

Pittsburgh Associated Presses



Department of Educational Foundations and Leadership
GREV 510 -- Statistics I
Fall 2010
Dr. Gibbs Y. Kanyongo

Submitted by:
Adam Wasilko
Douglas Strahler
Michaela Noakes

December 10, 2010

Table of Contents

Executive Summary	2
<i>Confidence Level Results</i>	3
Individualized Outcomes	6
Q 1: How much confidence do you have in Major Companies?	4
Q 2: How much confidence do you have in Organized Religion?	5
Q 3: How much confidence do you have in Education?	6
Q 4: How much confidence do you have in Executive Branch of the Federal Government?	7
Q 5: How much confidence do you have in Organized Labor?	8
Q 6: How much confidence do you have in the Press?	9
Q 7: How much confidence do you have in Medicine?	10
Q 8: How much confidence do you have in Television?	11
Q 9: How much confidence do you have in the United States Supreme Court?	12
Q 10: How much confidence do you have in Scientific Community?	13
<i>Continuous Variables: AGE, WORDSUM, and TVHOURS</i>	14
<i>AGE</i>	15
<i>WORDSUM</i>	16
<i>TVHOURS</i>	17
<i>TVHOURS</i> : (8) Extreme Values	17-18
Respondents by Sex	19
<i>AGE</i> using Box Plot Analysis.....	19-21
<i>WORDSUM</i> using Box Plot Analysis.....	21-22
<i>TVHOURS</i> using Box Plot Analysis.....	23-24
<i>Part II: Response to Specifics raised by PAP</i>	25
<i>Any differences by gender (SEX) with respect to the variables AGE, WORDSUM, and TVHOURS</i>	25
Group Statistics and Independent Samples Test.....	26
Independent Samples Test.....	27
<i>May we say the average age of the population is 46?</i>	28

Stem-and-Leaf Plot.....	29
<i>Are the hours spent watching TV related to a subject's age?</i>	30
Correlations.....	31
<i>Is smoking behavior related to gender?</i>	31
Correlations.....	32
<i>Part II: Additional points of interest</i>	
<i>Is there a correlation between interest in protecting the environment and the nation's health?</i>	33
Correlations.....	34
<i>Is there a correlation between interest in improving the education system and the nation's health?</i>	34
Correlations.....	34
<i>Is there a correlation between interest in improving the education system and the environment?</i>	35
Correlations.....	36
Appendix.....	37

Executive Summary

Statisticians, Wasilko, Strahler and Noakes were contracted by Pittsburgh Associated Presses (PAP) to analyze and evaluate the data results of a recent poll on areas of interest to the paper's readership. The survey poll was presented to a random sample of 418 participants. The findings presented in this report are derived by Descriptive Statistical analysis and interpretation inclusive of Hypothesis (assumption) Testing. The report will be presented in two distinct sections: describing what the data shows in a summary format and inferring general conclusions from the sample data as to the considerations of the population.

Confidence Level Results

The ordinal variable, confidence, was tested to determine the level of assurance the readers indicated in response to a Likert scale ranking of mutually exclusive categories. Table 1 provides a listing of topical areas that were selected for analysis with the cumulative number of responses indicated.

Table 1

Confidence Variables

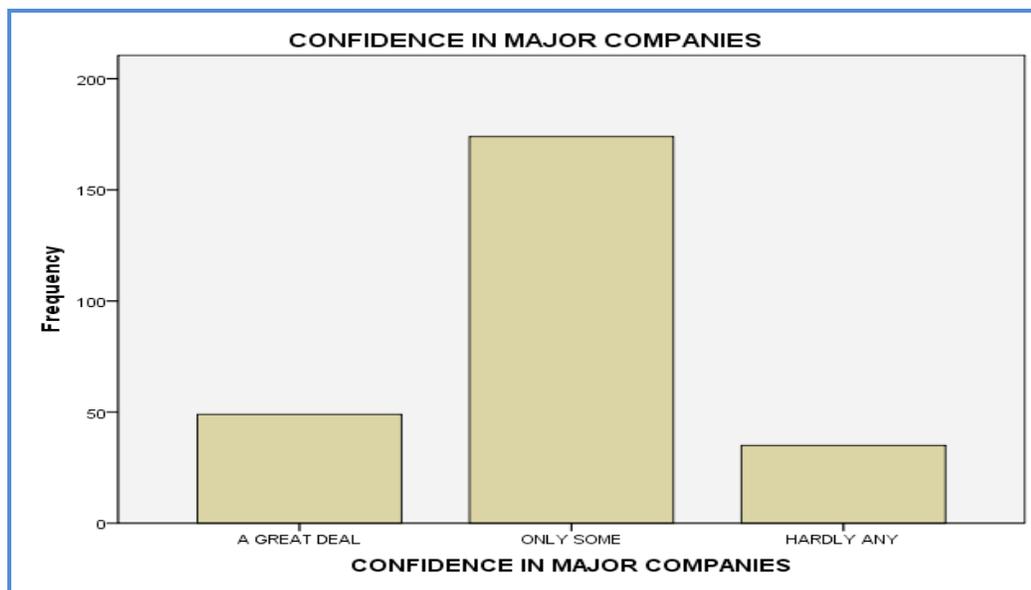
Confidence Statistics										
	Major Companies	Organized Religion	Education	Executive Branch of Government	Organized Labor	Press	Medicine	Television	United States Supreme Court	Scientific Community
N										
Valid	258	262	265	266	251	263	263	268	259	248
Missing	160	156	153	152	167	155	155	150	159	170

Each category variable with its subsequent analysis will be presented in this report based on the questions that were presented in the survey tool. The survey sample population consisted of 418 participants. Each participant was requested to rank their response based upon mutually exclusive categories: A GREAT DEAL (1), ONLY SOME (2), HARDLY ANY (3), DON'T KNOW (8) and NOT APPLICABLE (9). The frequency of each valid response was tallied. Any missing data will be addressed in conjunction with any outliers (extreme scores in the distribution) that are identified during data analysis. Supporting documentation for these findings will be presented as graphical output in the form of a bar graph representing the discrete data. The height of the bars will indicate the frequency of the components in each set by comparing the relative frequencies in each category. The discrete categories are represented by separations between the bars.

Q 1: How much confidence do you have in Major Companies?

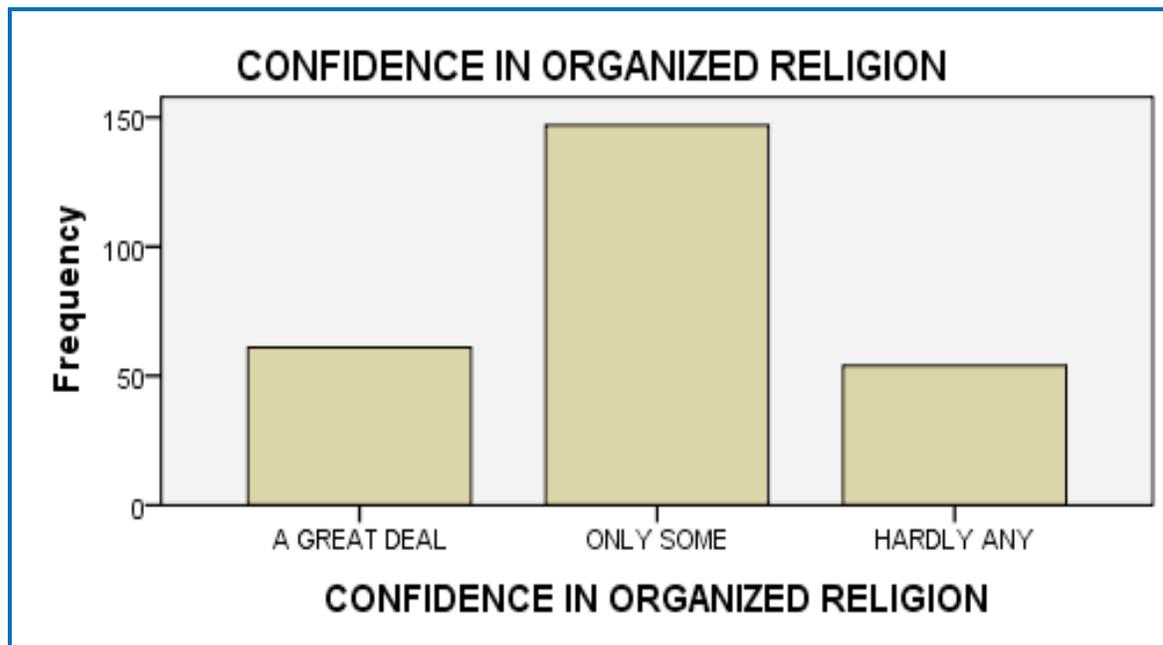
Of the sample population of 418 people who were presented with the survey tool, 258 provided a valid response to this question, 15 answered NA (NOT APPLICABLE), and 145 were evaluated as a DK (DON'T KNOW) response. Missing data to this question is attributed to the combined scores of 15 and 145 which equates to a total of 160 participants or 38.3% who responded in this manner. The most frequently recorded response to this question was ONLY SOME selected by 174 participants for a 41.6% response rate. A GREAT DEAL was selected by 49 participants for an 11.7% response rate, and HARDLY ANY was selected by 35 participants for an 8.4% response rate. The relative frequency of the 258 respondents out of the 418 participants indicates a proportional response of 61.7%.

Cumulative Percentage represents the scores at or below each level. For example, 86.4% of the reported scores were at the ONLY SOME confidence level and 19% were below this level. The bar graph provides a visual representation to each of the categories clearly indicating that the majority of respondent's confidence levels in regards to Major Companies are only some of the time.



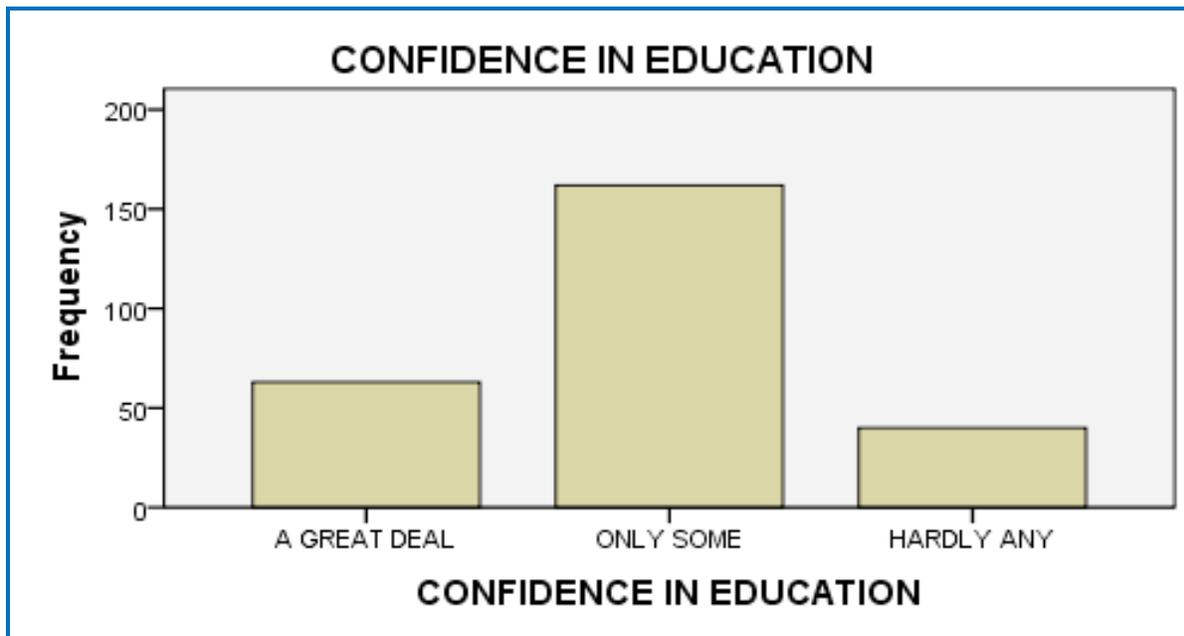
Q 2: How much confidence do you have in Organized Religion?

Of the sample population of 418 people who were presented with the survey tool, 262 provided a valid response to this question, 11 answered NA (NOT APPLICABLE), and 145 were evaluated as a DK (DON'T KNOW) response. Missing data to this question is attributed to the combined scores of 11 and 145 which equates to a total of 156 participants or 37.3% who responded in this manner. The most frequently recorded response to this question was ONLY SOME selected by 147 participants for a 35.2% response rate. A GREAT DEAL was selected by 61 participants for a 14.6% response rate, and HARDLY ANY was selected by 54 participants for a 12.9% response rate. The relative frequency of the 262 respondents out of the 418 participants indicates a proportional response of 62.7%. Cumulative Percentage represents the scores at or below each level. For example, 79.4% of the reported scores were at the ONLY SOME confidence level and 23.3% were below this level. The bar graph provides a visual representation to each of the categories clearly indicating that the majority of respondent's confidence levels in regards to Organized Religion are only some of the time.



Q 3: How much confidence do you have in Education?

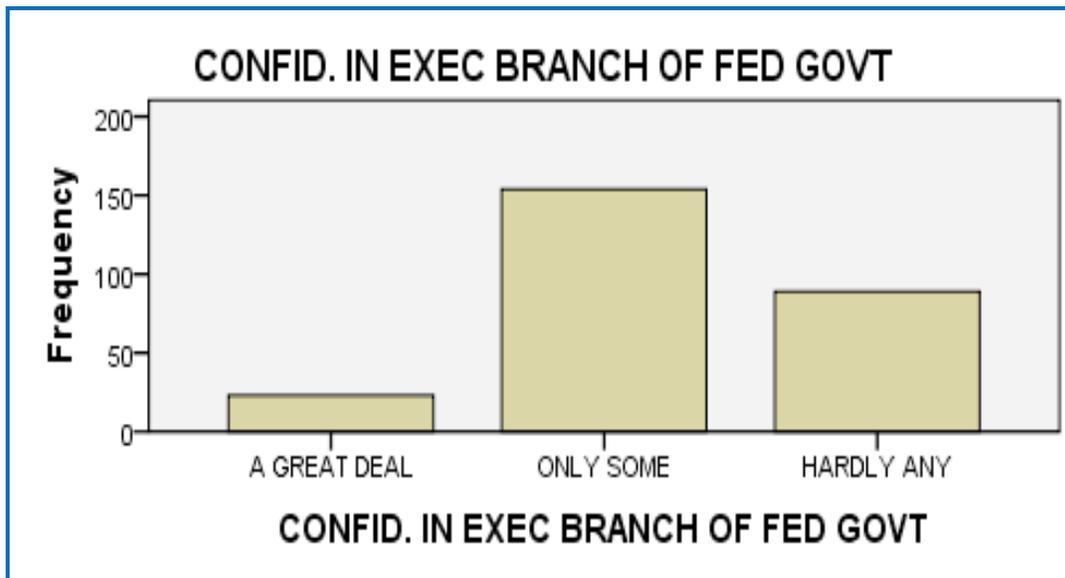
Of the sample population of 418 people who were presented with the survey tool, 265 provided a valid response to this question, 8 answered NA (NOT APPLICABLE), and 145 were evaluated as a DK (DON'T KNOW) response. Missing data to this question is attributed to the combined scores of 8 and 145 which equates to a total of 153 participants or 36.6% who responded in this manner. The most frequently recorded response to this question was ONLY SOME selected by 162 participants for a 38.8% response rate. A GREAT DEAL was selected by 63 participants for a 15.1% response rate, and HARDLY ANY was selected by 40 participants for a 9.6% response rate. The relative frequency of the 265 respondents out of the 418 participants indicates a proportional response of 63.4%. Cumulative Percentage represents the scores at or below each level. For example, 84.9% of the reported scores were at the ONLY SOME confidence level and 23.8% were below this level. The bar graph provides a visual representation to each of the categories clearly indicating that the majority of respondent's confidence levels in regards to Education are only some of the time.



Q 4: How much confidence do you have in the Executive Branch of Federal Govt.?

Of the sample population of 418 people who were presented with the survey tool, 266 provided a valid response to this question, 7 answered NA (NOT APPLICABLE), and 145 were evaluated as a DK (DON'T KNOW) response. Missing data to this question is attributed to the combined scores of 7 and 145 which equates to a total of 152 participants or 36.4% who responded in this manner. The most frequently recorded response to this question was ONLY SOME selected by 154 participants for a 36.8% response rate. A GREAT DEAL was selected by 23 participants for a 5.5% response rate, and HARDLY ANY was selected by 89 participants for a 21.3% response rate. The relative frequency of the 266 respondents out of the 418 participants indicates a proportional response of 63.6%.

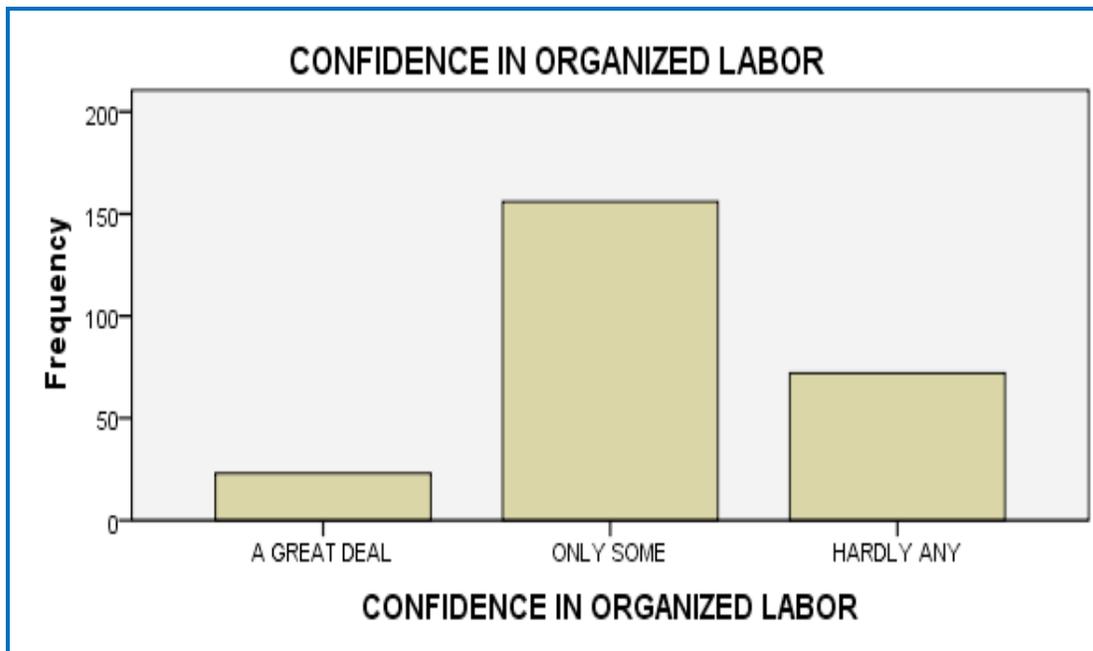
Cumulative Percentage represents the scores at or below each level. For example, 66.5% of the reported scores were at the ONLY SOME confidence level and 8.6% were below this level. The bar graph provides a visual representation to each of the categories clearly indicating that the majority of respondent's confidence levels in regards to the Executive Branch of the Federal Government are only some of the time.



Q 5: How much confidence do you have in Organized Labor?

Of the sample population of 418 people who were presented with the survey tool, 251 provided a valid response to this question, 22 answered NA (NOT APPLICABLE), and 145 were evaluated as a DK (DON'T KNOW) response. Missing data to this question is attributed to the combined scores of 22 and 145 which equates to a total of 167 participants or 40.0% who responded in this manner. The most frequently recorded response to this question was ONLY SOME selected by 156 participants for a 37.3% response rate. A GREAT DEAL was selected by 23 participants for a 5.5% response rate, and HARDLY ANY was selected by 72 participants for a 17.2% response rate. The relative frequency of the 251 respondents out of the 418 participants indicates a proportional response of 60.0%.

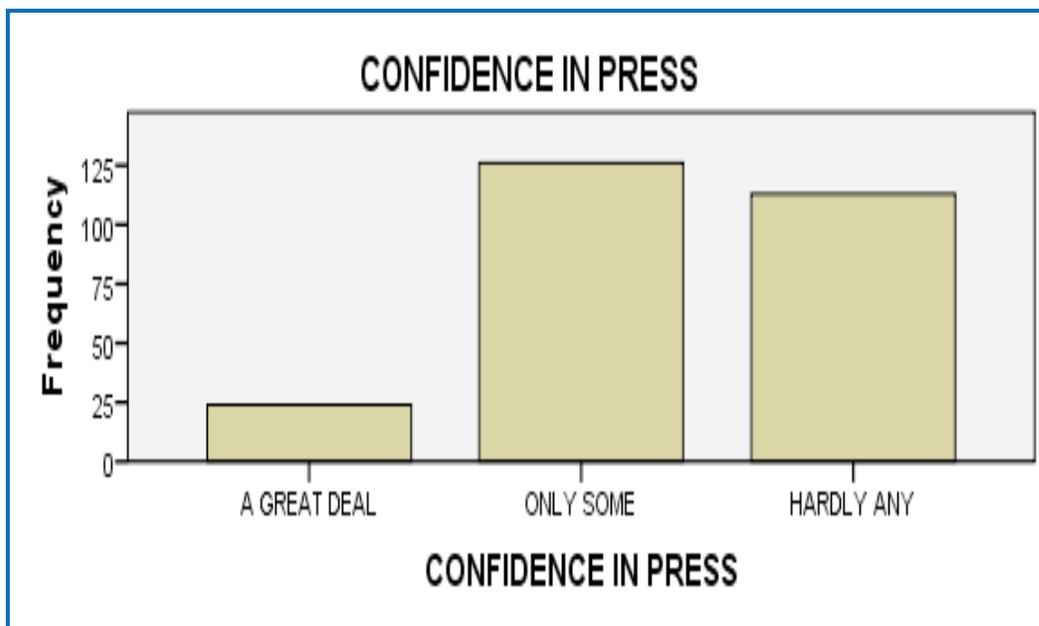
Cumulative Percentage represents the scores at or below each level. For example, 71.3% of the reported scores were at the ONLY SOME confidence level and 9.2% were below this level. The bar graph provides a visual representation to each of the categories clearly indicating that the majority of respondent's confidence levels in regards to Organized Labor are only some of the time.



Q 6: How much confidence do you have in the Press?

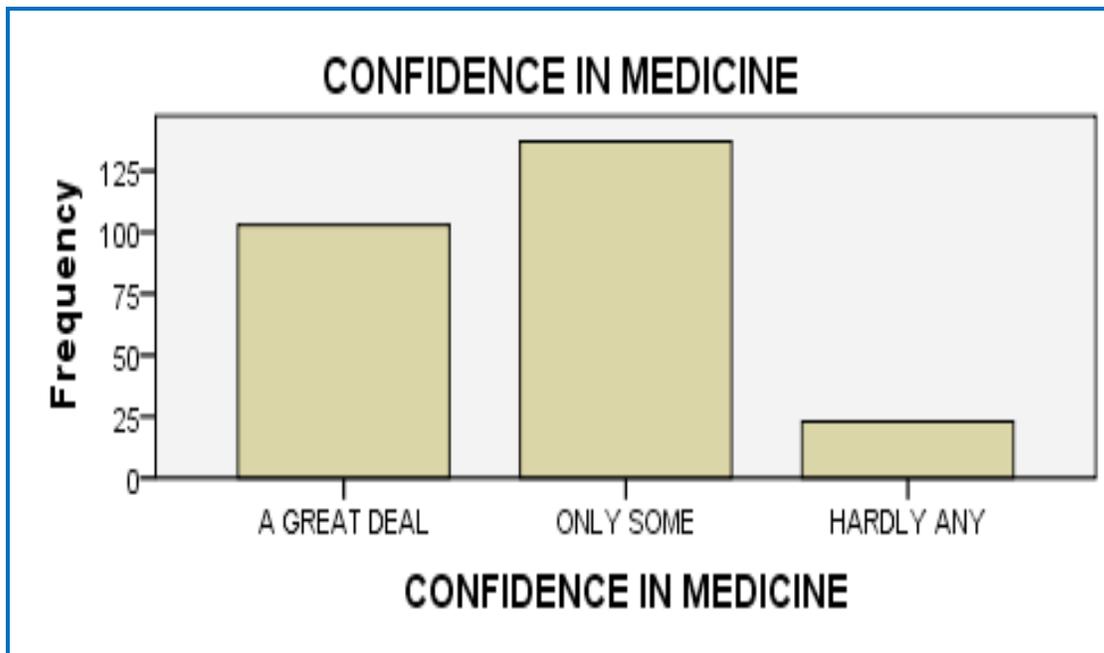
Of the sample population of 418 people who were presented with the survey tool, 263 provided a valid response to this question, 10 answered NA (NOT APPLICABLE), and 145 were evaluated as a DK (DON'T KNOW) response. Missing data to this question is attributed to the combined scores of 10 and 145 which equates to a total of 155 participants or 37.1% who responded in this manner. The most frequently recorded response to this question was ONLY SOME selected by 126 participants for a 30.1% response rate. A GREAT DEAL was selected by 24 participants for a 5.7% response rate, and HARDLY ANY was selected by 113 participants for a 27.0% response rate. The relative frequency of the 263 respondents out of the 418 participants indicates a proportional response of 62.9%.

Cumulative Percentage represents the scores at or below each level. For example, 57.0% of the reported scores were at the ONLY SOME confidence level and 9.1% were below this level. The bar graph provides a visual representation to each of the categories clearly indicating that the majority of respondent's confidence levels in regards to the Press are only some of the time.



Q 7: How much confidence do you have in Medicine?

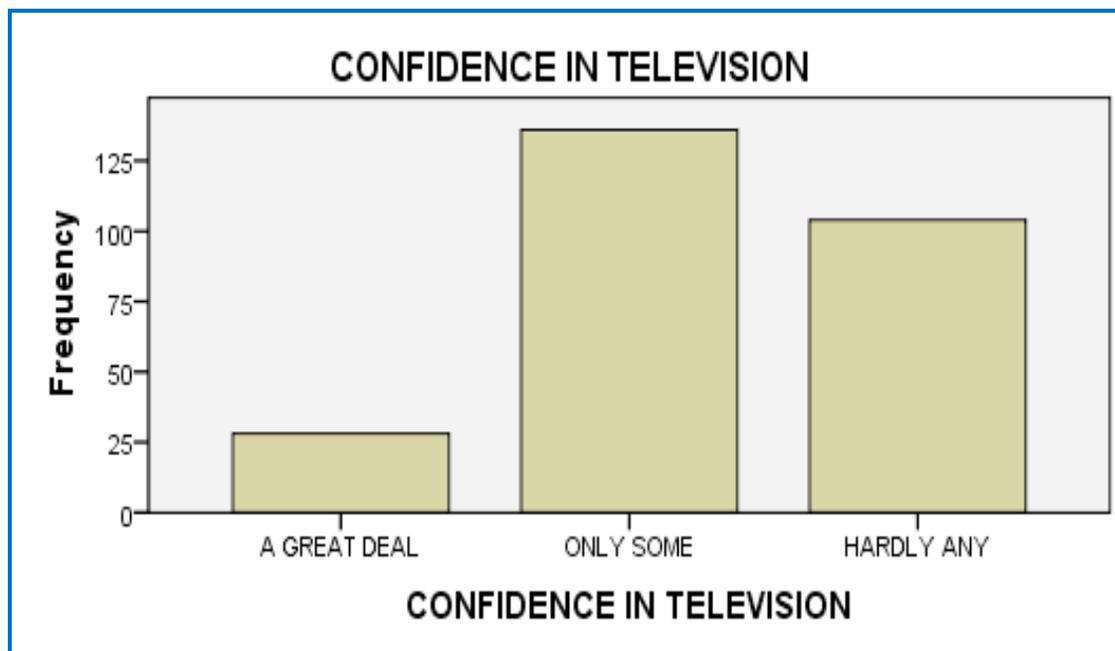
Of the sample population of 418 people who were presented with the survey tool, 263 provided a valid response to this question, 10 answered NA (NOT APPLICABLE), and 145 were evaluated as a DK (DON'T KNOW) response. Missing data to this question is attributed to the combined scores of 10 and 145 which equates to a total of 155 participants or 37.1% who responded in this manner. The most frequently recorded response to this question was ONLY SOME selected by 137 participants for a 32.8% response rate. A GREAT DEAL was selected by 103 participants for a 24.6% response rate, and HARDLY ANY was selected by 23 participants for a 5.5% response rate. The relative frequency of the 263 respondents out of the 418 participants indicates a proportional response of 62.9%. Cumulative Percentage represents the scores at or below each level. For example, 91.3% of the reported scores were at the ONLY SOME confidence level and 39.2% were below this level. The bar graph provides a visual representation to each of the categories clearly indicating that the majority of respondent's confidence levels in regards to Medicine are only some of the time.



Q 8: How much confidence do you have in Television?

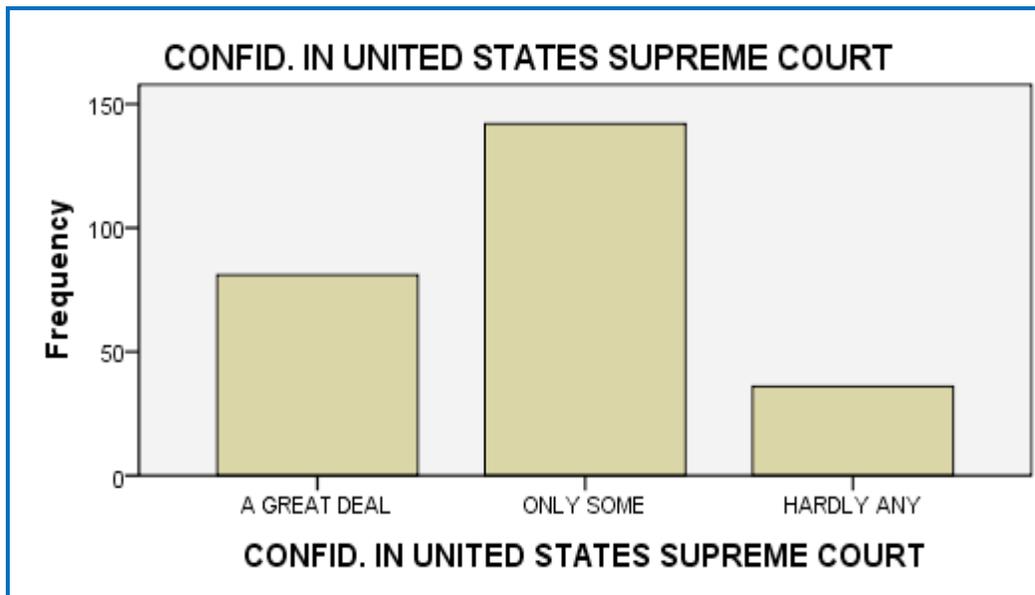
Of the sample population of 418 people who were presented with the survey tool, 268 provided a valid response to this question, 5 answered NA (NOT APPLICABLE), and 145 were evaluated as a DK (DON'T KNOW) response. Missing data to this question is attributed to the combined scores of 5 and 145 which equates to a total of 150 participants or 35.9% who responded in this manner. The most frequently recorded response to this question was ONLY SOME selected by 136 participants for a 32.5% response rate. A GREAT DEAL was selected by 28 participants for a 6.7% response rate, and HARDLY ANY was selected by 104 participants for a 24.9% response rate. The relative frequency of the 268 respondents out of the 418 participants indicates a proportional response of 64.1%.

Cumulative Percentage represents the scores at or below each level. For example, 61.2% of the reported scores were at the ONLY SOME confidence level and 10.4% were below this level. The bar graph provides a visual representation to each of the categories clearly indicating that the majority of respondent's confidence levels in regards to Television are only some of the time.



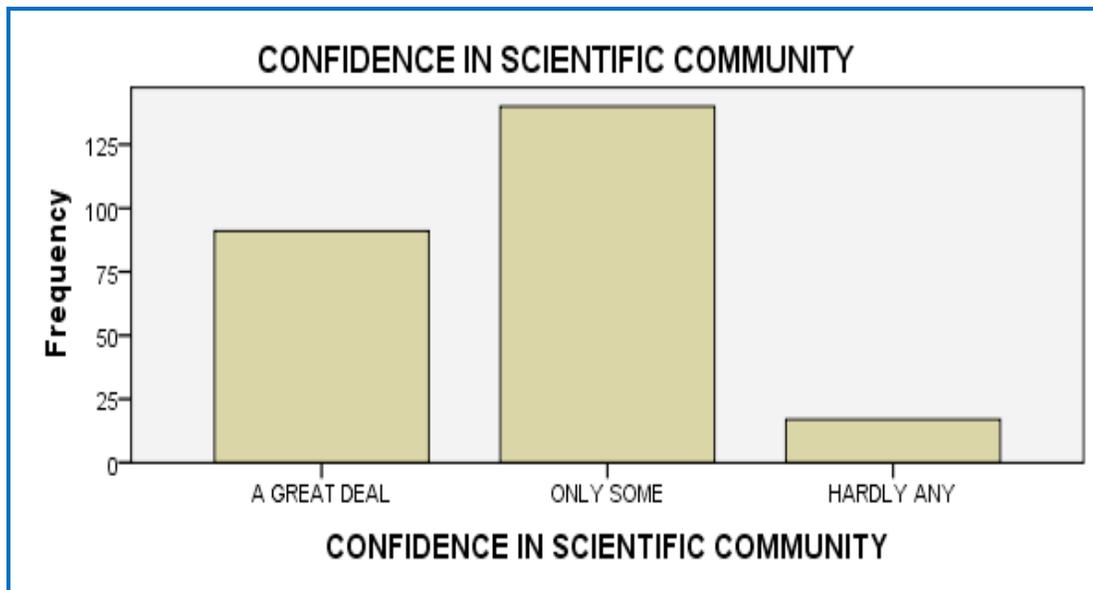
Q 9: How much confidence do you have in the United States Supreme Court?

Of the sample population of 418 people who were presented with the survey tool, 259 provided a valid response to this question, 14 answered NA (NOT APPLICABLE), and 145 were evaluated as a DK (DON'T KNOW) response. Missing data to this question is attributed to the combined scores of 14 and 145 which equates to a total of 159 participants or 38.0% who responded in this manner. The most frequently recorded response to this question was ONLY SOME selected by 142 participants for a 34.0% response rate. A GREAT DEAL was selected by 81 participants for a 19.4% response rate, and HARDLY ANY was selected by 36 participants for an 8.6% response rate. The relative frequency of the 259 respondents out of the 418 participants indicates a proportional response of 62.0%. Cumulative Percentage represents the scores at or below each level. For example, 86.1% of the reported scores were at the ONLY SOME confidence level and 31.3% were below this level. The bar graph provides a visual representation to each of the categories clearly indicating that the majority of respondent's confidence levels in regards to the United States Supreme Court are only some of the time.



Q 10: How much confidence do you have in the Scientific Community?

Of the sample population of 418 people who were presented with the survey tool, 248 provided a valid response to this question, 25 answered NA (NOT APPLICABLE), and 145 were evaluated as a DK (DON'T KNOW) response. Missing data to this question is attributed to the combined scores of 25 and 145 which equates to a total of 170 participants or 40.7% who responded in this manner. The most frequently recorded response to this question was ONLY SOME selected by 140 participants for a 33.5% response rate. A GREAT DEAL was selected by 91 participants for a 21.8% response rate, and HARDLY ANY was selected by 17 participants for a 4.1% response rate. The relative frequency of the 248 respondents out of the 418 participants indicates a proportional response of 59.3%. Cumulative Percentage represents the scores at or below each level. For example, 93.1% of the reported scores were at the ONLY SOME confidence level and 36.7% were below this level. The bar graph provides a visual representation to each of the categories clearly indicating that the majority of respondent's confidence levels in regards to the Scientific Community are only some of the time.



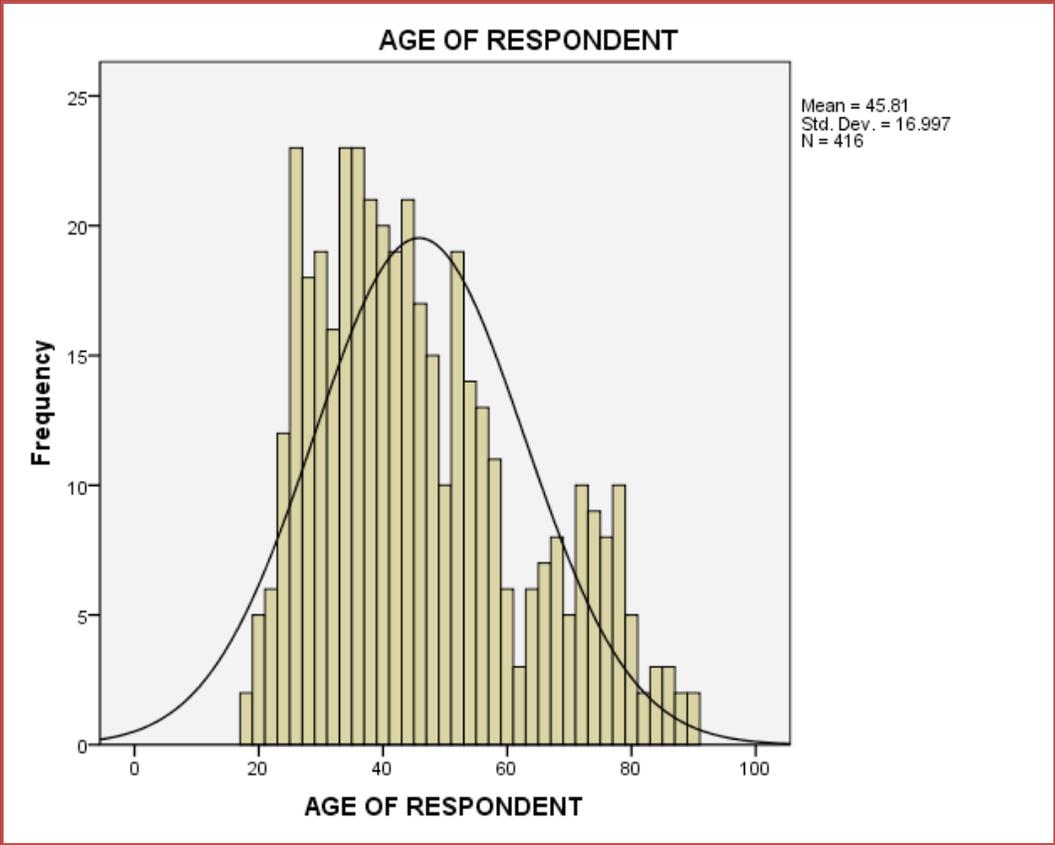
Continuous Variables: AGE, WORDSUM, and TVHOURS

As part of the overall analysis for PAP, the next step is to review three specific variables (characteristics) from the survey poll identified as *AGE*, *WORDSUM* and *TVHOURS*. These variables are considered mutually exclusive categories (classes). *AGE* represents the average chronological age of the respondents to the survey poll; *WORDSUM* denotes the number of words correctly selected in the vocabulary test, and, *TVHOURS* indicate the hours each participant invests per day watching Television. The data will be presented in two formats: the results of the statistical analysis will be identified and explained in a textual format, and, secondary analysis will be presented in a graphical format known as a histogram.

A histogram is a bar chart signifying frequency distribution. The height of each bar indicates the density (frequency) or concentration of the responses; the bars are placed next to each other demonstrating that the data that has been gathered is continuous. The width of each bar represents the lowest score and the highest score in the interval. The distance between the scores is known as the lower and upper real limits of the score and are reflected by the touching of the bars to eliminate any gaps between them.

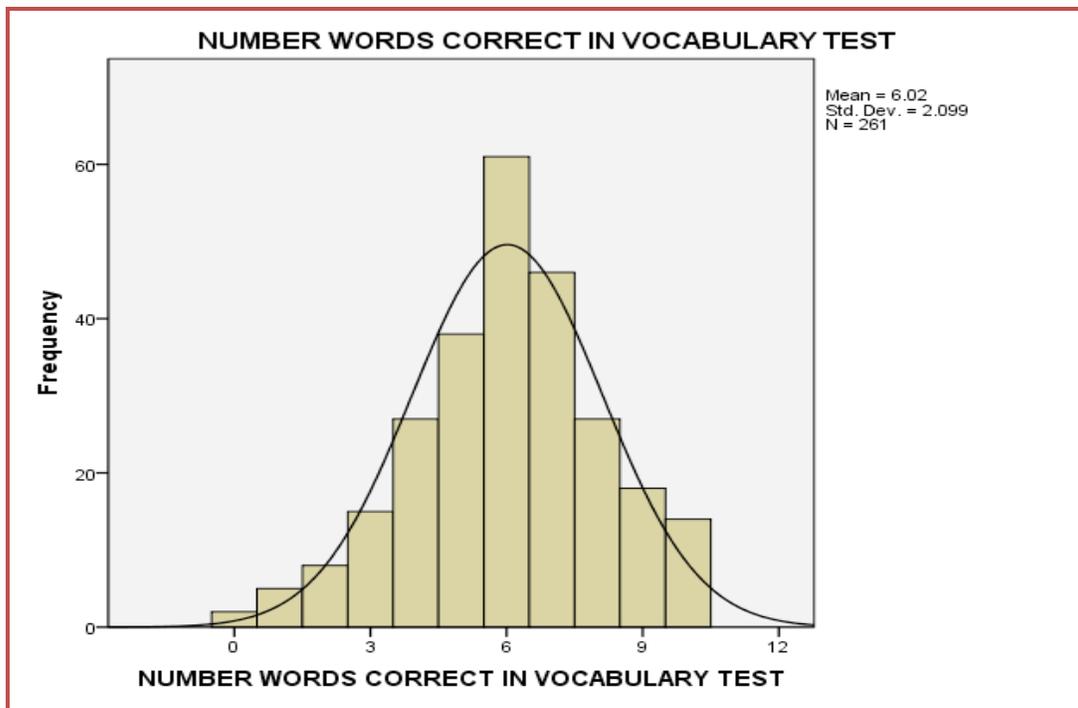
Each of the variables will be evaluated using the following descriptive statistics: Mean (average), Standard Deviation (how far plus or minus from the mean), Skewness (asymmetry/disproportionateness of the distribution) and Kurtosis (the flatness or peakedness of the curve). The number of observations (classifications per category) will also be provided for the entire sample. Of the 418 participants in the survey, only 416 responded to the question *What is your age?* The average age (mean) was 45.81, the standard deviation was +/- 16.997 (above/below the mean) which indicates the age range was between 33.994 and 62.807 years of age.

The Skewness score for this variable was a positive .617. A positive asymmetrical skew indicates that tail on the right side of the curve is longer than the left side and the majority of the values lie to the left of the mean as is evidence in the histogram.



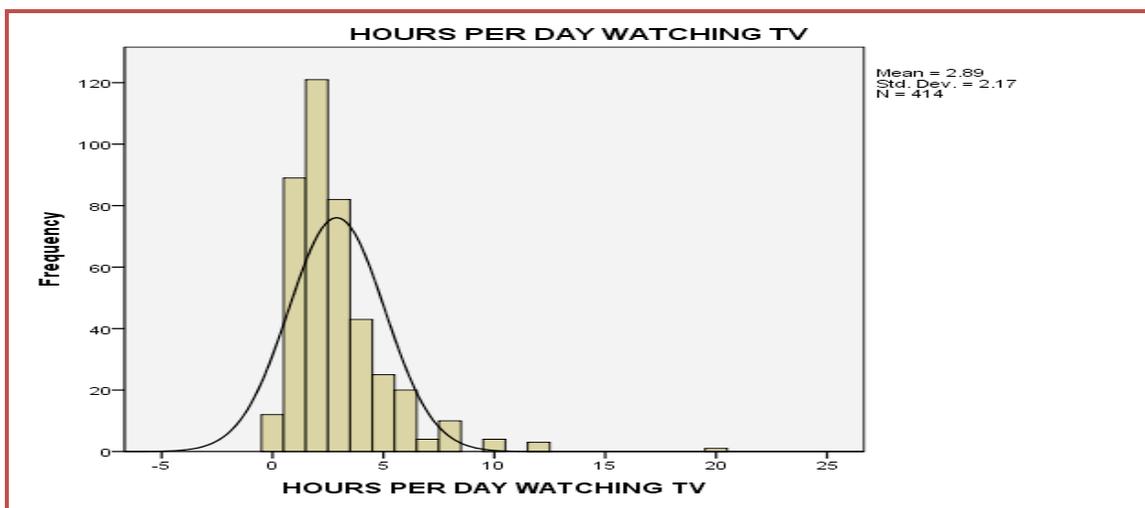
The Kurtosis statistic is $-.508$ since it is less than zero ($\beta_2 < 0$), it is considered a flat-topped curve. The flatness of the curve indicates that there are moderate frequency deviations.

Of the 418 participants in the survey, only 261 responded to the vocabulary test which calculated the number of words they selected correctly (*WORDSUM*). The average (mean) was 6.02, the standard deviation was +/- 2.099 (above/below the mean) which indicates the correct word range was between 3.921 and 8.119 of the words selected. The Skewness score for this variable was a negative -.267. A negative asymmetrical skew indicates that tail on the left side of the curve is longer than the right side and the majority of the values (including the median) lie to the right of the mean as is evidence in the histogram.



The Kurtosis statistic is .006 and since it is greater than zero ($\beta_2 > 0$), it is considered to be a distribution with a high peak. A higher Kurtosis indicates there is more of a variance which results from infrequent extreme deviations.

Of the 418 participants in the survey, only 414 responded to the question *How much confidence do you have in Television?* The answer was based upon the hours of television watched per day by the participant. The average (mean) was 2.89 hours per day, the standard deviation was +/- 2.170 (above/below the mean) which indicates a lower limit of .72 hours and upper limit of 5.06 hours of television watching per participant. The Skewness score for this variable was a positive 2.434. A positive asymmetrical skew indicates that tail on the right side of the curve is longer than the left side and the majority of the values lie to the left of the mean as is evidence in the histogram.



The Kurtosis statistic is 11.249 and since it is significantly greater than zero ($\beta_2 > 0$), it is considered to be a distribution with a high peak. A higher Kurtosis indicates there is more of a variance which results from infrequent extreme deviations. It should be noted, *TVHOURS* had **(8)** extreme values as noted in the following z-scores (standard scores):

- 3.27432 (4)
- 4.10585 (3)
- 7.88197 (1)

These extremes scores (in the right tail) correspond to exact positions in the distribution above the mean (the average of the scores). Four scores of 3.27432 were exactly (3) standard deviations above the

mean, three scores of 4.10585 were exactly (4) standard deviations above the mean, and one score of 7.88197 was exactly (1) standard deviation above the mean. The standard deviation is the most important measure of variability and is used to determine the distance of the scores relative to the mean (average). Generally, this measures (similar to a ruler) how near or far the scores are from the mean representing their average distance. The value to PAP is recognizing the scores in the tail of the distribution (those in excess of + or - 3) is that these results are unlikely due to sampling error, the results are therefore, statistically significant.

Respondents by Sex

When considering the sample reader base, it is also important to understand the effect gender may have on the recorded observations. Each of the three prior continuous variables were re-run with Sex (gender) as a co-variant. Table 1 is a summation of valid participants who responded to the question: *What is your age?*

Table 1

Gender Comparison: Age of the Respondent

AGE					
Male			Female		
Mean	42.26		Mean	47.42	
95% Confidence Interval for Mean	Lower Bound	39.5	95% Confidence Interval for Mean	Lower Bound	44.52
	Upper Bound	45.02		Upper Bound	50.31
5% Trimmed Mean	41.47		5% Trimmed Mean	46.89	
Median	39		Median	44	
Variance	220.992		Variance	308.93	
Std. Deviation	14.866		Std. Deviation	17.576	
Minimum	19		Minimum	21	
Maximum	86		Maximum	89	
Range	67		Range	68	
Interquartile Range	20		Interquartile Range	27	
Skewness	0.786		Skewness	0.48	
Kurtosis	0.073		Kurtosis	-0.888	

In this side by side comparison, PAP can quickly focus on several important characteristics of their sample population: With regards to those participants who answered this question (Male – 114 valid responses, 57 missing data and Female – 144 valid responses, 103 missing data), Females were on average 5 years older, the Skewness score for both Males and Females was positive, however there was an opposite effect with the Kurtosis score: Males had a positive 0.073 Kurtosis score and Females had a negative Kurtosis score of -0.888. Since the Male Kurtosis statistic is greater than zero ($\beta_2 > 0$),

it is considered to be a distribution with a high peak; since the Kurtosis statistic for Females is less than zero ($\beta_2 < 0$), it is considered a flat-topped curve.

In this table, a 95% confidence level was incorporated to estimate the unknown population mean. By increasing the confidence level to 95%, we also increased the width of the interval. In doing so, we traded precision (with a point estimate – single value) for a higher level of confidence (interval estimate – range of values). For Males, the lower bound was 39.5 and the upper bound was 45.02; and, for Females the lower bound is 44.52 and the upper bound is 50.31. We are therefore 95% confident that the unknown population means for Males and Females are located within their respective intervals.

Notice also the shape of the boxes and the whiskers (the two outer lines). The box represents the middle 50% of the data sample; the other 50% are represented in the area between the box and the whiskers. The exceptions are the outliers 10 and 105 located in the Male data which are the extreme values that deviate significantly from the rest of the data set. The location of the box between the whiskers provides insight into the Skewness (the normality) of the sample. For Males and Females, it is shifted significantly to the lower end of the whisker, so it is positively skewed. An estimate of Kurtosis (the flatness or peakedness of the distribution) is determined by the size of the box. The thinness or wideness of the box relative to the whiskers determines the shape of the peak.



The Male box result is thinner than the Female box, and as such, indicates a concentration of cases in a small segment of the sample resulting in a thinner peak. In contrast, the Female box result is considerably broader, indicating a wider peak. Lastly, the black line through the boxes represents the median (the middle value of the entire sample); 39 years of age for Males and 44 years of age for Females.

Table 2 represents a summation of the valid participants who responded to the vocabulary part of the survey: *Number of correct words obtained on the vocabulary test.*

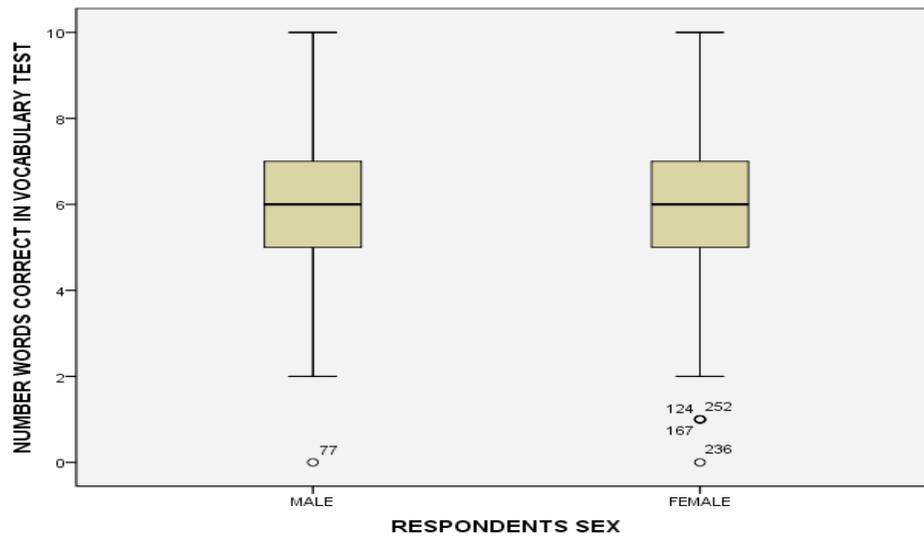
Table 2

Number of correct words obtained on the vocabulary test

WORDSUM					
Male			Female		
Mean		6.15	Mean		5.88
95% Confidence Interval for Mean	Lower Bound	5.78	95% Confidence Interval for Mean	Lower Bound	5.52
	Upper Bound	6.52		Upper Bound	6.23
5% Trimmed Mean		6.17	5% Trimmed Mean		5.92
Median		6	Median		6
Variance		4.04	Variance		4.726
Std. Deviation		2.01	Std. Deviation		2.174
Minimum		0	Minimum		0
Maximum		10	Maximum		10
Range		10	Range		10
Interquartile Range		2	Interquartile Range		2
Skewness		-0.142	Skewness		-0.293
Kurtosis		0.113	Kurtosis		-0.095

As with the previous question, the number of male and female respondents to this part of the survey remained equivalent. On average, males had a higher number of correct words (6.15) compared to females (5.88). The Skewness scores for both males and females are negative which means the scores tend to pile up on the right side of the (mean) scale and taper off gradually to the left side (tail) of the

distribution. PAP should interpret this data to mean the word choice for the vocabulary level for the paper was relatively understood by both males and females. With regards to Kurtosis, males had a positive result of 0.113 while females had a negative result of -0.095. Since the Male Kurtosis statistic is greater than zero ($\beta_2 > 0$), it is considered to be a distribution with a high peak; since the Kurtosis statistic for Females is less than zero ($\beta_2 < 0$), it is considered a flat-topped curve.



For the variable, *WORDSUM*, we are reporting on a higher level of confidence (95%) rather than on the precision of the estimate with regards to the unknown population means for males and females. The intervals for males was lower bound at 5.78 and upper bound at 6.52 and for females, the lower bound was 5.52 and the upper bound was 6.23. At at 95% confidence level, we assert the unknown population means for Males and Females are located within their respective intervals. As previously discussed, we now turn our attention to the location of the boxes between their respective whiskers to provide insight into the normality (ordinariness) of the sample. In regards to males and females, their respective boxes are both shifted slightly to the top of the whiskers which means both of their results are negatively skewed. The width of both boxes is relatively close in size. However, the female peak (-0.095) will be a slightly flatter curve than the wider, higher male peak (0.113). In the male responses,

there was one exception (outlier) 77 identified. The female responses had four exceptions (outliers): 124,167,236, and 252. The median (middle value for this sample is 6) for both the males and females.

The final variable, *TVHOURS*, is to be evaluated in this section. Table 3 represents a summation of the valid participants who responded to the question: *How much confidence do you have in Television?*

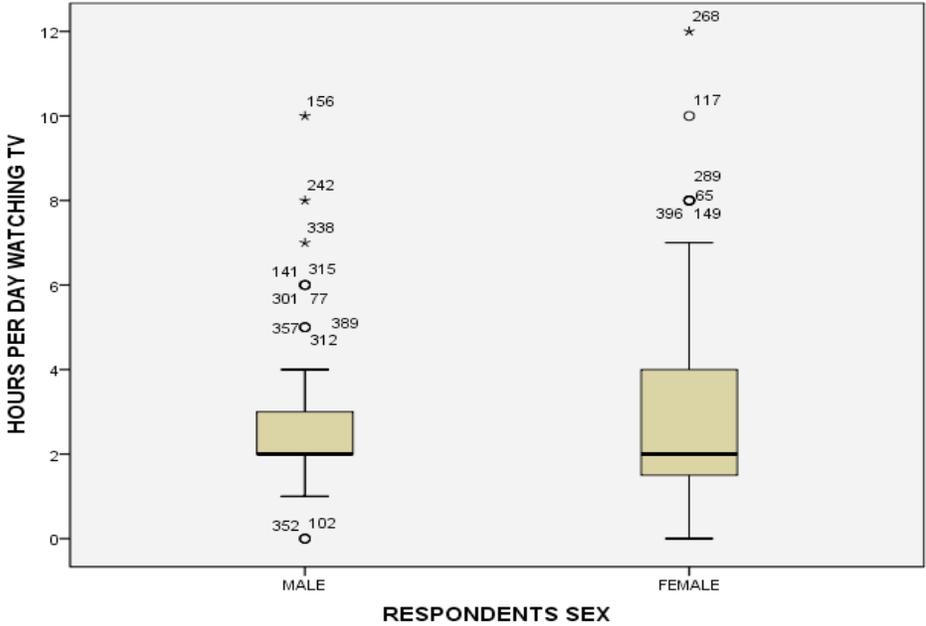
Table 3

How much confidence do you have in Television?

TVHOURS					
Male			Female		
Mean		2.59	Mean		2.9
95% Confidence Interval for Mean	Lower Bound	2.29	95% Confidence Interval for Mean	Lower Bound	2.56
	Upper Bound	2.88		Upper Bound	3.23
5% Trimmed Mean		2.45	5% Trimmed Mean		2.69
Median		2	Median		2
Variance		2.545	Variance		4.052
Std. Deviation		1.595	Std. Deviation		2.013
Minimum		0	Minimum		0
Maximum		10	Maximum		12
Range		10	Range		12
Interquartile Range		1	Interquartile Range		3
Skewness		1.647	Skewness		1.6
Kurtosis		4.454	Kurtosis		3.376

The average (mean) time spent watching TV is slightly higher for females (2.9) vs. males (2.59) hours. When expressing their confidence in TV, the level of skewness for males was 1.647 and for females it was 1.6. Again, relatively close results. Since both of these statistics are positively skewed, PAP can assume this to mean the majority of the scores stacked up on the left side of the mean and the extreme scores were in the right tail of the distribution. Another way PAP can view these results is in relationship to the asymmetry/disproportionateness of the normal curve when considering

how these outcomes effect the shape of the curve. In considering bounds of the intervals, males had a lower bound of 2.29 and an upper bound of 2.88 while females had a lower bound of 2.56 and an upper bound of 3.23. With a 95% confidence level, we again affirm that the unknown population means for Males and Females are located within their respective intervals. The Kurtosis statistic is 4.454 for males and 3.376 for females. Since these scores are significantly greater than zero ($\beta_2 > 0$), it is considered to be a distribution with a high peak. A higher Kurtosis indicates there is more of a variance which results from infrequent extreme deviations. Of the three variables analyzed, *TVHOURS* is the variable (phenomenon) with the most extreme (outliers).



As evidenced above, this box plot is the most interesting of the three because of the number of outliers and the shape and size of each of the boxes in relationship to their relative location with their respective whiskers. Clearly, there are more outliers in the male response than in the female responses. The male box is considerably narrower than the female box. Again, the indication is the concentration of cases in a small sample resulting in a more narrow peak and the variance is higher for males than for females. The median (middle value for this sample is 2) for both the males and females.

1. We would like to know if there are any differences by gender (SEX) with respect to the variables AGE, WORDSUM, and TVHOURS.

To evaluate the difference by gender with respect to age of respondent, number of correct words obtained on a vocabulary test, and the number of hours of TV watched per day. In this survey 114 males were surveyed and 144 females were surveyed. For males an average age of 42 was surveyed and for females an average age of 47 was observed. Males were observed to achieve slightly over 6 words correctly in a vocabulary test and females were observed to achieve slightly fewer than 6 words correctly in a vocabulary test. Males were found to watch 2.6 hours per day of TV and females were found to watch 2.9 hours per day of TV.

In this survey two sets of data (male and female) have each come from the same survey sample but are two separate samples. We have used Levene's test for equality of variances as an inferential statistic to assess the different samples. We have assumed the variances of the populations from which our samples are taken are equal. The used test, Levene's test, puts this assumption to statistical test. The null hypothesis, that population variances are equal, is tested in this way. If the critical value obtained is less than our critical value (0.05) it is determined that the differences in sample variance is unlikely to have occurred based on random sampling.

That being said, a null hypothesis is rejected and it is said that there is a difference between the variances in the population. The following levels of significance were obtained .012 (age of respondent), .411 (wordsum), and .035 (hours spent watching TV). The critical values for age of respondent and hours spent watching TV are less than the critical value of .05 which imply a difference by gender while the critical value for wordsum is above our critical value and our null hypothesis is not rejected.

Group Statistics

RESPONDENTS SEX		N	Mean	Std. Deviation	Std. Error Mean
AGE OF RESPONDENT	MALE	114	42.26	14.866	1.392
	FEMALE	144	47.42	17.576	1.465
NUMBER WORDS CORRECT IN VOCABULARY TEST	MALE	114	6.15	2.010	.188
	FEMALE	144	5.88	2.174	.181
HOURS PER DAY WATCHING TV	MALE	114	2.59	1.595	.149
	FEMALE	144	2.90	2.013	.168

Independent Samples Test

		Levene's Test for Equality of Variances				t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
AGE OF RESPONDENT	Equal variances assumed	6.424	.012	-2.501	256	.013	-5.154	2.060	-9.211	-1.096
	Equal variances not assumed			-2.550	254.856	.011	-5.154	2.021	-9.133	-1.174
NUMBER WORDS CORRECT IN VOCABULARY TEST	Equal variances assumed	.678	.411	1.040	256	.299	.274	.264	-.245	.793
	Equal variances not assumed			1.049	249.870	.295	.274	.261	-.240	.789
HOURS PER DAY WATCHING TV	Equal variances assumed	4.484	.035	-1.335	256	.183	-.308	.231	-.762	.146

Independent Samples Test

		Levene's Test for Equality of Variances				t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
AGE OF RESPONDENT	Equal variances assumed	6.424	.012	-2.501	256	.013	-5.154	2.060	-9.211	-1.096
	Equal variances not assumed			-2.550	254.856	.011	-5.154	2.021	-9.133	-1.174
NUMBER WORDS CORRECT IN VOCABULARY TEST	Equal variances assumed	.678	.411	1.040	256	.299	.274	.264	-.245	.793
	Equal variances not assumed			1.049	249.870	.295	.274	.261	-.240	.789
	Equal variances assumed	4.484	.035	-1.335	256	.183	-.308	.231	-.762	.146
	Equal variances not assumed			-1.372	255.999	.171	-.308	.225	-.751	.134

2. May we say that the average age of the population is 46 years?

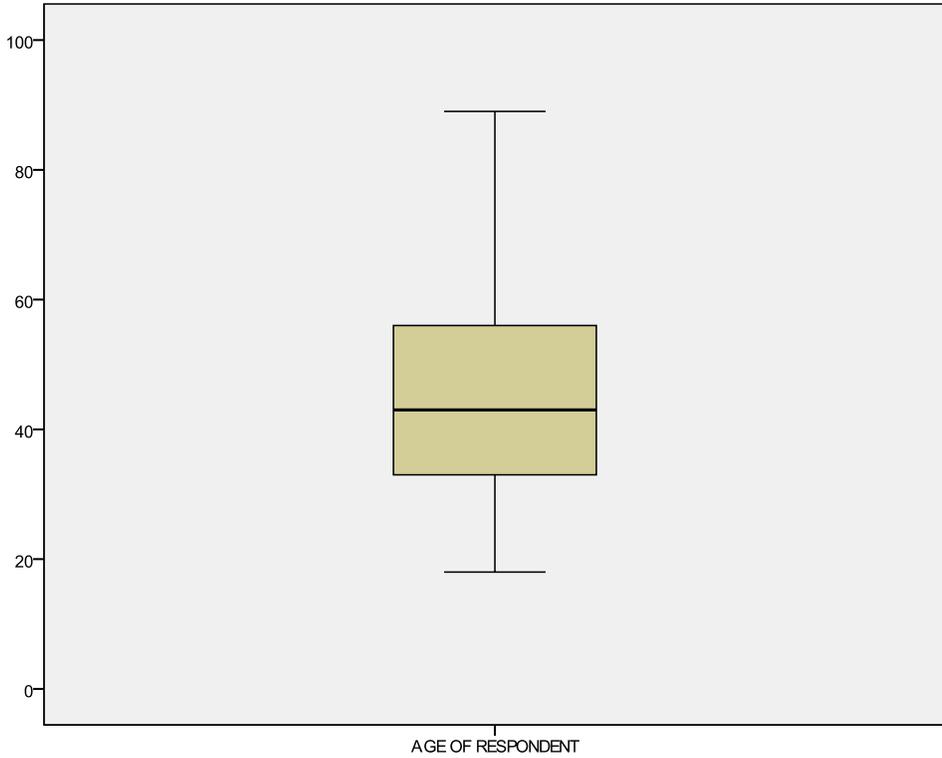
Given the results calculated from the survey, which is representative of the general population, it is possible to say that the average age of the population is approximately 46 years of age. We are 95% confident that the average age of the population is between 44.18 and 47.45. Our confidence interval is used to show how reliable our sample estimates the population in that respect. A stem and leaf plot is included to give some dimension to the values, from this the grouping of the values can easily be discerned. A bow plot is also included to give those figures a more graphical representation. In each it is visible that the mean is close to 46 years of age.

Descriptives			Statistic	Std. Error
AGE OF RESPONDENT	Mean		45.81	.833
	95% Confidence Interval for Mean	Lower Bound	44.18	
		Upper Bound	47.45	
	5% Trimmed Mean		45.14	
	Median		43.00	
	Variance		288.893	
	Std. Deviation		16.997	
	Minimum		18	
	Maximum		89	
	Range		71	
	Interquartile Range		23	
	Skewness		.617	.120
	Kurtosis		-.508	.239

AGE OF RESPONDENT Stem-and-Leaf Plot

Frequency	Stem &	Leaf
5.00	1 .	88999
20.00	2 .	00111222333333344444
53.00	2 .	
5555555555556666666666777777888888888888999999999999	3 .	
46.00	3 .	
000000011111111222222233333333333333344444444444	3 .	
54.00	3 .	
55555555555555556666666666777777777788888888888899999999999	4 .	
50.00	4 .	
0000000000111111111222222222333333333333333333344444444	4 .	55555555566666666677777778888888888999
35.00	4 .	
40.00	5 .	000000011111111111122222233333333334444444
26.00	5 .	55555566666677788888888899
13.00	6 .	0000112333344
18.00	6 .	5556666778888888999
21.00	7 .	00112222222333333444
20.00	7 .	555666667777788888899
8.00	8 .	00011344
7.00	8 .	5567899

Stem width: 10
 Each leaf: 1 case(s)



3. Are hours spent watching TV related to a subject's age?

To determine if there is a relation between the two variables of interest, here age of the respondent and the hours per day they watch TV, we ran a correlation. Correlation tests give values from negative one to positive one and are representative of how related two variables are. A large value (closer to positive one) implies a positive correlation. In the case a value close to positive one would imply a positive relationship: as age increases so does the hours of day spent watching TV. Conversely, a very small value (closer to negative 1), would imply an inverse relation between age and hours per day spent watching TV. In the case low age would be associated with a large number of hours spent watching TV, or the opposite scenario were high age values are associated with a low amount of hours being spent watching TV.

For our survey we have determined these correlations in the following way:

- -1.0 to -0.7 strong negative association.
- -0.7 to -0.3 weak negative association.
- -0.3 to +0.3 little or no association.
- +0.3 to +0.7 weak positive association.
- +0.7 to +1.0 strong positive association.

Given the value of .141 we have determined there is ***little or no association*** between the age of the respondent and the hours per day they watch TV.

Correlations

		AGE OF RESPONDENT	HOURS PER DAY WATCHING TV
AGE OF RESPONDENT	Pearson Correlation	1	.141**
	Sig. (2-tailed)		.004
	N	416	412
HOURS PER DAY WATCHING TV	Pearson Correlation	.141**	1
	Sig. (2-tailed)	.004	
	N	412	414

Correlations

		AGE OF RESPONDENT	HOURS PER DAY WATCHING TV
AGE OF RESPONDENT	Pearson Correlation	1	.141**
	Sig. (2-tailed)		.004
	N	416	412
HOURS PER DAY WATCHING TV	Pearson Correlation	.141**	1
	Sig. (2-tailed)	.004	
	N	412	414

**. Correlation is significant at the 0.01 level (2-tailed).

4. Is smoking behavior related to gender?

For our survey we have determined these correlations in the following way:

- -1.0 to -0.7 strong negative association between gender and smoking status.
- -0.7 to -0.3 weak negative association between gender and smoking status.
- -0.3 to +0.3 little or no association between gender and smoking status.
- +0.3 to +0.7 weak positive association between gender and smoking status.
- +0.7 to +1.0 strong positive association between gender and smoking status.

Given the value of .072 we have determined there is ***no association*** between the gender of the respondent and if they chose to smoke or not.

Correlations

		RESPONDENTS SEX	DOES R SMOKE
RESPONDENTS SEX	Pearson Correlation	1	.072
	Sig. (2-tailed)		.238
	N	418	271
DOES R SMOKE	Pearson Correlation	.072	1
	Sig. (2-tailed)	.238	
	N	271	271

Correlations

		RESPONDENTS SEX	DOES R SMOKE
RESPONDENTS SEX	Pearson	1	.072
	Correlation		
	Sig. (2-tailed)		.238
	N	418	271
	Pearson	.072	1
	Correlation		
	Sig. (2-tailed)	.238	

5. Additional points of interest: Is there a correlation between interest in protecting the environment and the nation's health?

We have also identified two additional points of interest from the survey conducted. The respondents were asked if they were concerned with improving and protecting the environment, improving and protecting the nation's health, and interest in improving the education system.

- -1.0 to -0.7 strong negative association between improving and protecting the environment and the nation's health
- -0.7 to -0.3 weak negative association between improving and protecting the environment and the nation's health
- -0.3 to +0.3 little or no association between improving and protecting the environment and the nation's health
- +0.3 to +0.7 weak positive association between improving and protecting the environment and the nation's health
- +0.7 to +1.0 strong positive association between improving and protecting the environment and the nation's health

Based on the determined value (+.305) we have determined that there is a weak positive association between improving and protecting the environment and the nation's health. There is a weak association that shows that as one becomes more concerned with improving and protecting the environment they become more interested in improving and protecting the nation's health.

Correlations

		IMPROVING & PROTECTING ENVIRONMEN T	IMPROVING & PROTECTING NATIONS HEALTH
IMPROVING & PROTECTING ENVIRONMENT	Pearson Correlation	1	.305**
	Sig. (2-tailed)		.000
	N	200	196
IMPROVING & PROTECTING NATIONS HEALTH	Pearson Correlation	.305**	1
	Sig. (2-tailed)	.000	
	N	196	206

** . Correlation is significant at the 0.01 level (2-tailed).

Additional point of interest: Is there a correlation between interest in improving the education system and the nation's health?

- -1.0 to -0.7 strong negative association between improving the education system and the nation's health
- -0.7 to -0.3 weak negative association between the education system and the nation's health
- -0.3 to +0.3 little or no association between improving the education system and the nation's health
- +0.3 to +0.7 weak positive association between improving the education system and the nation's health
- +0.7 to +1.0 strong positive association between improving the education system and the nation's health

Based on the determined value (+.308) we have determined that there is a weak positive association between improving the nation's education system and the nation's health. There is a weak association that shows that as one becomes more concerned with improving the education system they become more interested in improving and protecting the nation's health.

Correlations

		IMPROVING & PROTECTING NATIONS HEALTH	IMPROVING NATIONS EDUCATION SYSTEM
IMPROVING & PROTECTING NATIONS HEALTH	Pearson Correlation Sig. (2-tailed) N	1 206	.308** .000 200
IMPROVING NATIONS EDUCATION SYSTEM	Pearson Correlation Sig. (2-tailed) N	.308** .000 200	1 204

** . Correlation is significant at the 0.01 level (2-tailed).

Additional point of interest: Is there a correlation between interest in improving the education system and the environment?

- -1.0 to -0.7 strong negative association between improving the education system and the environment
- -0.7 to -0.3 weak negative association between the education system and the environment
- -0.3 to +0.3 little or no association between improving the education system and the environment
- +0.3 to +0.7 weak positive association between improving the education system and the environment
- +0.7 to +1.0 strong positive association between improving the education system and the environment

Based on the determined value (+.319) we have determined that there is a weak positive association between improving the nation’s education system and improving and protecting the environment. There is a weak association that shows that as one becomes more concerned with improving the education system they become more interested in improving and protecting the environment.

Correlations

		IMPROVING NATIONS EDUCATION SYSTEM	IMPROVING & PROTECTING ENVIRONMEN T
IMPROVING NATIONS EDUCATION SYSTEM	Pearson Correlation	1	.319**
	Sig. (2-tailed)		.000
	N	204	195
IMPROVING & PROTECTING ENVIRONMENT	Pearson Correlation	.319**	1
	Sig. (2-tailed)	.000	
	N	195	200

** . Correlation is significant at the 0.01 level (2-tailed).

Appendix

CONFIDENCE IN MAJOR COMPANIES

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A GREAT DEAL	49	11.7	19.0	19.0
	ONLY SOME	174	41.6	67.4	86.4
	HARDLY ANY	35	8.4	13.6	100.0
	Total	258	61.7	100.0	
Missing	NA	15	3.6		
	System	145	34.7		
	Total	160	38.3		
Total		418	100.0		

CONFIDENCE IN ORGANIZED RELIGION

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A GREAT DEAL	61	14.6	23.3	23.3
	ONLY SOME	147	35.2	56.1	79.4
	HARDLY ANY	54	12.9	20.6	100.0
	Total	262	62.7	100.0	
Missing	NA	11	2.6		
	System	145	34.7		
	Total	156	37.3		
Total		418	100.0		

CONFIDENCE IN EDUCATION

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A GREAT DEAL	63	15.1	23.8	23.8
	ONLY SOME	162	38.8	61.1	84.9
	HARDLY ANY	40	9.6	15.1	100.0
	Total	265	63.4	100.0	
Missing	NA	8	1.9		
	System	145	34.7		
	Total	153	36.6		
Total		418	100.0		

CONFID. IN EXEC BRANCH OF FED GOVT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A GREAT DEAL	23	5.5	8.6	8.6
	ONLY SOME	154	36.8	57.9	66.5
	HARDLY ANY	89	21.3	33.5	100.0
	Total	266	63.6	100.0	
Missing	NA	7	1.7		
	System	145	34.7		
	Total	152	36.4		
Total		418	100.0		

CONFIDENCE IN ORGANIZED LABOR

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A GREAT DEAL	23	5.5	9.2	9.2
	ONLY SOME	156	37.3	62.2	71.3
	HARDLY ANY	72	17.2	28.7	100.0
	Total	251	60.0	100.0	
Missing	NA	22	5.3		
	System	145	34.7		
	Total	167	40.0		
Total		418	100.0		

CONFIDENCE IN PRESS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A GREAT DEAL	24	5.7	9.1	9.1
	ONLY SOME	126	30.1	47.9	57.0
	HARDLY ANY	113	27.0	43.0	100.0
	Total	263	62.9	100.0	
Missing	NA	10	2.4		
	System	145	34.7		
	Total	155	37.1		
Total		418	100.0		

CONFIDENCE IN MEDICINE

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A GREAT DEAL	103	24.6	39.2	39.2
	ONLY SOME	137	32.8	52.1	91.3
	HARDLY ANY	23	5.5	8.7	100.0
	Total	263	62.9	100.0	
Missing	NA	10	2.4		
	System	145	34.7		
	Total	155	37.1		
Total		418	100.0		

CONFIDENCE IN TELEVISION

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A GREAT DEAL	28	6.7	10.4	10.4
	ONLY SOME	136	32.5	50.7	61.2
	HARDLY ANY	104	24.9	38.8	100.0
	Total	268	64.1	100.0	
Missing	NA	5	1.2		
	System	145	34.7		
	Total	150	35.9		
Total		418	100.0		

CONFID. IN UNITED STATES SUPREME COURT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A GREAT DEAL	81	19.4	31.3	31.3
	ONLY SOME	142	34.0	54.8	86.1
	HARDLY ANY	36	8.6	13.9	100.0
	Total	259	62.0	100.0	
Missing	NA	14	3.3		
	System	145	34.7		
	Total	159	38.0		
Total		418	100.0		

CONFIDENCE IN SCIENTIFIC COMMUNITY

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A GREAT DEAL	91	21.8	36.7	36.7
	ONLY SOME	140	33.5	56.5	93.1
	HARDLY ANY	17	4.1	6.9	100.0
	Total	248	59.3	100.0	
Missing	NA	25	6.0		
	System	145	34.7		
	Total	170	40.7		
Total		418	100.0		

Descriptives

Descriptive Statistics

	N	Sum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
AGE OF RESPONDENT	416	19059	45.81	16.997	.617	.120
Valid N (listwise)	416					

Descriptive Statistics

	Kurtosis	
	Statistic	Std. Error
AGE OF RESPONDENT	-.508	.239
Valid N (listwise)		

Descriptive Statistics

	N	Sum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
NUMBER WORDS CORRECT IN VOCABULARY TEST	261	1570	6.02	2.099	-.267	.151
Valid N (listwise)	261					

Descriptive Statistics

	Kurtosis	
	Statistic	Std. Error
NUMBER WORDS CORRECT IN VOCABULARY TEST	.006	.300
Valid N (listwise)		

Descriptive Statistics

	N	Sum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
HOURS PER DAY WATCHING TV	414	1198	2.89	2.170	2.434	.120
Valid N (listwise)	414					

Descriptive Statistics

	Kurtosis	
	Statistic	Std. Error
HOURS PER DAY WATCHING TV	11.249	.239
Valid N (listwise)		

Provide the same description by SEX

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
AGE OF RESPONDENT * RESPONDENTS SEX	416	99.5%	2	.5%	418	100.0%

Report

AGE OF RESPONDENT

RESPONDENTS SEX	Mean	N	Std. Deviation	Skewness	Kurtosis
MALE	43.13	171	15.669	.801	.152
_ FEMALE	47.69	245	17.656	.484	-.814
Total	45.81	416	16.997	.617	-.508

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
NUMBER WORDS CORRECT IN VOCABULARY TEST * RESPONDENTS SEX	261	62.4%	157	37.6%	418	100.0%

Report

NUMBER WORDS CORRECT IN VOCABULARY TEST

RESPONDENTS SEX	Mean	N	Std. Deviation	Skewness	Kurtosis
MALE	6.15	114	2.010	-.142	.113
FEMALE	5.91	147	2.167	-.327	-.088
Total	6.02	261	2.099	-.267	.006

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
HOURS PER DAY WATCHING TV * RESPONDENTS SEX	414	99.0%	4	1.0%	418	100.0%

Report

HOURS PER DAY WATCHING TV

RESPONDENTS SEX	Mean	N	Std. Deviation	Skewness	Kurtosis
MALE	2.61	171	1.617	1.466	3.473
FEMALE	3.09	243	2.470	2.398	10.064
Total	2.89	414	2.170	2.434	11.249

Construct a 95% confidence interval about the mean of the variable

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
AGE OF RESPONDENT	416	45.81	16.997	.833
NUMBER WORDS CORRECT IN VOCABULARY TEST	261	6.02	2.099	.130
HOURS PER DAY WATCHING TV	414	2.89	2.170	.107

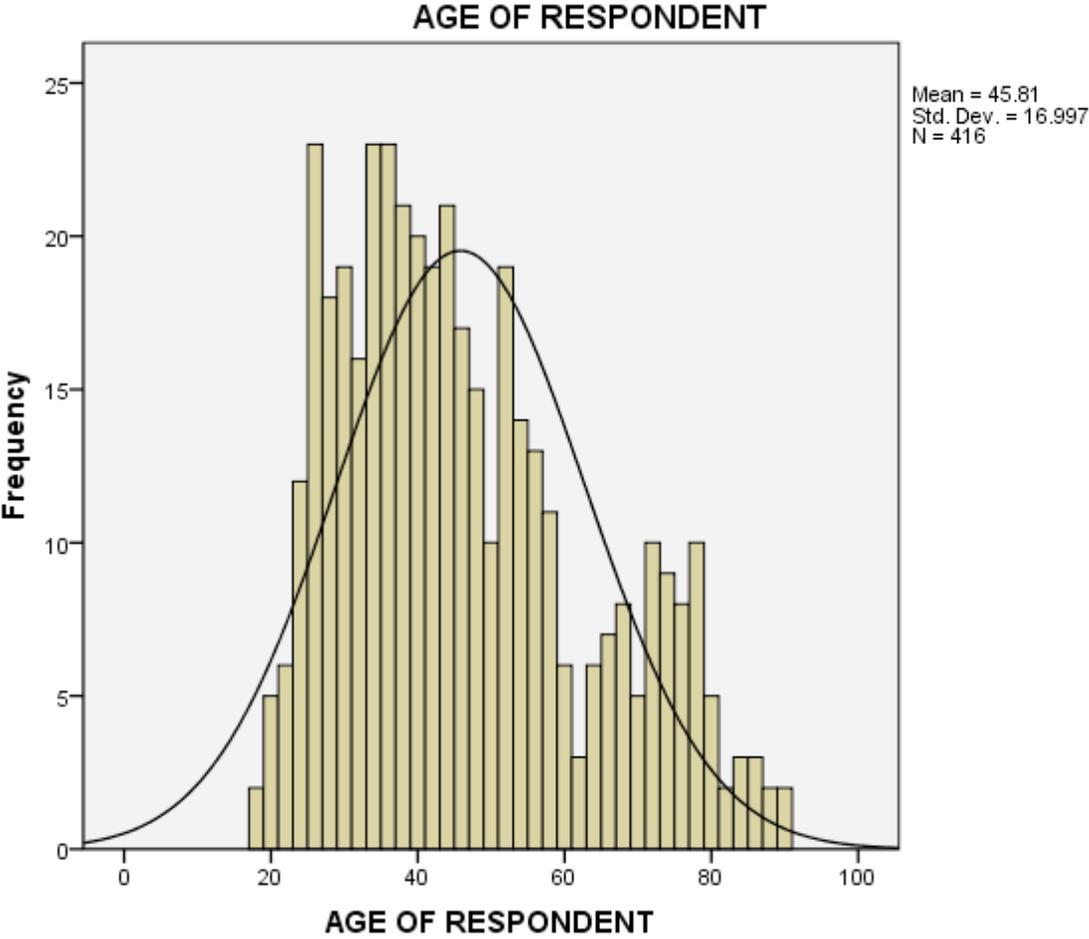
One-Sample Test

	Test Value = 0			
	t	df	Sig. (2-tailed)	Mean Difference
AGE OF RESPONDENT	54.977	415	.000	45.815
NUMBER WORDS CORRECT IN VOCABULARY TEST	46.290	260	.000	6.015
HOURS PER DAY WATCHING TV	27.129	413	.000	2.894

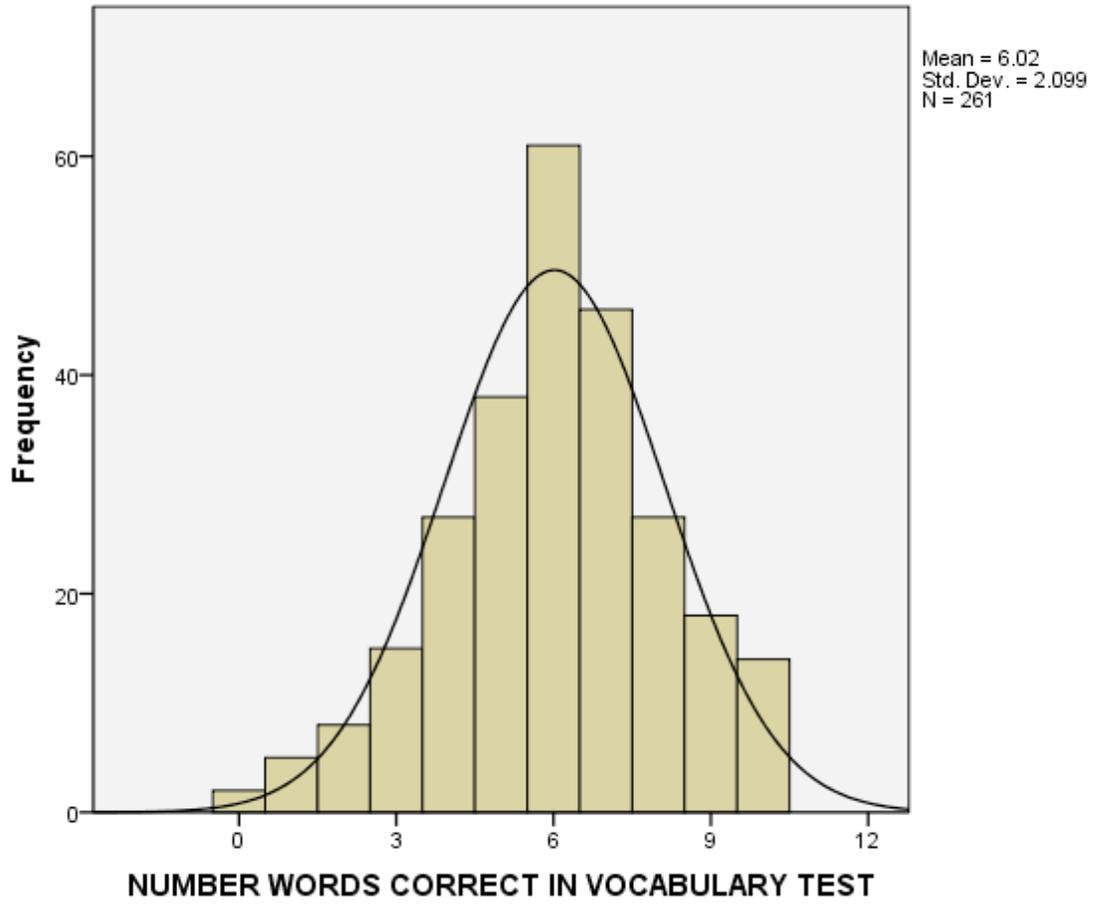
One-Sample Test

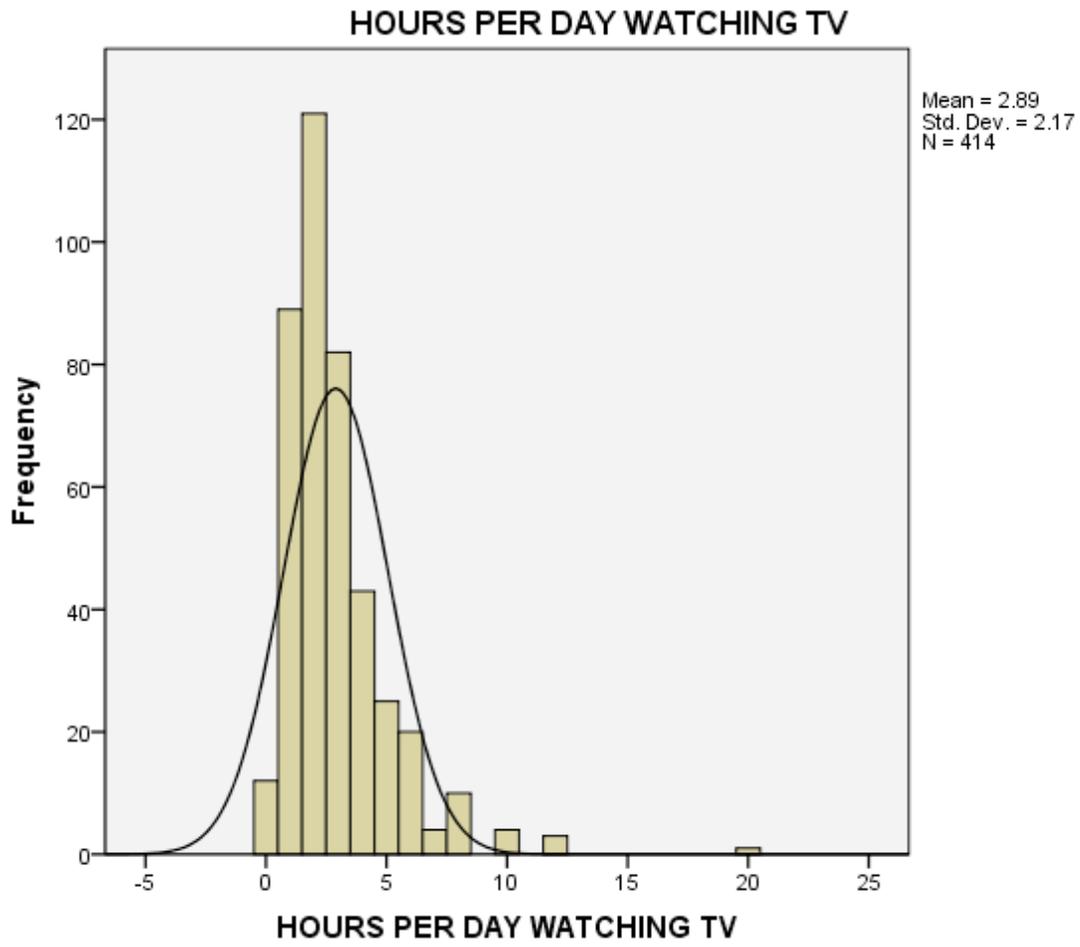
	Test Value = 0	
	95% Confidence Interval of the Difference	
	Lower	Upper
AGE OF RESPONDENT	44.18	47.45
NUMBER WORDS CORRECT IN VOCABULARY TEST	5.76	6.27
HOURS PER DAY WATCHING TV	2.68	3.10

Graphically illustrate the data distribution with a histogram for all subjects



NUMBER WORDS CORRECT IN VOCABULARY TEST





RESPONDENTS SEX

Case Processing Summary

RESPONDENTS SEX		Cases			
		Valid		Missing	
		N	Percent	N	Percent
AGE OF RESPONDENT	MALE	114	66.7%	57	33.3%
	— FEMALE	144	58.3%	103	41.7%
NUMBER WORDS CORRECT IN VOCABULARY TEST	MALE	114	66.7%	57	33.3%
	— FEMALE	144	58.3%	103	41.7%
HOURS PER DAY WATCHING TV	MALE	114	66.7%	57	33.3%
	— FEMALE	144	58.3%	103	41.7%

Case Processing Summary

RESPONDENTS SEX		Cases	
		Total	
		N	Percent
AGE OF RESPONDENT	MALE	171	100.0%
	— FEMALE	247	100.0%
NUMBER WORDS CORRECT IN VOCABULARY TEST	MALE	171	100.0%
	— FEMALE	247	100.0%
HOURS PER DAY WATCHING TV	MALE	171	100.0%
	— FEMALE	247	100.0%

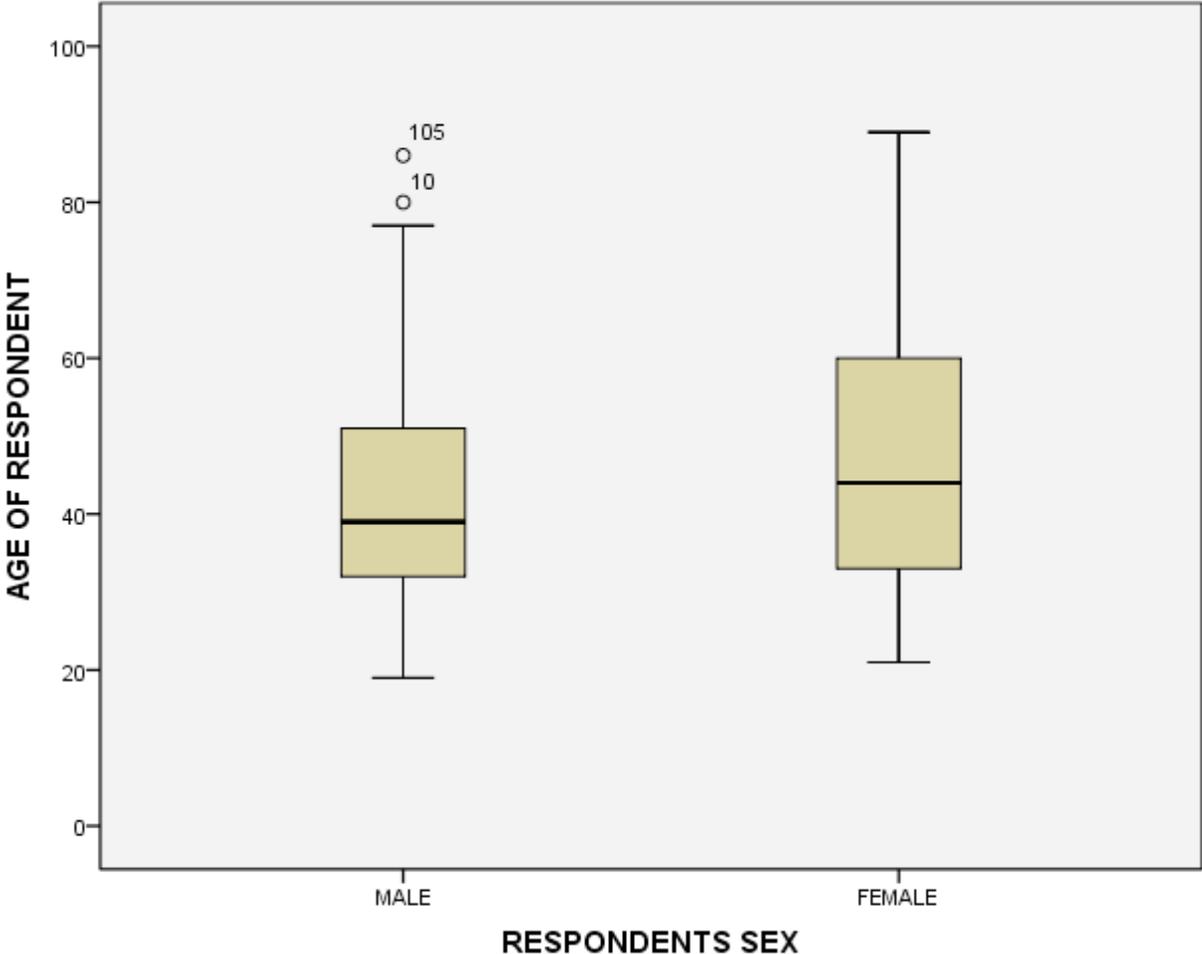
Descriptives

RESPONDENTS SEX			Statistic	Std. Error	
AGE OF RESPONDENT	MALE	Mean	42.26	1.392	
		95% Confidence Interval	Lower Bound	39.50	
		for Mean	Upper Bound	45.02	
		5% Trimmed Mean		41.47	
		Median		39.00	
		Variance		220.992	
		Std. Deviation		14.866	
		Minimum		19	
		Maximum		86	
		Range		67	
		Interquartile Range		20	
		Skewness		.786	.226
		Kurtosis		.073	.449
			FEMALE	Mean	47.42
95% Confidence Interval	Lower Bound			44.52	
for Mean	Upper Bound			50.31	
5% Trimmed Mean				46.89	
Median				44.00	
Variance				308.930	
Std. Deviation				17.576	
Minimum				21	
Maximum				89	
Range				68	
Interquartile Range				27	
Skewness				.480	.202
Kurtosis				-.888	.401
NUMBER WORDS CORRECT IN VOCABULARY TEST	MALE			Mean	6.15
		95% Confidence Interval	Lower Bound	5.78	
		for Mean	Upper Bound	6.52	
		5% Trimmed Mean		6.17	
		Median		6.00	
		Variance		4.040	
		Std. Deviation		2.010	

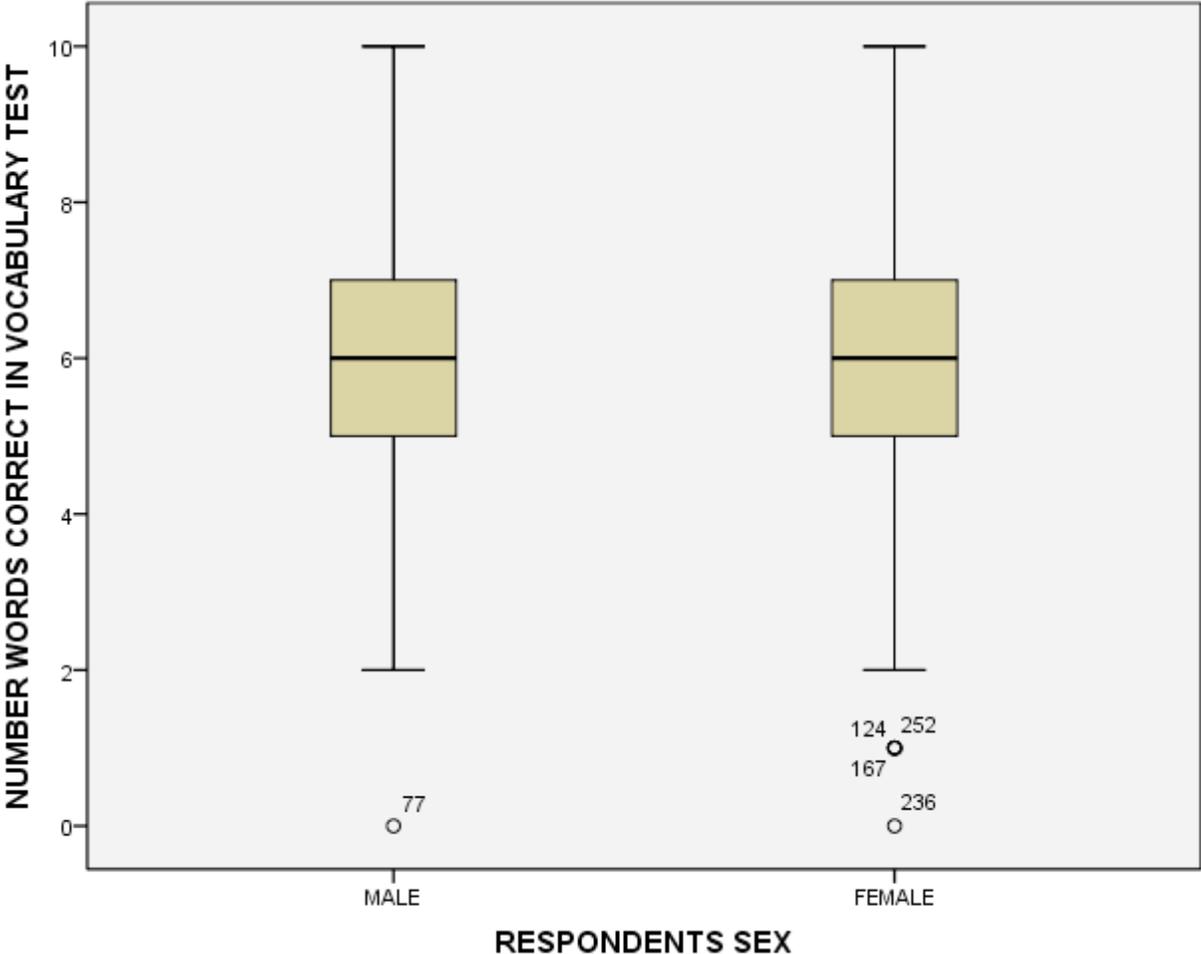
		Minimum		0	
		Maximum		10	
		Range		10	
		Interquartile Range		2	
		Skewness		-.142	.226
		Kurtosis		.113	.449
	FEMALE	Mean		5.88	.181
		95% Confidence Interval	Lower Bound	5.52	
		for Mean	Upper Bound	6.23	
		5% Trimmed Mean		5.92	
		Median		6.00	
		Variance		4.726	
		Std. Deviation		2.174	
		Minimum		0	
		Maximum		10	
		Range		10	
		Interquartile Range		2	
		Skewness		-.293	.202
		Kurtosis		-.095	.401
HOURS PER DAY	MALE	Mean		2.59	.149
WATCHING TV		95% Confidence Interval	Lower Bound	2.29	
		for Mean	Upper Bound	2.88	
		5% Trimmed Mean		2.45	
		Median		2.00	
		Variance		2.545	
		Std. Deviation		1.595	
		Minimum		0	
		Maximum		10	
		Range		10	
		Interquartile Range		1	
		Skewness		1.647	.226
		Kurtosis		4.454	.449
	FEMALE	Mean		2.90	.168
		95% Confidence Interval	Lower Bound	2.56	
		for Mean	Upper Bound	3.23	
		5% Trimmed Mean		2.69	

Median	2.00	
Variance	4.052	
Std. Deviation	2.013	
Minimum	0	
Maximum	12	
Range	12	
Interquartile Range	3	
Skewness	1.600	.202
Kurtosis	3.376	.401

AGE OF RESPONDENT

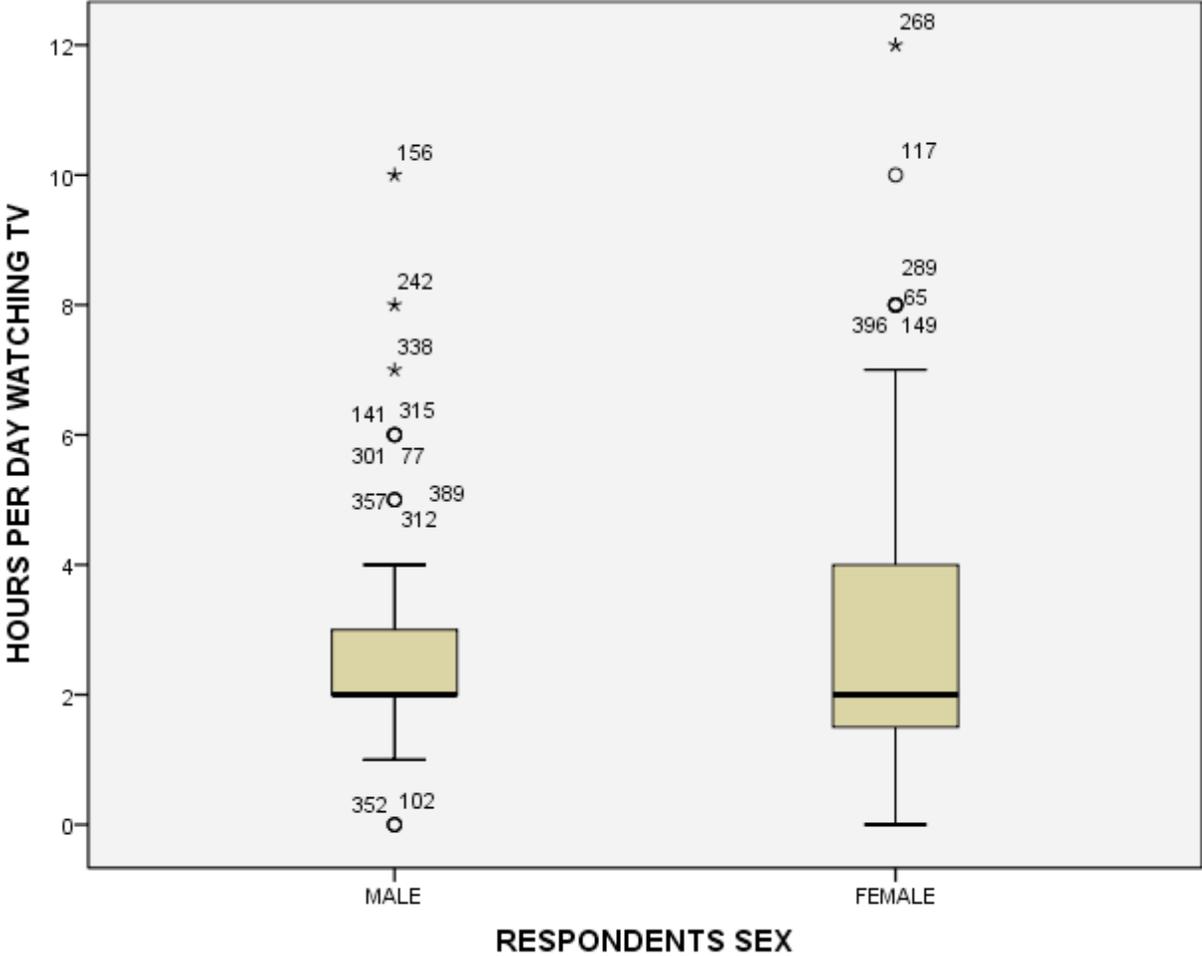


NUMBER WORDS CORRECT IN VOCABULARY TEST



HOURS PER DAY WATCHING TV

Stem-and-Leaf Plots



Group Statistics

RESPONDENTS		N	Mean	Std. Deviation	Std. Error Mean
SEX					
-	MALE	114	42.26	14.866	1.392
	FEMALE	144	47.42	17.576	1.465
-	MALE	114	6.15	2.010	.188
	FEMALE	144	5.88	2.174	.181
-	MALE	114	2.59	1.595	.149
	FEMALE	144	2.90	2.013	.168

Independent Samples Test

		Levene's Test for Equality of Variances				t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
AGE OF RESPONDENT	Equal variances assumed	6.424	.012	-2.501	256	.013	-5.154	2.060	-9.211	-1.096
	Equal variances not assumed			-2.550	254.856	.011	-5.154	2.021	-9.133	-1.174
NUMBER WORDS CORRECT IN VOCABULARY TEST	Equal variances assumed	.678	.411	1.040	256	.299	.274	.264	-.245	.793
	Equal variances not assumed			1.049	249.870	.295	.274	.261	-.240	.789
HOURS PER DAY WATCHING TV	Equal variances assumed	4.484	.035	-1.335	256	.183	-.308	.231	-.762	.146
	Equal variances not assumed			-1.372	255.999	.171	-.308	.225	-.751	.134

