**Measures of location (central tendency)**

A **central tendency** (**measure of central tendency**) is a central value or a **typical value** for a probability distribution. It is occasionally called an **average** or just the **center** of the distribution. The most common measures of central tendency are the arithmetic mean, median and the mode.

**Example:** find if you have the following data set

**Mean**

The sum of all measurements divided by the number of observations in the data set, most widely used measure of central location.

1. **Un-grouped data**

In case of frequency is not given, it's simply calculated by taking summation then divide it by total number of observations.

1. **Grouped data**

But when the frequency is given, it could be calculated as follow:

|  |  |  |  |
| --- | --- | --- | --- |
| **Time** | **Midpoint ()** | **Frequency ()** |  |
| 3.1-3.5 | 3.3 | 5 | 16.5 |
| 3.6-4.0 | 3.8 | 18 | 68.4 |
| 4.1-4.5 | 4.3 | 25 | 107.5 |
| 4.6-5.0 | 4.8 | 27 | 129.6 |
| 5.1-5.5 | 5.3 | 20 | 106 |
| 5.6-6.0 | 5.8 | 5 | 29 |
|  |  |  |  |

**Where**

**Median (middle)**

The value of the middle item in a set of observations which has been arranged in an ascending or descending order of magnitude and is the centermost value in a distribution.

1. **Un-grouped data**
2. **Odd number**
3. **Even number**
4. **Grouped data**

**Where**

= Median of the distribution

= lower limit of the median class.

= Summation of total frequencies.

= frequency of median class.

= cumulative frequency before **median class** exactly.

= class interval.

**Example**: Find the of the continuous frequency distribution below

|  |  |  |
| --- | --- | --- |
| **Class** | **Frequency** | **Cumulative Frequency** |
| 5-10 | 80 | 80 |
| 10-15 | 70 | 150 |
| 15-20 | 120 | 270 |
| **20-25** | **210** | 480 |
| 25-30 | 180 | 660 |
| 30-35 | 40 | 700 |
|  |  | The greater value from cumulative frequency is **480** (class 20-25) |

**Where**

= ?

= 20

= 700

= 210

= 270

= 5

**Mode**

The mode of a set of values is that value which occurs most frequently. If all the values are different there is no mode; on the other hand, a set of values may have more than one mode.

**Where**

= Mode of the distribution

= lower limit of the modal class.

= frequency of modal class.

= frequency just before the modal class.

= frequency just after the modal class.

= Class interval.

The approximate value of the mode can be computed using the following relationship:

**Example**: Find the of the continuous frequency distribution below

|  |  |  |
| --- | --- | --- |
| **Class** | **Frequency** | The most frequent class number is representing the modal class |
| 5-10 | 80 |
| 10-15 | 70 |
| 15-20 | 120 |
| **20-25** | 210 |
| 25-30 | 180 |
| 30-35 | 40 |
|  |  |

**Where**

= ?

= 20

= 210

= 120

= 180

= 5

**Example**: If you have Plasma Glucose Values (mg/dl) for a Sample of 100 Adults Aged 20–74 Years, find mean, mode and median?

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **70** | 81 | 83 | 87 | 90 | 91 | 94 | 97 | 101 | 106 |
| 75 | 81 | 83 | 87 | 90 | 92 | 95 | 98 | 102 | 107 |
| 77 | 81 | 83 | 88 | 90 | 92 | 95 | 98 | 102 | 109 |
| 78 | 82 | 84 | 88 | 90 | 92 | 95 | 99 | 103 | 111 |
| 78 | 82 | 84 | 89 | 90 | 92 | 95 | 99 | 103 | 112 |
| 78 | 82 | 84 | 89 | 90 | 93 | 95 | 99 | 103 | 114 |
| 80 | 82 | 84 | 89 | 90 | 93 | 96 | 100 | 104 | 115 |
| 80 | 83 | 85 | 89 | 90 | 94 | 96 | 100 | 104 | 116 |
| 80 | 83 | 86 | 89 | 90 | 94 | 97 | 101 | 105 | 117 |
| 80 | 83 | 86 | 90 | 91 | 94 | 97 | 101 | 106 | **117** |

Class number = 1 + 3.322 log n

= 1 + 3.322 \* log 100

= 1 + 3.322 \* 2

= 1 + 6.644

= 7.644

**Class number = 8**

**Class interval =** 6.14 or 6

|  |  |  |  |
| --- | --- | --- | --- |
| **Class** | **Frequency** | **Midpoint** |  |
| 70 – 75 | 2 | 73 | 146 |
| 76 – 81 | 15 | 79 | 1185 |
| 82 – 87 | 17 | 85 | 1445 |
| 88 – 93 | 27 | 91 | 2457 |
| 94 – 99 | 17 | 97 | 1649 |
| 100 – 105 | 13 | 103 | 1339 |
| 106 – 111 | 4 | 109 | 436 |
| 112 – 117 | 5 | 115 | 575 |
|  |  |  |  |

**Where**

= ?

= 88

= 100

= 27

= 34

= 6

|  |  |  |
| --- | --- | --- |
| **Class** | **Frequency** |  |
| 70 – 75 | 2 | 2 |
| 76 – 81 | 15 | 17 |
| 82 – 87 | 17 | 34 |
| 88 – 93 | 27 | 61 |
| 94 – 99 | 17 | 78 |
| 100 – 105 | 13 | 91 |
| 106 – 111 | 4 | 95 |
| 112 – 117 | 5 | 100 |
|  |  |  |

= ?

= 88

= 27

= 17

=17

= 6

|  |  |
| --- | --- |
| **Class** | **Frequency** |
| 70 – 75 | 2 |
| 76 – 81 | 15 |
| 82 – 87 | 17 |
| 88 – 93 | 27 |
| 94 – 99 | 17 |
| 100 – 105 | 13 |
| 106 – 111 | 4 |
| 112 – 117 | 5 |
|  |  |

**Question**: Construct a frequency distribution with suitable class interval size of marks obtained by 50 students of a class is given below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 12 | 34 | 43 | 50 | 55 | 62 | 75 |
| 15 | 35 | 44 | 51 | 56 | 63 |  |
| 21 | 36 | 46 | 52 | 56 | 64 |  |
| 23 | 38 | 47 | 52 | 57 | 65 |  |
| 26 | 39 | 47 | 52 | 58 | 65 |  |
| 27 | 41 | 48 | 53 | 59 | 67 |  |
| 30 | 42 | 48 | 54 | 59 | 68 |  |
| 33 | 43 | 50 | 54 | 60 | 72 |  |

Class number = 1 + 3.322 log n

= 1 + 3.322 \* log 50

= 1 + 3.322 \* 1.698

= 1 + 5.643

= 6.643

**Class number = 7**

**Class interval =** 9.78 or 10

|  |  |
| --- | --- |
| **Class** | **Frequency** |
| 12 – 21 |  |
| 22 – 31 |  |
| 32 – 41 |  |
| 42 – 51 |  |
| 52 – 61 |  |
| 62 – 71 |  |
| 72 – 81 |  |
|  |  |