3/20/25, 1:21 PM ASSIGNMENT3

ASSIGNMENT 3

Descriptive Statistics –

Measures of Central Tendency and variability Perform the following operations on any open source dataset (e.g., data.csv)

- 1. Provide summary statistics (mean, median, minimum, maximum, standard deviation) for a dataset (age, income etc.) with numeric variables grouped by one of the qualitative (categorical) variable. For example, if your categorical variable is age groups and quantitative variable is income, then provide summary statistics of income grouped by the age groups. Create a list that contains a numeric value for each response to the categorical variable.
- 2. Write a Python program to display some basic statistical details like percentile, mean, standard deviation etc. of the species of 'Iris-setosa', 'Iris-versicolor' and 'Iris-versicolor' of iris.csv dataset. Provide the codes with outputs and explain everything that you do in this step.

```
import numpy as np
In [2]:
         import pandas as pd
         import matplotlib.pyplot as plt
         import seaborn as sns
        iris = sns.load_dataset('iris')
In [3]:
        iris.head()
In [4]:
Out[4]:
            sepal_length sepal_width petal_length petal_width species
         0
                      5.1
                                   3.5
                                                 1.4
                                                              0.2
                                                                    setosa
         1
                      4.9
                                   3.0
                                                              0.2
                                                 1.4
                                                                    setosa
         2
                      4.7
                                   3.2
                                                 1.3
                                                              0.2
                                                                    setosa
         3
                      4.6
                                   3.1
                                                 1.5
                                                              0.2
                                                                    setosa
                                                              0.2
         4
                      5.0
                                   3.6
                                                 1.4
                                                                    setosa
        iris.describe()
In [5]:
```

3/20/25, 1:21 PM ASSIGNMENT3

Out[5]:		sepal_length	sepal_width	petal_length	petal_width				
	count	150.000000	150.000000	150.000000	150.000000				
	mean	5.843333	3.057333	3.758000	1.199333				
	std	0.828066	0.435866	1.765298	0.762238				
	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				
. [6].	inic do	senibo/inel	uda labiaati	\					
n [6]:	iris.de	<pre>iris.describe(include='object')</pre>							
ut[6]:		species							
	count	150							
	unique	3							
	top	setosa							
	freq	50							
[7]:	iris_gr	roupby = iri	s.groupby(by	='species')					
[8]:	iris_gr	oupby.std()							
ut[8]:		sepal_len	gth sepal_wi	dth petal_leng	th petal_wid				
	speci	es							
	seto	sa 0.352	490 0.379	0.1736	64 0.1053				
	versicol	or 0.516	171 0.313	798 0.4699	11 0.1977				
	virgini	ca 0.635	880 0.322	497 0.5518	95 0.2746				
n [9]:	inic an								
[] .	II.IZ_gr	oupby.mean()						
	TLT2 BL			dth petal_leng	th petal_wid				
	speci	sepal_len		dth petal_leng	th petal_wid				
out[9]:		sepal_len	gth sepal_wi	dth petal_leng					
	speci	sepal_len es sa 5.	gth sepal_wid		62 0.2				
	speci seto	sepal_len es sa 5. or 5.	gth sepal_wid	428 1.4	62 0.2 60 1.3				

3/20/25, 1:21 PM ASSIGNMENT3

In [10]:	iris_group	pby.median()							
Out[10]:		sepal_length	sepal_width	petal_length	petal_width				
	species								
	setosa	5.0	3.4	1.50	0.2				
	versicolor	5.9	2.8	4.35	1.3				
	virginica	6.5	3.0	5.55	2.0				
In [11]:]: iris_groupby.min()								
Out[11]:		senal length	senal width	petal_length	netal width				
00.0[]	species	sepai_leligtii	sepai_wiatii	petai_ieiigtii	petai_wiatii				
	setosa	4.3	2.3	1.0	0.1				
	versicolor	4.9	2.0	3.0	1.0				
	virginica	4.9	2.2	4.5	1.4				
	3	5		1.5					
	<pre>iris_groupby.max()</pre>								
In [12]:	iris_group	pby.max()							
	iris_group		sepal_width	petal_length	petal_width				
	iris_group		sepal_width	petal_length	petal_width				
			sepal_width	petal_length	petal_width				
<pre>In [12]: Out[12]:</pre>	species	sepal_length	•						
	species setosa	sepal_length	4.4	1.9	0.6				
	species setosa versicolor virginica	5.8 7.0 7.9	4.4 3.4 3.8	1.9	0.6				
Out[12]:	species setosa versicolor virginica	sepal_length 5.8 7.0	4.4 3.4 3.8	1.9	0.6				
Out[12]: In [13]:	species setosa versicolor virginica	sepal_length 5.8 7.0 7.9 pby.quantile(4.4 3.4 3.8	1.9	0.6 1.8 2.5				
	species setosa versicolor virginica	sepal_length 5.8 7.0 7.9 pby.quantile(4.4 3.4 3.8	1.9 5.1 6.9	0.6 1.8 2.5				
Out[12]: In [13]:	species setosa versicolor virginica iris_grou	sepal_length 5.8 7.0 7.9 pby.quantile(4.4 3.4 3.8	1.9 5.1 6.9	0.6 1.8 2.5				
Out[12]: In [13]:	species setosa versicolor virginica iris_grou	5.8 7.0 7.9 pby.quantile(sepal_length	4.4 3.4 3.8) sepal_width	1.9 5.1 6.9 petal_length	0.6 1.8 2.5 petal_width				