

pvp

Autogenerated data summary from dataMaid

2020-11-21 10:57:26

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	2352
Number of variables	8

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.

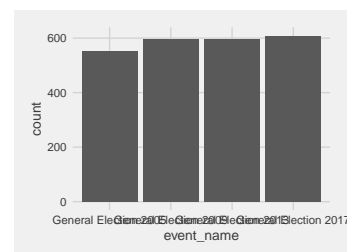
Summary table

	Variable class	# unique values	Missing observations	Any problems?
event_name	character	4	0.00 %	
event_year	numeric	4	0.00 %	
ed_abbreviation	character	132	0.00 %	×
ed_name	character	124	0.00 %	
age_group	character	7	0.00 %	
participation	numeric	1905	0.00 %	×
registered_voters	numeric	2033	0.00 %	×
event_date_text	character	4	0.00 %	

Variable list

event__name

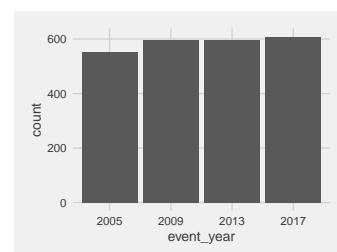
Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	4
Mode	“General Election 2017”



event__year

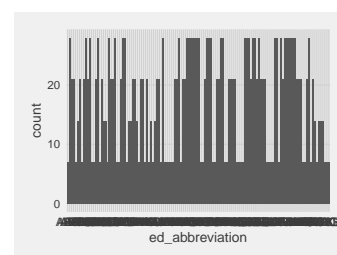
- Note that this variable is treated as a factor variable below, as it only takes a few unique values.

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	4
Mode	“2017”
Reference category	2005



ed__abbreviation

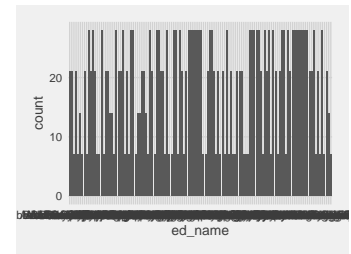
Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	132
Mode	“ABM”



- The following suspected missing value codes enter as regular values: "NAN".

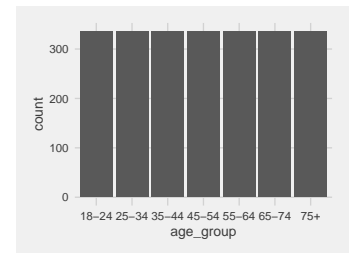
ed_name

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	124
Mode	"Burnaby North"



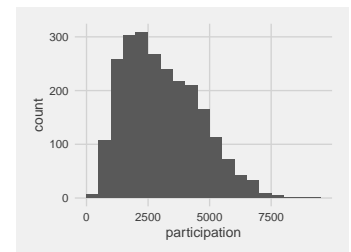
age_group

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	7
Mode	"18-24"



participation

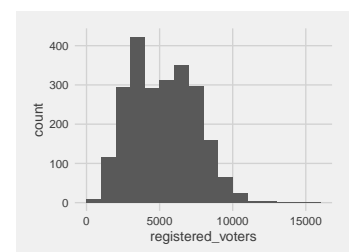
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	1905
Median	2866
1st and 3rd quartiles	1875.75; 4135.5
Min. and max.	417; 9498



- Note that the following possible outlier values were detected: "9498".

registered_voters

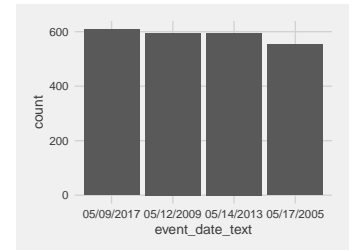
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	2033
Median	5139
1st and 3rd quartiles	3415; 6895.75
Min. and max.	768; 15027



- Note that the following possible outlier values were detected: "12555", "12581", "12688", "13176", "13627", "15027".

event__date__text

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	4
Mode	"05/09/2017"



Report generation information:

- Created by: Rafael Pilliard-Hellwig (username: `raf`).
- Report creation time: Sat Nov 21 2020 10:57:26
- Report was run from directory: `/Users/raf/Dropbox/mds/b3/522_dwf/bc_election_turnout`
- dataMaid v1.4.0 [Pkg: 2019-12-10 from CRAN (R 4.0.0)]
- R version 4.0.3 (2020-10-10).
- Platform: x86_64-apple-darwin17.0 (64-bit)(macOS Catalina 10.15.7).
- Function call: `dataMaid::makeDataReport(data = pvp, file = here::here("eda", "profile_pvp.Rmd"), replace = TRUE)`