

SAS Project 1 -College bound

Statistics 521



xxx

xxx@apu.edu

November 6, 2020

**Table of Contents**

Part 1: Introduction………………………………………………………………………………………………..3

Part 2: Descriptive Statistics of Numerical Variables…………………………………………………………..3-4

* 1. Descriptive Statistics of the Base Year Composite Test Score………………………………………..3
  2. Descriptive Statistics of the State Hourly Wage in Manufacturing…………………………………...3
  3. Descriptive Statistics of the Distance from Four-year College in 10’s of Miles……………………...3
  4. Descriptive Statistics of the Average State Four-year College Tuition in $1000’s…………………...4
  5. Descriptive Statistics of the Years of Education Completed………………………………………….4

Part 3: Descriptive Statistics of Categorical Variables………………………………………………………….4-6

* 1. Descriptive Statistics by Gender………………………………………………………………………4
  2. Descriptive Statistics by Homeownership………………………………………………………….…5
  3. Descriptive Statistics by Urban Area………………………………………………………………….5
  4. Descriptive Statistics by Family Income……………………………………………………………...5
  5. Descriptive Statistics by Homeownership and Income……………………………………………….6
  6. Descriptive Statistics by Gender and Urban Area…………………………………………………….6

Part 4: Histogram and Density plots of Numerical Variable…………………………………………………..7-11

4.1 Histogram of Test Score……………………………………………………………………………...7

4.2 Density Plot of Test Score………………………………………………………………………….…7

4.3 Histogram of the State Hourly Wage in Manufacturing………………………………………………7

4.4 Density Plot of the State Hourly Wage in Manufacturing…………………………………….………8

4.5 Histogram of the Distance from 4yr College in 10’s of miles………………………………………...8

4.6 Density Plot of the Distance from 4yr College in 10’s of miles………………………………………9

4.7 Histogram of the Average State 4yr College Tuition in $1000’s………………………………..……9

4.8 Density Plot of the Average State 4yr College Tuition in $1000’s………………………….………10

4.9 Histogram of the Years of Education Completed……………………………………………………10

4.10 Density Plot of the Years of Education Completed………………………………………...………11

Part 5: Partial Outputs……………………………………………………………………………………………11

Part 6: Works Cited………………………………………………………………………………………………12

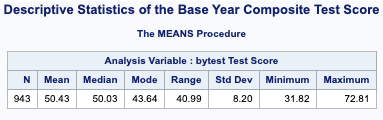
**Part 1|Introduction**

This data set was derived from a survey response of high school students from 1100 high schools. The data set has 943 observations and 9 variables. The variables are comprised of 4 categorical and 5 numerical. The categorical variables are defined as female, ownhome, urban and incomehi. The numerical variables are defined as bytest, stwmfg, distribution, tuition and ed.

This project is a simple compilation of the Descriptive Statistics of both the categorical and numerical variables and the various interpretations that can be made and gathered.

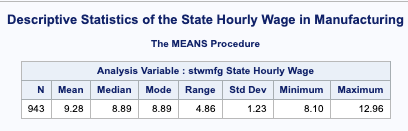
**Part 2 | Descriptive Statistics of Numerical Variables**

* 1. Descriptive Statistics of the Base Year Composite Test Score



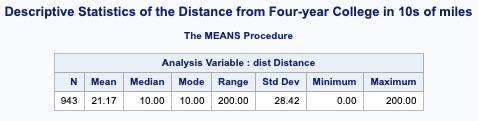
The Descriptive Statistics of the Base Year Composite Test Score is comprised of 943 observations. The test scores range from the lowest value of 31.82 and 72.81 with an average of 50.43.

* 1. Descriptive Statistics of the State Hourly Wage in Manufacturing



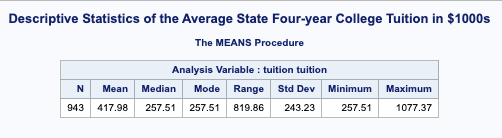
The Descriptive Statistics of the State Hourly Wage in Manufacturing is comprised of 943 observations. The State Hourly Wage ranges from the lowest value of 8.10 to 12.96 with an average of 9.28.

* 1. Descriptive Statistics of the Distance from Four-year College in 10’s of Miles



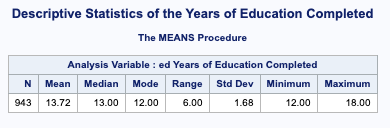
The Descriptive Statistics of the Distance from Four-year College on 10’s of Miles is comprised of 943 observations. The Distances from the Four-year colleges range from the smallest distance of 0.00 and the largest distance of 200.00 miles with an average distance of 21.17 miles.

* 1. Descriptive Statistics of the Average State Four-year College Tuition in $1000’s.



The Descriptive Statistics of the Average State Four-year College Tuition in $1000’s is comprised of 943 observations. The Average State 4-year college tuition cost ranges from the lowest value of 257.51 and 1077.37 with an average of 417.98.

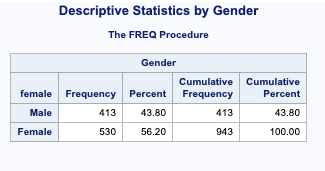
2.5 Descriptive Statistics of the Years of Education Completed



The Descriptive Statistics of the Average State Four-year College Tuition in $1000’s is comprised of 943 observations. The Average State 4-year college tuition cost ranges from the lowest value of 257.51 and 1077.37 with an average of 417.98.

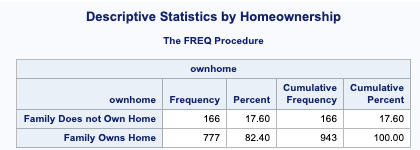
**Part 3 | Descriptive Statistics of Categorical Variables**

3.1 Descriptive Statistics by Gender



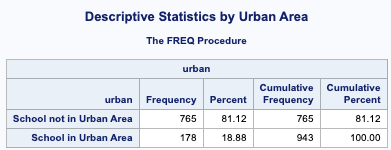
The Descriptive Statistics by Gender show the male category containing 413 participants and the female category containing 530 participants. The Male gender makes up 43 % of the population while the female gender makes up 56.20% of the population. It is evident that there are more women than men.

3.2 Descriptive Statistics by Homeownership



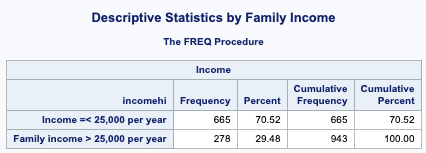
The Descriptive Statistics by Homeownership shows that student many fall in the category of belonging to a family that owns a home or a family that does not own a home. Students whose families own a home make up 82.40% while the students who’s families do not own a home is 17.60%

3.3 Descriptive Statistics by Urban Area



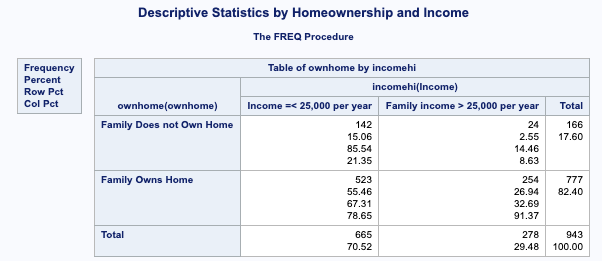
The Descriptive Statistics by Urban Area can be characterized as belonging to a school that is located in an Urban Area or belonging to a school that is not in an Urban Area. There are 765 or 81.12% of students that attend a school that is not located in an Urban Area while 178 or 18.88% of students attend a school in an Urban area.

3.4 Descriptive Statistics by Family Income



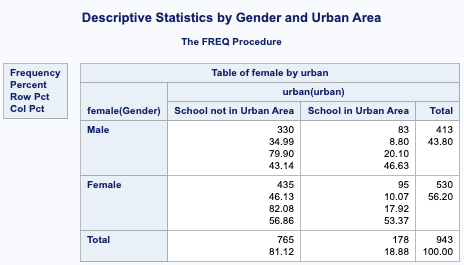
The Descriptive Statistics by family income shows students that identify having a household income that was less than or equal to $25,000 as well as students in households that made more than $25,000. The output shows that 665 or 70.52% students identify as having household incomes that are less than or equal to $25,000 while 278 or 29.48 % of students identify as having household incomes that are greater than $25,000.

3.5 Descriptive Statistics by Homeownership and Income



Descriptive Statistics by Homeownership and Income shows the relationship between students whose families own homes and have an income of less than and equal to $25,000 is comprised as 15.06%.

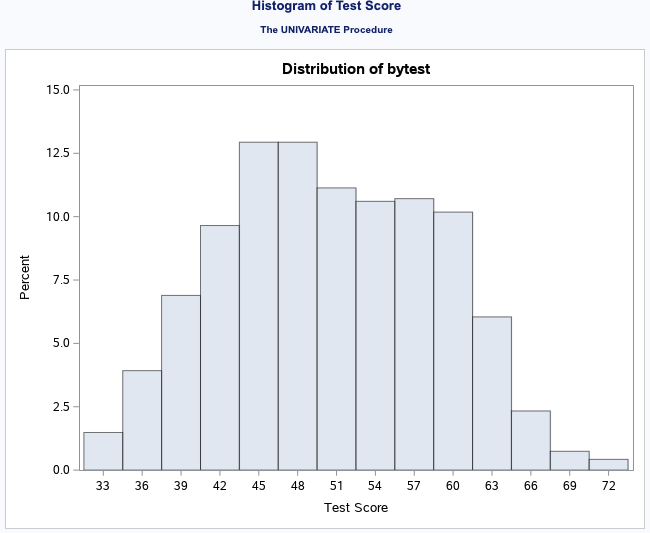
3.6 Descriptive Statistics by Gender and Urban Area



The Descriptive Statistics by Gender and Urban Area demonstrates that 46.13% of those classified as female attended schools that are not in Urban Areas.

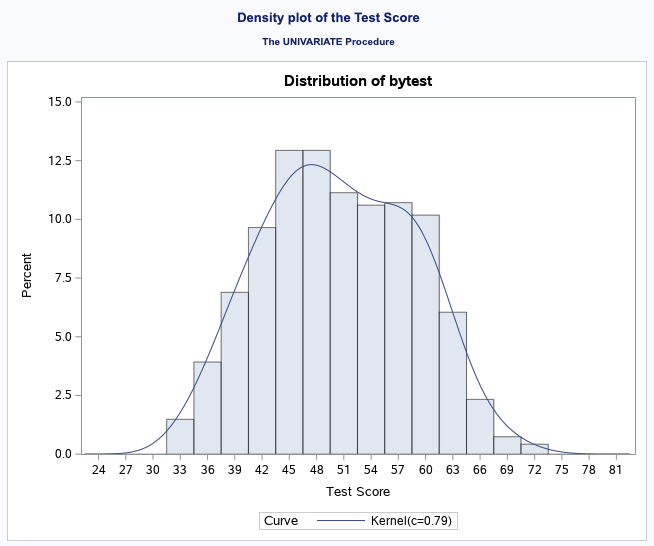
**Part 4 | Histogram and Density plots of Numerical Variable**

4.1 Histogram of Test Score



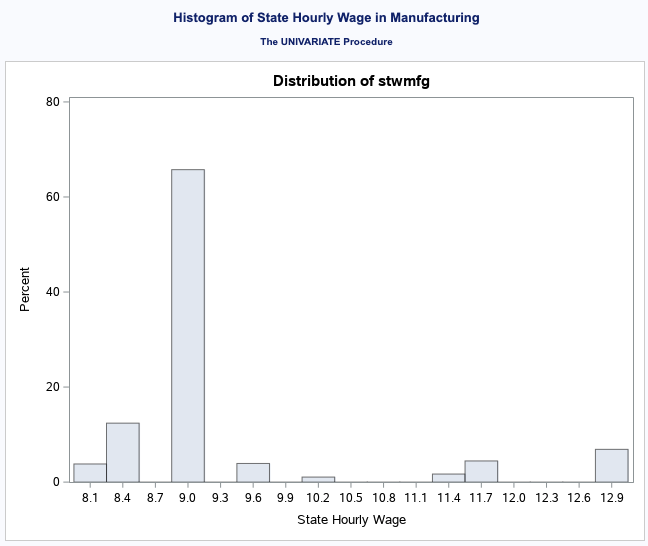
The Test Score histogram shows a slight skewness of the data with 12.5% of the test scores falling between 44 – 50.

4.2 Density Plot of Test Score



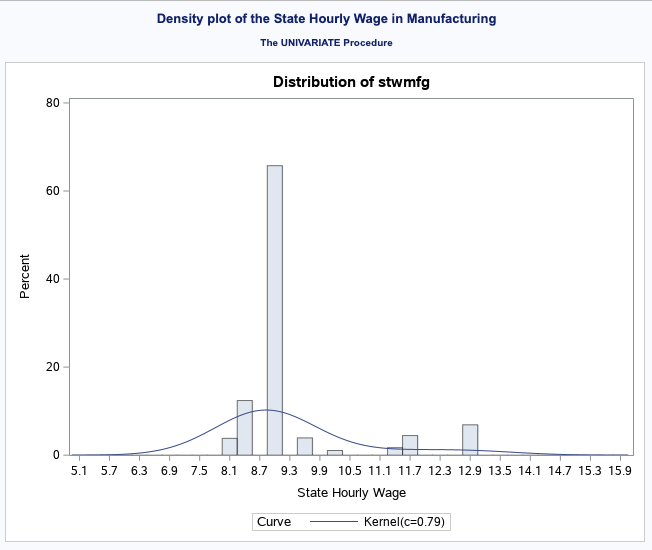
The density plot of the Test scores shows a slight skewness of the data with 12.5% of the test scores falling between 44 – 50.

4.3 Histogram of the State Hourly Wage in Manufacturing



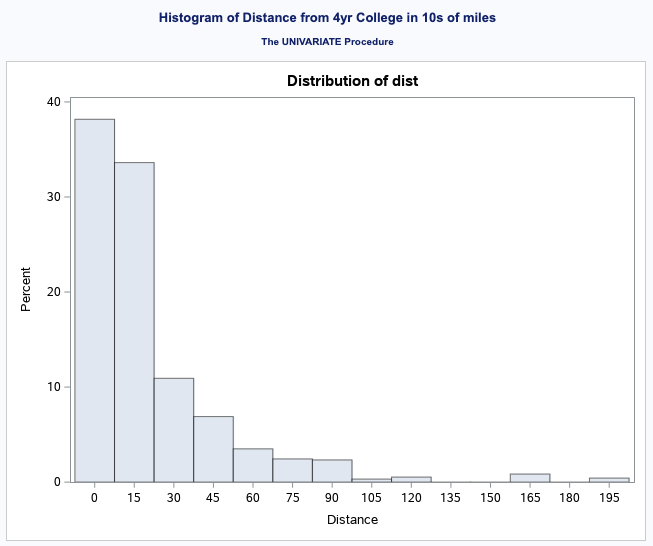
The Histogram of the State Hourly Wage in Manufacturing shows that about 65% of workers earn a State Hourly Wage of 9.0.

4.4 Density Plot of the State Hourly Wage in Manufacturing



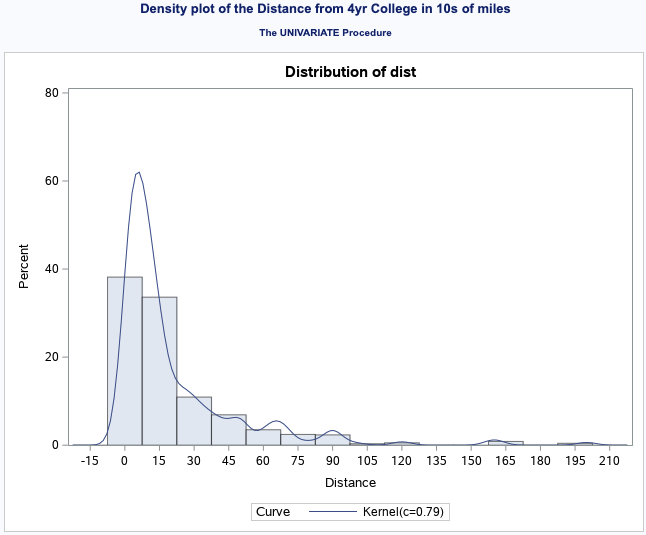
The density plot shows that over 65% of workers earn a state hourly wage of 9.0 while the kernel shows that we have a normal distribution.

4.5 Histogram of the Distance from 4yr College in 10’s of miles



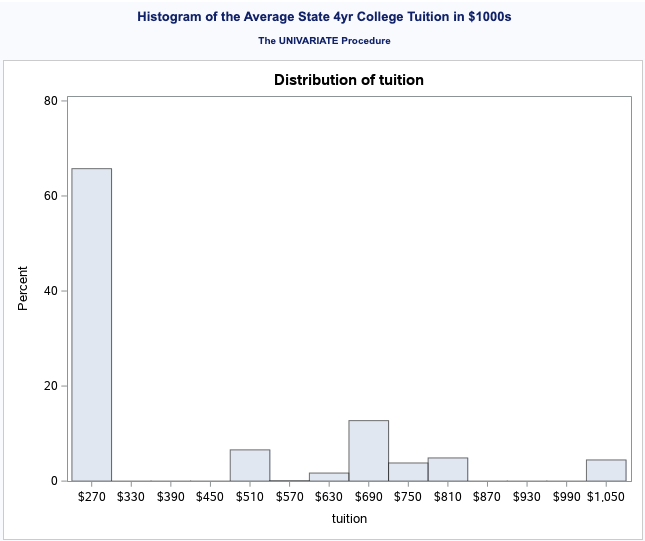
The Histogram of Distance from 4 year college in 10’s of miles is skewed right with the most prominent distance traveled being between 0 to 10 mile. The largest distances traveled fall at 165, 120 and 195 miles and it is apparent that this choice is not popular.

4.6 Density Plot of the Distance from 4yr College in 10’s of miles



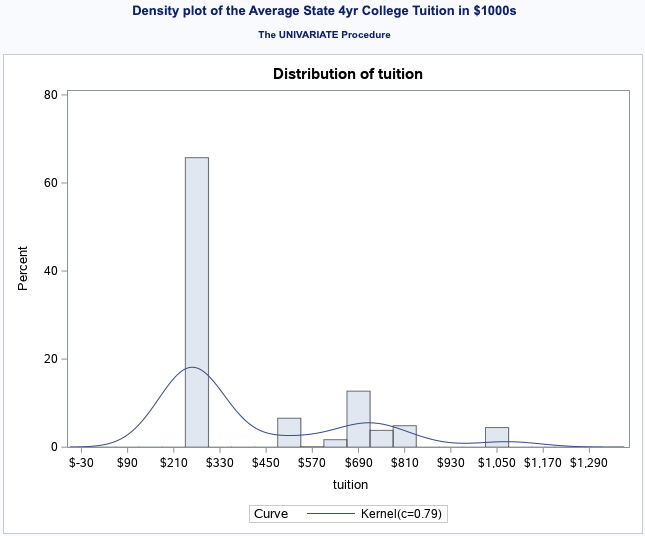
The Density plot from 4 year college in 10’s of miles is skewed right with the most prominent distance traveled being between 0 to 10 mile. The largest distances traveled fall at 165, 120 and 195 miles and it is apparent that this choice is not popular. The Kernel also support the claim that the data is skewed right.

4.7 Histogram of the Average State 4yr College Tuition in $1000’s



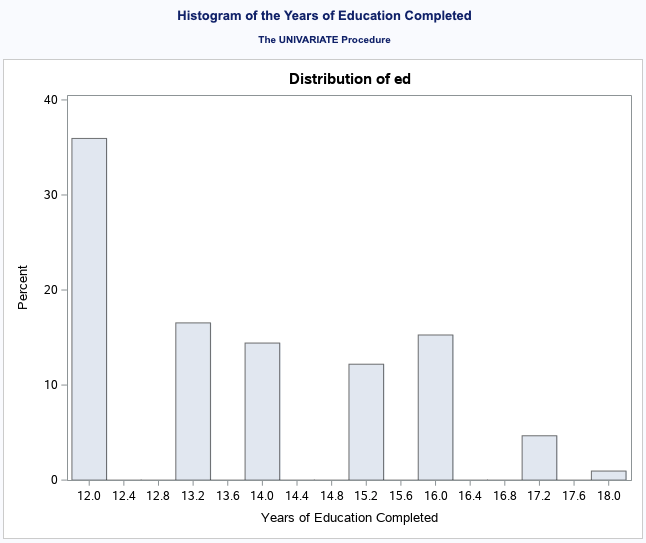
The Histogram of the Average State 4yr College Tuition in 1000’s shows a multimodal distribution with a large percentage, about 65% of tuition cost ranging from 0 to $270. We also see an apparent outlier at $1050 accounting for less than 10%.

4.8 Density Plot of the Average State 4yr College Tuition in $1000’s



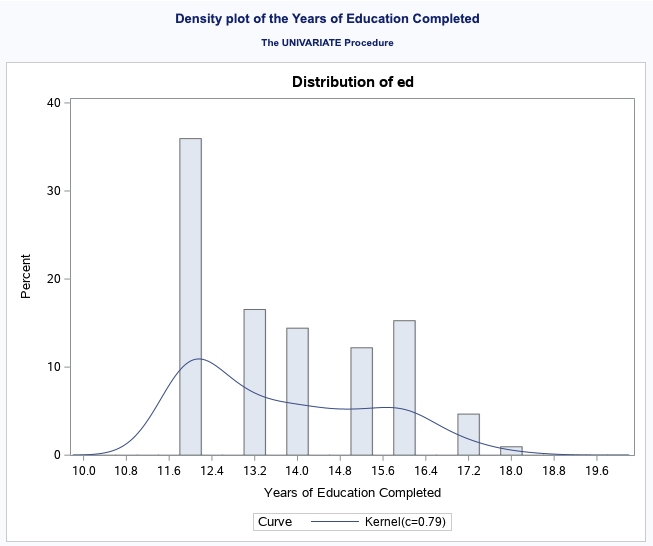
The Density Plot of the Average State 4yr College Tuition in 1000’s shows a multimodal distribution with a large percentage, about 65% of tuition cost ranging between 230 to $270. A second cluster of data shows tuition costs that range from $470 - $820. We also see an apparent outlier at $1050 accounting for less than 10%.

4.9 Histogram of the Years of Education Completed



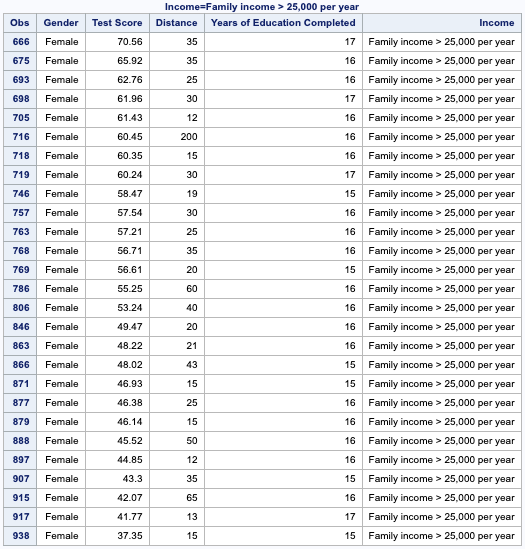
The histogram of the years of education completed shows a slight bimodality within the distribution. The data shows that less than ten percent of the population has 18 years or education while more than 35% and less than 40% of students have completed no more than 12 years of education.

4.10 Density Plot of the Years of Education Completed



The density plot of the years of education completed shows a slight bimodality within the distribution. The data shows that less than ten percent of the population has 18 years or education while more than 35% and less than 40% of students have completed no more than 12 years of education.

**Part 5: Partial Outputs**



From the 943 observations, 27 women identify as having income higher than or equal to $25,000. These women have completed between 15-17 years of education yet have test scores that are no higher than 70.56 yet no lower than 37.35.

**Works Cited**

Rouse, C. E. (1995). Democratization or Diversion? The Effect of Community Colleges on Educational Attainment. *Journal of Business & Economic Statistics*, *13*(2), 217. https://doi.org/10.2307/1392376