USP 634 Data Analysis I   Spring, 2017
Assignment 2: Descriptive Statistics; Probability and Distribution
Due Tuesday, May 9th, 2017 on D2L dropbox

In this assignment you will ask and answer the questions below. The objectives of the
assignment are to:
- Practice computing confidence intervals and sample size calculation;
- Employ t-tests for hypothesis testing of means.

1. State law requires every county in the state to average a minimum expenditure of $5,000 per
student across all public schools in the county. A random survey of 200 schools in a county found an
average expenditure of $4,782, with a standard deviation of $600. Can you inform the state's
Secretary of Education at a 95 percent confidence level that the county is meeting the standard?
Show how you arrive at the conclusion. (15 pts)

2. You have been asked to manage a regional study on attitudes toward regional growth management.
You've been given a budget of $75,000 to conduct a regional mail survey of households. Your Board
wants to know the proportion of residents in favor of an urban growth boundary within an accuracy
of plus or minus two percent and at a confidence level of 90 percent. (Recall that when you don't
know the true population proportion, to be conservative you should assume a maximum standard
error). What is the minimum number of households you need to randomly sample? Can you conduct
the survey within budget given an estimated cost of $25 for administering each survey? (15 pts)

3. A random sample of adults living in 60 traditional, mixed-use, pedestrian-friendly
neighborhoods and adults living in 55 postwar, auto-oriented neighborhoods, all with
comparable household income levels, revealed the following:
- Traditional neighborhoods averaged 15.8 daily vehicle miles traveled (VMT) per
  adult household member, with a standard deviation of 5.3 VMT.
- Auto-oriented neighborhoods averaged 18.3 VMT per adult household member and a
  standard deviation of 7.5 VMT.
Create 90% confidence intervals for VMT for each type of neighborhoods. Do the intervals
overlap with each other? Using the hypothesis testing process, test at α = .10 level the
hypothesis of New Urbanists that people living in traditional neighborhoods have lower
automobile usage. (20 pts)

4. Housing values were compared between residences with a view of Mount Hood and otherwise
comparable residences (e.g., amenities, neighborhood quality) though without a view. The
analysis sought to measure the imputed value of a view on home sales prices. Two factors were
also controlled for in the analysis: age of home and distance to downtown Portland. Thus,
randomly selected homes that just sold were matched on these two factors, yielding the data
shown in the excel spreadsheet view.csv. Test at α = .05 level that homes with a view enjoy
significant value premiums, as reflected by differences in housing values. (20 pts)

5. Using your own data set and test a hypothesis with t-test. First select (or create by recoding) a
nominal variable V1 with two categories (for example, male and female, income above or below
poverty line, before or after a "treatment"), then choose a continuous V2 (or a discrete or even an
ordinal variable if there is no continuous variable in your dataset), and finish the following tasks

(30 points):
1.) Calculate confidence intervals for V2, first for all observations and then for each of the two groups (categories) in variable V1;
2.) Formulate a hypothesis for the relationship between the two groups (categories), conduct a t-test for your hypothesis, and interpret your hypothesis testing results;
3.) Describe and discuss your findings, as appropriate.