Option 1:
Background and Motivation

Although time spent with peers is a known risk factor for delinquency (Osgood and Anderson, 2004; Osgood et al., 1996), time with peers matters most in the absence of responsible authority figures (i.e., “unstructured socializing” or unsupervised time with peers). In other words, parental supervision reduces delinquency and the impacts of criminogenic variables. However, females are generally supervised more so than males, so the effect of parental supervision might vary by sex.

Estimate an ordinary least squares multiple regression model to test the following hypothesis:

Parental supervision has a negative effect on criminal offending, but this effect varies with the sex of the respondent. The effect of parental supervision on delinquency is thought to be larger for females.

Hold constant the potentially confounding effects of: (1) time spent with peers; (2) age of the respondent; and (3) race of the respondent.

You may use a calculator.
You will be assessed based on your responses to the following items:

1. Using the data file described below, use SPSS to estimate a multivariate ordinary least squares regression equation with crime as the dependent variable. The primary independent variables are: (1) the respondent’s sex; (2) the extent of the respondent being supervised by parents; and (3) the statistical interaction (i.e., the product term) of the respondent’s sex and the level of parental supervision they receive. The secondary variables (that is, the control variables) are: time spent with peers, age of the respondent, and race of the respondent.
2. Based on the model you estimated, interpret (a) the y-intercept; (b) the slopes (i.e., the coefficients) for the primary independent variables; and (c) their tests of statistical significance. Assume that all continuous variables have been centered on their means.

3. Explain whether and how the results support (or reject) the motivating hypothesis.

4. Explain and discuss the error-term assumptions of the estimated model. Also, explain and discuss what the residuals from the estimated model indicate with regard to these assumptions. If evidence for problems exists, do not address problems with additional analysis.

5. Explain and discuss (1) collinearity, which is also known as multicollinearity; (2) the estimated model’s assumptions regarding it; and (3) the consequences of collinearity.

6. Another researcher takes issue with the hypothesis on this exam. The researcher is interested in the effect of time spent with peers on delinquency. They argue that time with peers increases delinquency, but as parental supervision increases, the effect of time with peers decreases. Does the estimated model (as it stands and without modification) provide evidence for their hypothesis?
The data file contains data from 750 individual respondents. The variables relevant to the exam are named and described below.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Variable Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>crime</td>
<td>A continuous measure for criminal offending. Higher values indicate more offending. Refer to units of this variable as points on the crime scale.</td>
</tr>
</tbody>
</table>
| male          | A dummy-coded variable for sex of respondent.  
0 = not male  
1 = male |
| parent        | A mean-centered and continuous measure of supervision of a parent (or responsible authority figure). Higher values indicate higher levels of parental supervision. Refer to units on this scale as points on parental supervision scale. |
| product       | The product-term for the statistical interaction between the sex of the respondent and their level of parental supervision (that is, product = male × parent). |
| peertime      | A mean-centered and continuous measure for time spent with peers. Higher values indicate more time with peers. Refer to units of this variable as points on the time with peers scale. |
| age           | A mean-centered and continuous measure for the respondent’s age in years. Higher values indicate older respondent. Refer to units of this variable as years of age. |
| white         | A dummy-coded variable for race of respondent.  
0 = not white  
1 = white |

End of Option 1
Option 2:  
Background and Motivation

Whether the American justice system delivers consistent and equal punishments across offenders is a widely studied topic. However, there is disagreement regarding whether the objective is to issue a punishment that “fits the crime” or a punishment that “fits the criminal.” Furthermore, regarding the possibility of gender bias in the system, there is a chivalry hypothesis (Pollock, 1999) whereby females are treated more leniently by the system.

The investigator has collected sentencing data and has estimated a standard linear regression model. Both male and female offenders are in the sample, and the investigator is interested in whether the gender gap in sentencing changes based on the severity of the crime. The results of the estimation are presented in the table below.

The dependent variable is length of criminal sentence. The primary independent variables are: (1) a dummy-coded variable for respondent sex; (2) a mean-centered continuous variable for severity of the offense; and (3) the product-term for the statistical interaction between sex and severity of the offense.

You may use a calculator.
You will be assessed based on your responses to the following items:

1. Based on the results, interpret the model-fit statistics.

2. Interpret (a) the y-intercept; (b) the slopes (i.e., the coefficients) for the primary independent variables; and (c) their tests of statistical significance.

3. Discuss if and how the results support or reject the motivating hypothesis.

4. The standard deviation of severity of offense is 3.25. What is the gender gap in sentencing when the severity of the offense is two standard deviations above the mean severity? Report the actual numeric value.

5. Assume the standard error for the gender coefficient presented in the table remains constant. Is the quantity provided for question 4 statistically significant? In other words, is the gender gap statistically significant when severity of the offense is two standard deviations above the mean severity?

6. Explain and discuss the error-term assumptions of the estimated model.

7. Explain and discuss (1) collinearity, which is also known as multicollinearity; (2) the estimated model’s assumptions regarding it; and (3) the consequences of collinearity.
### Table for statistics exam, option 2

Ordinary least squares model explaining sentence length

N = 870

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-1.157</td>
<td>8.674</td>
<td>0.894</td>
</tr>
<tr>
<td>Severity of offense&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.681</td>
<td>0.257</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Female&lt;sup&gt;a&lt;/sup&gt; × Severity of offense&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-1.230</td>
<td>0.378</td>
<td>0.001</td>
</tr>
<tr>
<td>First-time offender&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-24.079</td>
<td>12.398</td>
<td>0.052</td>
</tr>
<tr>
<td>Private legal counsel&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-3.451</td>
<td>1.168</td>
<td>0.003</td>
</tr>
<tr>
<td>Mitigating factors&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-9.476</td>
<td>5.236</td>
<td>0.071</td>
</tr>
<tr>
<td>Education&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.123</td>
<td>4.282</td>
<td>0.977</td>
</tr>
<tr>
<td>Age&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.000</td>
<td>1.596</td>
<td>0.210</td>
</tr>
<tr>
<td>White&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-3.967</td>
<td>7.683</td>
<td>0.606</td>
</tr>
<tr>
<td>Constant</td>
<td>13.541</td>
<td>35.835</td>
<td>0.706</td>
</tr>
</tbody>
</table>

R² = 0.1140  
Root MSE = 89.508  

Model-fit Statistics

\[ F_{df_1=9, df_2=860} = 12.30, p < .05 \]

<sup>a</sup> A dummy-coded variable where zero indicates absence of characteristic.

<sup>b</sup> A continuous variable that is mean centered.

---

**End of Option 2**