



Department of Census and Statistics

Training Division

Training Module Description

Module: TABULATION

Version: 1.0

Duration: 05 days

Tabulation

1.0 Module Description

The main objective this module is to introduce different tabulation methods in statistics. Tabulation refers to the systematic arrangement of the information in rows and columns. Tabulation is as statistical table with logical listing of related quantitative data in vertical columns and horizontal rows of numbers with sufficient explanatory and qualifying words, phrases and statements in the form of titles, headings and footnotes to make clear the full meaning of the data and their origin.” The main purpose of tabulation is to summarize and present data is to see whether there are patterns in them that can prompt effective decisions. The module covers simple frequency tables, group frequency tables, multivariate frequency tables, cumulative frequency tables and cross tabulation.

1.1 Objectives of the Module are;

- 1.1.1 To discuss the rules of tabulation
- 1.1.2 To discuss different tabulation methods
- 1.1.3 To improve the quality of data presentation

1.2 Target Participants

Statistical Officers of the Department of Census and Statistics

1.3 Brief Module Contents (Syllabus)

- 1.3.1 Introduction to tabulation
 - 1.3.1.1 Objectives and importance of tabulation
 - 1.3.1.2 Rules of tabulation
 - 1.3.1.3 Parts of a table
 - 1.3.1.4 Limitation of tabulation
- 1.3.2 Methods of tabulation
 - 1.3.2.1 Simple tabulation

1.3.2.2 Double tabulation

1.3.2.3 Complex tabulation

1.3.3 Frequency tables

1.3.3.1 Simple frequency tables

1.3.3.2 Grouped frequency tables

1.3.3.3 Multivariate frequency table

1.3.3.4 Cumulative frequency tables

1.3.4 Cross tabulation

1.3.4.1 Bivariate cross tabulation

1.3.4.2 Multivariate cross tabulation

1.4 Method of delivery

Lectures and discussions are the key methods in delivering the course. In addition to lectures and discussions, participants are encouraged to improve their knowledge through group activities.

2.0 Session 1 (Day 1: 0900 – 1200hrs): Introduction to tabulation

2.1 Session Description

The main objective of this session is to make an overview of tabulation techniques. The session is basically devoted to discuss objectives and importance of tabulation techniques, rules in tabulating data, essential components of a standard table and limitation of tabulation.

2.2 Session Learning Outcomes

By the end of this session participants will be able to;

- 2.2.1 understand the objectives and importance of tabulation techniques.
- 2.2.2 tabulate data according to standard rules in statistics.
- 2.2.3 create a table with all essential components
- 2.2.4 understand limitations in tabulating data.

2.3 Brief outline of the session

- 2.3.1 The importance of tabulation
- 2.3.2 The objectives of tabulation
- 2.3.3 The standard rules in tabulating data
- 2.3.4 The practice in tabulating data
- 2.3.5 Limitation in tabulating data

3.0 Session 2 (Day 1: 1300 – 1600hrs): Methods of tabulation

3.1 Session Description

The main objective of this session is to discuss deferent tabulation methods in statistics. There are three methods of tabulation; Simple, double and complex. This section is basically devoted to discuss characteristics of each method. Participants are trained to apply most appropriate method in proper manner in tabulating data for different occasions.

3.2 Session Learning Outcomes

By the end of this session participants will be able to;

- 3.2.1 understand different methods in tabulating data
- 3.2.2 understand the relevant method of tabulating data for different occasions
- 3.2.3 understand alternative ways of crating simple, double and complex tables with considerations of characteristics of the set of data.

3.3 Brief outline of the session

- 3.3.1 Simple tabulation
 - 3.3.1.1 Overview of simple tabulation
 - 3.3.1.2 Alternative ways of creating a simple table
 - 3.3.1.3 Comparative advantages and disadvantages of simple tabulation
- 3.3.2 Double tabulation
 - 3.3.2.1 Overview of double tabulation
 - 3.3.2.2 Alternative ways of creating a double table
 - 3.3.2.3 Comparative advantages and disadvantages of double tabulation
- 3.3.3 Complex tabulation
 - 3.3.3.1 Overview of complex tabulation
 - 3.3.3.2 Alternative ways of creating a double table
 - 3.3.3.3 Comparative advantages and disadvantages of double tabulation

4.0 Session 3 (Day 2: 0900 – 1200hrs): Simple frequency tables

4.1 Session Description

The main objective of this session is to discuss and practice simple frequency tables. The session covers characteristics of a simple table and way of creating an interactive simple table through particle examples.

4.2 Session Learning Outcomes

By the end of this session participants will be able to;

- 4.2.1 understand the nature of a simple table
- 4.2.2 create an interactive simple table
- 4.2.3 to understand the scope and limitations of a simple table

4.3 Brief outline of the session

- 4.3.1 Introduction to simple frequency tables
- 4.3.2 Components of a simple frequency tables
- 4.3.3 Creating of simple tables through practical examples.
- 4.3.4 Scope and limitations of simple frequency tables

5.0 Session 4 (Day 2: 1300 – 1600hrs): Grouped frequency tables

5.1 Session Description

The grouped frequency table is a statistic method to organize and simplify a large set of data in to smaller "groups." When a data consists of large number of values, it is preferable to group them in a smaller chunk to make it more understandable. This session is basically devoted to discuss various concepts of grouped frequency tables such as class limits, class boundaries, class marks, relative frequency, percentage frequency, frequency density etc and to practice different types of grouped frequency tables.

5.2 Session Learning Outcomes

By the end of this session participants will be able to;

- 5.2.1 understand the nature of a grouped frequency table
- 5.2.2 create an interactive grouped frequency table
- 5.2.3 to understand the scope and limitations of a grouped frequency table

5.3 Brief outline of the session

- 5.3.1 Introduction to grouped frequency tables
- 5.3.2 Related concepts of grouped frequency tables
 - 5.3.2.1 Class limits
 - 5.3.2.2 Class boundaries
 - 5.3.2.3 Class marks
 - 5.3.2.4 Class intervals
 - 5.3.2.5 Relative frequency
 - 5.3.2.6 Percentage frequency
 - 5.3.2.7 Frequency density
- 5.3.3 Types of grouped frequency tables
 - 5.3.3.1 Exclusive types
 - 5.3.3.2 Inclusive types
 - 5.3.3.3 Open and end type
 - 5.3.3.4 Unequal class intervals
- 5.3.4 Creating of grouped tables through practical examples.
- 5.3.5 Scope and limitations of grouped frequency tables

6.0 Session 5 (Day 3: 0900 – 1200hrs): Multivariate frequency tables

6.1 Session Description

The multivariate frequency table is a statistic method to organize and simplify a large set of data of two or more variables in a single table. This session is basically devoted to discuss characteristics of a multivariate frequency table, steps of creating a multivariate frequency tables and to practice various types of multivariate frequency tables.

6.2 Session Learning Outcomes

By the end of this session participants will be able to;

- 6.2.1 understand the nature of a multivariate frequency table
- 6.2.2 create an interactive multivariate frequency table
- 6.2.3 to understand the scope and limitations of a multivariate frequency table

6.3 Brief outline of the session

- 6.3.1 Introduction to multivariate e frequency tables
- 6.3.2 Steps of construction of a multivariate table
- 6.3.3 Marginal distribution of a multivariate frequency distribution
- 6.3.4 Conditional distribution of a multivariate frequency distribution
- 6.3.5 Creating of multivariate tables through practical examples.
- 6.3.6 Scope and limitations of multivariate frequency tables

7.0 Session 6 (Day 3: 1300 – 1600hrs): Cumulative frequency tables

7.1 Session Description

The cumulative frequency table of a set of data is a table which indicates the sum of the frequencies of the data up to a required level. It can be used to determine the number of items that have values below a particular level. This session covers cumulative frequency tables for ungrouped data, cumulative frequency tables for grouped data and calculations of quartiles through cumulative frequency tables.

7.2 Session Learning Outcomes

By the end of this session participants will be able to;

- 7.2.1 understand the nature of a cumulative frequency table.
- 7.2.2 create an interactive cumulative frequency table
- 7.2.3 understand the scope and limitations of a cumulative frequency tables

7.3 Brief outline of the session

- 7.3.1 Introduction to cumulative frequency tables
- 7.3.2 Types of cumulative frequency tables
 - 7.3.2.1 Cumulative frequency tables for ungrouped data
 - 7.3.2.2 Cumulative frequency tables for grouped data
- 7.3.3 Creating of bivariate tables through practical examples.
- 7.3.4 Summery measurements of cumulative frequency tables
- 7.3.5 Scope and limitations of bivariate frequency tables

8.0 Session 7 (Day 4: 0900 – 1200hrs): Bivariate cross tabulation

8.1 Session Description

Bivariate cross tabulation shows various dimensional cross observations of two variables. This session covers Bivariate cross tabulation for qualitative data and quantitative data.

8.2 Session Learning Outcomes

By the end of this session participants will be able to;

- 8.2.1 Understand rules and importance of bivariate cross tabulation
- 8.2.2 create an attractive bivariate cross table.
- 8.2.3 understand the scope and limitations of use in bivariate cross tabulation

8.3 Brief outline of the session

- 8.3.1 Introduction to bivariate cross tabulation
- 8.3.2 Types of cross tables
 - 8.3.2.1 Cross tabulation for qualitative data
 - 8.3.2.2 Cross tables of quantitative data
- 8.3.3 Creating of cross tables through practical examples.
- 8.3.4 Scope and limitations of bivariate cross tabulation

9.0 Session 8 (Day 4: 1300 – 1600hrs): Multivariate cross tabulation

9.1 Session Description

Multivariate cross tabulation shows various dimensional cross observations of more than two variables. This session covers multivariate cross tabulation for qualitative data and quantitative data.

9.2 Session Learning Outcomes

By the end of this session participants will be able to;

- 9.2.1 Understand rules and importance of multivariat cross tabulation
- 9.2.2 create an attractive multivariat cross table.
- 9.2.3 understand the scope and limitations of use in multivariat cross tabulation

9.3 Brief outline of the session

- 9.3.1 Introduction to multivariat cross tabulation
- 9.3.2 Types of cross tables
 - 9.3.2.1 Cross tabulation for qualitative data
 - 9.3.2.2 Cross tables of quantitative data
- 9.3.3 Creating of cross tables through practical examples.
- 9.3.4 Scope and limitations of multivariat cross tabulation

10.0 Session 9 (Day 5: 0900 – 1200hrs): Computer practical I

10.1 Session Description

The main objective of this session is to practice basic tabulation methods by using computer software. This session covers computer practical for frequency tables.

10.2 Session Learning Outcomes

By the end of this session participants will be able to;

- 10.2.1 Understand the way of using computer software for basic tabulation
- 10.2.2 critically analyze information through tabulation

10.3 Brief outline of the session

- 10.3.1 Simple frequency tables
- 10.3.2 Grouped frequency tables
- 10.3.3 Cumulative frequency tables
- 10.3.4 Descriptive and graphical analysis on tabulation

11.0 Session 10 (Day 5: 1300 – 1600hrs): Computer practical II

11.1 Session Description

The main objective of this session is to practice advanced tabulation methods by using computer software. This session covers computer practical for cross tabulations and its applications.

11.2 Session Learning Outcomes

By the end of this session participants will be able to;

11.2.1 Understand the way of using computer software for basic tabulation

11.2.2 critically analyze information through tabulation

11.3 Brief outline of the session

11.3.1 Constructing of cross tables

11.3.2 Graphical analysis on cross tabulation

11.3.3 Statistical analysis on cross tabulation

12.0 Material for further reading and useful web links

12.1 Tabulation of Farad Comparative and Veterinary Pharmacokinetic Data by Arthur L. Craigmill, Jim E.

12.2 Progressive Tabulation Workbook by Jennifer Iles (Paperback - Aug 1985)

12.3 www.measuredhs.com/data/Data-Tabulation-Plan.cfm

12.4 www.popcouncil.org/Horizons/ORToolkit

12.5 www.federalreserve.gov/communitydev/files/cra_datatab.pdf

12.6 www.discoveryresearchgroup.com/data-tabulation/

12.7 www.cfmc.com/services/data-tabulation

13.0 Model MCQ Paper

13.1 Which one of the followings is not one of the advantages in tabulating data?

13.1.1 Concise

13.1.2 No repetition

13.1.3 Comparisons can be done easily

13.1.4 No graphical presentations are further required

13.2 Which one of the followings should not be an essential item of an ideal table?

13.2.1 The date of preparation of the table

13.2.2 The title of the table

13.2.3 The persons who support to create the table

13.2.4 Source of the table

- 13.3 A limitation of data tabulation.
- 13.3.1 Requirement of large space for storing data.
 - 13.3.2 Lack of clarity
 - 13.3.3 Inability of doing comparisons
 - 13.3.4 Not comprehensive
- 13.4 Given the following grouped frequency table, in which interval does the median fall?

Class intervals	Frequency
10 – 14	12
15 – 19	38
20 – 24	20
25 – 29	22
30 – 34	08

- 13.4.1 15 – 19
 - 13.4.2 20 – 24
 - 13.4.3 25 – 29
 - 13.4.4 30 – 34
- 13.5 Which of the following cannot be easily observed in a simple frequency table?
- 13.5.1 the most frequently occurring observations (the mode)
 - 13.5.2 the highest and lowest values scored (the range)
 - 13.5.3 the middlemost score (the median)
 - 13.5.4 the average score (the mean)
- 13.6 Cumulative frequency refers to
- 13.6.1 the total number of observations in a data set.
 - 13.6.2 the total number of frequencies falling at or below a given point in the distribution.
 - 13.6.3 the total number of frequencies falling at or above a given point in the distribution.
 - 13.6.4 the proportion of scores below the median.
- 13.7 An open-ended table is constructed when
- 13.7.1 limited number of observations is scattered over a long interval.
 - 13.7.2 large number of observations is scattered over a short interval.
 - 13.7.3 limited number of observations is scattered over a short interval.

13.7.4 Observations are taken from large number of variables

13.8 Class size is:

13.8.1 the difference between upper and lower class boundaries

13.8.2 the difference between upper and lower class limits

13.8.3 the average between upper and lower class limits

13.8.4 number of observations belonged to a class

13.9 Maximum and minimum values of a sample of data are 150 and 30 respectively. The grouped table with equal class intervals should contain at least 5 and at most 20 class intervals. The size of the class should be:

13.9.1 in between 6 and 24

13.9.2 in between 30 and 150

13.9.3 less than 6

13.9.4 greater than 30

13.10 Class mark is:

13.10.1 the difference between upper and lower class boundaries

13.10.2 the difference between upper and lower class limits

13.10.3 the average between upper and lower class limits

13.10.4 number of observations belonged to a class

13.11 Class mark is a good representative for the:

13.11.1 entire grouped table

13.11.2 mean of the sample

13.11.3 variance of the sample

13.11.4 corresponding class intervals

13.12 Annual incomes of 30 business firms are given below (figures are in Rs. millions)
40 50 51 63 55 70 48 66 77 49 59 63 72 80 53 70 5 99 50 63 77 77 49 50
80 49 55 49 55 61. What is the most appropriate type of grouped table for the given set of data?

13.12.1 Exclusive type

13.12.2 Inclusive type

13.12.3 Open-end type

13.12.4 All of the above

13.13 Answer the questions from 1.8 to 1.9 by using the following table.

Monthly Income (Rs. 000)	Number of households
5000 – 14999	40
15000 – 24999	50
25000 – 34999	100
35000 – 44999	30
45000 – 54999	25
55000 – 64999	15
65000 – 74999	05

Relative frequency of the class interval 35000 - 44999

- 13.13.1 100
- 13.13.2 0.11
- 13.13.3 30
- 13.13.4 05

13.14 Frequency density of the class interval 25000 - 34999

- 13.14.1 0.378
- 13.14.2 100
- 13.14.3 0.1
- 13.14.4 0.01

13.15 Complex tabulation of data that include:

- 13.15.1 only one characteristic
- 13.15.2 two characteristics
- 13.15.3 more than two characteristics
- 13.15.4 hundreds of characteristics

13.16 Which one of the followings is an advantage of textual presentation of data?

- 13.16.1 Too short
- 13.16.2 Comparison can be easily done
- 13.16.3 There is no repetition of words
- 13.16.4 Comprehensive analysis can be done

13.17 Secondary data consists of information:

- 13.17.1 used by competitors
- 13.17.2 that the researcher can obtain through surveys and observation
- 13.17.3 that already exists somewhere and is outdated
- 13.17.4 that already exists somewhere and was collected for another purpose

13.18 Answer the questions from 1.13 to 1.15 by using the following table. (Hint: you are required to make the cumulative frequency table before your calculations)

Class boundaries	90-100	100-110	110-120	120-130	130-140	140-150	150-160
Cumulative frequency (less than)	16	22	45	60	50	24	10

Calculate the number of observation between 112 and 134

- 13.18.1 105
- 13.18.2 177
- 13.18.3 120
- 13.18.4 116

13.19 Calculate the number of observations less than 116

- 13.19.1 83
- 13.19.2 65
- 13.19.3 38
- 13.19.4 22

13.20 Calculate the number of observations greater than 138

- 13.20.1 84
- 13.20.2 34
- 13.20.3 44
- 13.20.4 14

13.21 The number of times a certain event has happened is called:

- 13.21.1 probability
- 13.21.2 fate
- 13.21.3 frequency
- 13.21.4 tally

13.22 'Frequency density' is calculated in the case of:

- 13.22.1 Component bar chart
- 13.22.2 Unequal width histogram
- 13.22.3 Lorenz
- 13.22.4 Equal width histogram

13.23 A key advantage of cross tabulation data.

- 13.23.1 Simplicity
- 13.23.2 Ability in calculating quartiles
- 13.23.3 Ability in comparing more variables through various dimensions.
- 13.23.4 Ability to create easy graphical presentations.

13.24 A teacher needs to identify 10 percent of students who got lowest marks 5 percent of students who got highest marks. What should be the most appropriate data tabulation method for this purpose?

- 13.24.1 A cross table.
- 13.24.2 A grouped frequency table
- 13.24.3 Simple frequency table
- 13.24.4 Cumulative frequency table

13.25 A researcher needs to identify whether there is any disparity in suffering from cancers between male and female populations. What should be the most appropriate data tabulation technique to present information collected from a sample?

- 13.25.1 A cross table.
- 13.25.2 A grouped frequency table
- 13.25.3 Simple frequency table
- 13.25.4 Cumulative frequency table

14.0 Model Take-home project (assignment)

14.1 Nature of the assignment

Participants should create a data presentation report including at least 20 A4 size papers. Relevant tabulation techniques should be applied to present data in a proper manner.

14.2 Source of data

A sample should be selected from your residential area, your working place etc including at least 40 persons.

14.3 Variables

The variables can be selected by your preference such as gender, age, income, occupation, attitudes etc.