

# diamonds

## Autogenerated data summary from dataMaid

2021-02-25 21:35:15

## Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	53940
Number of variables	10

## Checks performed

The following variable checks were performed, depending on the data type of each variable:

	character	factor	labelled	haven labelled	numeric	integer	logical	Date
Identify miscoded missing values	x	x	x	x	x	x		x
Identify prefixed and suffixed whitespace	x	x	x	x				
Identify levels with < 6 obs.	x	x	x	x				
Identify case issues	x	x	x	x				
Identify misclassified numeric or integer variables	x	x	x	x				
Identify outliers					x	x		x

Please note that all numerical values in the following have been rounded to 2 decimals.

## Summary table

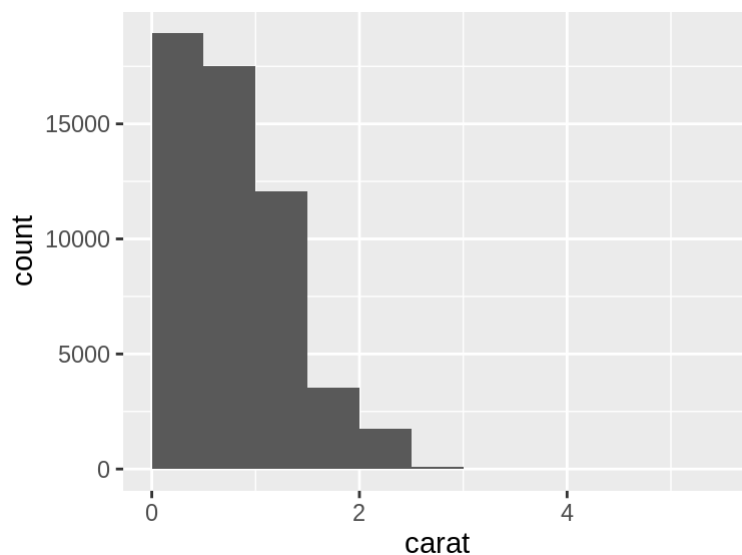
	Variable class	# unique values	Missing observations	Any problems?
carat	numeric	273	0.00 %	x
cut	ordered	5	0.00 %	
color	ordered	7	0.00 %	
clarity	ordered	8	0.00 %	

	Variable class	# unique values	Missing observations	Any problems?
depth	numeric	184	0.00 %	×
table	numeric	127	0.00 %	×
price	integer	11602	0.00 %	
x	numeric	554	0.00 %	×
y	numeric	552	0.00 %	×
z	numeric	375	0.00 %	×

## Variable list

### carat

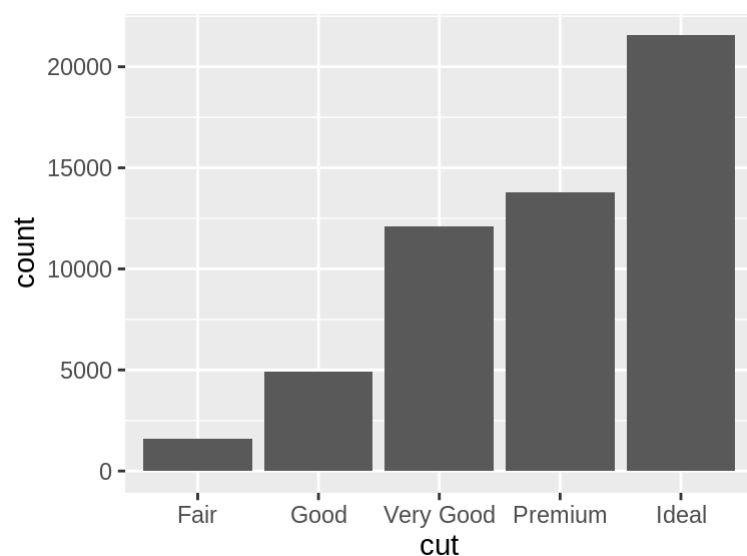
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	273
Median	0.7
1st and 3rd quartiles	0.4; 1.04
Min. and max.	0.2; 5.01



- Note that the following possible outlier values were detected: "2.47", "2.48", "2.49", "2.5", "2.51", "2.52", "2.53", "2.54", "2.55", "2.56" (36 additional values omitted).

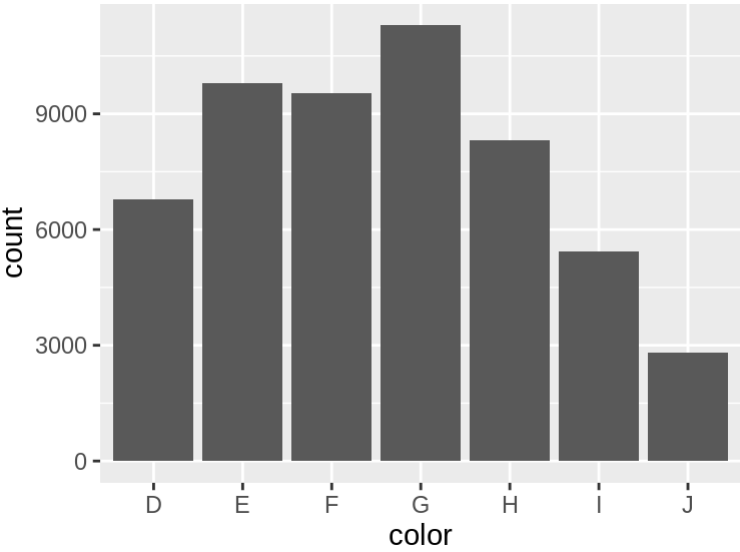
# cut

Feature	Result
Variable type	ordered
Number of missing obs.	0 (0 %)
Number of unique values	5
Mode	"Ideal"
Reference category	Fair



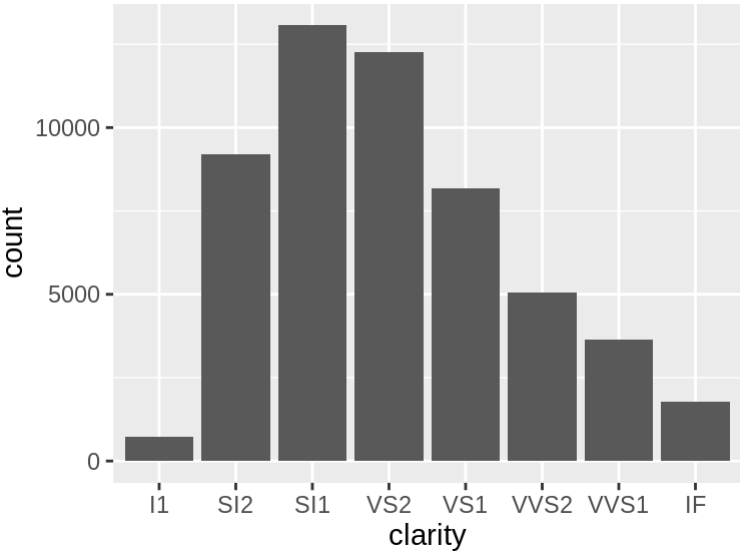
# color

Feature	Result
Variable type	ordered
Number of missing obs.	0 (0 %)
Number of unique values	7
Mode	"G"
Reference category	D



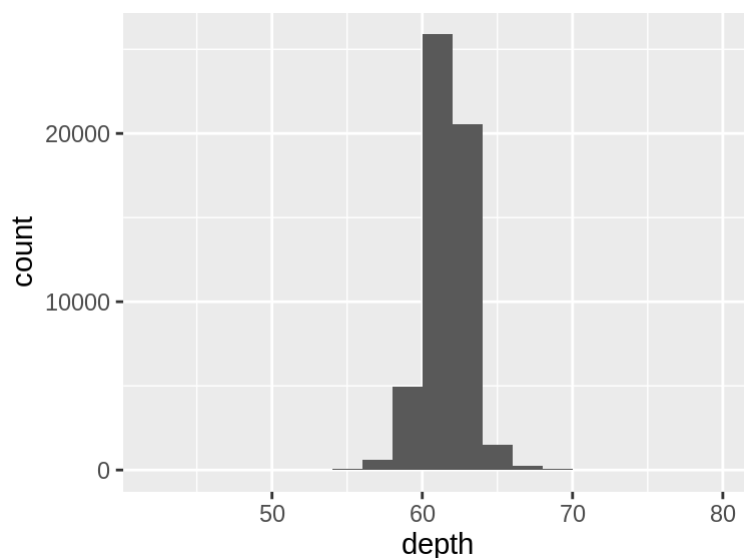
clarity

Feature	Result
Variable type	ordered
Number of missing obs.	0 (0 %)
Number of unique values	8
Mode	"SI1"
Reference category	I1



depth

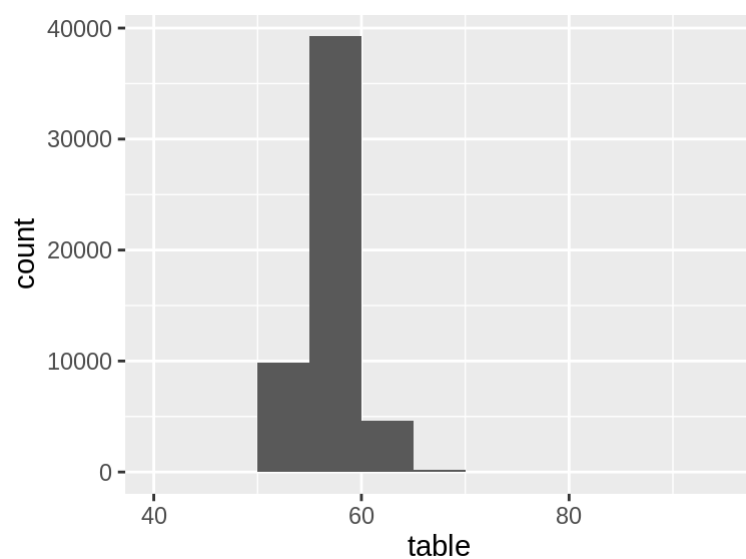
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	184
Median	61.8
1st and 3rd quartiles	61; 62.5
Min. and max.	43; 79



- Note that the following possible outlier values were detected: "43", "44", "50.8", "51", "52.2", "52.3", "52.7", "53", "53.1", "53.2" (114 additional values omitted).

## table

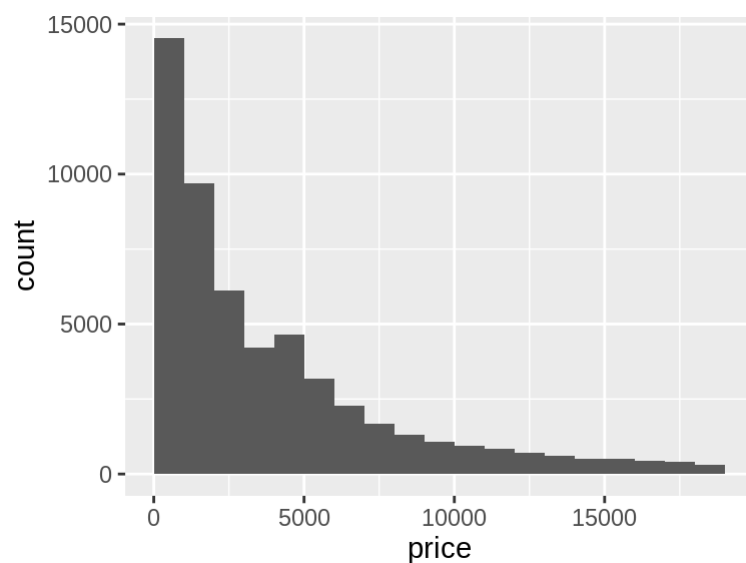
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	127
Median	57
1st and 3rd quartiles	56; 59
Min. and max.	43; 95



- Note that the following possible outlier values were detected: "43", "44", "49", "50", "50.1", "51", "51.6", "52", "52.4", "52.8" (18 additional values omitted).

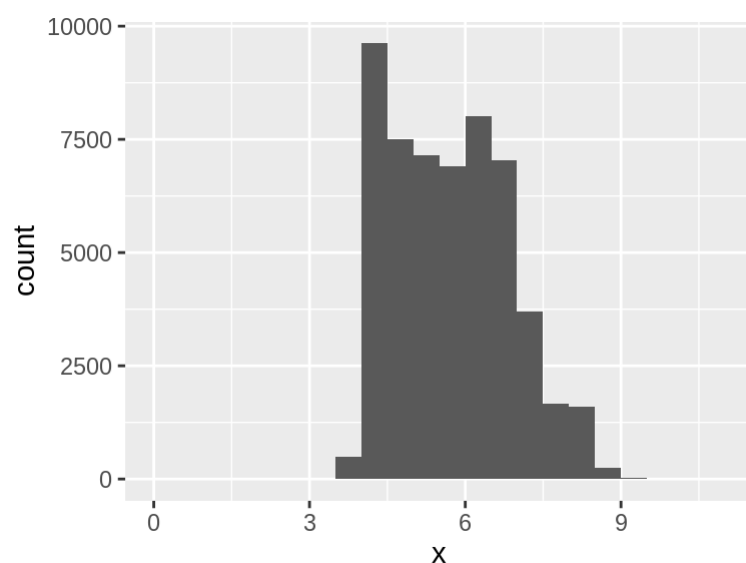
## price

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	11602
Median	2401
1st and 3rd quartiles	950; 5324.25
Min. and max.	326; 18823



## X

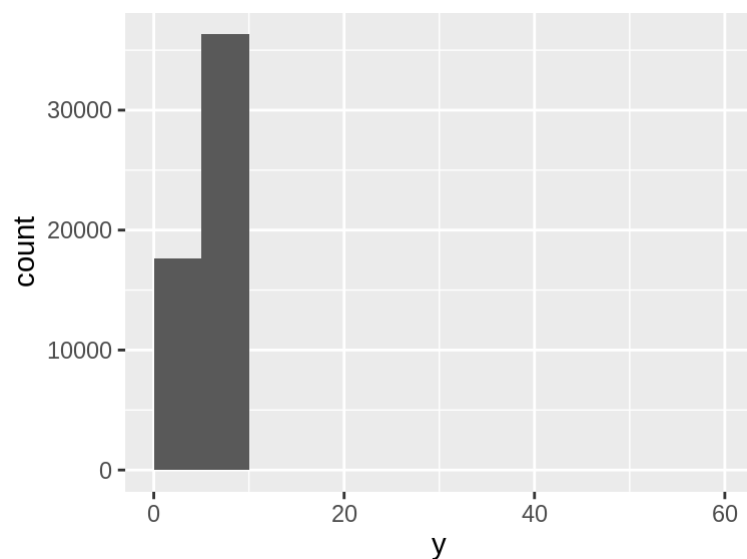
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	554
Median	5.7
1st and 3rd quartiles	4.71; 6.54
Min. and max.	0; 10.74



- Note that the following possible outlier values were detected: "0", "9.17", "9.23", "9.24", "9.25", "9.26", "9.3", "9.32", "9.35", "9.36" (17 additional values omitted).

## y

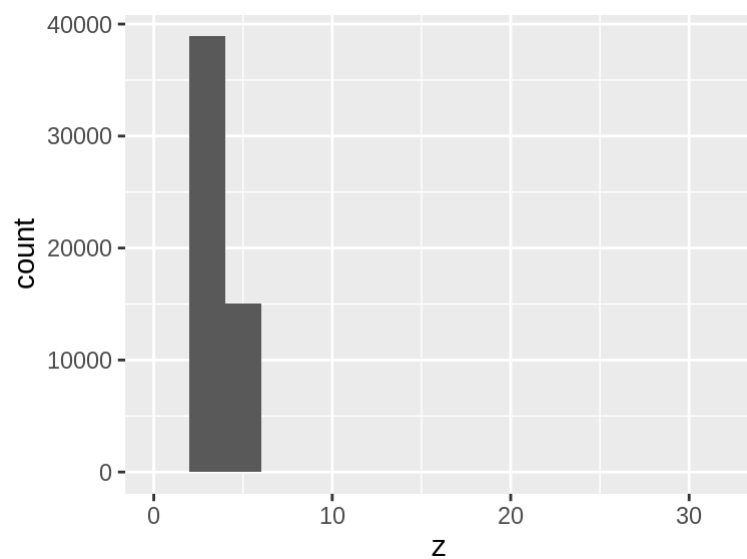
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	552
Median	5.71
1st and 3rd quartiles	4.72; 6.54
Min. and max.	0; 58.9



- Note that the following possible outlier values were detected: "0", "9.01", "9.02", "9.03", "9.07", "9.1", "9.13", "9.14", "9.19", "9.2" (22 additional values omitted).

## Z

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	375
Median	3.53
1st and 3rd quartiles	2.91; 4.04
Min. and max.	0; 31.8





- Note that the following possible outlier values were detected: "0", "5.56", "5.57", "5.58", "5.59", "5.6", "5.61", "5.62", "5.64", "5.65" (29 additional values omitted).
- 

#### Report generation information:

- Created by: Could not determine from system (username: root ).
- Report creation time: Thu Feb 25 2021 21:35:16
- Report was run from directory: /content
- dataMaid v1.4.0 [Pkg: 2019-12-10 from CRAN (R 4.0.4)]
- R version 4.0.4 (2021-02-15).
- Platform: x86\_64-pc-linux-gnu (64-bit)(Ubuntu 18.04.5 LTS).
- Function call: makeDataReport(data = diamonds)