In [1]:
```
import numpy as np
import pandas as pd
```

In [2]:
```
wine_review = pd.read_csv('Wine_review.csv')
wine_review
```

Out[2]:

	Unnamed: 0	country	description	designation	points	price	province	region_1	reg
0	0	US	This tremendous 100% varietal wine hails from ...	Martha's Vineyard	96	235.0	California	Napa Valley	
1	1	Spain	Ripe aromas of fig, blackberry and cassis are ...	Carodorum Selección Especial Reserva	96	110.0	Northern Spain	Toro	
2	2	US	Mac Watson honors the memory of a wine once ma...	Special Selected Late Harvest	96	90.0	California	Knights Valley	Sc
3	3	US	This spent 20 months in 30% new French oak, an...	Reserve	96	65.0	Oregon	Willamette Valley	Willa
4	4	France	This is the top wine from La Bégude, named aft...	La Brûlade	95	66.0	Provence	Bandol	
...	...	...	...	...	...	...	...	...	
150925	150925	Italy	Many people feel Fiano represents southern Ita...	NaN	91	20.0	Southern Italy	Fiano di Avellino	
150926	150926	France	Offers an intriguing nose with ginger, lime an...	Cuvée Prestige	91	27.0	Champagne	Champagne	
150927	150927	Italy	This classic example comes from a cru vineyard...	Terre di Dora	91	20.0	Southern Italy	Fiano di Avellino	
150928	150928	France	A perfect salmon shade, with scents of peaches...	Grand Brut Rosé	90	52.0	Champagne	Champagne	

	Unnamed: 0	country	description	designation	points	price	province	region_1	reg
150929	150929	Italy	More Pinot Grigios should taste like this. A r...	NaN	90	15.0	Northeastern Italy	Alto Adige	

150930 rows × 11 columns

In [ ]:

#QUESTION 1

Set the index of the dataframe to country

In [3]:
```
wine_review2 = pd.read_csv('Wine_review.csv', index_col = 0)
wine_review2
```

Out[3]:

	country	description	designation	points	price	province	region_1	region_2	
0	US	This tremendous 100% varietal wine hails from ...	Martha's Vineyard	96	235.0	California	Napa Valley	Napa	Ca Sau
1	Spain	Ripe aromas of fig, blackberry and cassis are ...	Carodorum Selección Especial Reserva	96	110.0	Northern Spain	Toro	NaN	T
2	US	Mac Watson honors the memory of a wine once ma...	Special Selected Late Harvest	96	90.0	California	Knights Valley	Sonoma	Sau
3	US	This spent 20 months in 30% new French oak, an...	Reserve	96	65.0	Oregon	Willamette Valley	Willamette Valley	Pir
4	France	This is the top wine from La Bégude, named aft...	La Brûlade	95	66.0	Provence	Bandol	NaN	Pr re
...	...	...	...	...	...	...	...	...	
150925	Italy	Many people feel Fiano represents southern Ita...	NaN	91	20.0	Southern Italy	Fiano di Avellino	NaN	Whit
150926	France	Offers an intriguing nose with ginger, lime an...	Cuvée Prestige	91	27.0	Champagne	Champagne	NaN	Char
150927	Italy	This classic example comes from a cru vineyard...	Terre di Dora	91	20.0	Southern Italy	Fiano di Avellino	NaN	Whit
150928	France	A perfect salmon shade, with scents of peaches...	Grand Brut Rosé	90	52.0	Champagne	Champagne	NaN	Char

	country	description	designation	points	price	province	region_1	region_2	
150929	Italy	More Pinot Grigios should taste like this. A r...	NaN	90	15.0	Northeastern Italy	Alto Adige	NaN	Pino

150930 rows × 10 columns

In [4]:
```
wine_review.columns
```

Out[4]:
```
Index(['Unnamed: 0', 'country', 'description', 'designation', 'points',
       'price', 'province', 'region_1', 'region_2', 'variety', 'winery'],
      dtype='object')
```

In [ ]:
```

```

#Question 2

- Using the index and any country of your choice,
- show the reviews for those country Sample code
- df.loc[df.country == 'Italy']
- can you glean any insight from this?
- Perhaps the percentage of wines from the selected country.

In [5]:
```
#show the reviews for those country Sample code

Country_France = wine_review2.loc[wine_review2.country == 'France']
Country_France
```

Out[5]:

	country	description	designation	points	price	province	region_1	region_2	
4	France	This is the top wine from La Bégude, named aft...	La Brûlade	95	66.0	Provence	Bandol	NaN	Pro rec
13	France	This wine is in peak condition. The tannins an...	Château Montus Prestige	95	90.0	Southwest France	Madiran	NaN	
18	France	Coming from a seven-acre vineyard named after ...	Le Pigeonnier	95	290.0	Southwest France	Cahors	NaN	I
33	France	Pale in color, this is nutty in character, wit...	Nonpareil Trésor Rosé Brut	90	22.0	France Other	Vin Mousseux	NaN	Sp
36	France	Gingery spice notes accent fresh pear and melo...	NaN	90	60.0	Rhône Valley	Châteauneuf-du-Pape	NaN	Rhôn White
...	...	...	...	...	...	...	...	...	
150921	France	Shows some older notes: a bouquet of toasted w...	Blanc de Blancs Brut Mosaïque	91	38.0	Champagne	Champagne	NaN	Cham
150923	France	Rich and toasty, with tiny bubbles. The bouque...	Demi-Sec	91	30.0	Champagne	Champagne	NaN	Cham
150924	France	Really fine for a low-acid vintage, there's an...	Diamant Bleu	91	70.0	Champagne	Champagne	NaN	Cham
150926	France	Offers an intriguing nose with ginger, lime an...	Cuvée Prestige	91	27.0	Champagne	Champagne	NaN	Cham

	country	description	designation	points	price	province	region_1	region_2	v
150928	France	A perfect salmon shade, with scents of peaches...	Grand Brut Rosé	90	52.0	Champagne	Champagne	NaN	Cham

21098 rows × 10 columns

In [6]:
```
#percentage of wines from the selected country

Country_France['points_percentage'] = (Country_France['points']/Country_France['p
Country_France
```

```
C:\Users\user\AppData\Local\Temp\ipykernel_4308\2752624048.py:3: 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 (https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)
  Country_France['points_percentage'] = (Country_France['points']/Country_France['points'].sum())*100
```

Out[6]:

	country	description	designation	points	price	province	region_1	region_2	
4	France	This is the top wine from La Bégude, named aft...	La Brûlade	95	66.0	Provence	Bandol	NaN	
13	France	This wine is in peak condition. The tannins an...	Château Montus Prestige	95	90.0	Southwest France	Madiran	NaN	
18	France	Coming from a seven-acre vineyard named after ...	Le Pigeonnier	95	290.0	Southwest France	Cahors	NaN	
33	France	Pale in color, this is nutty in character, wit...	Nonpareil Trésor Rosé Brut	90	22.0	France Other	Vin Mousseux	NaN	
36	France	Gingery spice notes accent fresh pear and melo...	NaN	90	60.0	Rhône Valley	Châteauneuf-du-Pape	NaN	Rl W
...	...	...	...	...	...	...	...	...	
150921	France	Shows some older notes: a bouquet of toasted w...	Blanc de Blancs Brut Mosaïque	91	38.0	Champagne	Champagne	NaN	Ch
150923	France	Rich and toasty, with tiny bubbles. The bouque...	Demi-Sec	91	30.0	Champagne	Champagne	NaN	Ch

	country	description	designation	points	price	province	region_1	region_2	
150924	France	Really fine for a low-acid vintage, there's an...	Diamant Bleu	91	70.0	Champagne	Champagne	NaN	Ch
150926	France	Offers an intriguing nose with ginger, lime an...	Cuvée Prestige	91	27.0	Champagne	Champagne	NaN	Ch
150928	France	A perfect salmon shade, with scents of peaches...	Grand Brut Rosé	90	52.0	Champagne	Champagne	NaN	Ch

21098 rows × 11 columns

In [ ]:

#Question

- Supposing wines are reviewed on a 80-to-100 point scale,
- and wines with greater than 90 points are average,
- can you check what percentage of the total wines are above average?
- What about the percentage in the country you selected from question 2?

In [8]:
```
#wines greater than 90 points are average.

Average_wine = wine_review2.loc[wine_review2.points > 90]
Average_wine
```

Out[8]:

	country	description	designation	points	price	province	region_1	region_2	v
0	US	This tremendous 100% varietal wine hails from ...	Martha's Vineyard	96	235.0	California	Napa Valley	Napa	Ca Sau
1	Spain	Ripe aromas of fig, blackberry and cassis are ...	Carodorum Selección Especial Reserva	96	110.0	Northern Spain	Toro	NaN	T
2	US	Mac Watson honors the memory of a wine once ma...	Special Selected Late Harvest	96	90.0	California	Knights Valley	Sonoma	Sau
3	US	This spent 20 months in 30% new French oak, an...	Reserve	96	65.0	Oregon	Willamette Valley	Willamette Valley	Pin
4	France	This is the top wine from La Bégude, named aft...	La Brûlade	95	66.0	Provence	Bandol	NaN	Pro rec
...	...	...	...	...	...	...	...	...	
150923	France	Rich and toasty, with tiny bubbles. The bouque...	Demi-Sec	91	30.0	Champagne	Champagne	NaN	Cham
150924	France	Really fine for a low-acid vintage, there's an...	Diamant Bleu	91	70.0	Champagne	Champagne	NaN	Cham
150925	Italy	Many people feel Fiano represents southern Ita...	NaN	91	20.0	Southern Italy	Fiano di Avellino	NaN	White
150926	France	Offers an intriguing nose with ginger, lime an...	Cuvée Prestige	91	27.0	Champagne	Champagne	NaN	Cham

	country	description	designation	points	price	province	region_1	region_2	v
150927	Italy	This classic example comes from a cru vineyard...	Terre di Dora	91	20.0	Southern Italy	Fiano di Avellino	NaN	White

32237 rows × 10 columns

In [10]:
```
#percentage of the total wines are above average.

Average_wine['points_percentage'] = (Average_wine['points']/Average_wine['points'
Average_wine['price_percentage'] = (Average_wine['price']/Average_wine['price'].s
Average_wine
Average_wine
```

```
C:\Users\user\AppData\Local\Temp\ipykernel_4308\1096799435.py:3: SettingWithCop
yWarning:
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/sta
ble/user_guide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pyd
ata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-c
opy)
  Average_wine['points_percentage'] = (Average_wine['points']/Average_wine['poi
nts'].sum())*100
C:\Users\user\AppData\Local\Temp\ipykernel_4308\1096799435.py:4: SettingWithCop
yWarning:
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/sta
ble/user_guide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pyd
ata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-c
opy)
  Average_wine['price_percentage'] = (Average_wine['price']/Average_wine['pric
e'].sum())*100
```

Out[10]:

	country	description	designation	points	price	province	region_1	region_2	
0	US	This tremendous 100% varietal wine hails from ...	Martha's Vineyard	96	235.0	California	Napa Valley	Napa	S
1	Spain	Ripe aromas of fig, blackberry and cassis are ...	Carodorum Selección Especial Reserva	96	110.0	Northern Spain	Toro	NaN	
2	US	Mac Watson honors the memory of a wine once ma...	Special Selected Late Harvest	96	90.0	California	Knights Valley	Sonoma	S
3	US	This spent 20 months in 30% new French oak, an...	Reserve	96	65.0	Oregon	Willamette Valley	Willamette Valley	

	country	description	designation	points	price	province	region_1	region_2	
4	France	This is the top wine from La Bégude, named aft...	La Brûlade	95	66.0	Provence	Bandol	NaN	
...	...	...	...	...	...	...	...	...	
150923	France	Rich and toasty, with tiny bubbles. The bouque...	Demi-Sec	91	30.0	Champagne	Champagne	NaN	Ch
150924	France	Really fine for a low-acid vintage, there's an...	Diamant Bleu	91	70.0	Champagne	Champagne	NaN	Ch
150925	Italy	Many people feel Fiano represents southern Ita...	NaN	91	20.0	Southern Italy	Fiano di Avellino	NaN	Wl
150926	France	Offers an intriguing nose with ginger, lime an...	Cuvée Prestige	91	27.0	Champagne	Champagne	NaN	Ch
150927	Italy	This classic example comes from a cru vineyard...	Terre di Dora	91	20.0	Southern Italy	Fiano di Avellino	NaN	Wl

32237 rows × 12 columns

In [11]:
```
#percentage in the country you selected from question 2?

Country_France['price_percentage'] = (Country_France['price']/Country_France['pri
Country_France
```

```
C:\Users\user\AppData\Local\Temp\ipykernel_4308\2493061992.py:3: SettingWithCop
yWarning:
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/sta
ble/user_guide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pyd
ata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-c
opy)
  Country_France['price_percentage'] = (Country_France['price']/Country_France
['price'].sum())*100
```

Out[11]:

	country	description	designation	points	price	province	region_1	region_2	
4	France	This is the top wine from La Bégude, named aft...	La Brûlade	95	66.0	Provence	Bandol	NaN	
13	France	This wine is in peak condition. The tannins an...	Château Montus Prestige	95	90.0	Southwest France	Madiran	NaN	
18	France	Coming from a seven-acre vineyard named after ...	Le Pigeonnier	95	290.0	Southwest France	Cahors	NaN	
33	France	Pale in color, this is nutty in character, wit...	Nonpareil Trésor Rosé Brut	90	22.0	France Other	Vin Mousseux	NaN	
36	France	Gingery spice notes accent fresh pear and melo...	NaN	90	60.0	Rhône Valley	Châteauneuf-du-Pape	NaN	Rl W
...	...	...	...	...	...	...	...	...	
150921	France	Shows some older notes: a bouquet of toasted w...	Blanc de Blancs Brut Mosaïque	91	38.0	Champagne	Champagne	NaN	Ch
150923	France	Rich and toasty, with tiny bubbles. The bouque...	Demi-Sec	91	30.0	Champagne	Champagne	NaN	Ch

	country	description	designation	points	price	province	region_1	region_2	
150924	France	Really fine for a low-acid vintage, there's an...	Diamant Bleu	91	70.0	Champagne	Champagne	NaN	Ch
150926	France	Offers an intriguing nose with ginger, lime an...	Cuvée Prestige	91	27.0	Champagne	Champagne	NaN	Ch
150928	France	A perfect salmon shade, with scents of peaches...	Grand Brut Rosé	90	52.0	Champagne	Champagne	NaN	Ch

21098 rows × 12 columns

In [ ]:

#Question 4

- Check the pandas ‘isin’ operator
- and use it to perform any operation of your choice on the dataset

In [14]:
```
new = wine_review2.loc[(wine_review2.country.isin(['Italy']))]
new
```

Out[14]:

	country	description	designation	points	price	province	region_1	region_2	
10	Italy	Elegance, complexity and structure come togeth...	Ronco della Chiesa	95	80.0	Northeastern Italy	Collio	NaN	
32	Italy	Underbrush, scorched earth, menthol and plum s...	Vigna Piaggia	90	NaN	Tuscany	Brunello di Montalcino	NaN	Sa
35	Italy	Forest floor, tilled soil, mature berry and a ...	Riserva	90	135.0	Tuscany	Brunello di Montalcino	NaN	Sa
37	Italy	Aromas of forest floor, violet, red berry and ...	NaN	90	29.0	Tuscany	Vino Nobile di Montepulciano	NaN	Sa
38	Italy	This has a charming nose that boasts rose, vio...	NaN	90	23.0	Tuscany	Chianti Classico	NaN	Sa
...	...	...	...	...	...	...	...	...	
150920	Italy	Rich and mature aromas of smoke, earth and her...	Brut Riserva	91	19.0	Northeastern Italy	Trento	NaN	Cha
150922	Italy	Made by 30-ish Roberta Borghese high above Man...	Superiore	91	NaN	Northeastern Italy	Colli Orientali del Friuli	NaN	
150925	Italy	Many people feel Fiano represents southern Ita...	NaN	91	20.0	Southern Italy	Fiano di Avellino	NaN	Wh
150927	Italy	This classic example comes from a cru vineyard...	Terre di Dora	91	20.0	Southern Italy	Fiano di Avellino	NaN	Wh
150929	Italy	More Pinot Grigios should taste like this. A r...	NaN	90	15.0	Northeastern Italy	Alto Adige	NaN	Pir

23478 rows × 10 columns

In [ ]:

#Question 5

- Suppose we'll buy any wine that's made in Italy or which is rated above average.
- Can you select or show these wines using the or (|) operator?

In [18]:
```
buy_specific_wines = wine_review2.loc[(wine_review2.country.isin(['Italy'])) | (w
buy_specific_wines
```

Out[18]:

	country	description	designation	points	price	province	region_1	region_2
0	US	This tremendous 100% varietal wine hails from ...	Martha's Vineyard	96	235.0	California	Napa Valley	Napa
1	Spain	Ripe aromas of fig, blackberry and cassis are ...	Carodorum Selección Especial Reserva	96	110.0	Northern Spain	Toro	NaN
2	US	Mac Watson honors the memory of a wine once ma...	Special Selected Late Harvest	96	90.0	California	Knights Valley	Sonoma

In [15]:
```
buy_fine_wines = [new| (wine_review2.points >90)]
buy_fine_wines
```

```
   8870     )
   8871 elif isinstance(other, ABCSeries):
-> 8872     return self._align_series(
   8873         other,
   8874         join=join,
   8875         axis=axis,
   8876         level=level,
   8877         copy=copy,
   8878         fill_value=fill_value,
   8879         method=method,
   8880         limit=limit,
   8881         fill_axis=fill_axis,
   8882     )
   8883 else:  # pragma: no cover
   8884     raise TypeError(f"unsupported type: {type(other)}")

File ~\anaconda3\lib\site-packages\pandas\core\generic.py:8995, in NDFrame._a
lign_series(self, other, join, axis, level, copy, fill_value, method, limit,
 fill_axis)
   8993     if lidx is not None:
```

In [ ]:
```
```

#Question 6 Read about the notnull operator and use it to check for wines whose prices are not null

In [16]:
```
list = ({'Apples': [30, 15, 25, 14, 11], 'Bananas': [15, 18, 42, 22, 17]})
Fruits = pd.DataFrame(list, index = ['2017 sales', '2018 sales', '2019 sales', '2
print(Fruits)
```

```
            Apples  Bananas
2017 sales      30       15
2018 sales      15       18
2019 sales      25       42
2020 sales      14       22
2021 sales      11       17
```

In [17]:
```
Fruits.notnull()
```

Out[17]:

	Apples	Bananas
2017 sales	True	True
2018 sales	True	True
2019 sales	True	True
2020 sales	True	True
2021 sales	True	True

In [ ]:
```
```