General Education Common Graded Assignment: ***MATH 153 Spring 2017***
Inferences from Data on the "Working Poor" in the U.S.

MATH 153: Introductory Statistics is a general education course designed to assist students in the development of critical life skills. One of the goals of this assignment is to assess student competence for each of these objectives:

I. **Written and Oral Communication**— interpret statistical graphs and effectively communicate the results of a statistical analysis.
II. **Critical Analysis and Reasoning**— analyze bivariate data and perform hypothesis test to draw inference regarding parameter.
III. **Technological Competence—** apply technology to manage data, explore data with graphs, perform inference, and check condition**s**
IV. **Scientific and Quantitative or Logical Reasoning—** analyze bivariate data using scatter plots and linear regression
V. **Local and Global Diversity**— apply statistical methods to data from diverse cultural and global populations.
VI. **Personal and Professional Ethics**— construct a solution to real world problems using problem methods individually and in teams.

In addition to the above general education objectives, this assignment assesses students' understanding and application of the following skills and knowledge specific to the topics of linear regression, sampling methods, confidence intervals, hypothesis testing, and critical and quantitative reasoning covered in ***MATH 153- Introductory Statistics:***

- ***Demonstrate statistical reasoning in everyday life using real-world data***
- ***Apply technology to manage data, explore data, perform inference, and check conditions***
- ***Describe data with appropriate measures of central tendency and variability***
- ***Generate and interpret statistical graphs***
- ***Analyze bivariate data using linear regression***
- ***Apply statistical methods to data from diverse cultural and global populations***
- ***Construct and interpret confidence intervals in order to make inferences about parameters Perform hypothesis testing to draw inferences regarding parameters***
- ***Effectively communicate the results of a statistical analysis***

ASSIGNMENT: There is a population of workers in this country who are employed full- or part-time who do not earn enough in wages to securely cover the cost of their family's basic needs. These individuals exist in each of the U.S. states and the District of Columbia. Although the numbers vary by state, these individuals and their families are similar in that they cross racial, cultural, and local boundaries. They are considered the nation's "Working Poor." This common graded assignment was designed to allow you, the Introductory Statistics student, to explore some of the data related to this population. Through the use of multiple statistical concepts covered in your MATH 153 course you will complete the provided assignment below to demonstrate your understanding of the manipulation of data, the use of technology in statistical analysis, the relationships between related variables, estimating population parameters, and hypothesis testing.

Purpose: To assess student understanding of selected key statistical concepts through the use of real-world data related to the working poor in the U.S.
Audience: Students in all MATH 153 courses in the spring 2017 semester, including all 7-week and online sections.
Directions: Please complete this assignment by responding specifically to each of the items 1-18, provided. According to the instructions you will be expected to use technology to make calculations, to create and include graphical displays, and to support your responses with complete, well-organized, and clear statements and interpretations.

ASSIGNMENT SPECIFICATIONS:

- All assignments must be completed using word processing software like Word.
- All responses must be typed in Times New Roman 12-pt font.
- Page margins should be set to 1" all around.
- Minimum of 5 pages typed, and double-spaced.
- Include requested graphical displays using cut and paste from statistical software.
- You will use the included data file on the Working Poor to complete the project. It is provided in hard-copy and, upon request, will be made available in electronic format for your convenience.

GRADING:

- This assignment will account for ***10%*** of the total course grade.
- See attached rubric for details about how your assignment will be graded.

SUBMISSION GUIDELINES:

- **DUE DATE: For full semester the due date is Wednesday, May 10. For the 1st 7-week sections, the due date is Friday, March 17. For the 2nd 7-week sections, the due date is Friday, May 12.**
- Submit your completed assignment in hard copy, or electronic format by the selected due date.
- Submit TWO copies. One copy should include your student ID, course name, and section; it should omit student and faculty names.

MATH 153 – Spring 2017 **Student ID_____**
Section_____ **Date_____**

GREAT PROJECT

Correlation and Linear Regression

It is widely believed that the more education one receives the higher the income earned at the time of first employment and over the course of a career. However, due to varying reasons, many people never complete high school and, thus, never receive their high-school diploma. Although individuals without a high-school diploma are often able to find employment, they experience economic outcomes quite different from those who finish high school before entering the workforce to earn a living. Across the nation, there are millions of individuals with families who are now working but do not possess the credentials of a high-school diploma. Many of these individuals and their families are considered to be a part of the working poor that make up a considerable portion of this nation's labor force.

1. Use technology to create a scatterplot of the percent of low-income working families and the percent of 18-64 yr-olds with no high school diploma information (The Working Poor Families Project, 2011) in each jurisdiction. Print or copy and paste the scatterplot and be sure to clearly identify the predictor and response variables based on the believed association between education and income.

2. Write one or two sentences to describe the association between the predictor and response variables in your scatterplot. Be sure to use the actual names of the variables in their appropriate places in your response.

3. Use technology to find the regression equation for the linear association between the percent of low-income working families and the percent of 18-64 yr-olds with no high school diploma. Provide this equation and write a brief interpretation of its slope (Be sure to use the same predictor and response variable assignments that you chose for the scatterplot).

4. Identify the R-squared value for this regression equation. Provide this value and write a sentence to interpret its meaning.

5. A student states that a decrease in the percent of 18-64 yr-olds with no high school diploma will no doubt lead to a decrease in the percent of low-income working families. Based on what you know about correlation, write a brief response to this statement.

6. The student in number 5 also believes that the majority of those with no high-school diploma are similar in terms of their nationality, native language, and disability status. The student, therefore, believes that it is the responsibility of these sub-groups of the population and their advocates to address the working poor issue themselves. Use the key rules about samples and extrapolation when making inferences from correlations to write a socially responsible and statistically appropriate response regarding the validity of the student's beliefs.

Confidence Intervals

During the recent recovery from the Great Recession of 2007-2009, the economic situation for many families has improved. However, in 2011 the recovery was slow and it was uncertain as to how much had really changed on the national level. To estimate the population parameter of the percent of low-income working families, a representative simple random sample could be used to calculate a point estimate and a confidence interval.

7. Write an appropriate sampling procedure that will enable a researcher to collect a simple random sample of size n=20 from the full list of jurisdictions.

8. Carry out a sampling of the jurisdictions by using a random number generator, random number table, or the sampling capabilities of technology to select a simple random sample of size n=20 from the full list of the jurisdictions. Provide the selected sample of 20 in the table included below, making sure to include the corresponding percent of low-income working families (%LIWF) information for each jurisdiction selected.

Jurisdiction	%LIWF	Jurisdiction	%LIWF	Jurisdiction	%LIWF	Jurisdiction	%LIWF

9. Use technology to calculate the sample mean and sample standard deviation of the percent of low-income working families for the selected 20 jurisdictions in your sample. Provide your mean and standard deviation and express in detail why it is unlikely that any two samples would produce the same results.

10. A different sample of size n=20 produced a sample mean of 31.08% and a sample standard deviation of 5.48%. Use these values to calculate a 90% confidence interval (without the use of statistical technology) for the national mean percent of low-income working families. Please show all calculations and provide the upper and lower limits that make up the confidence interval. (Round the limits to two decimal places.)

11. Express an appropriate statistical interpretation of the 90% confidence interval included in the answer to number 10.

12. In any of the 20-value samples selected, some jurisdictions may have a percent of low-income working families that is not included in the 90% confidence interval calculated above. If federal funds are available only for jurisdictions whose percent of low-income working families falls within the reported confidence interval, explain whether an interval with a *higher* or *lower* confidence level would be more advantageous to the jurisdictions.

13. If a lawmaker reports only research results that ensure that the jurisdiction he or she represents gets federal aid, explain whether such an action would constitute a misuse of statistics. If the

action is not a misuse, state why it is appropriate. If the action is a misuse, state an alternate reporting practice that would be more appropriate.

Hypothesis Testing

In 2011, the national percent of low-income working families had an approximately normal distribution with a mean of 31.3% (The Working Poor Families Project, 2011). Although it has remained slow, some politicians now claim that the recovery from the Great Recession has been steady and noticeable. As a result, it is believed that the national percent of low-income working families is significantly lower now in 2014 than it was in 2011. To support this belief, a recent spring 2014 sample of n=16 jurisdictions produced a sample mean of 29.8% for the percent of low income working families, with a sample standard deviation of 4.1%. Using a 0.10 significance level, test the claim that the national average percent of low-income working families has improved since 2011.

14. Write a few brief sentences to state the type of test that should be performed and state the assumptions and conditions that justify its appropriateness.

15. Clearly identify and state the null and alternate hypothesis for this test.

16. Use technology to identify the test statistic and the P-value associated with the hypothesis test. Provide these values.

17. State the decision of the hypothesis test based on a 0.10 significance level.

18. Provide the appropriate conclusion about the claim that the national average percent of low income working families has improved since 2011.

Reference(s)

The Working Poor Families Project. (2011). Indicators and Data. Retrieved from http://www.workingpoorfamilies.org/indicators/

2011 Data		
Jurisdiction	Percent of low income working families (<200% poverty level)	Percent of 18-64 year olds with no HS diploma
Alabama	37.3	15.3
Alaska	25.9	8.6
Arizona	38.9	14.8
Arkansas	41.8	14
California	34.3	17.6
Colorado	27.6	10.1
Connecticut	21.1	9.5
Delaware	27.8	11.9
District of Columbia	23.2	10.8
Florida	37.3	13.1
Georgia	36.6	14.9
Hawaii	25.8	7.2
Idaho	38.6	10.7
Illinois	30.4	11.5
Indiana	31.9	12.2
Iowa	28.8	8.1
Kansas	32	9.7
Kentucky	34.1	13.6
Louisiana	36.3	16.1
Maine	30.4	7.1
Maryland	19.5	9.7
Massachusetts	20.1	9.1
Michigan	31.6	10
Minnesota	24.2	7.3
Mississippi	43.6	17
Missouri	32.7	11.1
Montana	36	7
Nebraska	31.1	8.7
Nevada	37.4	16.6
New Hampshire	19.7	7.3
New Jersey	21.2	10.1
New Mexico	43	16.2
New York	30.2	13
North Carolina	36.2	13.6
North Dakota	27.2	5.9
Ohio	31.8	10.3
Oklahoma	37.4	13.2
Oregon	33.9	10.8
Pennsylvania	26	9.4
Rhode Island	26.9	12
South Carolina	38.3	14.2
South Dakota	31	8.7
Tennessee	36.6	12.7
Texas	38.3	17.8
Utah	32.3	9.9
Vermont	26.2	6.6
Virginia	23.3	10.2
Washington	26.4	10.2

West Virginia	36.1	12.9
Wisconsin	28.7	8.5
Wyoming	28.1	8