articles_programs_used

Autogenerated data summary from dataMaid

2018-10-10 09:21:01

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	1462
Number of variables	25

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 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.

Codebook summary table

Label	Variable	Class	# unique values	Missing	Description
	compressed_file_path	character	1462	0.00 %	
	article_url	character	1462	0.00 %	
	issue_name	character	112	0.00 %	
	issue_url	character	112	0.00 %	
	data_text	character	1371	0.00 %	
	data_url	character	1462	0.00 %	
	year	integer	18	0.07 %	
	author	character	1389	0.00 %	
	title	character	1460	0.00 %	
	journal	character	1	0.00 %	
	vol	character	1462	0.00 %	
	doi	character	1462	0.00 %	

Label	Variable	Class	# unique values	Missing	Description
	doi_clean	character	1462	0.00 %	
	data_url_fixed	character	1462	0.00 %	
	size_string	character	1371	0.00 %	
	size_bytes	numeric	1371	0.00 %	
	stata	logical	2	0.00 %	
	julia	logical	2	0.00 %	
	python	logical	2	0.00 %	
	R	logical	2	0.00 %	
	C	logical	2	0.00 %	
	C++	logical	1	0.00 %	
	matlab	logical	2	0.00 %	
	fortran	logical	2	0.00 %	
	sas	logical	2	0.00 %	

Variable list

compressed_file_path

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

---

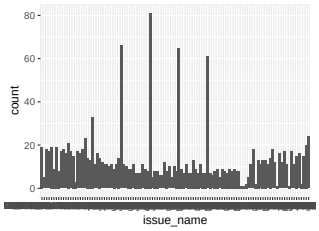
article_url

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

---

issue_name

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	112
Mode	"Vol. 105 No. 5 May 2015"

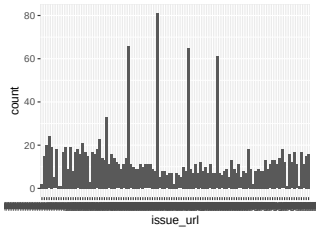


- Note that the following levels have at most five observations: "Vol. 100 No. 2 May 2010", "Vol. 102 No. 3 May 2012", "Vol. 105 No. 6 June 2015", "Vol. 106 No. 2 February 2016", "Vol. 108 No. 1 January 2018", "Vol. 89 No. 3 June 1999", "Vol. 90 No. 3 June 2000", "Vol. 92 No. 5 December 2002", "Vol. 94 No. 4 September 2004", "Vol. 95 No. 2 May 2005" (3 additional values omitted).

---

issue_url

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	112
Mode	“https://www.aeaweb.org/issues/373”

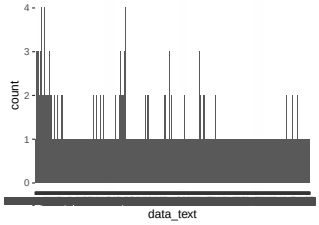


- Note that the following levels have at most five observations: "https://www.aeaweb.org/issues/100/“, "https://www.aeaweb.org/issues/127/”, "https://www.aeaweb.org/issues/153/“, "https://www.aeaweb.org/issues/158/”, "https://www.aeaweb.org/issues/247/“, "https://www.aeaweb.org/issues/374/”, "https://www.aeaweb.org/issues/39/“, "https://www.aeaweb.org/issues/399/”, "https://www.aeaweb.org/issues/48/“, "https://www.aeaweb.org/issues/491/” (3 additional values omitted).

---

data_text

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	1371
Mode	“Data Set (1.28 MB)”



- Note that the following levels have at most five observations: "Data Set (1,000.40 KB)", "Data Set (1,006.64 KB)", "Data Set (1,013.55 KB)", "Data Set (1,015.66 KB)", "Data Set (1,022.74 KB)", "Data Set (1.01 MB)", "Data Set (1.02 MB)", "Data Set (1.03 MB)", "Data Set (1.04 GB)", "Data Set (1.04 MB)" (1361 additional values omitted).

---

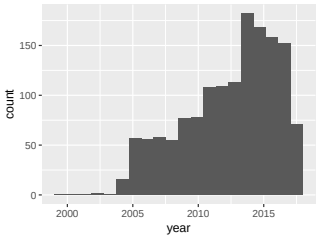
data_url

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

---

year

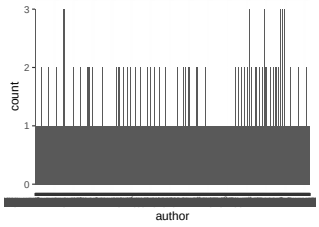
Feature	Result
Variable type	integer
Number of missing obs.	1 (0.07 %)
Number of unique values	17
Median	2014
1st and 3rd quartiles	2010; 2016
Min. and max.	1999; 2018



---

author

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	1389
Mode	“Abaluck, Jason, and Jonathan Gruber.”

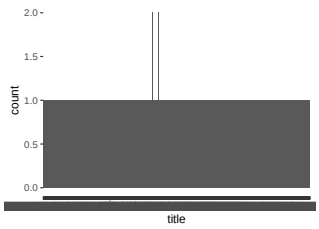


- Note that the following levels have at most five observations: "Aaronson, Daniel, Fabian Lange, and Bhashkar Mazumder.", "Aaronson, Daniel, Sumit Agarwal, and Eric French.", "Abaluck, Jason, and Jonathan Gruber.", "Abaluck, Jason, Leila Agha, Chris Kabrhel, Ali Raja, and Arjun Venkatesh.", "Abbink, Klaus, Jordi Brandts, Benedikt Herrmann, and Henrik Orzen.", "Abdellaoui, Mohammed, Aurélien Baillon, Laetitia Placido, and Peter P. Wakker.", "Abdulkadiroğlu, Atila, Joshua D. Angrist, Peter D. Hull, and Parag A. Pathak.", "Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A. Pathak.", "Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth.", "Abdulkadroğlu, Atila, Joshua D. Angrist, Yusuke Narita, Parag A. Pathak, and Roman A. Zarate." (1379 additional values omitted).

---

title

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	1460
Mode	“”Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Comment.“”



- Note that the following levels have at most five observations: "“ ‘Acting Wife’: Marriage Market Incentives and Labor Market Investments.”", "“”Momma’s Got the Pill“: How Anthony Comstock and Griswold v. Connecticut Shaped US Childbearing.”", "“(Dis)organization and Success in an Economics MOOC.”", "“(Indirect) Input Linkages.”", "“$1,000 Cash Back: The Pass-Through of Auto Manufacturer Promotions.”", "“A Balls-and-Bins Model of Trade: Comment.”", "“A Balls-and-Bins Model of Trade.”", "“A Change Would Do You Good .... An Experimental Study on How to Overcome Coordination Failure in Organizations.”", "“A Comment on the Economics of Labor Adjustment: Mind the Gap: Evidence from a Monte Carlo Experiment: Reply.”", "“A Comment on the Economics of Labor Adjustment: Mind the Gap: Evidence from a Monte Carlo Experiment.”" (1450 additional values omitted).

---

journal

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

---

vol

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

---

doi

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

---

doi_clean

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

---

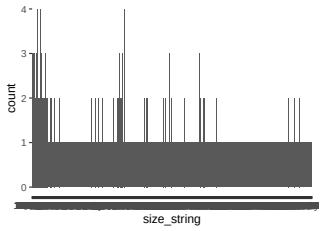
data_url_fixed

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

---

size_string

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	1371
Mode	“1.28 MB”

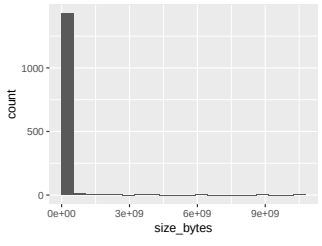


- Note that the following levels have at most five observations: "1.01 MB", "1.02 MB", "1.03 MB", "1.04 GB", "1.04 MB", "1.05 MB", "1.06 GB", "1.06 MB", "1.08 MB", "1.09 MB" (1361 additional values omitted).

---

size_bytes

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	1371
Median	470604.8
1st and 3rd quartiles	98304; 4550819.84
Min. and max.	158; 10769630494.72

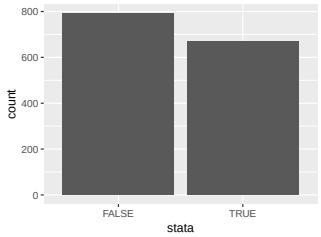


- Note that the following possible outlier values were detected: "97171537.92", "97664368.64", "99845406.72", "101669928.96", "102047416.32", "102162759.68", "104134082.56", "106052976.64", "107342725.12", "107594383.36" (94 additional values omitted).

---

stata

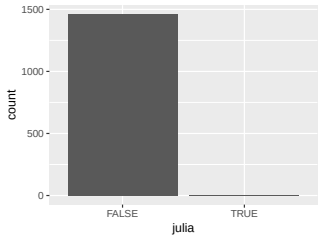
Feature	Result
Variable type	logical
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“FALSE”



---

julia

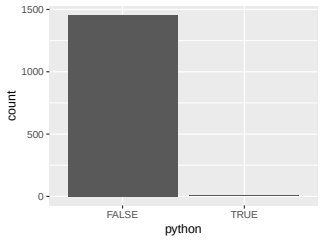
Feature	Result
Variable type	logical
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“FALSE”



---

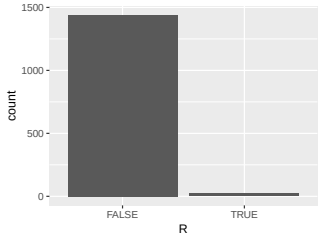
python

Feature	Result
Variable type	logical
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“FALSE”



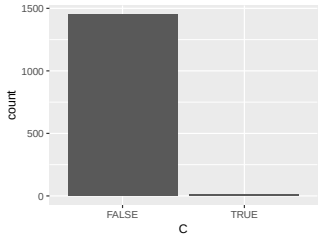
R

Feature	Result
Variable type	logical
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“FALSE”



C

Feature	Result
Variable type	logical
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“FALSE”

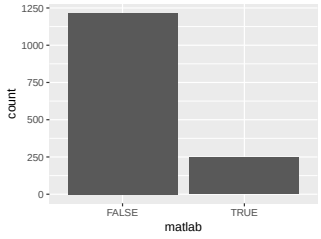


C++

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

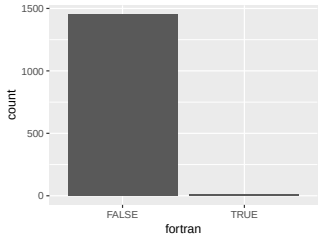
matlab

Feature	Result
Variable type	logical
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“FALSE”



fortran

Feature	Result
Variable type	logical
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“FALSE”



sas

Feature	Result
Variable type	logical
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“FALSE”

Report generation information:

- Created by Lars Vilhuber (username: `vilhuber`).
- Report creation time: Wed Oct 10 2018 09:21:01
- Report Was run from directory: `/mnt/local/slow_home/vilhuber/Workspace-non-encrypted/git/AEA/econ-program`
- 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-suse-linux-gnu (64-bit)(openSUSE Leap 42.3).
- Function call: `makeDataReport(data = articles_programs_used, file = "codebook_articles_programs_used.Rm`
  `replace = TRUE, openResult = FALSE, codebook = TRUE, clean = FALSE)`