Program 22

October 25, 2022

1 PROGRAM 22

Aim : a)Load iris dataset into a dataframe b)Take its descriptive statistics c)delete all rows which have null values d)select the Species classes,Count the no.of datas in each class e)Map the species classes into 1,2 and 3 f)Print the standard deviation and mean of petal length g)select all columns except species ##### Date : 31/08/2022 ##### By : Anu C Scharia

```
[29]: import pandas as pd
      df=pd.read_csv('iris.csv')
      df=pd.DataFrame(df)
      print(df)
```

```
     sepal_length  sepal_width  petal_length  petal_width    species
0             5.1          3.5           1.4          0.2     setosa
1             4.9          3.0           1.4          0.2     setosa
2             4.7          3.2           1.3          0.2     setosa
3             4.6          3.1           1.5          0.2     setosa
4             5.0          3.6           1.4          0.2     setosa
..            ...          ...           ...          ...        ...
145           6.7          3.0           5.2          2.3  virginica
146           6.3          2.5           5.0          1.9  virginica
147           6.5          3.0           5.2          2.0  virginica
148           6.2          3.4           5.4          2.3  virginica
149           5.9          3.0           5.1          1.8  virginica

[150 rows x 5 columns]
```

```
[16]: df.describe()
```

```
[16]:        sepal_length  sepal_width  petal_length  petal_width
count    150.000000   150.000000    150.000000   150.000000
mean       5.843333     3.054000      3.758667     1.198667
std        0.828066     0.433594      1.764420     0.763161
min        4.300000     2.000000      1.000000     0.100000
25%        5.100000     2.800000      1.600000     0.300000
50%        5.800000     3.000000      4.350000     1.300000
75%        6.400000     3.300000      5.100000     1.800000
max        7.900000     4.400000      6.900000     2.500000
```

```
[5]: #df.isnull().sum()
```

```
[5]: sepal_length    0
     sepal_width     0
     petal_length    0
     petal_width     0
     species         0
     dtype: int64
```

```
[6]: df=df.dropna()
     print(df)
```

```
     sepal_length  sepal_width  petal_length  petal_width    species
0             5.1          3.5           1.4          0.2     setosa
1             4.9          3.0           1.4          0.2     setosa
2             4.7          3.2           1.3          0.2     setosa
3             4.6          3.1           1.5          0.2     setosa
4             5.0          3.6           1.4          0.2     setosa
..            ...          ...           ...          ...        ...
145           6.7          3.0           5.2          2.3  virginica
146           6.3          2.5           5.0          1.9  virginica
147           6.5          3.0           5.2          2.0  virginica
148           6.2          3.4           5.4          2.3  virginica
149           5.9          3.0           5.1          1.8  virginica

[150 rows x 5 columns]
```

```
[7]: df.value_counts("species")
```

```
[7]: species
     setosa        50
     versicolor    50
     virginica     50
     dtype: int64
```

```
[30]: df['species']=df['species'].map({'setosa':1,'versicolor':2,'virginica':3})
      print(df)
```

```
     sepal_length  sepal_width  petal_length  petal_width  species
0             5.1          3.5           1.4          0.2        1
1             4.9          3.0           1.4          0.2        1
2             4.7          3.2           1.3          0.2        1
3             4.6          3.1           1.5          0.2        1
4             5.0          3.6           1.4          0.2        1
..            ...          ...           ...          ...      ...
145           6.7          3.0           5.2          2.3        3
146           6.3          2.5           5.0          1.9        3
147           6.5          3.0           5.2          2.0        3
```

```
148             6.2         3.4         5.4         2.3         3
149             5.9         3.0         5.1         1.8         3

[150 rows x 5 columns]
```

[24]:
```python
mean=df["petal_length"].mean()
std=df["petal_length"].std()
print("mean is:",mean)
print("Standard deviation :",std)
```

```
mean is: 3.7586666666666693
Standard deviation : 1.7644204199522617
```

[31]:
```python
df1=df.loc[:,df.columns!="species"]
print(df1)
```

```
     sepal_length  sepal_width  petal_length  petal_width
0             5.1          3.5           1.4          0.2
1             4.9          3.0           1.4          0.2
2             4.7          3.2           1.3          0.2
3             4.6          3.1           1.5          0.2
4             5.0          3.6           1.4          0.2
..            ...          ...           ...          ...
145           6.7          3.0           5.2          2.3
146           6.3          2.5           5.0          1.9
147           6.5          3.0           5.2          2.0
148           6.2          3.4           5.4          2.3
149           5.9          3.0           5.1          1.8

[150 rows x 4 columns]
```

[ ]: