Semester Project – Spring 2020

Caution   Grading for the Lab Project is subjective. You may not get the grade you expect.

Purpose   To provide students an opportunity to showcase their knowledge and skills in the theory and practice of statistics in a *formal report*.

Due Date   The date due will be shown on Moodle. Late projects will not be accepted. You must submit your project in a MS Word file or a PDF file by email to me (mmckenn@lsu.edu) by the due date.

Super Important

1. The *Subject line* on the email MUST be:

   EXST 2201 Semester Project

2. The *File name* of your pdf file or MS Word file MUST be:

   your-section-number_your-LSU-username_Project

   (examples: 001_mmckenn_Project ; 044_wcoxe1_Project)

   I need the section number (all three digits first, as in this example) to be able to sort the files to be able to grade the large number of projects.

3. *Do Not* **paste** any information, from any source, into your report, as that is academic dishonesty.

Preparation

Individual Effort: You must do all the work by yourself. You may talk (only!) to others, including TA's and Instructors, but every keystroke of doing the statistical analysis and writing the report must be done by you. You may ask for help.

Situation: The situation (The Titanic) and the dataset is shown below.

Questions: Form three questions of the data that can be answered using the statistical methods discussed in Unit 03 (two-sample *t*-test, ANOVA, regression) of lecture and as well as discussed in Topic 02 of the lab. Make sure that your questions are appropriate for this dataset (refer to the labs if you need help).

Analysis: You may not have to use all three variables in your analysis, but you do need to use observations from all of the 48 individuals. You *must* use *SAS* Enterprise Guide to do the calculations.

Formal Report

Notes: 1. Three (3) pages *only*! *No* cover page. One page per method *in order* (two-sample *t*-test, ANOVA, regression).
2. 12 point Calibri font, one-and-one-half line spacing (tables can be single-spaced).
3. Verboseness is *not* an asset. Be succinct and to the point. Your grade will depend more on what you say, not on how long you take to say it.

Page 1

Two-Sample *t*-Test   **FName LName (LSU username)**   **Date**

Descriptive Statistics:

Give one *table* of the descriptive statistics describing your columns of data.

Theory:

In one *paragraph*, briefly discuss the theory of this method.

Results:

Give one *table* showing the main inferential results of this method.

Conclusion:

In one *paragraph*, give the *p*-value and state what this means to the researcher.

Page 2

ANOVA   **FName LName (LSU username)**   **Date**

Descriptive Statistics:

Give one *table* of the descriptive statistics describing your columns of data.

Theory:

In one *paragraph*, briefly discuss the theory of this method.

Results:

Give one *table* showing the main inferential results of this method.

Conclusion:

In one *paragraph*, give the *p*-value and state what this means to the researcher.

Page 3

Regression   **FName LName (LSU username)**   **Date**

Descriptive Statistics:

Give one table of the descriptive statistics describing your columns of data.

Theory:

In one paragraph, briefly discuss the theory of this method.

Results:

Give one table showing the main inferential results of this method.

Conclusion:

In one paragraph, give the p-value and state what this means to the researcher.

Description

The RMS Titanic was an Olympic-class passenger ship that at the time of her construction in 1912 was the largest ship in the world. She was also the most technologically advanced ship that was described as being 'unsinkable'. Four days into her maiden voyage, the Titanic hit an iceberg and sank two hours and forty minutes later. Of the 2,223 people on board at the time, 1,517 died.

Data

The data below is a random sample from each passenger class on the Titanic. The original dataset included information from only half of the passengers. Passenger Class and whether the passenger Survived, or not, is given. The variable Fare gives the cost in British pounds paid by each passenger. The variable Age gives the age in years of each passenger.

First Class				Second Class				Third Class			
Survived		Died		Survived		Died		Survived		Died	
Fare	Age	Fare	Age	Fare	Age	Fare	Age	Fare	Age	Fare	Age
51	53	61	46	30	25	12	18	22	4	20	13
91	19	71	39	30	45	26	25	8	22	7	25
27	28	212	33	37	8	26	37	10	63	8	38
31	40	26	30	33	6	12	21			18	18
152	33	94	23	37	1	13	23			31	13
82	53	27	33	20	13	14	24			8	19
89	49			13	33	74	24			14	34
77	48					11	16			34	21
109	39					14	59			8	30
31	30									8	35
										16	26
										8	42
										14	1

Example Page

An example page of this report follows on the next page.

Descriptive Statistics:

N	Mean	Median	Std	Max	Q_3	Q_1	Min	SW-*p*
48	28.7	27	14.7	63	38.5	19	1	0.81

Theory:

A one-sample *t*-test is a hypothesis test that compares probability information assuming the null hypothesis is true, in the form of a critical value, to sample information, in the form of a test statistic, to get population information about the mean. The null hypothesis is that the population mean equals 30, and the alternative hypothesis is that the population mean is less than 30 to test if the mean age of those who traveled on the Titanic were young. The alpha used was 0.05.

Results:

df	*t*-Value	*p*-Value
47	-0.61	0.5459

Conclusion:

A *p*-value of 0.54 is bigger than an alpha value of 0.05 so the null hypothesis is not rejected. This means that, based on this sample, it is reasonable that the mean age of all of the travelers on the Titanic was 30 years old. This result might indicate that there were not a lot of children on board, which makes sense as this was the Titanic's first cruise and parents did not want to risk their children. Or, maybe it was that people did not think that this trip would be an adventure and so it did not attract younger adults.