Question 1

Register an application with the Github API here https://github.com/settings/applications. Access the API to get information on your instructors repositories (hint: this is the url you want "https://api.github.com/users/jtleek/repos"). Use this data to find the time that the datasharing repo was created. What time was it created?

This tutorial may be useful (https://github.com/hadley/httr/blob/master/demo/oauth2-github.r). You may also need to run the code in the base R package and not R studio.

Answer: Persistent issues on authentification. Could not get past "Bad credentials". Similar problems escalated without responses yet.

Question 2

Which of the following commands will select only the data for the probability weights pwgtp1 with ages less than 50?

sqldf("select pwgtp1 from acs")

sqldf("select * from acs where AGEP < 50")

sqldf("select pwgtp1 from acs where AGEP < 50")

sqldf("select * from acs")

Answer: commonsensically – only the 4th answer makes sense.

Question 3

Using the same data frame you created in the previous problem, what is the equivalent function to unique(acs$AGEP)

sqldf("select distinct AGEP from acs")

sqldf("select unique AGEP from acs")

sqldf("select AGEP where unique from acs")

sqldf("select distinct pwgtp1 from acs")

Answer: unique -> distinct when it comes to mysql, and select the option that mentions AGEP

Question 4

How many characters are in the 10th, 20th, 30th and 100th lines of HTML from this page:

http://biostat.jhsph.edu/~jleek/contact.html

(Hint: the nchar() function in R may be helpful)

```
url4 <- "http://biostat.jhsph.edu/~jleek/contact.html"
con <- url(url4)
htmlcode <-readline(con)
```

```
## 5
```

```
close(con)
nchar(htmlcode[c(10,20,30,100)])
```

```
## [1] 45 21 7 25
```

Question 5

Read this data set into R and report the sum of the numbers in the fourth of the nine columns.

https://d396qusza40orc.cloudfront.net/getdata%2Fwksst8110.for

Original source of the data: http://www.cpc.ncep.noaa.gov/data/indices/wksst8110.for

(Hint this is a fixed width file format)

28893.3

35824.9

32426.7

222243.1

101.83

36.5

```
x <- read.fwf(
  file=url("https://d396qusza40orc.cloudfront.net/getdata%2Fwksst8110.for"),
  skip=4,
  widths=c(12, 7, 4, 9, 4, 9, 4, 9, 4))
sum(x[,4])
```

```
## [1] 32426.7
```