Homework 01
Joseph Blubaugh
jblubau1@tamu.edu
STAT 636-720

1)

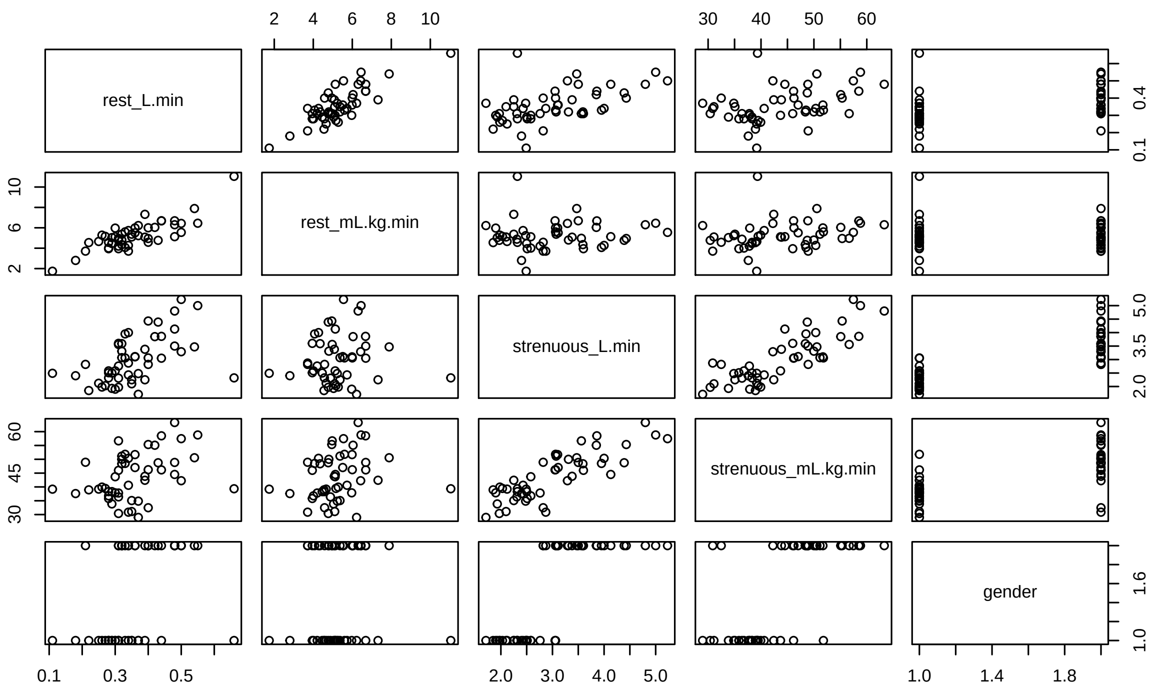
```r
oxygen =
  read.table("C:/Users/Joseph/Projects/learning/Statistics/STAT_636/Data/textbook/T6-12.
             quote="\"",
             comment.char="")
colnames(oxygen) = c("rest_L.min", "rest_mL.kg.min", "strenuous_L.min",
                     "strenuous_mL.kg.min", "gender")

## a)
## Sample Averages
apply(X = oxygen[, 1:4], MARGIN = 2, FUN = mean)
```

```
         rest_L.min      rest_mL.kg.min     strenuous_L.min
             0.3554              5.2542              3.0014
strenuous_mL.kg.min
            43.7876
```

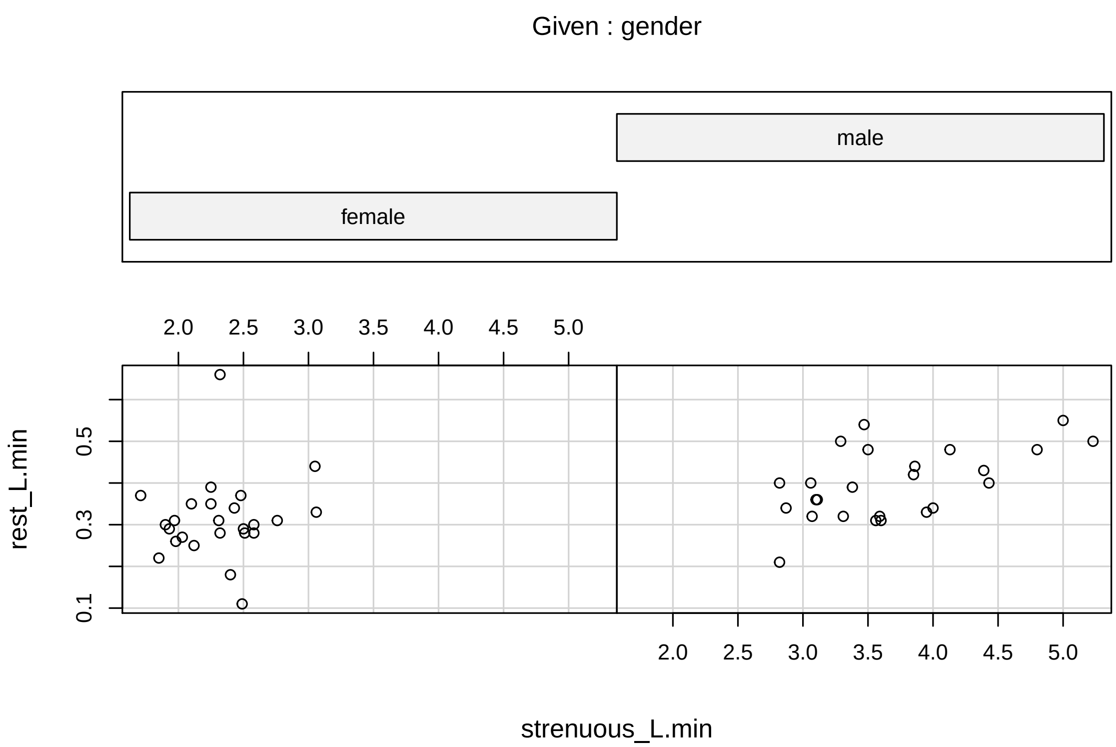
```r
## Sample Standard Deviations
apply(X = oxygen[, 1:4], MARGIN = 2, FUN = sd)
```

```
         rest_L.min      rest_mL.kg.min     strenuous_L.min
          0.1001878           1.3885863           0.8733796
strenuous_mL.kg.min
          8.4161326
```

```r
## b)
pairs(oxygen)
```




```r
## c)
coplot(rest_L.min ~ strenuous_L.min | gender, data = oxygen)
```



a) The resting oxygen measurement means and standard deviation are lower than that measurements for the strenuous activity. Strenuous mL/kg/min is much higher and has more variation than the Resting mL/kg/min measurement.

b) Resting L/min appears to be positively correlated with both Rest mL/kg/min and Strenuous L/min. Strenuous L/min also appears to be positively correlated with Strenuous mL/kg/min The observations of the female with a resting mL/kg/min of 11.05 appears to be an outlier and shows up as such on several plots.

c) Visually there appears to be no obvious relationship between Rest L/min and Strenuous L/min for females, but the relationship appears to be positively correlated for men.

2)

a)

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$$

$$d(x, \mu) = \sqrt{(x-\mu)'\Sigma^{-1}(x-\mu)}$$

$$d(x, \mu) = \sqrt{\begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}^{-1} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}}$$

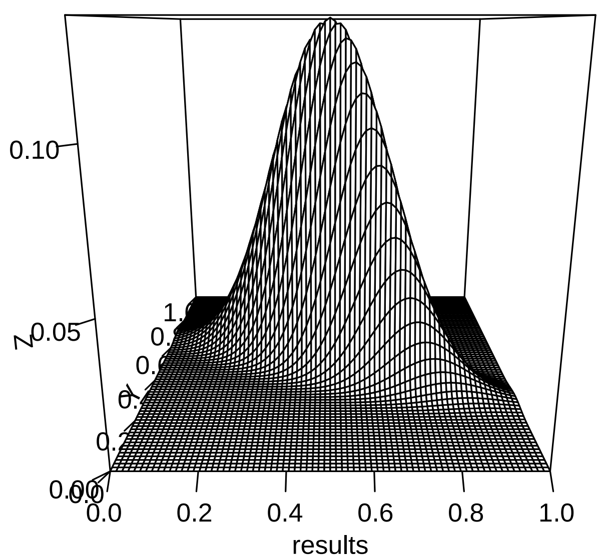
$$= \sqrt{\sigma_{11}(x_1-\mu_1)^2 + 2\sigma_{12}(x_1-\mu_1)(x_2-\mu_2) + \sigma_{22}(x_2-\mu_2)^2}$$

$$a_{11} = \sigma_{11}, a_{12} = \sigma_{12}, a_{22} = \sigma 22$$

b i.

```
mu = c(1, -1)
Sigma = matrix(c(1, -1.6, -1.6, 4), nrow = 2)

mvn = function(x) {
  constant = 1/((2 * pi)^(2/2) * sqrt(det(Sigma)))
  scalar = exp(-t(x - mu) %*% solve(Sigma) %*% (x - mu)/2)

  out = constant * scalar

  return(as.numeric(out))
}
```

```r
x1 = seq(-2, 4, length = 75)
x2 = seq(-13, 11, length = 75)

results = matrix(NA, nrow = 75, ncol = 75)

for (i in 1:75) {
  for (j in 1:75) {
    results[i, j] = mvn(c(x1[i], x2[j]))
  }
}

persp(results, ticktype = "detailed")
```



ii.

$$\mu = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \Sigma = \begin{bmatrix} 1.0 & -1.6 \\ -1.6 & 4 \end{bmatrix}$$

$$d(P,Q) = \sqrt{\sigma_{11}(x_1-\mu_1)^2 + \sigma_{12}(x_1-\mu_1)(x_2-\mu_2) + \sigma_{22}(x_2-\mu_2)^2}$$

$$O = (0,0), P = (0,-2)$$

Statistical Distance

$$\begin{aligned} d(O,Q) =& \sqrt{1(0-1)^2 - 1.6(0-1)(0+1) + 4(0+1)^2)} \\ =& \sqrt{1+1.6+4} \\ =& \sqrt{6.6} = 2.569 \end{aligned}$$

$$\begin{aligned} d(P,Q) =& \sqrt{1(0-1)^2 - 1.6(0-1)(-2+1) + 4(-2+1)^2)} \\ =& \sqrt{1-1.6+4} \\ =& \sqrt{3.4} = 1.843 \end{aligned}$$

Straightline Distance

$$\begin{aligned} d(O,Q) =& \sqrt{(0-1)^2 + (0+1)^2} \\ =& \sqrt{(2)} \\ =& 1.41 \end{aligned}$$

$$\begin{aligned} d(P,Q) =& \sqrt{0-1)^2 + (-2+1)^2} \\ =& \sqrt{2} \\ =& 1.41 \end{aligned}$$

Conclusion: Using the straightline distance both P and O are equally distant from Q, but when taking into accoun the variances, P is actually closer to Q than O is.

iii.

For R_P, x_2 appears to no longer by centered on P and as such will have less volume under the curve than R_O which is centered on O for both x_1 and x_2, therefore $P(x \in R_O) > P(x \in R_P)$.