin,Angelo Gaja	2010	750ml	https://www.wine-searcher.com/find/barbaresco+...	$455.00	
22	11337	Barbaresco,Sori Tildin,Angelo Gaja	2010	750ml	https://www.wine-searcher.com/find/barbaresco+...	$455.00	
24	11337	Barbaresco,Sori Tildin,Angelo Gaja	2010	750ml	https://www.wine-searcher.com/find/barbaresco+...	$489.97	
28	11338	Barolo Bussia Colonnello,Aldo Conterno	1989	750ml	https://www.wine-searcher.com/find/barolo+buss...	$156.00	
29	11338	Barolo Bussia Colonnello,Aldo Conterno	1989	750ml	https://www.wine-searcher.com/find/barolo+buss...	$278.00	
30	11339	Barolo Bussia Colonnello,Aldo Conterno	1990	750ml	https://www.wine-searcher.com/find/barolo+buss...	$266.00	
31	11340	Barolo Monprivato,Giuseppe Mascarello	1989	750ml	https://www.wine-searcher.com/find/barolo+monp...	$949.00	
32	11342	Barolo,Gran Bussia,Aldo Conterno	1990	750ml	https://www.wine-searcher.com/find/barolo+gran...	$499.00	
33	11342	Barolo,Gran Bussia,Aldo Conterno	1990	750ml	https://www.wine-searcher.com/find/barolo+gran...	$607.00	
34	11343	Barolo,Sperss,Angelo Gaja	2000	750ml	https://www.wine-searcher.com/find/barolo+sper...	$228.00	
35	11344	Beaune,Closdes Mouches,rouge,Domaine Joseph Dr...	2005	750ml	https://www.wine-searcher.com/find/beaune+clos...	$145.00	
36	11345	Bienvenues-Batard-Montrachet,Domaine Vincent L...	2015	750ml	https://www.wine-searcher.com/find/bienvenue+b...	$626.00	
37	11345	Bienvenues-Batard-Montrachet,Domaine Vincent L...	2015	750ml	https://www.wine-searcher.com/find/bienvenue+b...	$2499.00	
38	11346	Bonnes Mares,Domaine Comte Georgesde Vogue	2011	750ml	https://www.wine-searcher.com/find/bienvenue+b...	$499.97	
39	11347	Bonnes Mares,Domaine Drouhin Laroze	2010	750ml	https://www.wine-searcher.com/find/bonnes+mare...	$158.00	
40	11348	Carruades De Lafite	1999	750ml	https://www.wine-searcher.com/find/carruades+d...	$255.00	
41	11348	Carruades De Lafite	1999	750ml	https://www.wine-searcher.com/find/carruades+d...	$325.55	
42	11348	Carruades De Lafite	1999	750ml	https://www.wine-searcher.com/find/carruades+d...	$377.00	
45	11349	Carruades De Lafite	2003	750ml	https://www.wine-searcher.com/find/carruades+d...	$380.00	

	SKU	DESCRIPTION	YEAR	CAPACITY	URL	PRICE	S
44	11349	Carruades De Lafite	2003	750ml	https://www.wine-searcher.com/find/carruades+d...	$305.00	
43	11349	Carruades De Lafite	2003	750ml	https://www.wine-searcher.com/find/carruades+d...	$249.00	
46	11350	Ch Angelus	1994	750ml	https://www.wine-searcher.com/find/ch+angelus/...	$232.89	
47	11350	Ch Angelus	1994	750ml	https://www.wine-searcher.com/find/ch+angelus/...	$373.00	
48	11350	Ch Angelus	1994	750ml	https://www.wine-searcher.com/find/ch+angelus/...	$456.89	
49	11350	Ch Angelus	1994	750ml	https://www.wine-searcher.com/find/ch+angelus/...	$946.89	

In [514]:
```python
df  = data.iloc[:,[0,5]]
df
```

Out[514]:

	SKU	PRICE
0	11331	$329.00
1	11332	$175.00
2	11333	$304.00
3	11333	$379.00
4	11333	$379.98
...	...	...
1485	11624	$699.97
1486	11624	$739.89
1487	11625	$949.00
1488	11625	$980.00
1489	11625	$5875.00

1490 rows × 2 columns

In [ ]:

In [515]:
```python
count = 0
value12 = {}
list1 = []
for i in df.index:
    if not df['SKU'][i] in value12.keys():
        value12[df['SKU'][i]] =[df['PRICE'][i]]
    else:
        value12[df['SKU'][i]].append(df['PRICE'][i])
    list1.append(df['PRICE'][i])
```

In [516]:
```python
for i in value12:
    for j in range(len(value12[i])):
        value12[i][j] = value12[i][j].replace("?","")
        value12[i][j] = value12[i][j].replace("$","")
        value12[i][j] = float(value12[i][j].replace(",",""))
```

In [ ]:

In [517]:
```python
correct_data = {}
for i in value12:
    if len(value12[i]) > 1:
        correct_data[i] = value12[i]
        correct_data[i].sort()

lowestPrice = ""
for i in correct_data:
    lowestPrice = lowestPrice+str(i)+","+str(correct_data[i][0]) + ","+str(cor
rect_data[i][1])+","
```

In [518]: lowestPrice

Out[518]: '11333,304.0,379.0,11334,339.0,348.0,11335,304.0,315.0,11336,395.0,455.0,1133
7,395.0,455.0,11338,156.0,278.0,11342,499.0,607.0,11345,626.0,2499.0,11348,25
5.0,325.55,11349,249.0,305.0,11350,232.89,373.0,11351,449.98,475.0,11352,320.
0,411.0,11353,375.0,430.0,11355,390.0,390.0,11356,250.0,259.99,11357,383.0,38
5.0,11358,675.0,999.99,11360,139.0,145.0,11361,79.99,127.0,11362,132.0,224.7
5,11363,96.95,129.94,11364,120.0,139.0,11365,475.0,538.0,11366,475.0,538.0,11
368,395.0,455.0,11369,435.0,450.0,11370,525.0,531.0,11371,873.0,875.0,11372,1
050.0,1069.0,11373,492.0,493.0,11374,59.0,65.0,11376,295.0,304.0,11377,185.0,
220.0,11378,215.0,215.0,11380,158.0,200.0,11381,120.0,125.0,11382,210.0,217.
0,11383,312.0,315.0,11384,218.0,225.0,11385,721.0,808.65,11386,125.0,269.98,1
1388,260.0,297.0,11389,316.0,354.0,11390,189.0,241.89,11391,189.0,241.89,1139
2,260.0,318.89,11393,258.0,325.0,11394,165.0,219.0,11395,285.0,311.0,11397,17
1.0,225.0,11398,218134.0,219252.0,11399,189.98,190.0,11400,215.0,240.0,11402,
162.0,219.98,11403,229.0,230.0,11404,155.0,158.0,11405,269.94,294.94,11406,22
5.0,228.0,11407,114.98,119.0,11408,163.89,245.0,11409,228.0,242.0,11410,215.
0,219.0,11411,129.0,129.98,11412,135.0,146.0,11413,195.0,195.0,11414,69.0,74.
0,11415,234.0,245.0,11416,152.0,152.0,11417,525.0,574.95,11418,849.0,858.0,11
419,819.0,825.0,11420,407.0,459.0,11421,644.98,663.0,11422,233.0,749.99,1142
3,188.0,221.0,11424,177.0,177.0,11425,195.0,199.0,11426,278.0,278.0,11427,37
5.0,380.0,11428,129.94,133.0,11429,200.0,210.0,11430,250.0,399.97,11432,245.
0,259.0,11433,278.0,284.0,11435,506.75,575.0,11436,621.0,625.0,11437,400.0,41
5.0,11438,1099.0,1195.0,11439,875.0,987.0,11440,399.0,961.0,11442,825.0,850.
0,11444,750.0,835.0,11445,750.0,835.0,11446,799.0,899.0,11448,898.0,1500.0,11
449,725.0,725.0,11451,846.0,849.0,11452,835.0,863.0,11453,425.0,435.0,11454,8
99.0,974.0,11455,899.0,925.0,11456,633.0,800.0,11458,392.0,899.99,11459,75.0,
83.0,11460,450.0,499.99,11461,495.0,499.98,11464,668.0,689.0,11465,795.0,795.
0,11467,475.0,475.0,11468,499.98,525.0,11469,549.99,582.0,11470,549.0,617.0,1
1471,871.0,874.95,11472,348.0,349.95,11473,149.94,149.99,11476,185.0,190.0,11
477,152.0,159.98,11478,133.0,140.0,11479,400.0,419.0,11480,459.0,481.0,11481,
345.0,380.0,11482,155.0,184.49,11483,155.0,184.49,11484,195.0,215.0,11485,19
5.0,215.0,11486,309.0,317.0,11487,310.0,323.0,11488,175.0,179.0,11489,188.0,1
99.0,11491,248.99,275.0,11492,123.95,199.99,11493,375.0,392.0,11494,325.0,35
0.0,11495,139.0,180.0,11496,185.0,195.0,11497,139.97,230.0,11498,139.0,145.0,
11499,195.0,200.0,11500,375.0,440.0,11502,164.0,169.0,11503,135.0,137.0,1150
5,175.0,179.0,11506,135.0,139.0,11507,177.0,185.0,11508,1000.0,1040.0,11509,4
50.0,475.0,11510,493.0,513.0,11511,525.0,575.0,11512,850.0,895.0,11513,550.0,
699.99,11514,506.0,509.0,11515,475.89,550.0,11516,450.0,500.0,11517,848.0,84
8.0,11518,799.0,830.0,11519,296.0,304.0,11520,463.65,588.0,11522,147.0,159.0,
11523,190.0,196.0,11524,225.0,235.0,11525,162.0,162.0,11526,288.0,288.0,1152
7,234.95,241.0,11528,1199.0,1249.99,11529,425.0,500.0,11530,557.0,575.0,1153
1,419.99,557.0,11532,495.0,499.0,11533,545.0,549.97,11534,549.0,599.97,11535,
685.0,695.0,11536,685.0,695.0,11537,575.0,582.0,11539,538.89,720.0,11540,250.
0,259.0,11541,354.0,365.0,11542,225.0,234.0,11543,248.0,249.0,11544,139.0,15
9.0,11545,175.0,188.95,11546,127.0,179.99,11547,209.0,213.0,11548,170.0,177.9
8,11549,170.0,177.98,11550,218.0,225.0,11551,250.0,275.0,11552,475.0,495.0,11
553,285.0,315.0,11554,180.0,215.0,11555,324.0,389.0,11556,390.0,405.0,11557,3
85.0,385.0,11559,139.0,152.0,11561,6319.0,7528.89,11562,2525.0,2530.0,11563,2
300.0,2395.0,11564,3500.0,3595.0,11565,135.0,135.0,11567,132.0,135.94,11568,1
77.0,188.0,11570,245.0,245.0,11571,190.0,190.0,11572,73.0,126.0,11573,73.0,12
6.0,11574,241.0,245.0,11575,142.0,151.0,11576,150.0,152.0,11577,129.0,162.0,1
1578,120.0,130.0,11579,135.0,152.0,11580,152.0,155.0,11581,74.98,137.0,11582,
171.0,179.0,11583,209.0,209.0,11584,209.0,214.94,11585,112.95,133.0,11586,13
9.0,145.0,11587,150.0,152.0,11588,250.0,254.94,11589,155.0,158.0,11590,13982
9.0,148778.0,11592,158.0,163.0,11593,389.0,405.0,11594,325.0,349.97,11595,99.
0,184.94,11596,194.94,196.0,11597,202.0,626.75,11598,169.0,189.0,11600,240.0,
247.0,11601,182.0,219.65,11603,135.0,247.0,11604,449.99,506.0,11605,430.0,50
6.0,11606,158.0,191.0,11607,239.0,608.0,11609,455.0,499.0,11610,470.0,475.0,1

```
1612,145.0,274.97,11613,149.99,166.0,11614,180.0,184.94,11615,259.99,274.94,1
1616,455.0,459.98,11618,316.0,499.99,11622,519.0,519.0,11624,500.0,525.0,1162
5,949.0,980.0,'
```

In [519]:
```
lowestPrice = lowestPrice.split(',');lowestPrice.remove("")
```

In [520]:
```python
import numpy as np
lowestPrice = pd.DataFrame(np.array(lowestPrice).reshape(int(len(lowestPrice)/
3),3), columns = ["SKU","FIRST_MINIMUM_PRICE","SECOND_MINIMUM_PRICE"] )
```

In [522]:
```
lowestPrice.head(10)
```

Out[522]:

	SKU	FIRST_MINIMUM_PRICE	SECOND_MINIMUM_PRICE
0	11333	304.0	379.0
1	11334	339.0	348.0
2	11335	304.0	315.0
3	11336	395.0	455.0
4	11337	395.0	455.0
5	11338	156.0	278.0
6	11342	499.0	607.0
7	11345	626.0	2499.0
8	11348	255.0	325.55
9	11349	249.0	305.0

In [523]:
```
lowestPrice.tail(10)
```

Out[523]:

	SKU	FIRST_MINIMUM_PRICE	SECOND_MINIMUM_PRICE
237	11610	470.0	475.0
238	11612	145.0	274.97
239	11613	149.99	166.0
240	11614	180.0	184.94
241	11615	259.99	274.94
242	11616	455.0	459.98
243	11618	316.0	499.99
244	11622	519.0	519.0
245	11624	500.0	525.0
246	11625	949.0	980.0

In [526]:
```python
lowestPrice.to_csv('output/lowestPrice.csv', index=False)
```

```
In [527]: lowestPrice.columns
```

Out[527]: Index(['SKU', 'FIRST_MINIMUM_PRICE', 'SECOND_MINIMUM_PRICE'], dtype='object')

In [ ]:

In [ ]:

In [ ]:

In [ ]: