**I apologize that it is not as well put together this week. I did not have time to make it look better. Thanks!**

**Problem 1**

1. The variable sex is significant at a 0.05 significance level. The variable income is significant at a 0.01 significance level.
2. As the variable “sex” is increased by one unit, it is estimated that the gamble variable will change by -22.11833 units (as long as all other variables are held constant). Because the variable only takes values of 0 and 1, this means that it is estimated to drop -22.11833 as it switches from sex = 0 to sex = 1 (as long as all other variables are held constant).
3. F = 4.1338, p-value = 0.01177. Statistically significant at significance level 0.05.

**Problem 2**

T = -1.878

Pval = 0.0667

Statistically significant at the 0.10 significance level

F = 4.0662

Pval = 0.01209

Statistically significant at the 0.05 significance level

T = 0.686

Pval = 0.496

Fail to reject null. We do not have statistically significant evidence that

F = 52.88

Pval is approximately 0.

Reject the null, statistically significant at and reasonable significance level.

The F-Test and t-test are not equivalent. Different decisions are made between rejecting the null and failing to reject the null at different significance levels.

**Problem 3**

Lcavol only: Residual standard error = 0.7875 and R-squared = 0.5394

Lweight added: Residual standard error = 0.7506 and R-squared = 0.5859

SVI added: Residual standard error = 0.7168 and R-squared = 0.6264

Lpbh added: Residual standard error = 0.7108 and R-squared = 0.6366

Age added: Residual standard error = 0.7073 and R-squared = 0.6441

LCP added: Residual standard error = 0.7102 and R-squared = 0.6451

PGG45 added: Residual standard error = 0.7048 and R-squared = 0.6544

Gleason added: Residual standard error = 0.7084 and R-squared = 0.6548

As the number of variables increased the RSE decreased on average, but R^2 increased. Generally, as one of the quantities went up, the other went down.

A graph with numbers and lines

Description automatically generatedA graph with numbers and lines

Description automatically generated

**Problem 4**

and

R Code:

library(faraway)

#Problem 1

out\_full <- lm(gamble ~ sex + status + income + verbal, data = teengamb)

summary(out\_full)

RSS\_full <- sum(out\_full$residuals^2)

out\_red <- lm(gamble ~ income, data = teengamb)

summary(out\_red)

RSS\_red <- sum(out\_red$residuals^2)

anova(out\_full, out\_red)

F <- ((RSS\_red - RSS\_full)/3) / (RSS\_full/42)

F

p\_value <- pf(F, lower.tail = FALSE, df1 = 3, df2 = 42)

p\_value

#Problem 2

out\_1 <- lm(total ~ expend + ratio + salary, data = sat)

out\_2 <- lm(total ~ 1, data = sat)

anova(out\_1, out\_2)

out\_3 <- lm(total ~ expend + ratio + salary + takers, data = sat)

summary(out\_3)

#Problem 3

out\_pro\_1 <- lm(lpsa ~ lcavol, data = prostate)

summary(out\_pro\_1)

out\_pro\_2 <- lm(lpsa ~ lcavol + lweight, data = prostate)

summary(out\_pro\_2)

out\_pro\_3 <- lm(lpsa ~ lcavol + lweight + svi, data = prostate)

summary(out\_pro\_3)

out\_pro\_4 <- lm(lpsa ~ lcavol + lweight + svi + lbph, data = prostate)

summary(out\_pro\_4)

out\_pro\_5 <- lm(lpsa ~ lcavol + lweight + svi + lbph + age, data = prostate)

summary(out\_pro\_5)

out\_pro\_6 <- lm(lpsa ~ lcavol + lweight + svi + lbph + age + lcp, data = prostate)

summary(out\_pro\_6)

out\_pro\_7 <- lm(lpsa ~ lcavol + lweight + svi + lbph + age + lcp + pgg45,

data = prostate)

summary(out\_pro\_7)

out\_pro\_8 <- lm(lpsa ~ lcavol + lweight + svi +

lbph + age + lcp + pgg45 + gleason,

data = prostate)

summary(out\_pro\_8)

RSE\_list <- c(0.7875, 0.7506, 0.7168, 0.7108, 0.7073, 0.7102, 0.7048, 0.7084)

plot(RSE\_list, type = "l")

R\_squared\_list <- c(0.5394, 0.5859, 0.6264, 0.6366, 0.6441, 0.6451,0.6544, 0.6548)

plot(R\_squared\_list, type = "l")