HW_12

izd3

Problem #01 - Chapter 44 Exercise #2A

```r
# Show your work here
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.2.3

## Warning: package 'ggplot2' was built under R version 4.2.3

## Warning: package 'tibble' was built under R version 4.2.3

## Warning: package 'purrr' was built under R version 4.2.3

## Warning: package 'dplyr' was built under R version 4.2.3

## Warning: package 'stringr' was built under R version 4.2.3

## Warning: package 'forcats' was built under R version 4.2.3

## Warning: package 'lubridate' was built under R version 4.2.3

## ── Attaching core tidyverse packages ──────────────────────── tidyverse
2.0.0 ──
## ✓ dplyr      1.1.2      ✓ readr      2.1.4
## ✓ forcats    1.0.0      ✓ stringr    1.5.0
## ✓ ggplot2    3.4.3      ✓ tibble     3.2.1
## ✓ lubridate  1.9.3      ✓ tidyr      1.3.0
## ✓ purrr      1.0.2
## ── 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
```

```r
Joins002.tib$FIRSTNAME=tolower(Joins002.tib$FIRSTNAME)
Joins002.tib$LASTNAME=tolower(Joins002.tib$LASTNAME)


Joins003.dat$firstname=tolower(Joins003.dat$firstname)
Joins003.dat$lastname=tolower(Joins003.dat$lastname)


inner_join(Joins002.tib,Joins003.dat,by=join_by(FIRSTNAME=="firstname",
                                                 LASTNAME=="lastname"))
```

```
## # A tibble: 1 × 4
##    LASTNAME  FIRSTNAME `Favorite Color` favoritecolor
##    <chr>     <chr>     <chr>            <chr>
## 1 deherrera jacob     honeydew4        navajowhite1
```

Problem #02 - Chapter 44 Exercise #02A

```
# Show your work here
Joins004.tib

## # A tibble: 20 × 4
##    LastNames        firstNames registrationCode dataStuff
##    <chr>            <chr>      <chr>                <dbl>
##  1 al-Azad          Jessica    282                   99.1
##  2 Morett           Matthew    244                   86.8
##  3 Nelson           Tasneem    371                   86.7
##  4 Hallam           Robert     174                   85.8
##  5 Munoz Torres     Wesley     469                   93.8
##  6 Joel             Draven     696                   94.0
##  7 el-Tariq         Taariq     683                   92.0
##  8 Nelson           Tasneem    328                   92.0
##  9 el-Tariq         Taariq     260                   84.5
## 10 Deherrera        Jacob      392                   84.9
## 11 Vogt             Chantelle  175                   87.4
## 12 Morett           Matthew    357                   94.6
## 13 Nelson           Tasneem    029                   82.5
## 14 Hallam           Robert     525                   88.1
## 15 Morett           Matthew    310                   99.8
## 16 Morett           Matthew    182                   81.0
## 17 Williams Sanders Albert     737                   94.3
## 18 al-Azer          Noel       783                   86.9
## 19 Conner           Deshaun    975                   94.7
## 20 Nelson           Tasneem    206                   97.1

Joins005.tib

## # A tibble: 21 × 5
##    LastNames        firstNames ID_NUMBER dataThings coloredThings
##    <chr>            <chr>      <chr>          <dbl> <chr>
##  1 Morett           Matthew    182             81.0 grey71
##  2 Conner           Deshaun    975             94.7 steelblue4
##  3 el-Tariq         Taariq     683             92.0 sienna1
##  4 Williams Sanders Albert     737             94.3 grey40
##  5 al-Azad          Jessica    282             99.1 goldenrod4
##  6 Munoz Torres     Wesley     161             98.0 cornsilk2
##  7 Munoz Torres     Wesley     469             93.8 darkslategray4
##  8 Villamil Buenfil Renita     525             85.0 navajowhite
##  9 Vogt             Chantelle  175             87.4 hotpink2
## 10 Deherrera        Jacob      169             85.7 slategray3
## # i 11 more rows

inner_join(Joins004.tib,Joins005.tib,
          by=join_by("registrationCode"=="ID_NUMBER",
                                      "LastNames"=="LastNames",
                                      "firstNames"=="firstNames"))
```

```
## # A tibble: 14 × 6
##    LastNames      firstNames registrationCode dataStuff dataThings
coloredThings
##    <chr>          <chr>      <chr>                <dbl>      <dbl> <chr>
##  1 al-Azad        Jessica    282                   99.1       99.1
goldenrod4
##  2 Morett         Matthew    244                   86.8       86.8 green1
##  3 Munoz Torres   Wesley     469                   93.8       93.8
darkslategra…
##  4 Joel           Draven     696                   94.0       94.0
cornflowerbl…
##  5 el-Tariq       Taariq     683                   92.0       92.0 sienna1
##  6 el-Tariq       Taariq     260                   84.5       84.5 bisque3
##  7 Deherrera      Jacob      392                   84.9       84.9 grey8
##  8 Vogt           Chantelle  175                   87.4       87.4
hotpink2
##  9 Morett         Matthew    357                   94.6       94.6 thistle
## 10 Nelson         Tasneem    029                   82.5       82.5
palevioletre…
## 11 Morett         Matthew    182                   81.0       81.0 grey71
## 12 Williams Sand… Albert     737                   94.3       94.3 grey40
## 13 Conner         Deshaun    975                   94.7       94.7
steelblue4
## 14 Nelson         Tasneem    206                   97.1       97.1
hotpink1
```
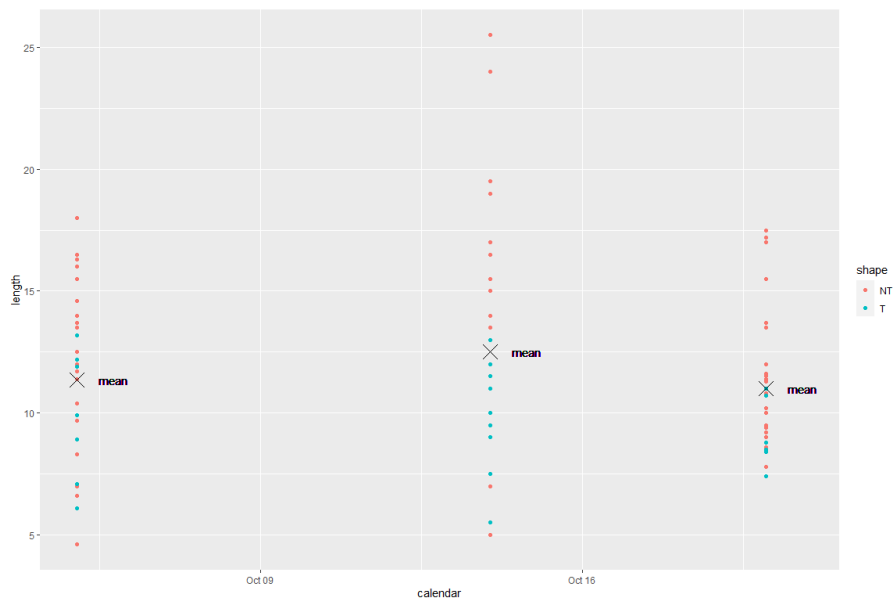
Problem #03 - Chapter 45 Exercise #04

```r
# Show your work here
joined<-full_join(Joins006.tib,Joins008.tib,by=c('X3','X1','X2'))

joined<-full_join(joined,Joins007.tib,
                  by=join_by('X2'=="shape",'X3'=='calendar',
                                              'X1'=='length'))
mean_vals<-joined|>
  group_by(X3)|>
  summarise(
    mean_len=mean(X1)
  )
mean_vals$X3=mdy(mean_vals$X3)

joined$X3=mdy(joined$X3)
joined<-joined|>
  full_join(mean_vals,by='X3')
joined|>
  ggplot(aes(x=X3,y=X1,color=X2))+geom_point()+
  scale_x_date(breaks = make_date(month = 10,day = c(9,16),year = 2023),
               date_labels = "%b %d")+
  labs(
    x='calendar',
    y='length',
    color='shape'
  )+
  geom_point(aes(x=X3,y=mean_len),size=6,shape=4,color='black')+
  geom_text(aes(x=X3,y=mean_len),
            label="mean",color="black",nudge_x = 0.8)
```

Problem #04 - Chapter 48 Exercise #2

```
# Show your work here
Long001b.tib

## # A tibble: 6 × 3
##    COLOUR          TYPE alphabet
##    <chr>          <int> <chr>
## 1 mediumturquoise    1 E
## 2 mediumturquoise    2 P
## 3 mediumturquoise    3 G
## 4 peru               1 C
## 5 peru               2 N
## 6 peru               3 U

MyLongTibble<-Wide001.tib|>
  pivot_longer(
    cols = !COLOUR,
    names_to=c("Discard","TYPE"),
    names_sep = "e",
    values_to = c("alphabet")
  )|>
  select(!Discard)|>
  mutate(TYPE=as.integer(TYPE))

MyLongTibble

## # A tibble: 6 × 3
##    COLOUR          TYPE alphabet
##    <chr>          <int> <chr>
## 1 mediumturquoise    1 E
## 2 mediumturquoise    2 P
## 3 mediumturquoise    3 G
## 4 peru               1 C
## 5 peru               2 N
## 6 peru               3 U
```
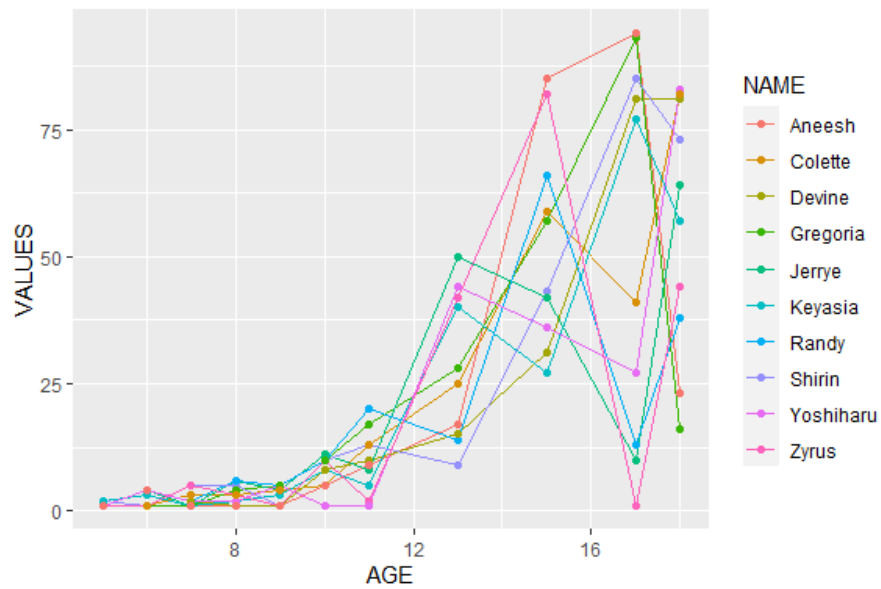
Problem #05 - Chapter 48 Exercise #06

```r
# Show your work here
newtib<-Wide003.tib|>
  pivot_longer(cols = !NAME,
               names_to = c("Discard","AGE"),
               names_sep = "E",
               values_to = "VALUES")|>
  select(!Discard)|>
  mutate(AGE=as.integer(AGE))

newtib|>
  ggplot(aes(x=AGE,y=VALUES,color=NAME))+geom_line()+geom_point()
```

Problem #06 - Chapter 49 Exercise #2

```
# Show your work here
Wide001.tib

## # A tibble: 2 × 4
##   COLOUR          type1 type2 type3
##   <chr>           <chr> <chr> <chr>
## 1 mediumturquoise E     P     G
## 2 peru            C     N     U

MyWideTibble<-Long001a.tib|>
  pivot_wider(id_cols = COLOUR,names_from = TYPE,values_from = alphabet)

MyWideTibble

## # A tibble: 2 × 4
##   COLOUR          type1 type2 type3
##   <chr>           <chr> <chr> <chr>
## 1 mediumturquoise E     P     G
## 2 peru            C     N     U
```

Problem #07 - Chapter 50 Exercise #01

```r
# Show your work here
MyLongTibble<-Split001.tib|>
  separate(col = stuff,
           into = c("NUMBER","UPPER","LOWER","LOGICAL"),sep = "-",convert =
T)
MyLongTibble

## # A tibble: 10 × 4
##    NUMBER UPPER LOWER LOGICAL
##     <int> <chr> <chr> <lgl>
##  1     26 C     d     FALSE
##  2     37 J     m     FALSE
##  3      9 K     v     FALSE
##  4     55 A     r     TRUE
##  5     41 N     p     TRUE
##  6     62 X     x     TRUE
##  7     59 Z     o     TRUE
##  8      1 R     c     FALSE
##  9     20 O     s     TRUE
## 10     87 F     z     TRUE

Separated001.tib

## # A tibble: 10 × 4
##    NUMBER UPPER LOWER LOGICAL
##     <chr>  <chr> <chr> <chr>
##  1 26     C     d     FALSE
##  2 37     J     m     FALSE
##  3 9      K     v     FALSE
##  4 55     A     r     TRUE
##  5 41     N     p     TRUE
##  6 62     X     x     TRUE
##  7 59     Z     o     TRUE
##  8 1      R     c     FALSE
##  9 20     O     s     TRUE
## 10 87     F     z     TRUE
```

Problem #08 - Chapter 50 Exercise #03

```
# Show your work here
Split002.tib|>
  separate(col = LotsOfData,into = c("Discard","LOCATION"),sep ="LEAF_" )|>
  select(!Discard)|>
  separate(col = LOCATION,into = c("DISCARD","LOCATION"),
           sep ="[0-9][0-9][0-9]_[0-9][0-9][0-9]_" )|>
  select(!DISCARD)|>
  separate(col = LOCATION,into = c("LOCATION","RID"))|>
  select(!RID)|>
  ggplot(aes(x=shapes,fill=LOCATION))+geom_bar()+
  scale_fill_manual(breaks = c("AGRI","ARTS"),
                    labels=c("Ag Quad","Arts Quad"),
                    values = c("red","white"),
                    name="location")+
  labs(
    title="This is the title"
  )
```