

Analyze the internet data of www.datadb.com

Business Analytic Foundation with R Tools- Solutions



Solutions

Disclaimer: In Business Analytics, there are different ways of solving the same set of problems, we are just presenting one. Feel free to explore other ways of answering these questions.

1. The team wants to analyze each variable of the data collected through data summarization to get a basic understanding of the dataset and to prepare for further analysis.

To understand how the data is distributed in the dataset and the kind of variables present along with their count, what is their maximum and minimum value, etc., a summarization of the data is done.

Code:

```
summary(mydata)
```

Result:

From the result of summarized dataset, it is observed that the numerical data includes information related to the maximum, minimum, and mean data. The categorical data like continent includes the data of the number of times the category has been repeated in the dataset. We can see that there is a maximum value of 30 bounces for the website. This site was accessed maximum number of times by visitors from North America.

```
> summary(mydata)
```

Bounces		Exits		Continent		Source.Group		Time.on.Page		Unique.Pageviews		Visits	
Min. :	0.000	Min. :	0.000	AF :	321	google	:11542	Min. :	0.00	Min. :	1.000	Min. :	0.000
1st Qu.:	0.000	1st Qu.:	1.000	AS :	3171	(direct)	: 7532	1st Qu.:	0.00	1st Qu.:	1.000	1st Qu.:	1.000
Median :	1.000	Median :	1.000	EU :	6470	Others	: 5360	Median :	0.00	Median :	1.000	Median :	1.000
Mean :	0.713	Mean :	0.906	N.America:	20043	tableausoftware.com	: 2388	Mean :	73.18	Mean :	1.114	Mean :	0.906
3rd Qu.:	1.000	3rd Qu.:	1.000	OC :	1356	t.co	: 2249	3rd Qu.:	10.00	3rd Qu.:	1.000	3rd Qu.:	1.000
Max. :	30.000	Max. :	36.000	SA :	748	public.tableausoftware.com:	1354	Max. :	46745.00	Max. :	45.000	Max. :	45.000
						(Other)	: 1684						