Preparing data for analysis

The first stage in research and data analysis is to make it for the analysis so that the nominal data can be converted into something meaningful. Data preparation consists of the below phases.

Phase I: Data Validation

Data validation is done to understand if the collected data sample is per the pre-set standards, or it is a biased data sample again divided into four different stages

1. **Fraud:** To ensure an actual human being records each response to the survey or the questionnaire
2. **Screening:** To make sure each participant or respondent is selected or chosen in compliance with the research criteria
3. **Procedure:** To ensure ethical standards were maintained while collecting the data sample
4. **Completeness:** To ensure that the respondent has answered all the questions in an online survey. Else, the interviewer had asked all the questions devised in the questionnaire.

Phase II: Data Editing

More often, an extensive research data sample comes loaded with errors. Respondents sometimes fill in some fields incorrectly or sometimes skip them accidentally. Data editing is a process wherein the researchers have to confirm that the provided data is free of such errors. They need to conduct necessary checks and outlier checks to edit the raw edit and make it ready for analysis.

Phase III: Data Coding

Out of all three, this is the most critical phase of data preparation associated with grouping and assigning values to the survey responses. If a survey is completed with a 1000 sample size, the researcher will create an age bracket to distinguish the respondents based on their age. Thus, it becomes easier to analyze small data buckets rather than deal with the massive data pile.

Methods used for data analysis in quantitative research

After the data is prepared for analysis, researchers are open to using different research and data analysis methods to derive meaningful insights. For sure, statistical techniques are the most favored to analyze numerical data. The method is

again classified into two groups. First, 'Descriptive Statistics' used to describe data. Second, 'Inferential statistics' that helps in comparing the data.

Descriptive statistics

This method is used to describe the basic features of versatile types of data in research. It presents the data in such a meaningful way that pattern in the data starts making sense. Nevertheless, the descriptive analysis does not go beyond making conclusions. The conclusions are again based on the hypothesis researchers have formulated so far. Here are a few major types of descriptive analysis methods.

Measures of Frequency

1. Count, Percent, Frequency
2. It is used to denote home often a particular event occurs.
3. Researchers use it when they want to showcase how often a response is given.

Measures of Central Tendency

1. Mean, Median, Mode
2. The method is widely used to demonstrate distribution by various points.
3. Researchers use this method when they want to showcase the most commonly or averagely indicated response.

Measures of Dispersion or Variation

1. Range, Variance, Standard deviation
2. Here the field equals high/low points.
3. Variance standard deviation = difference between the observed score and mean
4. It is used to identify the spread of scores by stating intervals.
5. Researchers use this method to showcase data spread out. It helps them identify the depth until which the data is spread out that it directly affects the mean.

Measures of Position

1. Percentile ranks, Quartile ranks
2. It relies on standardized scores helping researchers to identify the relationship between different scores.
3. It is often used when researchers want to compare scores with the average count.

For quantitative market research use of descriptive analysis often give absolute numbers, but the analysis is never sufficient to demonstrate the rationale behind those numbers. Nevertheless, it is necessary to think of the best method for research and data analysis suiting your survey questionnaire and what story researchers want to tell. For example, the mean is the best way to demonstrate the students' average scores in schools. It is better to rely on the descriptive statistics when the researchers intend to keep the research or outcome limited to the provided sample without generalizing it. For example, when you want to compare average voting done in two different cities, differential statistics are enough.

Descriptive analysis is also called a 'univariate analysis' since it is commonly used to analyze a single variable.

Inferential statistics

Inferential statistics are used to make predictions about a larger population after research and data analysis of the representing population's collected sample. For example, you can ask some odd 100 audiences at a movie theater if they like the movie they are watching. Researchers then use inferential statistics on the collected sample to reason that about 80-90% of people like the movie.

Here are two significant areas of inferential statistics.

Estimating parameters: It takes statistics from the sample research data and demonstrates something about the population parameter.

Hypothesis test: It's about sampling research data to answer the survey research questions. For example, researchers might be interested to understand if the new shade of lipstick recently launched is good or not, or if the multivitamin capsules help children to perform better at games.

These are sophisticated analysis methods used to showcase the relationship between different variables instead of describing a single variable. It is often used when researchers want something beyond absolute numbers to understand the relationship between variables.

Here are some of the commonly used methods for data analysis in research.

Correlation: When researchers are not conducting experimental research wherein the researchers are interested to understand the relationship between two or more variables, they opt for correlational research methods.

Cross-tabulation: Also called contingency tables, cross-tabulation is used to analyze the relationship between multiple variables. Suppose provided data has age and

gender categories presented in rows and columns. A two-dimensional cross-tabulation helps for seamless data analysis and research by showing the number of males and females in each age category.

Regression analysis: For understanding the strong relationship between two variables, researchers do not look beyond the primary and commonly used regression analysis method, which is also a type of predictive analysis used. In this method, you have an essential factor called the dependent variable. You also have multiple independent variables in regression analysis. You undertake efforts to find out the impact of independent variables on the dependent variable. The values of both independent and dependent variables are assumed as being ascertained in an error-free random manner.

Frequency tables: The statistical procedure is used for testing the degree to which two or more vary or differ in an experiment. A considerable degree of variation means research findings were significant. In many contexts, ANOVA testing and variance analysis are similar.

Analysis of variance: The statistical procedure is used for testing the degree to which two or more vary or differ in an experiment. A considerable degree of variation means research findings were significant. In many contexts, ANOVA testing and variance analysis are similar.