

Week 1 Case Study - SPSS

Friday, August 12, 2022 2:36 PM

Question 1.

Data science can add value to any business who can use their data well. By utilizing the data to construct algorithms and creating programs that help in proving optimal solutions to individual problems, lot of business problems can be solved. Organizations upper management can hire a good data scientist and trust them to be an advisor to them by maximizing their analytics capabilities. A data scientist can demonstrate the value of the institutions data to facilitate improved decision making processes across the entire organization, through measuring, tracking, and recording performance metrics and other information

Question 2.

Our data has variables for the package offered by telecommunication company and their revenue for chosen months. Numerical variable can be used to separate them and numerical variables, ARPUs can be used to determine which packages performs well and how to proceed in the future with those packages. So, let's say, company needs to remove one of the packages, together with numerical data, analyst can determine which package performs worse than the others.

Question 3.

COS variable is a categorical nominal variable as its values represent categories with no intrinsic ranking. ARPU variables are numerical continuous variables as there is infinite range of possible values.

Question 4.

In business, statistical research enables managers to analyze best performance, predict future business environments and make data-driven decision making. Statistics can help professionals to understand markets, make advertising decisions, set prices and respond to changes in consumer demand. Descriptive analytics looks at what has happened. Diagnostic analytics helps explain why. Some common applications of descriptive and diagnostic analytics include sales, marketing, finance and operations. Predictive analytics uses a variety of statistical

techniques to predict future probabilities and trends based on historical data. This goes beyond reporting what has happened to create best estimates for what will happen. Some common applications are fraud detection, security, risk assessment, marketing and operations.

Question 5.

After splitting the data, we analyze using frequencies, and get the required information in our report.

Statistics

average_revenue

Gencol N Valid 1525

Missing 0

Mean 9.8599

Median 7.4176

Mode .20

Std. Deviation 9.44670

Percentiles 25 4.5468

50 7.4176

75 12.0723

Sade N Valid 13030

Missing 3

Mean 12.9219

Median 9.5064

Mode .20

Std. Deviation 12.08166

Percentiles 25 5.3974

50 9.5064

75 16.1164

Serbest N Valid 605

Missing 0

Mean 26.4819

Median 19.5135

Mode .10a

Std. Deviation 23.37864

Percentiles 25 10.5506

50 19.5135

75 34.4681

a Multiple modes exist. The smallest value is shown

We can see that the "Serbest" has the highest mean which means it performs the best out of all three. "Serbest" is the 2nd best performing, and the worse performing is "Gencol".

Question 6.

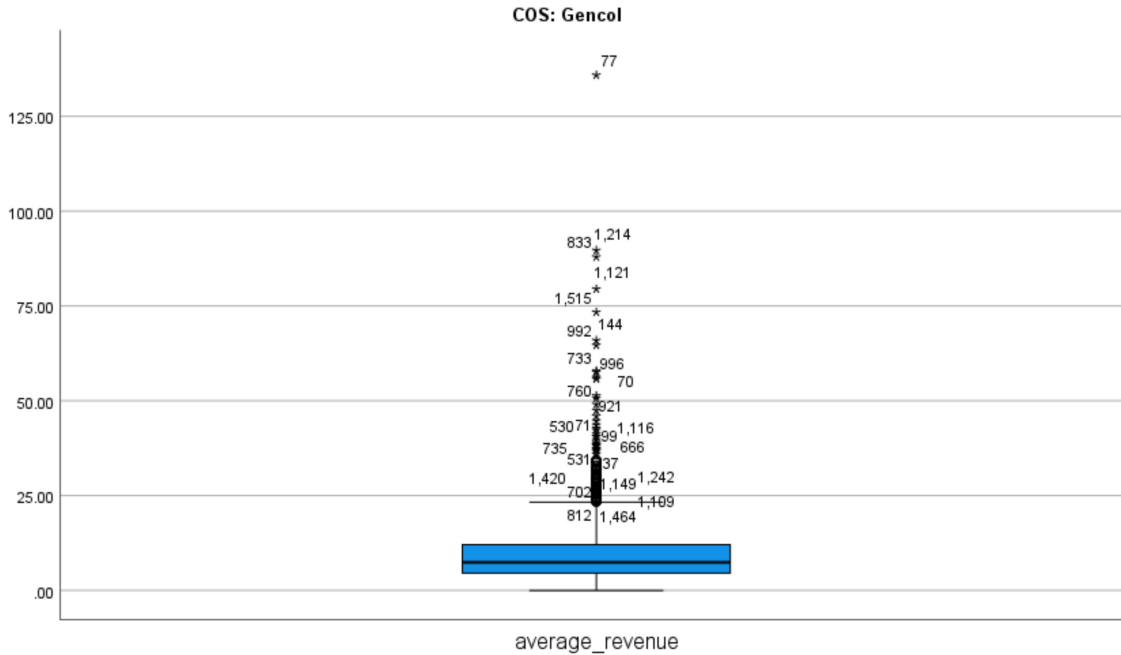
Gencol:

Inter quartile range: $12.0723 - 4.5468 = 7.5255$

Outlier Ranges: $7.5255 * 1.5 = 11.2882$

Upper outlier range: $12.0723 + 11.2882 = 23.3605$

Lower outlier range: $4.5468 - 11.2882 = -6.7414$



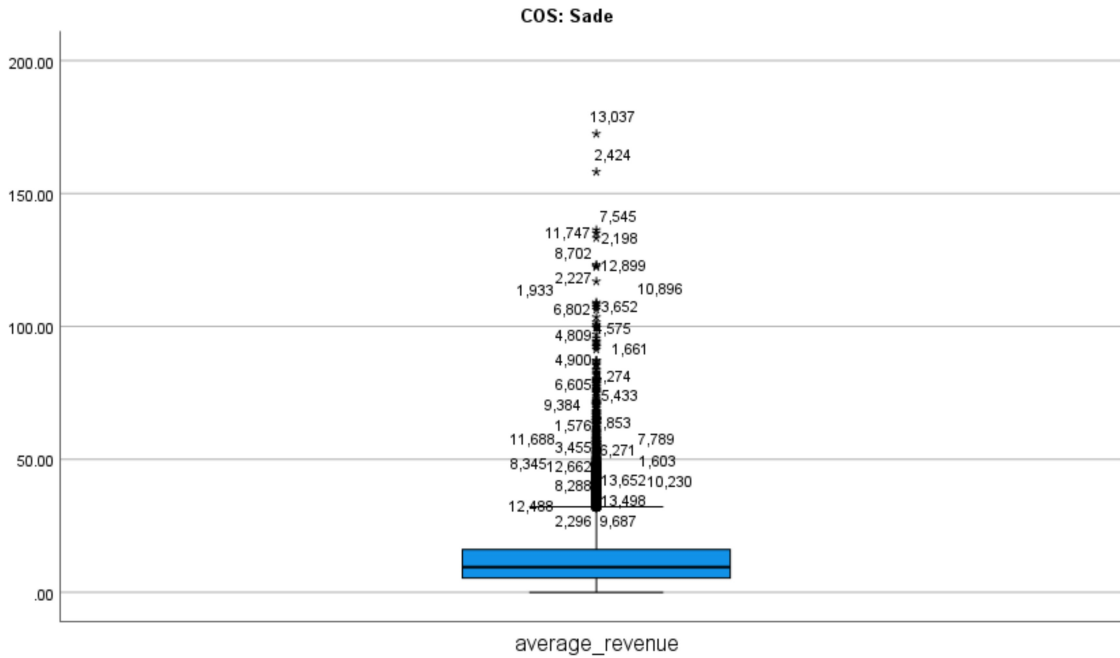
Sade:

Inter quartile range: $16.1164 - 5.3974 = 10.719$

Outlier ranges: $10.719 * 1.5 = 16.0785$

Upper outlier range: $16.1164 + 16.0785 = 32.1949$

Lower outlier range: $5.3974 - 16.0785 = -10.6811$



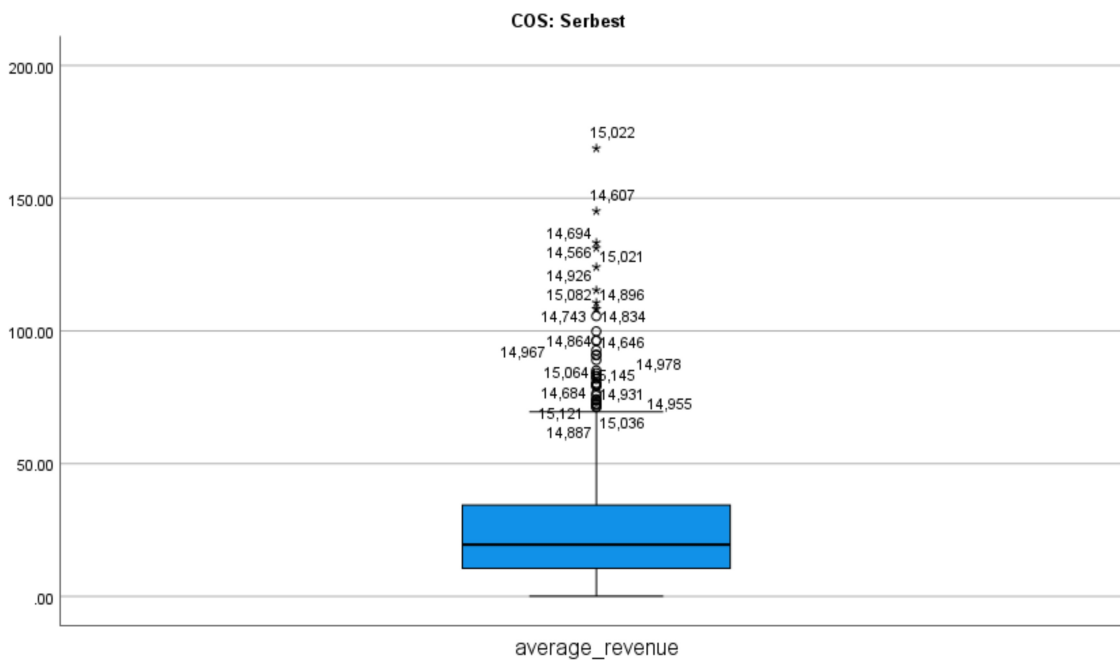
Serbest:

Inter quartile range: $34.4681 - 10.5506 = 23.9175$

Outlier range: $23.9175 * 1.5 = 35.8763$

Upper outlier range: $34.4681 + 35.8763 = 70.3444$

Lower outlier range: $10.5506 - 35.8763 = -25.3257$



Question 7.

Gencol:

Range - 135.87

Dividing into 5 subcategories: $135.87/5 = 27.174$

0 -- 27.174 -- 54.348 -- 81.522 -- 108.696 -- 135.87

Sade:

Range - 172.48

Dividing into 5 subcategories: $172.48/5 = 34.496$

0 -- 34.496 -- 68.992 -- 103.488 -- 137.984 -- 172.48

Serbest:

Range - 168.69

Dividing into 5 subcategories: $168.69/5 = 33.738$

0 -- 33.738 -- 67.476 -- 101.214 -- 134.952 -- 168.69

Question 8.



Table (8 fields, 15,163 records)								
File Edit Generate								
Table Annotations								
	COS	ARPU_Sep	ARPU_Oct	ARPU_Nov	ARPU_Dec	ARPU_Jan	average_revenue	ar_sub
1	Gencol	1.260	0.460	0.820	31.530	4.908	7.796	cox pis
2	Gencol	8.530	8.030	8.360	3.990	8.966	7.575	cox pis
3	Gencol	25.330	33.430	25.310	\$null\$	21.186	21.051	cox pis
4	Gencol	13.900	20.820	20.730	19.370	5.606	16.085	cox pis
5	Gencol	6.000	22.990	46.000	23.750	\$null\$	19.748	cox pis
6	Gencol	5.000	5.000	5.000	20.000	5.923	8.185	cox pis
7	Gencol	\$null\$	4.120	27.170	2.890	\$null\$	6.836	cox pis
8	Gencol	10.950	4.560	4.490	9.310	2.133	6.289	cox pis
9	Gencol	\$null\$	\$null\$	5.810	8.640	\$null\$	2.890	cox pis
10	Gencol	4.180	0.320	3.560	6.480	1.587	3.225	cox pis
11	Gencol	\$null\$	\$null\$	6.710	6.160	\$null\$	2.574	cox pis
12	Gencol	8.730	5.990	5.990	6.740	4.613	6.413	cox pis
13	Gencol	7.600	27.780	35.480	32.200	0.462	20.704	cox pis
14	Gencol	19.190	26.840	18.520	\$null\$	16.230	16.156	cox pis
15	Gencol	10.560	2.860	2.910	2.160	5.136	4.725	cox pis
16	Gencol	0.780	5.100	9.590	0.190	8.475	4.827	cox pis
17	Gencol	30.040	26.090	10.090	23.860	17.838	21.584	cox pis
18	Gencol	0.570	1.840	0.670	7.370	0.102	2.110	cox pis
19	Gencol	\$null\$	4.300	0.100	10.010	6.597	4.201	cox pis
20	Gencol	4.430	6.580	5.810	16.860	3.708	7.478	cox pis
21	Gencol	9.930	7.010	8.630	6.010	4.975	7.311	cox pis
22	Gencol	\$null\$	\$null\$	1.000	8.380	\$null\$	1.876	cox pis
23	Gencol	2.140	3.910	12.400	6.000	\$null\$	4.890	cox pis
24	Gencol	10.000	8.000	5.990	1.250	\$null\$	5.048	cox pis
25	Gencol	10.370	8.500	15.340	11.150	11.304	11.333	cox pis

Question 9.



File Edit Generate									
Audit Quality Annotations									
Field	Sample Graph	Measurement	Min	Max	Mean	Std. Dev	Skewness	Unique	Valid
COS		Nominal	--	--	--	--	--	3	15163
ARPU_Sep		Continuous	-6.580	250.260	15.607	17.605	3.259	--	12946
ARPU_Oct		Continuous	-5.530	242.820	14.861	16.660	3.352	--	13505
ARPU_Nov		Continuous	-8.630	259.730	15.318	17.153	3.382	--	14455
ARPU_Dec		Continuous	0.000	245.940	16.962	14.223	3.544	--	14745
ARPU_Jan		Continuous	0.000	185.573	11.008	12.476	3.609	--	11163
average_reve...		Continuous	0.000	172.477	13.155	12.823	3.234	--	15160
ar_sub		Nominal	--	--	--	--	--	7	15158

* Indicates a multimode result * Indicates a sampled result

OK