

## DATA ANALYSIS (ISOM 5510), FALL 2020

**Instructor:** Yingying Li

**Email:** [yyli@ust.hk](mailto:yyli@ust.hk)

**Office Hours:** By appointment (email). Zoom ID: 983 0239 0866

### Course Materials

- Lecture notes and exercise questions are downloadable from course website (<https://canvas.ust.hk>);
- Reference book: “Statistics for Business, Decision Making and Analysis”, by R. Stine and D. Foster;
- Software: Excel.

### Evaluation

- A. Participation 5%.
- B. In-class exercises 40%. Group work (about 2-3 students per group)
- C. Final exam 55%. (All MC questions)

### Course Organization

#### Accelerator

- Data and Variation (Ch2 - Ch6)
  - Types of data, §2.1,2.2,2.4
  - Describing categorical data, §3.1-3.4
  - Association between categorical variables, §5.1-5.2
  - Describing numerical data, §4.1-4.4
  - Association between numerical variables, §6.1-6.3, 6.5
- Probability Model (Ch7 - Ch9)
  - Basic concepts of probability, §7.1-7.2
  - Independence, conditional probability, §7.3, 8.1-8.3
  - Random variable, mean and variance §9.1-9.4
- Association between Two Random Variables (Ch10)
  - Portfolios, joint probability, §10.1-10.2
  - Dependence between random variables, §10.3
  - Mean and variance of weighted variables, covariance, §10.4, 10.6

## Main Topics

- Normal Probability Model (Ch12)
  - Normal random variable and normal Model, §12.1-12.2
  - Normal quantile plot, §12.3-12.4
- Sampling and Sampling Distributions (Ch13 - Ch14)
  - Concepts of population, sample and sampling, §13.1-13.2
  - Sampling distribution, §14.1
- Standard Errors and Confidence Intervals (Ch15)
  - Confidence interval for  $\mu$ , §15.1-15.4
  - Determine the sample size, §15.5
- Statistical Hypothesis Testing (Ch16)
  - Concepts, §16.1
  - One sample  $z$ - and  $t$ - tests, §16.2 - 16.3
- Simple Linear Regression Model (Ch19 - Ch21)
  - Linear pattern of data, residual plot, §19.1-19.3,19.5,
  - Nonlinear pattern of data and transformation, §20.1-20.4
  - The model and conditions, §21.1-21.2
  - Inference in regression, §21.3
  - Prediction interval, §21.4
- Regression Diagnostics (Ch22)
  - Detecting differences in variation, §22.1
  - Outlier and influential points, §22.2
  - Dependent errors and time series, §22.3
- Multiple Regression (Ch23)
  - The model and interpretation, §23.1-23.2, 23.5
  - Checking conditions, §23.3
  - Inference in multiple regression, §23.4