Lab 4

1. One theory of modernization argues that foreign investments are never altruistically motivated. Instead, foreign-investment-benefits to citizens of developing countries are entirely due to the attractiveness of their countries' natural resources to foreign powers. You wish to evaluate this theory using data from 100 developing countries on the following three variables:

\[ Y = \text{median income of the country's citizens (in thousands of dollars)} \]
\[ X_1 = \text{foreign investment (in millions of dollars) in the country} \]
\[ X_2 = \text{oil, coal, and minerals (in thousands of tons) exported from the country} \]

The following quantities were calculated from the data:

<table>
<thead>
<tr>
<th></th>
<th>(Y)</th>
<th>(X_1)</th>
<th>(X_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>5.1</td>
<td>12.3</td>
<td>51</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>1.9</td>
<td>6.2</td>
<td>12</td>
</tr>
</tbody>
</table>

\[
(X^T X)^{-1} = \begin{bmatrix}
0.192718 & 0.00030657 & -0.00365664 \\
0.00030657 & 0.00035037 & -0.000090511 \\
-0.00365664 & -0.000090511 & 0.0000935279
\end{bmatrix}
\]

\[
X^T y = \begin{bmatrix} 510.00 \\ 6506.24 \\ 26912.90 \end{bmatrix}
\]

NOTE: The design matrix, \( X \), used in computing in the above matrices was NOT a matrix of standardized variables!

a. Find the appropriate standardized and unstandardized regression equations. Express the unstandardized partial slope, \( \hat{b}_1 \), in words.

b. If the metric on \( Y \) was changed from thousands of dollars to dollars (e.g., resulting in the average median income changing from 5.1 thousand dollars to 5,100 dollars), what would be the resulting standardized and unstandardized regression equations?

c. What proportion of the variance in median income that is not explained by foreign investment is explained by resources exported? Is this proportion statistically significant at the .05 level?

Hints: In finding \( r_{YX_1} \) and \( r_{YX_2} \), \( \sum YX_1 \) and \( \sum YX_2 \) can be obtained...
from $X^T y$. In finding $r_{x_1 x_2}$, recall from the lecture notes (p. 23) what the second and third diagonal elements (i.e., $x_{11}$ and $x_{22}$) of the $(X^T X)^{-1}$ matrix are. To determine the sign of $r_{x_1 x_2}$, you can work from the formula for either standardized slope.

d. What is the increment (specifically, over the explanatory role of exported resources) in the proportion of explained variance in median income that is due to foreign investment? Is this a statistically significant amount of variance at the .05 level? Does this finding provide support for (or against) the theory that foreign investment is never altruistically motivated? (EXPLAIN YOUR ANSWER.)

2. With the fall of the Soviet Union in 1991, Uzbekistan (a small country located a few hundred miles east of the Caspian Sea) became an independent state. Political independence brought the country fundamental economic changes as a controlled communist economy was replaced by a free market economy. Independence also allowed the emergence of cultural norms that had been illegal under Soviet rule. For example, in rural towns such as Namangan "religious patrols" began roaming the streets enforcing the veiling of women, the ban of alcohol, and other Islamic practices. In Uzbekistan's cities Islamic norms are still prevalent, but are less militantly enforced. An Islamic norm of particular interest to you in your research is the secondary role of women to men in economic matters. It is because of this role, you believe, that Uzbek women are less likely to be employed in rural areas than in urban areas. In contrast to your position, other researchers have argued that this tendency for urban women to be employed results simply because there are more jobs available in cities than in rural areas. You decide to investigate whether cultural biases or a lack of jobs undermines Uzbek women's labor market participation. You obtain data on the following three variables for each of 21 Uzbek cities and towns:

RELIG number of persons per 1,000 who visit a mosque at least weekly (This is your measure of the prevalence of Islamic norms in a region.)

NOEMPLOY number of unemployed persons per 1,000 population (This is your measure of the availability of jobs. Your position is that if a region's unemployment rate is high, the region will have fewer jobs available.)

WOMEN number of women employed per 1,000 workers (This your measure of women's participation in the labor force.)
Correlations, means, and standard deviations among these variables are as follows:

<table>
<thead>
<tr>
<th></th>
<th>RELIG</th>
<th>NOEMPLOY</th>
<th>WOMEN</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>RELIG</td>
<td>1.00</td>
<td>.80</td>
<td>-.40</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td>NOEMPLOY</td>
<td>.80</td>
<td>1.00</td>
<td>-.60</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>WOMEN</td>
<td>-.40</td>
<td>-.60</td>
<td>1.00</td>
<td>200</td>
<td>72</td>
</tr>
</tbody>
</table>

(Be sure that you read these data correctly. For example, the zero-order correlation between RELIG and NOEMPLOY is .80; the mean RELIG score is 400 persons; the standard deviation for RELIG is 100 persons; etc.)

a. Find the unstandardized regression equation (i.e., compute numbers, please) for the regression of WOMEN on RELIG and NOEMPLOY. Express in words the meaning of the partial slope associated with NOEMPLOY.

b. Find the total, regression, and error sums of squares for the regression described in part a. Do the marginal effects of both RELIG and NOEMPLOY explain (at \( \alpha = .05 \)) a significantly large proportion of the variance in WOMEN? Be sure to state your conclusion in words. (Hints: The phrase "marginal effects of both" refers to the total effects of two variables, not part of the effect of one, or of some single measure constructed out to the two. The total sum of squares can be obtained based on your knowledge of the sample size and the standard deviation of the variable, WOMEN.)

c. At the .05 significance level and based on coefficients computed in part a, determine if either the partial effects of cultural biases or those of a lack of jobs undermines Uzbek women's labor force participation. (Hints: In your answer you will need to perform one or more hypothesis tests related to coefficient[s] found in part a. Be sure and remember to state your conclusion in words.)

d. Using the data provided at the beginning of this problem, obtain an equation (with numbers, please) that would convert NOEMPLOY into a new variable, NOEMPLOYADJUST, such that this new variable has the linear effects of RELIG removed. What would be the value of the slope (no calculation, just a number is called for) in the bivariate regression of WOMEN on NOEMPLOYADJUST? Explain your answer, please.

3. This problem involves your use of matrix algebra: When a matrix is not square (i.e., when it is rectangular), it may have a right-inverse or a left-inverse. For example, if A has a right-inverse, \( AA^{-1} = I \).
If $B$ has a left-inverse, $B^{-1}B = I$. Moreover, if a rectangular matrix has more columns than rows, it may have a right-inverse, but it does not have a left-inverse. Likewise, if a rectangular matrix has more rows than columns, it may have a left-inverse, but it does not have a right-inverse. Use the following two matrices to illustrate these characteristics of matrix inverses:

\[
C = \begin{bmatrix}
-1 & -1 \\
-1 & 0 \\
1 & 1
\end{bmatrix} \quad D = \begin{bmatrix}
1 & -1 & 1 \\
1 & 1 & 2
\end{bmatrix}
\]

a. Multiply the two above matrices twice, first by computing $CD$ and then by computing $DC$. (Be sure to show all intermediate steps in these multiplications. That is, be sure and show your work!)

b. Which matrix is the left-inverse of the other? Which is the right-inverse of the other?

c. What is the dimension of the identity matrix that is produced when the two matrices are multiplied in the sequence, left-inverse times right-inverse?