lab3yg

Question1:

```
met <- read.csv("~/Downloads/met_all.gz")
```

Question 2:

```
dim(met)
```

```
[1] 2377343      30
```

```
head(met)
```

```
  USAFID  WBAN year month day hour min  lat      lon elev wind.dir wind.dir.qc
1 690150 93121 2019     8   1    0  56 34.3 -116.166  696      220           5
2 690150 93121 2019     8   1    1  56 34.3 -116.166  696      230           5
3 690150 93121 2019     8   1    2  56 34.3 -116.166  696      230           5
4 690150 93121 2019     8   1    3  56 34.3 -116.166  696      210           5
5 690150 93121 2019     8   1    4  56 34.3 -116.166  696      120           5
6 690150 93121 2019     8   1    5  56 34.3 -116.166  696       NA           9
  wind.type.code wind.sp wind.sp.qc ceiling.ht ceiling.ht.qc ceiling.ht.method
1              N     5.7          5      22000             5                 9
2              N     8.2          5      22000             5                 9
3              N     6.7          5      22000             5                 9
4              N     5.1          5      22000             5                 9
5              N     2.1          5      22000             5                 9
6              C     0.0          5      22000             5                 9
  sky.cond vis.dist vis.dist.qc vis.var vis.var.qc temp temp.qc dew.point
1        N    16093           5       N          5 37.2       5      10.6
2        N    16093           5       N          5 35.6       5      10.6
3        N    16093           5       N          5 34.4       5       7.2
4        N    16093           5       N          5 33.3       5       5.0
5        N    16093           5       N          5 32.8       5       5.0
6        N    16093           5       N          5 31.1       5       5.6
  dew.point.qc atm.press atm.press.qc       rh
1            5    1009.9           5 19.88127
2            5    1010.3           5 21.76098
3            5    1010.6           5 18.48212
4            5    1011.6           5 16.88862
5            5    1012.7           5 17.38410
6            5    1012.7           5 20.01540
```

```
tail(met)
```

```
        USAFID  WBAN year month day hour min    lat      lon elev wind.dir
2377338 726813 94195 2019     8  31   18  56 43.650 -116.633  741       NA
```

```
2377339 726813 94195 2019     8  31   19  56 43.650 -116.633  741       70
2377340 726813 94195 2019     8  31   20  56 43.650 -116.633  741       NA
2377341 726813 94195 2019     8  31   21  56 43.650 -116.633  741       10
2377342 726813 94195 2019     8  31   22  56 43.642 -116.636  741       10
2377343 726813 94195 2019     8  31   23  56 43.642 -116.636  741       40
        wind.dir.qc wind.type.code wind.sp wind.sp.qc ceiling.ht ceiling.ht.qc
2377338           9              C     0.0          5      22000             5
2377339           5              N     2.1          5      22000             5
2377340           9              C     0.0          5      22000             5
2377341           5              N     2.6          5      22000             5
2377342           1              N     2.1          1      22000             1
2377343           1              N     2.1          1      22000             1
        ceiling.ht.method sky.cond vis.dist vis.dist.qc vis.var vis.var.qc temp
2377338                 9        N    16093           5       N          5 30.0
2377339                 9        N    16093           5       N          5 32.2
2377340                 9        N    16093           5       N          5 33.3
2377341                 9        N    14484           5       N          5 35.0
2377342                 9        N    16093           1       9          9 34.4
2377343                 9        N    16093           1       9          9 34.4
        temp.qc dew.point dew.point.qc atm.press atm.press.qc       rh
2377338       5      11.7            5    1013.6            5 32.32509
2377339       5      12.2            5    1012.8            5 29.40686
2377340       5      12.2            5    1011.6            5 27.60422
2377341       5       9.4            5    1010.8            5 20.76325
2377342       1       9.4            1    1010.1            1 21.48631
2377343       1       9.4            1    1009.6            1 21.48631
```

there are 2377343 rows and 30 column.

Question 3:

```
str(met)
```

```
'data.frame':   2377343 obs. of  30 variables:
 $ USAFID          : int  690150 690150 690150 690150 690150 690150 690150 690150 690150
690150 ...
 $ WBAN            : int  93121 93121 93121 93121 93121 93121 93121 93121 93121 93121
...
 $ year            : int  2019 2019 2019 2019 2019 2019 2019 2019 2019 2019 ...
 $ month           : int  8 8 8 8 8 8 8 8 8 ...
 $ day             : int  1 1 1 1 1 1 1 1 1 ...
 $ hour            : int  0 1 2 3 4 5 6 7 8 9 ...
 $ min             : int  56 56 56 56 56 56 56 56 56 56 ...
 $ lat             : num  34.3 34.3 34.3 34.3 34.3 34.3 34.3 34.3 34.3 34.3 ...
 $ lon             : num  -116 -116 -116 -116 -116 ...
 $ elev            : int  696 696 696 696 696 696 696 696 696 696 ...
 $ wind.dir        : int  220 230 230 210 120 NA 320 10 320 350 ...
 $ wind.dir.qc     : chr  "5" "5" "5" "5" ...
 $ wind.type.code  : chr  "N" "N" "N" "N" ...
 $ wind.sp         : num  5.7 8.2 6.7 5.1 2.1 0 1.5 2.1 2.6 1.5 ...
 $ wind.sp.qc      : chr  "5" "5" "5" "5" ...
```

```
 $ ceiling.ht       : int  22000 22000 22000 22000 22000 22000 22000 22000 22000 22000
...
 $ ceiling.ht.qc    : int  5 5 5 5 5 5 5 5 5 ...
 $ ceiling.ht.method: chr  "9" "9" "9" "9" ...
 $ sky.cond         : chr  "N" "N" "N" "N" ...
 $ vis.dist         : int  16093 16093 16093 16093 16093 16093 16093 16093 16093 16093
...
 $ vis.dist.qc      : chr  "5" "5" "5" "5" ...
 $ vis.var          : chr  "N" "N" "N" "N" ...
 $ vis.var.qc       : chr  "5" "5" "5" "5" ...
 $ temp             : num  37.2 35.6 34.4 33.3 32.8 31.1 29.4 28.9 27.2 26.7 ...
 $ temp.qc          : chr  "5" "5" "5" "5" ...
 $ dew.point        : num  10.6 10.6 7.2 5 5 5.6 6.1 6.7 7.8 7.8 ...
 $ dew.point.qc     : chr  "5" "5" "5" "5" ...
 $ atm.press        : num  1010 1010 1011 1012 1013 ...
 $ atm.press.qc     : int  5 5 5 5 5 5 5 5 5 5 ...
 $ rh               : num  19.9 21.8 18.5 16.9 17.4 ...
```

Question4:

```
table(met$year)
```

```
   2019
2377343
```

```
table(met$day)
```

```
    1     2     3     4     5     6     7     8     9    10    11    12    13
75975 75923 76915 76594 76332 76734 77677 77766 75366 75450 76187 75052 76906
   14    15    16    17    18    19    20    21    22    23    24    25    26
77852 76217 78015 78219 79191 76709 75527 75786 78312 77413 76965 76806 79114
   27    28    29    30    31
79789 77059 71712 74931 74849
```

```
table(met$hour)
```

```
    0     1     2     3     4     5     6     7     8     9    10
 99434  93482  93770  96703 110504 112128 106235 101985 100310 102915 101880
   11    12    13    14    15    16    17    18    19    20    21
100470 103605  97004  96507  97635  94942  94184 100179  94604  94928  96070
   22    23
 94046  93823
```

```
summary(met$temp)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
 -40.00   19.60   23.50   23.59   27.80   56.00   60089
```

```
summary(met$elev)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  -13.0   101.0   252.0   415.8   400.0  9999.0
```

```
summary(met$wind.sp)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
  0.000   0.000   2.100   2.459   3.600  36.000   79693
```

```
met$elev[met$elev==9999.0] <- NA

summary(met$elev)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
    -13     101     252     413     400    4113     710
```

The highest weather station is 4113 meters.

```
met <- met[met$temp>-40,]
met2 <- met[order(met$temp),]
head(met2)
```

```
        USAFID WBAN year month day hour min    lat    lon elev wind.dir
1203053 722817 3068 2019     8   1    0  56 38.767 -104.3 1838      190
1203055 722817 3068 2019     8   1    1  56 38.767 -104.3 1838      180
1203128 722817 3068 2019     8   3   11  56 38.767 -104.3 1838       NA
1203129 722817 3068 2019     8   3   12  56 38.767 -104.3 1838       NA
1203222 722817 3068 2019     8   6   21  56 38.767 -104.3 1838      280
1203225 722817 3068 2019     8   6   22  56 38.767 -104.3 1838      240
        wind.dir.qc wind.type.code wind.sp wind.sp.qc ceiling.ht ceiling.ht.qc
1203053           5              N     7.2          5         NA             9
1203055           5              N     7.7          5         NA             9
1203128           9              C     0.0          5         NA             9
1203129           9              C     0.0          5         NA             9
1203222           5              N     2.6          5         NA             9
1203225           5              N     7.7          5         NA             9
        ceiling.ht.method sky.cond vis.dist vis.dist.qc vis.var vis.var.qc
1203053                 9        N       NA          9       N          5
1203055                 9        N       NA          9       N          5
1203128                 9        N       NA          9       N          5
1203129                 9        N       NA          9       N          5
1203222                 9        N       NA          9       N          5
1203225                 9        N       NA          9       N          5
         temp temp.qc dew.point dew.point.qc atm.press atm.press.qc rh
```

```
1203053 -17.2       5        NA           9        NA           9 NA
1203055 -17.2       5        NA           9        NA           9 NA
1203128 -17.2       5        NA           9        NA           9 NA
1203129 -17.2       5        NA           9        NA           9 NA
1203222 -17.2       5        NA           9        NA           9 NA
1203225 -17.2       5        NA           9        NA           9 NA
```

Question 5

```
met$elev[which.min(met$temp)]
```

```
[1] 1838
```

```
range(met$elev,na.rm = TRUE)
```

```
[1]  -13 4113
```

```
met <- met[met$temp>-15,]
met2 <- met[order(met$temp),]
head(met2)
```

```
        USAFID  WBAN year month day hour min    lat      lon elev wind.dir
2370758 726764 94163 2019     8  27   11  50 44.683 -111.116 2025       NA
2370759 726764 94163 2019     8  27   12  10 44.683 -111.116 2025       NA
2370760 726764 94163 2019     8  27   12  30 44.683 -111.116 2025       NA
2370761 726764 94163 2019     8  27   12  50 44.683 -111.116 2025       NA
252489  720411   137 2019     8  18   12  35 36.422 -105.290 2554       NA
2370688 726764 94163 2019     8  26   12  30 44.683 -111.116 2025       NA
        wind.dir.qc wind.type.code wind.sp wind.sp.qc ceiling.ht ceiling.ht.qc
2370758           9              C       0          5      22000             5
2370759           9              C       0          5      22000             5
2370760           9              C       0          5      22000             5
2370761           9              C       0          5      22000             5
252489            9              C       0          5      22000             5
2370688           9              C       0          5      22000             5
        ceiling.ht.method sky.cond vis.dist vis.dist.qc vis.var vis.var.qc temp
2370758                 9        N    16093           5       N          5 -3.0
2370759                 9        N    16093           5       N          5 -3.0
2370760                 9        N    16093           5       N          5 -3.0
2370761                 9        N    16093           5       N          5 -3.0
252489                  9        N    16093           5       N          5 -2.4
2370688                 9        N    16093           5       N          5 -2.0
        temp.qc dew.point dew.point.qc atm.press atm.press.qc       rh
2370758       C      -5.0            C        NA           9 86.26537
2370759       5      -4.0            5        NA           9 92.91083
2370760       5      -4.0            5        NA           9 92.91083
2370761       C      -4.0            C        NA           9 92.91083
252489        5      -3.7            5        NA           9 90.91475
2370688       5      -3.0            5        NA           9 92.96690
```

Question 6:

```
elev <- met[met$elev==max(met$elev, na.rm=TRUE), ]
summary(elev)
```

```
    USAFID            WBAN            year          month
Min.   :720385   Min.   :419     Min.   :2019    Min.   :8
1st Qu.:720385   1st Qu.:419     1st Qu.:2019    1st Qu.:8
Median :720385   Median :419     Median :2019    Median :8
Mean   :720385   Mean   :419     Mean   :2019    Mean   :8
3rd Qu.:720385   3rd Qu.:419     3rd Qu.:2019    3rd Qu.:8
Max.   :720385   Max.   :419     Max.   :2019    Max.   :8
NA's   :60271    NA's   :60271   NA's   :60271   NA's   :60271
     day             hour            min             lat
Min.   : 1.0    Min.   : 0.00   Min.   : 6.00   Min.   :39.8
1st Qu.: 8.0    1st Qu.: 6.00   1st Qu.:13.00   1st Qu.:39.8
Median :16.0    Median :12.00   Median :36.00   Median :39.8
Mean   :16.1    Mean   :11.66   Mean   :34.38   Mean   :39.8
3rd Qu.:24.0    3rd Qu.:18.00   3rd Qu.:53.00   3rd Qu.:39.8
Max.   :31.0    Max.   :23.00   Max.   :59.00   Max.   :39.8
NA's   :60271   NA's   :60271   NA's   :60271   NA's   :60271
     lon              elev          wind.dir     wind.dir.qc
Min.   :-105.8   Min.   :4113    Min.   : 10.0   Length:62388
1st Qu.:-105.8   1st Qu.:4113    1st Qu.:250.0   Class :character
Median :-105.8   Median :4113    Median :300.0   Mode  :character
Mean   :-105.8   Mean   :4113    Mean   :261.5
3rd Qu.:-105.8   3rd Qu.:4113    3rd Qu.:310.0
Max.   :-105.8   Max.   :4113    Max.   :360.0
NA's   :60271    NA's   :60271   NA's   :60508
wind.type.code       wind.sp        wind.sp.qc          ceiling.ht
Length:62388       Min.   : 0.000   Length:62388       Min.   :   30
Class :character   1st Qu.: 4.100   Class :character   1st Qu.: 2591
Mode  :character   Median : 6.700   Mode  :character   Median :22000
                   Mean   : 7.245                      Mean   :15145
                   3rd Qu.: 9.800                      3rd Qu.:22000
                   Max.   :21.100                      Max.   :22000
                   NA's   :60439                       NA's   :60275
ceiling.ht.qc   ceiling.ht.method    sky.cond            vis.dist
Min.   :5.000   Length:62388       Length:62388       Min.   :    0
1st Qu.:5.000   Class :character   Class :character   1st Qu.:16093
Median :5.000   Mode  :character   Mode  :character   Median :16093
Mean   :5.008                                         Mean   :15913
3rd Qu.:5.000                                         3rd Qu.:16093
Max.   :9.000                                         Max.   :16093
NA's   :60271                                         NA's   :60954
vis.dist.qc          vis.var           vis.var.qc             temp
Length:62388       Length:62388       Length:62388       Min.   : 1.00
Class :character   Class :character   Class :character   1st Qu.: 6.00
Mode  :character   Mode  :character   Mode  :character   Median : 8.00
                                                         Mean   : 8.13
```

```
                                                    3rd Qu.:10.00
                                                    Max.   :15.00
                                                    NA's   :60271
  temp.qc            dew.point      dew.point.qc        atm.press
Length:62388       Min.   :-6.000   Length:62388       Min.   : NA
Class :character   1st Qu.: 0.000   Class :character   1st Qu.: NA
Mode  :character   Median : 0.000   Mode  :character   Median : NA
                   Mean   : 0.873                      Mean   :NaN
                   3rd Qu.: 2.000                      3rd Qu.: NA
                   Max.   : 7.000                      Max.   : NA
                   NA's   :60271                       NA's   :62388
 atm.press.qc         rh
Min.   :9       Min.   :53.62
1st Qu.:9       1st Qu.:58.10
Median :9       Median :61.39
Mean   :9       Mean   :60.62
3rd Qu.:9       3rd Qu.:61.85
Max.   :9       Max.   :70.01
NA's   :60271   NA's   :60271
```

```r
cor(elev$temp, elev$wind.sp, use="complete")
```

```
[1] -0.09373843
```

```r
cor(elev$temp, elev$hour, use="complete")
```

```
[1] 0.4397261
```

```r
cor(elev$wind.sp, elev$day, use="complete")
```

```
[1] 0.3643079
```

```r
cor(elev$wind.sp, elev$hour, use="complete")
```

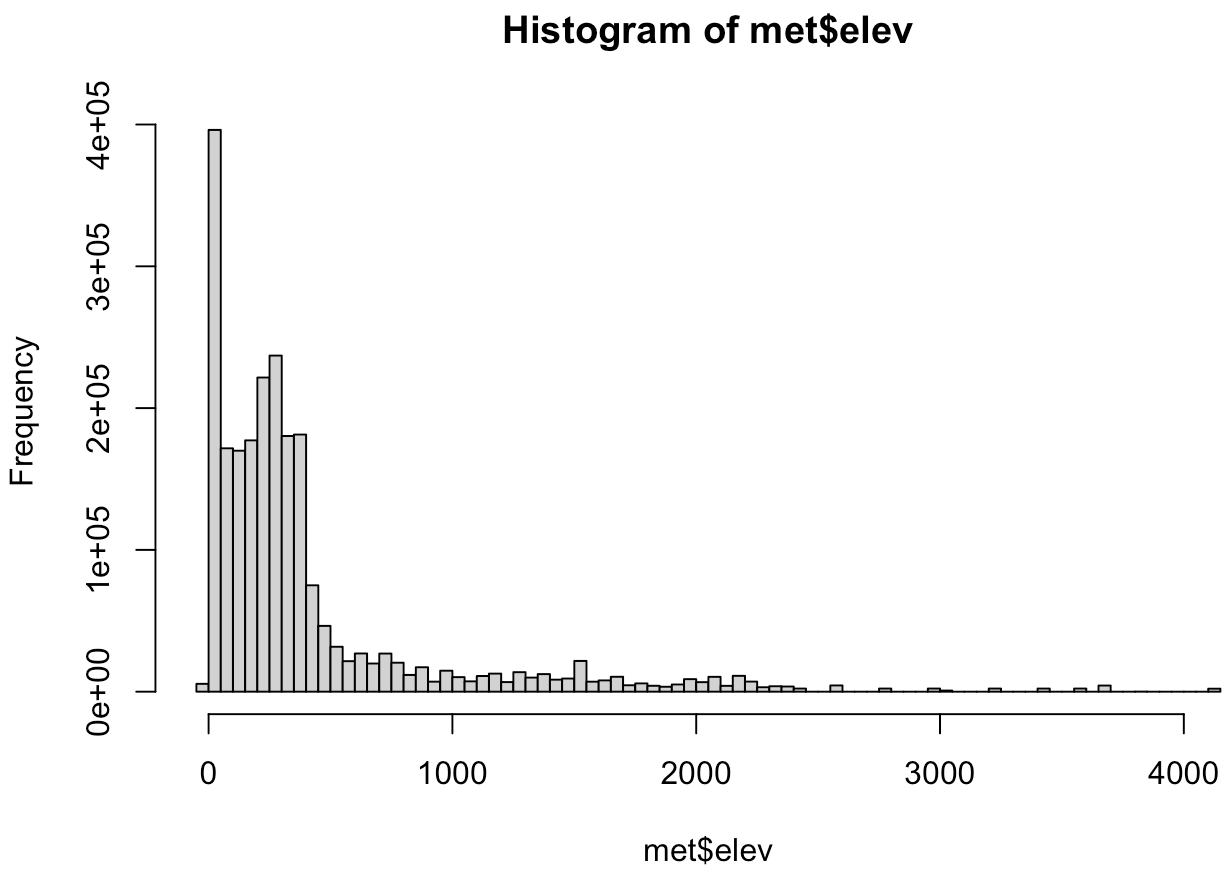
```
[1] 0.08807315
```

```r
cor(elev$temp, elev$day, use="complete")
```

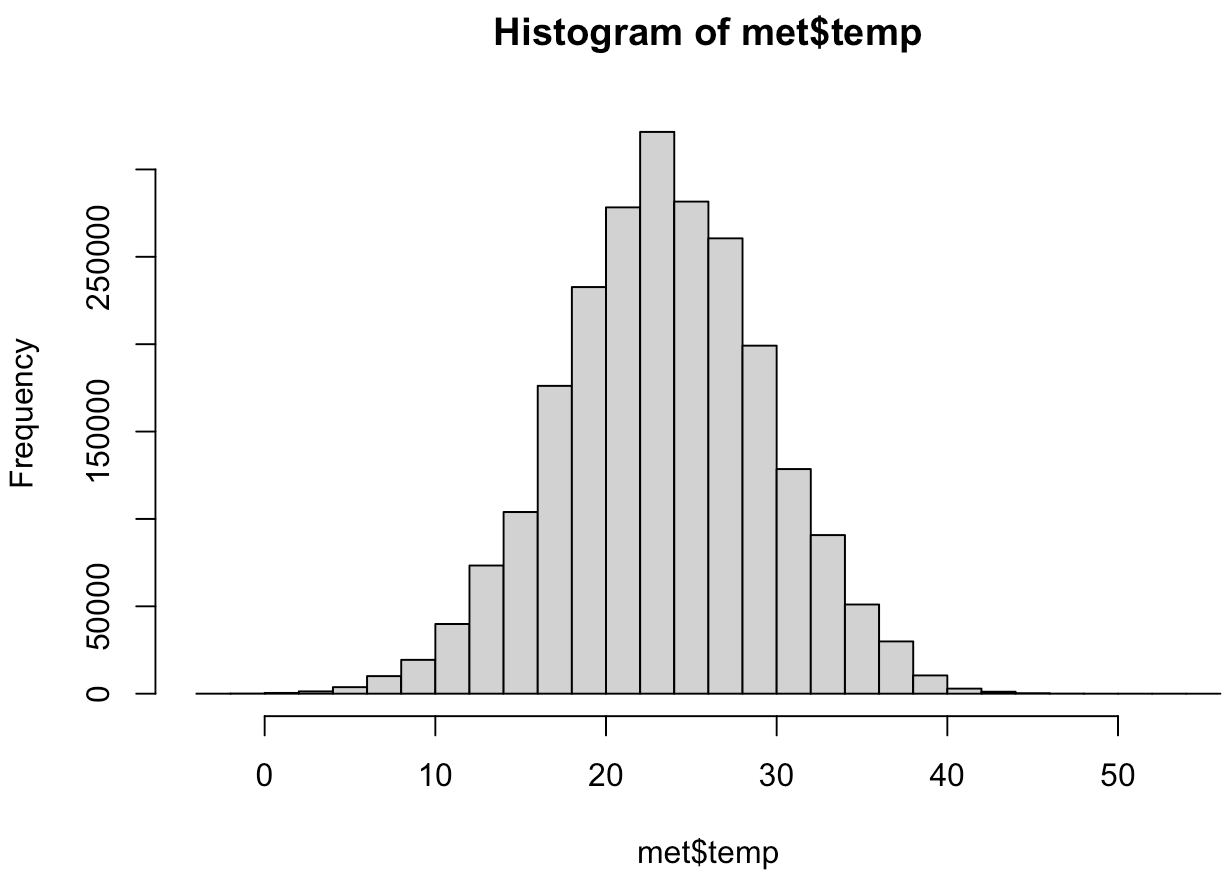
```
[1] -0.003857766
```

Question 7;

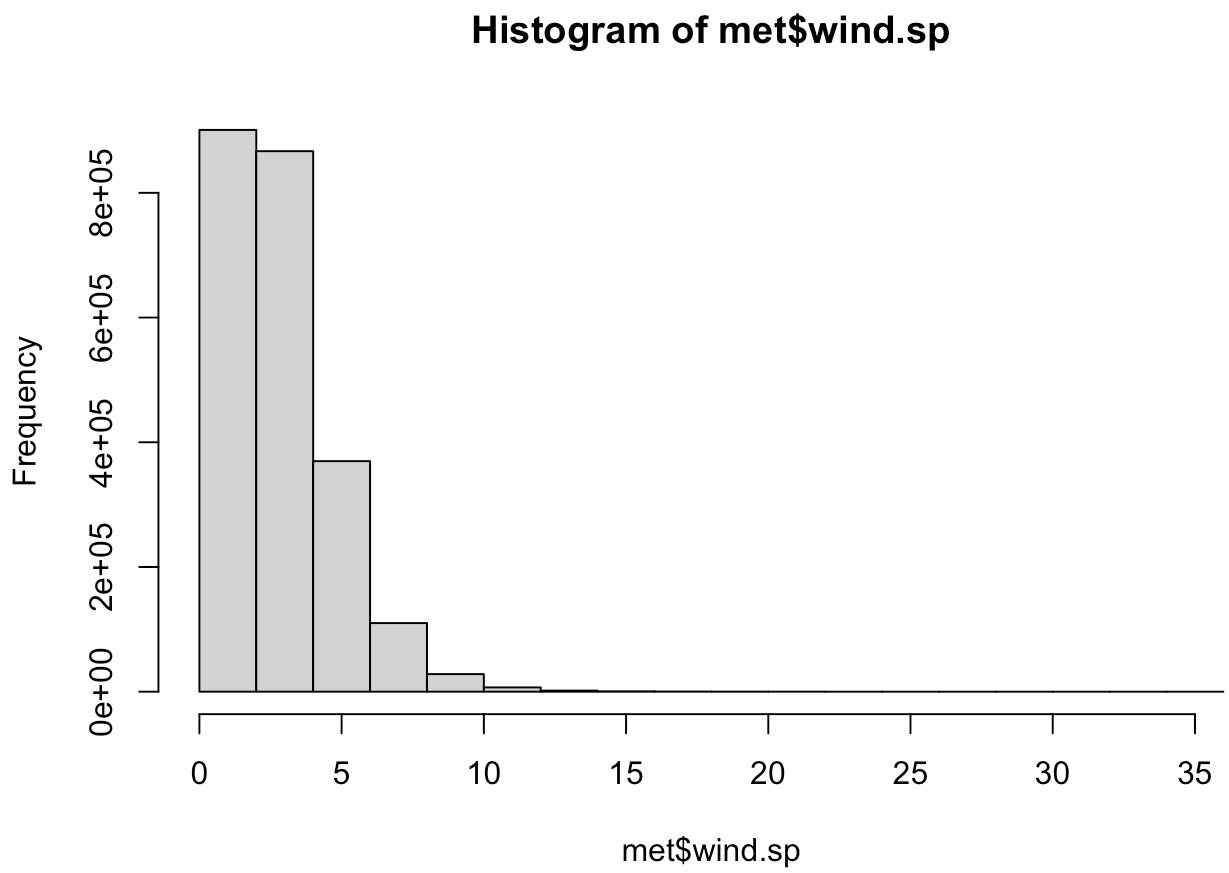
```r
hist(met$elev, breaks=100)
```




```
hist(met$temp)
```



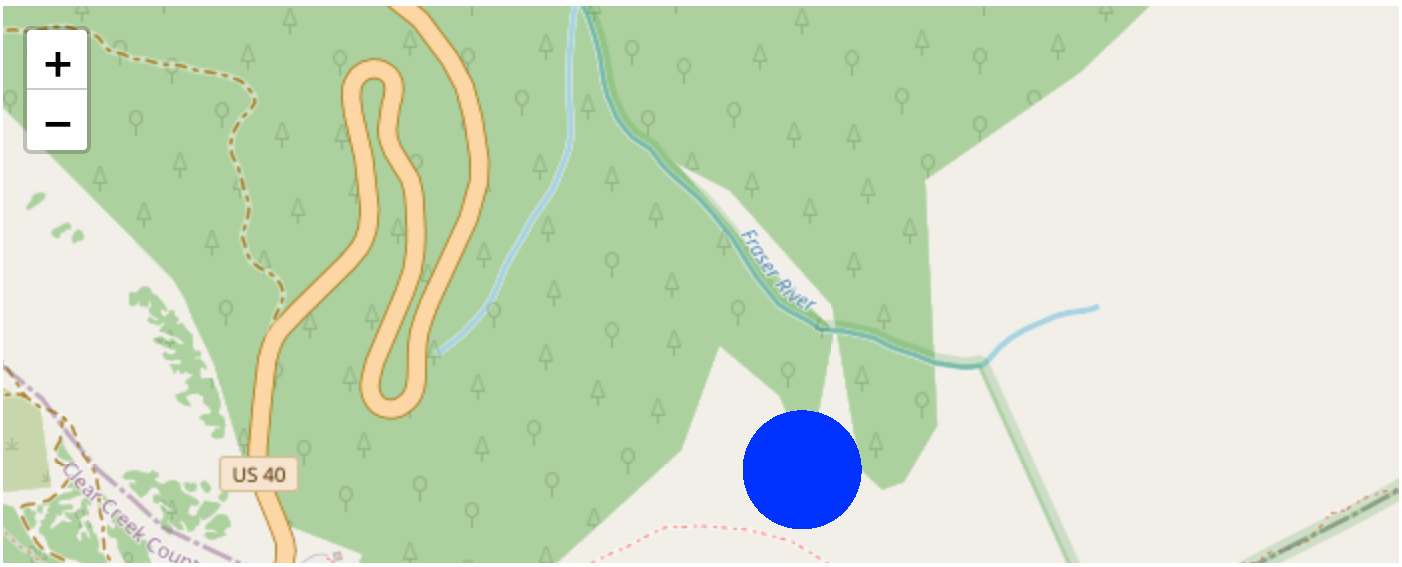

```
hist(met$wind.sp)
```



library(leaflet)

```
library(leaflet)
leaflet(elev) %>%
  addProviderTiles('OpenStreetMap') %>%
  addCircles(lat=~lat,lng=~lon, opacity=1, fillOpacity=1, radius=100)
```

```
Warning in validateCoords(lng, lat, funcName): Data contains 60271 rows with
either missing or invalid lat/lon values and will be ignored
```




```
library(lubridate)
```

```
Attaching package: 'lubridate'

The following objects are masked from 'package:base':

    date, intersect, setdiff, union
```

```
elev$date <- with(elev, ymd_h(paste(year, month, day, hour, sep= ' ')))
```

```
Warning: 60271 failed to parse.
```

```
summary(elev$date)
```

```
                Min.               1st Qu.                Median
"2019-08-01 00:00:00" "2019-08-08 11:00:00" "2019-08-16 22:00:00"
                Mean               3rd Qu.                   Max.
"2019-08-16 14:09:56" "2019-08-24 11:00:00" "2019-08-31 22:00:00"
                NA's
              "60271"
```

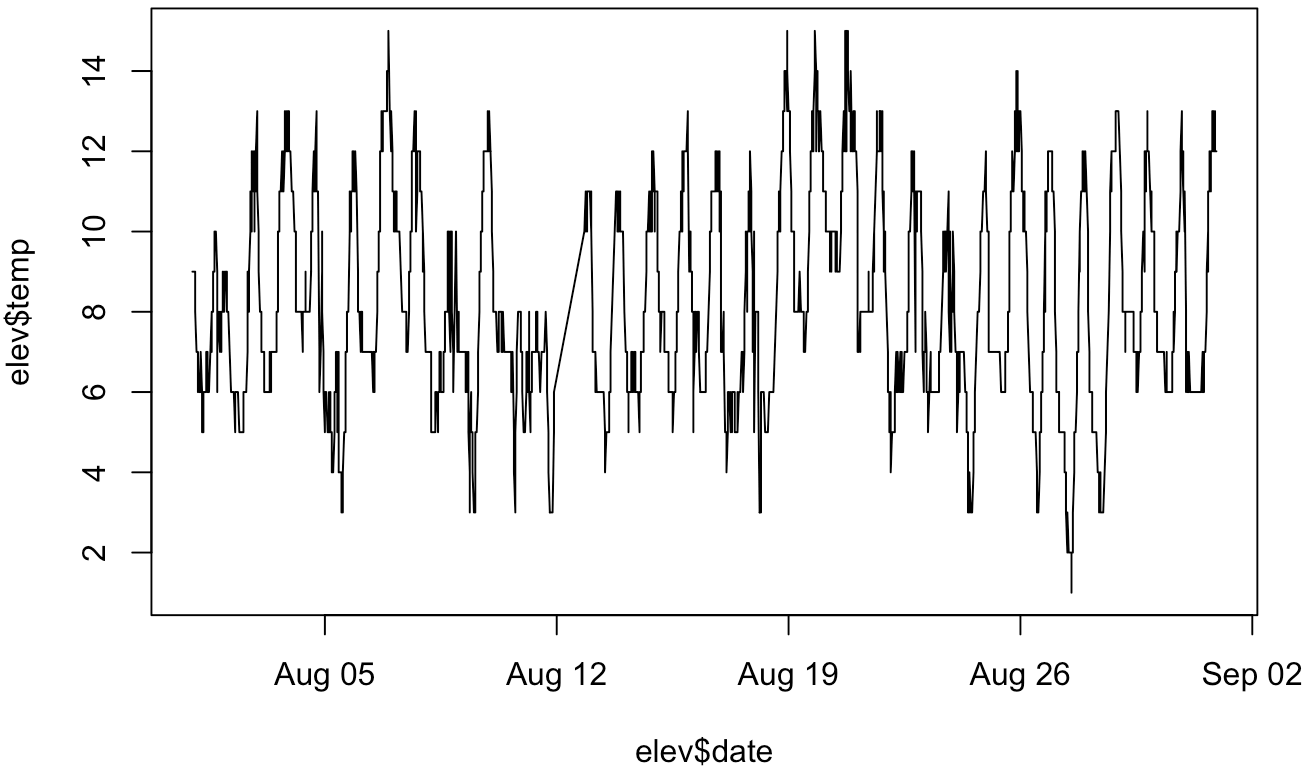
```
elev <- elev[order(elev$date),]
head(elev)
```

```
       USAFID WBAN year month day hour min  lat      lon elev wind.dir
221697 720385  419 2019     8   1    0  36 39.8 -105.766 4113      170
221698 720385  419 2019     8   1    0  54 39.8 -105.766 4113      100
221699 720385  419 2019     8   1    1  12 39.8 -105.766 4113       90
221700 720385  419 2019     8   1    1  35 39.8 -105.766 4113      110
221701 720385  419 2019     8   1    1  53 39.8 -105.766 4113      120
221702 720385  419 2019     8   1    2  12 39.8 -105.766 4113      120
       wind.dir.qc wind.type.code wind.sp wind.sp.qc ceiling.ht ceiling.ht.qc
221697           5              N     8.8          5       1372             5
221698           5              N     2.6          5       1372             5
221699           5              N     3.1          5       1981             5
221700           5              N     4.1          5       2134             5
221701           5              N     4.6          5       2134             5
```

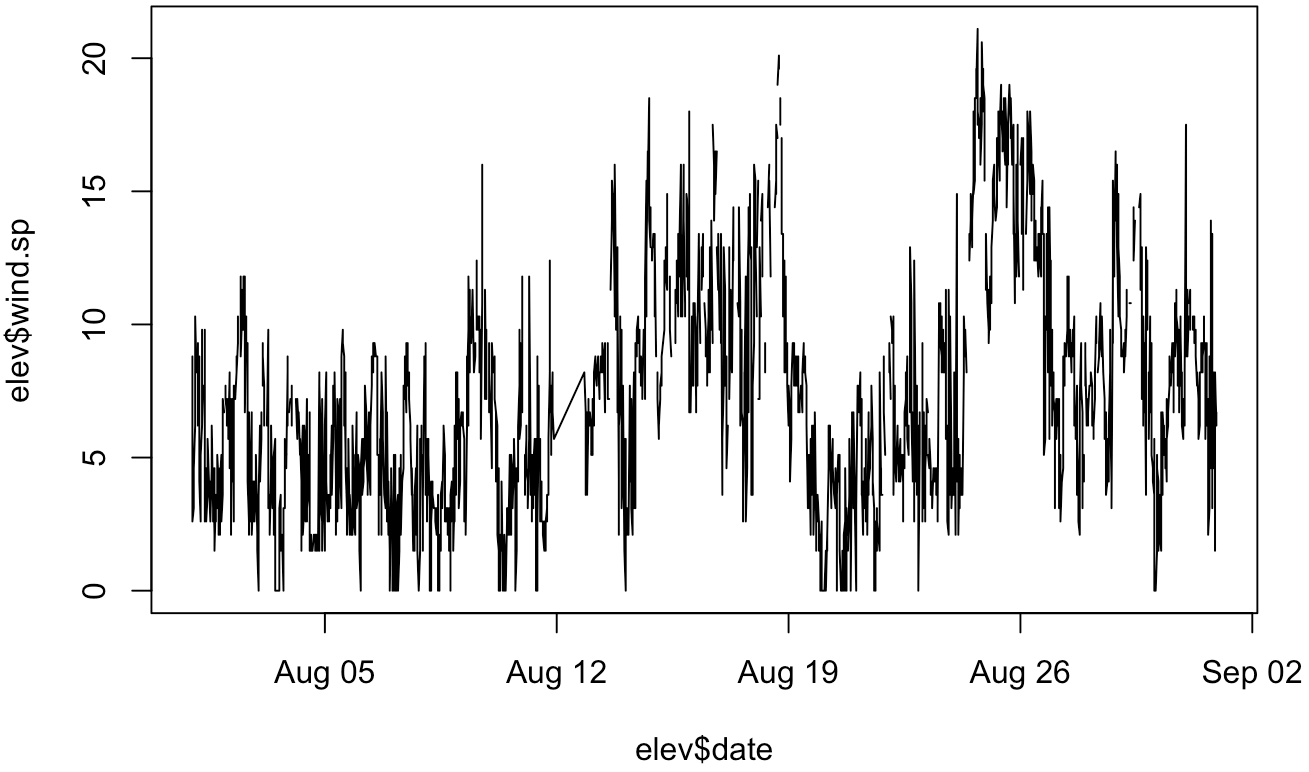
```
221702           5              N     6.2          5      22000             5
       ceiling.ht.method sky.cond vis.dist vis.dist.qc vis.var vis.var.qc temp
221697                 M        N       NA           9       N          5    9
221698                 M        N       NA           9       N          5    9
221699                 M        N       NA           9       N          5    9
221700                 M        N       NA           9       N          5    9
221701                 M        N       NA           9       N          5    9
221702                 9        N       NA           9       N          5    9
       temp.qc dew.point dew.point.qc atm.press atm.press.qc       rh
221697       5         1            5        NA            9 57.61039
221698       5         1            5        NA            9 57.61039
221699       5         2            5        NA            9 61.85243
221700       5         2            5        NA            9 61.85243
221701       5         2            5        NA            9 61.85243
221702       5         2            5        NA            9 61.85243
                    date
221697 2019-08-01 00:00:00
221698 2019-08-01 00:00:00
221699 2019-08-01 01:00:00
221700 2019-08-01 01:00:00
221701 2019-08-01 01:00:00
221702 2019-08-01 02:00:00
```

Question 8:

```
plot(elev$date, elev$temp, type='l')
```




```
plot(elev$date, elev$wind.sp, type='l')
```



Tem: -axis (horizontal): Dates ranging from early August to early September (Aug 01 – Sep 02). Y-axis (vertical): Temperature values (elev$temp), ranging from about 2 to 15. Temperatures oscillate daily (short-term cycles), likely reflecting day–night variation.Values generally fall between 5–12°C, with occasional lows around 2–4°C and highs peaking at 13–15°C. Early August (Aug 01–10): Strong fluctuations, but temperatures mostly in mid-range (6–12°C). Mid-August (Aug 11–20): Slight dips observed, some lows near 2–4°C. Late August (Aug 21–31): More pronounced variability, with frequent swings between 5–13°C.

Wind: X-axis (horizontal): Dates ranging from early August to early September (Aug 01 – Sep 02).Y-axis (vertical): Wind speed values (elev$wind.sp), ranging approximately from 0 to a little over 20. Wind speed fluctuates frequently with short-term variability. Early August (Aug 01–10): Wind speed is mostly lower, with values generally between 0–10. Mid-August (Aug 11–20): Increased variability, with some peaks approaching 15–20. Late August (Aug 21–30): More frequent high wind speeds, often above 10 and peaking above 20.

Question 8: