

ggplot2::txhousing

Autogenerated data summary from dataReporter

2021-10-21 04:36:14

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	8602
Number of variables	9

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.

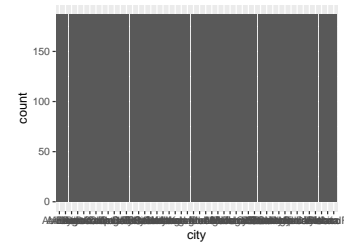
Summary table

	Variable class	# unique values	Missing observations	Any problems?
city	character	46	0.00 %	
year	integer	16	0.00 %	
month	integer	12	0.00 %	
sales	numeric	1712	6.60 %	×
volume	numeric	7495	6.60 %	×
median	numeric	1538	7.16 %	×
listings	numeric	3703	16.55 %	×
inventory	numeric	296	17.05 %	×
date	numeric	187	0.00 %	

Variable list

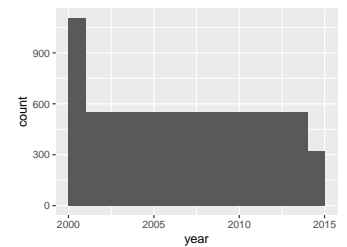
city

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	46
Mode	“Abilene”



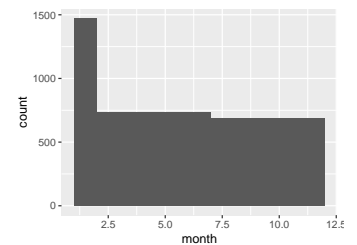
year

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	16
Median	2007
1st and 3rd quartiles	2003; 2011
Min. and max.	2000; 2015



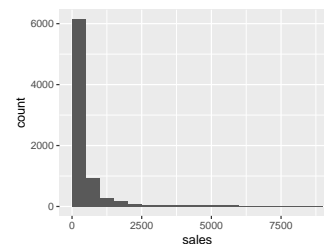
month

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	12
Median	6
1st and 3rd quartiles	3; 9
Min. and max.	1; 12



sales

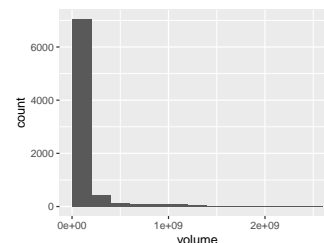
Feature	Result
Variable type	numeric
Number of missing obs.	568 (6.6 %)
Number of unique values	1711
Median	169
1st and 3rd quartiles	86; 467
Min. and max.	6; 8945



- Note that the following possible outlier values were detected: "6", "9", "10", "11", "12", ..., "8439", "8449", "8468", "8628", "8945" (289 values omitted).

volume

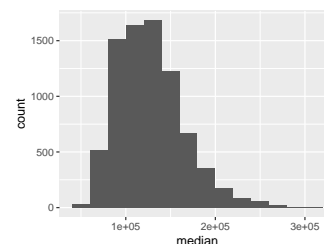
Feature	Result
Variable type	numeric
Number of missing obs.	568 (6.6 %)
Number of unique values	7494
Median	22986824
1st and 3rd quartiles	10840000; 75121388.75
Min. and max.	835000; 2568156780



- Note that the following possible outlier values were detected: "835000", "1018825", "1110000", "1156999", "1165000", ..., "2195184825", "2278932511", "2342443127", "2490238594", "2568156780" (689 values omitted).

median

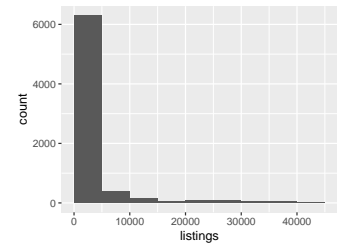
Feature	Result
Variable type	numeric
Number of missing obs.	616 (7.16 %)
Number of unique values	1537
Median	123800
1st and 3rd quartiles	1e+05; 150000
Min. and max.	50000; 304200



- Note that the following possible outlier values were detected: "246800", "246900", "247900", "248500", "248900", ..., "285800", "291400", "292600", "300400", "304200" (45 values omitted).

listings

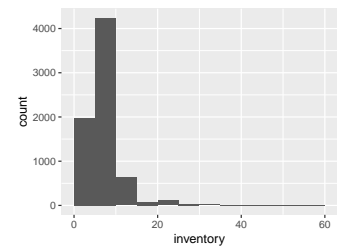
Feature	Result
Variable type	numeric
Number of missing obs.	1424 (16.55 %)
Number of unique values	3702
Median	1283
1st and 3rd quartiles	682; 2953.75
Min. and max.	0; 43107



- Note that the following possible outlier values were detected: "0", "35", "54", "61", "62", ..., "42332", "42766", "42919", "43038", "43107" (375 values omitted).

inventory

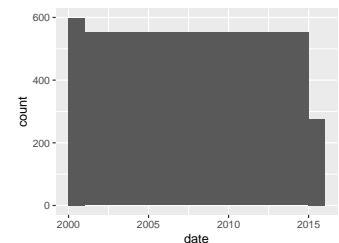
Feature	Result
Variable type	numeric
Number of missing obs.	1467 (17.05 %)
Number of unique values	295
Median	6.2
1st and 3rd quartiles	4.9; 8.15
Min. and max.	0; 55.9



- Note that the following possible outlier values were detected: "0", "0.8", "1", "1.1", "1.2", ..., "52", "54.5", "54.9", "55.7", "55.9" (149 values omitted).

date

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	187
Median	2007.75
1st and 3rd quartiles	2003.83; 2011.67
Min. and max.	2000; 2015.5



Report generation information:

- Created by: Priyanka Gagneja (username: `pgagneja`).
- Report creation time: Thu Oct 21 2021 04:36:15
- Report was run from directory: `/Users/pgagneja/OneDrive - Infor/Documents/Work/Side Projects/EDA_Workflow`
- dataReporter v1.0.0 [Pkg: 2020-12-14 from CRAN (R 4.1.0)]
- R version 4.1.0 (2021-05-18).
- Platform: x86_64-apple-darwin17.0 (64-bit)(macOS Mojave 10.14.6).
- Function call: `dataReporter::makeDataReport(data = ggplot2::txhousing)`