304a2

Part 1

(a)

```
N1 <- 120
n1 <- 6
N2 <- 180
n2 <- 9
total_x1 <- 24500
total_x2 <- 21200
x1 <- c(204, 143, 82, 256, 275, 198)
x2 <- c(137, 189, 119, 63, 103, 107, 159, 63, 87)
y1 <- c(210, 160, 75, 280, 300, 190)
y2 <- c(150, 200, 125, 60, 110, 100, 180, 75, 90)
x_mean1 <- mean(x1)
x_mean2 <- mean(x2)
y_mean1 <- mean(y1)
y_mean2 <- mean(y2)

total_y_hat <- N1 * y_mean1 + N2 * y_mean2
total_y_hat
```

```
## [1] 46100
```

```
s2_1 <- var(y1)
s2_2 <- var(y2)

var_hat_total_y_hat <- N1 ^ 2 * (1 - n1 / N1) * (s2_1 / n1) + N2 ^ 2 * (1 - n2 / N2) * (s2_2 / n2)
var_hat_total_y_hat
```

```
## [1] 23075975
```

Therefore, the basic estimate of the total potential is 46100 The Variance of my estimator is: 23075975

(b)

```
r1 <- y_mean1 / x_mean1
r2 <- y_mean2 / x_mean2
total_y_sr_hat <- r1 * total_x1 + r2 * total_x2
total_y_sr_hat
```

```
## [1] 48206.45
```

```
sr2_1 <- var(y1 - r1 * x1)
sr2_2 <- var(y2 - r2 * x2)
var_total_y_hat_sr_hat <- N1 ^ 2 * (1 - n1 / N1) * sr2_1 / n1 + N2 ^ 2 * (1 - n2 / N2) * sr2_2 / n2
var_total_y_hat_sr_hat
```

```
## [1] 564612.8
```

Therefore, the ratio estimate of the total potential sales is: 48206.45 The variance of my estimator is: 564612.8

(c)

```
fitreg1 <- lm(y1~x1)
mu_y_reg1_hat <- y_mean1 + coef(fitreg1)[2] * ((total_x1 / N1) - x_mean1)
```

```r
fitreg2 <- lm(y2~x2)
mu_y_reg2_hat <- y_mean2 + coef(fitreg2)[2] * ((total_x2 / N2) - x_mean2)

total_y_reg_hat <- N1 * mu_y_reg1_hat + N2 * mu_y_reg2_hat
total_y_reg_hat
```

```
##       x1
## 48353.87
```

```r
y_hat1 <- fitted(fitreg1)
MSE1 <- sum((y1 - y_hat1) ^ 2)/(n1 - 2)
y_hat2 <- fitted(fitreg2)
MSE2 <- sum((y2 - y_hat2) ^ 2)/(n2 - 2)

var_hat_total_y_reg_hat <- N1 ^ 2 * (1 - n1 / N1) * MSE1 / n1 + N2 ^ 2 * (1 - n2 / N2) * MSE2 / n2
var_hat_total_y_reg_hat
```

```
## [1] 558098.7
```

The regression estimate of the total potential sales is: 48353.87 The Variance of this estimate is: 558098.7

(d)

```r
realative_efficency_ratio2basic <- var_hat_total_y_hat/var_total_y_hat_sr_hat
realative_efficency_ratio2basic
```

```
## [1] 40.87044
```

The ratio is greater than 1. Ratio estimation is better than the basic estimation.

```r
realative_efficency_ratio2reg <- var_hat_total_y_reg_hat/var_total_y_hat_sr_hat
realative_efficency_ratio2reg
```

```
## [1] 0.9884627
```

The ratio is less than 1. Regression estimation is better than the ratio estimation.

```r
relative_efficiency_reg2basic <- var_hat_total_y_hat/var_hat_total_y_reg_hat
relative_efficiency_reg2basic
```

```
## [1] 41.34748
```

The ratio is greater than 1. Regression estimation is better than the basic estimation.

(e) We can see from part (d) that overall regression estimation behaves better. However, there is not many differences between the performance of regression estimation and ratio estimation. Therefore, when their results are similar, I would recommand the ratio estimation. Because there are too many calculations during the regression estimation.

Part2

(a)

```r
mu_y_hat <- total_y_hat / (N1 + N2)
mu_y_hat
```

```
## [1] 153.6667
```

```r
var_hat_mu_y_hat <- var_hat_total_y_hat / ((N1 + N2) ^ 2)
var_hat_mu_y_hat
```

```
## [1] 256.3997
```

The basic estimate of the mean potential sales is: 153.6667 The variance of this estimate is: 256.3997

(b)

```
x_barst <- N1 / (N1 + N2) * x_mean1 + N2 / (N1 + N2) * x_mean2
y_barst <- N1 / (N1 + N2) * y_mean1 + N2 / (N1 + N2) * y_mean2
mu_x <- (total_x1 + total_x2) / (N1 + N2)
r <- y_barst / x_barst
mu_y_hat <- r * mu_x
mu_y_hat
```

```
## [1] 160.6995
```

```
sr2_1 <- var(y1 - r * x1)
sr2_2 <- var(y2 - r * x2)

var_mu_y_hat_sr_hat <- (N1 / (N1 + N2))  ^  2 * (1 - n1 / N1) * sr2_1 / n1 + (N2 / (N1 + N2)) ^ 2 * (1 -
var_mu_y_hat_sr_hat
```

```
## [1] 6.189945
```

The ratio estimate of the mean potential sales is: 160.6995 The variance of the estimate is: 6.189945

(d)

```
relative_efficiency <- var_hat_mu_y_hat/var_mu_y_hat_sr_hat
relative_efficiency
```

```
## [1] 41.42197
```

The relative efficiency of ratio estimation to basic estimation is: 41.42197

(e) Since the relative efficiency is greater than 1. Ratio estimation is better.