

Load the data

```
##load recoded ces files
load(here("Data", "recoded_cesdata.Rdata"))
```

This is just a bare-bones recode of the union variables.

```
#just check the 04s
ces0411 %>%
  mutate(union_both04=case_when(
    ces04_CPS_S6A==1 | ces04_CPS_S6B==1 ~ 1,
    ces04_CPS_S6A==5 ~ 0,
    ces04_CPS_S6B==5 ~ 0,
    TRUE ~ NA_real_
  )
  # ces04_CPS_S6A==8 & ces04_CPS_S6B==8 ~ NA_real_,
  # ces04_CPS_S6A==9 & ces04_CPS_S6B==9 ~ NA_real_,
  # ces06_CPS_S6A==1 | ces06_CPS_S6B==1 & ces04_rtype1==1 ~ 1,
  # ces06_CPS_S6A==5 & ces04_rtype1==1 ~ 0,
  # ces06_CPS_S6B==5 & ces04_rtype1==1 ~ 0,
  )->ces0411
```

Now we just filter only the respondents who took the PES04 and print.

```
library(flextable)
ces0411 %>%
  filter(str_detect(ces0411$survey, "PES04")) %>%
  #select employment status, respondent union, household union variables
  select(ces04_CPS_S4, ces04_CPS_S6A, ces04_CPS_S6B, union_both04) %>%
  #gtroup by each one
  group_by(ces04_CPS_S4, ces04_CPS_S6A, ces04_CPS_S6B, union_both04) %>%
  #count the numbers in each group
  summarize(n=n()) %>%
  #turn to factors for the value labels
  as_factor() %>%
  #print the table in a pretty way
  knitr::kable(., caption="Employment status, respondent union, household union and combined union counts")
```

Table 1: Employment status, respondent union, household union and combined union counts

ces04_CPS_S4	ces04_CPS_S6A	ces04_CPS_S6B	union_both04	n
Self-employed (with/without employees)	NA	NA	NA	389
Working for pay (full- or part-time)	Yes	NA	1	581
Working for pay (full- or part-time)	No	Yes	1	196
Working for pay (full- or part-time)	No	No	0	693
Working for pay (full- or part-time)	No	don't know	0	8
Working for pay (full- or part-time)	don't know	No	0	2
Working for pay (full- or part-time)	don't know	don't know	NA	3
Working for pay (full- or part-time)	refused	refused	NA	1
Working for pay (full- or part-time)	NA	NA	NA	3
Retired	NA	NA	NA	766
Unemployed/Looking for work	NA	NA	NA	111
Student	NA	NA	NA	95
Caring for a family	NA	NA	NA	120
Disabled	NA	NA	NA	51
R volunteers: Work at 2 or more jobs	Yes	NA	1	1

ces04_CPS_S4	ces04_CPS_S6A	ces04_CPS_S6B	union_both04	n
R volunteers: Work at 2 or more jobs	NA	NA	NA	19
Student and working for pay	NA	NA	NA	27
Caring for family and working for pay	NA	NA	NA	16
Retired and working for pay	NA	NA	NA	24
Other [specify]	NA	NA	NA	19
don't know	No	No	0	1
don't know	NA	NA	NA	4
refused	NA	NA	NA	11

Now as we found, there were a bunch of ces04 people who were reinterviewed in ces06 and we tried this trick of using their ces06 union status and applying it to ces04. But most people in ces04 who were NA on the union variables were actually retired. We can just look at what happened to them in this table that shows ces04 employment status, by ces06 union variables (respondent and household.)

```
#take the data
ces0411 %>%
  #filter only people who took ces06
  filter(str_detect(ces0411$survey, "PES06")) %>%
  #Select the ces04_rtype1, the ces04 employment status variable and the CES06 UNION VARIABLES
  select(ces04_rtype1, ces04_CPS_S4, ces06_CPS_S6A, ces06_CPS_S6B, union_both04) %>%
#Group
  group_by(ces04_rtype1, ces04_CPS_S4, ces06_CPS_S6A, ces06_CPS_S6B) %>%
  #count
  summarize(n=n()) %>%
as_factor() %>%
  knitr::kable(., caption="CES04 survey status, 04 employment status, 06 Respondent union and 06 household union counts")
```

Table 2: CES04 survey status, 04 employment status, 06 Respondent union and 06 household union counts

ces04_rtype1	ces04_CPS_S4	ces06_CPS_S6A	ces06_CPS_S6B	n
Yes-CPS	Self-employed (with/without employees)	Yes	NA	3
Yes-CPS	Self-employed (with/without employees)	No	Yes	2
Yes-CPS	Self-employed (with/without employees)	No	No	21
Yes-CPS	Self-employed (with/without employees)	No	don't know	1
Yes-CPS	Self-employed (with/without employees)	NA	NA	194
Yes-CPS	Working for pay (full- or part-time)	Yes	NA	295
Yes-CPS	Working for pay (full- or part-time)	No	Yes	80
Yes-CPS	Working for pay (full- or part-time)	No	No	291
Yes-CPS	Working for pay (full- or part-time)	No	don't know	1
Yes-CPS	Working for pay (full- or part-time)	don't know	No	1
Yes-CPS	Working for pay (full- or part-time)	NA	NA	103
Yes-CPS	Retired	Yes	NA	1
Yes-CPS	Retired	No	Yes	2
Yes-CPS	Retired	No	No	1
Yes-CPS	Retired	NA	NA	449
Yes-CPS	Unemployed/Looking for work	Yes	NA	4
Yes-CPS	Unemployed/Looking for work	No	Yes	2
Yes-CPS	Unemployed/Looking for work	No	No	13
Yes-CPS	Unemployed/Looking for work	NA	NA	32
Yes-CPS	Student	Yes	NA	4
Yes-CPS	Student	No	No	7

ces04_rtype1	ces04_CPS_S4	ces06_CPS_S6A	ces06_CPS_S6B	n
Yes-CPS	Student	NA	NA	20
Yes-CPS	Caring for a family	Yes	NA	1
Yes-CPS	Caring for a family	No	Yes	2
Yes-CPS	Caring for a family	No	No	4
Yes-CPS	Caring for a family	NA	NA	56
Yes-CPS	Disabled	No	Yes	1
Yes-CPS	Disabled	No	No	3
Yes-CPS	Disabled	NA	NA	21
Yes-CPS	R volunteers: Work at 2 or more jobs	Yes	NA	4
Yes-CPS	R volunteers: Work at 2 or more jobs	No	No	3
Yes-CPS	R volunteers: Work at 2 or more jobs	NA	NA	8
Yes-CPS	Student and working for pay	Yes	NA	1
Yes-CPS	Student and working for pay	No	Yes	2
Yes-CPS	Student and working for pay	No	No	3
Yes-CPS	Student and working for pay	NA	NA	6
Yes-CPS	Caring for family and working for pay	Yes	NA	3
Yes-CPS	Caring for family and working for pay	No	No	3
Yes-CPS	Caring for family and working for pay	NA	NA	2
Yes-CPS	Retired and working for pay	No	No	2
Yes-CPS	Retired and working for pay	NA	NA	13
Yes-CPS	Other [specify]	Yes	NA	1
Yes-CPS	Other [specify]	No	Yes	1
Yes-CPS	Other [specify]	No	No	1
Yes-CPS	Other [specify]	NA	NA	9
Yes-CPS	don't know	NA	NA	1
Yes-CPS	refused	No	No	1
Yes-CPS	refused	NA	NA	4
NA	NA	Yes	NA	318
NA	NA	No	Yes	79
NA	NA	No	No	370
NA	NA	No	don't know	3
NA	NA	don't know	don't know	1
NA	NA	NA	NA	795

Let's just see the differences between the stripped down recode and your recode.

```
ces0411 %>%
  mutate(union_both042=case_when(
    ces04_CPS_S6A==1 | ces04_CPS_S6B==1 ~ 1,
    ces04_CPS_S6A==5 ~ 0,
    ces04_CPS_S6B==5 ~ 0,
    ces04_CPS_S6A==8 & ces04_CPS_S6B==8 ~ NA_real_,
    ces04_CPS_S6A==9 & ces04_CPS_S6B==9 ~ NA_real_,
    ces06_CPS_S6A==1 | ces06_CPS_S6B==1 & ces04_rtype1==1 ~ 1,
    ces06_CPS_S6A==5 & ces04_rtype1==1 ~ 0,
    ces06_CPS_S6B==5 & ces04_rtype1==1 ~ 0,
  ))->ces0411

#count the stripped down recode
knitr::kable(table(ces0411$union_both04), caption="Frequency of union_both with stripped down recode")
```

Table 3: Frequency of union\_both with stripped down recode

Var1	Freq
0	994
1	1061

```
#count your recode
knitr::kable(table(ces0411$union_both042), caption="Frequency of union_both with your 06 recodes")
```

Table 4: Frequency of union\_both with your 06 recodes

Var1	Freq
0	1073
1	1494

So there are about 400 extra cases , but where did these 400 cases come from?

```
ces0411 %>%
  select(ces04_rtype1, ces04_CPS_S4, union_both04, union_both042) %>%
  group_by(ces04_rtype1, ces04_CPS_S4, union_both04, union_both042) %>%
  summarize(n=n()) %>%
  #filtering out those with missing values on union_both042 for readability
  filter(is.na(union_both042)==F) %>%
  as_factor() %>%
  knitr::kable(., caption="CES04 survey status, 04 employment status, combined union with stripped down")
```

Table 5: CES04 survey status, 04 employment status, combined union with stripped down code and Matt's 06 to 04 recode

ces04_rtype1	ces04_CPS_S4	union_both04	union_both042	n
Yes-CPS	Self-employed (with/without employees)	NA	0	27
Yes-CPS	Self-employed (with/without employees)	NA	1	6
Yes-CPS	Working for pay (full- or part-time)	0	0	991
Yes-CPS	Working for pay (full- or part-time)	1	1	1060
Yes-CPS	Working for pay (full- or part-time)	NA	1	1
Yes-CPS	Retired	NA	0	3
Yes-CPS	Retired	NA	1	3
Yes-CPS	Unemployed/Looking for work	NA	0	15
Yes-CPS	Unemployed/Looking for work	NA	1	8
Yes-CPS	Student	NA	0	9
Yes-CPS	Student	NA	1	5
Yes-CPS	Caring for a family	NA	0	5
Yes-CPS	Caring for a family	NA	1	3
Yes-CPS	Disabled	NA	0	3
Yes-CPS	Disabled	NA	1	1
Yes-CPS	R volunteers: Work at 2 or more jobs	0	0	1
Yes-CPS	R volunteers: Work at 2 or more jobs	1	1	1
Yes-CPS	R volunteers: Work at 2 or more jobs	NA	0	4
Yes-CPS	R volunteers: Work at 2 or more jobs	NA	1	4
Yes-CPS	Student and working for pay	NA	0	4
Yes-CPS	Student and working for pay	NA	1	4

ces04_rtype1	ces04_CPS_S4	union_both04	union_both042	n
Yes-CPS	Caring for family and working for pay	NA	0	3
Yes-CPS	Caring for family and working for pay	NA	1	3
Yes-CPS	Retired and working for pay	NA	0	4
Yes-CPS	Other [specify]	NA	0	1
Yes-CPS	Other [specify]	NA	1	2
Yes-CPS	don't know	0	0	2
Yes-CPS	refused	NA	0	1
NA	NA	NA	1	393

So by far most of these cases (393) came from people who had an NA on their CPS04\_rtype1, so we're not even sure they took part in CPS04, which is why they are also NA on the 2004 self-employment question, but their union06 variable response has been added to their union04. So I don't think this is very good.