

in class activity-7

Jiaqi Wang

2025-09-23

```
library(tidyverse)
```

```
## — Attaching core tidyverse packages — tidyverse 2.0.0 —
## ✓ dplyr      1.1.4      ✓ readr      2.1.5
## ✓ forcats    1.0.0      ✓ stringr    1.5.1
## ✓ ggplot2     3.5.2      ✓ tibble     3.3.0
## ✓ lubridate  1.9.4      ✓ tidyr      1.3.1
## ✓ purrr      1.1.0
## — Conflicts — tidyverse_conflicts() —
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(here)
```

```
## here() starts at /Users/seventh/Desktop/All class-25 fall
```

```
pat_demo <- read.csv(here("data","pat_demo-2.csv"))
pat_comorbid <- read.csv(here("data", "pat_comorbid-2.csv"))
dat <- pat_demo |> left_join(pat_comorbid, by = "PATID")
```

```
## Warning in left_join(pat_demo, pat_comorbid, by = "PATID"): Detected an unexpected
many-to-many relationship between `x` and `y`.
## i Row 25 of `x` matches multiple rows in `y`.
## i Row 25 of `y` matches multiple rows in `x`.
## i If a many-to-many relationship is expected, set `relationship =
## "many-to-many"` to silence this warning.
```

```
dat <- dat |> distinct()

dat <- dat |>
  mutate(
    birth_date = parse_date_time(DOB,
                                orders = c("mdy", "ymd", "dmy", "m/d/y", "Y-m-d")),
    birth_date = if_else(birth_date > today(), birth_date - year(100), birth_date)
  )
```

```
## Warning: There was 1 warning in `mutate()`.  
## i In argument: `birth_date = if_else(birth_date > today(), birth_date -  
##   year(100), birth_date)`.  
## Caused by warning:  
## ! tz(): Don't know how to compute timezone for object of class numeric; returning  
"UTC".
```

```
dat <- dat |> select(-DOB)
```

```
dat |>  
  select(PATID, birth_date, race, ethnicity, financialclass,  
         hypertension, CHF, diabetes, CKD) |>  
  arrange(PATID)
```

##	PATID	birth_date	race	ethnicity	financialclass	hypertension	CHF	diabetes
## 1	Z1267	1972-02-29	White	non-Hispanic	Medicare	N	N	N
## 2	Z1349	1972-12-25	White	non-Hispanic	Medicare	N	N	N
## 3	Z1392	1957-08-15	White	non-Hispanic	Private	N	N	N
## 4	Z1572	1987-10-19	White	non-Hispanic	Private	N	N	N
## 5	Z1649	1982-01-21	White	non-Hispanic	Medicare	N	N	N
## 6	Z1884	1959-04-13	White	non-Hispanic	Medicare	N	N	N
## 7	Z2041	1964-04-19	White	non-Hispanic	Private	N	N	N
## 8	Z2359	1994-08-19	White	non-Hispanic	Private	N	N	N
## 9	Z2914	1962-06-29	Black	non-Hispanic	Medicare	N	N	N
## 10	Z2979	1959-02-21	White	non-Hispanic	Private	Y	N	N
## 11	Z3108	1995-09-27	White	non-Hispanic	Private	N	Y	N
## 12	Z3307	1998-01-01	Black	non-Hispanic	Medicare	Y	N	N
## 13	Z3359	1951-08-27	White	non-Hispanic	Private	Y	N	N
## 14	Z3471	1999-08-16	Black	Hispanic	Private	N	N	N
## 15	Z3946	1950-01-30	White	non-Hispanic	Medicare	Y	N	Y
## 16	Z4245	2000-05-02	White	non-Hispanic	Private	Y	N	N
## 17	Z4318	1973-10-23	White	non-Hispanic	Medicare	Y	N	N
## 18	Z4582	1989-09-10	White	non-Hispanic	Private	Y	N	N
## 19	Z4776	1987-10-02	White	non-Hispanic	Private	Y	N	N
## 20	Z4877	1998-01-27	White	non-Hispanic	Medicare	N	N	N
## 21	Z4978	1996-03-31	Black	non-Hispanic	Medicare	Y	N	N
## 22	Z5222	1973-08-30	Black	non-Hispanic	Medicare	N	N	N
## 23	Z5278	1958-07-01	Other	non-Hispanic	Private	N	N	N
## 24	Z5552	1986-03-17	Other	non-Hispanic	Medicare	N	N	N
## 25	Z5601	1975-08-24	White	non-Hispanic	Private	Y	N	N
## 26	Z5835	1960-12-04	White	non-Hispanic	Private	Y	N	N
## 27	Z6198	1986-06-23	White	Hispanic	Medicare	N	N	N
## 28	Z6400	1966-04-04	White	non-Hispanic	Medicare	N	N	N
## 29	Z6514	1967-02-12	Black	non-Hispanic	Medicare	N	Y	N
## 30	Z6790	1967-06-28	White	non-Hispanic	Private	N	N	N
## 31	Z7184	1983-01-25	Black	non-Hispanic	Medicare	N	Y	N
## 32	Z7599	1971-02-02	White	non-Hispanic	Private	N	N	N
## 33	Z7851	1949-04-13	Black	non-Hispanic	Private	Y	N	N
## 34	Z7891	1993-12-23	White	non-Hispanic	Medicare	Y	N	N
## 35	Z7940	1996-03-30	Black	non-Hispanic	Medicare	N	N	N
## 36	Z8015	1966-11-04	White	non-Hispanic	Private	N	N	N
## 37	Z8517	1977-07-17	White	non-Hispanic	Medicare	Y	N	N
## 38	Z8590	1993-01-16	White	non-Hispanic	Medicare	Y	N	N
## 39	Z8728	1979-10-22	White	Hispanic	Private	Y	N	N
## 40	Z8747	1961-07-31	White	non-Hispanic	Medicare	N	Y	N
## 41	Z8812	1977-08-05	White	non-Hispanic	Private	N	N	N
## 42	Z8814	1989-05-09	White	non-Hispanic	Medicare	Y	N	N
## 43	Z9101	1965-10-21	White	non-Hispanic	Private	Y	N	N
## 44	Z9306	1984-10-29	Other	non-Hispanic	Medicare	N	N	Y
## 45	Z9525	1975-02-06	White	non-Hispanic	Medicare	N	N	N
## 46	Z9594	1967-03-04	White	non-Hispanic	Medicare	Y	Y	N
## 47	Z9634	1989-06-25	Black	non-Hispanic	Medicare	N	N	N
## 48	Z9637	1971-04-17	White	non-Hispanic	Medicare	N	N	N
## 49	Z9801	1997-04-05	White	non-Hispanic	Medicare	Y	N	N
##	CKD							
## 1	Y							
## 2	N							
## 3	N							
## 4	N							

```
## 5      N
## 6      N
## 7      Y
## 8      N
## 9      N
## 10     Y
## 11     N
## 12     N
## 13     N
## 14     N
## 15     Y
## 16     N
## 17     N
## 18     N
## 19     Y
## 20     N
## 21     N
## 22     N
## 23     N
## 24     N
## 25     N
## 26     N
## 27     N
## 28     N
## 29     N
## 30     N
## 31     Y
## 32     N
## 33     Y
## 34     Y
## 35     N
## 36     Y
## 37     N
## 38     N
## 39     N
## 40     N
## 41     Y
## 42     N
## 43     N
## 44     N
## 45     N
## 46     Y
## 47     N
## 48     N
## 49     N
```

```
dat <- as_tibble(dat)
nrow(dat)
```

```
## [1] 49
```

```
print(dat, n = 10)
```

```
## # A tibble: 49 × 9
##   PATID race ethnicity financialclass hypertension CHF diabetes CKD
##   <chr> <chr> <chr>          <chr>          <chr>      <chr> <chr> <chr>
## 1 Z5278 Other non-Hispanic Private          N          N      N      N
## 2 Z9101 White non-Hispanic Private          Y          N      N      N
## 3 Z1649 White non-Hispanic Medicare         N          N      N      N
## 4 Z3946 White non-Hispanic Medicare         Y          N      Y      Y
## 5 Z4582 White non-Hispanic Private          Y          N      N      N
## 6 Z1392 White non-Hispanic Private          N          N      N      N
## 7 Z8517 White non-Hispanic Medicare         Y          N      N      N
## 8 Z4318 White non-Hispanic Medicare         Y          N      N      N
## 9 Z3359 White non-Hispanic Private          Y          N      N      N
## 10 Z9306 Other non-Hispanic Medicare         N          N      Y      N
## # i 39 more rows
## # i 1 more variable: birth_date <dtm>
```