Research Question: Which age of veterans are most likely to be homeless?

Hypothesis:

For this research, we can either find that there is no relationship between homelessness and age, or we can find that there is in fact a relationship. As such, we will have a null hypothesis and an alternative hypothesis that correspond to these two ideas.

H0 = There is no relationship between age and homelessness.

H1 = There is a relationship between age and homelessness.

Data:

Numerical summaries of Age of Homeless Veterans

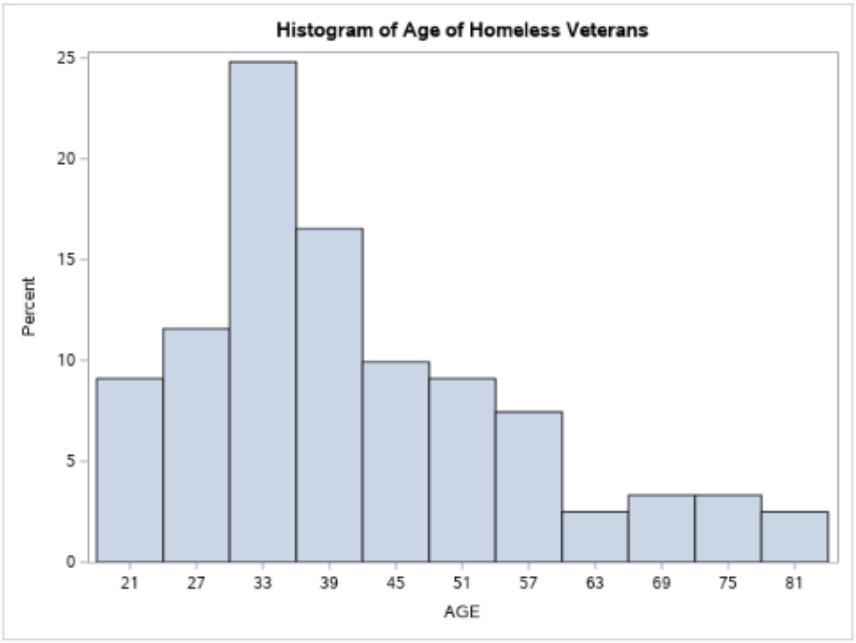
The MEANS Procedure

Analysis Variable : AGE								
N	Mean	Mode	Std Dev	Minimum	Median	Maximum	Skewness	Kurtosis
242	40.9	33.0	14.8	20.0	37.0	79.0	0.9	0.2

Frequencies of Age of Homeless Veterans

The FREQ Procedure

AGE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
20	4	1.65	4	1.65
21	4	1.65	8	3.31
22	8	3.31	16	6.61
23	6	2.48	22	9.09
24	4	1.65	26	10.74
25	8	3.31	34	14.05
26	2	0.83	36	14.88
27	2	0.83	38	15.70
28	6	2.48	44	18.18
29	6	2.48	50	20.66
30	14	5.79	64	26.45
31	6	2.48	70	28.93
32	8	3.31	78	32.23
33	16	6.61	94	38.84



Using the numeral summaries table, we can see that the mode for the variable AGE, or the most common value for a veteran's age, is 33. The included section of the frequency table shows us that this value occurs 16 times out of a valid N of 242, which means that homeless 33 year old veterans accounted for a total of 6.61% of our dataset. Our dataset has an age range of 59, with a minimum age of 20 and a maximum age of 79. The average age of our data is 40.9, and the standard deviation of our dataset is 14.8. Judging by these values, we can conclude that the majority of our data falls within one standard deviation on either side of the mean, which ranges from 26.1 to 55.7.

The numerical summaries table also shows us the skewness of the age distribution in our data, which is 0.9. This means that there is a positive skew and the tail of the histogram will be on the right side of the chart. From the histogram, we can see that there is a tail on the right side, but the majority of the data points are concentrated around the mean, which is 40.9 according to the numeric summaries table. Overall, this demonstrates that while there are homeless veterans higher on the age scale, the majority of them within our dataset fall within the 30-40 years old range.

Conclusion:

Overall, my data shows that we can reject our null hypothesis and accept our alternative hypothesis that there is a relationship between age and homelessness within our dataset. Specifically, I found that the most frequently appearing age within our dataset is 33, and a histogram shows that the distribution of age is positively skewed, meaning that the majority of homeless veterans are younger rather than older. The variability of age within our dataset also supports this idea, as the majority of data falls within one standard deviation of the means, which is 40.9, which is nearly half the size of the maximum age included in the dataset.