Parth Vohra
102016044
3CS10

Assignment 6

```
country<-c('india', 'usa', 'china', 'germany', 'russia', 'india', 'usa', 'china', 'germany', 'russia','india', 'usa', 'china', 'german
continent<-c('asia', 'america', 'asia', 'europe', 'europe', 'asia', 'america', 'asia', 'europe', 'europe', 'asia', 'america', 'asia',
year<-seq(1:20)
lifeexp<-c(11.12, 81, 90, 16.12, 16.22, 11.12, 99, 18.25, 16.12, 16.22, 11.12, 19.2, 18.25, 16.12, 16.22, 11.12, 19.2, 18.25, 16.12, 1
pop<-c(1080, 2200, 3900, 4100, 5300, 1100, 4100, 3001, 4019, 5201, 1254, 2164, 3000, 4000, 5000, 1100, 9000, 3200, 4001, 5890)
gdpPerc<-c(1.1, 5.5, 2.2, 4.4, 30.3, 1.1, 5.5, 2.2, 4.4, 3.3, 1.1, 5.5, 2.2, 4.4, 3.3, 1.1, 5.5, 2.2, 4.4, 3.3)
df=data.frame(country, continent, year, lifeexp, pop, gdpPerc)
head(df,2)

# Ques: 1
aggregate(df$country, by=list(continent=df$continent), FUN=table)

# Ques: 2
df_raw=filter(df, continent=='europe')
filter(df_raw, gdpPerc==min(gdpPerc))

# Ques: 3
aa<- df %>% group_by(continent, year) %>% summarise(avg=mean(lifeexp))
aa

# Ques: 4
bb<-df %>% group_by(country) %>% summarise(total=sum(gdpPerc))
bb

sort(bb, total)


# Ques: 5
aa1<-filter(df, lifeexp>80)
aa1$country
aa1$year


# Ques: 6
aa2<-df %>% group_by(country) %>% summarise(correlation=cor(df$lifeexp, df$gdpPerc))
aa2

# Ques: 7
df1=subset(df, continent != 'asia')
df1

arrange(df1, desc=df1$pop)
tail(df1, 1)

# Ques: 8
as<-tapply(df$pop, df$country, sd)
as<-sort(as)
as[1:3]

# Ques: 9
#-------


## Ques: Q2
# Ques: 1
med<-read.csv("med.csv")
med
med$exp_date<-as.Date(med$exp_date, "%m/%d/%Y")

# Ques: 2
head(med, 4)

# Ques: 3
tail(med, 4)

# Ques: 4
days_left=(as.numeric)(med$exp_date-Sys.Date())
cor(med$quantity_in_stock, days_left)

# Ques: 5
barplot(med$sales, med$manf_year)

# Ques: 6
for(i in unique(med$company)){
  if(sum(med$company == i) > 1){
    print(i)
  }
}
```

```
# Ques: 7
med=mutate(med, days_left=days_left)
filter(med, days_left>0)

# Ques: 8
boxplot(med$days_left, data=med)
print("Medicine below 0 are expired")

# Ques: 9
mean(med$quantity_in_stock)

# Ques: 10
plot(med$manf_year, med$sales)
abline(lm(med$medID ~ med$sales))
```

▼ Screenshots Q 1

Console output (screenshot 1):

```
> usa          0.00223
> # Ques: 7
> df1=subset(df, continent != 'asia')
> df1
   country continent year lifeexp  pop gdpPerc
2      usa   america    2   81.00 2200     5.5
4  germany    europe    4   16.12 4100     4.4
5   russia    europe    5   16.22 5300    30.3
7      usa   america    7   99.00 4100     5.5
9  germany    europe    9   16.12 4019     4.4
10  russia    europe   10   16.22 5201     3.3
12     usa   america   12   19.20 2164     5.5
14 germany    europe   14   16.12 4000     4.4
15  russia    europe   15   16.22 5000     3.3
17     usa   america   17   19.20 9000     5.5
19 germany    europe   19   16.12 4001     4.4
20  russia    europe   20   16.22 5890     3.3
> arrange(df1, desc=df1$pop)
   country continent year lifeexp  pop gdpPerc
1      usa   america   12   19.20 2164     5.5
2      usa   america    2   81.00 2200     5.5
3  germany    europe   14   16.12 4000     4.4
4  germany    europe   19   16.12 4001     4.4
5  germany    europe    9   16.12 4019     4.4
6  germany    europe    4   16.12 4100     4.4
7      usa   america    7   99.00 4100     5.5
8   russia    europe   15   16.22 5000     3.3
9   russia    europe   10   16.22 5201     3.3
10  russia    europe    5   16.22 5300    30.3
11  russia    europe   20   16.22 5890     3.3
12     usa   america   17   19.20 9000     5.5
> tail(df1, 1)
```

Console output (screenshot 2):

```
15  russia    europe   15   16.22 5000     3.3
17     usa   america   17   19.20 9000     5.5
19 germany    europe   19   16.12 4001     4.4
20  russia    europe   20   16.22 5890     3.3
> arrange(df1, desc=df1$pop)
   country continent year lifeexp  pop gdpPerc
1      usa   america   12   19.20 2164     5.5
2      usa   america    2   81.00 2200     5.5
3  germany    europe   14   16.12 4000     4.4
4  germany    europe   19   16.12 4001     4.4
5  germany    europe    9   16.12 4019     4.4
6  germany    europe    4   16.12 4100     4.4
7      usa   america    7   99.00 4100     5.5
8   russia    europe   15   16.22 5000     3.3
9   russia    europe   10   16.22 5201     3.3
10  russia    europe    5   16.22 5300    30.3
11  russia    europe   20   16.22 5890     3.3
12     usa   america   17   19.20 9000     5.5
> tail(df1, 1)
   country continent year lifeexp  pop gdpPerc
20  russia    europe   20   16.22 5890     3.3
> # Ques: 8
> as<-tapply(df$pop, df$country, sd)
> as<-sort(as)
> as[1:3]
  germany     india    russia
 47.47631  80.88469 382.43986
> ## Ques: Q2
> # Ques: 1
> med<-read.csv("med.csv")
> |
```

▼ Screenshots Q 2

First RStudio window console:

```
> ## Ques: Q2
Warning message:
In doTryCatch(return(expr), name, parentenv, handler) :
  display list redraw incomplete
> # Ques: 1
> med<-read.csv("med.csv")
> med
   medID    med_name   company manf_year   exp_date quantity_in_stock sales
1      1       crocin  prajjwal      2018   1/1/2019                 8    22
2      2   paracetamol prajjwal      2017  11/8/2020                 7    23
3      3    soframycin    ishit      2020  9/14/2022                27     3
4      4      honeytus    ishit      2023  7/22/2024                 3    27
5      5 benzyl peroxide prajjwal    2022  5/28/2026                13    17
6      6         vicks   deepak      2025   4/4/2027                29     1
7      7      laughter  prajjwal      2021   2/8/2029                23     7
8      8          dolo   deepak      2017 12/16/2020                29     1
9      9      betadine    ishit      2012 10/22/2018                 7    23
10    10      chericof   deepak      2021  8/29/2022                30     0
> med$exp_date<-as.Date(med$exp_date, "%m/%d/%Y")
>
> # Ques: 2
> head(med, 4)
   medID    med_name  company manf_year   exp_date quantity_in_stock sales
1      1      crocin prajjwal      2018 2019-01-01                 8    22
2      2 paracetamol prajjwal      2017 2020-11-08                 7    23
3      3  soframycin    ishit      2020 2022-09-14                27     3
4      4    honeytus    ishit      2023 2024-07-22                 3    27
>
> # Ques: 3
> tail(med, 4)
   medID med_name  company manf_year   exp_date quantity_in_stock sales
7      7 laughter prajjwal      2021 2029-02-08                23     7
8      8     dolo   deepak      2017 2020-12-16                29     1
```


Second RStudio window console:

```
4      4    honeytus    ishit      2023  7/22/2024                 3    27
5      5 benzyl peroxide prajjwal    2022  5/28/2026                13    17
6      6         vicks   deepak      2025   4/4/2027                29     1
7      7      laughter  prajjwal      2021   2/8/2029                23     7
8      8          dolo   deepak      2017 12/16/2020                29     1
9      9      betadine    ishit      2012 10/22/2018                 7    23
10    10      chericof   deepak      2021  8/29/2022                30     0
> med$exp_date<-as.Date(med$exp_date, "%m/%d/%Y")
>
> # Ques: 2
> head(med, 4)
   medID    med_name  company manf_year   exp_date quantity_in_stock sales
1      1      crocin prajjwal      2018 2019-01-01                 8    22
2      2 paracetamol prajjwal      2017 2020-11-08                 7    23
3      3  soframycin    ishit      2020 2022-09-14                27     3
4      4    honeytus    ishit      2023 2024-07-22                 3    27
>
> # Ques: 3
> tail(med, 4)
   medID med_name  company manf_year   exp_date quantity_in_stock sales
7      7 laughter prajjwal      2021 2029-02-08                23     7
8      8     dolo   deepak      2017 2020-12-16                29     1
9      9 betadine    ishit      2012 2018-10-22                 7    23
10    10 chericof   deepak      2021 2022-08-29                30     0
>
> # Ques: 4
> days_left=(as.numeric)(med$exp_date-Sys.Date())
> cor(med$quantity_in_stock, days_left)
[1] 0.3409058
>
> # Ques: 5
> barplot(med$sales, med$manf_year)
>
```