Experiment 10

Aim: Data Sampling and Stratification: Implement data sampling techniques to generate representative subsets of large datasets, and stratify the data based on specific criteria for balanced sampling.

Dataset Sampling and Visualization:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
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

	target
0	0
1	0
2	0
3	0
4	0

Fig 1: Dataset Printing

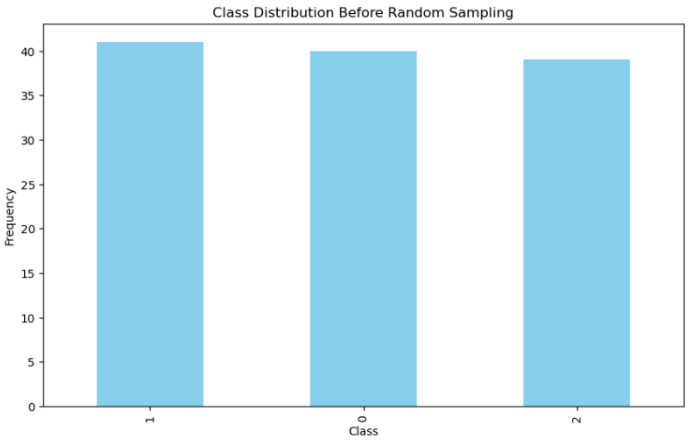

Fig 2: Class Distribution before Random Sampling

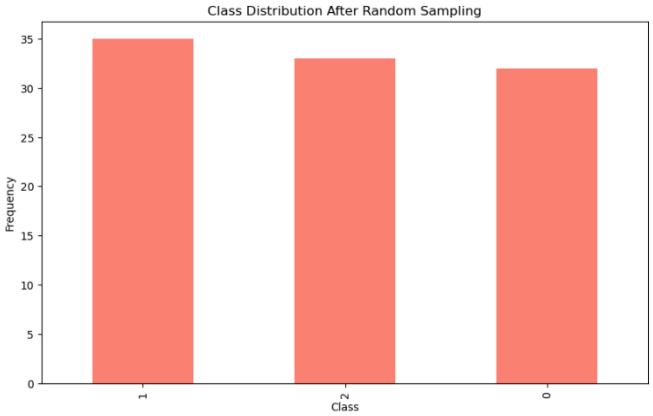

Fig 3: Class Distribution after Random Sampling

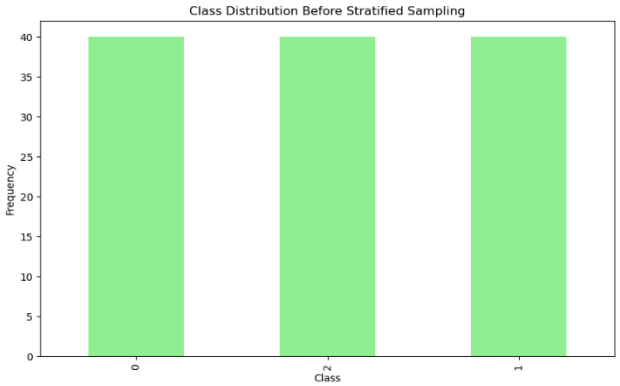

Fig 4: Class Distribution before Stratified Sampling

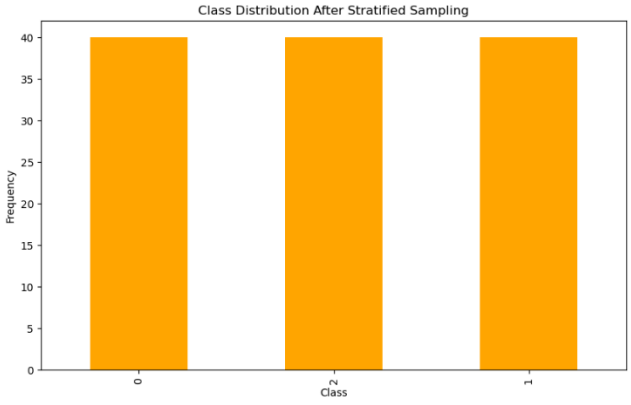

Fig 5: Class Distribution after Stratified Sampling

```
Original Data - Variance:
sepal length (cm)     0.685694
sepal width (cm)      0.189979
petal length (cm)     3.116278
petal width (cm)      0.581006
target                0.671141
dtype: float64

Resampled Data - Variance:
sepal length (cm)     0.605918
sepal width (cm)      0.234637
petal length (cm)     3.071572
petal width (cm)      0.586622
dtype: float64
```

Fig 6: Variance of the Dataset