Week 1 HW - SPSS

11 avqust 2022, cümə axşamı 15:23

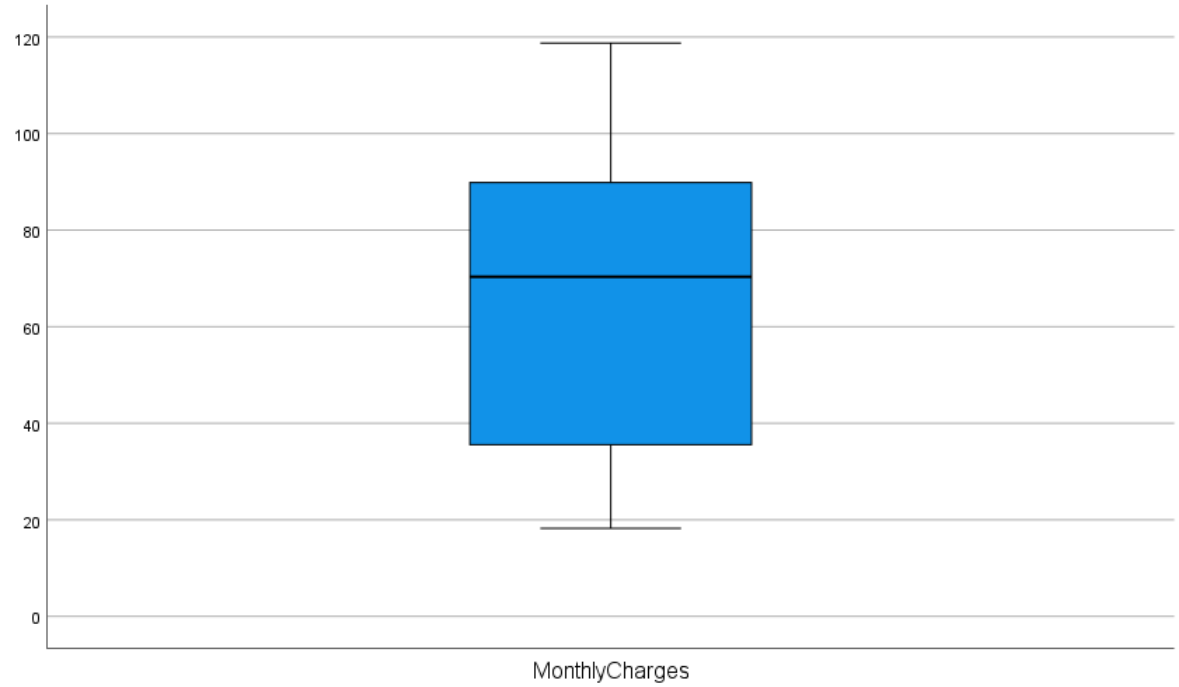
Question 3.

First we calculate the interquartile range and multiplying the value by 1.5. Then by adding the value to 75% percentile and subtracting from 25% percentile, we get upper and lower outlier range. Let's first take a look at **MonthlyCharges** variable:

Percentiles

		Percentiles						
		5	10	25	50	75	90	95
Weighted Average (Definition 1)	MonthlyCharges	19.6500	20.0500	35.5000	70.3500	89.8500	102.6300	107.4400
Tukey's Hinges	MonthlyCharges			35.5000	70.3500	89.8500		

Inter Quartile Range = 89.85 - 35.5 = 54.35
Outlier Ranges = 54.35 * 1.5 = 81.525
Upper Outlier Range = 89.85 + 81.525 = 171.375
Lower Outlier Range = 35.5 - 81.525 = -46.025

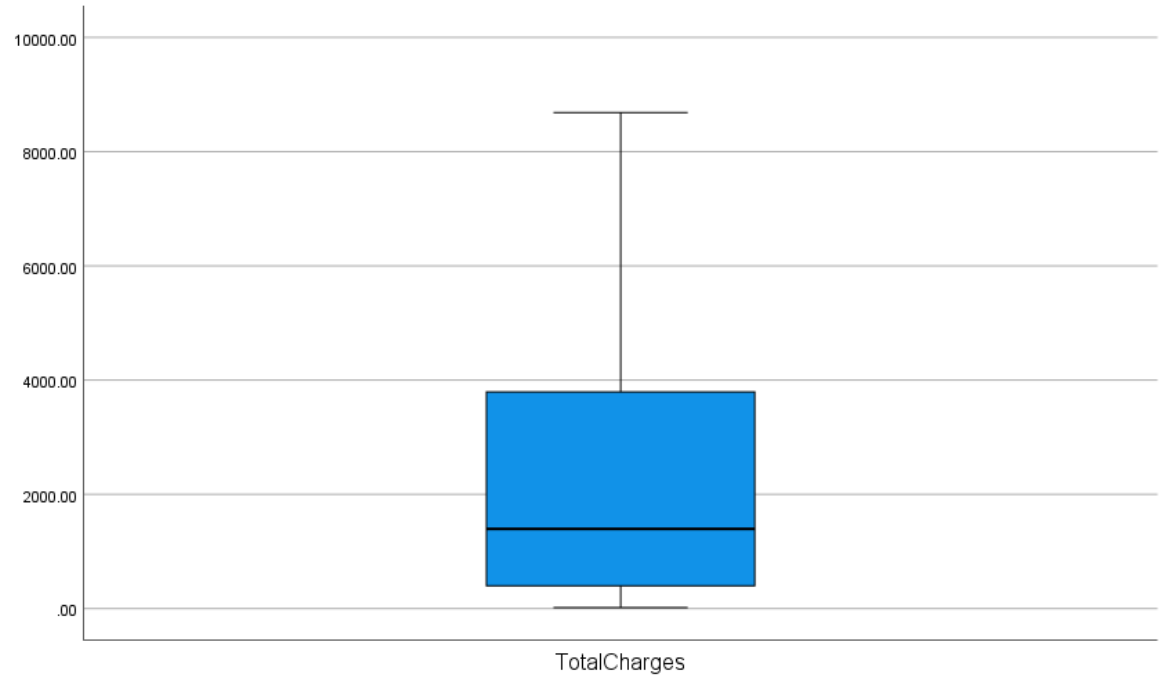


TotalCharges

Percentiles

		Percentiles						
		5	10	25	50	75	90	95
Weighted Average (Definition 1)	TotalCharges	49.5500	84.5300	401.3500	1397.4750	3795.2125	5978.8600	6927.1250
Tukey's Hinges	TotalCharges			401.4000	1397.4750	3794.9750		

Inter Quartile Range = 3795.2125 - 401.35 = 3393.8625
Outlier Ranges = 3393.8625 * 1.5 = 5090.79
Upper Outlier Range = 5090.79 + 3795.2125 = 8886.006
Lower Outlier Range = 401.35 - 5090.79 = -4689.44



Question 4.

MonthlyCharges variable was split into 5 groups by first determining the minimum and maximum values.

- Minimum – 18.25
- Maximum – 118.75

We subtract minimum from maximum and divide it by 5, we get 20.1. We can then create an interval consisting of ranges:

- 18.25
- 38.35
- 58.45
- 78.55
- 98.65
- 118.75

We created a new variable with conditional ranges to automatically determine if the value is going to be "cox pis, pis, normal, yaxshi, ela".

Question 5.

Statistics

		MonthlyCharges	tenure	TotalCharges
N	Valid	7043	7043	7032
	Missing	0	0	11
Mean		64.7616	32.37	2283.3004
Median		70.3500	29.00	1397.4750
Mode		20.05	1	20.20
Minimum		18.25	0	18.80
Maximum		118.75	72	8684.80
Percentiles	25	35.5000	9.00	401.3500
	50	70.3500	29.00	1397.4750
	75	89.8500	55.00	3795.2125

Question 7.

After importing the file from demo folder, we connect the table from output section to our sav file to see the variables and values of the "bankloan.sav" file. We can see that the variables consists of age, education, credit debt, income, etc.

age	ed	employ	address	income	debtinc	creddebt	othdebt	default	preddef1	preddef2	preddef3
41.000	3.000	17.000	12.000	176.000	9.300	11.359	5.009	1.000	0.808	0.789	0.213
27.000	1.000	10.000	6.000	31.000	17.300	1.362	4.001	0.000	0.198	0.128	0.437
40.000	1.000	15.000	14.000	55.000	5.500	0.856	2.169	0.000	0.010	0.003	0.141
41.000	1.000	15.000	14.000	120.000	2.900	2.659	0.821	0.000	0.022	0.010	0.104
24.000	2.000	2.000	0.000	28.000	17.300	1.787	3.057	1.000	0.782	0.738	0.437
41.000	2.000	5.000	5.000	25.000	10.200	0.393	2.157	0.000	0.217	0.328	0.234
39.000	1.000	20.000	9.000	67.000	30.600	3.834	16.668	0.000	0.186	0.179	0.817
43.000	1.000	12.000	11.000	38.000	3.600	0.129	1.239	0.000	0.015	0.011	0.113
24.000	1.000	3.000	4.000	19.000	24.400	1.358	3.278	1.000	0.748	0.619	0.664
36.000	1.000	0.000	13.000	25.000	19.700	2.778	2.147	0.000	0.815	0.797	0.516
27.000	1.000	0.000	1.000	16.000	1.700	0.183	0.089	0.000	0.350	0.611	0.091
25.000	1.000	4.000	0.000	23.000	5.200	0.252	0.944	0.000	0.239	0.219	0.136
52.000	1.000	24.000	14.000	64.000	10.000	3.930	2.470	0.000	0.010	0.006	0.229