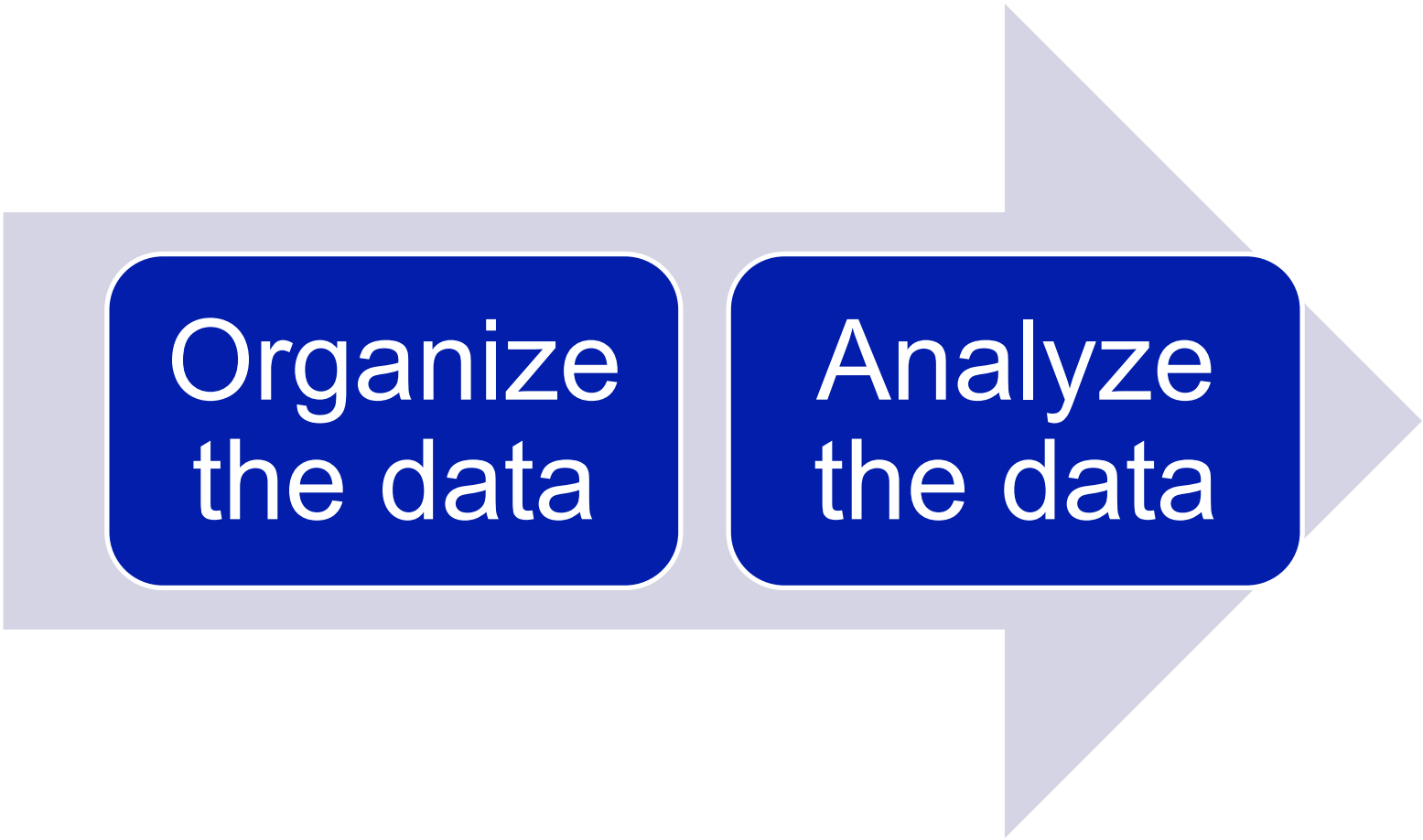
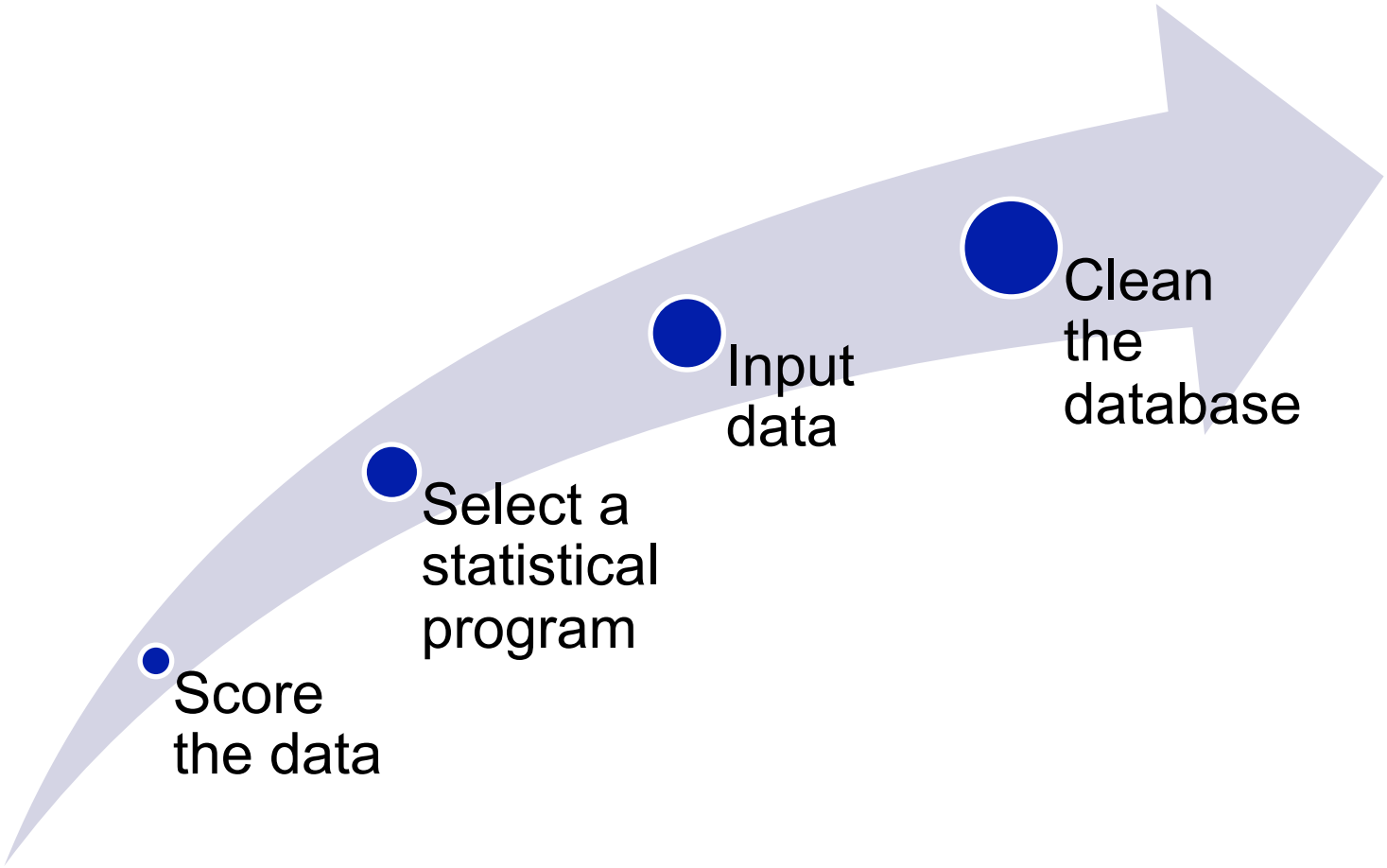
Steps for Quantitative Data Analysis



Yuhei Inoue, Ph.D.
The University of Memphis

THE UNIVERSITY OF MEMPHIS

Step 1: Organize the Data



Score the Data

- Assign a numeric score to each response category, such as:
  - 1 = Male, 2 = Female
  - 1 = strongly disagree, 2 = disagree, 3 = neutral, 4= agree, 5 = strongly agree
- Use the consistent numbering system for the same scale

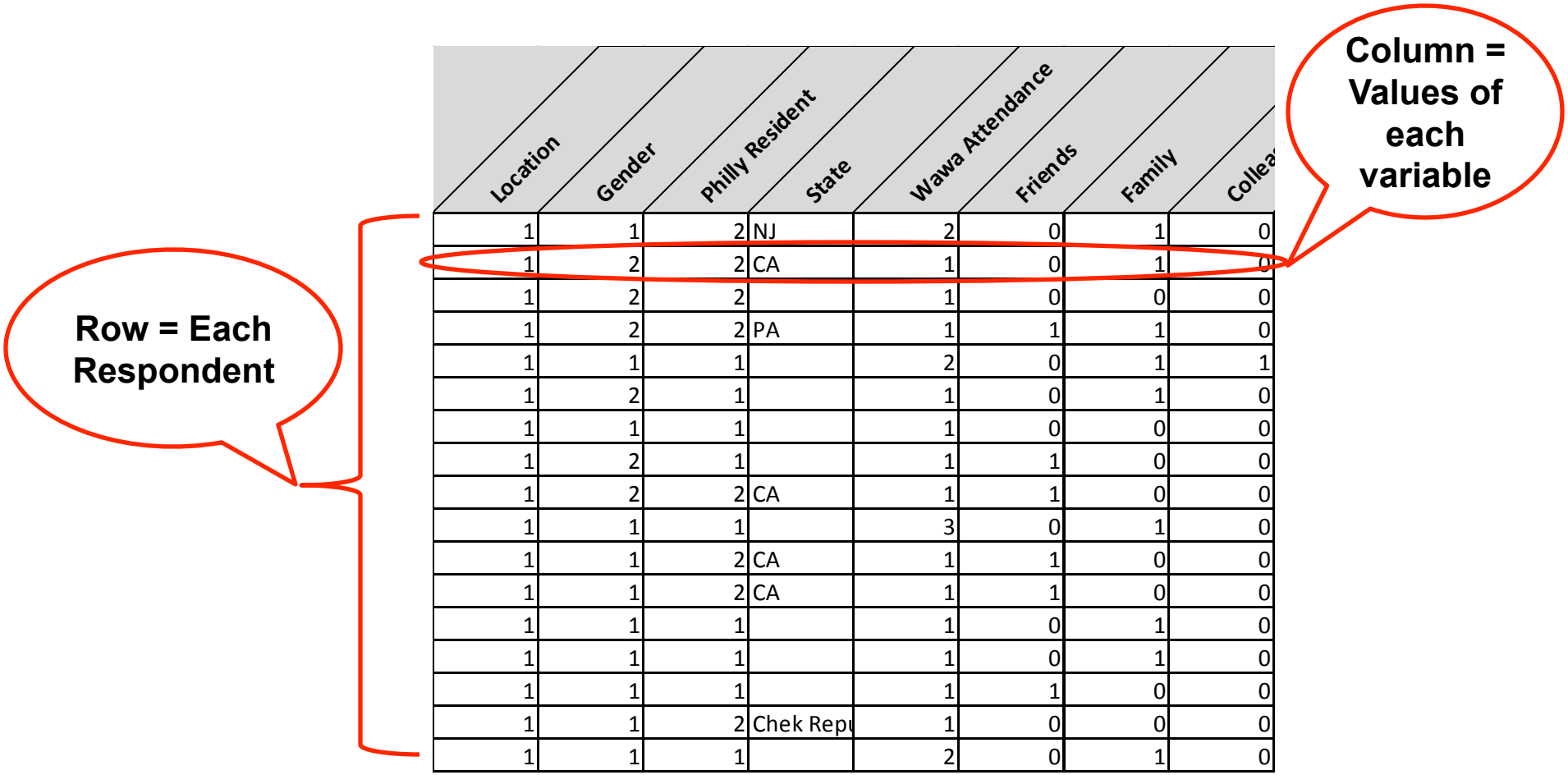
Select a Statistical Program

- Excel
- SPSS
- STATA
- SAS
- AMOS
- Mplus

Input Data

- Assign names to variables (or items)
  - e.g., “How old are you?” → “AGE”
- Enter the data from scores on the questionnaires
  - Rows – each individual / Columns – values of each variable

Sample Data File



Location	Gender	Philly Resident	State	Wawa Attendance	Friends	Family	Collea
1	1	2	NJ	2	0	1	0
1	2	2	CA	1	0	1	0
1	2	2		1	0	0	0
1	2	2	PA	1	1	1	0
1	1	1		2	0	1	1
1	2	1		1	0	1	0
1	1	1		1	0	0	0
1	2	1		1	1	0	0
1	2	2	CA	1	1	0	0
1	1	1		3	0	1	0
1	1	2	CA	1	1	0	0
1	1	2	CA	1	1	0	0
1	1	1		1	0	1	0
1	1	1		1	0	1	0
1	1	1		1	1	0	0
1	1	2	Chek Repu	1	0	0	0
1	1	1		2	0	1	0

Clean the Database

- Find scores that are outside the accepted range
  - e.g., Gender = 3; IQ = 300
- Address missing data
  - Eliminate participants with many missing scores
  - Replace each missing value with an average score for the question for all participants (continuous scores only)

Step 2: Analyze the Data

- Choose a statistical analysis that MOST APPROPRIATELY answers your research question
- Types of statistical analysis
  - **Descriptive analysis:** Describe general characteristics of participants
  - **Inferential statistics:** Examine the relationship among variables

Descriptive Analysis

Frequency analysis

- e.g., How many male and female participants were in the dataset? What is the proportion of male and female respondents?

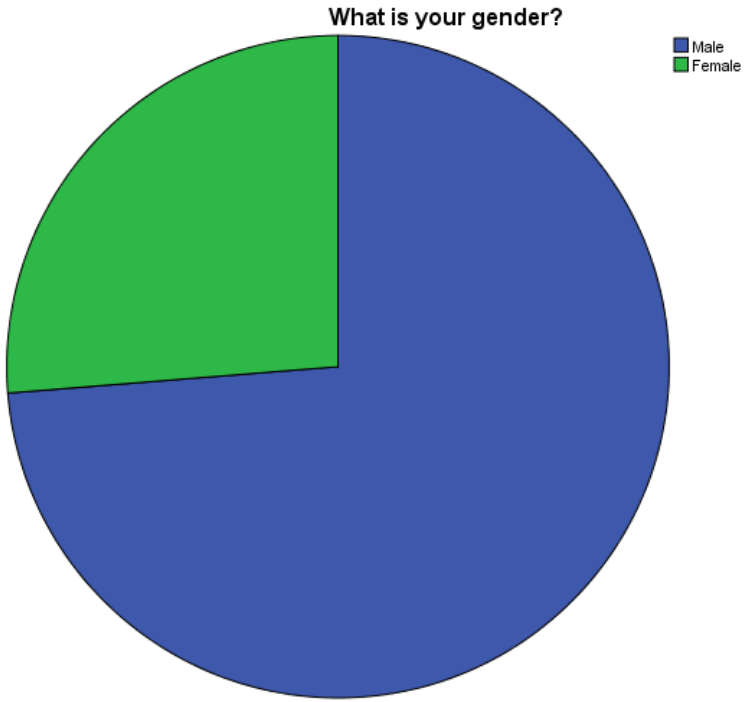
Descriptive analysis

- e.g., What is the average score for this question? How much do scores vary among participants?

Frequency Analysis

SEX What is your gender?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	536	73.7	73.7	73.7
	Female	191	26.3	26.3	100.0
	Total	727	100.0	100.0	



Descriptive Analysis

Central tendency

- Mean = "Average"
- Median = "Midpoint"
- Mode = "Most frequent value"

Variability

- Standard deviation = indicator of how much the data distribution varies
- Range = (maximum value) – (minimum value)

Inferential Statistics

- **Test for group difference**: T-test, ANOVA, MANOVA
  - Question asked – Does annual income differ based on gender?
- **Test for relationships**: Correlation, regression, structural equation modeling
  - Question asked – Is annual income associated with years of education?

Test for Group Difference

Tests	Number of Groups for Comparison	Number of Dependent Variables Compared
T-test	2	1
ANOVA	3 or more	1
MANOVA	2 or more	2 or more

Test for Relationships

Tests	Questions Answered
Correlation	How are two variables related to each other?
Bivariate regression	How does ONE independent variable predict ONE dependent variable?
Multiple regression	How do TWO OR MORE independent variables predict ONE dependent variable?
Structural equation modeling	How do TWO OR MORE independent variables predict TWO OR MORE dependent variables?

Some Points to Consider

- Assumption check
- Sample size
  - Consider power analysis: Determine sample size based on pre-determined effect size and statistical power to ensure statistical significance
- Statistical significance vs. practical significance
- Bias / subjectivity in quantitative analysis