Case Study 1: IPL Review

Uploadings:

The .ipynb file is uploaded on this link kindly go through the links:

https://drive.google.com/file/d/1hDIn5A8-Q93NVNcfOUmbwRojhBpixfdP/view?usp=sharing

Questions Asked:

- Q1. Question-1: Suppose in 'ipl18', you want to filter out the teams that have an NRR greater than zero, and for which the 'For' score exceeds the 'Against' score, i.e. both the conditions should be satisfied. Which teams will be left after you perform the above filtration?
- a) CSK, MI b) SRH,CSK, MI c) SRH,CSK, RCB d) SRK,CSK, MI,RCB
- Q2. If all the stats are taken for both 'ipl17' and 'ipl18', which team with its total points greater than 25 will have the highest win percentage

In [1]:

```
import numpy as np
import pandas as pd
```

In [2]:

Out[2]:

	Team	Matches	Won	Lost	Tied	N/R	Points	NRR	For	Against
1	SRH	14	9	5	0	0	18	0.284	2230	2193
2	CSK	14	9	5	0	0	18	0.253	2488	2433
3	KKR	14	8	6	0	0	16	-0.070	2363	2425
4	RR	14	7	7	0	0	14	-0.250	2130	2141
5	MI	14	6	8	0	0	12	0.317	2380	2282
6	RCB	14	6	8	0	0	12	0.129	2322	2383
7	KXIP	14	6	8	0	0	12	-0.502	2210	2259
8	DD	14	5	9	0	0	10	-0.222	2297	2304

In [3]:

Out[3]:

	Team	Matches	Won	Lost	Tied	N/R	Points	NRR	For	Against
1	MI	14	10	4	0	0	20	0.784	2407	2242
2	RPS	14	9	5	0	0	18	0.176	2180	2165
3	SRH	14	8	5	0	1	17	0.469	2221	2118
4	KKR	14	8	6	0	0	16	0.641	2329	2300
5	KXIP	14	7	7	0	0	14	0.123	2207	2229
6	DD	14	6	8	0	0	12	-0.512	2219	2255
7	GL	14	4	10	0	0	8	-0.412	2406	2472
8	RCB	14	3	10	0	1	7	-1.299	1845	2033

Question-1:

Suppose in 'ipl18', you want to filter out the teams that have an NRR greater than zero, and for which the 'For' score exceeds the 'Against' score, i.e. both the conditions should be satisfied. Which teams will be left after you perform the above filtration?**

a) CSK, MI b) SRH, CSK, MI c) SRH, CSK, RCB d) SRK, CSK, MI, RCB

Answer-1: b)

Solution 1:

```
In [4]:
```

```
ipl18
```

Out[4]:

	Team	Matches	Won	Lost	Tied	N/R	Points	NRR	For	Against
1	SRH	14	9	5	0	0	18	0.284	2230	2193
2	CSK	14	9	5	0	0	18	0.253	2488	2433
3	KKR	14	8	6	0	0	16	-0.070	2363	2425
4	RR	14	7	7	0	0	14	-0.250	2130	2141
5	MI	14	6	8	0	0	12	0.317	2380	2282
6	RCB	14	6	8	0	0	12	0.129	2322	2383
7	KXIP	14	6	8	0	0	12	-0.502	2210	2259
8	DD	14	5	9	0	0	10	-0.222	2297	2304

In [5]:

```
for x in range(1,8):
    if ipl18.loc[x].at['For']>ipl18.loc[x].at['Against']:
        if ipl18.loc[x].at['NRR']>0:
            print(ipl18.loc[x].at['Team'])
```

SRH CSK MI

Question-2:

If all the stats are taken for both 'ipl17' and 'ipl18', which team with its total points greater than 25 will have the highest win percentage

In [6]:

ipl18

Out[6]:

	Team	Matches	Won	Lost	Tied	N/R	Points	NRR	For	Against
1	SRH	14	9	5	0	0	18	0.284	2230	2193
2	CSK	14	9	5	0	0	18	0.253	2488	2433
3	KKR	14	8	6	0	0	16	-0.070	2363	2425
4	RR	14	7	7	0	0	14	-0.250	2130	2141
5	MI	14	6	8	0	0	12	0.317	2380	2282
6	RCB	14	6	8	0	0	12	0.129	2322	2383
7	KXIP	14	6	8	0	0	12	-0.502	2210	2259
8	DD	14	5	9	0	0	10	-0.222	2297	2304

In [7]:

ipl17

Out[7]:

	Team	Matches	Won	Lost	Tied	N/R	Points	NRR	For	Against
1	MI	14	10	4	0	0	20	0.784	2407	2242
2	RPS	14	9	5	0	0	18	0.176	2180	2165
3	SRH	14	8	5	0	1	17	0.469	2221	2118
4	KKR	14	8	6	0	0	16	0.641	2329	2300
5	KXIP	14	7	7	0	0	14	0.123	2207	2229
6	DD	14	6	8	0	0	12	-0.512	2219	2255
7	GL	14	4	10	0	0	8	-0.412	2406	2472
8	RCB	14	3	10	0	1	7	-1.299	1845	2033

In [8]:

```
total_points=ipl17['Points'].sum()+ipl18['Points'].sum()
```

In [9]:

```
for x in range(1,8):
    for y in range(1,8):
        if ipl18.loc[x].at['Team']==ipl17.loc[y].at['Team']:
              if ipl18.loc[x].at['Points']+ipl17.loc[y].at['Points']>25:
                   t=ipl18.loc[x].at['Points']+ipl17.loc[y].at['Points']
                   print((str)(ipl18.loc[x].at['Team'])+' has win percentage of '+(str)(round())
```

SRH has win percentage of 15.625% KKR has win percentage of 14.286% MI has win percentage of 14.286% KXIP has win percentage of 11.607%

Case Study 2: Wine Review

Uploadings:

The .ipynb file is uploaded on this link kindly go through the links:

https://drive.google.com/file/d/1O3zLohSNBZD2qc--JxSCrBjBvkQ-S8ui/view?usp=sharing

Questions Asked:

- 1. Look at an overview of your data.
- 2. Select the `description` column from `reviews` and assign the result to the variable `desc`.
- **3.** Select the first value from the description column of `**reviews**`, assigning it to variable `**first_description**`.
- **4.** Select the first row of data (the first record) from `**reviews**`, assigning it to the variable `**first_row**`.
- **5.** Select the first 10 values from the `description` column in `reviews`, assigning the result to variable `first_descriptions`.
- **6.** Select the records with index labels `1`, `2`, `3`, `5`, and `8`, assigning the result to the variable **`sample_reviews`.**
- 7. Create a variable `df` containing the `country`, `province`, `region_1`, and `region_2` columns of the records with the index labels `0`, `1`, `10`, and `100`.
- **8.** Create a variable `df` containing the `country` and `variety` columns of the first 100 records
- 9. Create a DataFrame `italian_wines` containing reviews of wines made in `Italy`.

In [1]:

```
import numpy as np
import pandas as pd
```

In [2]:

```
review=pd.read_csv('winemag-data-130k-v2.csv')
```

Ques-1: Overview of the DataFrame:

In [3]:

review

Out[3]:

	Unnamed: 0	country	description	designation	points	price	province	region_1	region_2	taster
0	0	Italy	Aromas include tropical fruit, broom, brimston	Vulkà Bianco	87	NaN	Sicily & Sardinia	Etna	NaN	(
1	1	Portugal	This is ripe and fruity, a wine that is smooth	Avidagos	87	15.0	Douro	NaN	NaN	Rog
2	2	US	Tart and snappy, the flavors of lime flesh and	NaN	87	14.0	Oregon	Willamette Valley	Willamette Valley	Paul
(•

In [4]:

review.shape

Out[4]:

(129971, 14)

In [5]:

```
review.columns
```

Out[5]:

In [6]:

review.head()

Out[6]:

Aromas include Vulkà 87 NaN Sicily & Etna NaN Sicily & Sardinia 1 Portugal This is ripe and fruity, a wine that is	
1 1 Portugal and fruity, a Avidagos 87 15.0 Douro NaN NaN	
smooth	N
Tart and snappy, the 2 2 US flavors of NaN 87 14.0 Oregon Willamette Willamette Willamette Walley Valley and	
Pineapple rind, lemon Reserve Lake 3 3 US pith and Late 87 13.0 Michigan Michigan NaN orange Harvest Shore blossom	١
Much like the regular Reserve 4 4 US bottling Wild Child from 2012, this Block	
4	

Ques-2:

In [7]:

```
desc=review['description']
desc
```

Out[7]:

0	Aromas include tropical fruit, broom, brimston
1	This is ripe and fruity, a wine that is smooth
2	Tart and snappy, the flavors of lime flesh and
3	Pineapple rind, lemon pith and orange blossom
4	Much like the regular bottling from 2012, this
	•••
129966	Notes of honeysuckle and cantaloupe sweeten th
129967	Citation is given as much as a decade of bottl
129968	Well-drained gravel soil gives this wine its c
129969	A dry style of Pinot Gris, this is crisp with
129970	Big, rich and off-dry, this is powered by inte
Name: des	cription, Length: 129971, dtype: object

Ques-3:

In [8]:

```
first_description=review['description'][0]
first_description
```

Out[8]:

"Aromas include tropical fruit, broom, brimstone and dried herb. The palate isn't overly expressive, offering unripened apple, citrus and dried sage alo ngside brisk acidity."

Ques-4:

In [9]:

```
first_row=review.loc[0]
first_row
```

Out[9]:

```
Unnamed: 0
                                                                            0
country
                                                                        Italy
                          Aromas include tropical fruit, broom, brimston...
description
designation
                                                                 Vulkà Bianco
points
                                                                           87
price
                                                                          NaN
province
                                                           Sicily & Sardinia
region_1
                                                                         Etna
                                                                          NaN
region_2
taster_name
                                                                Kerin O'Keefe
taster_twitter_handle
                                                                 @kerinokeefe
                                          Nicosia 2013 Vulkà Bianco (Etna)
title
variety
                                                                 White Blend
                                                                      Nicosia
winery
Name: 0, dtype: object
```

Ques-5:

```
In [10]:
```

```
first_descriptions=review['description'].head(10)
first_descriptions
```

Out[10]:

```
Aromas include tropical fruit, broom, brimston...
0
1
     This is ripe and fruity, a wine that is smooth...
     Tart and snappy, the flavors of lime flesh and...
2
3
     Pineapple rind, lemon pith and orange blossom ...
4
     Much like the regular bottling from 2012, this...
     Blackberry and raspberry aromas show a typical...
5
     Here's a bright, informal red that opens with ...
6
     This dry and restrained wine offers spice in p...
7
     Savory dried thyme notes accent sunnier flavor...
Я
     This has great depth of flavor with its fresh ...
Name: description, dtype: object
```

Ques-6:

In [11]:

```
arr=review.values review.columns
```

Out[11]:

In [12]:

```
numpy_arr=np.array([arr[1],arr[2],arr[3],arr[5],arr[8]])
numpy_arr
Out[12]:
array([[1, 'Portugal',
```

"This is ripe and fruity, a wine that is smooth while still structur ed. Firm tannins are filled out with juicy red berry fruits and freshened wi th acidity. It's already drinkable, although it will certainly be better fr om 2016.", 'Avidagos', 87, 15.0, 'Douro', nan, nan, 'Roger Voss', '@vossroger', 'Quinta dos Avidagos 2011 Avidagos Red (Douro)', 'Portuguese Red', 'Quinta dos Avidagos'], [2, 'US', Tart and snappy, the flavors of lime flesh and rind dominate. Some green pineapple pokes through, with crisp acidity underscoring the flavors. The wine was all stainless-steel fermented.', nan, 87, 14.0, 'Oregon', 'Willamette Valley', 'Willamette Valley', 'Paul Gregutt', '@paulgwine\xa0', 'Rainstorm 2013 Pinot Gris (Willamette Valley)', 'Pinot Gris', 'Rainstorm'], [3, 'US', 'Pineapple rind, lemon pith and orange blossom start off the aromas. The palate is a bit more opulent, with notes of honey-drizzled guava and man go giving way to a slightly astringent, semidry finish.', 'Reserve Late Harvest', 87, 13.0, 'Michigan', 'Lake Michigan Shore', nan, 'Alexander Peartree', nan, 'St. Julian 2013 Reserve Late Harvest Riesling (Lake Michigan Shor e)', 'Riesling', 'St. Julian'], [5, 'Spain', 'Blackberry and raspberry aromas show a typical Navarran whiff of gr een herbs and, in this case, horseradish. In the mouth, this is fairly full bodied, with tomatoey acidity. Spicy, herbal flavors complement dark plum fr uit, while the finish is fresh but grabby.', 'Ars In Vitro', 87, 15.0, 'Northern Spain', 'Navarra', nan, 'Michael Schachner', '@wineschach', 'Tandem 2011 Ars In Vitro Tempranillo-Merlot (Navarra)', 'Tempranillo-Merlot', 'Tandem'], [8, 'Germany', "Savory dried thyme notes accent sunnier flavors of preserved peach in this brisk, off-dry wine. It's fruity and fresh, with an elegant, spright ly footprint.", 'Shine', 87, 12.0, 'Rheinhessen', nan, nan, 'Anna Lee C. Iijima', nan, 'Heinz Eifel 2013 Shine Gewürztraminer (Rheinhessen)',

'Gewürztraminer', 'Heinz Eifel']], dtype=object)

In [13]:

```
col=review.columns
sample_reviews=pd.DataFrame(data=numpy_arr,columns=col)
sample_reviews
```

Out[13]:

	Unnamed: 0	country	description	designation	points	price	province	region_1	region_2	taster_na
0	1	Portugal	This is ripe and fruity, a wine that is smooth	Avidagos	87	15.0	Douro	NaN	NaN	Roger V
1	2	US	Tart and snappy, the flavors of lime flesh and	NaN	87	14.0	Oregon	Willamette Valley	Willamette Valley	Paul Greç
2	3	US	Pineapple rind, lemon pith and orange blossom	Reserve Late Harvest	87	13.0	Michigan	Lake Michigan Shore	NaN	Alexan Peart
4										•

Ques-7:

In [14]:

review

Out[14]:

	Unnamed: 0	country	description	designation	points	price	province	region_1	region_2	tastei
0	0	Italy	Aromas include tropical fruit, broom, brimston	Vulkà Bianco	87	NaN	Sicily & Sardinia	Etna	NaN	(
1	1	Portugal	This is ripe and fruity, a wine that is smooth	Avidagos	87	15.0	Douro	NaN	NaN	Rog
2	2	US	Tart and snappy, the flavors of lime flesh and	NaN	87	14.0	Oregon	Willamette Valley	Willamette Valley	Paul
4										•

In [15]:

```
def select_cols(df, colm):
   new = df.loc[:, colm]
   return new
```

In [16]:

```
colm=['country','province','region_1','region_2']
df=select_cols(review,colm)
df
```

Out[16]:

	country	province	region_1	region_2
0	Italy	Sicily & Sardinia	Etna	NaN
1	Portugal	Douro	NaN	NaN
2	US	Oregon	Willamette Valley	Willamette Valley
3	US	Michigan	Lake Michigan Shore	NaN
4	US	Oregon	Willamette Valley	Willamette Valley
129966	Germany	Mosel	NaN	NaN
129967	US	Oregon	Oregon	Oregon Other
129968	France	Alsace	Alsace	NaN
129969	France	Alsace	Alsace	NaN
129970	France	Alsace	Alsace	NaN

129971 rows × 4 columns

Ques-8:

In [17]:

```
def select_cols(df, colm):
   new = df.loc[:, colm]
   return new
```

In [18]:

```
col_span=['country','variety']
dp1=select_cols(review,col_span)
dp1=dp1.head(100)
dp1
```

Out[18]:

	country	variety
0	Italy	White Blend
1	Portugal	Portuguese Red
2	US	Pinot Gris
3	US	Riesling
4	US	Pinot Noir
95	France	Gamay
96	France	Gamay
97	US	Riesling
98	Italy	Sangiovese
99	US	Bordeaux-style Red Blend

100 rows × 2 columns

Ques-9:

In [19]:

review.head()

Out[19]:

	Unnamed: 0	country	description	designation	points	price	province	region_1	region_2
0	0	Italy	Aromas include tropical fruit, broom, brimston	Vulkà Bianco	87	NaN	Sicily & Sardinia	Etna	NaN
1	1	Portugal	This is ripe and fruity, a wine that is smooth	Avidagos	87	15.0	Douro	NaN	NaN
2	2	US	Tart and snappy, the flavors of lime flesh and	NaN	87	14.0	Oregon	Willamette Valley	Willamette Valley
3	3	US	Pineapple rind, lemon pith and orange blossom	Reserve Late Harvest	87	13.0	Michigan	Lake Michigan Shore	NaN
4	4	US	Much like the regular bottling from 2012, this	Vintner's Reserve Wild Child Block	87	65.0	Oregon	Willamette Valley	Willamette Valley

localhost:8888/notebooks/Wine Review/Module2.ipynb

```
In [20]:
```

```
arry=[]
for x in range(0,len(review)-1):
    if(review.loc[x].at['country']=='Italy'):
        arry.append(review.loc[x].values)
arry
Out[20]:
```

```
[array([0, 'Italy',
```

"Aromas include tropical fruit, broom, brimstone and dried herb. T he palate isn't overly expressive, offering unripened apple, citrus and dr ied sage alongside brisk acidity.",

```
'Vulkà Bianco', 87, nan, 'Sicily & Sardinia', 'Etna', nan,
       'Kerin O'Keefe', '@kerinokeefe',
       'Nicosia 2013 Vulkà Bianco (Etna)', 'White Blend', 'Nicosia'],
      dtype=object),
array([6, 'Italy',
```

"Here's a bright, informal red that opens with aromas of candied b erry, white pepper and savory herb that carry over to the palate. It's bal anced with fresh acidity and soft tannins.",

```
'Belsito', 87, 16.0, 'Sicily & Sardinia', 'Vittoria', nan,
       'Kerin O'Keefe', '@kerinokeefe',
       'Terre di Giurfo 2013 Belsito Frappato (Vittoria)', 'Frappato',
       'Terre di Giurfo'], dtype=object),
array([13, 'Italy',
```

'This is dominated by oak and oak-driven aromas that include roast

In [21]:

len(arry)

Out[21]:

19540

```
In [22]:
```

```
numpy_array=np.array(arry)
numpy_array
```

Out[22]:

```
array([[0, 'Italy',
```

"Aromas include tropical fruit, broom, brimstone and dried herb. The palate isn't overly expressive, offering unripened apple, citrus and dried s age alongside brisk acidity.",

```
..., 'Nicosia 2013 Vulkà Bianco (Etna)', 'White Blend',
'Nicosia'],
[6, 'Italy',
```

"Here's a bright, informal red that opens with aromas of candied ber ry, white pepper and savory herb that carry over to the palate. It's balance d with fresh acidity and soft tannins.",

```
..., 'Terre di Giurfo 2013 Belsito Frappato (Vittoria)',
'Frappato', 'Terre di Giurfo'],
[13, 'Italy',
```

'This is dominated by oak and oak-driven aromas that include roasted coffee bean, espresso, coconut and vanilla that carry over to the palate, to gether with plum and chocolate. Astringent, drying tannins give it a rather abrupt finish.',

```
..., 'Masseria Setteporte 2012 Rosso (Etna)',
'Nerello Mascalese', 'Masseria Setteporte'],
...,
[129947, 'Italy',
```

'A blend of 65% Cabernet Sauvignon, 30% Merlot and 5% Petit Verdot, this bold wine features aromas of spiced plum, blackcurrant, leather and a h int of coconut. On the palate, assertive tannins support the mature blackber ry, espresso and licorice flavors.',

```
'Feudo Principi di Butera 2012 Symposio Red (Terre Siciliane)', 'Red Blend', 'Feudo Principi di Butera'], [129961, 'Italy',
```

'Intense aromas of wild cherry, baking spice, tilled soil and savory herb lead the nose on this soulful, silky red. The round, smooth palate dole s out juicy red cherry, strawberry jelly, mineral, white pepper and an intri guing note of zabaglione alongside soft, supple tannins and bright acidit y...',

```
..., 'COS 2013 Frappato (Sicilia)', 'Frappato', 'COS'],
[129962, 'Italy',
```

'Blackberry, cassis, grilled herb and toasted aromas come together in the glass. On the palate, espresso, mint and black pepper add depth to the core of black cherry and blackberry flavors. It finishes on a licorice not e.',

```
"Cusumano 2012 Sàgana Tenuta San Giacomo Nero d'Avola (Sicilia)", "Nero d'Avola", 'Cusumano']], dtype=object)
```

In [23]:

```
temp_dp=pd.DataFrame(data=numpy_array,columns=col)
temp_dp
```

Out[23]:

	Unnamed: 0	country	description	designation	points	price	province	region_1	region_2	taster_nar
0	0	Italy	Aromas include tropical fruit, broom, brimston	Vulkà Bianco	87	NaN	Sicily & Sardinia	Etna	NaN	Ke O'Kee
1	6	Italy	Here's a bright, informal red that opens with	Belsito	87	16.0	Sicily & Sardinia	Vittoria	NaN	Ke O'Ke∉
2	13	Italy	This is dominated by oak and oak-driven	Rosso	87	NaN	Sicily & Sardinia	Etna	NaN	Ke O'Kee
4										•

In [24]:

```
def select_cols(df, colm):
   new = df.loc[:, colm]
   return new
```

In [25]:

col_new=['description','price','points','taster_name','taster_twitter_handle','title','vari
last_dp=select_cols(temp_dp,col_new)
last_dp

Out[25]:

	description	price	points	taster_name	taster_twitter_handle	title	variety	
0	Aromas include tropical fruit, broom, brimston	NaN	87	Kerin O'Keefe	@kerinokeefe	Nicosia 2013 Vulkà Bianco (Etna)	White Blend	
1	Here's a bright, informal red that opens with	16.0	87	Kerin O'Keefe	@kerinokeefe	Terre di Giurfo 2013 Belsito Frappato (Vittoria)	Frappato	
2	This is dominated by oak and oak-driven aromas	NaN	87	Kerin O'Keefe	@kerinokeefe	Masseria Setteporte 2012 Rosso (Etna)	Nerello Mascalese	:
3	Delicate aromas recall white flower and citrus	19.0	87	Kerin O'Keefe	@kerinokeefe	Baglio di Pianetto 2007 Ficiligno White (Sicilia)	White Blend	
4	Aromas of prune, blackcurrant, toast and oak c	35.0	87	Kerin O'Keefe	@kerinokeefe	Canicattì 2009 Aynat Nero d'Avola (Sicilia)	Nero d'Avola	
19535	This luminous sparkler has a sweet, fruit-forw	38.0	91	NaN	NaN	Col Vetoraz Spumanti NV Prosecco Superiore di	Prosecco	Cı
19536	A blend of Nero d'Avola and Syrah, this convey	29.0	90	Kerin O'Keefe	@kerinokeefe	Baglio del Cristo di Campobello 2012 Adènzia R	Red Blend	C٤
19537	A blend of 65% Cabernet Sauvignon, 30% Merlot 	20.0	90	Kerin O'Keefe	@kerinokeefe	Feudo Principi di Butera 2012 Symposio Red (Te	Red Blend	
19538	Intense aromas of wild cherry, baking spice, t	30.0	90	Kerin O'Keefe	@kerinokeefe	COS 2013 Frappato (Sicilia)	Frappato	

	description	price	points	taster_name	taster_twitter_handle	title	variety	
1953	Blackberry, cassis, grilled herb and toasted a	40.0	90	Kerin O'Keefe	@kerinokeefe	Cusumano 2012 Sàgana Tenuta San Giacomo Nero d	Nero d'Avola	C

19540 rows × 8 columns