Codebook for agency_data

Autogenerated data summary from dataMaid

2019-08-29 12:18:48

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	2376
Number of variables	14

Codebook summary table

Label	Variable	Class	# unique values	Missing	Description
	account_name	factor	2375	0.00 %	
	account_type	factor	2	0.00 %	
	assigned_agent	factor	6	0.00 %	
	branch_name	factor	1	0.00 %	
	lob	factor	20	0.00 %	
	master_company	factor	20	0.00 %	
	effective_date	Date	752	0.00 %	
	policy_term	factor	3	0.00 %	
	policy_type	factor	2	0.00 %	
	annual_premium	numeric	1966	0.00 %	
	written_premium	numeric	1717	0.00 %	
	rating_state	factor	8	0.00 %	
	status	factor	2	0.00 %	
	transaction_type	factor	13	0.00 %	

Variable list

account_name

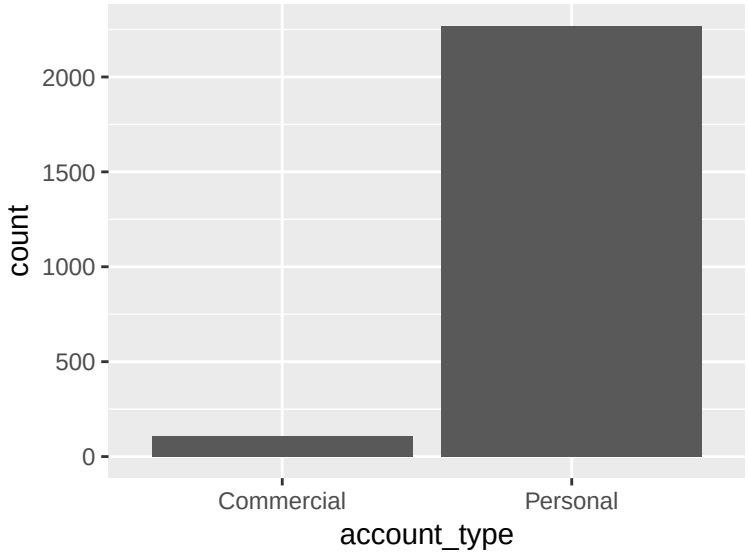
The account name is the name of the insured.

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2375
Mode	“Lucius Schoen”
Reference category	Abdul Kiehn

account_type

The account type really indicates the type of business, commercial lines or personal lines.

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“Personal”
Reference category	Commercial



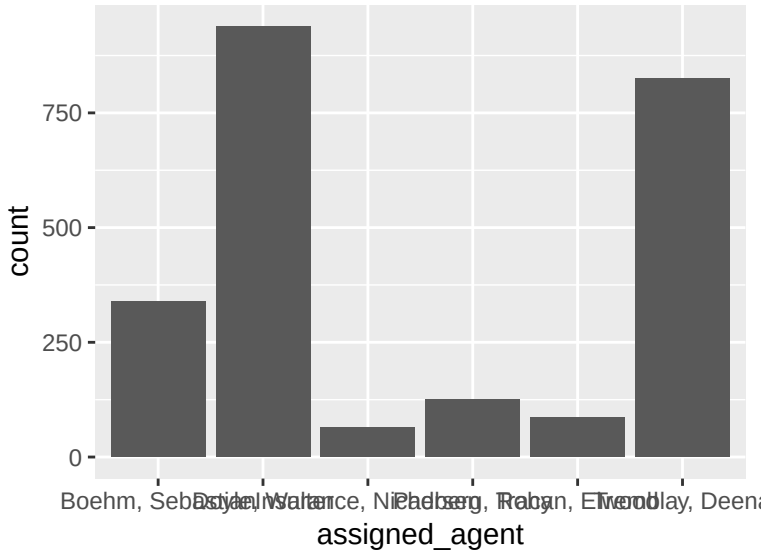
- Observed factor levels: "Commercial", "Personal".

---

assigned_agent

The name of the agency’s agent who manages the account.

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	6
Mode	“Doyle, Walter”
Reference category	Boehm, Sebastian



- Observed factor levels: "Boehm, Sebastian", "Doyle, Walter", "Insurance, Nichersen", "Padberg, Tracy", "Rohan, Elwood", "Tremblay, Deena".

---

branch_name

The name of the insurance agency from which the entire data set originated.

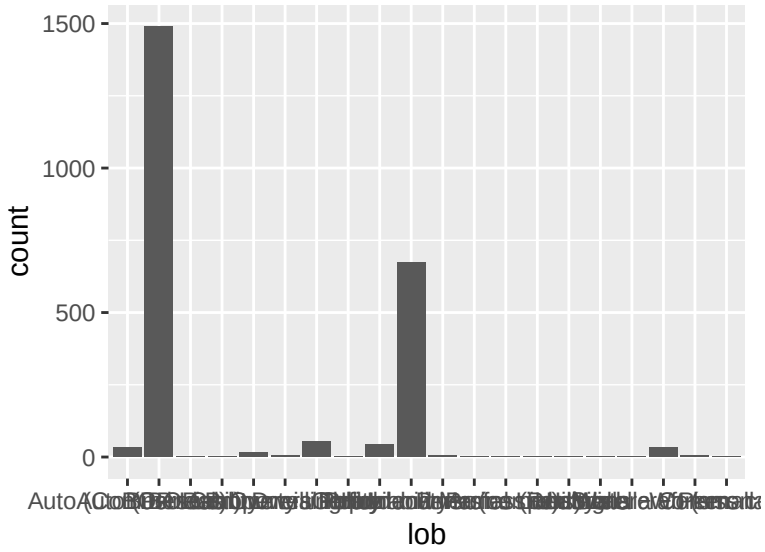
- The variable only takes one (non-missing) value: "Acme Group Insurance". The variable contains 0 % missing observations.

---

lob

Stands for line of business, and in insurance, there are many (many) lines and sub-lines. The agency doesn't have too many LOBs.

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	20
Mode	"Auto (Personal)"
Reference category	Auto (Commercial)



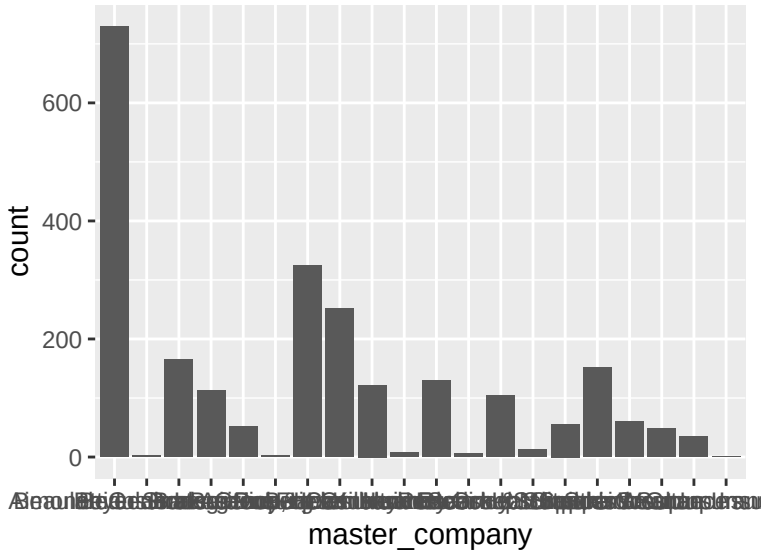
- Observed factor levels: "Auto (Commercial)", "Auto (Personal)", "BOP Liability", "BOP Property", "Business Owners Policy", "Commercial Prpty", "Dwelling fire", "Flood", "Genl Liability", "Homeowners", "Inland marine (comm)", "Inland Marine (pers)", "Personal Liability", "Professional Liab", "Renters", "Signs", "Umbrella - Comm", "Umbrella - Personal", "Watercraft (small boat)", "Workers comp".

---

master_company

The master company is the actual insurer that was used for a given transaction.

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	20
Mode	“Aimonetti Insurance”
Reference category	Aimonetti Insurance



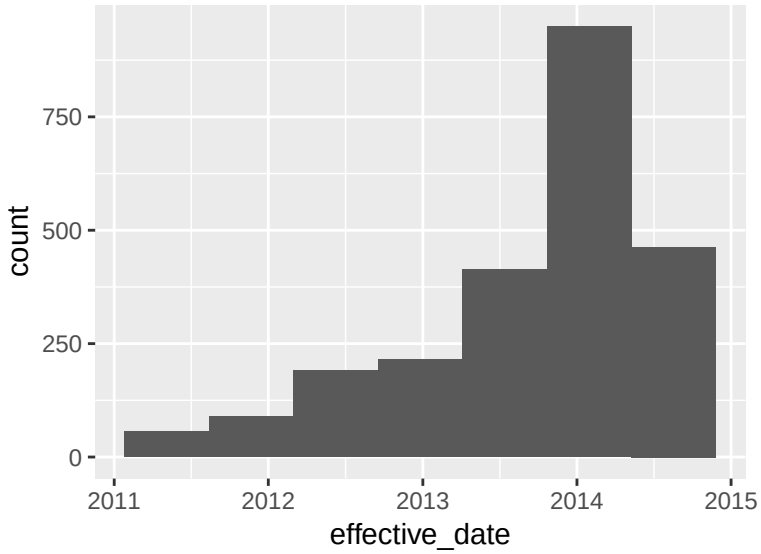
- Observed factor levels: "Aimonetti Insurance", "Beaulah General Agency", "Beaulah Insurance", "Beyer Strategic Insurance", "Borer Group, LLC", "Brakus Fire & Casualty", "Breitenberg Insurance", "DuBuque Insurance", "Friesen Insurance", "Helen Herman Group", "Hui Insurance", "Kilback Excess & Surplus", "Roderick Insurance", "Ryan Hyatt Insurance", "Schimel Mutual Insurance", "Schuppe Group", "Simonis Insurance", "Spinka West Insurance", "Streich Group Insurance", "Sultan Insurance".

---

effective_date

This is the data the insurance goes into effect and is *not* the same as "today" or whatever date the transaction was created. It's not uncommon to have future dated policies (and in rare cases even back dated policies).

Feature	Result
Variable type	Date
Number of missing obs.	0 (0 %)
Number of unique values	752
Mode	"2014-04-01"
Min. and max.	2011-02-01; 2014-10-07
1st and 3rd quartiles	2013-05-08; 2014-04-21

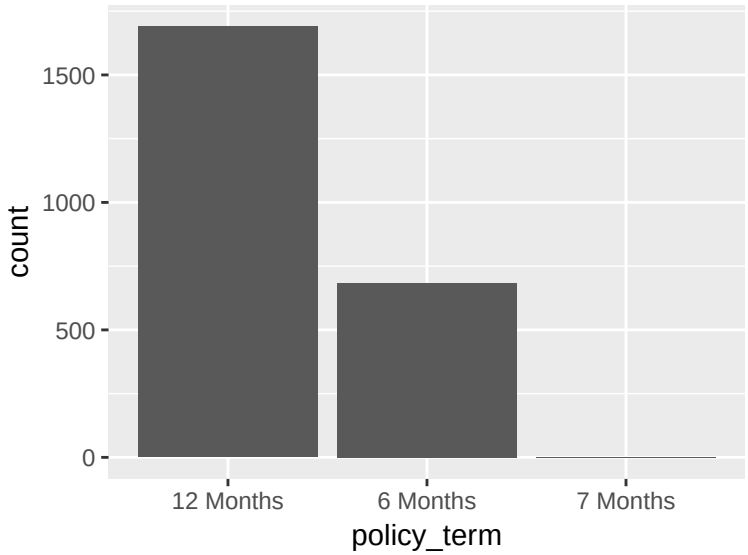


---

policy_term

The length of time the policy is written for.

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	3
Mode	“12 Months”
Reference category	12 Months



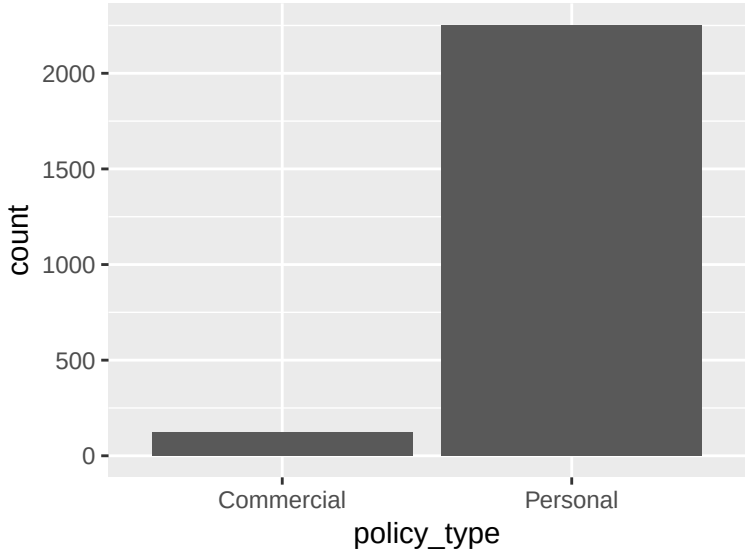
- Observed factor levels: "12 Months", "6 Months", "7 Months".

---

policy_type

Policy type is very much like account type in many instances but for some accounts, there are both personal and commercial policies, necessitating two variables.

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“Personal”
Reference category	Commercial



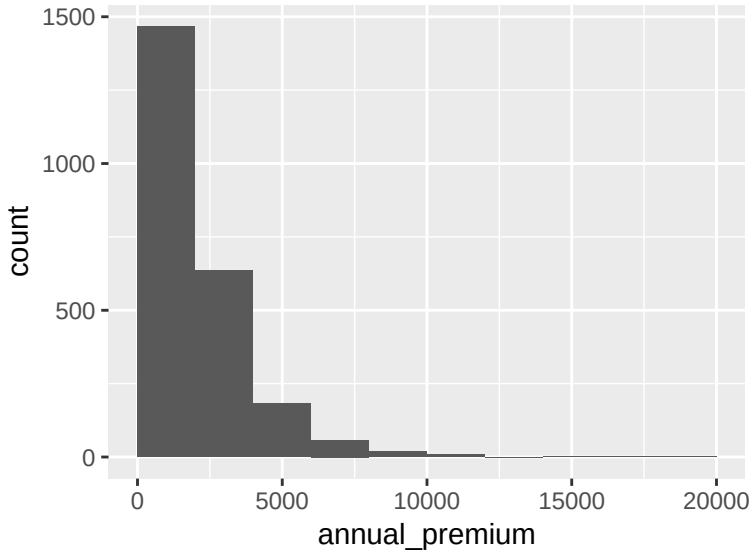
- Observed factor levels: "Commercial", "Personal".

---

annual_premium

This is the annualized premium, which in some cases is exactly equal to the written premium. However, since some policy are written for less than a year, and other transactions (e.g., policy changes) result in premium for less than a year, the annual and written premium differ. This value is always what the premium would be over the course of an entire year.

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	1966
Median	1559.34
1st and 3rd quartiles	812; 2728.72
Min. and max.	43.04; 19002.17

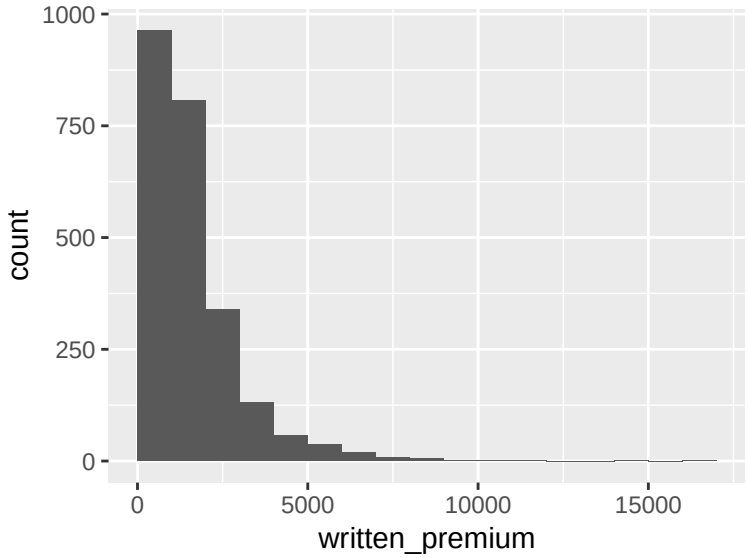


---

written_premium

As noted for `annual_premium`, the written premium is how much the client was actually charged. This amount is pro-rated for the policy term or the amount of the policy term remaining.

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	1717
Median	1218.73
1st and 3rd quartiles	724.97; 2022.17
Min. and max.	43.04; 16504.1

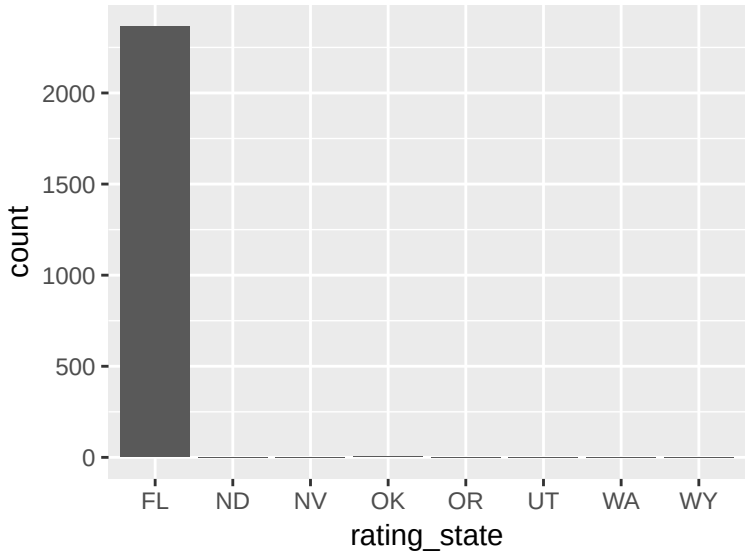


---

rating_state

The state where the insured is at, which is the insured state.

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	8
Mode	“FL”
Reference category	FL



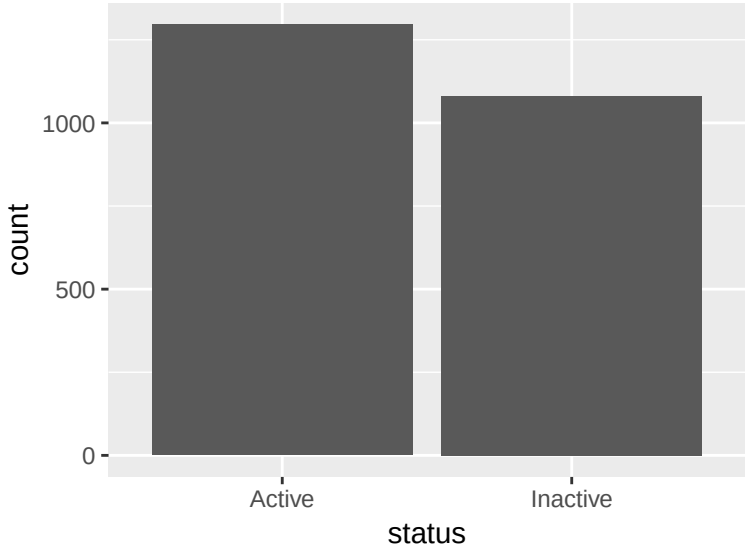
- Observed factor levels: "FL", "ND", "NV", "OK", "OR", "UT", "WA", "WY".

---

status

Status indicates the policy / account / insured / transaction status.

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“Active”
Reference category	Active



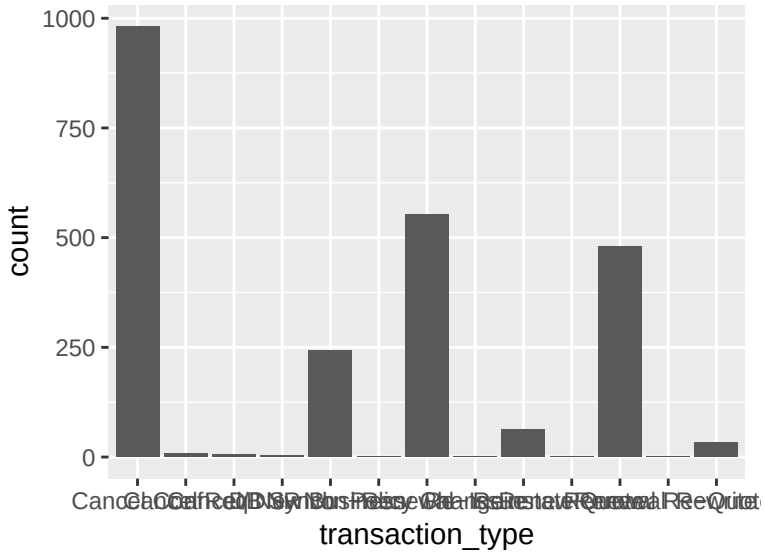
- Observed factor levels: "Active", "Inactive".

---

transaction_type

The transaction shows what was actually done within the agency management system.

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	13
Mode	“Cancel Conf”
Reference category	Cancel Conf



- Observed factor levels: "Cancel Conf", "Cancel Req", "Cancel/DNR", "DB Synch", "New Business", "Non-Renewal", "Policy Change", "Re-Issue", "Reinstate", "Renew Quote", "Renewal", "Renewal Re-Quote", "Rewrite".

Report generation information:

- Created by Frank Neugebauer (username: `neugg`).
- Report creation time: Thu Aug 29 2019 12:18:49
- Report was run from directory: `C:/Users/neugg/OneDrive/Documents/GitHub/agency_data/data`
- dataMaid v1.3.2 [Pkg: 2019-07-27 from CRAN (R 3.6.1)]
- R version 3.6.0 (2019-04-26).
- Platform: x86_64-w64-mingw32/x64 (64-bit)(Windows 10 x64 (build 17134)).
- Function call: `makeDataReport(data = agency_data, mode = c("summarize", "visualize", "check"), smartNum = FALSE, file = "codebook_agency_data.Rmd", checks = list(character = "showAllFactorLevels", factor = "showAllFactorLevels", labelled = "showAllFactorLevels", haven_labelled = "showAllFactorLevels", numeric = NULL, integer = NULL, logical = NULL, Date = NULL), listChecks = FALSE, maxProbVals = Inf, codebook = TRUE, reportTitle = "Codebook for agency_data")`