ANA 515 ASSIGNMENT 2

#Library Setup

```
knitr::opts_chunk$set(echo = TRUE)
library(ggplot2)
library(readr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
```

```
## v tibble  3.1.7     v stringr 1.4.0
## v tidyr   1.2.0     v forcats 0.5.1
## v purrr   0.3.4
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(knitr)
library(bslib)
```

```
##
## Attaching package: 'bslib'
##
## The following object is masked from 'package:utils':
##
##     page
```

```
library(magrittr)
```

```
##
## Attaching package: 'magrittr'
##
## The following object is masked from 'package:purrr':
##
##     set_names
##
## The following object is masked from 'package:tidyr':
##
##     extract
```

```
#This next chunk is inline code. Inline code puts the text with the output of the function in my docume
#Description


#This dataset is predominantly rooted in an article that 1057 participants responded to about their sle
```

#Loading Data

```
sleep_csv <- read.csv("/Users/harshsanghvi/Downloads/Sleepseparate.csv")

#The data was downloaded from https://github.com/fivethirtyeight/data/blob/master/sleeping-alone-data/s
```

#Data Cleaning

```
#Creating a custom summary for all variables and their type
custom_glimpse <- function(sleep_csv) {
  data.frame(
    col_name = colnames(sleep_csv),
    col_index = 1:ncol(sleep_csv),
    col_class = sapply(sleep_csv, class),
    row.names = NULL
  )
}
custom_glimpse(sleep_csv)
```

```
##
## 1
## 2
## 3
## 4                                                                                    Which.of.the.followin
## 5  How.long.have.you.been.in.your.current.relationship..If.you.are.not.currently.in.a.relationship..
## 6                                                                        When.both.you.and.your.partner.are
## 7
## 8                                                                  When.you.re.not.sleeping.in.the.same.
## 9
## 10                                                                      When.you.re.not.sleeping.in.the.
## 11
## 12
## 13                                                                                                    On
## 14
## 15
## 16                                                                                                We.ha
```

```
## 17
## 18
## 19
## 20
## 21
## 22                                                                            Wh
## 23                                  To.what.extent.do.you.agree.with.the.following.statement...sle
## 24                                       To.what.extent.do.you.agree.with.the.following.statement.
## 25                  To.what.extent.do.you.agree.with.the.following.statement.._.our.sex.life.has.impr
## 26                                                                     Which.of.the
## 27
## 28
## 29
## 30
## 31
## 32
##    col_index col_class
## 1          1 character
## 2          2 character
## 3          3 character
## 4          4 character
## 5          5 character
## 6          6 character
## 7          7   integer
## 8          8 character
## 9          9 character
## 10        10 character
## 11        11 character
## 12        12 character
## 13        13 character
## 14        14 character
## 15        15 character
## 16        16 character
## 17        17 character
## 18        18 character
## 19        19 character
## 20        20 character
## 21        21 character
## 22        22 character
## 23        23 character
## 24        24 character
## 25        25 character
## 26        26 character
## 27        27 character
## 28        28 character
## 29        29 character
## 30        30 character
## 31        31 character
## 32        32 character
```

```
#Renaming variables
cleansleep_csv <- sleep_csv %>%
       rename(StartTime = 1, EndTime = 2, TimeTaken = 3, RelationshipStatus = 4, LengthRelationship = 5
custom_glimpse(cleansleep_csv)
```

```
##                              col_name col_index col_class
## 1                           StartTime         1 character
## 2                             EndTime         2 character
## 3                           TimeTaken         3 character
## 4                  RelationshipStatus         4 character
## 5                  LengthRelationship         5 character
## 6              SleepSeparateFrequency         6 character
## 7          TotalCountForSleepSeparate         7   integer
## 8           SleepSeparateSelfLocation         8 character
## 9      SleepSeparateSelfLocationOther         9 character
## 10       SleepSeparatePartnerLocation        10 character
## 11 SleepSeparatePartnerLocationOther        11 character
## 12                         ReasonSnores        12 character
## 13                ReasonBathroomUsage        13 character
## 14                          ReasonSick        14 character
## 15                   ReasonNoIntimacy        15 character
## 16               ReasonTempPreference        16 character
## 17              ReasonFightOrArgument        17 character
## 18               ReasonNotEnoughSpace        18 character
## 19                ReasonNotShareCover        19 character
## 20             ReasonChildNeedsParent        20 character
## 21      ReasonDifferentSleepSchedules        21 character
## 22             SleptSeparateFirstTime        22 character
## 23       SeparateHelpedToStayTogether        23 character
## 24         SeparateHelpedSleepQuality        24 character
## 25            SeparateHelpedToSexLife        25 character
## 26                         Occupation        26 character
## 27                   OccupactionOther        27 character
## 28                             Gender        28 character
## 29                                Age        29 character
## 30                    HouseholdIncome        30 character
## 31                          Education        31 character
## 32                             Region        32 character
```

```
#Inline Code
```

This dataset has 1093 rows and 32 variables.

#Column Description

custom_glimpse2 <- function(description_csv) { data.frame(col_name = colnames(cleansleep_csv), description = colnames(sleep_csv), row.names = NULL) } custom_glimpse2(description_csv)

```
#Subset and Exclude
```

```
```r
data_pick3 <- select(cleansleep_csv, LengthRelationship, SleepSeparateFrequency, Age)

data_pick3$Age %<>% factor
data_pick3$LengthRelationship %<>% factor
data_pick3$SleepSeparateFrequency %<>% factor
view(data_pick3)
str(data_pick3$LengthRelationship)
```
```

```
##  Factor w/ 7 levels "","1-5 years",..: 2 2 2 2 2 2 6 7 7 7 ...
```

```
unique (data_pick3$LengthRelationship)
```

```
## [1] 1-5 years          Less than 1 year   More than 20 years 6-10 years
## [5] 16-20 years        11-15 years
## 7 Levels:  1-5 years 11-15 years 16-20 years 6-10 years ... More than 20 years
```

```
unique (data_pick3$SleepSeparateFrequency)
```

```
## [1] Once a year or less   A few times per month Never
## [4] Every night           Once a month or less  A few times per week
## [7]
## 7 Levels:  A few times per month A few times per week Every night ... Once a year or less
```

```
unique (data_pick3$Age)
```

```
## [1] 18-29       45-60 > 60  30-44
## Levels:  > 60 18-29 30-44 45-60
```

```
excludeempty_csv <- data_pick3[!(data_pick3$LengthRelationship=="" |data_pick3$SleepSeparateFrequency==
```

#Summary

```
summary(data_pick3)
```

```
##           LengthRelationship           SleepSeparateFrequency    Age
##                    : 14                          : 14             : 88
##  1-5 years         :162      A few times per month: 62        > 60 :283
##  11-15 years       :159      A few times per week : 46        18-29:122
##  16-20 years       :104      Every night          :147        30-44:288
##  6-10 years        :165      Never                :586        45-60:312
##  Less than 1 year  : 21      Once a month or less :101
##  More than 20 years:468      Once a year or less  :137
```

```
summarise_all(data_pick3["SleepSeparateFrequency"], funs(nlevels(.), nmiss=sum(is.na(.))))
```

```
## Warning: `funs()` was deprecated in dplyr 0.8.0.
## Please use a list of either functions or lambdas:
##
##   # Simple named list:
##   list(mean = mean, median = median)
##
##   # Auto named with `tibble::lst()`:
##   tibble::lst(mean, median)
##
##   # Using lambdas
##   list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was generated.
```

```
##   nlevels nmiss
## 1       7     0
```

```
summarise_all(data_pick3["LengthRelationship"], funs(nlevels(.), nmiss=sum(is.na(.))))
```

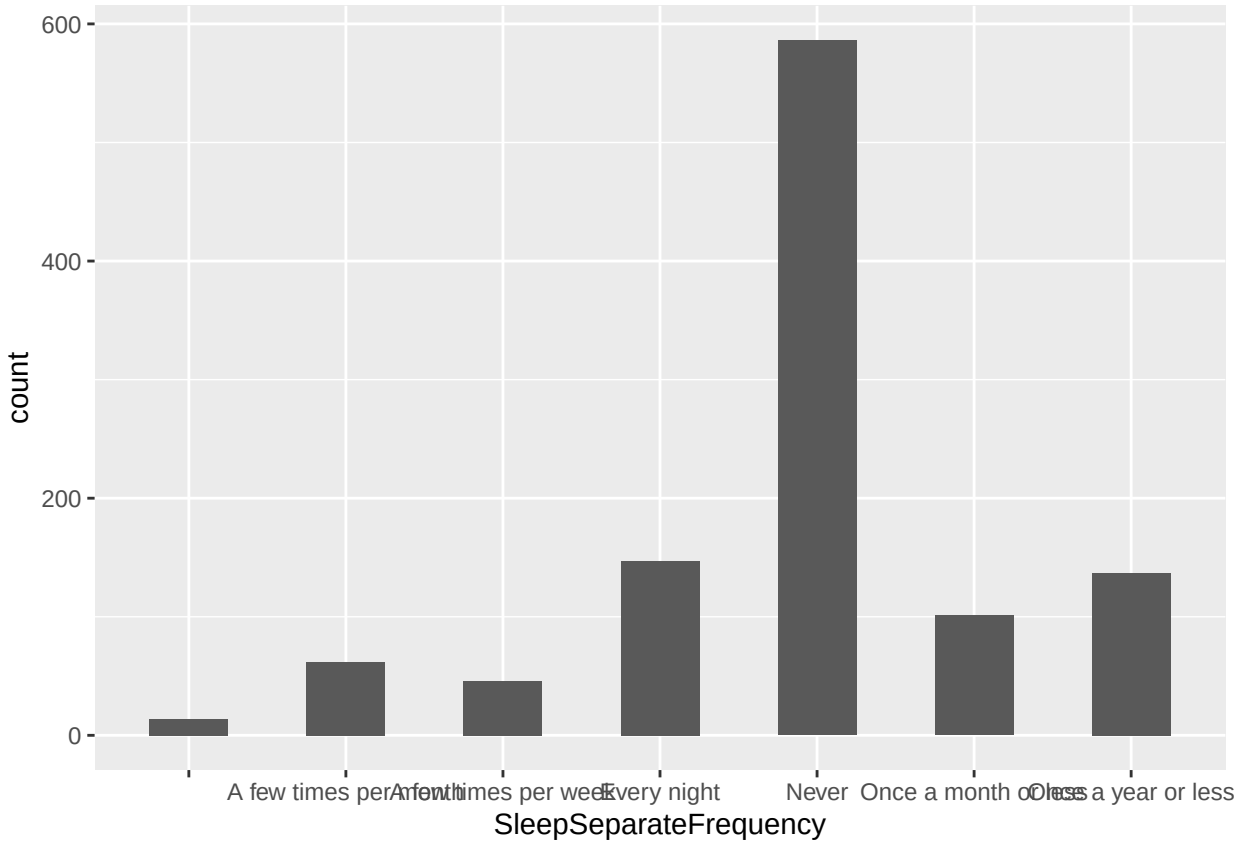
```
##   nlevels nmiss
## 1       7     0
```

```
summarise_all(data_pick3["Age"], funs(nlevels(.), nmiss=sum(is.na(.))))
```

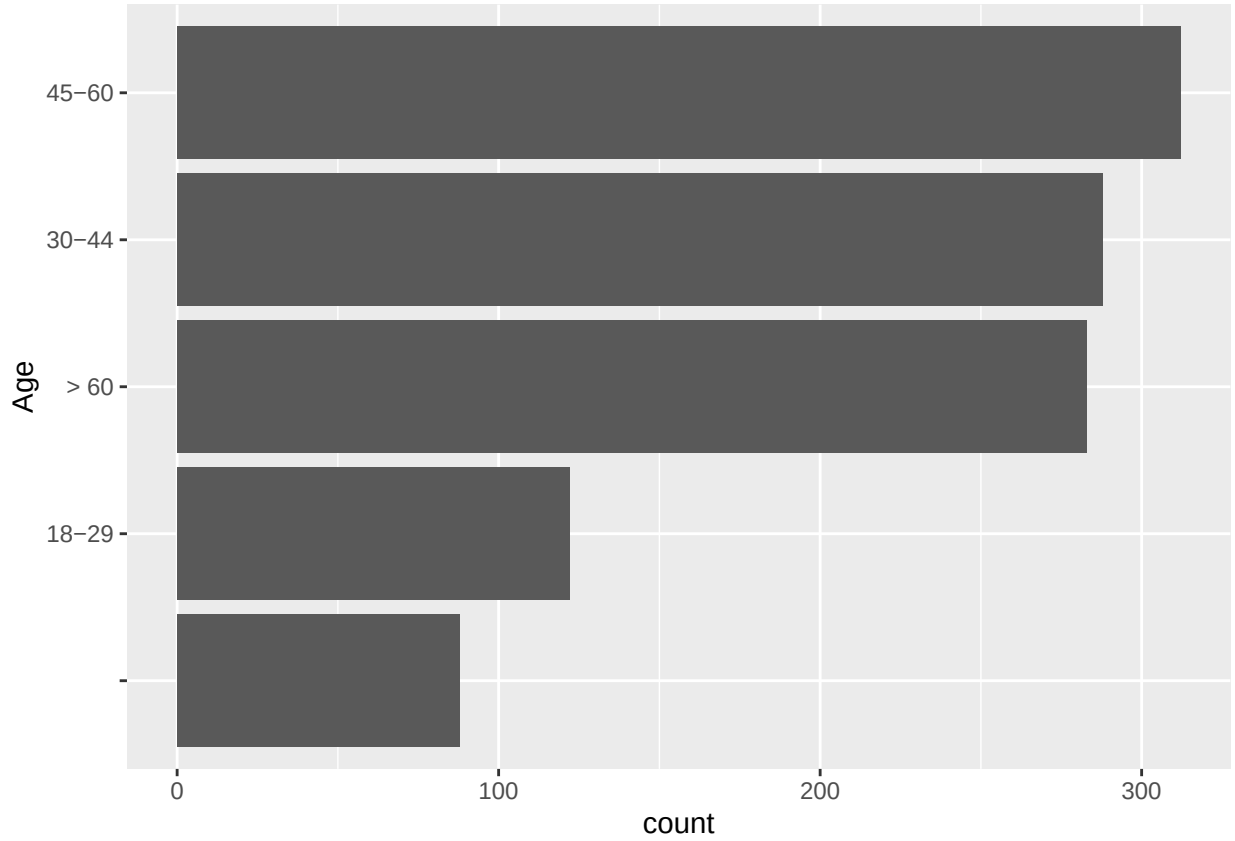
```
##   nlevels nmiss
## 1       5     0
```

Including Plots

```
data_pick3 %>%
ggplot(aes(SleepSeparateFrequency)) +
stat_count(width = 0.5)
```




```
data_pick3 %>%
ggplot(aes(fct_infreq(Age) %>% fct_rev())) +
geom_bar() + coord_flip() +
labs(x = "Age")
```


45−60
30−44
> 60
18−29
Age
0
100
200
300
count