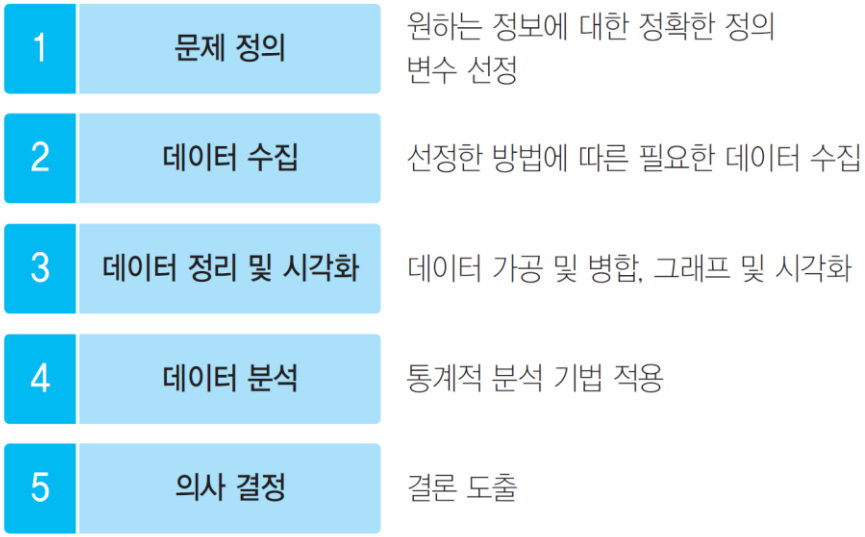
Basic Descriptive Statistics

What is statistics?

- **A set of methods of inquiry used to collect or process or interpret quantitative data**
- Collections of data gathered by those methods
- Certain specially calculated figures (e.g., average) that characterize a collection of data



Types of statistics

- Descriptive statistics
  - Summarizing or describing a sample
  - Characterizing a collection of data
- Inferential statistics
  - Generalizing a sample to make estimates and inferences about a wider population
  - Reliability of generalization depends on how well the sample mirrors the population

Population & Sample

- Population
  - A collection of data whose properties are of interest and to be described
- Sample
  - A part of the population of interest selected to represent the population
- Some errors are inevitable in samples

Data & information

- Examples of Data
  - Weight, height
  - Sales volume
  - Sales and net profit

Types of data (revisited)

- Numerical data
  - Continuous data
    - Weight, height, score… (measurement)
  - Discrete data
    - The number of something… (count)
- Categorical data
  - Ordinal data
    - Grade (A, B, C…), satisfaction survey…
  - Nominal data
    - Gender, (birth)place…

Question

- 스마트폰을 사용하고 있는 기간 : 7,120일
- 지금까지 고장 횟수 : 3회
- 스마트폰을 구매한 가격 : 12만 원
- 스마트폰의 종류 : 갤럭시 S
- 스마트폰 등급 : 저가
- 스마트폰 색 : 은색

- For the data above, guess correct data types.
- 고장횟수: discrete
- 구매 가격: continuous
- 종류: nominal
- 등급: ordinal
- 색: nominal

Variables: input vs output

- Independent variable
  - a.k.a. predictor variable, treatment variable, manipulated variable
  - input, conditions
  - the presumed cause in an experiment
  - its values are under control by experimenter
  - e.g., gender, age, speaker types, designated contexts
- Dependent variable
  - a.k.a. response variable
  - output, observations
  - the presumed effect in an experiment
  - its values change as a result
  - e.g., pitch, loudness, duration

Questions

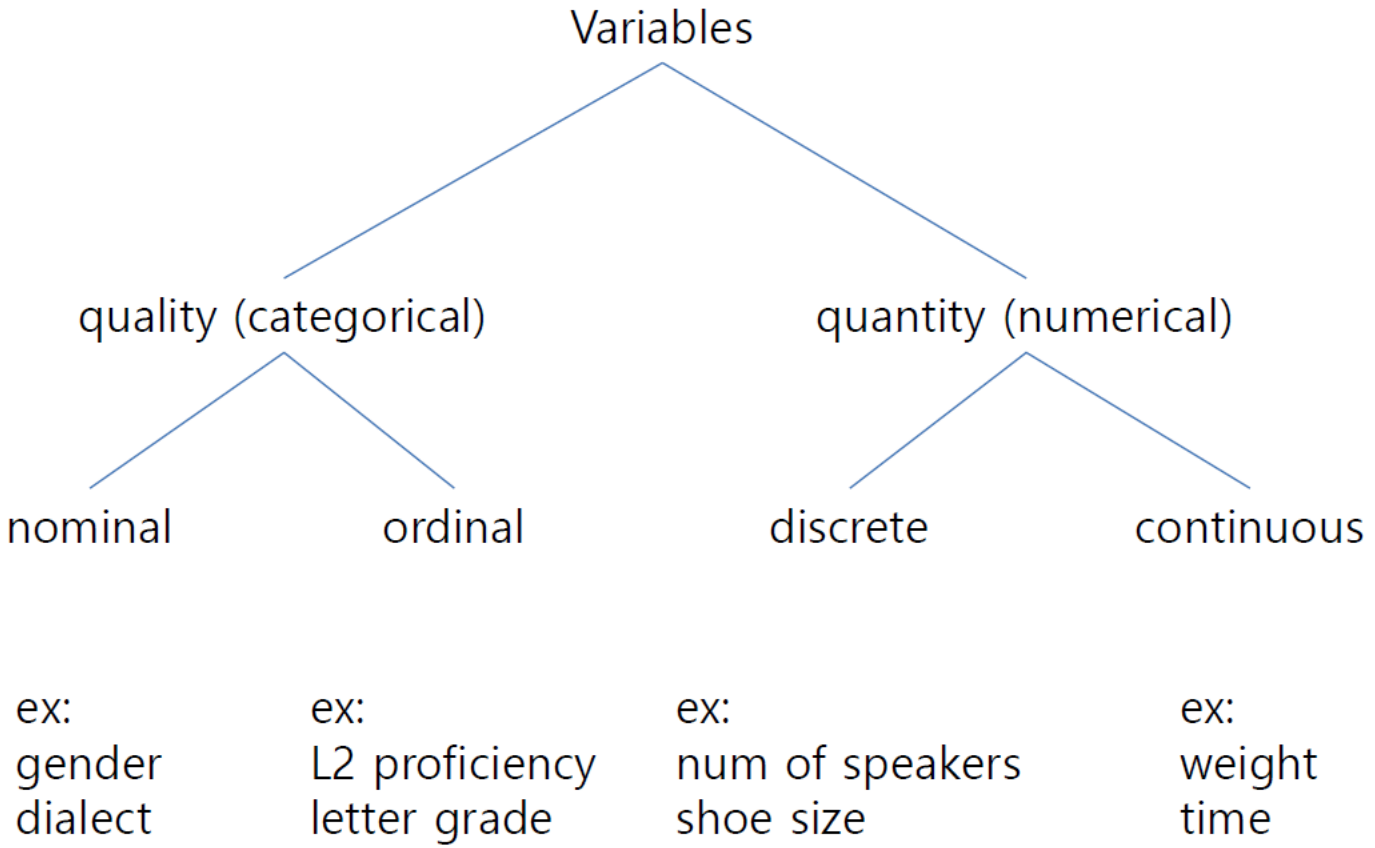
- An experiment is designed to verify whether any of gender, weight, age and profession influences the pitch of a speaker. What are independent variables and what are dependent variable?

- For a quantitative analysis, which variable needs to be a numerical format, predictors, or responses?

Variables: types

- Qualitative (categorical)
  - Groups or categories
    - Ordinal
      - Example: age span, L2 proficiency (novice, intermediate, advanced)
    - Nominal
      - Example: gender, nationality
- Quantitative (numerical)
  - Numerical values
    - Continuous
      - Example: age, weight
    - Discrete
      - Example: number of speakers, shoe size (8, 8.5, 9…)

Variables



- Which type does each of the following variables belong to?: month, age, clothes' size