

17.871: PS1 Key

Part I

```
. log using ps1.log
. set more off           // Always set this to avoid keying through results
. insheet using scores.dat, clear
. list
```

```
+-----+
| name      age  test1 test2 |
+-----+
| Bob       18   95    18    |
| Carol     21   43    27    |
| Ted       14   67     9    |
| Alice     12   23    31    |
+-----+
```

```
. save scores, replace
```

Part II

```
. use spaesubset2012.dta, clear
```

```
/* Use tabulate to quickly view the coding scheme for the line length variable. */
```

```
. tab q10, m    // m displays missing values
. tab q10, m nol
```

```
/* Generate a dummy variable that contains a 1 if the respondent reported waiting more than 30
minutes (when the variable== 4 or 5), and 0 otherwise (1,2,3). Respondents who did not answer
or answered '6' (don't know) should be coded as missing ".", not as 0. */
```

```
. gen thirtyplus_2012=1 if q10==4|q10==5
. replace thirtyplus_2012=0 if q10<=3
. tab q10 thirtyplus, m    // to check variable assignment
```

```

               |               thirtyplus_2012
               |               0               1               . |               Total
-----+-----+-----+-----+-----+-----+-----+
Not at all | 3,037               0               0 | 3,037
Less than 10 minutes | 2,394               0               0 | 2,394
10-30 minutes | 1,375               0               0 | 1,375
31 minutes ñ 1 hour | 0               593               0 | 593
More than 1 hour | 0               211               0 | 211
I don't know | 0               0               24 | 24
. | 0               0               2,566 | 2,566
-----+-----+-----+-----+-----+
Total | 6,806               804               2,590 | 10,200
```

```
/* Collapse the dataset to (1) the percentage of respondents who waited 30 minutes or longer to
vote in each state and (2) the number of respondents who were included in the calculation. */
```

```
. collapse (mean) thirtyplus_2012 (count) n = thirtyplus_2012 [aweight=weight], by(regstate)
```

```
// First 10 observations
```

```
. list in 1/10
```

regstate	thirtyplus_2012	n
Alabama	.0760943	174
Alaska	.0222348	165
Arizona	.0761444	67
Arkansas	.1117	170
Californ	.0210003	95
Colorado	.0249651	67
Connecti	.0667891	171
Delaware	.0139119	190
District	.3852806	178
Florida	.3867461	131

```
. save wait_by_state_2012, replace
```

Part III

```
. use spaesubset2008.dta, clear
```

```
. describe // use describe to view any differences between datasets
```

```
. tab q12, m
```

```
. tab q12, m nol
```

```
/* Generate a variable that equals 1 if the respondent reported waiting more than 30 minutes */
```

```
. gen thirtyplus_2008=1 if q12==4| q12==5
```

```
. replace thirtyplus_2008=0 if q12<=3
```

```
/* Collapse the dataset */
```

```
. collapse (mean) thirtyplus_2008 (count) n= thirtyplus_2008 [aweight=weight], by(inputstate)
```

```
// First 10 obs
```

```
. list in 1/10
```

	inputs~e	thi~2008	n
1.	alabama	.135972	175
2.	alaska	.0192774	169
3.	arizona	.239623	92

```

4. | arkansas    .2076403    178 |
5. | californ   .0814993     97 |
   |-----|
6. | colorado    .123161     73 |
7. | connecti   .0963251    181 |
8. | delaware    .102089    187 |
9. | florida     .2867942    144 |
10. | georgia     .3538881    171 |
   +-----+

```

/* Rename the 'inputstate' variable to match the variable in the 2012 dataset (a common identifier is needed to merge datasets) */

```

. rename inputstate regstate
. save wait_by_state_2008, replace

```

/* Merge datasets using the regstate identifier. Note that 1 observation is not matched because there is no data for the District of Columbia in the 2008 dataset */

```

. merge 1:1 regstate using wait_by_state_2012.dta

```

List states where the proportion that waited 30+ minutes was greater in 2012 than in 2008

```

. list regstate if thirtyplus_2012>thirtyplus_2008

```

```

   +-----+
   | regstate |
   |-----|
2. | alaska   |
9. | florida  |
11. | hawaii   |
15. | iowa     |
20. | maryland |
   |-----|
21. | massachu |
22. | michigan |
26. | montana  |
29. | new hamp |
32. | new york |
   |-----|
34. | north da |
39. | rhode is |
41. | south da |
   +-----+

```

```

. save wait_by_state_merge, replace

```

Part IV

```

. use spaesubset2012.dta, clear

```

```
. gen thirtyplus_2012=1 if q10==4| q10==5
. replace thirtyplus_2012=0 if q10<=3
```

```
/* Collapse to create a dataset with the mean of thirtyplus_2012 by different types of voting and
state, as well as a count of the number of observations */
```

```
. collapse (mean) thirtyplus_2012 (count) n=thirtyplus_2012 [aweight=weight], by(regstate q4)
```

```
/* Drop observations where mode of voting is missing or where the mode is outside the scope of
our study */
```

```
. drop if q4==. | q4>=3
```

```
/* Reshape to create a state-level dataset with separate variables for thirty_plus2012 by each
mode */
```

```
. reshape wide thirtyplus_2012 n, i(regstate) j(q4)
. list
```

```
(First 5 values)
+-----+
| regstate  th~20121    n1    th~20122    n2 |
+-----+
1. | Alabama    .0760943    174      .      . |
2. |  Alaska      0    136    .1432126    29 |
3. |  Arizona    .0881236    61      0      6 |
4. | Arkansas    .0971476    87    .1297159    83 |
5. | Californ    .0213493    94      0      1 |
```

```
. log close
```