

# breweries\_\_clean

Autogenerated data summary from dataMaid

*2018-06-19 15:12:39*

## Part 1

# Data cleaning summary

The dataset examined has the following dimensions:

| Feature             | Result |
|---------------------|--------|
| Number of rows      | 555    |
| Number of variables | 4      |

### Checks performed

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

|  | character | factor | labelled | numeric | integer | logical | Date |
|--|-----------|--------|----------|---------|---------|---------|------|
| Identify miscoded missing values                       | ×         | ×      | ×        | ×       | ×       |         |      |
| Identify prefixed and suffixed<br>whitespace           | ×         | ×      | ×        |         |         |         |      |
| Identify levels with < 6 obs.                          | ×         | ×      | ×        |         |         |         |      |
| Identify case issues                                   | ×         | ×      | ×        |         |         |         |      |
| Identify misclassified numeric or<br>integer variables | ×         | ×      | ×        |         |         |         |      |
| Identify outliers                                      |           |        |          | ×       | ×       |         | ×    |

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

## Part 2

### Summary table

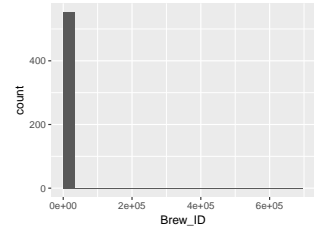
|              | Variable class | # unique values | Missing observations | Any problems? |
|--------------|----------------|-----------------|----------------------|---------------|
| Brew_ID      | integer        | 555             | 0.00 %               | ×             |
| Brewery_Name | factor         | 550             | 0.00 %               | ×             |
| City         | character      | 381             | 0.00 %               | ×             |
| State        | character      | 51              | 0.00 %               | ×             |

## Part 3

# Variable list

### Brew\_ID

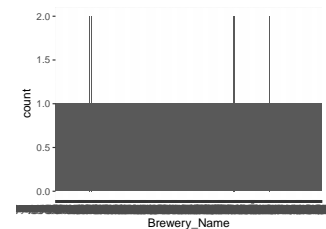
| Feature                 | Result       |
|-------------------------|--------------|
| Variable type           | integer      |
| Number of missing obs.  | 0 (0 %)      |
| Number of unique values | 555          |
| Median                  | 282          |
| 1st and 3rd quartiles   | 143.5; 421.5 |
| Min. and max.           | 1; 697225    |



- Note that the following possible outlier values were detected: "11881", "39204", "697225".

### Brewery\_Name

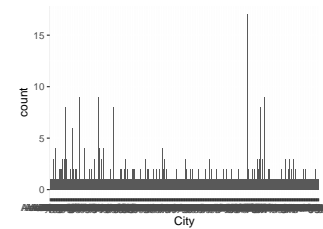
| Feature                 | Result               |
|-------------------------|----------------------|
| Variable type           | factor               |
| Number of missing obs.  | 0 (0 %)              |
| Number of unique values | 550                  |
| Mode                    | "Blackrocks Brewery" |



- Note that the following levels have at most five observations: "10 Barrel Brewing Company", "18th Street Brewery", "2 Towns Ciderhouse", "21st Amendment Brewery", "3 Daughters Brewing", "4 Hands Brewing Company", "450 North Brewing Company", "7 Seas Brewing Company", "7venth Sun", "Abita Brewing Company" (540 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           | character  |
| Number of missing obs.  | 0 (0 %)    |
| Number of unique values | 381        |
| Mode                    | "Portland" |

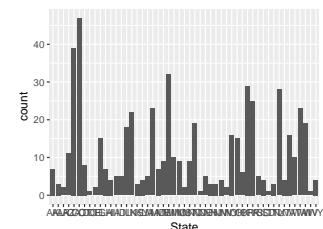


- Note that the following levels have at most five observations: "Abingdon", "Abita Springs", "Ada", "Afton", "Airway Heights", "Albuquerque", "Alpine", "Anchorage", "Angola", "Ann Arbor" (363 additional values omitted).

---

## State

| Feature                 | Result    |
|-------------------------|-----------|
| Variable type           | character |
| Number of missing obs.  | 0 (0 %)   |
| Number of unique values | 51        |
| Mode                    | "CO"      |



- Note that the following levels have at most five observations: "AL", "AR", "DC", "DE", "HI", "IA", "ID", "KS", "KY", "LA" (14 additional values omitted).

---

### Report generation information:

- Created by No Instance(s) Available..
- Report creation time: Tue Jun 19 2018 15:12:39
- dataMaid v1.0.0 [Pkg: 2017-11-10 from CRAN (R 3.4.3)]
- R version 3.4.3 (2017-11-30).
- Platform: x86\_64-w64-mingw32/x64 (64-bit)(Windows >= 8 x64 (build 9200)).
- Function call: `clean(data = breweries_clean)`