Assignment 3

Nicole Mis

February 23, 2018

10.5 Exercises

5. What does tibble::enframe() do? When might you use it?

It turns a vector or list into a tibble. If the vector is named, then it assigns the name as the column name.For unnamed vectors, the sequence of the vector is used for the name column. You can use enframe when you want to turn a named list into a data frame. This allows a list of data to be transformed into a tibble. These lists could be data sets or regression results. You can also use enframe when your data is a vector. For instance, if you have data that is contained in separate vectors, then you can use enframe to convert it to a data frame and then manipulate that data and turn it into a more easily readible set of data.

```
#binstall.packages("kableExtra")

library(kableExtra)

library(knitr)

library(tidyverse)
```

```
-- Attaching packages ----------------------------------------------------- tidyverse 1.2.1 --
```

```
v ggplot2 2.2.1     v purrr   0.2.4
v tibble  1.4.2     v dplyr   0.7.4
v tidyr   0.8.0     v stringr 1.2.0
v readr   1.1.1     v forcats 0.2.0
```

```
-- Conflicts -------------------------------------------------------- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
```

```
x <- c("a", "b", "c", "d", "e", "f", "g", "h")
x <- enframe(x)
x
```

```
# A tibble: 8 x 2
   name value
  <int> <chr>
1     1 a
2     2 b
3     3 c
4     4 d
5     5 e
6     6 f
7     7 g
8     8 h
```

12.6.1

3. I claimed that iso2 and iso3 were redundant with country. Confirm this claim.

I checked to see how many unique values there were for country, iso2 and iso3 and I checked for unique combinations of these rows to confirm the above claim.

```
library(tidyverse)

who1 <- tidyr::who %>% gather(new_sp_m014:newrel_f65, key = "key", value = "cases", na.rm = TRUE)

who1 %>% count(key)
```

```
# A tibble: 56 x 2
   key              n
   <chr>        <int>
 1 new_ep_f014   1032
 2 new_ep_f1524  1021
 3 new_ep_f2534  1021
 4 new_ep_f3544  1021
 5 new_ep_f4554  1017
 6 new_ep_f5564  1017
 7 new_ep_f65    1014
 8 new_ep_m014   1038
 9 new_ep_m1524  1026
10 new_ep_m2534  1020
# ... with 46 more rows
```

```
who2 <- who1 %>% mutate(key = stringr::str_replace(key, "newrel", "new_rel"))

who3 <- who2 %>% separate(key, c("new", "type", "sexage"), sep = "_")

who3 %>% select(1:3) %>% sapply(function(x){length(unique(x))})
```

```
country    iso2    iso3
    219     219     219
```

```
 who3 %>% select(1:3) %>%
   unite(combined, 1:3) %>%
   select(combined) %>%
   distinct() %>%
   nrow()
```

```
[1] 219
```

4. For each country, year, and sex compute the total number of cases of TB. Make an informative visualisation of the data.

```
 tidyr::who %>%
   gather(code, value, new_sp_m014:newrel_f65, na.rm = TRUE) %>%
   mutate(code = stringr::str_replace(code, "newrel", "new_rel")) %>%
   separate(code, c("new", "var", "sexage")) %>%
   select(-new, -iso2, -iso3) %>%
```

```
  separate(sexage, c("sex", "age"), sep = 1) %>%
  group_by(country, year, sex) %>%
  summarize(total_case = sum(value)) %>%
  unite(country_sex, country, sex, remove = FALSE) %>%
  ggplot() +
  geom_line(mapping = aes(x = year, y = total_case, color = sex,
                          group = country_sex))
```

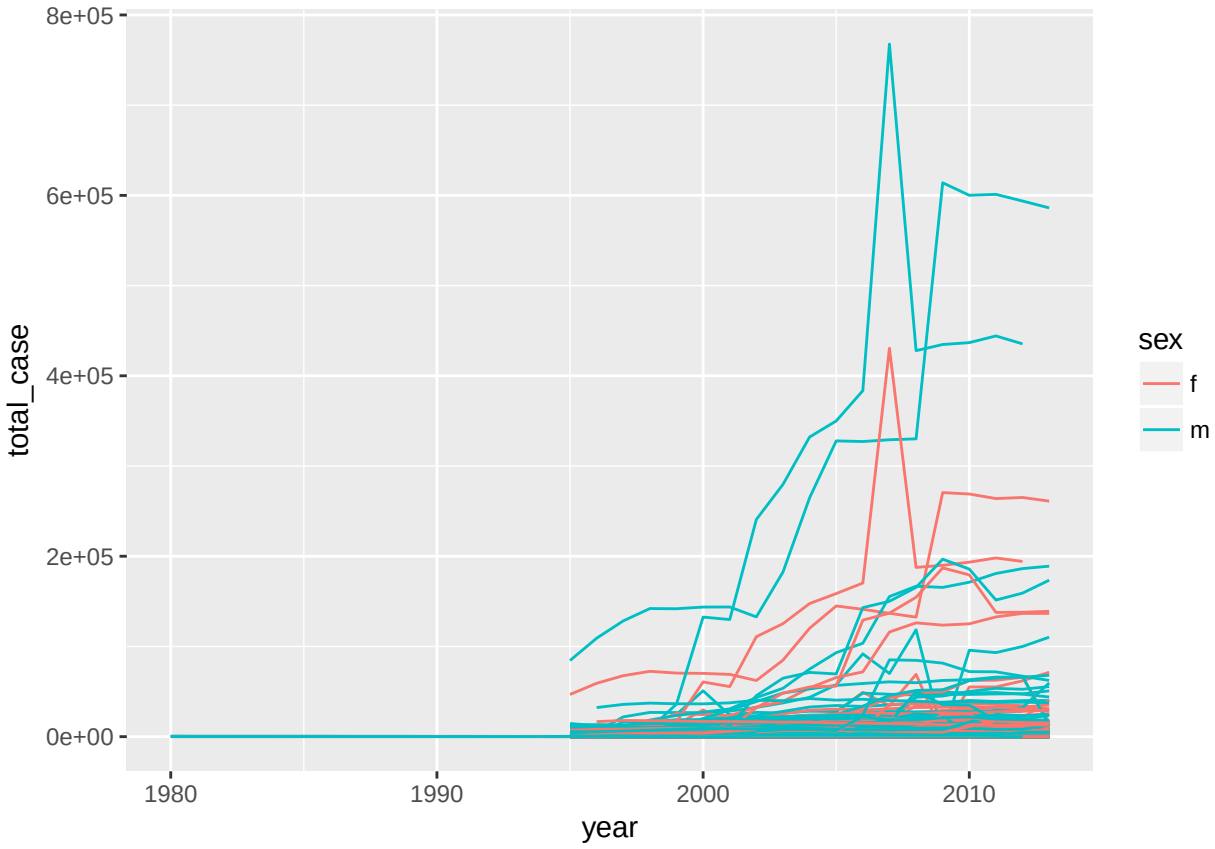

Table 4 -> table 5

```
# install.packages("kableExtra")
library(tidyverse)
library(kableExtra)
library(knitr)

### HW added line
raw <- read_rds("rawdat.RDS")

raw.tidy <- raw %>% gather(key = "Income", value = "freq", -religion)
raw2 <- raw.tidy %>% arrange(religion)

knitr::kable(head(raw2, n=10), "latex", booktabs = T, linesep = "") %>% kable_styling(position = "cente
```

religion	Income	freq
Agnostic	$75-100k	826
Atheist	$75-100k	515
Buddhist	$75-100k	411
Catholic	$75-100k	8054
Don't know/refused	$75-100k	272
Evangelical Prot	$75-100k	9472
Hindu	$75-100k	257
Historically Black Prot	$75-100k	1995
Jehovah's Witness	$75-100k	215
Jewish	$75-100k	682

Table 7 -> table 8

```
library(tidyverse)
library(lubridate)
```

```
Attaching package: 'lubridate'
```

```
The following object is masked from 'package:base':

    date
```

```
library(knitr)
library(kableExtra)

bb <- read_csv("billboard.csv")
```

```
Parsed with column specification:
cols(
  .default = col_integer(),
  artist.inverted = col_character(),
  track = col_character(),
  time = col_time(format = ""),
  genre = col_character(),
  date.entered = col_date(format = ""),
  date.peaked = col_date(format = ""),
  x66th.week = col_character(),
  x67th.week = col_character(),
  x68th.week = col_character(),
  x69th.week = col_character(),
  x70th.week = col_character(),
  x71st.week = col_character(),
  x72nd.week = col_character(),
  x73rd.week = col_character(),
  x74th.week = col_character(),
  x75th.week = col_character(),
  x76th.week = col_character()
)
```

```
See spec(...) for full column specifications.
```

```
bb.1 <- bb %>% gather(key="week", value = "rank", -year, -artist.inverted, -track, -time, -genre, -date

bb.2 <- bb.1 %>% select(year, artist=artist.inverted, time, track, date = date.entered, week, rank )

bb.3 <- bb.2 %>% arrange(track)

bb.4 <- bb.3 %>% filter(!is.na(rank))

bb.5 <- bb.4 %>% separate(week, into=c("A", "B", "C"), sep=c(1, -8), convert=TRUE)

bb.6 <- bb.5 %>% select(-A, -C)

bb.7 <- bb.6 %>% dplyr::rename(week = B)

bb.8 <- bb.7 %>% arrange(artist, track)

bb.9 <- bb.8 %>% mutate(rank = as.integer(rank))

bb.10 <- bb.9 %>% group_by(track) %>% mutate(week=seq(n()))

bb.11 <- bb.10 %>% group_by(track) %>% mutate(date = date+(week-1)*7)

knitr::kable(head(bb.11, n=15), "latex", booktabs = T, linesep = "") %>% kable_styling(position = "cent
```

year	artist	time	track	date	week	rank
2000	2 Pac	04:22:00	Baby Don't Cry (Keep Ya Head Up II)	2000-02-26	1	87
2000	2 Pac	04:22:00	Baby Don't Cry (Keep Ya Head Up II)	2000-03-04	2	82
2000	2 Pac	04:22:00	Baby Don't Cry (Keep Ya Head Up II)	2000-03-11	3	72
2000	2 Pac	04:22:00	Baby Don't Cry (Keep Ya Head Up II)	2000-03-18	4	77
2000	2 Pac	04:22:00	Baby Don't Cry (Keep Ya Head Up II)	2000-03-25	5	87
2000	2 Pac	04:22:00	Baby Don't Cry (Keep Ya Head Up II)	2000-04-01	6	94
2000	2 Pac	04:22:00	Baby Don't Cry (Keep Ya Head Up II)	2000-04-08	7	99
2000	2Ge+her	03:15:00	The Hardest Part Of Breaking Up (Is Getting Back Your Stuff)	2000-09-02	1	91
2000	2Ge+her	03:15:00	The Hardest Part Of Breaking Up (Is Getting Back Your Stuff)	2000-09-09	2	87
2000	2Ge+her	03:15:00	The Hardest Part Of Breaking Up (Is Getting Back Your Stuff)	2000-09-16	3	92
2000	3 Doors Down	03:53:00	Kryptonite	2000-04-08	1	81
2000	3 Doors Down	03:53:00	Kryptonite	2000-04-15	2	70
2000	3 Doors Down	03:53:00	Kryptonite	2000-04-22	3	68
2000	3 Doors Down	03:53:00	Kryptonite	2000-04-29	4	67
2000	3 Doors Down	03:53:00	Kryptonite	2000-05-06	5	66