

CHAPTER SIX

Data Analysis Interpretation



Content

7.1 Data analysis techniques

7.2 Data interpretation

Data Analysis Techniques

- Data analysis is the next step once data gathering is done
- Data analysis is preceded by data preparation and presentation
- **Data preparation** includes editing, coding, and data entry and is the activity that ensures the accuracy of the data and their conversion from raw form to reduced and classified forms that are more appropriate for analysis
- Preparing a descriptive statistical summary is another preliminary step leading to an understanding of the collected data
- Frequency tables, pie charts, bar charts &c are among the common **data presentation** methods

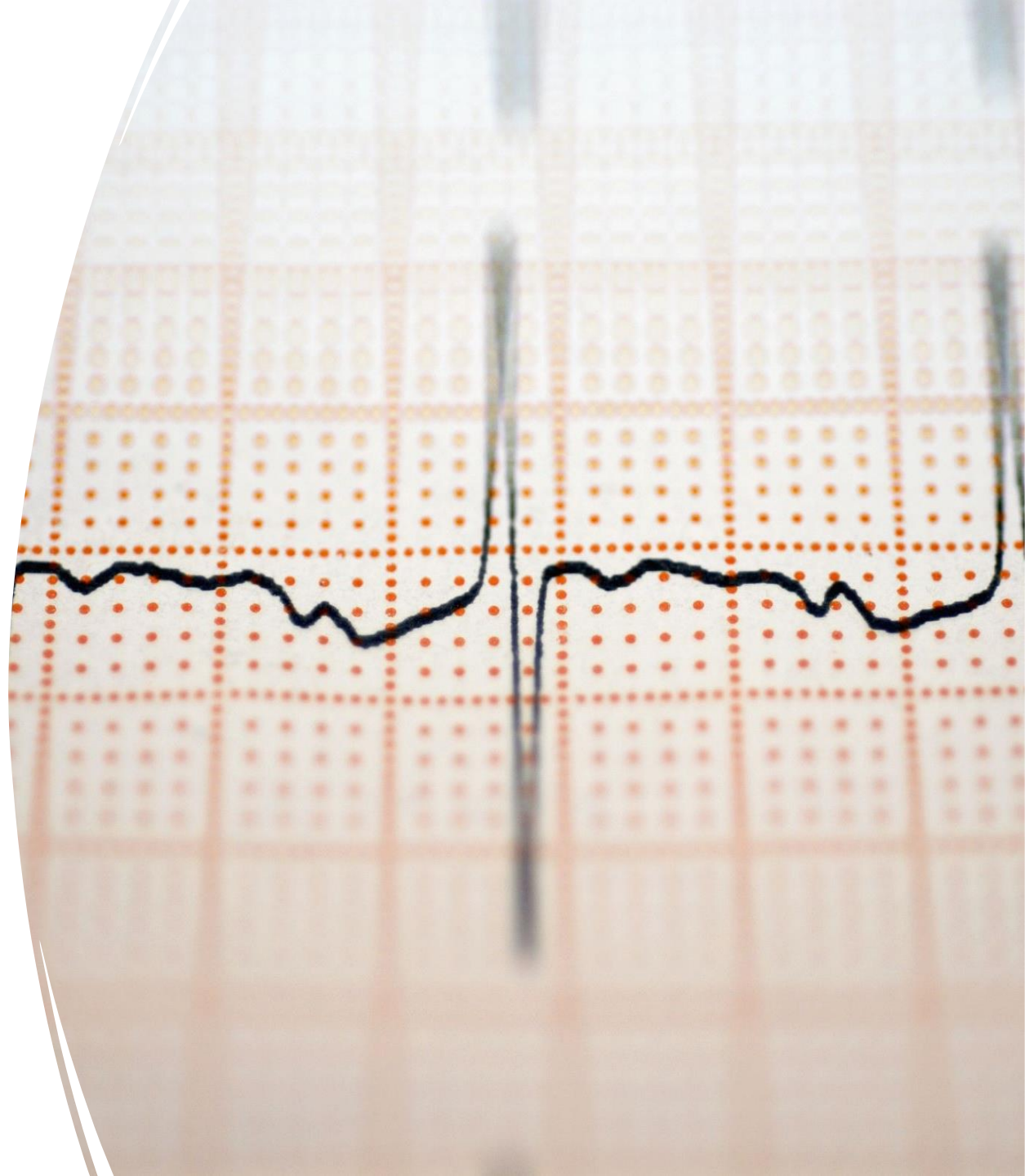


Data Analysis ...

- Different data analysis techniques are applied for quantitative and qualitative data types
- Quantitative data are analyzed basically using statistical methods
- Descriptive statistics is applied mainly as a part of data presentation
 - Part of a preliminary analysis
- Inferential statistics is the second of two major categories of statistical procedures, the other being descriptive statistics
 - **Analytical researches usually rely on inferential statistics to test hypotheses**

Data Analysis...

- The technique of statistical data analysis that are applied to test hypotheses varies depending on the nature of quantitative data we gathered:
 - Data generated through nominal and ordinal scale = nonparametric tests
 - Interval and ratio scale data = parametric tests





When do we use parametric tests?

- Assumptions for parametric tests include the following:
 - The observations must be independent—that is, the selection of any one case should not affect the chances for any other case to be included in the sample.
 - The observations should be drawn from normally distributed populations.
 - These populations should have equal variances.
 - The measurement scales should be at least interval so that arithmetic operations can be used with them.
-

Nonparametric tests

- Nonparametric tests have fewer and less stringent assumptions.
- They do not specify normally distributed populations or equality of variance. Some tests require independence of cases; others are expressly designed for situations with related cases.
- Nonparametric tests are the only ones usable with nominal data; they are the only technically correct tests to use with ordinal data, although parametric tests are sometimes employed in this case.
- A variety of nonparametric tests may be used, depending on the measurement scale and other conditions.
- If the measurement scale is nominal (classificatory only), it is possible to use either the binomial test or the chi-square (χ^2) one-sample test.

Nonparametric tests...

- The binomial test is appropriate when the population is viewed as only two classes, such as male and female, buyer and nonbuyer, and successful and unsuccessful, and all observations fall into one or the other of these categories.
- The binomial test is particularly useful when the size of the sample is so small that the x^2 test cannot be used.
- Typical uses of chi-square are cases where persons, events, or objects are grouped in two or more nominal categories such as “yes–no,” “favor–undecided–against,” or class “A, B, C, or D.”
- Using this technique, we test for significant differences between the *observed* distribution of data among categories and the *expected* distribution based on the null hypothesis.

Relational Hypotheses

- Management questions frequently revolve around the study of relationships between two or more variables. Then, a *relational hypothesis* is necessary.
- For example, you may be interested to see what happens in level of job satisfaction and employee absenteeism in workplace, to analyze the perception of footwears consumer in Ethiopia for domestic products and imported products etc...
- Relational hypotheses are commonly handled with measures of association (correlation) and causation (or regression)

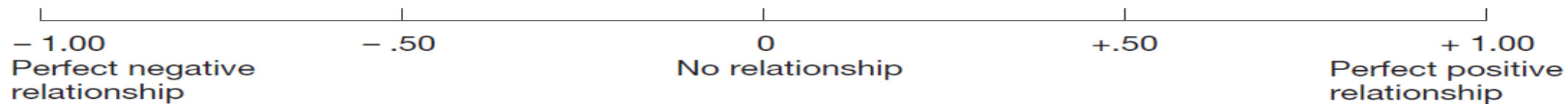
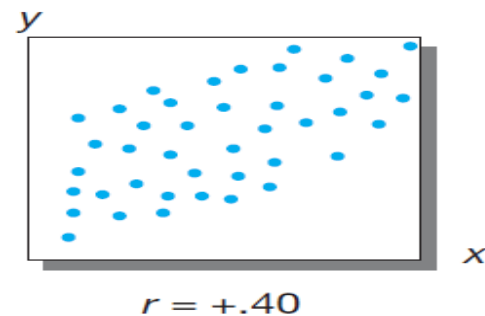
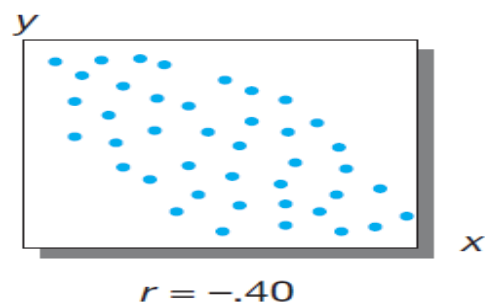
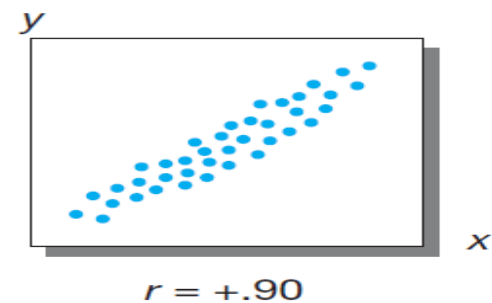
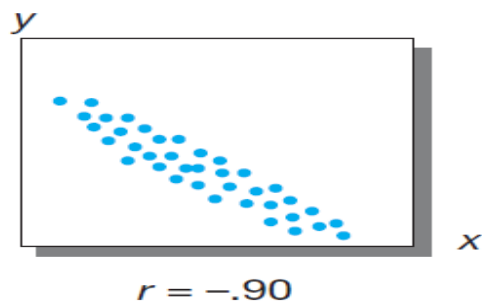
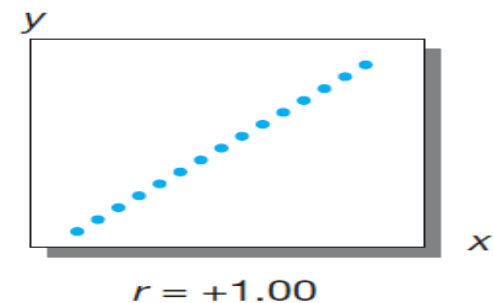
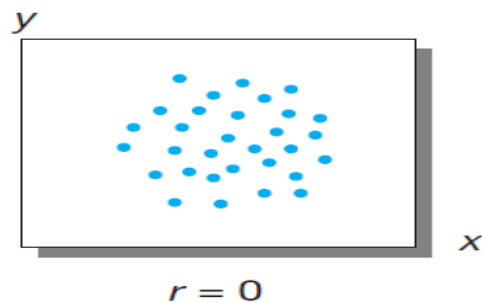
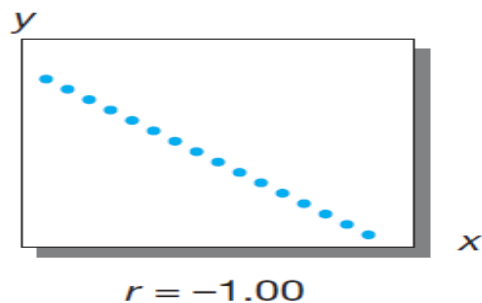
Measures of Association

- Measure of association in a correct relational hypothesis would state that the variables occur together in some specified manner without implying that one causes the other.
- Correlation analysis would test the strength, direction, shape, and other features of the relationships but not causation.
- In Pearson (product moment) correlation analysis, the correlation coefficients reveal the magnitude and direction of relationships.
- The *magnitude* is the degree to which variables move in unison or opposition.

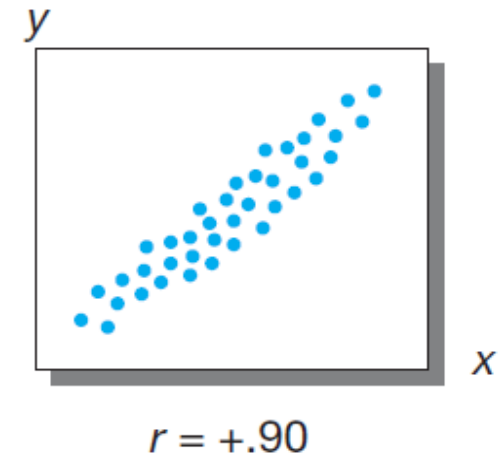
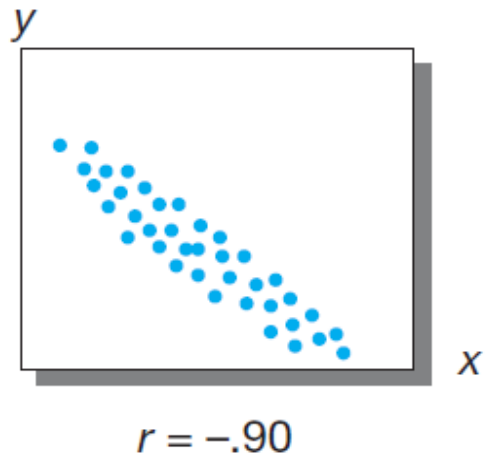
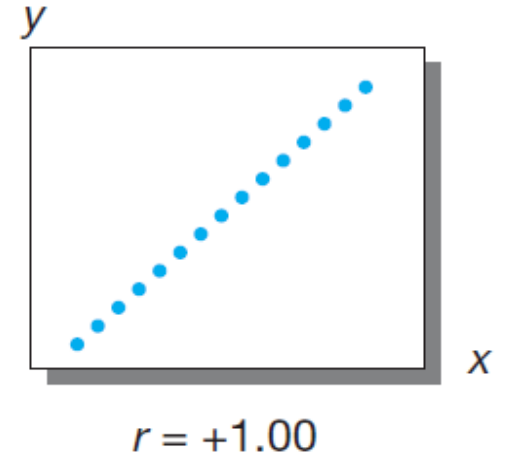
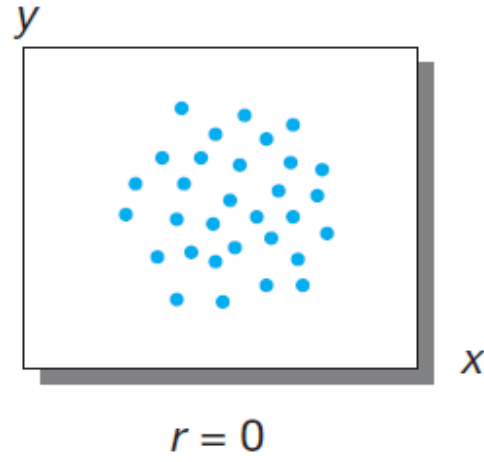
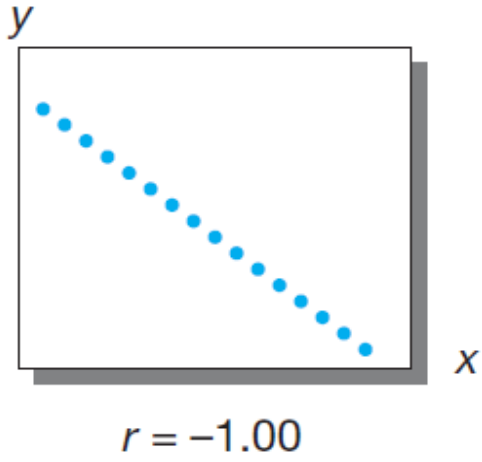
Measures of Association...

- The size of a correlation of $+0.40$ is the same as one of -0.40 .
- The sign says nothing about size. The coefficient's sign signifies the *direction* of the relationship.
- The **Pearson** (product moment) **correlation coefficient** varies over a range of -1 through 0 to $+1$. The designation r symbolizes the coefficient's estimate of linear association based on sampling data. The coefficient ρ represents the population correlation

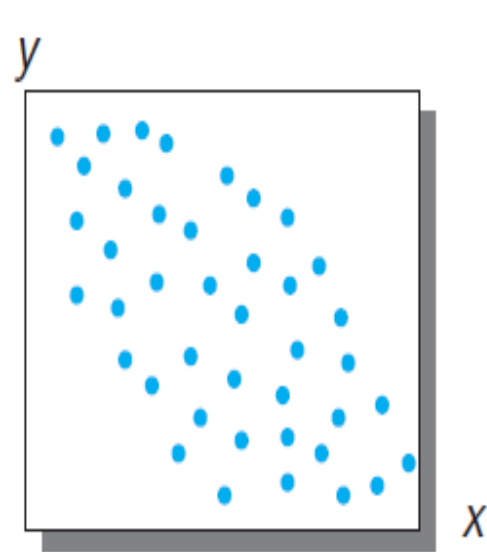
Scatter Plot



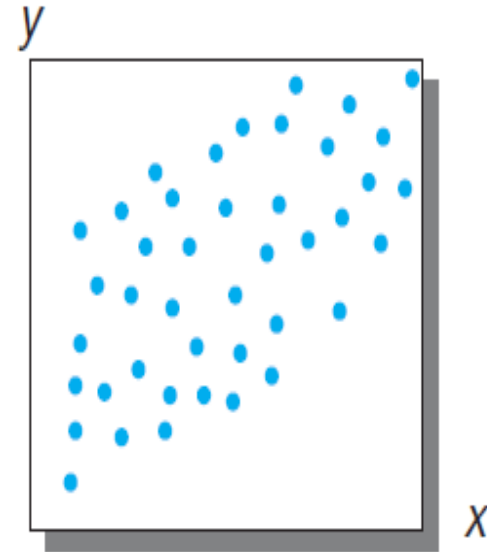
Scatter Plot



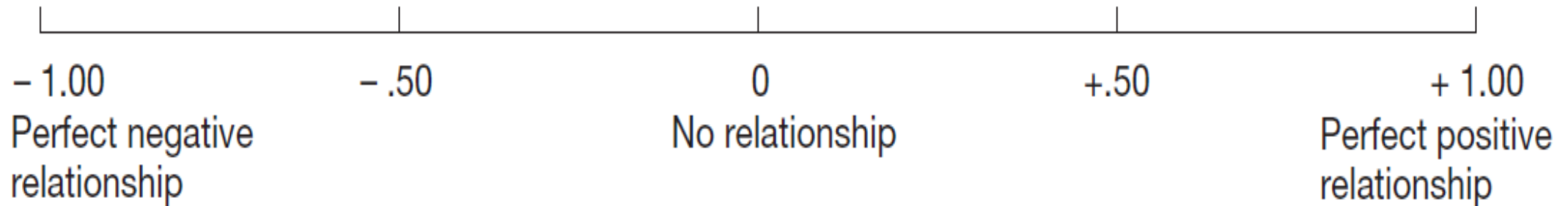
Scatter Plot



$$r = -.40$$

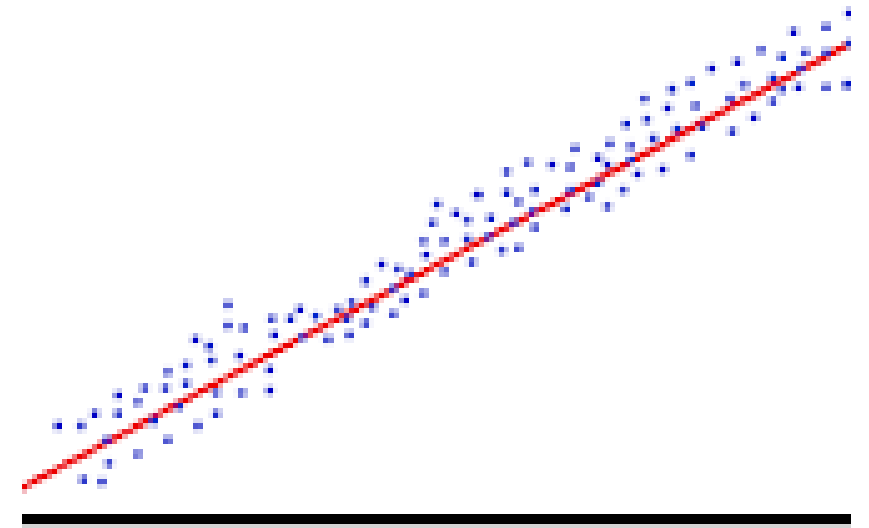


$$r = +.40$$



Regression Analysis

- When we take the observed values of X to estimate or predict corresponding Y values, the process is called **simple prediction**.
- When more than one X variable is used, the outcome is a function of multiple predictors.
- Simple and multiple predictions are made with a technique called **regression analysis**.
- **The basic model of the regression analysis:**
- A straight line is fundamentally the best way to model the relationship between two continuous variables.



Regression Analysis...

- The bivariate linear regression may be expressed as:

$$Y = \beta_0 + \beta_1 X_i$$

- where the value of the dependent variable Y is a linear function of the corresponding value of the independent variable X_i in the i th observation.
- The slope and the Y intercept are known as **regression coefficients**.
- The **slope**, b_1 , is the change in Y for a 1-unit change in X . It is sometimes called the “rise over run.”

Interpretation

- After collecting and analyzing the data, the researcher has to accomplish the task of drawing inferences followed by report writing.
- The whole purpose of doing research may get vitiated if careful interpretation of the findings and conclusions are not made.
- It is only through interpretation that the researcher can expose relations and processes that underlie his findings.
- In case of hypotheses testing studies, if hypotheses are tested and upheld several times, the researcher may arrive at generalizations. But in case the researcher had no hypothesis to start with, he would try to explain his findings on the basis of some theory.
- This may at times result in new questions, leading to further researches

Interpretation...

- Interpretation refers to the task of drawing inferences from the collected facts after an analytical and/or experimental study. In fact, it is a search for broader meaning of research findings.
- The task of interpretation has two major aspects viz., (i) the effort to establish continuity in research through linking the results of a given study with those of another, and (ii) the establishment of some explanatory concepts.
- Interpretation also extends beyond the data of the study to include the results of other research, theory and hypotheses