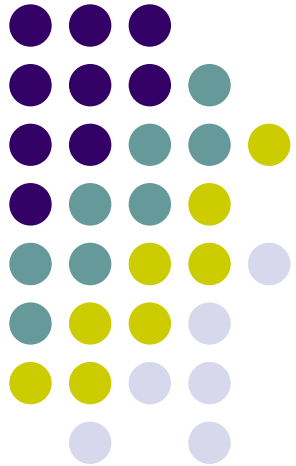
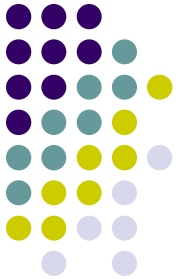
Econ 3790: Business and Economics Statistics

Instructor: Yogesh Uppal
Email: yuppal@ysu.edu

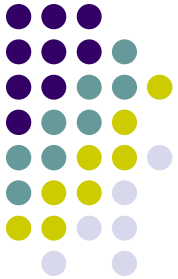


Chapter 1

- Goals of the course
- Data and statistics

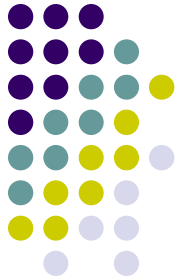


Why use Statistics?



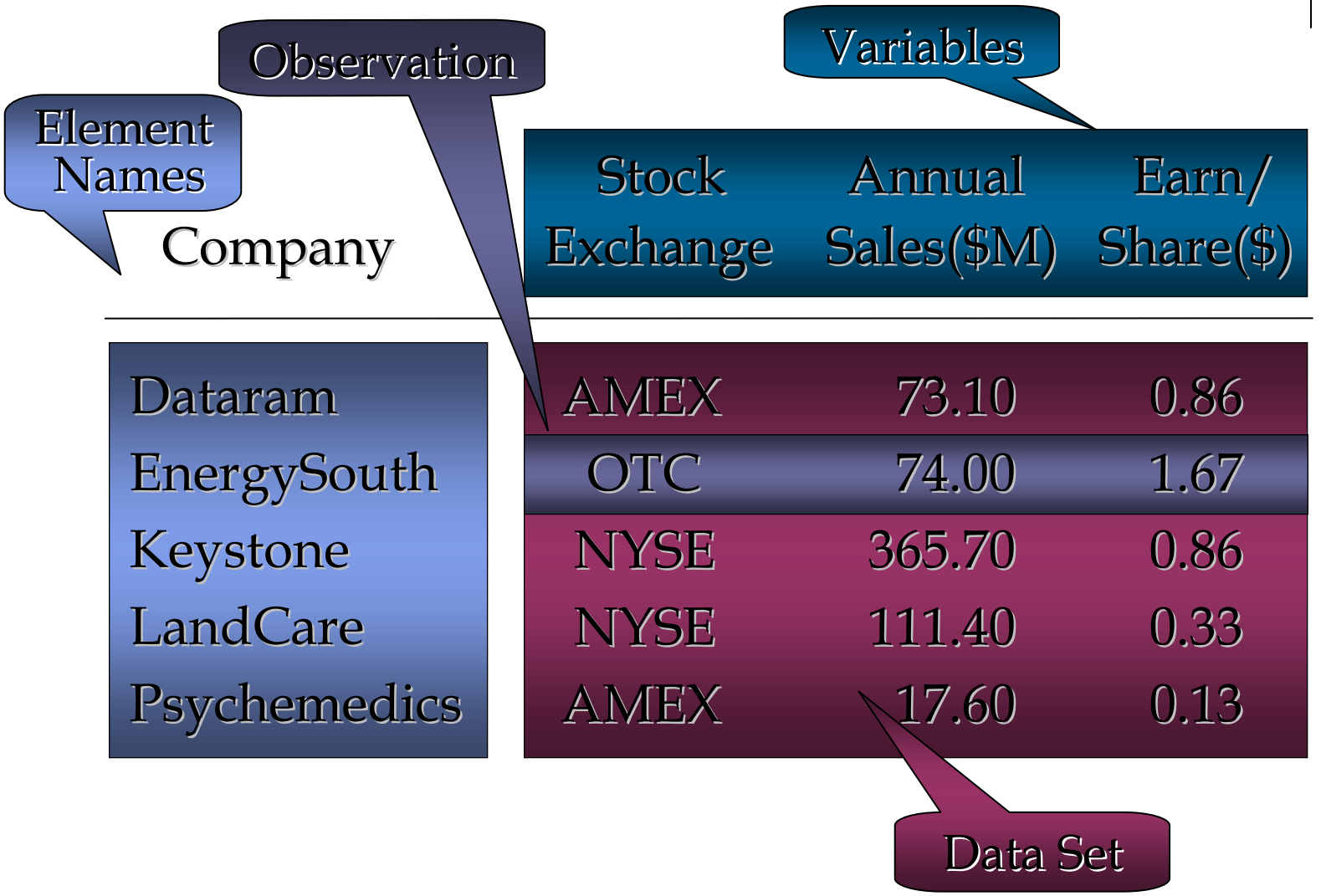
- To make sense of large amounts of data:
  - What are the demographics of Youngstown in 2000?
  - Have U.S. wages increased since 1975?
- To test hypotheses:
  - Is demand curve downward sloping?
  - Are GDP and Saving Rate positively correlated?
- To make predictions:
  - What might happen to savings behavior after a large tax cut?

Data: Basic Definitions

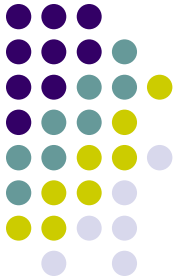


- Data: a set of measurements
- Dataset: all data collected for one study
- Element, or unit: an entity on which data are collected
- Variable: a property or attribute of each unit
- Observation: the values of all variables for one unit

Data: Basic Definitions



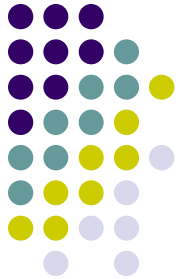
Data: Scales of Measurement



- Four scales of measurement:
  - Nominal, ordinal, interval, and ratio scales

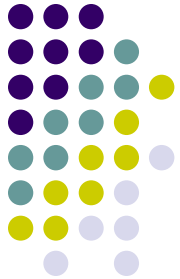
- Scale determines which methods of summarization and analysis are appropriate for any given variable

Data: Scales of Measurement



- Characteristic
  - Nominal, like a label or name for a characteristic
    - e.g., color: red, green, blue
    - race: black, Hispanic, white, Asian
    - binary: (male, female), (yes, no), (0, 1)
  - Ordinal, still a characteristic, but having a natural order
    - e.g., how was service?: poor, average, good

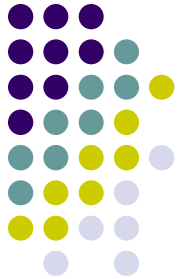
Data: Scales of Measurement



- Numeric
  - Interval scale
    - Numeric data showing the properties of ordinal data
    - e.g., SAT scores, Fahrenheit temperature
  - Ratio scale
    - Ordered, numeric data with real zero
    - e.g., income, distance, price, quantity

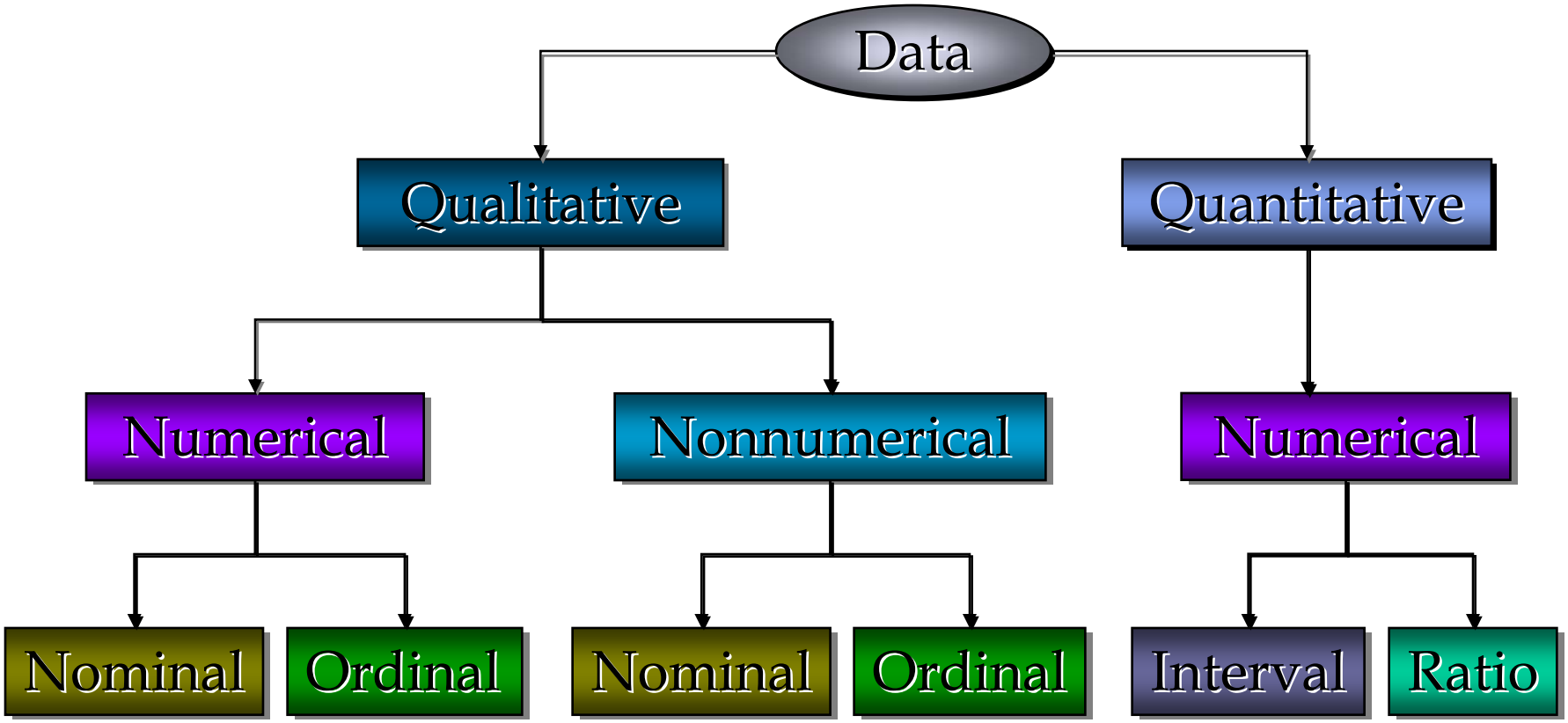
    - http://www.math.sfu.ca/~cschwarz/Stat-301/Handouts/node5.html

Data: Other Classifications

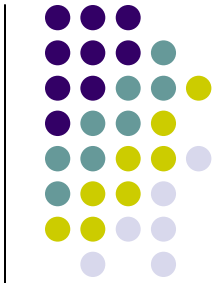


- Qualitative, or categorical: measures a quality
- Quantitative: numeric values that indicate how much or how many
- Cross-sectional: data collected at one point in time
- Time series: data collected over several time periods
- Panel or longitudinal: combination of cross-sectional and time series

Data: Summary of Definitions

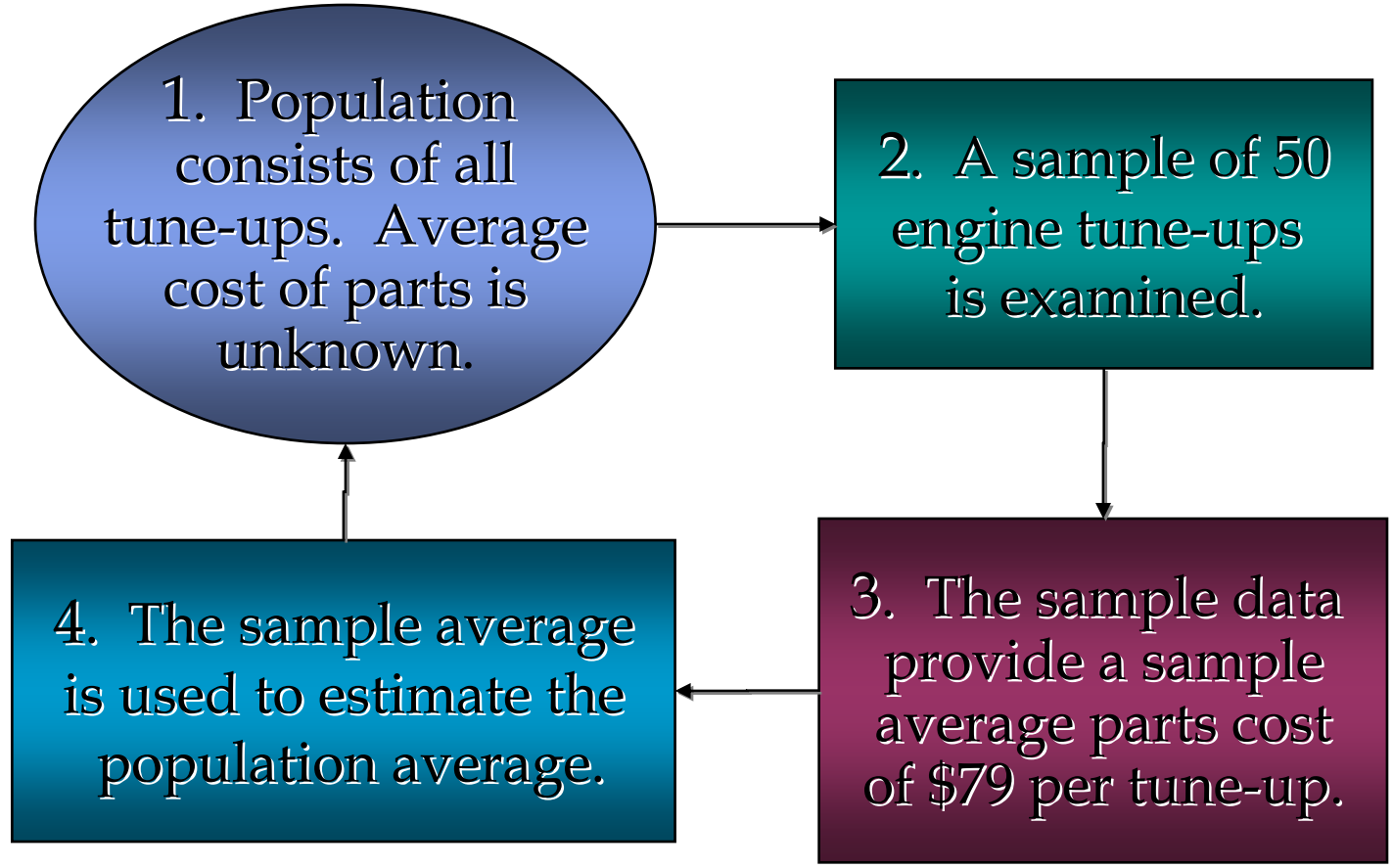


Statistical Inference: Definitions

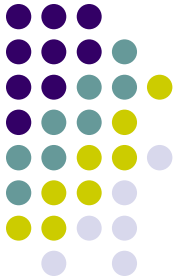


- Population: the set of all elements of interest in a study
- Sample: a subset of the population
- Statistical Inference: the process of using data obtained from a sample to make estimates and test hypotheses about the characteristics of a population

Statistical Inference: Process

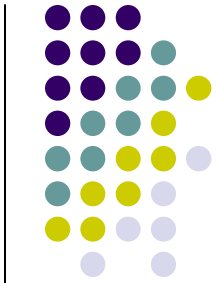


Descriptive Statistics: Definition



- Descriptive statistics are the tabular, graphical, and numerical methods used to summarize data

Descriptive Statistics: Common Methods



- Some common methods:
  - Tabular
    - Frequency table (for one variable)
    - Crosstabulation, or crosstab (for more than one variable)
  - Graphical
    - Bar graph (for categorical variables)
    - Histogram (for interval- or ratio-scaled variables)
    - Scatterplot (for two variables)
  - Numerical
    - Mean (arithmetic average)