

CLASS XII - IP (2020-2021)

BBPSPP

Chapter 2 - Data Visualisation using Pyplot

By Beena Nair

BBPSPP

LEARNING OBJECTIVES

This presentation will help you to analyse and comprehend about the following topics:

1. **Histogram/Frequency Distribution**
2. **Frequency Polygon**
3. **Box Plot**

- Histogram charts are a graphical display of frequencies represented as bars.
- Bar graph are usually represented using rectangular bars with lengths proportional to the values that they represent.

- Histogram charts can be drawn using `hist()` function.
- A histogram is used to represent the frequency of various values in a data set using bars , also known as bins. The bins are usually specified as consecutive, non-overlapping intervals of a variable.

Watch this video to understand the way
to create histogram using
MATPLOTLIB -

https://www.youtube.com/watch?list=RDCMUC79Gv3mYp6zKiSwYemEik9A&v=snkkKrek7TU&feature=emb_rel_end

To plot a histogram

```
Import matplotlib.pyplot as plt
```

```
h=[1,11,24,94,1,52,2,63,74,46,52,32,58,43, 98, 85, 77]
```

```
plt.hist(h)
```

```
plt.xlabel("value")
```

```
plt.ylabel("frequency")
```

```
plt.title("HISTOGRAM")
```

```
plt.show()
```

To plot a histogram with bin=25

```
Import matplotlib.pyplot as plt
```

```
h=[1,11,24,94,1,52,2,63,74,46,52,32,58,43, 98, 85, 77]
```

```
plt.hist(h, bins=25, color="red", edgecolor="black")
```

```
plt.xlabel("value")
```

```
plt.ylabel("frequency")
```

```
plt.title("HISTOGRAM")
```

```
plt.show()
```

Try yourself :

https://www.youtube.com/watch?list=RDCMUC79Gv3mYp6zKiSwYemEik9A&v=snkkKrek7TU&feature=emb_rel_end

- **Frequency Polygon is a graphical device for understanding the shapes of distribution.**
- **If we connect the top center points of each bin then we obtain relative frequency polygon.**

- The `hist()` function is used to create frequency polygon for a given set of data set. The argument `histtype='step'` is used for creating frequency polygon.

Steps to Draw a Frequency Polygon

- Mark the class intervals for each class on the horizontal axis. We will plot the frequency on the vertical axis.
- Calculate the classmark for each class interval. The formula for class mark is:

$$\text{Classmark} = (\text{Upper limit} + \text{Lower limit}) / 2$$

- Mark all the class marks on the horizontal axis. It is also known as the mid-value of every class.

- Corresponding to each class mark, plot the frequency as given to you. The height always depicts the frequency. Make sure that the frequency is plotted against the class mark and not the upper or lower limit of any class.
- Join all the plotted points using a line segment. The curve obtained will be kinked.
- This resulting curve is called the frequency polygon.

To plot a frequency polygon

```
Import matplotlib.pyplot as plt
```

```
h=[1,11,24,94,1,52,2,63,74,46,52,32,58,43, 98, 85, 77]
```

```
plt.hist(h,color='r', histtype='step')
```

```
plt.xlabel("value")
```

```
plt.ylabel("frequency")
```

```
plt.title("FREQUENCY POLYGON")
```

```
plt.show()
```

Watch this video to understand the way
to create FREQUENCY POLYGON
using MATPLOTLIB -

<https://www.youtube.com/watch?v=UO98IJQ3QGI>

BBPSPP

- **BOX PLOT** - is a visual representation of the statistical five number summary of a given data set which includes minimum, first quartile, median (second quartile), third quartile and maximum.

To plot a box plot

```
Import matplotlib.pyplot as plt
```

```
h=[1,3, 4, 7, 8, 8, 9]
```

```
plt.boxplot(h)
```

```
plt.show()
```


ASSIGNMENT

Given the data :

Weight measurements for 14 values of muffins

78, 72, 69, 81, 63, 67, 65, 79, 74, 71, 83, 71, 79, 80

- 1. Create a simple histogram for the above data.**
- 2. Create a horizontal box plot for the above data.**
- 3. Create a vertical box plot for the above data.**
- 4. What is a the difference between histogram and frequency polygon?**

ALL THE ABOVE QUESTIONS TO BE PART OF PRACTICAL FILE ALSO.YOU WRITE THE CODE USING W3SCHOOL.COM OR JUPYTER.ORG SITES