

```
In [49]: import pandas as pd
import numpy as np

dictionary = {"Day":["D1", "D2", "D3", "D4", "D5", "D6", "D7", "D8", "D9", "D10", "D11", "D12", "D13", "D14"],
             "OUTLOOK": ["Sunny", "Sunny", "Overcast", "Rain", "Rain", "Rain", "Overcast", "Sunny", "Sunny",
                        "Rain", "Sunny", "Overcast", "Overcast", "Rain"],
             "Temperature":["Hot", "Hot", "Hot", "Mild", "Cool", "Cool", "Cool", "Cool", "Mild", "Cool",
                          "Mild", "Mild", "Mild", "Hot", "Mild"],
             "Humidity": ["High", "High", "High", "High", "Normal", "Normal", "Normal", "High", "Normal",
                        "Normal", "Normal", "High", "Normal", "High"],
             "Wind":["Weak", "Strong", "Weak", "Weak", "Weak", "Strong", "Strong", "Weak", "Weak", "Weak",
                   "Strong", "Strong", "Weak", "Strong"],
             "PlayTennis":["No", "No", "Yes", "Yes", "Yes", "No", "Yes", "No", "Yes", "Yes", "Yes",
                          "Yes", "Yes", "No"]

df = pd.DataFrame(dictionary)
df
```

Out[49]:

	Day	OUTLOOK	Temperature	Humidity	Wind	PlayTennis
0	D1	Sunny	Hot	High	Weak	No
1	D2	Sunny	Hot	High	Strong	No
2	D3	Overcast	Hot	High	Weak	Yes
3	D4	Rain	Mild	High	Weak	Yes
4	D5	Rain	Cool	Normal	Weak	Yes
5	D6	Rain	Cool	Normal	Strong	No
6	D7	Overcast	Cool	Normal	Strong	Yes
7	D8	Sunny	Mild	High	Weak	No
8	D9	Sunny	Cool	Normal	Weak	Yes
9	D10	Rain	Mild	Normal	Weak	Yes
10	D11	Sunny	Mild	Normal	Strong	Yes
11	D12	Overcast	Mild	High	Strong	Yes
12	D13	Overcast	Hot	Normal	Weak	Yes
13	D14	Rain	Mild	High	Strong	No

```
In [50]: part1 = len(df.loc[(df.PlayTennis=="Yes")&(df.Humidity == "High"), :])
part2 = len(df.loc[(df.PlayTennis=="Yes")&(df.Humidity == "Normal"), :])

len_df = len(df)
def frac(a, b):
    return r'$\frac{' + str(a) + '}{ ' + str(b) + '}$'
```

Question 1

$$\sum_i p(\text{PlayTennis} = \text{Yes} \cap \text{Humidity}_i) = p(\text{PlayTennis} = \text{Yes} \cap \text{Humidity} = \text{High}) + p(\text{PlayTennis} = \text{Yes} \cap \text{Humidity} = \text{Normal})$$

$$= \frac{3}{14} + \frac{6}{14} = \frac{9}{14}$$

```
In [51]: #yes
Result_yes1 = len(df.loc[(df.OUTLOOK=="Sunny")&(df.PlayTennis == "Yes"), :])
Result_yes2 = len(df.loc[(df.Humidity=="Normal")&(df.PlayTennis == "Yes"), :])
Result_yes3 = len(df.loc[(df.PlayTennis=="Yes"), :])

#no
Result_no1 = len(df.loc[(df.OUTLOOK=="Sunny")&(df.PlayTennis == "No"), :])
Result_no2 = len(df.loc[(df.Humidity=="Normal")&(df.PlayTennis == "No"), :])
Result_no3 = len(df.loc[(df.PlayTennis=="No"), :])
```

Question 2

- Classify: (Outlook = Sunny, Humidity = Normal)

$$\bullet \text{Result}_{\text{yes}} = \frac{p(\text{Outlook}=\text{Sunny} \cap \text{PlayTennis}=\text{Yes})}{p(\text{PlayTennis}=\text{Yes})} \times \frac{p(\text{Humidity}=\text{Normal} \cap \text{PlayTennis}=\text{Yes})}{p(\text{PlayTennis}=\text{Yes})} \times p(\text{PlayTennis} = \text{Yes})$$

$$\text{Result}_{\text{yes}} = \frac{2}{9} \times \frac{6}{9} \times \frac{9}{14} = \frac{12}{126} = \frac{2}{21}$$

$$\bullet \text{Result}_{\text{no}} = \frac{p(\text{Outlook}=\text{Sunny} \cap \text{PlayTennis}=\text{No})}{p(\text{PlayTennis}=\text{No})} \times \frac{p(\text{Humidity}=\text{Normal} \cap \text{PlayTennis}=\text{No})}{p(\text{PlayTennis}=\text{No})} \times p(\text{PlayTennis} = \text{No})$$

$$\text{Result}_{\text{No}} = \frac{3}{5} \times \frac{1}{5} \times \frac{5}{14} = \frac{3}{70}$$

In []: