Tidyverse Problem Set

Qi Huang

October 4, 2019

The purpose of this problem set is to provide data contexts in which to exercise the capabilitiues of the tidyverse. While some questons require specific answers, other parts of the problems have been written to be purposely ambiguous, requiring you to think through the presentation details of your answer.

HOLD THE PRESSES!
As I was preparing to post these problems yesterday, I noticed that tidyr had been updata in the last few weeks. I was looking for more exercises on gather() and spread() – which are always difficult to master. And I found that they have been superceded!! Why do I love working with R as the tidyversie is on a path of continuous improvement? Because the improvements come from developers who write things like this:

For some time, it's been obvious that there is something fundamentally wrong with the design of spread() and gather(). Many people don't find the names intuitive and find it hard to remember which direction corresponds to spreading and which to gathering. It also seems surprisingly hard to remember the arguments to these functions, meaning that many people (including me!) have to consult the documentation every time. Hadley Wickham, Pivot Vingette

So... before you do anymore tidyverse exercises, Read this tidyr 1.0.0.

Then go to the tidyr cran page and to the examples and exercies in the new vignettes.

In your solutions to the problems below, if you need to use table reshaping functions from TidyR, be sure that you use pivot_longer(), and pivot_wider().

Problem 1

Load the gapminder data from the gapminder package.

```
library(gapminder)
data(gapminder)
```

1.1 How many continents are included in the data set?

```
summary(gapminder$continent)
```

```
##   Africa Americas     Asia   Europe  Oceania 
##      624      300      396      360       24
```

Answer: 5 continents included in the dataset.

1.2 How many countrys are included? How many countries per continent? summary(gapminder$country)

```
summary(gapminder$country)
```

```
##              Afghanistan                  Albania                  Algeria 
##                       12                       12                       12 
##                   Angola                Argentina                Australia 
##                       12                       12                       12 
##                  Austria                  Bahrain               Bangladesh 
##                       12                       12                       12 
##                  Belgium                    Benin                  Bolivia 
##                       12                       12                       12 
##   Bosnia and Herzegovina                 Botswana                   Brazil 
##                       12                       12                       12
##                 Bulgaria             Burkina Faso                  Burundi
##                       12                       12                       12
##                 Cambodia                 Cameroon                   Canada
##                       12                       12                       12
## Central African Republic                     Chad                    Chile
##                       12                       12                       12
##                    China                 Colombia                  Comoros
##                       12                       12                       12
##         Congo, Dem. Rep.              Congo, Rep.               Costa Rica
##                       12                       12                       12
##            Cote d'Ivoire                  Croatia                     Cuba
##                       12                       12                       12
##           Czech Republic                  Denmark                 Djibouti
##                       12                       12                       12
##       Dominican Republic                  Ecuador                    Egypt
##                       12                       12                       12
##              El Salvador        Equatorial Guinea                  Eritrea
##                       12                       12                       12
##                 Ethiopia                  Finland                   France
##                       12                       12                       12
##                    Gabon                   Gambia                  Germany
##                       12                       12                       12
##                    Ghana                   Greece                Guatemala
##                       12                       12                       12
##                   Guinea            Guinea-Bissau                    Haiti
##                       12                       12                       12
##                 Honduras         Hong Kong, China                  Hungary
##                       12                       12                       12
##                  Iceland                    India                Indonesia
##                       12                       12                       12
##                     Iran                     Iraq                  Ireland
##                       12                       12                       12
##                   Israel                    Italy                  Jamaica
##                       12                       12                       12
##                    Japan                   Jordan                    Kenya
##                       12                       12                       12
##         Korea, Dem. Rep.              Korea, Rep.                   Kuwait
##                       12                       12                       12
##                  Lebanon                  Lesotho                  Liberia
##                       12                       12                       12
##                    Libya               Madagascar                   Malawi
##                       12                       12                       12
##                 Malaysia                     Mali               Mauritania
##                       12                       12                       12
##                Mauritius                   Mexico                 Mongolia
##                       12                       12                       12
##               Montenegro                  Morocco               Mozambique
##                       12                       12                       12
##                  Myanmar                  Namibia                    Nepal
##                       12                       12                       12
##              Netherlands              New Zealand                Nicaragua
##                       12                       12                       12
##                    Niger                  Nigeria                   Norway
##                       12                       12                       12
##                     Oman                 Pakistan                   Panama
##                       12                       12                       12
##                  (Other)
##                      516
```

```
observe1<-gapminder %>% group_by(continent) %>% summarize(num_obs = n(), num_countries = n_distinct(cou
observe1
```

```
## # A tibble: 5 x 3
##   continent num_obs num_countries
##   <fct>       <int>         <int>
## 1 Africa        624            52
## 2 Americas      300            25
## 3 Asia          396            33
## 4 Europe        360            30
## 5 Oceania        24             2
```

Answer: 52 countries in Africa, 25 countries in America, 33 countries in Asia, 30 countries in Europe and 2 countries in Oceania.

1.3 Using the gapminder data, produce a report showing the continents in the dataset, total population per continent, and GDP per capita. Be sure that the table is properly labeled and suitable for inclusion in a printed report.

```
report2<-select(gapminder,continent,pop,gdpPercap)
report<-cbind(aggregate(pop~continent,report2,sum),aggregate(gdpPercap~continent,report2,sum))
report<-subset(report,select=-3)
report
```

```
##   continent         pop gdpPercap
## 1    Africa  6187585961 1368902.9
## 2  Americas  7351438499 2140833.1
## 3      Asia 30507333901 3129251.6
## 4    Europe  6181115304 5209011.2
## 5   Oceania   212992136  446918.6
```

1.4 Produce a well-labeled table that summarizes GDP per capita for the countries in each continent, contrasting the years 1952 and 2007.

```
Summary_1952 = gapminder %>%
  filter(year == 1952) %>%
  group_by(continent) %>%
  summarise(average = mean(gdpPercap),
            max = max(gdpPercap),
            min = min(gdpPercap),
            var = var(gdpPercap)) %>%
  arrange(average)
kable(Summary_1952, caption = "Summary for gdpPercap in 1952", align = "c", booktab = T, format = "late
  kable_styling(latex_options = c("HOLD_position"))
```

Table 2: Summary for gdpPercap in 2007

continent	average	max	min	var
Africa	3089.033	13206.48	277.5519	13091107
Americas	11003.032	42951.65	1201.6372	94346435
Asia	12473.027	47306.99	944.0000	200362251
Europe	25054.482	49357.19	5937.0295	139248020
Oceania	29810.188	34435.37	25185.0091	42784565

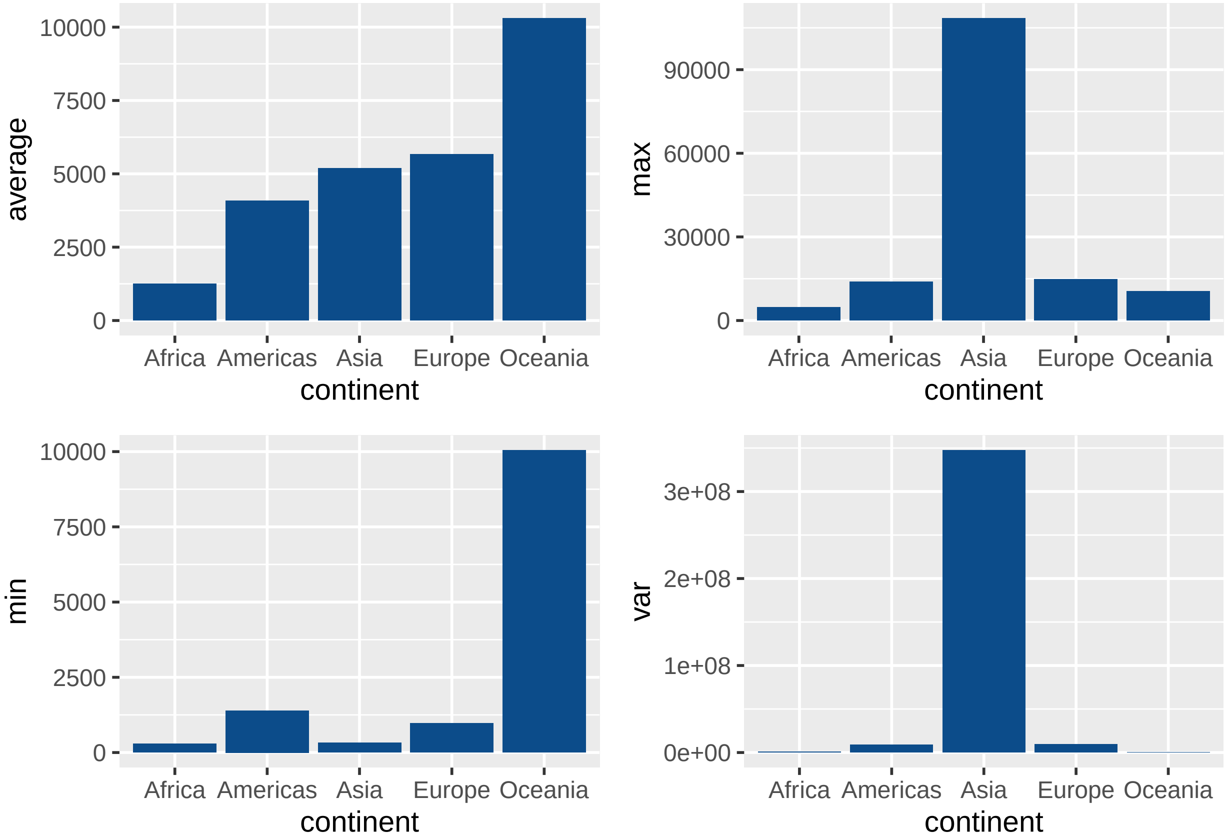
Table 1: Summary for gdpPercap in 1952

continent	average	max	min	var
Africa	1252.572	4725.296	298.8462	966194.9
Americas	4079.063	13990.482	1397.7171	9010368.1
Asia	5195.484	108382.353	331.0000	347259157.6
Europe	5661.057	14734.233	973.5332	9697372.8
Oceania	10298.086	10556.576	10039.5956	133634.2

```r
Summary_2007 = gapminder %>%
  filter(year == 2007) %>%
  group_by(continent) %>%
  summarise(average = mean(gdpPercap),
            max = max(gdpPercap),
            min = min(gdpPercap),
            var = var(gdpPercap)) %>%
  arrange(average)
kable(Summary_2007, caption = "Summary for gdpPercap in 2007", align = "c", booktab = T, format = "late
  kable_styling(latex_options = c("Hold_position", "scale_down"))
```

1.5 Product a plot that summarizes the same data as the table. There should be two plots per continent.

```r
a1 = ggplot(Summary_1952) +
  aes(x = continent, weight = average) +
  geom_bar(fill = "#0c4c8a") +
  labs(y = "average")

a2= ggplot(Summary_1952) +
 aes(x = continent, weight = max) +
 geom_bar(fill = "#0c4c8a") +
 labs(y = "max")

a3 = ggplot(Summary_1952) +
 aes(x = continent, weight = min) +
 geom_bar(fill = "#0c4c8a") +
 labs(y = "min")

a4 = ggplot(Summary_1952) +
 aes(x = continent, weight = var) +
 geom_bar(fill = "#0c4c8a") +
 labs(y = "var")

gridExtra::grid.arrange(a1, a2, a3, a4, ncol = 2)
```



1.6 Which countries in the dataset have had periods of negative population growth? Illustrate your answer with a table or plot.

```
neginc = gapminder %>%
  group_by(country) %>%
  summarise(t = sum(diff(pop) > 0), l = length(pop), negnumber = 11 - t) %>%
              filter(t < 11) %>%
              arrange(negnumber)
colnames(neginc) = c("Country", "", "", "# of year of negative pop growth")
neginc = cbind(neginc[1:9, ], neginc[10:18, ], neginc[19:27, ])
kable(neginc[, c(1, 4, 5, 8, 9, 12)], caption = "Countries had periods of negative population growth", a
  kable_styling(latex_options = c("HOLD_position")) %>%
  column_spec(c(1, 2, 3, 4, 5, 6), width = "7em")
```

Table 3: Countries had periods of negative population growth

Country	# of year of negative pop growth	Country.1	# of year of negative pop growth.1	Country.2	# of year of negative pop growth.2
Afghanistan	1	Montenegro	1	Germany	2
Cambodia	1	Portugal	1	Ireland	2
Croatia	1	Rwanda	1	Poland	2
Equatorial Guinea	1	Serbia	1	Slovenia	2
Guinea-Bissau	1	Somalia	1	Czech Republic	3
Kuwait	1	South Africa	1	Romania	3
Lebanon	1	Switzerland	1	Bulgaria	4
Lesotho	1	West Bank and Gaza	1	Trinidad and Tobago	4
Liberia	1	Bosnia and Herzegovina	2	Hungary	5

```
neginc$t=NULL
neginc$l=NULL
```

1.7 Which countries in the dataset have had the highest rate of growth in per capita GDP? Illustrate your answer with a table or plot.

```
gapminder$'Log_gdpC' = log(gapminder$gdpPercap)
growthrate = gapminder %>%
  group_by(country) %>%
  summarise(Max_GR = max(diff(Log_gdpC))) %>%
  arrange(desc(Max_GR))
kable(growthrate[1:10, ], format = "latex", booktab=T, align = "c", caption = "Log Growth Rate") %>%
  kable_styling(latex_options = "HOLD_position")
```

Table 4: Log Growth Rate

country	Max_GR
Libya	1.0218229
Equatorial Guinea	1.0068965
Oman	0.8105458
Cambodia	0.6482633
Gabon	0.6456260
Bosnia and Herzegovina	0.6267517
Botswana	0.6224566
Angola	0.5480056
Singapore	0.5465899
Korea, Dem. Rep.	0.5463120

Problem 2

The data for Problem 2 is the Fertility data in the AER package. This data is from the 1980 US Census and is comprised of date on married women aged 21-35 with two or more children. The data report the gender of each woman's first and second child, the woman's race, age, number of weeks worked in 1979, and whether the woman had more than two children.

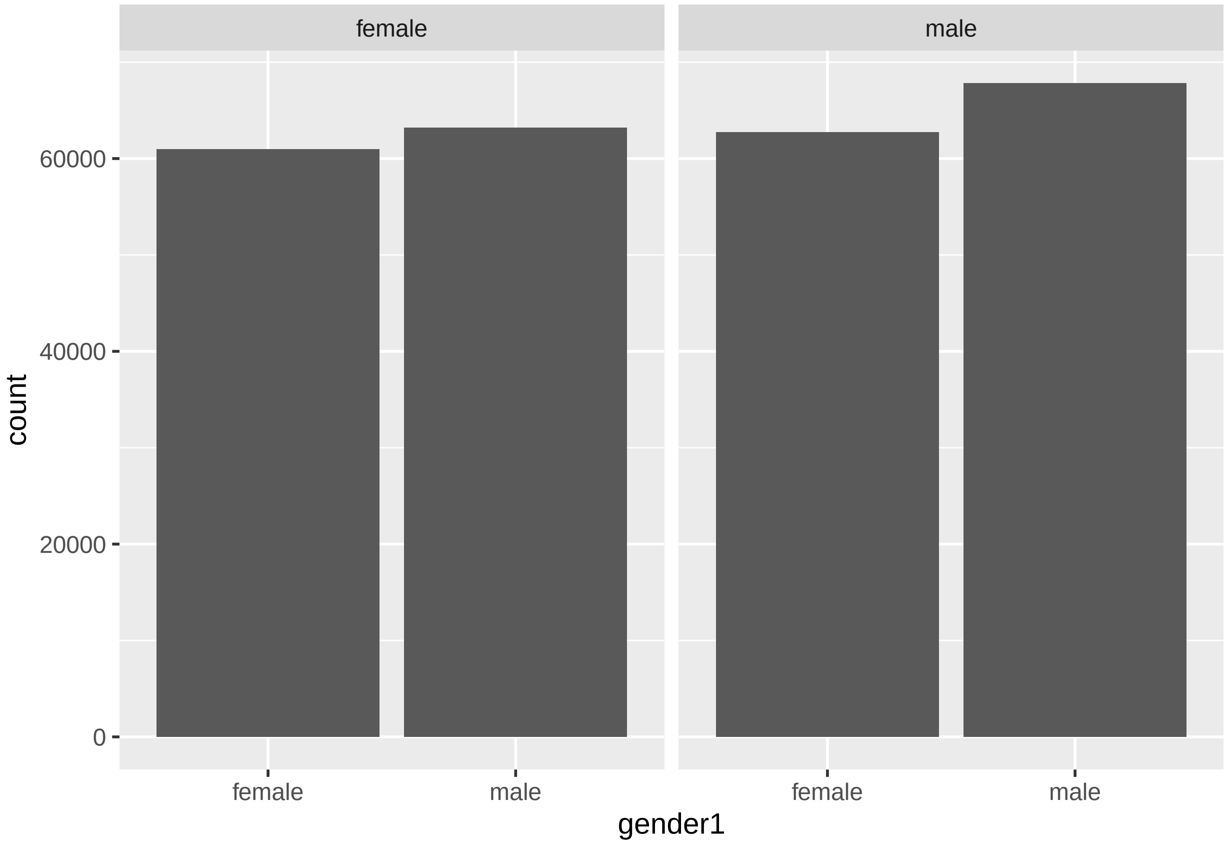
```
library(AER)
```

```
## Loading required package: car

## Loading required package: carData

##
## Attaching package: 'car'

## The following object is masked from 'package:expss':
##
##     recode

## The following object is masked from 'package:dplyr':
##
##     recode

## The following object is masked from 'package:purrr':
##
##     some

## Loading required package: lmtest

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

## Loading required package: sandwich

## Loading required package: survival
```

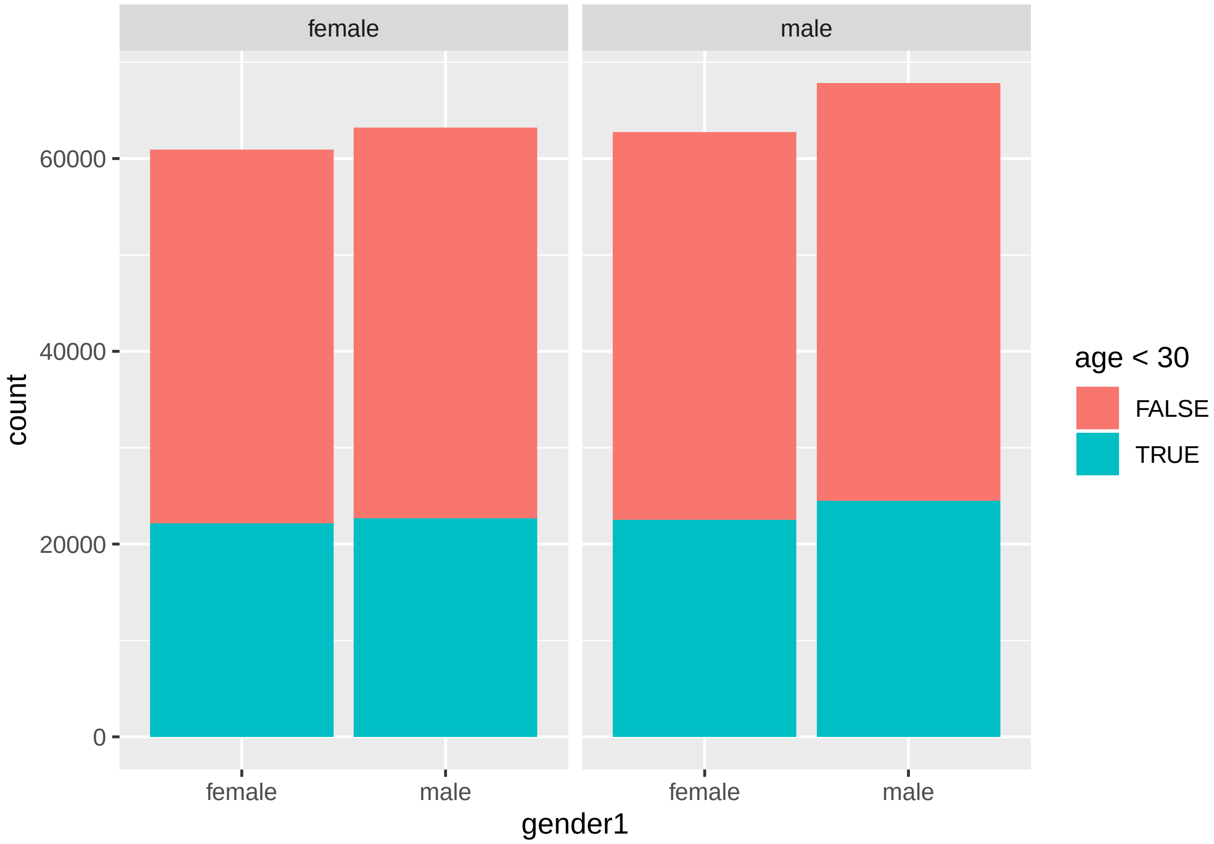
```
data("Fertility")
```

2.1 There are four possible gender combinations for the first two Children. Product a plot the contracts the frequency of these four combinations. Are the frequencies different for women in their 20s and wemen who are older than 29?

```
f_in20s<-Fertility %>% filter(age <30)
f_out20s<-Fertility %>% filter(age >=30)
ggplot(data = Fertility)+
  geom_bar(mapping = aes(x=gender1))+
  facet_grid(.~gender2)
```



```
ggplot(data = Fertility)+
  geom_bar(mapping = aes(x=gender1,fill = age <30))+
  facet_grid(.~gender2)
```

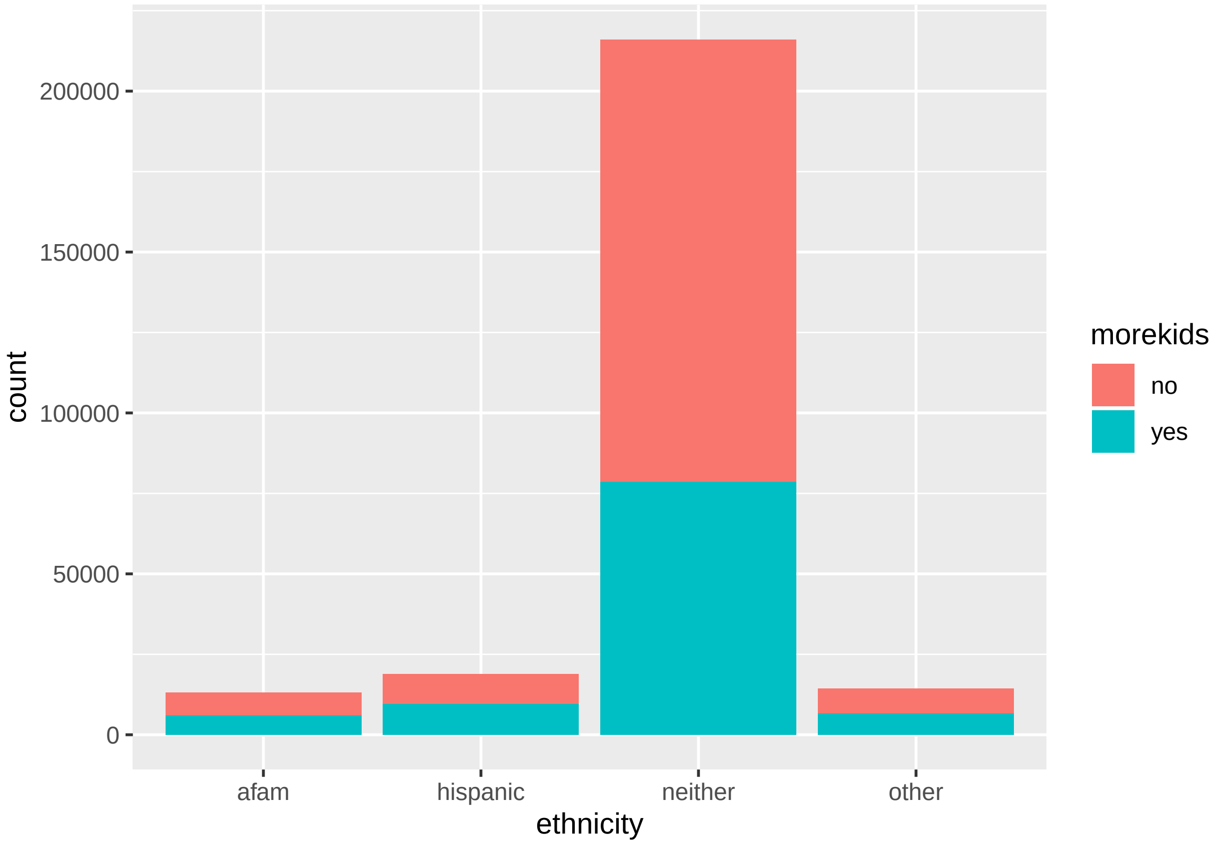


2.2 Produce a plot that contrasts the frequency of having more than two children by race and ethnicity.

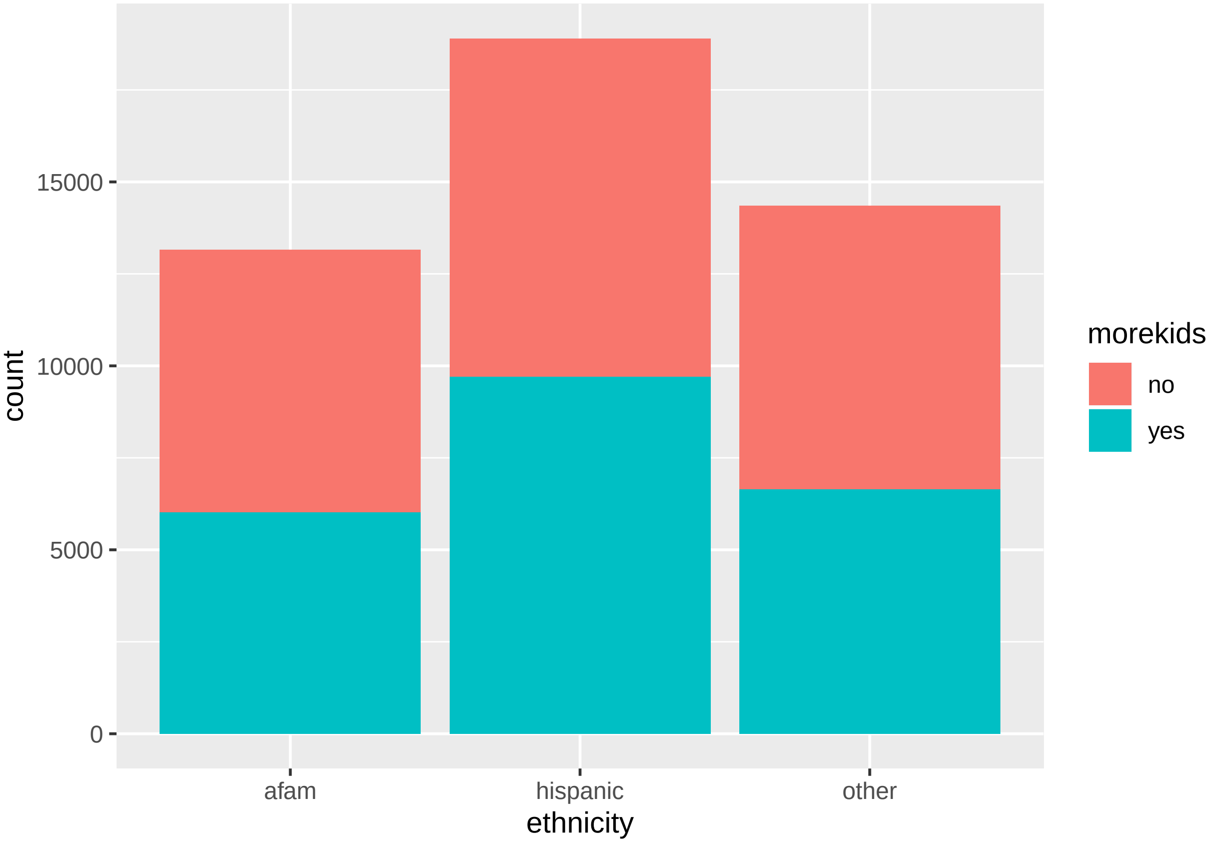
```
f3 <- Fertility %>%
  mutate(neither = (afam == "no" & hispanic == "no" & other == "no") )
f4 <- f3%>%
  within(neither[neither == TRUE]<- "yes")
f_race <-f4 %>% gather(`afam`,`hispanic`,`other`,`neither`, key = ethnicity, value = "yes")%>%
  filter(yes == "yes")
```

```
## Warning: attributes are not identical across measure variables;
## they will be dropped
```

```
ggplot(data = f_race)+
  geom_bar(mapping =aes(x=ethnicity,fill = morekids))
```



```
f_test <- f3 %>%
  filter(afam=="yes" & hispanic == "yes")

f_race_only_three <-Fertility %>% gather(`afam`,`hispanic`,`other`, key = ethnicity, value = "yes")%>%
  filter(yes == "yes")
ggplot(data = f_race_only_three)+
  geom_bar(mapping =aes(x=ethnicity,fill = morekids))
```



### Problem 3 Use the mtcars and mpg datasets.
3.1 How many times does the letter "e" occur in mtcars rownames?

```
data(mpg)
data("mtcars")
mtcars<-as_tibble(rownames_to_column(mtcars,var="Model"))
mtcars$ecars<-str_count(mtcars$Model,"e")
sum(mtcars$ecars)
```

```
## [1] 25
```

3.2 How many cars in mtcars have the brand Merc?

```
sum(str_count(mtcars$Model,"Merc"))
```

```
## [1] 7
```

3.3 How many cars in mpg have the brand("manufacturer" in mpg) Merc?

```
sum(str_count(mpg$manufacturer,"mercury"))
```

```
## [1] 4
```

3.4 Contrast the mileage data for Merc cars as reported in mtcars and mpg. Use tables, plots, and a short explaination.

"' ### Problem 4 Install the babynames package.

```
library(babynames)
data("babynames")
```

4.1 Draw a sample of 500,000 rows from the babynames data

```r
sample<-sample(1:1924665,500000,replace=F)
sampledata<-babynames[sample,]
##Produce a tabble that displays the five most popular boy names and girl names in the years 1880,1920,
y=c(1880,1920,1960,2000)
gender<-c("F","M")
re=NULL
for (i in y){
  for (j in gender){
    a=filter(sampledata,year==i,sex==j) %>% arrange(desc(n))
    re=rbind(re,as.matrix(a[1:5,]))
  }
}
kable(re,caption="most popular name each year", align = "c",booktab=T,format="latex") %>%
  kable_styling(latex_options = c("HOLD_position"))
```

Table 5: most popular name each year

year	sex	name	n	prop
1880	F	Mary	7065	0.07238359
1880	F	Clara	1226	0.01256083
1880	F	Martha	1040	0.01065519
1880	F	Nellie	995	0.01019415
1880	F	Jennie	793	0.00812458
1880	M	James	5927	0.05005912
1880	M	Edward	2364	0.01996622
1880	M	Albert	1493	0.01260980
1880	M	Joe	731	0.00617399
1880	M	Clarence	730	0.00616554
1920	F	Dorothy	36643	0.02945486
1920	F	Virginia	17314	0.01391757
1920	F	Marie	12743	0.01024325
1920	F	Martha	8709	0.00700058
1920	F	Marjorie	8659	0.00696039
1920	M	William	50147	0.04555435
1920	M	Thomas	14938	0.01356992
1920	M	Raymond	12194	0.01107723
1920	M	Arthur	10236	0.00929855
1920	M	Harry	9408	0.00854638
1960	F	Mary	51474	0.02474901
1960	F	Patricia	32102	0.01543483
1960	F	Debra	26737	0.01285531
1960	F	Nancy	21896	0.01052773
1960	F	Diane	17900	0.00860643
1960	M	Robert	72369	0.03341605
1960	M	Mark	58731	0.02711876
1960	M	Steven	33895	0.01565086
1960	M	Timothy	30484	0.01407584
1960	M	Charles	29676	0.01370275
2000	F	Jessica	15709	0.00787466
2000	F	Elizabeth	15094	0.00756637
2000	F	Abigail	13088	0.00656079
2000	F	Megan	11434	0.00573167
2000	F	Rachel	10673	0.00535019
2000	M	Christopher	24931	0.01194362
2000	M	Brandon	20336	0.00974231
2000	M	John	20092	0.00962541
2000	M	Anthony	19648	0.00941271
2000	M	Christian	16056	0.00769190

4.2 What names overlap boys and girls?

```
overlap<-sampledata %>% group_by(year,name) %>% summarise(count=length(sex)) %>% arrange(desc(count)) %
  filter(count>1)
unique(overlap$name)[1:10]
```

```
##  [1] "Augustine" "Clare"     "Edith"     "Ethel"     "Jennie"
##  [6] "Alice"     "Claude"    "Francis"   "Joseph"    "Odie"
```

4.3 What names were used in the 19th century but have not been used in the 21sth century?

```
l1 = sampledata %>%
  filter(year > 1999)
l1 = unique(l1$name)
l2 = sampledata %>%
  filter(year < 1900)
l2 = unique(l2$name)
Int = intersect(l1, l2)
Int[1:10]
```

```
##  [1] "Huey"      "Albion"    "Sebastian" "Gregoria"  "Yancy"
##  [6] "Samantha"  "Cruz"      "Williams"  "Leda"      "Leander"
```

4.4 Produce a chart that shows the relative frequency of the names "Donald", "Hilary", "Hillary", "Joe", "Barrack", over the years 1880 through 2017.

```
chart<-sampledata %>%
  filter(year>1879 & year <2018) %>%
  group_by(year,name) %>%
  summarise(cou=sum(n)) %>%
  filter(name==c("Donald","Hilary","Joe","Barrack")) %>%
  group_by(name) %>%
  summarise(count=sum(cou))
```

```
## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length
## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length
## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length
## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length
## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length
## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length
## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length

## Warning in name == c("Donald", "Hilary", "Joe", "Barrack"): longer object
## length is not a multiple of shorter object length
```

```
chart$"frequency"<-round(chart$count/sum(chart$count), 2)
kable(chart,align="c") %>%
  kable_styling(latex_options="HOLD_position")
```

name	count	frequency
Donald	112493	0.82
Hilary	801	0.01
Joe	23323	0.17