

# tidyDF

Autogenerated data summary from dataMaid

*2019-10-24 13:24:00*

## Part 1

# Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	2410
Number of variables	15

### 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	×	×	×	×	×	×		×
Identify prefixed and suffixed whitespace	×	×	×	×				
Identify levels with < 6 obs.	×	×	×	×				
Identify case issues	×	×	×	×				
Identify misclassified numeric or integer variables	×	×	×	×				
Identify outliers					×	×		×

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

## Part 2

### Summary table

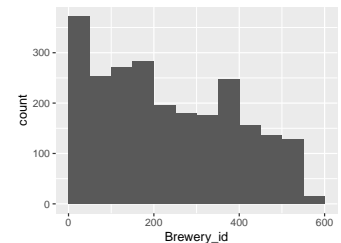
	Variable class	# unique values	Missing observations	Any problems?
Brewery_id	integer	558	0.00 %	
Beer_name	factor	2305	0.00 %	×
Beer_ID	integer	2410	0.00 %	
ABV	numeric	76	0.00 %	×
IBU	numeric	172	0.00 %	×
Style	factor	100	0.00 %	×
Ounces	numeric	7	0.00 %	×
Brewery_name	factor	551	0.00 %	×
City	factor	384	0.00 %	×
State	factor	51	0.00 %	×
ABVmedian	numeric	17	0.00 %	×
predict_IBU	numeric	76	0.00 %	×
IPA_Ale	character	3	0.00 %	
BudweiserOrOther	character	1	0.00 %	×
abb	factor	51	0.00 %	×

## Part 3

# Variable list

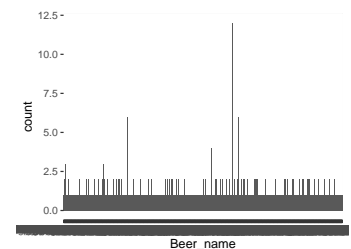
### Brewery\_id

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	558
Median	206
1st and 3rd quartiles	94; 367
Min. and max.	1; 558



### Beer\_name

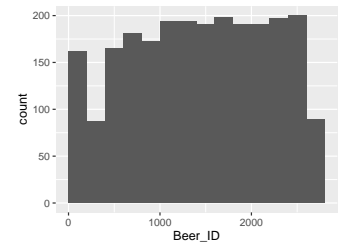
Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2305
Mode	"Nonstop Hef Hop"
Reference category	#001 Golden Amber Lager



- The following values appear with prefixed or suffixed white space: "14° ESB ", "32°/50° Kölsch ", "Alter Ego ", "Aurora ", "Bunker Hill Blueberry Ale ", "Lewbricator Wheat Dopplebock ", "Nordskye ", "Port Barrel Wee Mac ", "Texicali ", "White Rabbit ".
- Note that the following levels have at most five observations: "#001 Golden Amber Lager", "#002 American I.P.A.", "#003 Brown & Robust Porter", "#004 Session I.P.A.", "#9", "077XX", "10 Degrees of Separation", "10 Ton", "113 IPA", "11th Hour IPA" (2292 additional values omitted).

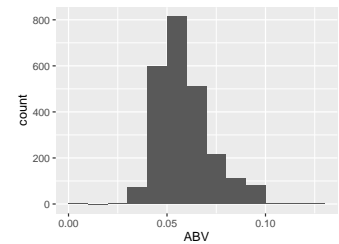
## Beer\_ID

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	2410
Median	1453.5
1st and 3rd quartiles	808.25; 2075.75
Min. and max.	1; 2692



## ABV

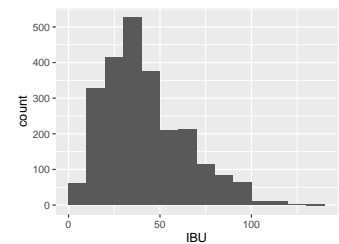
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	76
Median	0.06
1st and 3rd quartiles	0.05; 0.07
Min. and max.	0; 0.13



- Note that the following possible outlier values were detected: "0", "0.03", "0.03", "0.03", "0.03", "0.04", "0.04", "0.04", "0.04", "0.12" (2 additional values omitted).

## IBU

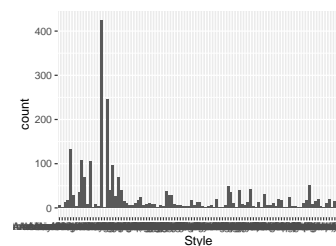
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	172
Median	36.41
1st and 3rd quartiles	26; 55.64
Min. and max.	0; 138



- Note that the following possible outlier values were detected: "0", "0.52", "4", "5", "6", "7", "8", "9", "10", "10.77" (2 additional values omitted).

## Style

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	100
Mode	"American IPA"
Reference category	

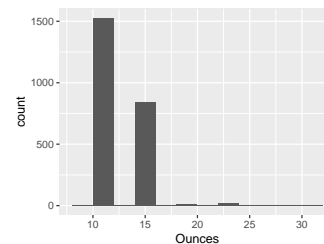


- The following suspected missing value codes enter as regular values: "".
- Note that the following levels have at most five observations: "", "Abbey Single Ale", "American Barleywine", "American Double / Imperial Pilsner", "American India Pale Lager", "American Malt Liquor", "Braggot", "Chile Beer", "Dubbel", "Dunkelweizen" (26 additional values omitted).

---

## Ounces

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	7
Median	12
1st and 3rd quartiles	12; 16
Min. and max.	8.4; 32

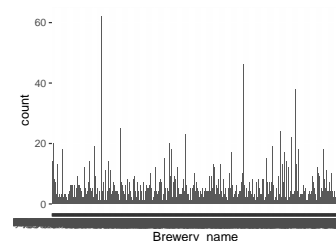


- Note that the following possible outlier values were detected: "8.4".

---

## Brewery\_name

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	551
Mode	"Brewery Vivant"
Reference category	10 Barrel Brewing Company

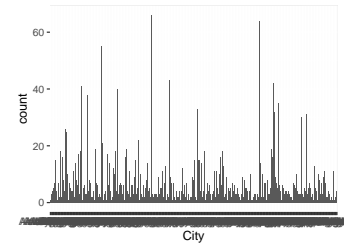


- The following values appear with prefixed or suffixed white space: "Four Fathers Brewing ", "Mountain Town Brewing Company ", "NorthGate Brewing ".
- Note that the following levels have at most five observations: "10 Barrel Brewing Company", "2 Towns Ciderhouse", "3 Daughters Brewing", "7 Seas Brewing Company", "7venth Sun", "Abita Brewing Company", "AC Golden Brewing Company", "Against The Grain Brewery", "Airways Brewing Company", "Alameda Brewing" (418 additional values omitted).

- Note that there might be case problems with the following levels: "Against the Grain Brewery", "Against The Grain Brewery".

## City

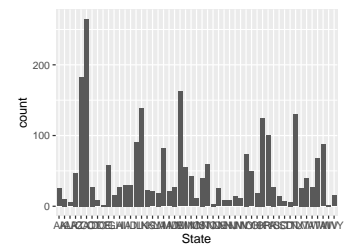
Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	384
Mode	"Grand Rapids"
Reference category	Abingdon



- Note that the following levels have at most five observations: "Abingdon", "Abita Springs", "Ada", "Afton", "Airway Heights", "Alpine", "Angola", "Ann Arbor", "Arrington", "Ashburn" (252 additional values omitted).

## State

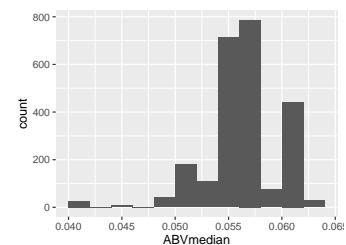
Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	51
Mode	" CO"
Reference category	AK



- The following values appear with prefixed or suffixed white space: " AK", " AL", " AR", " AZ", " CA", " CO", " CT", " DC", " DE", " FL" (41 additional values omitted).
- Note that the following levels have at most five observations: " AR", " DE", " ND", " WV".

## ABVmedian

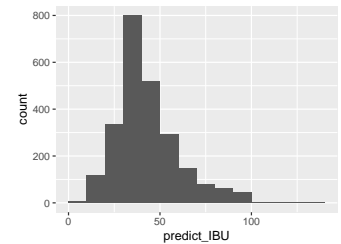
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	17
Median	0.06
1st and 3rd quartiles	0.06; 0.06
Min. and max.	0.04; 0.06



- Note that the following possible outlier values were detected: "0.04", "0.05", "0.05".

## predict\_IBU

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	76
Median	38.98
1st and 3rd quartiles	30; 51.8
Min. and max.	0; 130

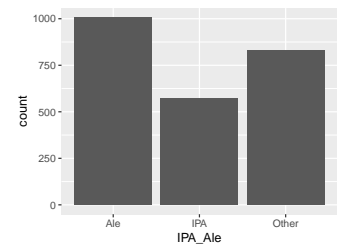


- Note that the following possible outlier values were detected: "0", "0.52", "1.8", "6.93", "9.49", "10.77", "13.34", "14.62", "15.9", "119.75" (2 additional values omitted).

---

## IPA\_Ale

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	3
Mode	"Ale"



---

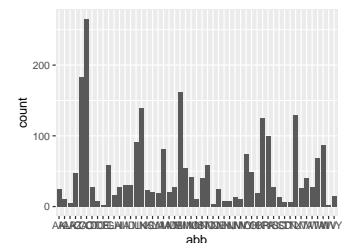
## BudweiserOrOther

- The variable only takes one (non-missing) value: "Other". The variable contains 0 % missing observations.

---

## abb

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	51
Mode	"CO"
Reference category	AK



- Note that the following levels have at most five observations: "AR", "DE", "ND", "WV".

---

Report generation information:



- Created by 48120778 (username: q1608).
- Report creation time: Thu Oct 24 2019 13:24:00
- Report was run from directory: C:/Edu/GitHub\_DDS/Unit8
- dataMaid v1.3.2 [Pkg: 2019-07-27 from CRAN (R 3.5.3)]
- R version 3.5.3 (2019-03-11).
- Platform: x86\_64-w64-mingw32/x64 (64-bit)(Windows 10 x64 (build 17763)).
- Function call: makeDataReport(data = tidyDF)