Practical No. 1

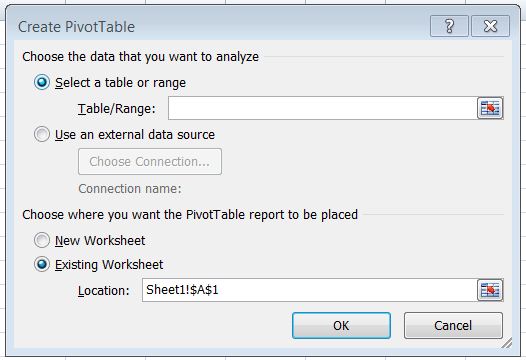
***Title - Formation of Frequency Distribution***

**Case 1 : For Qualitative and Discrete Information:**

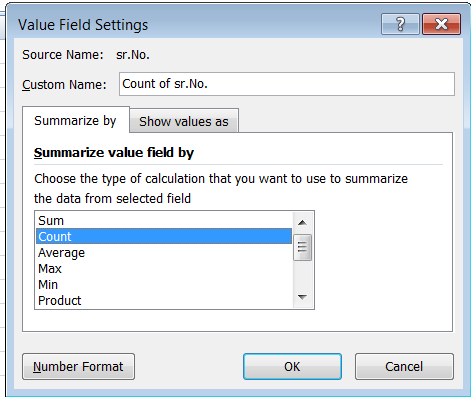
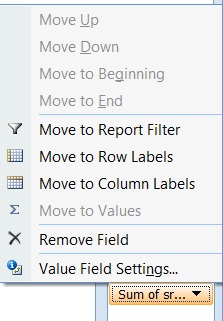
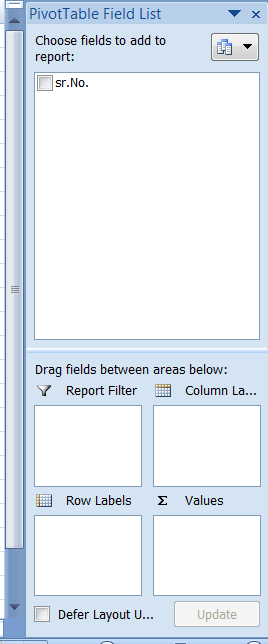
When there are small number of distinct elements or observations available in the data and data is in qualitative nature, we cannot form grouped class interval 0-10, 10-20, 20-30,… In this case, individual observation itself is to be treated as a class, while formulating the table. In such case, following methodology should be followed:

***Stepwise Procedure:***

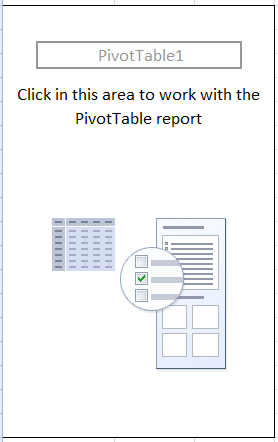
1. List out all distinct observations from the data. Here we can use filter option to identify all discrete observations and duplicate cases (where total no. of observations are large) to maintain uniformity in the data.
2. Go to insert, choose option as Pivot table. The dialogue box for the pivoting table looks like as follow;



1. Select table or data range in above dialogue box. Choose where you want to the pivot table report to be placed either by selecting New Worksheet or Existing Worksheet. Click on OK.
2. From the ‘pivot table field list’ choose appropriate variable whose frequency is to be calculated. Drag the variable whose frequency is to be calculate in Row Labels. Drag the same variable in values. Note that in value field setting there must be option ‘Count’ Selected



1. We get required frequency table for qualitative and discrete information in place of Pivot Table1



When we have discrete information, we also use the following stepwise procedure to form a frequency table;

1. Enter the all information in one column. Use filter to identify all distinct observation.
2. Enter all distinct observation in Bins array. Hence created all distinct observations are treated as a class in a frequency table.
3. To obtain frequency corresponding to class in Bins array, follow the steps given below:
4. Give the column title Frequency to the immediate right column of Bins array.
5. Select blank sales in frequency column corresponding to the classes available in the Bins array.
6. Click on insert function button which is available in formula bar (Insert function, button is also available in formula menu)
7. In insert function window, select the statistical option as the function category.
8. Select option- ‘frequency’ from the select function option, where many statistical functions are displayed → Click on OK button.
9. A function arguments window, select all observations in data array.
10. Press F2.
11. Click on Bins array option. In a function argument window and select all classes from the Bins array.
12. Press F2 twice.
13. Press Ctrl +Shift + Enter.

**Note:**

If one of the above three ctrl+ Shift+ enter is not working properly, one may get wrong frequency. To overcome this problem, one should press f2 and use another key combination of Ctrl + Shift + Enter available in the keyboard.

**Case 2: For Grouped Classes**

Whenever there is a large no. of distinct observations available in the data, it is not available to compute frequency. By treating each distinct observation as a class in a frequency table. Instead of this, we form class intervals to obtain frequency for the observations belonging to that particular class interval. In such case, follow the methodology given below:

1. Obtain the maximum observation from the data by using formula MAX.

Maximum = MAX(data range).

Obtain the minimum observation from the data by using the formula MIN.

Minimum= MIN (data range).

1. Calculate Range= MAX-MIN.
2. Compute the value of k : number of classes by using the formula given below:

k= 1 + 3.322 \* ;

1. Compute the length of the class interval by using the formula given below:

Length of CI =

1. Construct the Bins array to the immediate right column of the data array by using appropriate class interval in such a way that both, minimum and maximum observation must be covered by first and last class interval respectively.

Here the value of Bins array will be upper class limit of the respective class intervals for inclusive method of classification while for the exclusive method of classification. The value of Bins array will be lesser than the upper limit of respective class interval because when we compute the frequency of the value of Bins array included.

1. Obtain frequency in the immediate right column the Bins array by using methodology, already discussed in Step 4, Case 1 (i.e. methodology for ungrouped class interval).
2. Generate the class interval and frequency column separately. To generate class interval column, enter the lower class limit in a single column in the next column, enter all upper class limit. In the immediate right column use the function “Concatenate”.

= CONCATENATE(insert lower class limit, “-”, insert upper class limit).

***Examples:***

1. Following is the data obtained from strawberry orchard. Researcher collects data from 95 strawberry plants, to know frequency distribution of no. of strawberries per tree. Help him to form frequency table from collected information:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 64 | 53 | 32 | 94 | 30 | 69 | 27 | 34 | 74 | 78 | 15 | 78 | 33 | 61 | 19 | 74 | 37 | 76 | 52 |
| 67 | 98 | 48 | 54 | 14 | 49 | 92 | 63 | 41 | 58 | 42 | 25 | 79 | 62 | 54 | 68 | 40 | 60 | 87 |
| 91 | 39 | 98 | 30 | 52 | 10 | 62 | 62 | 79 | 96 | 78 | 26 | 42 | 96 | 47 | 40 | 66 | 52 | 13 |
| 94 | 36 | 86 | 61 | 21 | 21 | 43 | 38 | 48 | 78 | 39 | 68 | 89 | 98 | 84 | 92 | 21 | 16 | 44 |
| 72 | 76 | 87 | 66 | 70 | 21 | 30 | 27 | 67 | 87 | 64 | 30 | 14 | 95 | 88 | 45 | 93 | 75 | 20 |

**Solution:**

|  |  |
| --- | --- |
| Max | 98 |
| Min | 10 |
| Range | 88 |
| K | 8 |
| C.I | 12 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sr. No.** | **Lower bond** | **Upper bond** | **CLASS** | **BINS** | **Frequency** | **GCF** |
| 1 | 8 | 20 | 8-20 | 19 | 7 | 95 |
| 2 | 20 | 32 | 20-32 | 31 | 13 | 88 |
| 3 | 32 | 44 | 32-44 | 43 | 14 | 75 |
| 4 | 44 | 56 | 44-56 | 55 | 12 | 61 |
| 5 | 56 | 68 | 56-68 | 67 | 14 | 49 |
| 6 | 68 | 80 | 68-80 | 79 | 16 | 35 |
| 7 | 80 | 92 | 80-92 | 91 | 8 | 19 |
| 8 | 92 | 104 | 92-104 | 103 | 11 | 11 |
| **Total** | **---** | **---** | **---** | **---** | **95** | **---** |

**Interpretation:**

From the above frequency table, we can say that there are near about 62% plants having production per plant is more than 44 strawberries.

1. Following data were observed to know popularity of bike models in Vidyanagar. To address above question, the data collected from the Bhaikaka circle for an hour on a holiday.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 7 | 10 | 9 | 2 | 11 | 7 | 3 | 6 | 11 | 7 | 11 | 9 | 1 | 1 | 9 | 10 | 11 | 5 | 6 | 2 |
| 7 | 10 | 7 | 6 | 9 | 9 | 9 | 4 | 1 | 9 | 5 | 7 | 9 | 6 | 1 | 7 | 6 | 9 | 9 | 8 |
| 6 | 11 | 2 | 7 | 11 | 10 | 11 | 1 | 5 | 7 | 4 | 3 | 11 | 7 | 6 | 6 | 9 | 1 | 10 | 6 |
| 8 | 4 | 2 | 10 | 10 | 9 | 1 | 11 | 9 | 6 | 9 | 3 | 4 | 6 | 1 | 6 | 8 | 4 | 3 | 11 |
| 5 | 1 | 5 | 1 | 3 | 1 | 5 | 4 | 2 | 2 | 9 | 8 | 1 | 6 | 9 | 2 | 5 | 7 | 1 | 11 |
| 5 | 2 | 7 | 3 | 6 | 3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |
| --- | --- |
| **Codes** | **Bikes** |
| 1 | Discover |
| 2 | passion |
| 3 | Splendor |
| 4 | Delux |
| 5 | Glamour |
| 6 | Activa |
| 7 | CT100 |
| 8 | Dreamyuga |
| 9 | Livo |
| 10 | Sports |
| 11 | Centure |

**Solution:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr. No.** | **Models** | **# Bikes** | **% frequency** |
| 1 | Activa | 14 | 13.20 |
| 2 | Centure | 11 | 10.37 |
| 3 | CT100 | 12 | 11.32 |
| 4 | Delux | 6 | 5.66 |
| 5 | Discover | 13 | 12.26 |
| 6 | Dreamyuga | 4 | 3.77 |
| 7 | Glamour | 8 | 7.54 |
| 8 | Livo | 16 | 15.09 |
| 9 | Passion | 8 | 7.54 |
| 10 | Splendor | 7 | 6.60 |
| 11 | Sports | 7 | 6.60 |
|  | **Grand Total** | **106** |  |

**Interpretation:**

From the above frequency table, we say that there are Livo is most frequently bike passed through Bhaikaka Circle.

1. Following data collected from hospital. To know in which age group most diabetics patient are given. Then investigator collect a data from a particular hospital.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 38 | 42 | 25 | 34 | 30 | 48 | 52 | 32 | 62 | 34 | 51 | 54 | 62 |
| 34 | 67 | 70 | 52 | 72 | 74 | 21 | 17 | 75 | 24 | 70 | 54 | 64 |
| 34 | 43 | 45 | 38 | 32 | 27 | 62 | 29 | 47 | 43 | 68 | 75 | 77 |
| 46 | 50 | 33 | 28 | 45 | 35 | 33 | 36 | 30 | 38 | 36 | 33 | 44 |
| 21 | 46 | 49 | 43 | 48 | 34 | 41 | 62 | 53 | 69 | 36 | 50 | 37 |

**Solution:**

|  |  |
| --- | --- |
| Max | 77 |
| Min | 17 |
| Range | 60 |
| K | 8 |
| C.I | 8 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sr. No.** | **Lower Bond** | **Upper Bound** | **Age Groups** | **Bins** | **# Patients** |
| 1 | 15 | 23 | 15-23 | 22 | 6 |
| 2 | 23 | 31 | 23-31 | 30 | 11 |
| 3 | 31 | 39 | 31-39 | 38 | 26 |
| 4 | 39 | 47 | 39-47 | 46 | 16 |
| 5 | 47 | 55 | 47-55 | 54 | 20 |
| 6 | 55 | 63 | 55-63 | 62 | 8 |
| 7 | 63 | 71 | 63-71 | 70 | 8 |
| 8 | 71 | 79 | 71-79 | 78 | 5 |
| **Total** | **---** | **---** | **---** | **---** | **100** |

**Interpretation:**

From the above frequency table, we can say that the in 31-39 age group most diabetics patients are given.

1. Following are the data obtained from Library. To know the frequency distribution of no. of readers meet to Library in a one day, The Investigator collects 96 days information as given below:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 66 | 79 | 48 | 46 | 41 | 62 | 57 | 55 | 47 | 45 | 55 | 53 |
| 49 | 60 | 31 | 22 | 54 | 53 | 56 | 47 | 45 | 42 | 67 | 38 |
| 85 | 52 | 48 | 51 | 69 | 52 | 51 | 42 | 43 | 37 | 59 | 58 |
| 46 | 39 | 41 | 27 | 53 | 36 | 39 | 56 | 49 | 63 | 38 | 82 |
| 44 | 59 | 38 | 41 | 47 | 61 | 48 | 68 | 43 | 63 | 47 | 46 |
| 38 | 39 | 46 | 27 | 48 | 39 |  |  |  |  |  |  |

**Solution:**

|  |  |
| --- | --- |
| Max | 85 |
| Min | 11 |
| Range | 74 |
| K | 7.585105 |
| Width | 9.755969 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sr. No.** | **Lower bound** | **Upper bond** | **# Readers** | **bins** | **# Days** | **GCF** |
| 1 | 10 | 20 | 10-20 | 19 | 2 | 96 |
| 2 | 20 | 30 | 20-30 | 29 | 5 | 94 |
| 3 | 30 | 40 | 30-40 | 39 | 14 | 89 |
| 4 | 40 | 50 | 40-50 | 49 | 35 | 75 |
| 5 | 50 | 60 | 50-60 | 59 | 23 | 40 |
| 6 | 60 | 70 | 60-70 | 69 | 12 | 17 |
| 7 | 70 | 80 | 70-80 | 79 | 3 | 5 |
| 8 | 80 | 90 | 80-90 | 89 | 2 | 2 |
|  | **Total** | **---** | **---** | **---** | **96** | **---** |

**Interpretation:**

From the above frequency table, we can say that there are near about 50% readers having meets per day more than 40 readers.

1. Following data collected from the hospital. To know the in which OPD more patients are given. Data are collected from investigator by each OPD have particular code and investigator write the code randomly.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 1 | 5 | 10 | 3 | 4 | 4 | 8 | 1 | 2 | 4 | 2 | 10 | 1 | 4 | 7 | 2 | 6 |
| 8 | 1 | 10 | 8 | 1 | 7 | 9 | 6 | 6 | 10 | 4 | 2 | 6 | 4 | 9 | 1 | 7 | 8 |
| 8 | 8 | 10 | 6 | 6 | 10 | 1 | 4 | 3 | 1 | 5 | 1 | 5 | 1 | 5 | 10 | 6 | 10 |
| 8 | 7 | 2 | 6 | 10 | 9 | 3 | 9 | 4 | 1 | 1 | 2 | 3 | 6 | 6 | 8 | 2 | 3 |
| 3 | 10 | 4 | 1 | 5 | 6 | 9 | 4 | 6 | 5 | 10 | 3 | 2 | 3 | 7 | 3 | 10 | 5 |
| 1 | 9 | 7 | 9 | 1 | 8 | 2 | 4 | 9 | 5 | 6 | 3 | 1 | 8 |  |  |  |  |

|  |  |
| --- | --- |
| **OPD Code** | **OPD Name** |
| 1 | Skin |
| 2 | Eyes |
| 3 | Dental |
| 4 | Neurology |
| 5 | Hair |
| 6 | Accident |
| 7 | Surgery |
| 8 | Orthopaedic |
| 9 | Outhaul |
| 10 | Urology |

**Solution:**

|  |  |  |
| --- | --- | --- |
| **Sr. No.** | **OPD Names** | **# OPD** |
| 1 | Accident | 13 |
| 2 | Dental | 10 |
| 3 | Eyes | 9 |
| 4 | Hair | 8 |
| 5 | Neurology | 11 |
| 6 | outhaul | 8 |
| 7 | Orthopaedic | 10 |
| 8 | Skin | 17 |
| 9 | Surgery | 6 |
| 10 | Urology | 12 |
|  | **Grand Total** | **104** |

**Interpretation:**

From the above frequency table, we say that the in the Accident OPD have more patients.