Table_Convert

Junpei Xiao

2018/02/28

Table 4 to table 6 convert

```r
# Data from http://pewforum.org/Datasets/Dataset-Download.aspx

# Load data

### Convert Table4 to Table6... The code for creating table 4 is from Author Hadley Wickham.

pew <- read.spss("pew.sav")

## Warning in read.spss("pew.sav"): Undeclared level(s) 2, 3, 4, 9 added in
## variable: density3

## Warning in read.spss("pew.sav"): Duplicated levels in factor denom:
## Electronic ministries

## Warning in read.spss("pew.sav"): Undeclared level(s) 1, 2, 3, 4, 5, 6, 7,
## 8, 9, 10, 11, 12, 14, 16, 23, 33 added in variable: children

## Warning in read.spss("pew.sav"): Undeclared level(s) 18, 19, 20, 21, 22,
## 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41,
## 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60,
## 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79,
## 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96 added in
## variable: age

pew <- as.data.frame(pew)


religion <- pew[c("q16", "reltrad", "income")]
religion$reltrad <- as.character(religion$reltrad)
religion$reltrad <- str_replace(religion$reltrad, " Churches", "")
religion$reltrad <- str_replace(religion$reltrad, " Protestant", " Prot")
religion$reltrad[religion$q16 == " Atheist (do not believe in God) "] <
```

```
 - "Atheist"
religion$reltrad[religion$q16 == " Agnostic (not sure if there is a God)
 "] <- "Agnostic"
religion$reltrad <- str_trim(religion$reltrad)
religion$reltrad <- str_replace_all(religion$reltrad, " \\(.*?\\)", "")

religion$income <- c("Less than $10,000" = "<$10k",
                     "10 to under $20,000" = "$10-20k",
                     "20 to under $30,000" = "$20-30k",
                     "30 to under $40,000" = "$30-40k",
                     "40 to under $50,000" = "$40-50k",
                     "50 to under $75,000" = "$50-75k",
                     "75 to under $100,000" = "$75-100k",
                     "100 to under $150,000" = "$100-150k",
                     "$150,000 or more" = ">150k",
                     "Don't know/Refused (VOL)" = "Don't know/refused")
[religion$income]

religion$income <- factor(religion$income, levels = c("<$10k", "$10-20k
", "$20-30k", "$30-40k", "$40-50k", "$50-75k",
                                                "$75-100k", "$100
-150k", ">150k", "Don't know/refused"))

clean_Table <- religion %>% group_by(reltrad,income)%>% summarise(count
=n())
clean_Table<- as.data.frame(clean_Table)
head(clean_Table,n=10)
```

reltrad	income	count
Agnostic	<$10k	27
Agnostic	$10-20k	34
Agnostic	$20-30k	60
Agnostic	$30-40k	81
Agnostic	$40-50k	76
Agnostic	$50-75k	137
Agnostic	$75-100k	122
Agnostic	$100-150k	109
Agnostic	>150k	84
Agnostic	Don't know/refused	96

Table 7 To Table 8 convert

```
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.entered, -date.peaked)
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", "week", "C"), sep=c(1, -7),
convert=TRUE)
bb.6 <-bb.5 %>% select(-A, -C)
bb.7 <-bb.6%>% filter(!is.na(week))
#must to specified rename in the dplyr
bb.8 <- bb.7 %>% arrange(artist, track)
bb.9 <-bb.8 %>% mutate(date = date+(week-1)*7)
bb.10 <- bb.9 %>% mutate(rank = as.numeric(bb.9$rank))
head(bb.10,n=10)
```

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