Problem 1

1)

```
f=file.choose()
ojjuice=read_table(f)
cols(
  run = col_double(),
  sweetness = col_double(),
  pectin = col_double()
)
ojjuice
# A tibble: 24 × 3
    run sweetness pectin
  <dbl>     <dbl>  <dbl>
 1     1       5.2    220
 2     2       5.5    227
 3     3       6      259
 4     4       5.9    210
 5     5       5.8    224
 6     6       6      215
 7     7       5.8    231
 8     8       5.6    268
 9     9       5.6    239
10    10       5.9    212
# … with 14 more rows
# ℹ Use `print(n = ...)` to see more rows
```

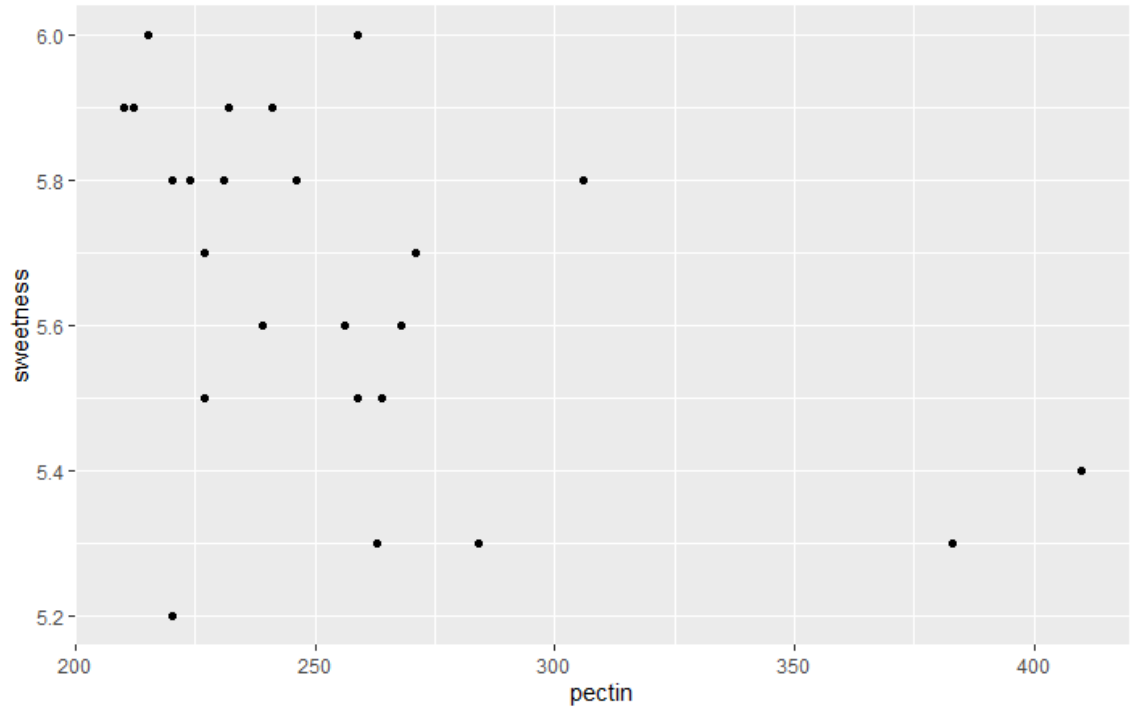
2)

```
> print(as_tibble((ojjuice)),n=24)
# A tibble: 24 × 3
    run sweetness pectin
  <dbl>     <dbl>  <dbl>
 1     1       5.2    220
 2     2       5.5    227
 3     3       6      259
 4     4       5.9    210
 5     5       5.8    224
 6     6       6      215
 7     7       5.8    231
 8     8       5.6    268
 9     9       5.6    239
```

10	10	5.9	212
11	11	5.4	410
12	12	5.6	256
13	13	5.8	306
14	14	5.5	259
15	15	5.3	284
16	16	5.3	383
17	17	5.7	271
18	18	5.5	264
19	19	5.7	227
20	20	5.3	263
21	21	5.9	232
22	22	5.8	220
23	23	5.8	246
24	24	5.9	241

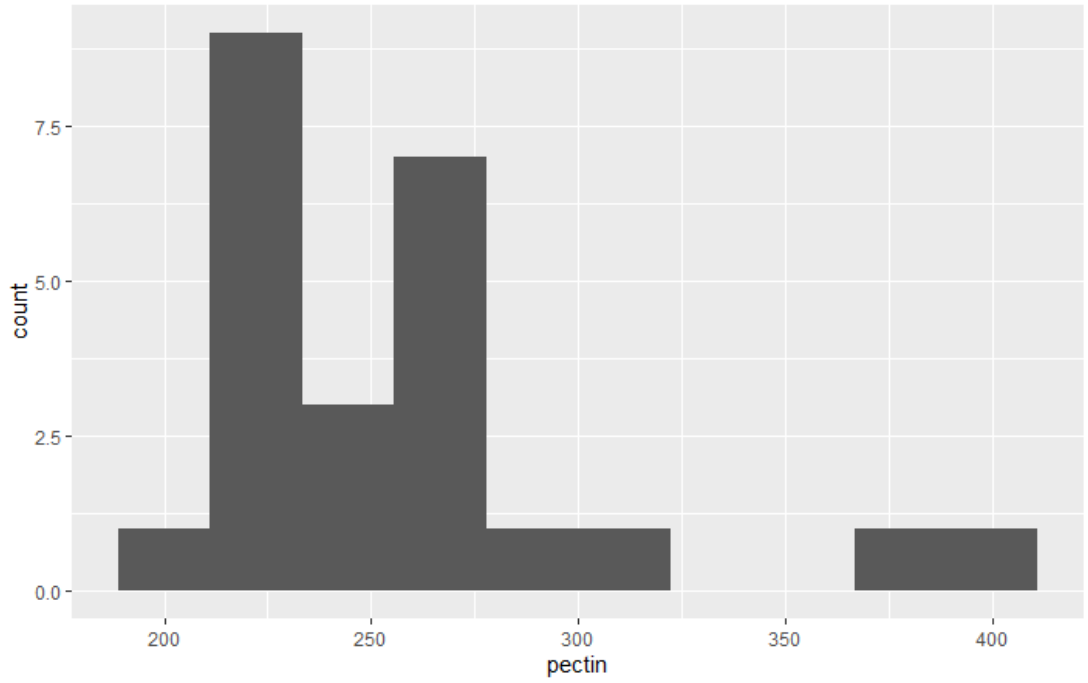
3)

```
ggplot(ojjuice, aes(x=pectin, y=sweetness)) + geom_point()
```



4)

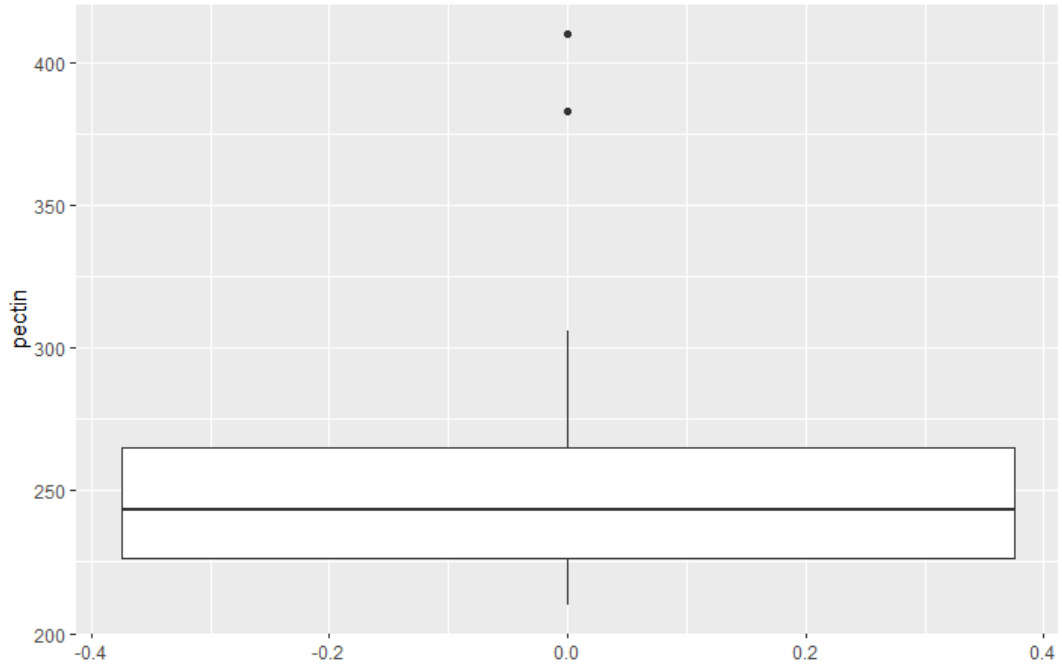
```
ggplot(ojjuice, aes(x = pectin)) + geom_histogram(bins = 10)
```



Based off the graph, this data is skewed to the right

5)

ggplot(ojjuice,aes(y=pectin))+geom_boxplot()



Problem 2

a)

```
g=file.choose()
tests=read_table(g)
cols(
  gender = col_character(),
  CAS = col_double(),
  CARS = col_double()
)
tests
# A tibble: 35 × 3
   gender   CAS  CARS
   <chr>  <dbl> <dbl>
 1 female  2.85  2.9
 2 male    2.6   2.32
 3 female  2.2   1
 4 male    2.65  2.58
 5 male    2.6   2.58
 6 male    3.2   3.05
 7 male    3.65  3.74
 8 female  2.55  1.9
 9 male    3.15  3.32
10 male    2.8   2.74
# … with 25 more rows
# ℹ Use `print(n = ...)` to see more rows
```

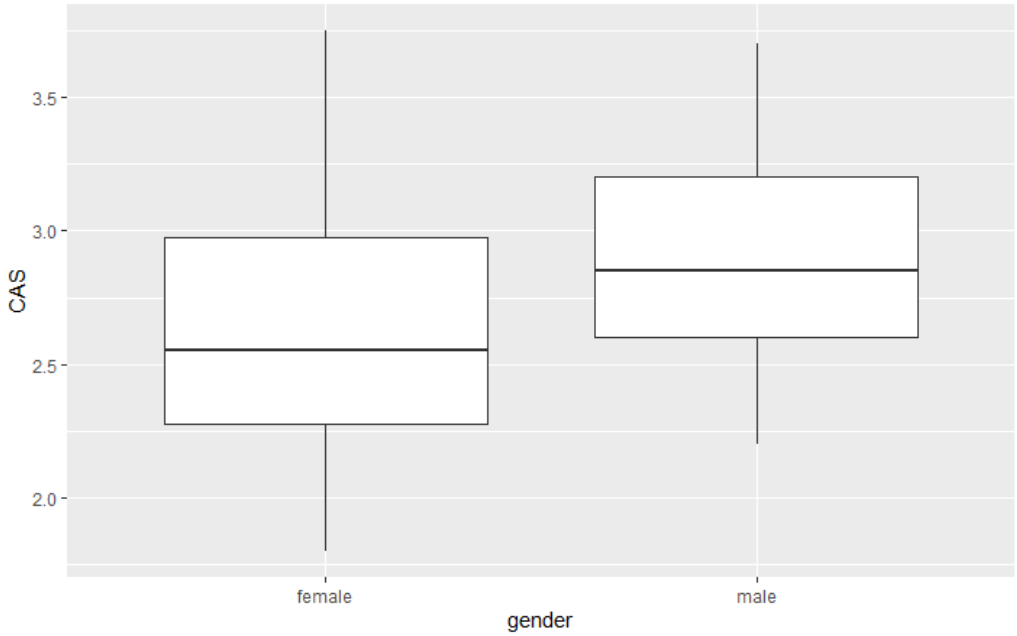
Yes, it contains the gender of the test taker and their score of each test from 35 participants.

b)

```
> tests %>% count(gender)
# A tibble: 2 × 2
  gender     n
  <chr>  <int>
1 female    15
2 male      20
```

c)

```
ggplot(tests,aes(x=gender,y=CAS))+geom_boxplot()
```



The CAS scores tend to be higher for males indicating they experience more anxiety around computers compared to the females on average.

d)

```
> tests %>% group_by(gender) %>% summarize(m=median(CAS))
# A tibble: 2 × 2
  gender     m
  <chr>  <dbl>
1 female  2.55
2 male    2.85
```

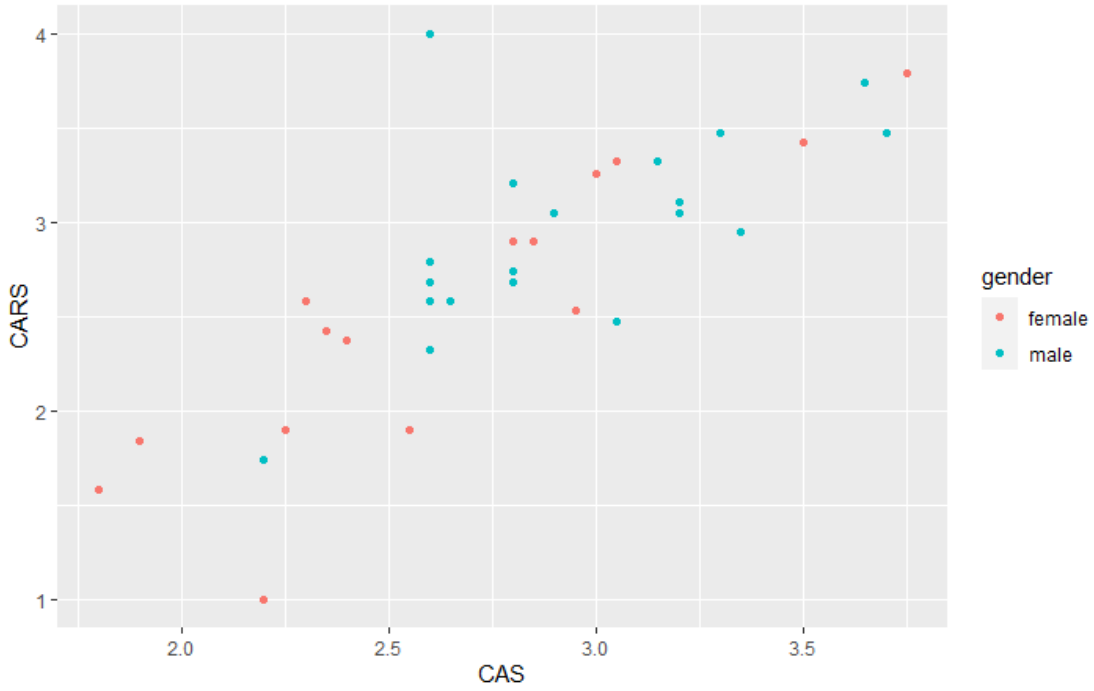
The output for the median of the test results support the information given by the boxplot.

e)

```
> tests %>% summarise_if(is.numeric,c(mean,sd))
# A tibble: 1 × 4
  CAS_fn1 CARS_fn1 CAS_fn2 CARS_fn2
    <dbl>    <dbl>   <dbl>    <dbl>
1    2.82     2.77   0.484    0.671
```

f)

```
ggplot(tests,aes(x=CAS,y=CARS,colour=gender))+geom_point()
```


4
3
2
1
CARS
2.0
2.5
3.0
3.5
CAS
gender
female
male