Deliverable 2:

Have Truck Prices Changed from 2004 to 2019?

Rebekah Sander

Research Question

Is the center of truck prices in 2004 different from $49,543?

Research Variable

WHO	WHAT measurement is made on each		TYPE OF MEASURE
	Name of Variable	Question Asked	
Trucks in 2004	Truck Price	How much does the truck cost?	Quantitative Variable Unit: dollars ($)
One Quantitative Variable being tested against an average truck price of $49,543.			
1.One Mean t Test 2.Sign Test			

SAS Code: Naming and Normality

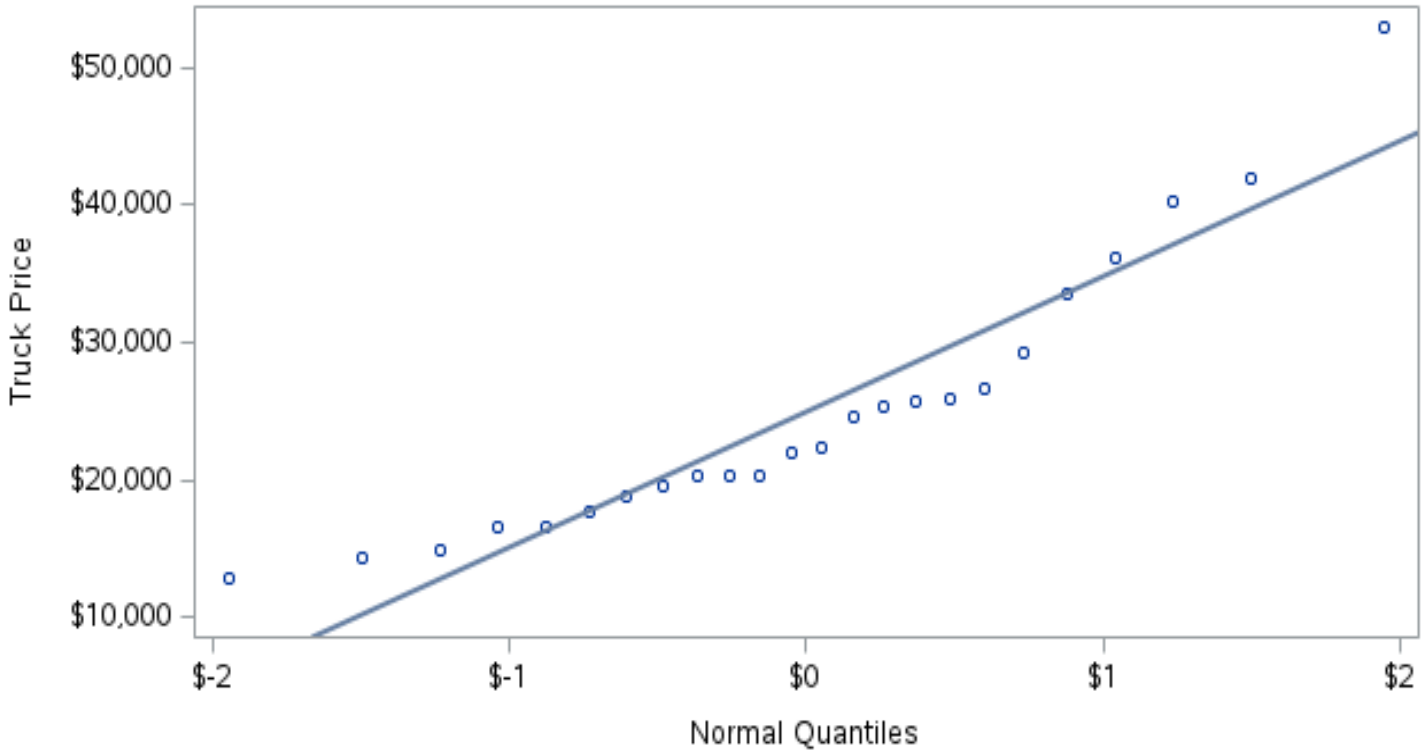
```
/*rename the variable.*/
data work.trucks;
    set work.trucks;
    rename MSRP='Truck Price'n;
run;
```

```
/*Check for the parametric assumptions*/
title "Figure 1: QQ Plot of Truck Price";
proc univariate data=work.trucks normaltest plots ;
var 'Truck Price'n;
title;
```

Assessing Normality

- Normality Tests: All 4 normality tests resulted in p-values < α=0.05, suggesting that we may not assume normality.
- QQ Plot: The data shows large amounts of deviation from the agreement line, contradicting normality. Further, this agrees with the normality tests that say that the data is not normal.

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.888421	Pr < W	0.0123
Kolmogorov-Smirnov	D	0.181295	Pr > D	0.0405
Cramer-von Mises	W-Sq	0.156423	Pr > W-Sq	0.0192
Anderson-Darling	A-Sq	0.908315	Pr > A-Sq	0.0188



Choosing Hypothesis Test: Sign Test

- Since the data is not normal, we will perform the sign test.

$$H_0: \eta = 49{,}543$$
$$H_A: \eta \neq 49{,}543$$
$$\alpha = 0.05$$

- The null hypothesis is that the price of trucks in 2004 is centered at \$49,543.
- The alternative hypothesis is that the price of trucks in 2004 is not centered at \$49,543.
- The level of significance, $\alpha = 0.05$, tells us that 5% of the time we will conclude $\eta \neq \$49{,}543$ when $\eta = \$49{,}543$ is actually true.

Performing Sign Test

- p-value: There is less than a 0.01% chance of getting a random sample of trucks in 2004 to have 1 truck above $49,543 and 23 trucks below $49,543, when the true median is $49,543.
- Conclusion: p-value = 0.001 < α = 0.05, we have evidence to suggest that the median price of trucks in 2004 is different from $49,543, because the sample results of 1 truck with a cost above and 23 trucks with a cost below $49,543 is not likely to happen when the true median cost of trucks is $49,543.

Tests for Location: Mu0=49543				
Test	Statistic		p Value	
Student's t	t	-12.2086	Pr > |t|	<.0001
Sign	M	-11	Pr >= |M|	<.0001
Signed Rank	S	-149	Pr >= |S|	<.0001

Location Counts: Mu0=49543.00	
Count	Value
Num Obs > Mu0	1
Num Obs ^= Mu0	24
Num Obs < Mu0	23

Post hoc test: Confidence Interval

- We are 95% confident that the true population median cost of all trucks in 2004 is between $19,479 and $26,650.
- We can conclude that $49,543 is different than the true median cost of trucks in 2004 from this confidence interval. In fact, $49,543 is too high as $49,543 is above all values contained in this confidence interval.
- The confidence interval agrees with the sign test.

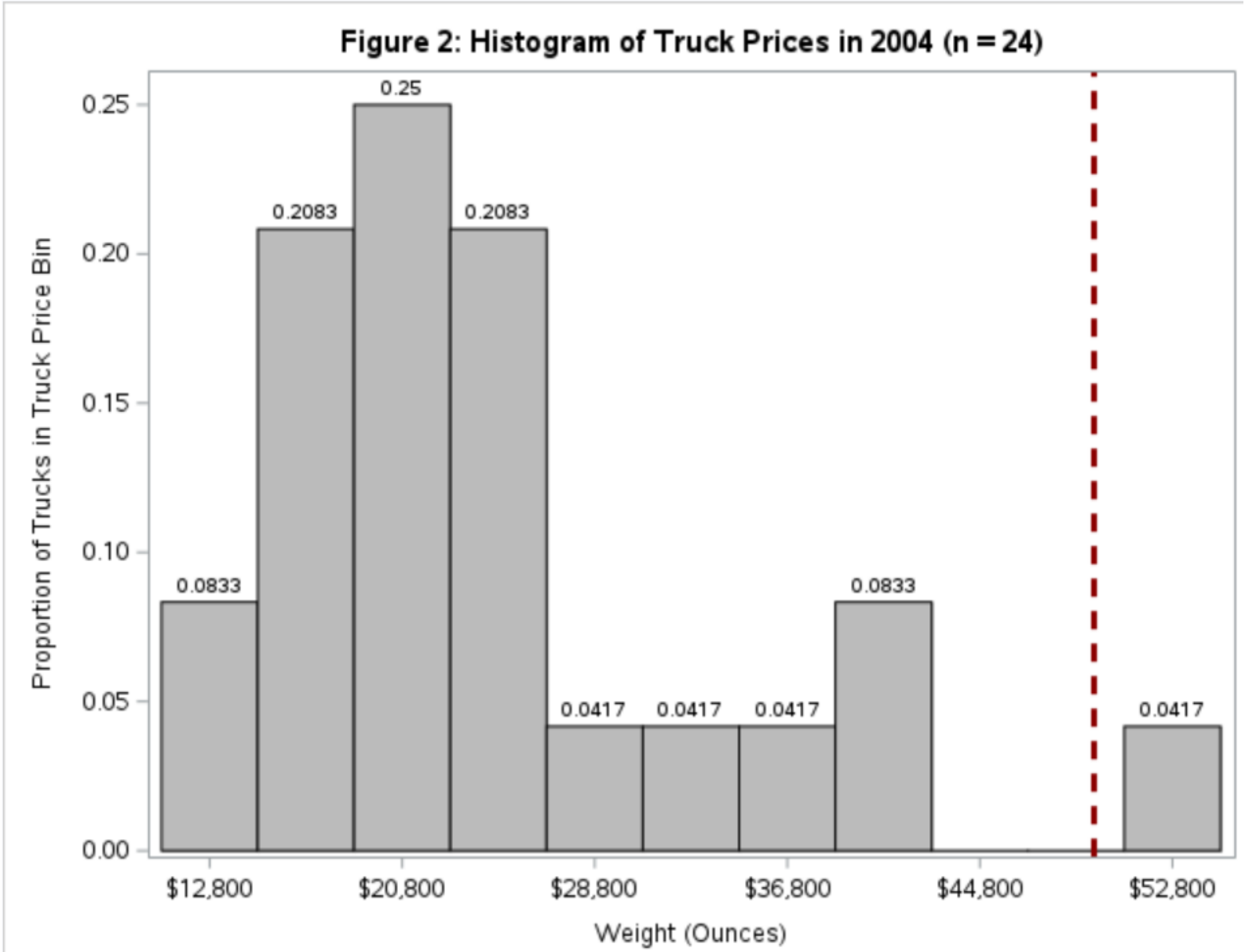
Quantiles (Definition 5)						
				Order Statistics		
Level	Quantile	95% Confidence Limits Distribution Free		LCL Rank	UCL Rank	Coverage
100% Max	52975					
99%	52975	.	.	.	.	.
95%	41995	40340	52975	22	24	59.21
90%	40340	33540	52975	20	24	83.52
75% Q3	27986	24520	41995	14	23	96.96
50% Median	22180	19479	26650	8	18	95.67
25% Q1	18195	14385	20310	2	11	96.96
10%	14840	12800	16530	1	5	83.52
5%	14385	12800	14840	1	3	59.21
1%	12800	.	.	.	.	.
0% Min	12800					

SAS Code: Testing and Histogram

```
/*    The sign test*/
Proc Univariate data=WORK.trucks
     mu0=49543
     loccount
     cipctldf
     alpha=.05;
var 'Truck Price'n;
Run;
```

```
TITLE1 "Figure 2: Histogram of Truck Prices in 2004 (n = 24)";
Proc SGPLOT data=work.trucks;
histogram 'Truck Price'n/
     datalabel=proportion
     scale=proportion showbins
     binwidth=4000 binstart=12800
     fillattrs=(color=H000BB00)
     boundary=upper;
     refline 49543 / axis=x lineattrs=(thickness=3
                                       color=darkred
                                       pattern=dash);
     xaxis label='Weight (Ounces)' labelattrs=(size=10);
     yaxis label='Proportion of Trucks in Truck Price Bin'
     labelattrs=(size=10);
run;
TITLE;
```

Supporting Graphic: Histogram



Taking Action

- From this analysis, we are able to gain insight when it comes to knowing how much a truck is worth.
- Suppose you are in the market to buy a truck new or used. Knowing recent prices and that trucks are costly, let's say $49,943, you may be more willing to buy a used truck from 2004. If a new truck back in 2004 cost lower than $49,543, you can only imagine how low it may cost now!
- The analysis tells us we are 95% confident that the true median truck price of a truck in 2004 is different than $49,543. Therefore, we should not settle on a used car from 2004 that does not reflect a better price.

SAS Code Screen Recording

SAS® Studio

SAS Program

SANDER Deliverable 2 SAS CODE.sas

CODE LOG RESULTS OUTPUT DATA

```
1  /*Research Question: Is the center of truck prices in 2004 different from $49,543?*/
2  /*Unit of observation: Trucks in 2004*/
3  /*Quantitative Variable: MSRP (Manufacturer Suggested Retail Price)*/
4
5
6
7  /*Determine hypothesis testing options for answering the question.*/
8  /*    Quantitative against a known value so either One Mean t Test (if normal) or
9        sign test (not normal)*/
10
11
12
13 /*Import the data set. Make a new data set. Keep ONLY the variables you are testing*/
14 data work.trucks;
15 set sashelp.cars (keep = MSRP type);
16 where type = "Truck";
17 run;
18
19
20
21 /*Check for miscoding (typos in the entries) and whether the variable types are coming
22   in as the correct variable type (quantitative or categorical).*/
23 /*      FOUND IN THE OUTPUT DATA and is the right type, numeric:)*/
24
25
```

/home/u62685438/sasuser.v94/stat3130/deliverables/SANDER Deliverable 2 SAS CODE.sas