SPRING 2014 : STAT - 1 , Section 10.

HANDOUT - Descriptive Statistics - CHAPTERS 1, 2 & 3

- Read Chapter 1. Technical terms such as Descriptive Statistics, Population, Sample, Inferential Statistics, the history of the Development of Statistics, Observational study, and Designed Experiment are introduced along with several examples in Section 1.1. TRY 1.1 - 1.6, 1.7, 1.13, 1.19, 1.21.

  In Section 1.2 and 1.3, various sampling methods such as Simple Random Sampling, Systematic Random Sampling, Cluster Sampling, Stratified Random Sampling and Multistage Sampling are discussed. In 1.4 Experimental Designs along with technical terms are introduced. Go over the worked examples in these sections. In 1.4, learn the definitions of technical terms of experimental design by reading pages 22 and 23. Read the biography of the "Lady of the Lamp" (pp 31−32).

- In Chapter 2, we introduce data sets and organization of statistical data. Key words are variable, observation, data, population, sample, univariate data, bivariate data, qualitative variable, quantitative variable, categorical variable, discrete variable, continuous variable, frequency, relative frequency, frequency distribution, relative frequency distribution, pie and bar charts; these are discussed in section 2.1 and 2.2. Go over Examples 2.5, 2.6, 2.7 and 2.8. TRY 2.9, 2.13, 2.21, 2.25.

- In Section 2.3, to organize quantitative discrete data (i) single valued grouping (Example 2.12) and (ii) limit grouping (Example 2.13, Definition 2.7)) are discussed. (iii) Cutpoint grouping is useful for organizing continuous data (Example 2.14, Definition 2.8). Histograms for each of these cases and how to identify distribution shapes are discussed next and in section 2.4. Procedure 2.6 and Example 2.16 deal with the construction of dotplots while procedure 2.7 and Example 2.17 show how to construct Stem-and Leaf diagrams. These can be used to summarize data as well as to identify distribution types. TRY 2.47, 2.49, 2.51, 2.53, 2.57, 2.63, 2.75, 2.99, 2.101, 2.103.
  Read about Adolphe Quetelet.

- Chapter 3 describes Descriptive measures. Measures of Center such as the sample mean $\overline{x}$, the sample median $\tilde{\text{x}}$, and the sample mode x_M and their comparison are given in Section 3.1.
  TRY 3.21, 3.23, 3.31, 3.33.

- Measures of Variation such as the sample range, sample variance s^2, sample standard deviation s and the Three Standard Deviation Rule are discussed in section 3.2. ALWAYS USE THE COMPUTING FORMULA (Formula 3.1, p. 106) for calculating s (see Example 3.13, p.107). The Three-Standard-Deviations Rule is a particular case of the *Empirical Rule (p.114)* which is applicable to data sets with roughly bell-shaped distributions. The *Chebyshev's Rule (p.114)* is applicable to all data sets. TRY 3.71, 3.77, 3.79, 3.81, 3.96, 3.97, 3.99, 3.101, 3.103.

- The Five Number Summary, Quartiles Q_1, Q_2, Q_3, their computation, IQR, Lower and Upper limits of a data set, determining outliers, Constructing Boxplots and using them are dealt with in section 3.3,
  TRY 3.121, 3.123, 3.130.

- The Population mean μ, the Population variance σ^2, the Population standard deviation σ, Parameter, Statistic, $z-$score and its use are described in Section 3.4. TRY 3.149, 3.159.

- MAKE SURE YOU HAVE A CALCULATOR IN WHICH YOU CAN
  (I) INPUT DATA AND
  (II) READ OFF n , $\sum x_i$, $\sum x_i^2$, $\overline{x}$, s
  BY PRESSING APPROPRIATE KEYS.
  IF NEEDED, GO TO THE MATH LAB TO FIGURE THIS OUT.

- Pay particular attention to the biography of John Tukey, a pioneer of EDA.