**FMS 102**

**FIRST CLASS**

**STATISTICS: DATA PRESENTATION**

Statistics is the study of the collection, analysis, interpretation, organization, and presentation of data.

**Uses: at different levels**

**The family**: for economic decisions

**The firm**: for management purposes

**The government**: welfare information

**The academia**: for research solutions

**The student**: add your reason

**Data**

Data is a collection of facts, such as numbers, words, measurements, observations etc.

**Types of Data-**

1. **Qualitative data-** it is descriptive data. One can only describe them
   * Example- She can run fast, He is thin.
2. **Quantitative data-** it is numerical information. Have numerical values
   * Example- exam score or grades.

**Types of quantitative data**

1. **Discrete data-** has a particular fixed value. It can be counted
2. **Continuous data-** is not fixed but has a range of data. It can be measured.

**Representation of Data**

By data representation, we mean the different ways in which we can present or represent data set to create meaning or understanding.

There are several ways including; Tables, graphs, charts, number lines (line graphs), pictographs, etc.

|  |  |
| --- | --- |
|  | **Tables**  Statistical table is a table showing required data in a convenient form. It can tabulate data from different time periods. |
| Statistics- Bar graph | **Bar Graph**  A Bar Graph represents grouped data with rectangular bars with lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally. |
| Statistics-Pie chart | **Pie Chart**  A type of graph in which a circle is divided into Sectors. Each of these sectors represents a proportion of the whole. |
| Statistics-Line graph | [Line graph](https://byjus.com/maths/line-graph/) The line chart is represented by a series of data points connected with a straight line. The series of data points are called ‘markers.’ |
| Statistics-Pictograph | [Pictograph](https://byjus.com/maths/pictograph/) A pictorial symbol for a word or phrase, i.e. showing data with the help of pictures. Such as Apple, Banana & Cherry can have different numbers, and it is just a representation of data. |
| Statistics- Histogram | [Histogram](https://byjus.com/maths/histogram/) A diagram is consisting of rectangles. Whose Area is proportional to the frequency of a variable and whose width is equal to the class interval. |

|  |  |
| --- | --- |
| Frequency distribution in Statistics | [Frequency Distribution](https://byjus.com/maths/frequency-distribution-table/) The frequency of a data value is often represented by “f.” A frequency table is constructed by arranging collected data values in ascending order of magnitude with their corresponding frequencies. |
|  | **A scatter diagram**  A scatter diagram (**Also known as scatter plot, scatter graph, and correlation chart**) is a tool for analyzing relationships between two variables for determining how closely the two variables are related. |

### Our concern for this present course would be Frequency distribution, the histogram, the pie chart.

**FREQUENCY DISTRIBUTION**

62 71 60 51 57 55 50 45 47 67 55 55 40

70 50 58 65 55 42 40 40 31 27 40 55 40

23 79 71 81 47 85 31 31 55 51 40 55 31

**1. UN-GROUPED DATA**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **MARKS** | **TALLY** | **FREQ** | **CF** |  |
| 23 | *I* | 1 | 1 |  |
| 27 | *I* | 1 | 2 |  |
| 31,31, 31,31 | *I III* | 4 | 6 |  |
| 40,40,40,40,40,40 | *~~I III~~ I* | 6 | 12 |  |
| 42 | *I* | 1 | 13 |  |
| 45 | *I* | 1 | 14 |  |
| 47, 47 | *I I* | 2 | 16 |  |
| 50,50 | *I I* | 2 | 18 |  |
| 51,51 | *I I* | 2 | 20 |  |
| 55,55,55,55,55,55,55 | *~~I I I I~~11* | 7 | 27 |  |
| 57 | *I* | 1 | 28 |  |
| 58 | *I* | 1 | 29 |  |
| 60 | *I* | 1 | 30 |  |
| 62 | *I* | 1 | 31 |  |
| 65 | *I* | 1 | 32 |  |
| 67 | *I* | 1 | 33 |  |
| 70 | *I* | 1 | 34 |  |
| 71,71 | *I I* | 2 | 36 |  |
| 79 | *I* | 1 | 37 |  |
| 81 | *I* | 1 | 38 |  |
| 85 | *I* | 1 | 39 |  |
| **TOTAL** |  | **39** | **-** |  |
|  |  |  |  |  |

**2. GROUPED DATA**

**FREQUENCY DISTRIBUTION TABLE**

Frequency (F)

C I (M)

A

B

C

DA

E

F

5

10

15

20

**A** **BAR** **CHART**

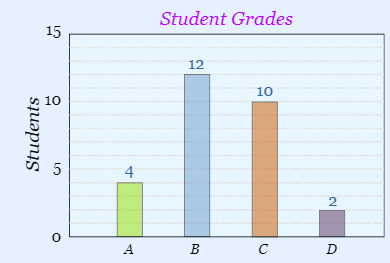
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Class interval** | **Tally** | **Freq.(F)** | **CF** |  |
| 0 – 39 (F) | ~~IIII~~ I | 6 | 6 |  |
| 40 – 44 (E) | ~~IIII~~ II | 7 | 13 |  |
| 45 – 49 (D) | III | 3 | 16 |  |
| 50 – 59 (C) | ~~IIII~~ ~~IIII~~ III | 13 | 29 |  |
| 60 – 69 (B) | ~~IIII~~ | 4 | 33 |  |
| 70 – 100 (A) | ~~IIII~~ I | 6 | 39 |  |
| **Total** |  | **39** | **-** |  |

**THE BAR CHART**

A Bar Chart (also called Bar Graph) is a graphical display of data using bars of different heights. It consists of rectangular bars with lengths proportional to the values that they represent (frequencies). The bars can be plotted vertically or horizontally.

**Example 2: Test scores for 20 students**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grade: | **A** | **B** | **C** | **D** |
| Students: | 4 | 12 | 10 | 2 |



**THE HISTOGRAM**

The histogram is a diagram which consists of rectangles. The height is proportional to the frequency of a variable and the width is equal to the class interval. The difference between a bar graph and a histogram is that there is no gap between the bars in the histogram.

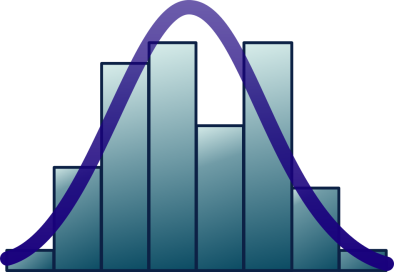
A **histogram** is **used** to summarize discrete or continuous data. In other words, it provides a visual interpretation. This requires focusing on the main points, facts of numerical data by showing the number of data points that fall within a specified range of values (called “bins” “intervals”).

### Importance of a Histogram

Creating a histogram provides a visual representation of data distribution. Histograms can display a large amount of data and the [frequency](https://corporatefinanceinstitute.com/resources/excel/functions/frequency-function/) of the data values. The [median](https://corporatefinanceinstitute.com/resources/excel/functions/median-function/) and distribution of the data can be determined by a histogram. In addition, it can show any outliers or gaps in the data.

**Types of Histogram?**

A histogram is used to summarize discrete or continuous data. In other words, it provides a [visual interpretation](https://corporatefinanceinstitute.com/resources/knowledge/other/data-presentation-guide/)of numerical data by showing the number of data points that fall within a specified range of values (called “bins”). It is similar to a vertical bar graph. However, a histogram, unlike a vertical bar graph, shows no gaps between the bars.



**Parts of a Histogram**

1. **The title:** The title describes the information included in the histogram.
2. **X-axis:** The [X-axis](https://en.wikipedia.org/wiki/Coordinate_system) are intervals that show the scale of values which the measurements fall under.
3. **Y-axis:** The Y-axis shows the number of times that the values occurred within the intervals set by the X-axis.
4. **The bars:** The height of the bar shows the number of times that the values occurred within the interval, while the width of the bar shows the interval that is covered. For a histogram with equal bins, the width should be the same across all bars.

**Importance of a Histogram**

Creating a histogram provides a visual representation of data distribution. Histograms can display a large amount of data and the [frequency](https://corporatefinanceinstitute.com/resources/excel/functions/frequency-function/) of the data values. The [median](https://corporatefinanceinstitute.com/resources/excel/functions/median-function/) and distribution of the data can be determined by a histogram. In addition, it can show any outliers or gaps in the data.

**Distributions of a Histogram**

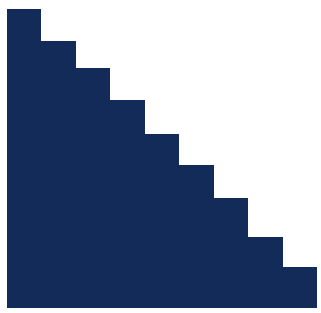
**A normal distribution (Symmetric, unimodal):**In a normal distribution, points on one side of the [average](https://corporatefinanceinstitute.com/resources/excel/functions/average-excel-function/) are as likely to occur as on the other side of the average.



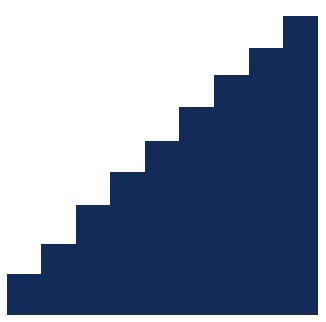
**A bimodal distribution:** In a bimodal distribution, there are two peaks. In a bimodal distribution, the data should be separated and analyzed as separate normal distributions.



**A right-skewed distribution:** A right-skewed distribution is also called a positively skewed distribution. In a right-skewed distribution, a large number of data values occur on the left side with a fewer number of data values on the right side. A right-skewed distribution usually occurs when the data has a range boundary on the left-hand side of the histogram. For example, a boundary of 0.



**A left-skewed distribution:** A left-skewed distribution is also called a negatively skewed distribution. In a left-skewed distribution, a large number of data values occur on the right side with a fewer number of data values on the left side. A right-skewed distribution usually occurs when the data has a range boundary on the right-hand side of the histogram. For example, a boundary such as 100.



**A random distribution:**A random distribution lacks an apparent pattern and has several peaks. In a random distribution histogram, it can be the case that different data properties were combined. Therefore, the data should be separated and analyzed separately.



**Steps in Creating a Grouped Frequency Distribution**

1. Find the largest and smallest values (Maximum and Minimum)
2. Compute the Range = Maximum - Minimum
3. Select the number of classes desired. This is usually between 5 and 20.
4. Find the class width by dividing the range by the number of classes. Usually between 5 and 20.
5. Pick a suitable starting point equal to the minimum value.
6. Your starting point is the lower limit of the first class.
7. Continue to add the class width to this lower limit to get the rest of the lower limits.
8. To find the upper limit of the first class, subtract one from the lower limit of the second class. Then continue to add the class width to this upper limit to find the rest of the upper limits.
9. Find the boundaries by subtracting 0.5 units from the lower limits and adding 0.5 units from the upper limits. The boundaries are also half-way between the upper limit of one class and the lower limit of the next class. Depending on what you're trying to accomplish, it may not be necessary to find the boundaries.
10. Tally the data.
11. Find the frequencies.
12. Find the cumulative frequencies. Depending on what you're trying to accomplish, it may not be necessary to find the cumulative frequencies.
13. If necessary, find the relative frequencies and/or relative cumulative frequencies.

**Working Example**

**The following observations are FMS 102 marks of 50 students in FUOYE**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 36 | 25 | 38 | 46 | 55 | 55 | 72 | 55 | 36 | 38 |
| 67 | 45 | 26 | 48 | 91 | 46 | 52 | 61 | 58 | 55 |
| 25 | 27 | 32 | 48 | 57 | 62 | 78 | 54 | 33 | 36 |
| 63 | 48 | 82 | 46 | 95 | 44 | 55 | 63 | 56 | 55 |
| 95 | 43 | 55 | 62 | 25 | 72 | 62 | 55 | 85 | 67 |

Construct a frequency distribution for the observation.

**Table XYZ: Frequency Distribution for FMS 102 Marks**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Intervals** | **Tally** | **Frequency (F)** | **Cumulative Freq. (CF)** | **Class Boundary** | **Class Mark (X)** |
| 25 – 33 | ~~IIII~~ II | 7 | 7 | 24.5 – 33.5 | 29 |
| 34 – 42 | ~~IIII~~ | 5 | 12 | 33.5 – 42.5 | 38 |
| 43 – 51 | ~~IIII~~ IIII | 9 | 21 | 42.2 – 51.5 | 47 |
| 52 – 60 | ~~IIII~~ ~~IIII~~ III | 13 | 34 | 51.5 – 60.5 | 56 |
| 61 – 69 | ~~IIII~~ III | 8 | 42 | 60.5 – 69.5 | 65 |
| 70 – 78 | III | 3 | 45 | 69.5 – 78.5 | 74 |
| 79 – 87 | II | 2 | 47 | 78.5 – 87.5 | 83 |
| 88 – 96 | III | 3 | 50 | 87.5 – 96.5 | 92 |
| Total |  | 50 | - |  |  |

**Histogram for FMS 102 marks of 50 students in FUOYE**

10

12

14

16

**Class Boundary**

**Frequency (F)**

24.5

33.5

42.5

2

51.5

60.5

69.5

78.5

87.5

96.5

4

6

8