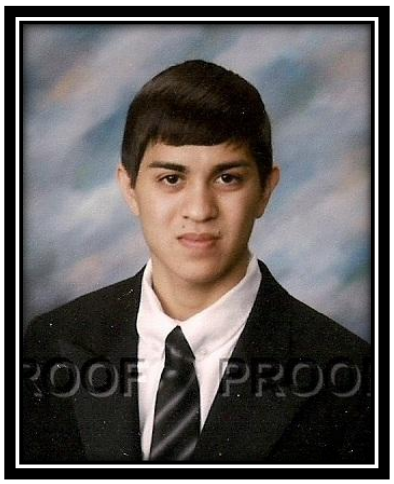
Republican Students

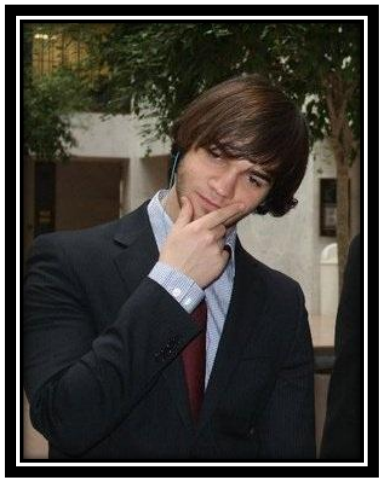
Plan To Earn More

Democratic Students

Our Team: Three students investigated this proposition.



Scott Semel



Tyler Ross

Pierre Maupus

Methods

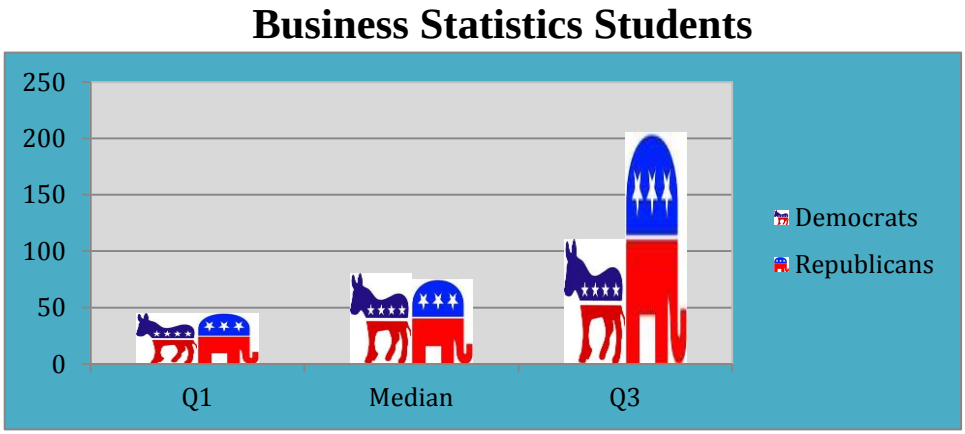
The data was gathered in a student survey administered to all Business Statistics students in the first week of the fall semester. Students completed the surveys at home, and were given assignment credit for their participation. No student was marked down for not answering individual questions. Our team used Excel's single variable data analysis functions and graphic displays to examine the data for patterns and relationships that would be most relevant to assessing the proposition. In the detailed distribution comparisons, hand-written parallel box plots were prepared to meet the remaining assignment requirements.

After completing our initial review of the data, we narrowed our investigations to the following three questions.

- #1 Did republican Business Statistics students project higher earnings in their thirties than democrat Business Statistics students?
- #2 Did republican business statistics students project higher earnings in their forties than democrat business statistics students?
- #3 Did republican business statistics students project higher earnings in their fifties than democrat business statistics students?

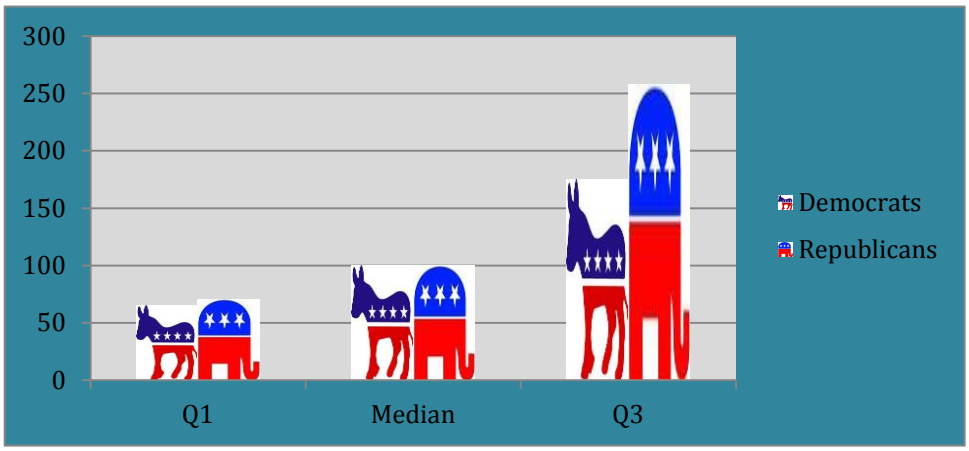
Summary Of Findings

Question #1: Among Business Statistics students, the first quartile value for planned earnings in their thirties reported by democrats was essentially the same as the first quartile value for planned earnings in their thirties reported by republicans. Among Business Statistics students, the median planned earnings in their thirties reported by democrats was $5,000 higher than the median planned earnings in their thirties reported by republicans. Among Business Statistics students, the third quartile value for planned earnings in their thirties reported by republicans was $95,000 higher than the third quartile value for planned earnings in their thirties reported by democrats.



Planned Earnings In Their Thirties ($1,000s)

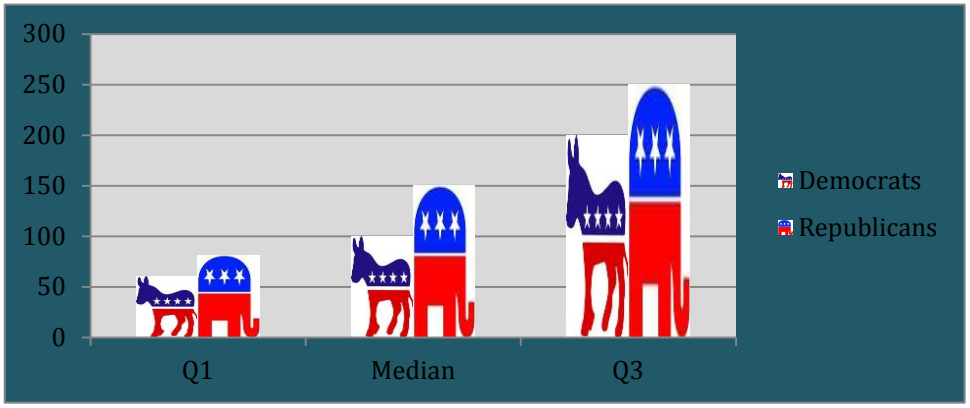
Question #2: Among Business Statistics students, the first quartile value for planned earnings in their forties reported by republicans was $5,000 higher than the first quartile value for planned earnings in their forties reported by democrats. Among Business Statistics students, the median planned earnings in their forties reported by republicans was essentially the same as the median planned earnings in their forties reported by democrats. Among Business Statistics students, the third quartile value planned earnings in their forties reported by republicans was $83,000 higher than the third quartile value for planned earnings in their forties reported by democrats.



Planned Earnings In Their Forties ($1,000s)

Question #3: Among Business Statistics students, the first quartile value for planned earnings in their fifties reported by republicans was $21,000 higher than the first quartile value for planned earnings in their fifties reported by democrats. Among Business Statistics students, the median planned earnings in their fifties reported by republicans was $50,000 higher than the median planned earnings in their fifties reported by democrats. Among Business Statistics students, the third quartile value planned earnings in their fifties reported by republicans was $50,000 higher than the third quartile value for planned earnings in their fifties reported by democrats.

Business Statistics Students



Planned Earnings In Their Fifties ($1,000s)

Conclusion : Based on these specific findings, we conclude the proposition is supported. Republican students do plan to have higher earnings than democratic students.

Detailed Findings :

The report has been organized as follows:

Descriptions of Distributions:

Comparisons of Distributions:

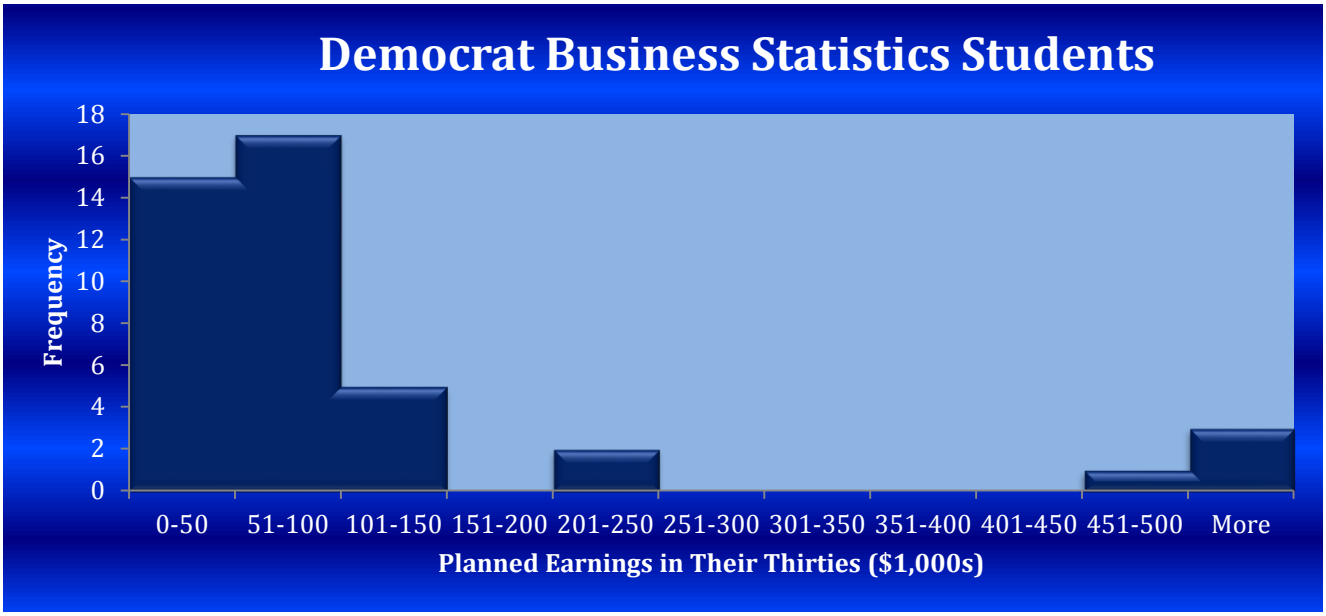
Population: Democratic Business Statistics Students

Variable: Expected Earnings in Their Thirties **Type:** Quantitative, Ratio, Continuous

Shape: A histogram was examined to determine the shape of the distribution. The histogram was displayed using a bin width of 50 ($1,000) increments.

This plot was found to be **uni**modal and highly skew right.

The Fisher skew statistic was 4.1 This statistic fell outside the computed range of **-.7 to .74** indicating that the distribution's shape is highly skew right.



Center: Mean = 175 ($1,000) Median = 80 ($1,000) Mode = **60** ($1,000)

The best measure of central tendency is the median because the distribution is skewed. This skew right shape causes the mean to be greater than the median.

Spread: Range = **1985** ($1,000) IQR = **65** ($1,000) σ = **348.6** ($1,000)

The best measure of spread is the standard deviation because the distribution is skewed.

Outliers: IQR Method: Adding 1.5 times the IQR to the third quartile value of **110** results in an upper outlier threshold of **207.5**($1,000). Subtracting 1.5 times the IQR from the first quartile value of **45** thousand results in a lower outlier threshold of **-52.5**($1,000). Examination of the data found six outliers that exceeded these thresholds, $500,000; $250,000 (2); $2,000,000; $1,000,000; $900,000.

Standard Deviation σ Method: Adding and subtracting three standard deviations from the mean of **175** ($1,000) establishes an upper outlier threshold of **1220.8** ($1,000) and a lower threshold of **-870.8** ($1,000). Examination of the data found one outlier that exceeded these thresholds, $2,000,000.

The best measure of outliers is the IQR Method because the distribution is skewed.

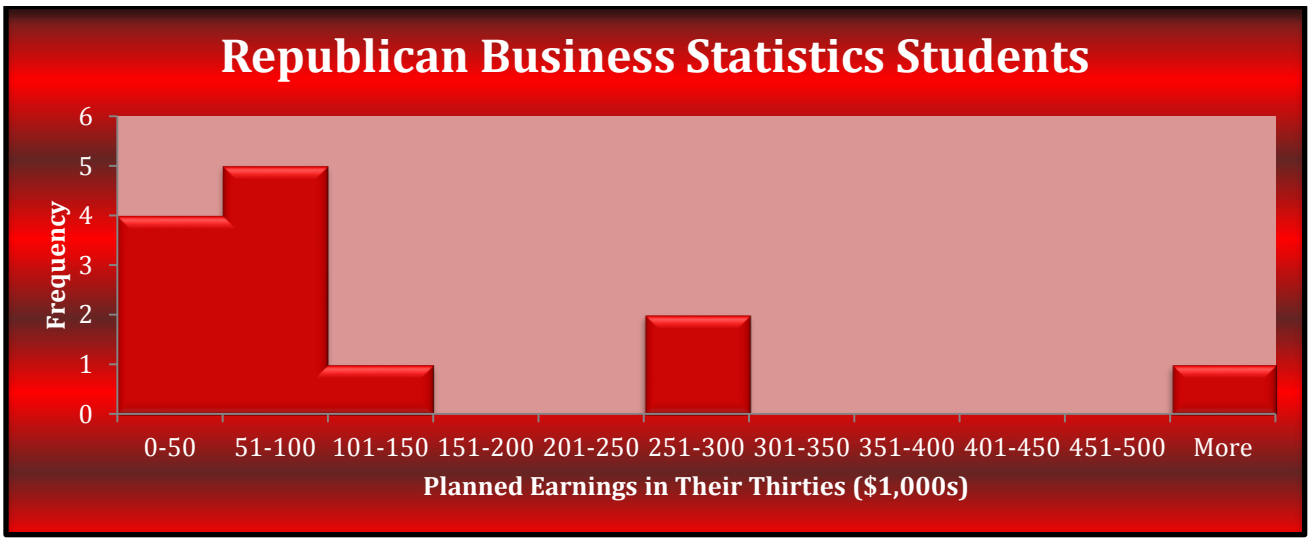
Population: Republican Business Statistics Students

Variable: Expected Earnings in Their Thirties **Type:** Quantitative, Ratio, Continuous

Shape: A histogram was examined to determine the shape of the distribution. The histogram was displayed using a bin width of 50(1,000s) increments.

This plot was found to be **uni**modal and highly skew right.

The Fisher skew statistic was 3.4. This statistic fell outside the computed range of **-1.4 to 1.35** indicating that the distribution's shape is highly skew right.



Center: Mean = 250.7 ($1,000) Median = 75 ($1,000) Mode = **75** ($1,000)

The best measure of central tendency is the median because the distribution is skewed. This skew right shape causes the mean to be greater than the median.

Spread: Range = **1970** ($1,000) IQR = **160** ($1,000) $\sigma =$ **533.3** ($1,000)

The best measure of spread is the standard deviation because the distribution is skewed.

Outliers: IQR Method: Adding 1.5 times the IQR to the third quartile value of **205** ($1,000) results in an upper outlier threshold of 445 ($1,000). Subtracting 1.5 times the IQR from the first quartile value of **45** ($1,000) results in a lower outlier threshold of -195 ($1,000) . Examination of the data found four outliers that exceeded these thresholds, $500,000; $2,000,000; $1,000,000; $900,000.

Standard Deviation σ Method: Adding and subtracting three standard deviations from the mean of **250.7** ($1,000) establishes an upper outlier threshold of **1850.6** ($1,000) and a lower threshold of **-1349.2** ($1,000). Examination of the data found one outlier that exceeded these thresholds, $2,000,000.

The best measure of outliers is the IQR Method because the distribution is skewed.

Population: Democratic Business Statistics Students

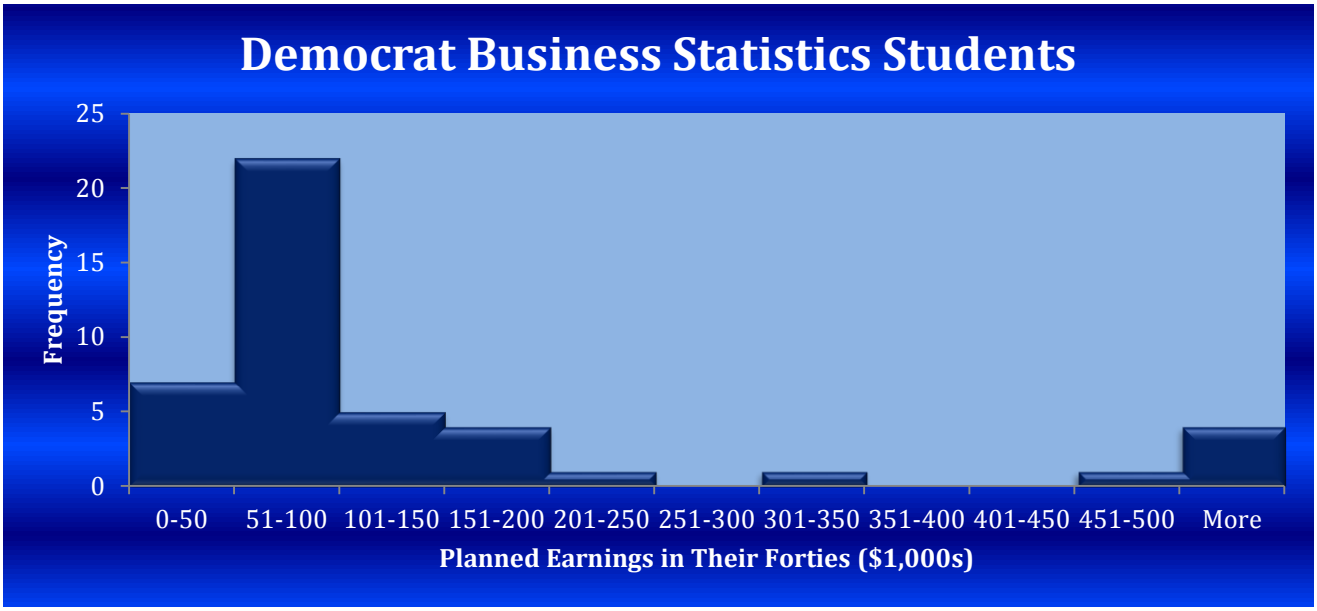
Variable: Planned Earnings in Forties **Type:** Quantitative, Ratio, Continuous

Shape:

A histogram was examined to determine the shape of the distribution. The histogram was displayed using a bin width of **50** ($1,000s) increments.

This plot was found to be **uni**modal and highly skew right.

The Fisher skew statistic was **3.17**. This statistic fell outside the computed range of **-0.73 to +0.73** indicating that the distribution's shape is highly skew right.



Center:

Mean = **221** ($1,000s), Median = **100** ($1,000s), Mode = **100** ($1,000s)

The best measure of central tendency is the median because the distribution is skewed. This skew right shape causes the mean to be greater than the median.

Spread:

Range = **1680** ($1,000s), IQR = **110** ($1,000s), $\sigma =$ **381** ($1,000s)

The best measure of spread is the range and interquartile range because the distribution is skewed.

Outliers:

IQR Method: Adding 1.5 times the IQR to the third quartile value of **175**($1,000s) results in an upper outlier threshold of **340** ($1,000s). Subtracting 1.5 times the IQR from the first quartile value of **65** ($1,000s) results in a lower outlier threshold of **-100** ($1,000s). Examination of the data found six outliers that exceeded these thresholds, $350,000; $1,500,000; $600,000; $500,000; $1,500,000; $1,700.

Standard Deviation σ Method: Adding and subtracting three standard deviations from the mean of **221** ($1,000s) establishes an upper outlier threshold of **1364** ($1,000s) and a lower threshold of **-922** ($1,000s). Examination of the data found three outliers that exceeded these thresholds, $1500,000;, $1,500,000; $1,700,000.

The best measure of outliers is the IQR Method because the distribution is skewed.

Population: Republican Business Statistics Students

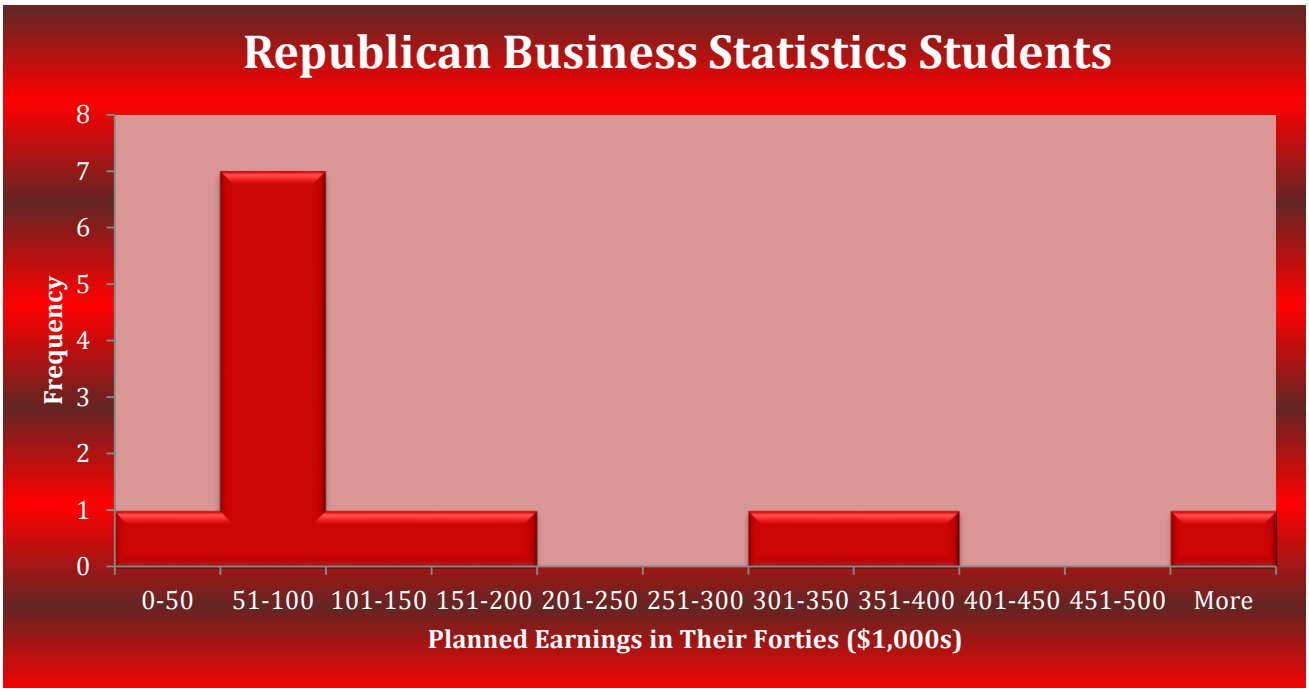
Variable: Planned Earnings in Forties **Type:** Quantitative, Ratio, Continuous

Shape:

A histogram was examined to determine the shape of the distribution. The histogram was displayed using a bin width of **50** ($1,000s) increments.

All of these plots were found to be **uni**modal and highly skew right.

The Fisher skew statistic was **3.6**. This statistic fell outside the computed range of **-1.36 to +1.36** indicating that the distribution's shape is highly skew right.



Center:

Mean = **2437** ($1,000s), Median = **100** ($1,000s), Mode = **70** ($1,000s)

The best measure of central tendency is the median because the distribution is skewed. This skew right shape causes the mean to be greater than the median.

Spread:

Range = **29960** ($1,000s), IQR = **188** ($1,000s), $\sigma =$ **8282** ($1,000s)

The best measure of spread is the range and interquartile range because the distribution is skewed.

Outliers:

IQR Method: Adding 1.5 times the IQR to the third quartile value of **258** ($1,000s) results in an upper outlier threshold of **539** ($1,000s). Subtracting 1.5 times the IQR from the first quartile value of **70** ($1,000s) results in a lower outlier threshold of **-211** ($1,000s). Examination of the data found one outlier that exceeded these thresholds, $3,000,000.

Standard Deviation (σ) Method: Adding and subtracting three standard deviations from the mean of **2437** ($1,000s) establishes an upper outlier threshold of **27283** ($1,000s) and a lower threshold of **-22409** ($1,000s). Examination of the data found one outlier that exceeded these thresholds, $3,000,000.

The best measure of outliers is the IQR Method because the distribution is skewed.

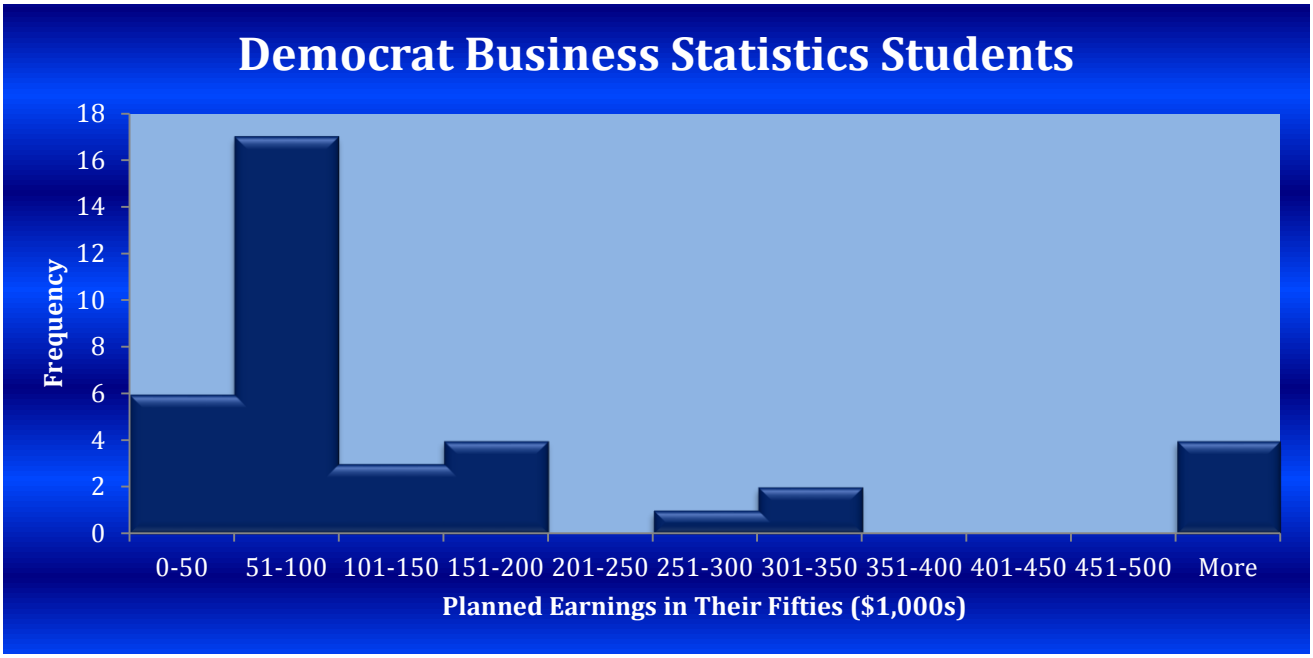
Population: Democratic Business Statistics Students

Variable: Planned Earnings in Fifties **Type:** Quantitative, Ratio, Continuous

Shape: A histogram was examined to determine the shape of the distribution. The histogram was displayed using a bin width of 50 ($1000's) increments.

This plot was found to be **uni**modal and highly skew right.

The Fisher skew statistic was **4.0**. This statistic fell outside the computed range of **-.805 to +.805** indicating that the distribution's shape is highly skew right.



Center: Mean = **273.7** ($1000's), Median = **100** ($1000's), Mode = **100** ($1000's)

The best measure of central tendency is the median because the distribution is symmetric. This skew right shape causes the mean to be greater than the median.

Spread: Range = **2960** ($1000's), IQR = **140** ($1000's), $\sigma =$ **557.7** ($1000's)

The best measure of spread is the range and interquartile range because the distribution is skewed.

Outliers: IQR Method: Adding 1.5 times the IQR to the third quartile value of **200** ($1000's) results in an upper outlier threshold of **410** ($1000's). Subtracting 1.5 times the IQR from the first quartile value of **60** ($1000's) results in a lower outlier threshold of **-150** ($1000's). Examination of the data found four outliers that exceeded these thresholds, $1,700,000; $1,000,000; $650,000; $3,000,000.

Standard Deviation (σ) Method: Adding and subtracting three standard deviations from the mean of **273.7** ($1000's) establishes an upper outlier threshold of **1946.8** ($1000's) and a lower threshold of **-1399.4** ($1000's). Examination of the data found one outliers that exceeded these thresholds, $3,000,000.

The best measure of outliers is the IQR Method because the distribution is skewed.

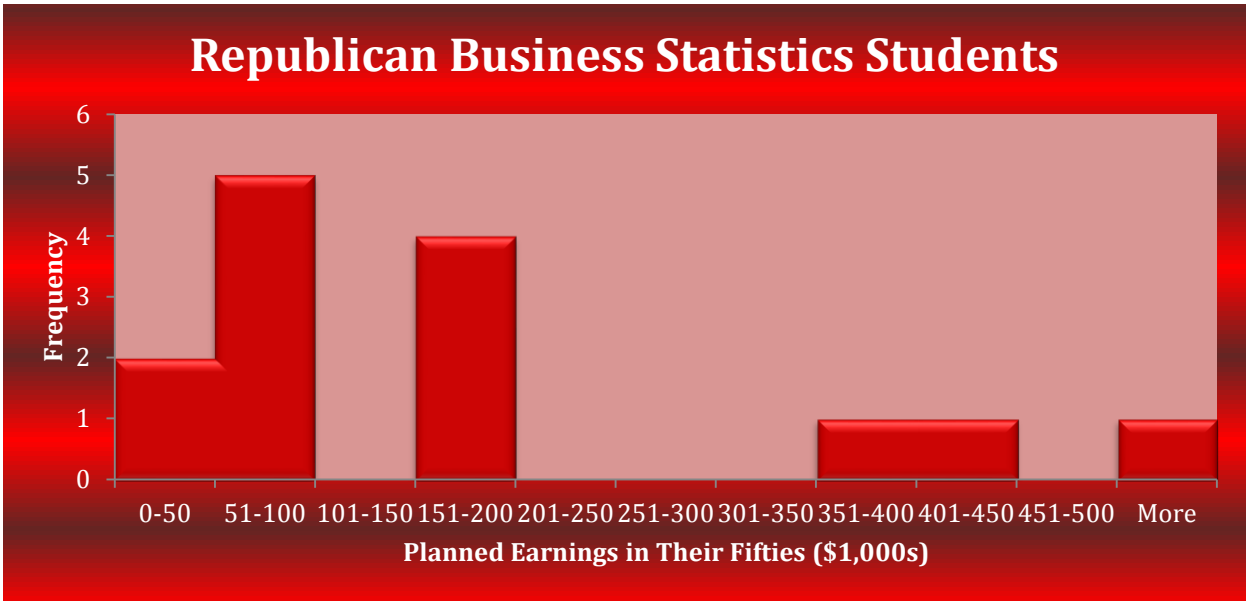
Population: Republican Business Statistics Students

Variable: Planned Earnings in Fifties **Type:** Quantitative, Ratio, Continuous

Shape: A histogram was examined to determine the shape of the distribution. The histogram was displayed using a bin width of 50 ($1000's) increments.

This plot was found to be **uni**modal and highly skew right.

The Fisher skew statistic was **3.8**. This statistic fell outside the computed range of **-1.3 to +1.3** indicating that the distribution's shape is highly skew right.



Center: Mean = **71585** ($1000's), Median = **150** ($1000's), Mode = **200** ($1000's)

The best measure of central tendency is the median because the distribution is symmetric. This skew right shape causes the mean to be greater than the median.

Spread: Range = **999950** ($1000's), IQR = **168.75** ($1000's), $\sigma =$ **267216.2** ($1000's)

The best measure of spread is the range and interquartile range because the distribution is skewed.

Outliers: IQR Method: Adding 1.5 times the IQR to the third quartile value of **250** ($1000's) results in an upper outlier threshold of **503** ($1000's). Subtracting 1.5 times the IQR from the first quartile value of **81.25** ($1000's) results in a lower outlier threshold of **-172** ($1000's). Examination of the data found three outliers that exceeded these thresholds, $1,000,000; $450,000; $400,000.

Standard Deviation σ Method: Adding and subtracting three standard deviations from the mean of **71585** ($1000's) establishes an upper outlier threshold of **873233.6** ($1000's) and a lower threshold of **-730063.6** ($1000's). Examination of the data found one outliers that exceeded these thresholds, $1,000,000.

The best measure of outliers is the IQR Method because the distribution is skewed.

Variable: Expected Earnings in Their Thirties in the population of Democrat Business Statistics Students

Variable: Thousands of Dollars in the population of Republicans Expected Earnings in Their Thirties

	Democrat	Republican	Comparison
Shape	Highly skew right	Highly skew right	The distributions have the same shape.
Center	Mean = 175 ($1,000) Median = 80 ($1,000)	Mean = 250.7 ($1,000) Median = 75 ($1,000)	Since both distributions are skewed, the best measure for comparing central tendencies is the median. The center of the distribution for Democrats planned earnings in their thirties is about five units higher than the distribution for Republicans planned earnings in their thirties
Spread	Range = 1985 ($1,000) IQR = 65 ($1,000) σ = 348.6 ($1,000)	Range = 1970 ($1,000) IQR = 160 ($1,000) σ = 533.3 ($1,000)	Since both distributions are skewed, the best measure for comparing spread are the range and interquartile range. Examination of these statistics shows the distribution for Republicans Expected Earnings in Their Thirties has more spread than the distribution for Democrats Expected Earnings in Their Thirties.
Outliers	$2,000,000 using the Standard Deviation Method	$2,000,000 using the Standard Deviation Method	The distribution for Democrats Expected Earnings in Their Thirties has 1 outlier, while the distribution for Republicans Expected Earnings in Their Thirties has one outlier.

Variable: Planned Earnings in Forties in the population of Democratic Business Statistics students.
Variable: Planned Earnings in Forties in the population of Republican Business Statistics students.

	Democrats	Republicans	Comparison
Shape	Highly Skew Right	Highly Skew Right	The distributions have the same shape.
Center	Mean = **221** ($1,000s) Median = **100** ($1,000s)	Mean = **2437** ($1,000s) Median = **100** ($1,000s)	Since both distributions are skewed, the best measure for comparing central tendencies is the median. The center of the distribution for Planned Earnings in Forties of Democrats is nearly the same as the distribution for Planned Earnings in Forties of Republicans.
Spread	Range = **1680** ($1,000s) IQR = **110** ($1,000s) σ = **381** ($1,000s)	Range = **29960** ($1,000s) IQR = **187.5** ($1,000s) σ = **8282** ($1,000s)	Since both distributions are skewed, the best measure for comparing spread are the range and interquartile range. Examination of these statistics shows the distribution for Planned Earnings in Forties of Republicans has more spread than the distribution for Planned Earnings in Forties of Democrats.
Outliers	350, 1500, 600, 500, 1500, 1700 ($1,000s) using the IQR Method	30000 ($1,000s) using the IQR Method	The distribution for Planned Earnings in Forties of Democrats has **six** outliers, while the distribution for Planned Earnings in Forties of Republicans has **one** outlier.

Variable: Planned Earnings in Fifties in the population of Democratic Business Statistics Students.
Variable: Earnings in Fifties in the population of Republican Business Statistics.

	Democrats	Republicans	Comparison
Shape	Highly skew right	Highly skew right	The distributions have the same shape.
Center	Mean = **273.7** ($1,000s) Median = **100** ($1,000s)	Mean = **71585** ($1,000s) Median = **150** ($1,000s)	Since both distributions are skewed, the best measure for comparing central tendencies is the median. The center of the distribution for Planned Earnings in Fifties of Republicans is about 50 units higher than the distribution for Planned Earnings in Fifties of Democrats.
Spread	Range = 2960($1,000) IQR = 140 ($1,000) σ = 557.7 ($1,000)	Range = 999950 ($1,000) IQR = 168.75 ($1,000) σ = 267216.2 ($1,000)	Since both distributions are skewed, the best measure for comparing spread are the range and interquartile range. Examination of these statistics shows the distribution for Planned Earnings in Fifties of Republicans has more spread than the distribution for Planned Earnings in Fifties of Democrats.
Outliers	1,700; 1,000; 650; 3,000 ($1,000) using the IQR Method	1,000,000; 450; 400 ($1,000) using the IQR Method	The distribution for Planned Earnings in Fifties of Democrats has **four** outliers, while the distribution for Planned Earnings in Fifties of Republicans has **three** outliers.