Test Exercise 2

a) Linear model: FGPA against a constant and SATV
    i. Coefficient of SATV: 0.06309
        i. Standard error: 0.02766
        ii. P-value: 0.0229
    ii. 95% confidence level (+/- 2 standard deviations): [0.00777, 0.11841]
b) Linear model: FGPA against a constant, SATV, SATM, and FEM
    i. Coefficients, standard errors, and p-values
        i. SATV
            1. Coefficient: 0.01416
            2. Standard Error: 0.02793
            3. P-value: 0.612
        ii. SATM
            1. Coefficient: 0.17274
            2. Standard error: 0.03193
            3. P-value: 0 [$9.07*10^{-8}$]
        iii. FEM
            1. Coefficient: 0.20027
            2. Standard error: 0.03738
            3. P-value: 0 [$1.2*10^{-7}$]
    ii. 95% confidence level (+/1 2 standard deviations)
        i. SATV: [-0.04169, 0.07002]
        ii. SATM: [0.10888, 0.23659]
        iii. FEM: [0.12551, 0.27503]
c) Correlation Matrix:

	FGPA	SATM	SATV	FEM
FGPA	1.000	0.195	0.092	0.176
SATM	0.195	1.000	0.288	-0.163
SATV	0.092	0.288	1.000	0.034
FEM	0.176	-0.163	0.034	1.000

This matrix explains the difference in the values for the coefficient of SATV in part (a) (where SATV alone explained FGPA) versus in part (b) (where SATV explained FGPA along with SATM and FEM). In part (a), the coefficient on SATV was 0.06309, meaning that a 1 point increase in SAT verbal score would be expected to result in a 0.06309 increase in the freshman GPA of a particular student. We can be confident that this coefficient is not 0 at the 5% significance level (p = 0.0229, and the 95% confidence interval for the coefficient is [0.00777, 0.11841]).

In part (b), the coefficient on SATV was just 0.01416, meaning that a 1 point increase in SAT verbal score would be expected to result in only a 0.01416 increase in the freshman GPA of a particular student. Additionally, we cannot be confident

that this coefficient is not 0—the p-value for this statistic is 0.612, and the 95% confidence interval of [-0.04169, 0.07002] contains the value 0.

This difference arises because the first model uses only SATV to predict freshman GPA, whereas the second model uses SATV along with SATM and FEM. We can see in the table that although SATV does correlate with FGPA, it has the weakest correlation with FGPA of all the explanatory factors (its correlation with FGPA is just 0.092, compared with 0.195 for SATM and 0.176 for FEM). So SATV is useful in predicting freshman GPA, but in a model that uses explanatory factors other than SATV, we ought to expect that SATV will be weighed less heavily in predicting FGPA than in a model that only uses SATV as an explanatory factor. We would expect the coefficient of SATV to be closer to 0 and the p-value for SATV to be larger in the model with more explanatory factors.

In addition, we can see that SATV is correlated very little with FEM ($r = 0.034$) and correlated somewhat with SATM ($r = 0.288$). This means that some of the predictive value of SATV for predicting FGPA may be accounted for with the inclusion of SATM and FEM, so in a model that incorporates SATM and FEM in addition to SATV, we would expect the coefficient of SATV to be closer to 0 and the p-value for SATV to be larger.

d)

i.

i. $\beta 2$: coefficient of SATV.
ii. H0: $\beta 2 = 0$
iii. H1: $\beta 2 \neq 0$
iv. Unrestricted regression: includes SATV, SATM, and FEM.
v. Restricted regression: includes SATM and FEM.
vi. Calculated F-value: F = 0.257580
vii. Because $F = 0.257580 < 3.9$, we fail to reject the null hypothesis at the 5% significance level that $\beta 2 = 0$.

ii. t-value of SATV = 0.507. $t^2 = 0.257 = F$