What data can be used for descriptive analysis?

Authored by

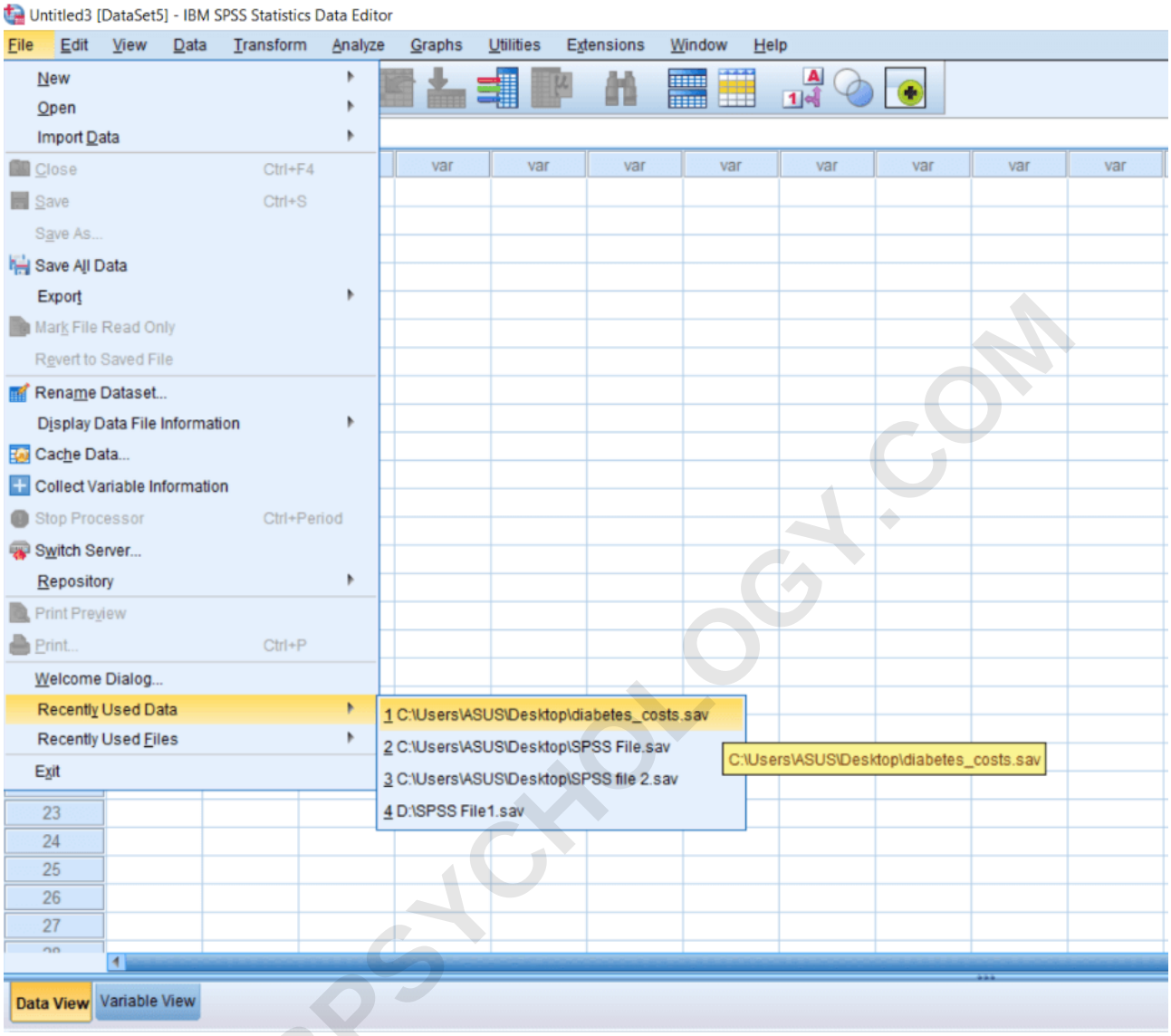
stats writer

June 23, 2024

Descriptive analysis is a statistical method used to summarize and describe a dataset to gain a better understanding of its characteristics. This method utilizes various data types such as numerical, categorical, and ordinal data. Numerical data, such as age, height, and income, can be used to calculate measures of central tendency and variability. Categorical data, like gender or education level, can be used to create frequency tables and bar charts to illustrate the distribution of the data. Ordinal data, such as ratings or rankings, can be used to determine the relative positioning of the data elements. In summary, descriptive analysis can be performed on a wide range of data types to provide valuable insights into a dataset.

Setting Data for Descriptive Analysis

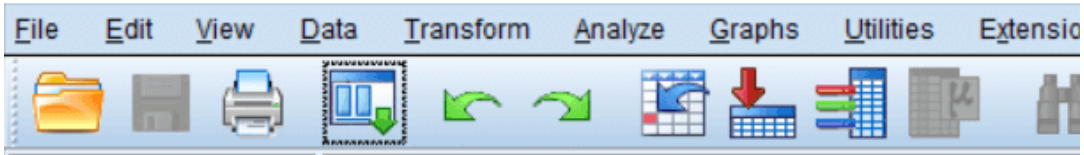
In this section, we will learn how we can calculate various shots of Descriptive stats using SPSS. SPSS is a wonderful software when we want to use descriptive analysis apart from the inferences analysis as well in our research. There are many ways through which we can calculate the Descriptive analysis in SPSS. Currently, we don't have any data. So first, we will import any data set in SPSS so that we can demonstrate how to use descriptive analysis. To import a new data set, we will go to the File menu, and then we can see many data set open here earlier. One of the data set is the diabetes_costs data set, as shown below:



It is a comparative new data set and less use. So we will use this data set. Click on this data set to open it. The following details will be open after click on this:

diabetes_costs.sav [DataSet7] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensio



	age	glucose	income	cost	var
1	43	6.3	22,445	3,748	
2	47	5.2	30,444	2,323	
3	34	7.2	35,333	2,145	
4	24	5.8	42,335	3,532	
5	23	6.4	47,323	5,332	
6	54	7.1	52,664	5,433	
7	36	6.8	68,333	2,465	
8	32	5.3	63,554	7,534	
9	14	7.5	57,754	5,937	
10	53	8.2	84,344	3,543	
11	64	6.9	80,335	6,423	
12	32	7.3	85,345	2,444	
13	76	7.6	79,443	7,653	
14	54	6.3	73,443	3,456	
15	34	6.9	68,543	12,453	
16	48	8.1	49,568	7,535	
17	34	8.5	65,635	8,435	
18	64	7.4	66,376	6,324	
19	51	7.5	78,854	7,256	
20	40	8.3	83,455	9,345	
21	36	8.5	84,345	9,433	
22	75	6.2	74,499	6,383	
23	38	6.8	68,536	5,643	
24	72	7.5	74,664	11,453	
25	75	8.3	83,674	5,324	
26	20	6.2	75,645	7,543	
27	25	8.2	64,644	9,433	

Data View Variable View

This is our data set that currently we are going to use for the demonstration purpose for calculating various sorts of descriptive stats. This data set contains the age of the subject, glucose level, income, and the treatment cost. So it is about calculating the treatment cost of the subject based on the age, glucose level, and the income of the subject. So we can use this for various sorts of descriptive stats.

Descriptive stats can be useful in communicating the overall picture of our data set. For example, we might wish to communicate with someone and want to find the average age of the subject analyzed in this data set and the percentage of the subjects below a particular age. For example, suppose we want to communicate that what percentage of subjects in this data set below 52 years or 60 years or any number of years. For this, we need to take the help of percentiles. If we need to communicate the average age of the subject, in that case, we wish to talk about the mean score, average scores. Then there are median, modes, etc.

We have a glucose level, income. So again, we might be interested in what is the average income of the subjects

which are there in the data set and what is the treatment cost.