趙友誠 H24101060

2024-09-26

Table of contents

```r
library(haven)
library(Hmisc)
pollsav <- read_sav("poll.sav")
write.csv(pollsav, file = "poll.csv", row.names = FALSE)
pollcsv <- read.csv("poll.csv")
```

Brief introduciotn to the data

Dimension of the Data : **1671 samples × 15 columns**

Variables	Explanation
V1、V2、V3	District and Li
V4_1~V4_8	Popularity
V5	Support level
V6	Age
V7	Education level
V8	Sex

```r
str(pollcsv)
```

```
'data.frame':    1671 obs. of  15 variables:
 $ v1  : int  2 2 2 1 1 2 1 2 2 1 ...
 $ v2  : int  99 99 99 9 7 99 10 99 99 9 ...
 $ v3  : int  5 11 5 99 99 44 99 1 44 99 ...
 $ v4_1: int  98 1 4 3 98 98 98 98 3 98 ...
 $ v4_2: int  99 4 99 10 99 99 99 99 99 99 ...
 $ v4_3: int  99 7 99 99 99 99 99 99 99 99 ...
 $ v4_4: int  99 10 99 99 99 99 99 99 99 99 ...
 $ v4_5: int  99 99 99 99 99 99 99 99 99 99 ...
 $ v4_6: int  99 99 99 99 99 99 99 99 99 99 ...
 $ v4_7: int  99 99 99 99 99 99 99 99 99 99 ...
 $ v4_8: int  99 99 99 99 99 99 99 99 99 99 ...
 $ v5  : int  98 7 6 7 99 10 7 99 7 99 ...
 $ v6  : int  5 5 5 5 5 5 5 5 5 3 ...
 $ v7  : int  1 2 5 2 2 3 2 5 1 3 ...
 $ v8  : int  2 1 2 2 2 2 1 2 1 2 ...
```

`summary(pollcsv)`

```
      v1                v2                v3              v4_1
 Min.   :1.000    Min.   : 1.00    Min.   : 1.00    Min.   : 1.00
 1st Qu.:1.000    1st Qu.:18.00    1st Qu.:19.00    1st Qu.: 3.00
 Median :1.000    Median :31.00    Median :99.00    Median :98.00
 Mean   :1.338    Mean   :48.96    Mean   :71.72    Mean   :57.03
 3rd Qu.:2.000    3rd Qu.:99.00    3rd Qu.:99.00    3rd Qu.:98.00
 Max.   :2.000    Max.   :99.00    Max.   :99.00    Max.   :98.00
      v4_2              v4_3              v4_4              v4_5              v4_6
 Min.   : 2.00    Min.   : 3.00    Min.   : 4.00    Min.   : 5.00    Min.   : 6.0
 1st Qu.: 7.00    1st Qu.:99.00    1st Qu.:99.00    1st Qu.:99.00    1st Qu.:99.0
 Median :99.00    Median :99.00    Median :99.00    Median :99.00    Median :99.0
 Mean   :71.89    Mean   :83.64    Mean   :90.61    Mean   :94.78    Mean   :96.4
 3rd Qu.:99.00    3rd Qu.:99.00    3rd Qu.:99.00    3rd Qu.:99.00    3rd Qu.:99.0
 Max.   :99.00    Max.   :99.00    Max.   :99.00    Max.   :99.00    Max.   :99.0
      v4_7              v4_8              v5                v6
 Min.   : 7.00    Min.   : 8.00    Min.   : 1.00    Min.   :1.00
 1st Qu.:99.00    1st Qu.:99.00    1st Qu.: 4.00    1st Qu.:4.00
 Median :99.00    Median :99.00    Median :10.00    Median :5.00
 Mean   :97.93    Mean   :98.73    Mean   :51.11    Mean   :4.29
 3rd Qu.:99.00    3rd Qu.:99.00    3rd Qu.:99.00    3rd Qu.:5.00
 Max.   :99.00    Max.   :99.00    Max.   :99.00    Max.   :6.00
      v7                v8
 Min.   : 1.000    Min.   :1.000
```

```
1st Qu.: 2.000     1st Qu.:1.000
Median : 3.000     Median :2.000
Mean   : 6.871     Mean   :1.592
3rd Qu.: 5.000     3rd Qu.:2.000
Max.   :95.000     Max.   :2.000
```

```
latex(describe(pollcsv), file = "", caption.placement = "top", title = "Public Opinion")
```

```
Warning in min(diff(uxinside)): no non-missing arguments to min; returning Inf
```

```
Warning in spikecomp(x, method = "grid", lumptails = lumptails, normalize =
FALSE, : program logic error in spikecomp; lengths: 2 1
```

pollcsv
15 Variables   1671 Observations

---

v1

	n	missing	distinct	Info	Mean	Gmd
	1671	0	2	0.671	1.338	0.4475

```
Value          1     2
Frequency   1107   564
Proportion 0.662 0.338
```

---

v2

	n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
	1671	0	36	0.96	48.96	40.81	7	8	18	31	99	99	99

```
lowest :  1  2  3  4  5, highest: 32 33 44 98 99
```

---

v3

	n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
	1671	0	23	0.709	71.72	38.21	4	6	19	99	99	99	99

```
lowest :  1  2  3  4  5, highest: 19 20 44 98 99
```

---

v4_1

	n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
	1671	0	12	0.813	57.03	47.11	1	1	3	98	98	98	98

```
Value           1     2     3     4     5     6     7     8     9    10    91    98
Frequency     328     5   214    43    27    38    47     4     1    11    14   939
Proportion 0.196 0.003 0.128 0.026 0.016 0.023 0.028 0.002 0.001 0.007 0.008 0.562

For the frequency table, variable is rounded to the nearest 0
```

---

v4_2

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1671	0	10	0.637	71.89	38.81	3	3	7	99	99	99	99

Value	2	3	4	5	6	7	8	9	10	99
Frequency	6	189	59	32	75	99	2	4	15	1190
Proportion	0.004	0.113	0.035	0.019	0.045	0.059	0.001	0.002	0.009	0.712

For the frequency table, variable is rounded to the nearest 0

v4_3

n	missing	distinct	Info	Mean	Gmd
1671	0	9	0.418	83.64	25.71

Value	3	4	5	6	7	8	9	10	99
Frequency	6	60	36	61	91	1	2	19	1395
Proportion	0.004	0.036	0.022	0.037	0.054	0.001	0.001	0.011	0.835

For the frequency table, variable is rounded to the nearest 0

v4_4

n	missing	distinct	Info	Mean	Gmd
1671	0	8	0.249	90.61	15.28

Value	4	5	6	7	8	9	10	99
Frequency	4	28	41	52	3	4	20	1519
Proportion	0.002	0.017	0.025	0.031	0.002	0.002	0.012	0.909

For the frequency table, variable is rounded to the nearest 0

v4_5

n	missing	distinct	Info	Mean	Gmd
1671	0	7	0.132	94.78	8.056

Value	5	6	7	8	9	10	99
Frequency	3	14	38	4	3	15	1594
Proportion	0.002	0.008	0.023	0.002	0.002	0.009	0.954

For the frequency table, variable is rounded to the nearest 0

v4_6

n	missing	distinct	Info	Mean	Gmd
1671	0	6	0.084	96.4	5.048

Value	6	7	8	9	10	99
Frequency	3	12	6	7	20	1623
Proportion	0.002	0.007	0.004	0.004	0.012	0.971

For the frequency table, variable is rounded to the nearest 0

v4_7

n	missing	distinct	Info	Mean	Gmd
1671	0	5	0.035	97.93	2.125

Value	7	8	9	10	99
Frequency	3	2	3	12	1651
Proportion	0.002	0.001	0.002	0.007	0.988

For the frequency table, variable is rounded to the nearest 0

---

v4_8

n	missing	distinct	Info	Mean	Gmd
1671	0	3	0.009	98.73	0.5337

Value	8	10	99
Frequency	1	4	1666
Proportion	0.001	0.002	0.997

For the frequency table, variable is rounded to the nearest 0

---

v5

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1671	0	13	0.956	51.11	47.91	1	3	4	10	99	99	99

Value	1	2	3	4	5	6	7	8	9	10	91	98	99
Frequency	158	9	205	79	33	98	195	6	8	53	10	269	548
Proportion	0.095	0.005	0.123	0.047	0.020	0.059	0.117	0.004	0.005	0.032	0.006	0.161	0.328

For the frequency table, variable is rounded to the nearest 0

---

v6

n	missing	distinct	Info	Mean	Gmd
1671	0	6	0.808	4.29	1.082

Value	1	2	3	4	5	6
Frequency	52	94	201	336	946	42
Proportion	0.031	0.056	0.120	0.201	0.566	0.025

For the frequency table, variable is rounded to the nearest 0

---

v7

n	missing	distinct	Info	Mean	Gmd
1671	0	6	0.945	6.871	8.373

Value	1	2	3	4	5	95
Frequency	292	165	431	198	520	65
Proportion	0.175	0.099	0.258	0.118	0.311	0.039

For the frequency table, variable is rounded to the nearest 0

v8

```
       n   missing  distinct    Info   Mean     Gmd
    1671         0         2   0.725  1.592  0.4834
```

```
Value          1     2
Frequency    682   989
Proportion 0.408 0.592
```