

Lecture 5 Notes

2024-01-30

Opening data and looking at it

```
require(tidyverse)
```

```
## Loading required package: tidyverse
```

```
## — Attaching packages — tidyverse 1.3.2 —
## ✓ ggplot2 3.3.6      ✓ purrr 0.3.4
## ✓ tibble 3.1.8       ✓ dplyr 1.0.10
## ✓ tidyr 1.2.1        ✓ stringr 1.4.1
## ✓ readr 2.1.2        ✓ forcats 0.5.2
## — Conflicts — tidyverse_conflicts() —
## ✗ dplyr::filter() masks stats::filter()
## ✗ dplyr::lag() masks stats::lag()
```

```
require(labelled)
```

```
## Loading required package: labelled
```

```
## Warning: package 'labelled' was built under R version 4.2.3
```

```
MI_raw <- read_rds("https://github.com/jbisbee1/DS1000_S2024/raw/main/data/MI2020_ExitPo
ll.rds")

view(MI_raw)
```

Ways of looking at data

```
# Simplest: look
MI_raw
```

```
## # A tibble: 1,231 × 63
##       ID WEIGHT LALMI6290... GROUP ZIP DISTR...1 Z1 S1 S2A
##   <dbl> <dbl> <chr> <dbl+lbl> <dbl> <dbl> <dbl> <dbl+lbl> <dbl+lbl>
## 1     9  0.405 LALMI6290... 3 [3] 49327 2 NA 1 [Yes] 2 [No]
## 2    66  1.81 LALMI2492... 1 [1] 48234 14 NA 1 [Yes] 1 [Yes]
## 3   225  0.860 LALMI5489... 4 [4] 48301 9 48322 1 [Yes] 1 [Yes]
## 4   243  0.199 LALMI5053... 1 [1] 48130 7 48130 1 [Yes] 1 [Yes]
## 5   286  0.177 LALMI6831... 1 [1] 49946 1 NA 1 [Yes] 1 [Yes]
## 6   293  0.492 LALMI4019... 1 [1] 48615 4 NA 1 [Yes] 1 [Yes]
## 7   365  1.37 LALMI4151... 1 [1] 48906 4 48813 1 [Yes] 1 [Yes]
## 8   367  1.15 LALMI5912... 1 [1] 49442 2 NA 1 [Yes] 1 [Yes]
## 9   388  1.50 LALMI6635... 1 [1] 48451 5 NA 1 [Yes] 2 [No]
## 10  417  1.30 LALMI3567... 1 [1] 48197 12 NA 1 [Yes] 1 [Yes]
## # ... with 1,221 more rows, 54 more variables: S2B <dbl+lbl>, S3 <dbl+lbl>,
## # S4 <dbl+lbl>, VERSION <dbl+lbl>, PRSMI20 <dbl+lbl>, SENMI20 <dbl+lbl>,
## # TIME16 <dbl+lbl>, ISSUE20 <dbl+lbl>, QLT20 <dbl+lbl>, TEMPBIDEN <dbl+lbl>,
## # TEMPTRUMP <dbl+lbl>, CONTROLSEN <dbl+lbl>, FINSIT <dbl+lbl>,
## # ECONVCORONA20 <dbl+lbl>, FAVBIDEN2 <dbl+lbl>, FAVTRUMP <dbl+lbl>,
## # FORCAND <dbl+lbl>, NEWVOTER <dbl+lbl>, NEC <dbl+lbl>,
## # HANDLEECON20 <dbl+lbl>, HANDLECORONA20 <dbl+lbl>, RACISM20 <dbl+lbl>, ...
```

```
# More advanced: glimpse()
glimpse(MI_raw)
```

```

## Rows: 1,231
## Columns: 63
## $ ID <dbl> 9, 66, 225, 243, 286, 293, 365, 367, 388, 417, 563, 572...
## $ WEIGHT <dbl> 0.4045421, 1.8052619, 0.8601966, 0.1991648, 0.1772090, ...
## $ LALVOTERID <chr> "LALMI6290066", "LALMI2492492", "LALMI548981440", "LALM...
## $ GROUP <dbl+lbl> 3, 1, 4, 1, 1, 1, 1, 1, 1, 1, 1, 1, 9, 1, 4, 2, 1, ...
## $ ZIP <dbl> 49327, 48234, 48301, 48130, 49946, 48615, 48906, 49442,...
## $ DISTRICT <dbl> 2, 14, 9, 7, 1, 4, 4, 2, 5, 12, 7, 2, 5, 9, 1, 5, 8, 2,...
## $ Z1 <dbl> NA, NA, 48322, 48130, NA, NA, 48813, NA, NA, NA, NA, NA...
## $ S1 <dbl+lbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ S2A <dbl+lbl> 2, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 2, 1, 1, 2, 2, 1, 2, ...
## $ S2B <dbl+lbl> NA, 1, 1, 1, 2, 1, 1, 1, NA, 1, NA, 1, 1, 1,...
## $ S3 <dbl+lbl> 1, NA, NA, NA, NA, NA, NA, NA, NA, 1, NA, 1, NA, NA,...
## $ S4 <dbl+lbl> 1, NA, NA, NA, NA, NA, NA, NA, NA, 1, NA, 1, NA, NA,...
## $ VERSION <dbl+lbl> 1, 2, 2, 1, 1, 2, 2, 2, 1, 1, 1, 1, 1, 2, 1, 2, 2, ...
## $ PRSMI20 <dbl+lbl> 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 2, 2, 1, 2, 1, 1, 2, ...
## $ SENMI20 <dbl+lbl> 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 2, 2, 1, 2, 1, 1, 2, ...
## $ TIME16 <dbl+lbl> 5, NA, NA, 5, 5, NA, NA, NA, 5, 5, 5, 5, 5, 5,...
## $ ISSUE20 <dbl+lbl> 5, NA, NA, 9, 1, NA, NA, NA, 4, 9, 3, 3, 1,...
## $ QLT20 <dbl+lbl> 4, NA, NA, 4, 3, NA, NA, NA, 3, 3, 2, 2, 3,...
## $ TEMPBIDEN <dbl+lbl> 1, NA, NA, 1, 1, NA, NA, NA, 2, 1, 2, 2, 1,...
## $ TEMPTRUMP <dbl+lbl> 2, NA, NA, 2, 2, NA, NA, NA, 1, 2, 1, 1, 2,...
## $ CONTROLSEN <dbl+lbl> 9, NA, NA, 1, 1, NA, NA, NA, 2, 1, 2, 2, 9,...
## $ FINSIT <dbl+lbl> 3, NA, NA, 3, 3, NA, NA, NA, 3, 3, 3, 3, 2,...
## $ ECONVCORONA20 <dbl+lbl> 1, NA, NA, 1, 1, NA, NA, NA, 2, 1, 9, 2, 1,...
## $ FAVBIDEN2 <dbl+lbl> 1, NA, NA, 1, 1, NA, NA, NA, 2, 1, 2, 2, 1,...
## $ FAVTRUMP <dbl+lbl> 2, NA, NA, 2, 2, NA, NA, NA, 1, 2, 1, 1, 2,...
## $ FORCAND <dbl+lbl> NA, 1, 1, NA, NA, 1, 2, 1, NA, NA, NA, NA, NA,...
## $ NEWVOTER <dbl+lbl> NA, 2, 2, NA, NA, 2, 2, 2, NA, NA, NA, NA, NA,...
## $ NEC <dbl+lbl> NA, 4, 4, NA, NA, 4, 4, 4, NA, NA, NA, NA, NA,...
## $ HANDLEECON20 <dbl+lbl> NA, 1, 1, NA, NA, 1, 1, 1, NA, NA, NA, NA, NA,...
## $ HANDLECORONA20 <dbl+lbl> NA, 1, 1, NA, NA, 1, 1, 1, NA, NA, NA, NA, NA,...
## $ RACISM20 <dbl+lbl> NA, 2, 2, NA, NA, 2, 2, 2, NA, NA, NA, NA, NA,...
## $ VOTE2016 <dbl+lbl> NA, 1, 1, NA, NA, 1, 1, 1, NA, NA, NA, NA, NA,...
## $ COUNTACC <dbl+lbl> NA, 2, 2, NA, NA, 1, 2, 1, NA, NA, NA, NA, NA,...
## $ TRUMP4 <dbl+lbl> NA, 4, 4, NA, NA, 4, 4, 4, NA, NA, NA, NA, NA,...
## $ CONTAINCOVID <dbl+lbl> NA, 4, 4, NA, NA, 4, 3, 4, NA, NA, NA, NA, NA,...
## $ COVIDHARDSHIP <dbl+lbl> NA, 3, 1, NA, NA, 2, 2, 2, NA, NA, NA, NA, NA,...
## $ AGE10 <dbl+lbl> 2, 10, 7, 9, 8, 7, 9, 8, 6, 8, 9, 10, 1,...
## $ SEX <dbl+lbl> 2, 2, 2, 1, 2, 2, 1, 1, 2, 1, 1, 1, 2, 1, 1, 2, 1, ...
## $ EDUC18 <dbl+lbl> 4, 1, 5, 4, 5, 3, 3, 3, 4, 4, 5, 5, 4, 1, 1, 1, 5, ...
## $ QRACEAI <dbl+lbl> 1, 2, 1, 1, 1, 1, 1, 1, 1, 2, 9, 1, 1, 1, 1, 1, 3, ...
## $ LATINOS <dbl+lbl> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, ...
## $ PARTYID <dbl+lbl> 3, 1, 1, 3, 3, 3, 1, 1, 2, 1, 3, 2, 4, 4, 1, 1, 3, ...
## $ PHIL3 <dbl+lbl> 2, 2, 1, 9, 1, 2, 9, 2, 3, 2, 3, 3, 1, 3, 9, 2, 2, ...
## $ INCOME20 <dbl+lbl> 3, NA, NA, 9, 4, NA, NA, NA, 4, 4, 4, 2, 1,...
## $ BRNAGAIN <dbl+lbl> NA, 1, 2, NA, NA, 2, 1, 2, NA, NA, NA, NA, NA,...
## $ CHILD12 <dbl+lbl> NA, 2, 2, NA, NA, 1, 2, 1, NA, NA, NA, NA, NA,...
## $ LGBT <dbl+lbl> NA, 2, 2, NA, NA, 2, 2, 2, NA, NA, NA, NA, NA,...
## $ UNIONHH12 <dbl+lbl> 2, NA, NA, 2, 2, NA, NA, NA, 2, 2, 2, 2, 2,...
## $ QN5 <dbl+lbl> 3, 2, 3, 2, 2, 3, 2, 2, 3, 2, 2, 1, 4, 2, 2, 1, 3, ...
## $ QN6A <dbl+lbl> NA, NA, 2, 1, NA, NA, 2, NA, NA, NA, NA, NA, 1,...

```

```
## $ QN6B <dbl+lbl> 1, 1, NA, NA, 1, 1, NA, 1, 2, 1, 1, 1, NA,...
## $ QN6C <dbl+lbl> 4, 0, 0, 0, 2, 1, 0, 2, 1, 1, 2, 1, 4, 0, 1, 1, 1, ...
## $ County <chr> "NEWAYGO", "WAYNE", "OAKLAND", "WASHTENAW", "BARAGA", "...
## $ SMPFIPS <dbl> 26123, 26163, 26125, 26161, 26013, 26057, 26037, 26121,...
## $ TTID <dbl> 94989, 94079, 94040, 94468, 94564, 93952, 93930, 94021,...
## $ GEOCODE <dbl> 5, 1, 2, 2, 5, 3, 2, 5, 2, 2, 3, 4, 3, 3, 5, 3, 2, 5, 2...
## $ alage5 <dbl> 1, 5, 4, 5, 4, 4, 5, 4, 3, 4, 5, 5, 1, 3, 5, 5, 4, 2, 1...
## $ A1RACE <dbl+lbl> 1, 2, 1, 1, 1, 1, 1, 1, 1, 2, 3, 1, 1, 1, 1, 1, 3, ...
## $ A1RACE3 <dbl+lbl> 1, 2, 1, 1, 1, 1, 1, 1, 1, 2, 3, 1, 1, 1, 1, 1, 3, ...
## $ A1SEX <dbl+lbl> 2, 2, 2, 1, 2, 2, 1, 1, 2, 1, 1, 1, 2, 1, 1, 2, 1, ...
## $ A1EDUC <dbl+lbl> 2, 1, 2, 2, 2, 1, 1, 1, 2, 2, 2, 2, 2, 1, 1, 1, 2, ...
## $ EDUCWHITE <dbl+lbl> 1, 4, 1, 1, 1, 2, 2, 2, 1, 3, 3, 1, 1, 2, 2, 2, 3, ...
## $ VOTEMETHOD <dbl> 1, 2, 2, 2, 3, 2, 2, 2, 1, 2, 1, 2, 2, 1, 1, 3, 3, 2, 1...
```

```
# More advanced: summary()
summary(MI_raw)
```

##	ID	WEIGHT	LALVOTERID	GROUP	
##	Min. : 9	Min. :0.1003	Length:1231	Min. :1.000	
##	1st Qu.: 21009	1st Qu.:0.3775	Class :character	1st Qu.:1.000	
##	Median : 52928	Median :0.8020	Mode :character	Median :2.000	
##	Mean : 68332	Mean :1.0000		Mean :2.648	
##	3rd Qu.:118105	3rd Qu.:1.4498		3rd Qu.:3.000	
##	Max. :190467	Max. :5.0853		Max. :9.000	
##					
##	ZIP	DISTRICT	Z1	S1	S2A
##	Min. :48001	Min. : 1.000	Min. :48002	Min. :1	Min. :1.000
##	1st Qu.:48221	1st Qu.: 3.000	1st Qu.:48188	1st Qu.:1	1st Qu.:1.000
##	Median :48744	Median : 6.000	Median :48502	Median :1	Median :1.000
##	Mean :48783	Mean : 6.707	Mean :48699	Mean :1	Mean :1.419
##	3rd Qu.:49327	3rd Qu.:10.000	3rd Qu.:49124	3rd Qu.:1	3rd Qu.:2.000
##	Max. :49968	Max. :14.000	Max. :49931	Max. :1	Max. :9.000
##			NA's :732		
##	S2B	S3	S4	VERSION	PRSMI20
##	Min. :1.000	Min. :1	Min. :1.000	Min. :1.0	Min. :0.00
##	1st Qu.:1.000	1st Qu.:1	1st Qu.:1.000	1st Qu.:1.0	1st Qu.:1.00
##	Median :1.000	Median :1	Median :1.000	Median :2.0	Median :1.00
##	Mean :1.328	Mean :1	Mean :1.285	Mean :1.5	Mean :1.63
##	3rd Qu.:1.000	3rd Qu.:1	3rd Qu.:1.000	3rd Qu.:2.0	3rd Qu.:2.00
##	Max. :9.000	Max. :1	Max. :3.000	Max. :2.0	Max. :9.00
##	NA's :508	NA's :722	NA's :722		
##	SENMI20	TIME16	ISSUE20	QLT20	TEMPBIDEN
##	Min. :0.000	Min. :2.00	Min. :1.000	Min. :1.000	Min. :1.000
##	1st Qu.:1.000	1st Qu.:5.00	1st Qu.:2.000	1st Qu.:2.000	1st Qu.:1.000
##	Median :1.000	Median :5.00	Median :3.000	Median :3.000	Median :1.000
##	Mean :1.526	Mean :4.85	Mean :3.489	Mean :2.956	Mean :1.496
##	3rd Qu.:2.000	3rd Qu.:5.00	3rd Qu.:5.000	3rd Qu.:4.000	3rd Qu.:2.000
##	Max. :9.000	Max. :9.00	Max. :9.000	Max. :9.000	Max. :9.000
##		NA's :616	NA's :616	NA's :616	NA's :616
##	TEMPTRUMP	CONTROLSN	FINSIT	ECONVCORONA20	
##	Min. :1.000	Min. :1.000	Min. :1.000	Min. :1.000	
##	1st Qu.:1.000	1st Qu.:1.000	1st Qu.:1.000	1st Qu.:1.000	
##	Median :2.000	Median :1.000	Median :2.000	Median :1.000	
##	Mean :1.686	Mean :1.873	Mean :2.211	Mean :1.901	
##	3rd Qu.:2.000	3rd Qu.:2.000	3rd Qu.:3.000	3rd Qu.:2.000	
##	Max. :9.000	Max. :9.000	Max. :9.000	Max. :9.000	
##	NA's :616	NA's :616	NA's :616	NA's :616	
##	FAVBIDEN2	FAVTRUMP	FORCAND	NEWVOTER	
##	Min. :1.000	Min. :1.000	Min. :1.000	Min. :1.000	
##	1st Qu.:1.000	1st Qu.:1.000	1st Qu.:1.000	1st Qu.:2.000	
##	Median :1.000	Median :2.000	Median :1.000	Median :2.000	
##	Mean :1.528	Mean :1.756	Mean :1.838	Mean :1.971	
##	3rd Qu.:2.000	3rd Qu.:2.000	3rd Qu.:2.000	3rd Qu.:2.000	
##	Max. :9.000	Max. :9.000	Max. :9.000	Max. :9.000	
##	NA's :616	NA's :616	NA's :615	NA's :615	
##	NEC	HANDLEECON20	HANDLECORONA20	RACISM20	
##	Min. :1.000	Min. :1.000	Min. :1.000	Min. :1.000	
##	1st Qu.:2.000	1st Qu.:1.000	1st Qu.:1.000	1st Qu.:2.000	
##	Median :3.000	Median :1.000	Median :1.000	Median :2.000	

##	Mean	:2.804	Mean	:1.799	Mean	:1.753	Mean	:2.325
##	3rd Qu.:	4.000	3rd Qu.:	2.000	3rd Qu.:	2.000	3rd Qu.:	3.000
##	Max.	:9.000	Max.	:9.000	Max.	:9.000	Max.	:9.000
##	NA's	:615	NA's	:615	NA's	:615	NA's	:615
##	VOTE2016		COUNTACC		TRUMP4		CONTAINCOVID	
##	Min.	:1.000	Min.	:1.000	Min.	:1.000	Min.	:1.000
##	1st Qu.:	1.000	1st Qu.:	1.000	1st Qu.:	1.000	1st Qu.:	2.000
##	Median	:2.000	Median	:1.500	Median	:4.000	Median	:3.000
##	Mean	:1.776	Mean	:1.808	Mean	:2.878	Mean	:3.092
##	3rd Qu.:	2.000	3rd Qu.:	2.000	3rd Qu.:	4.000	3rd Qu.:	4.000
##	Max.	:9.000	Max.	:9.000	Max.	:9.000	Max.	:9.000
##	NA's	:615	NA's	:615	NA's	:615	NA's	:615
##	COVIDHARDSHIP		AGE10		SEX		EDUC18	
##	Min.	:1.000	Min.	: 1.000	Min.	:1.00	Min.	:1.000
##	1st Qu.:	2.000	1st Qu.:	6.000	1st Qu.:	1.00	1st Qu.:	2.000
##	Median	:3.000	Median	: 8.000	Median	:2.00	Median	:3.000
##	Mean	:2.596	Mean	: 8.476	Mean	:1.53	Mean	:3.288
##	3rd Qu.:	3.000	3rd Qu.:	9.000	3rd Qu.:	2.00	3rd Qu.:	5.000
##	Max.	:9.000	Max.	:99.000	Max.	:2.00	Max.	:9.000
##	NA's	:615						
##	QACEAI		LATINOS		PARTYID		PHIL3	
##	Min.	:1.000	Min.	:1.000	Min.	:1.000	Min.	:1.000
##	1st Qu.:	1.000	1st Qu.:	2.000	1st Qu.:	1.000	1st Qu.:	1.000
##	Median	:1.000	Median	:2.000	Median	:2.000	Median	:2.000
##	Mean	:1.572	Mean	:2.175	Mean	:2.236	Mean	:2.295
##	3rd Qu.:	1.000	3rd Qu.:	2.000	3rd Qu.:	3.000	3rd Qu.:	3.000
##	Max.	:9.000	Max.	:9.000	Max.	:9.000	Max.	:9.000
##								
##	INCOME20		BRNAGAIN		CHILD12		LGBT	
##	Min.	:1.000	Min.	:1.000	Min.	:1.000	Min.	:1.000
##	1st Qu.:	2.000	1st Qu.:	1.000	1st Qu.:	2.000	1st Qu.:	2.000
##	Median	:3.000	Median	:2.000	Median	:2.000	Median	:2.000
##	Mean	:3.891	Mean	:1.907	Mean	:1.815	Mean	:2.224
##	3rd Qu.:	4.000	3rd Qu.:	2.000	3rd Qu.:	2.000	3rd Qu.:	2.000
##	Max.	:9.000	Max.	:9.000	Max.	:9.000	Max.	:9.000
##	NA's	:616	NA's	:615	NA's	:615	NA's	:615
##	UNIONHH12		QN5		QN6A		QN6B	
##	Min.	:1.000	Min.	: 1.000	Min.	:1.000	Min.	:1.000
##	1st Qu.:	2.000	1st Qu.:	2.000	1st Qu.:	2.000	1st Qu.:	1.000
##	Median	:2.000	Median	: 2.000	Median	:2.000	Median	:1.000
##	Mean	:1.837	Mean	: 3.219	Mean	:1.782	Mean	:1.178
##	3rd Qu.:	2.000	3rd Qu.:	3.000	3rd Qu.:	2.000	3rd Qu.:	1.000
##	Max.	:9.000	Max.	:99.000	Max.	:9.000	Max.	:9.000
##	NA's	:616			NA's	:732	NA's	:499
##	QN6C		County		SMPFIPS		TTID	
##	Min.	: 0.000	Length:1231		Min.	: -1	Min.	:93814
##	1st Qu.:	0.000	Class :character		1st Qu.:	-1	1st Qu.:	94122
##	Median	: 1.000	Mode :character		Median	:26063	Median	:94429
##	Mean	: 2.317			Mean	:16835	Mean	:94429
##	3rd Qu.:	2.000			3rd Qu.:	26125	3rd Qu.:	94737
##	Max.	:99.000			Max.	:26165	Max.	:95044
##								

```
##      GEOCODE      alage5      A1RACE      A1RACE3      A1SEX
##  Min.      :1.000    Min.      : 1.00    Min.      :1.000    Min.      :1.000    Min.      :1.00
## 1st Qu.:2.000    1st Qu.: 3.00    1st Qu.:1.000    1st Qu.:1.000    1st Qu.:1.00
## Median :3.000    Median : 4.00    Median :1.000    Median :1.000    Median :2.00
## Mean   :3.084    Mean   : 4.06    Mean   :1.348    Mean   :1.184    Mean   :1.53
## 3rd Qu.:4.000    3rd Qu.: 5.00    3rd Qu.:1.000    3rd Qu.:1.000    3rd Qu.:2.00
## Max.   :5.000    Max.   :11.00    Max.   :5.000    Max.   :3.000    Max.   :2.00
##
##      A1EDUC      EDUCWHITE      VOTEMETHOD
##  Min.      :1.000    Min.      :1.000    Min.      :1.00
## 1st Qu.:1.000    1st Qu.:1.000    1st Qu.:1.00
## Median :1.000    Median :2.000    Median :2.00
## Mean   :1.501    Mean   :1.861    Mean   :1.81
## 3rd Qu.:2.000    3rd Qu.:2.000    3rd Qu.:2.00
## Max.   :3.000    Max.   :4.000    Max.   :3.00
##
##      NA's      :11      NA's      :12
```

```
# Most advanced: count()
MI_raw %>%
  count(PRSMI20)
```

```
## # A tibble: 6 × 2
##   PRSMI20      n
##   <dbl> <dbl>
## 1 0 (NA) [Will/Did not vote for president]      6
## 2 1 [Joe Biden, the Democrat]                723
## 3 2 [Donald Trump, the Republican]            459
## 4 7 [Undecided/Don't know]                    4
## 5 8 [Refused]                                  14
## 6 9 [Another candidate]                       25
```

```
# Counting doesn't always work
MI_raw %>%
  count(WEIGHT)
```

```
## # A tibble: 411 × 2
##   WEIGHT      n
##   <dbl> <int>
## 1 0.100      1
## 2 0.113      1
## 3 0.119      1
## 4 0.133      2
## 5 0.141      1
## 6 0.142      1
## 7 0.144      1
## 8 0.146      1
## 9 0.147      1
## 10 0.149      5
## # ... with 401 more rows
```

```
summary(MI_raw %>% select(WEIGHT))
```

```
##      WEIGHT
##  Min.    :0.1003
##  1st Qu.:0.3775
##  Median :0.8020
##  Mean    :1.0000
##  3rd Qu.:1.4498
##  Max.    :5.0853
```

```
# Looking at PARTYID
```

```
glimpse(MI_raw %>% select(PARTYID))
```

```
## Rows: 1,231
## Columns: 1
## $ PARTYID <dbl+lbl> 3, 1, 1, 3, 3, 3, 1, 1, 2, 1, 3, 2, 4, 4, 1, 1, 3, 3, 3, 1...
```

```
MI_raw %>%
  count(PARTYID)
```

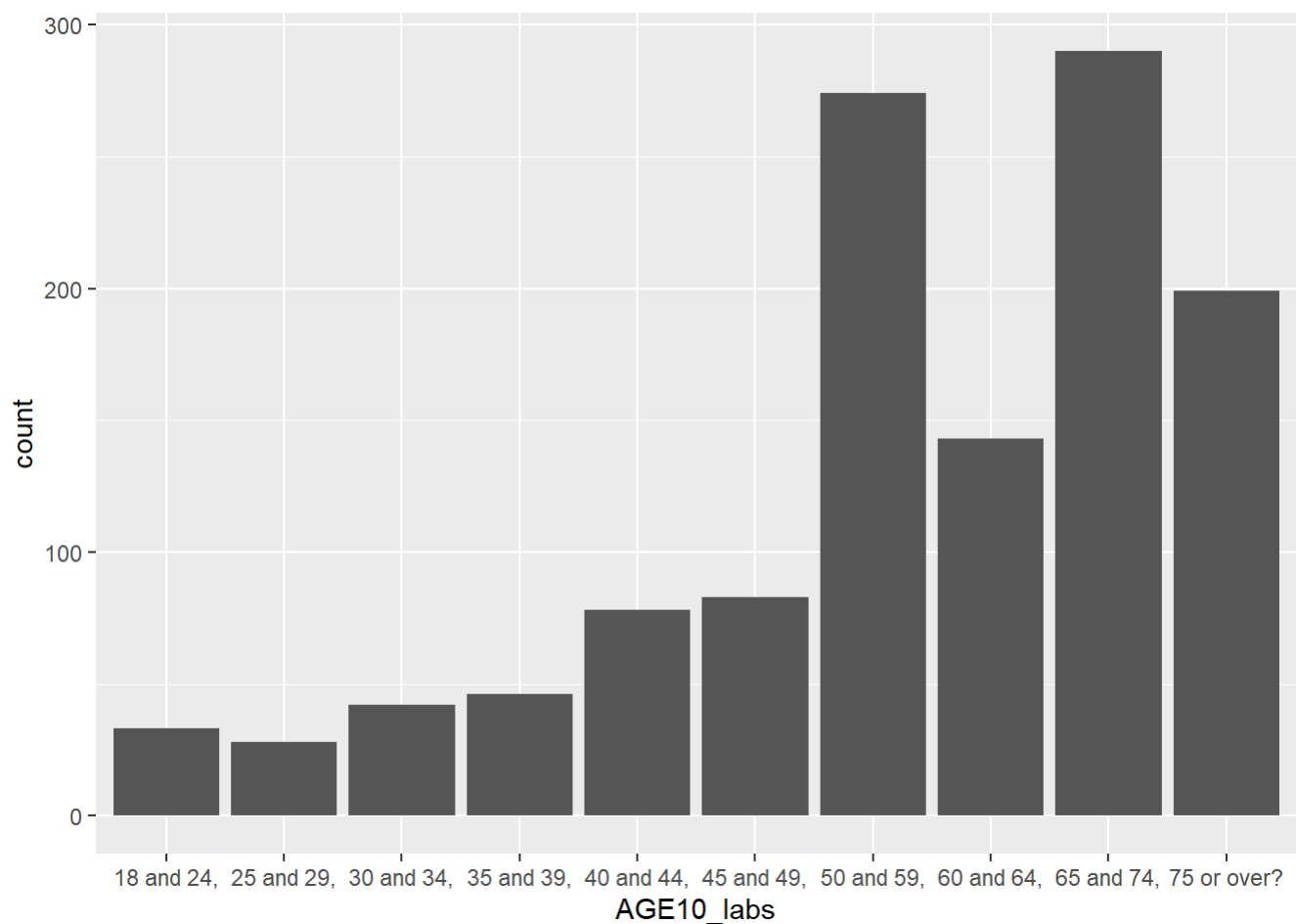
```
## # A tibble: 5 × 2
##   PARTYID          n
##   <dbl+lbl>      <int>
## 1 1 [Democrat]      425
## 2 2 [Republican]    280
## 3 3 [Independent]   416
## 4 4 [Something else]  94
## 5 9 [[DON'T READ] Don't know/refused] 16
```

Variable Types

```
# Combining as.character() with haven::as_factor()
MI_final <- MI_raw %>%
  mutate(AGE10_new = ifelse(AGE10 == 99, NA, AGE10))

MI_final <- MI_final %>%
  mutate(AGE10_labs = as.character(haven::as_factor(AGE10)))

MI_final %>%
  filter(!is.na(AGE10_new)) %>%
  ggplot(aes(x = AGE10_labs)) +
  geom_bar()
```

```
# Converting from labelled to character
MI_final <- MI_final %>%
  mutate(preschoice = as.character(haven::as_factor(PRSMI20)),
         Qlty = as.character(haven::as_factor(QLT20)),
         LGBT_lab = as.character(haven::as_factor(LGBT)))

MI_final %>%
  select(preschoice, Qlty, LGBT_lab)
```

```
## # A tibble: 1,231 × 3
##   preschoice          Qlty          LGBT_lab
##   <chr>             <chr>         <chr>
## 1 Joe Biden, the Democrat Has good judgment <NA>
## 2 Joe Biden, the Democrat <NA>          No
## 3 Joe Biden, the Democrat <NA>          No
## 4 Joe Biden, the Democrat Has good judgment <NA>
## 5 Joe Biden, the Democrat Cares about people like me <NA>
## 6 Joe Biden, the Democrat <NA>          No
## 7 Joe Biden, the Democrat <NA>          No
## 8 Joe Biden, the Democrat <NA>          No
## 9 Donald Trump, the Republican Cares about people like me <NA>
## 10 Joe Biden, the Democrat Cares about people like me <NA>
## # ... with 1,221 more rows
```

```
# Converting unit non-response to NA
MI_final %>%
  count(preschoice)
```

```
## # A tibble: 6 × 2
##   preschoice          n
##   <chr>          <int>
## 1 Another candidate      25
## 2 Donald Trump, the Republican 459
## 3 Joe Biden, the Democrat 723
## 4 Refused              14
## 5 Undecided/Don't know    4
## 6 Will/Did not vote for president 6
```

```
# Need to use either str_detect() or grepl()
MI_final %>%
  mutate(preschoice = ifelse(grepl('Refuse|know',preschoice),
                             NA,
                             preschoice)) %>%
  count(preschoice)
```

```
## # A tibble: 5 × 2
##   preschoice          n
##   <chr>          <int>
## 1 Another candidate      25
## 2 Donald Trump, the Republican 459
## 3 Joe Biden, the Democrat 723
## 4 Will/Did not vote for president 6
## 5 <NA>              18
```

Testing quiz possibilities

```
MI_final %>%
  summarise(preschoice)
```

```
## # A tibble: 1,231 × 1
##   preschoice
##   <chr>
## 1 Joe Biden, the Democrat
## 2 Joe Biden, the Democrat
## 3 Joe Biden, the Democrat
## 4 Joe Biden, the Democrat
## 5 Joe Biden, the Democrat
## 6 Joe Biden, the Democrat
## 7 Joe Biden, the Democrat
## 8 Joe Biden, the Democrat
## 9 Donald Trump, the Republican
## 10 Joe Biden, the Democrat
## # ... with 1,221 more rows
```

```
MI_final %>%
  mutate(preschoice)
```

```
## # A tibble: 1,231 × 68
##       ID WEIGHT LALVOTERID GROUP ZIP DISTR...1 Z1 S1 S2A
##   <dbl> <dbl> <chr>      <dbl+lbl> <dbl> <dbl> <dbl> <dbl+lbl> <dbl+lbl>
## 1     9  0.405 LALMI6290... 3 [3] 49327      2  NA 1 [Yes] 2 [No]
## 2    66  1.81 LALMI2492... 1 [1] 48234     14  NA 1 [Yes] 1 [Yes]
## 3   225  0.860 LALMI5489... 4 [4] 48301      9 48322 1 [Yes] 1 [Yes]
## 4   243  0.199 LALMI5053... 1 [1] 48130      7 48130 1 [Yes] 1 [Yes]
## 5   286  0.177 LALMI6831... 1 [1] 49946      1  NA 1 [Yes] 1 [Yes]
## 6   293  0.492 LALMI4019... 1 [1] 48615      4  NA 1 [Yes] 1 [Yes]
## 7   365  1.37 LALMI4151... 1 [1] 48906      4 48813 1 [Yes] 1 [Yes]
## 8   367  1.15 LALMI5912... 1 [1] 49442      2  NA 1 [Yes] 1 [Yes]
## 9   388  1.50 LALMI6635... 1 [1] 48451      5  NA 1 [Yes] 2 [No]
## 10  417  1.30 LALMI3567... 1 [1] 48197     12  NA 1 [Yes] 1 [Yes]
## # ... with 1,221 more rows, 59 more variables: S2B <dbl+lbl>, S3 <dbl+lbl>,
## # S4 <dbl+lbl>, VERSION <dbl+lbl>, PRSMI20 <dbl+lbl>, SENMI20 <dbl+lbl>,
## # TIME16 <dbl+lbl>, ISSUE20 <dbl+lbl>, QLT20 <dbl+lbl>, TEMPBIDEN <dbl+lbl>,
## # TEMPTRUMP <dbl+lbl>, CONTROLSEN <dbl+lbl>, FINSIT <dbl+lbl>,
## # ECONVCORONA20 <dbl+lbl>, FAVBIDEN2 <dbl+lbl>, FAVTRUMP <dbl+lbl>,
## # FORCAND <dbl+lbl>, NEWVOTER <dbl+lbl>, NEC <dbl+lbl>,
## # HANDLEECON20 <dbl+lbl>, HANDLECORONA20 <dbl+lbl>, RACISM20 <dbl+lbl>, ...
```

```
MI_final %>%
  select(preschoice)
```

```
## # A tibble: 1,231 × 1
##   preschoice
##   <chr>
## 1 Joe Biden, the Democrat
## 2 Joe Biden, the Democrat
## 3 Joe Biden, the Democrat
## 4 Joe Biden, the Democrat
## 5 Joe Biden, the Democrat
## 6 Joe Biden, the Democrat
## 7 Joe Biden, the Democrat
## 8 Joe Biden, the Democrat
## 9 Donald Trump, the Republican
## 10 Joe Biden, the Democrat
## # ... with 1,221 more rows
```

```
MI_final %>%
  count(preschoice)
```

```
## # A tibble: 6 × 2
##   preschoice      n
##   <chr>      <int>
## 1 Another candidate      25
## 2 Donald Trump, the Republican    459
## 3 Joe Biden, the Democrat    723
## 4 Refused      14
## 5 Undecided/Don't know      4
## 6 Will/Did not vote for president    6
```

```
MI_final %>%
  select(SEX | starts_with('P'))
```

```
## # A tibble: 1,231 × 5
##   SEX      PRSMI20      PARTYID      PHIL3      presc...1
##   <dbl+lbl> <dbl+lbl> <dbl+lbl> <dbl+lbl> <chr>
## 1 2 [Female] 1 [Joe Biden, the Democrat] 3 [Independe... 2 [Moderat... Joe Bi...
## 2 2 [Female] 1 [Joe Biden, the Democrat] 1 [Democrat] 2 [Moderat... Joe Bi...
## 3 2 [Female] 1 [Joe Biden, the Democrat] 1 [Democrat] 1 [Liberal] Joe Bi...
## 4 1 [Male] 1 [Joe Biden, the Democrat] 3 [Independe... 9 [[DON'T ... Joe Bi...
## 5 2 [Female] 1 [Joe Biden, the Democrat] 3 [Independe... 1 [Liberal] Joe Bi...
## 6 2 [Female] 1 [Joe Biden, the Democrat] 3 [Independe... 2 [Moderat... Joe Bi...
## 7 1 [Male] 1 [Joe Biden, the Democrat] 1 [Democrat] 9 [[DON'T ... Joe Bi...
## 8 1 [Male] 1 [Joe Biden, the Democrat] 1 [Democrat] 2 [Moderat... Joe Bi...
## 9 2 [Female] 2 [Donald Trump, the Republican] 2 [Republica... 3 [Conserv... Donald...
## 10 1 [Male] 1 [Joe Biden, the Democrat] 1 [Democrat] 2 [Moderat... Joe Bi...
## # ... with 1,221 more rows, and abbreviated variable name 'preschoice'
```