Assignment 1

Dr. Supaporn Erjongmanee

Department of Computer Engineering
Kasetsart University
fengspe@ku.ac.th

1

Assignment 1

One of your result from 2-5
must use shared transportation data

- What to do
  - Use your budget data to compute
    1. Visualization
       - At least three graphs (one for time-series, one for part-to-whole, one for deviation)
    2. Hypothesis test on two data sets
    3. Anova on single factor
    4. Anova on two factors (additive)
    5. Categorical data analysis (homogeneity)

2

Assignment 1 (cont.)

- Individual work
- Deadline: 20 February 2020, 9:00
- Submit in Google Classroom: Assignment 1
- What to submit
  1. Report (filename_format: studentID_report):
     - Can be Thai or English
     - Use written language , not spoken language, to write
  2. Computing_program (filename_format: studentID_program)
     - If you have multiple programs, compress into one file.

Assignment 1 (cont.)

- Topics in report
  - Part 1: Data collection (from … to … , duration)
  - Part 2: Basic findings (graphs and their description)
    - For each graph,
      - Explain objective of your graph
      - Explain why you choose to create such graph
      - How can you summarize your spending pattern from the constructed graph

Assignment 1 (cont.)

- Topics in report (cont.)
  - Part 3: Complex findings: report the following problem
    - 3.1 Hypothesis test on two data sets
    - 3.2 Anova on single factor
    - 3.3 Anova on two factors (additive)
    - 3.4 Categorical data analysis (homogeneity)

    You can also perform multiple comparisons.
  - For each complex finding problem in part 3,
    - a) Explain what you try to prove in words and write null and alternative hypotheses

5

Assignment 1 (cont.)

- Topics in report (cont.)
  - For each complex finding problem in part 3 (cont.)
    - b) Show your rearranged data to solve this problem
    - c) Show your computation steps (how to derive test statistics, how to decide whether you would reject null hypothesis or not (rejection region or p-value*), your rejection decision
    - d) Conclude your result (what you find out in this problem.)
  - Part 4: Summarize all of your findings

*** There are 4 problems, use rejection region for two problems, and p-value for two problems.

6