Quadrant II

Programme:	Bachelor of Arts
Subject:	History
Semester:	IV
Course Code:	HSS 104
Course Title:	Introduction to Research Skills in History
Title of the Unit:	Research Processes
Module Name:	Processing, Analysis and Interpretation of Data; Role of Statistics in Research
Presenter:	Murari P. Tapaswi,
Affiliation:	Chief Scientist (Retired), CSIR-National Institute of Oceanography, Dona Paula 403004 Goa

Outline of the e-module:

1. Explains different processes involved in makig data ready for the qualitative and quantitative analysis,
2. Highlights with examples the analysis of qualitative and quantitative data,
3. Shows how to interpret analysed data illustrated with examples, and
4. Exhibits the statistical and graphical methods that assist in the interpretation of quantitative data

Learning Outcomes:

At the end of the session, a student will be able to

1. Process, analyse and interpret data so collected for a given study
2. Use simple statistical and graphical methods for the analysis of quantitative data

Short Notes

Quantitative and Qualitative Data: What it means in general

- Quantitative: deal with numbers
  - Comparatively simple to process
- Qualitative: words, descriptions, concepts or ideas
  - Not so easy to document
  - 2 types: Nominal (grouping done based on definite variable e.g. country code attached to phone numbers) and Ordinal (grouping based on feelings of individual e.g. "satisfied, indifferent, dissatisfied")

Data Processing essentials

- Documentation of processes
  - Quality checks
  - Correct values
  - Lengths of rows and number of variables
  - Back-up
- Software: Spreadsheet?; Taguette, QualCoder, etc. in open source – specific for data processing

Quantitative Data Processing

- Simpler as compared to qualitative data process
- Matrix consists of rows and columns
- Rows and columns correspond to observations and variables respectively
- Online inputs automatically store data in matrix on submission. Offline collection would require manual recording in such a matrix

Qualitative Data Processing

- Dry run: experiment with 4-5 responses to list categories irt objectives of the study
- Leave space for unforeseen additions
- Codify: the process of assigning a codes to the categories (variables)
- Compile respective data per category for other resposnes with revisions on categories if needed
- Record specific comments of the respondents in 'voice sheets'

Qualitative Data Analysis

- Carefully listen/read the way respondent expressed to comprehend meaning in the data set
- Clear idea of objectives and familiarity of data
- Avoid bias
- Methods of analysis
  1. Content analysis: Frequency at which the idea is spoken, repeated, searched
  2. Narrative analysis: Listening to people and analysing what that means
     - e.g. Job to be done, no money but have time
       - = The respondent can Self-help
     - e.g. Job to be done, have money but no time
       - = The respondent can hire someone
  3. Discourse analysis: analysing language 'beyond the sentence' within its social context
     - e.g. Fruit is sweet but... (*'but...' in the sentence* hints its not a choice of the person!*)*
     - e.g. That was a nice meal (*look forward for your invitation again!*)
  4. Thematic analysis: Grouping data based on their similarities (themes)

Reactions of respondents	Pattern/Themes		
	Dislike	Like	Neutral
• Current ruling party is hopeless	x		
• Incorrigible fellows. They should be sent home	x		
• What to do? There is no choice			x
• I strongly support		x	
• They largely pose and don't work	x		

  5. Grounded theory (GT): Based on patterns, one can evolve hypotheses
     - e.g. The grounded theory could be built as "large population dislikes ruling party" from the thematic patterns studied in previous slide
  6. Interpretive phenomenological analysis (IPA): Personal experiences of a subject concerning a major life event, an experience or a situation

Phenomenon	As a positive response to the appeal	Only after increased number of pandemic cases
Individuals started wearing masks	10	100

Data Interpretation

- Helps in assigning meaning to the analysed data to arrive at a conclusion
- The size of data is always broad based and as one reaches to the conclusion using this data (with steps of analysis and interpretation) the same is more focused in the final stage

Qualitative Data Interpretation

- Based on the analysis in previous examples, some of the interpretations could be as follows:
- Number of searches on a given term reach to peak when large part of the population gets concerned with it (Content analysis)
- Large portion of the sample size is not in favour of current ruling party's functioning (Thematic analysis)
- Fear among the population resulted in observing preventive practices during pandemic (IPA)

Quantitative Data Interpretation

- Coding technique is not required for numbered data
- Use of statistical techniques is common
- Analysis and interpretation involves statistical modelling techniques such as standard deviation, mean and median
- These values are easily calculated using spreadsheets
- Good to know theory behind it – explained in following slides

- Mean, Median, Mode
  - Average (arithmetic mean) = sum of values/number of variables
  - Median = middle number in a sorted, ascending or descending, list of numbers
  - Mode = value that appears most often in a set of data values
  - From the following Table values, see how these are computed based upon above description

Item	No.
A	1
E	3
C	4
B	5
D	5
Sum	18

$$\text{Mean} = \frac{18}{5} = 3.6$$

← Median

} Mode

- Standard deviation
  - Technique is used to measure how well the responses align with or deviate from the mean (degree of consistency within the responses)

$$SD = \frac{\sum |x-\mu|^2}{N}$$

Where
∑ means "sum of", & √ means square root of
x is a value in the data set,
μ ("mu") is the mean of the data set, and
N is the number of data points in the population

Example,
  - Suppose there are 4 values in data set: 6, 2, 3 and 1
  - Mean = sum of these values no. of values (12/4 = 3)
  - Square of the distance from each value to the mean as in the following Table

x	$\lvert x-\mu\rvert^2$
6	$\lvert 6-3\rvert^2=32=9$
2	$\lvert 2-3\rvert^2=12=1$
3	$\lvert 3-3\rvert^2=02=0$
1	$\lvert 1-3\rvert^2=22=4$

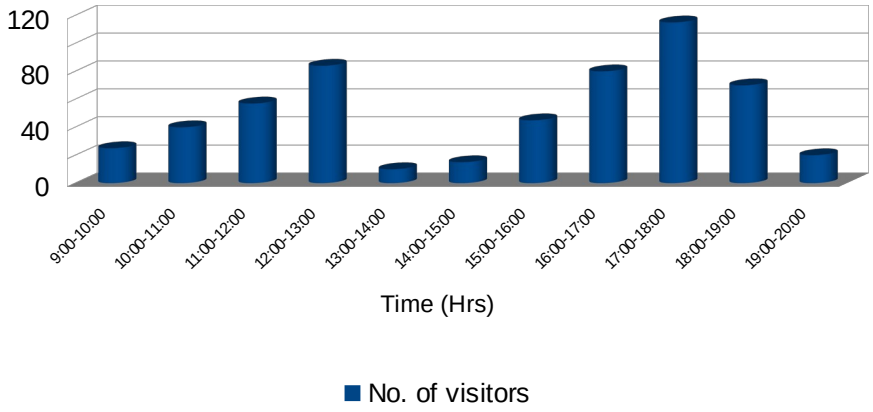
Putting these values in formula

$$SD = \sqrt{\frac{9+1+0+4}{4}} = \sqrt{\frac{14}{4}} = \sqrt{3.5} \approx 1.87$$

- Frequency distribution
  - Present raw data in an organized, easy-to-read format for easy interpretation
  - Identify the most frequently occurring scores, lower and upper limits, cases that are not common, outliers, and total number of observations in any given scores
  - example: Average number of visitors to the museum

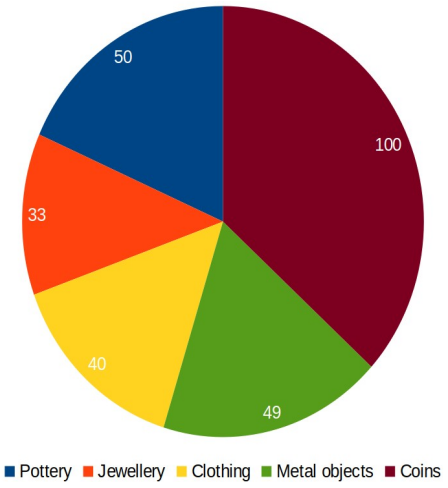
Time (Hrs)	No. of visitors	Peak hours Rank
9:00-10:00	25	8
10:00-11:00	40	7
11:00-12:00	57	5
12:00-13:00	84	2
13:00-14:00	10	11
14:00-15:00	15	10
15:00-16:00	45	6
16:00-17:00	80	3
17:00-18:00	115	1
18:00-19:00	70	4
19:00-20:00	20	9

- Data visualization: Frequency Distribution (Bar graph)

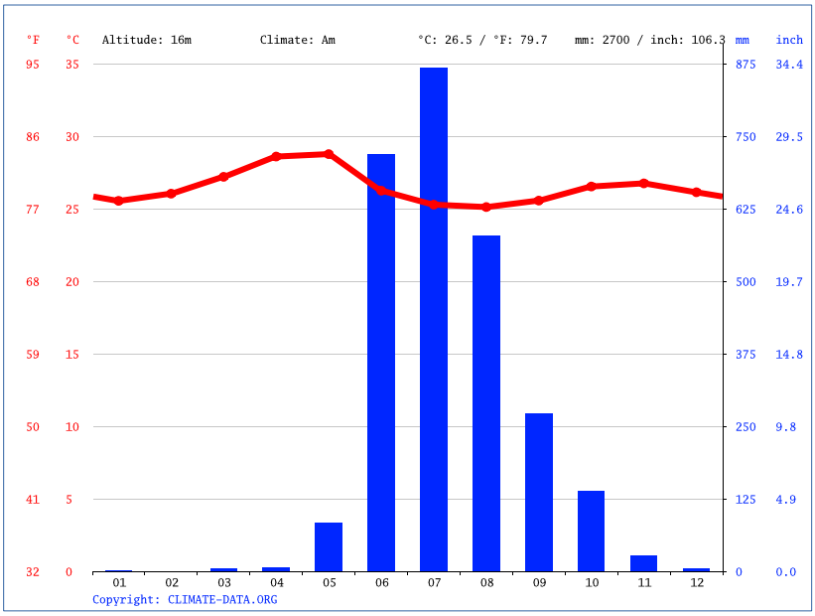


Data visualization: Number of articacts in museum (Distribution Pie chart)

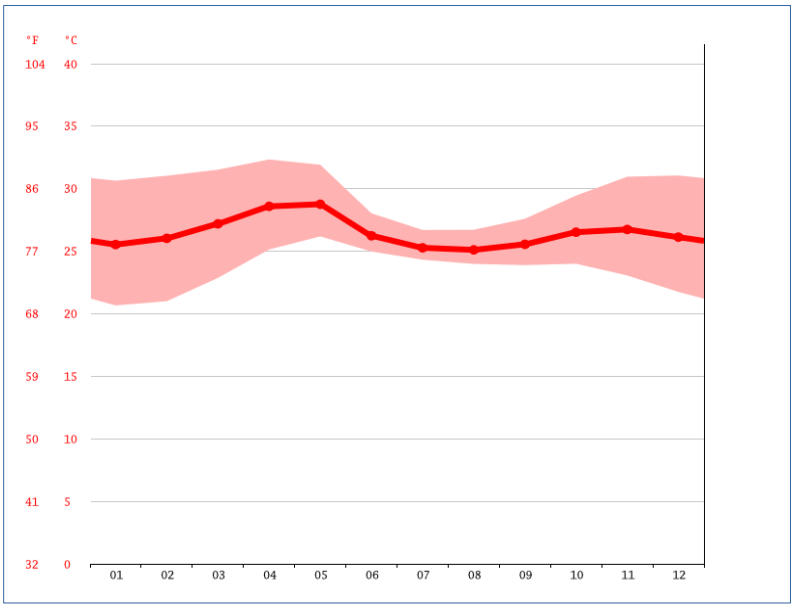
Artifact	**Number**
Pottery	50
Jewellery	33
Clothing	40
Metal objects	49
Coins	100



Data visualization: Distribution of variables
Graph credit: *https://en.climate-data.org/asia/india/goa/dona-paula-795617/#climate-graph (copied on 25 Jul 2021)*



Temperature and rainfall at Dona Paula



Maximum, minimum and average temperature at Dona Paula

Data visualization: Infographics

Researchers from 464 unique organizations in 51 countries including India with whom Goa University had a collaboration during 2014-18

***