Codebook summary for all tables for Case Study 1:

Codebook for breweries

Autogenerated data summary from dataMaid

2018-10-15 02:06:25

# Data report overview

The dataset examined has the following dimensions:

|  |  |
| --- | --- |
| Feature | Result |
| Number of observations | 558 |
| Number of variables | 4 |

# Codebook summary table

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Label | Variable | Class | # unique values | Missing | Description |
|  | [**Brew\_ID**](#brew_id) | integer | 558 | 0.00 % |  |
|  | [**Brewery**](#brewery) | factor | 551 | 0.00 % |  |
|  | [**City**](#city) | factor | 384 | 0.00 % |  |
|  | [**State**](#state) | factor | 51 | 0.00 % |  |

# Variable list

## Brew\_ID

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | integer |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 558 |
| Median | 279.5 |
| 1st and 3rd quartiles | 140.25; 418.75 |
| Min. and max. | 1; 558 |

## Brewery

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | factor |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 551 |
| Mode | “Blackrocks Brewery” |

## City

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | factor |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 384 |
| Mode | “Portland” |

## State

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

Report generation information:

* Created by Landcruiser87 (username: andyh).
* Report creation time: Mon Oct 15 2018 02:06:25
* Report Was run from directory: C:/Users/andyh/OneDrive/Documents
* dataMaid v1.1.2 [Pkg: 2018-05-03 from CRAN (R 3.5.1)]
* R version 3.5.1 (2018-07-02).
* Platform: x86\_64-w64-mingw32/x64 (64-bit)(Windows >= 8 x64 (build 9200)).
* Function call: makeDataReport(data = breweries, mode = "summarize", file = "codebook\_breweries.Rmd", checks = list(list("showAllFactorLevels")), listChecks = FALSE, maxProbVals = FALSE, codebook = TRUE, reportTitle = "Codebook for breweries")

Codebook for beer

Autogenerated data summary from dataMaid

2018-10-15 02:05:00

# Data report overview

The dataset examined has the following dimensions:

|  |  |
| --- | --- |
| Feature | Result |
| Number of observations | 2410 |
| Number of variables | 7 |

# Codebook summary table

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Label | Variable | Class | # unique values | Missing | Description |
|  | [**Name**](#name) | factor | 2305 | 0.00 % |  |
|  | [**Beer\_ID**](#beer_id) | integer | 2410 | 0.00 % |  |
|  | [**ABV**](#abv) | numeric | 75 | 2.57 % |  |
|  | [**IBU**](#ibu) | integer | 108 | 41.70 % |  |
|  | [**Brew\_ID**](#brew_id) | integer | 558 | 0.00 % |  |
|  | [**Style**](#style) | factor | 100 | 0.00 % |  |
|  | [**Ounces**](#ounces) | numeric | 7 | 0.00 % |  |

# Variable list

## Name

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | factor |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 2305 |
| Mode | “Nonstop Hef Hop” |

## 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. | 62 (2.57 %) |
| Number of unique values | 74 |
| Median | 0.06 |
| 1st and 3rd quartiles | 0.05; 0.07 |
| Min. and max. | 0; 0.13 |

## IBU

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | integer |
| Number of missing obs. | 1005 (41.7 %) |
| Number of unique values | 107 |
| Median | 35 |
| 1st and 3rd quartiles | 21; 64 |
| Min. and max. | 4; 138 |

## Brew\_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 |

## Style

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | factor |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 100 |
| Mode | “American IPA” |

## 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 |

Report generation information:

* Created by Landcruiser87 (username: andyh).
* Report creation time: Mon Oct 15 2018 02:05:00
* Report Was run from directory: C:/Users/andyh/OneDrive/Documents
* dataMaid v1.1.2 [Pkg: 2018-05-03 from CRAN (R 3.5.1)]
* R version 3.5.1 (2018-07-02).
* Platform: x86\_64-w64-mingw32/x64 (64-bit)(Windows >= 8 x64 (build 9200)).
* Function call: makeDataReport(data = beer, mode = "summarize", file = "codebook\_beer.Rmd", checks = list(list("showAllFactorLevels")), listChecks = FALSE, maxProbVals = FALSE, codebook = TRUE, reportTitle = "Codebook for beer")

Codebook for BreweriesPerState

Autogenerated data summary from dataMaid

2018-10-15 02:04:23

# Data report overview

The dataset examined has the following dimensions:

|  |  |
| --- | --- |
| Feature | Result |
| Number of observations | 50 |
| Number of variables | 2 |

# Codebook summary table

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Label | Variable | Class | # unique values | Missing | Description |
|  | [**State**](#state) | factor | 50 | 0.00 % |  |
|  | [**Breweries**](#breweries) | integer | 25 | 0.00 % |  |

# Variable list

## State

* The variable is a key (distinct values for each observation).

## Breweries

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | integer |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 25 |
| Median | 7 |
| 1st and 3rd quartiles | 4; 16 |
| Min. and max. | 1; 47 |

Report generation information:

* Created by Landcruiser87 (username: andyh).
* Report creation time: Mon Oct 15 2018 02:04:23
* Report Was run from directory: C:/Users/andyh/OneDrive/Documents
* dataMaid v1.1.2 [Pkg: 2018-05-03 from CRAN (R 3.5.1)]
* R version 3.5.1 (2018-07-02).
* Platform: x86\_64-w64-mingw32/x64 (64-bit)(Windows >= 8 x64 (build 9200)).
* Function call: makeDataReport(data = BreweriesPerState, mode = "summarize", file = "codebook\_BreweriesPerState.Rmd", checks = list(list( "showAllFactorLevels")), listChecks = FALSE, maxProbVals = FALSE, codebook = TRUE, reportTitle = "Codebook for BreweriesPerState")

Codebook for merged\_data

Autogenerated data summary from dataMaid

2018-10-15 02:02:44

# Data report overview

The dataset examined has the following dimensions:

|  |  |
| --- | --- |
| Feature | Result |
| Number of observations | 2410 |
| Number of variables | 10 |

# Codebook summary table

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Label | Variable | Class | # unique values | Missing | Description |
|  | [**Name**](#name) | factor | 2305 | 0.00 % |  |
|  | [**Beer\_ID**](#beer_id) | integer | 2410 | 0.00 % |  |
|  | [**ABV**](#abv) | numeric | 75 | 2.57 % |  |
|  | [**IBU**](#ibu) | integer | 108 | 41.70 % |  |
|  | [**Brew\_ID**](#brew_id) | integer | 558 | 0.00 % |  |
|  | [**Style**](#style) | factor | 100 | 0.00 % |  |
|  | [**Ounces**](#ounces) | numeric | 7 | 0.00 % |  |
|  | [**Brewery**](#brewery) | factor | 551 | 0.00 % |  |
|  | [**City**](#city) | factor | 384 | 0.00 % |  |
|  | [**State**](#state) | factor | 51 | 0.00 % |  |

# Variable list

## Name

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | factor |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 2305 |
| Mode | “Nonstop Hef Hop” |

## 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. | 62 (2.57 %) |
| Number of unique values | 74 |
| Median | 0.06 |
| 1st and 3rd quartiles | 0.05; 0.07 |
| Min. and max. | 0; 0.13 |

## IBU

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | integer |
| Number of missing obs. | 1005 (41.7 %) |
| Number of unique values | 107 |
| Median | 35 |
| 1st and 3rd quartiles | 21; 64 |
| Min. and max. | 4; 138 |

## Brew\_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 |

## Style

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | factor |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 100 |
| Mode | “American IPA” |

## 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 |

## Brewery

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | factor |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 551 |
| Mode | “Brewery Vivant” |

## City

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | factor |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 384 |
| Mode | “Grand Rapids” |

## State

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

Report generation information:

* Created by Landcruiser87 (username: andyh).
* Report creation time: Mon Oct 15 2018 02:02:45
* Report Was run from directory: C:/Users/andyh/OneDrive/Documents
* dataMaid v1.1.2 [Pkg: 2018-05-03 from CRAN (R 3.5.1)]
* R version 3.5.1 (2018-07-02).
* Platform: x86\_64-w64-mingw32/x64 (64-bit)(Windows >= 8 x64 (build 9200)).
* Function call: makeDataReport(data = merged\_data, mode = "summarize", file = "codebook\_merged\_data.Rmd", checks = list(list("showAllFactorLevels")), listChecks = FALSE, maxProbVals = FALSE, codebook = TRUE, reportTitle = "Codebook for merged\_data")

Codebook for MaxResults

Autogenerated data summary from dataMaid

2018-10-15 02:03:46

# Data report overview

The dataset examined has the following dimensions:

|  |  |
| --- | --- |
| Feature | Result |
| Number of observations | 51 |
| Number of variables | 3 |

# Codebook summary table

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Label | Variable | Class | # unique values | Missing | Description |
|  | [**State**](#state) | factor | 51 | 0.00 % |  |
|  | [**Max\_ABV**](#max_abv) | numeric | 27 | 0.00 % |  |
|  | [**Max\_IBU**](#max_ibu) | numeric | 30 | 0.00 % |  |

# Variable list

## State

* The variable is a key (distinct values for each observation).

## Max\_ABV

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 27 |
| Median | 0.09 |
| 1st and 3rd quartiles | 0.08; 0.1 |
| Min. and max. | 0.06; 0.13 |

## Max\_IBU

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 30 |
| Median | 90 |
| 1st and 3rd quartiles | 75; 110.5 |
| Min. and max. | -Inf; 138 |

Report generation information:

* Created by Landcruiser87 (username: andyh).
* Report creation time: Mon Oct 15 2018 02:03:46
* Report Was run from directory: C:/Users/andyh/OneDrive/Documents
* dataMaid v1.1.2 [Pkg: 2018-05-03 from CRAN (R 3.5.1)]
* R version 3.5.1 (2018-07-02).
* Platform: x86\_64-w64-mingw32/x64 (64-bit)(Windows >= 8 x64 (build 9200)).
* Function call: makeDataReport(data = MaxResults, mode = "summarize", file = "codebook\_MaxResults.Rmd", checks = list(list("showAllFactorLevels")), listChecks = FALSE, maxProbVals = FALSE, codebook = TRUE, reportTitle = "Codebook for MaxResults")

Codebook for MedianResults

Autogenerated data summary from dataMaid

2018-10-15 02:03:08

# Data report overview

The dataset examined has the following dimensions:

|  |  |
| --- | --- |
| Feature | Result |
| Number of observations | 51 |
| Number of variables | 3 |

# Codebook summary table

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Label | Variable | Class | # unique values | Missing | Description |
|  | [**State**](#state) | factor | 51 | 0.00 % |  |
|  | [**Median\_ABV**](#median_abv) | numeric | 17 | 0.00 % |  |
|  | [**Median\_IBU**](#median_ibu) | numeric | 35 | 1.96 % |  |

# Variable list

## State

* The variable is a key (distinct values for each observation).

## Median\_ABV

|  |  |
| --- | --- |
| 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 |

## Median\_IBU

|  |  |
| --- | --- |
| Feature | Result |
| Variable type | numeric |
| Number of missing obs. | 1 (1.96 %) |
| Number of unique values | 34 |
| Median | 35 |
| 1st and 3rd quartiles | 30; 42.75 |
| Min. and max. | 19; 61 |

Report generation information:

* Created by Landcruiser87 (username: andyh).
* Report creation time: Mon Oct 15 2018 02:03:08
* Report Was run from directory: C:/Users/andyh/OneDrive/Documents
* dataMaid v1.1.2 [Pkg: 2018-05-03 from CRAN (R 3.5.1)]
* R version 3.5.1 (2018-07-02).
* Platform: x86\_64-w64-mingw32/x64 (64-bit)(Windows >= 8 x64 (build 9200)).
* Function call: makeDataReport(data = MedianResults, mode = "summarize", file = "codebook\_MedianResults.Rmd", checks = list(list("showAllFactorLevels")), listChecks = FALSE, maxProbVals = FALSE, codebook = TRUE, reportTitle = "Codebook for MedianResults")