Based on your regression model from problem set #4,

\[
Wage = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + U.
\]

Where \( X_1 \) = Experience.

\( X_2 \) = Female.

\( X_3 \) = Interaction variable between experience and female.

(a) Construct a variable to account for a non-linear relationship between wages and experience. Write out your new regression model and explain and interpret the variables.

Adding \( X_4 = X_1^2 \) to the model to account for non linear returns to experience. The model becomes,

\[
Wage = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + U.
\]

(b) (i) What are the expected signs of the coefficients?

(ii) Write out the regression equation for men

\[
Wage = \beta_0 + \beta_1 X_1 + \beta_4 X_4 + U.
\]

(iii) Write out the regression equation for women.

\[
Wage = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) X_1 + \beta_4 X_4 + U.
\]

(c) Using the data set Wage1.dta, estimate the above equation:

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 526</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1492.95648</td>
<td>4</td>
<td>373.239119</td>
<td>F( 4, 521) = 34.31</td>
</tr>
<tr>
<td>Residual</td>
<td>5667.45781</td>
<td>521</td>
<td>10.878038</td>
<td>Prob &gt; F = 0.0000</td>
</tr>
<tr>
<td>Total</td>
<td>7160.41429</td>
<td>525</td>
<td>13.638844</td>
<td>Adj R-squared = 0.2024</td>
</tr>
</tbody>
</table>

| wage      | Coef.    | Std. Err. | t     | P>|t|  [95% Conf. Interval] |
|-----------|----------|-----------|-------|-----|-----------------------|
| exper     | .3101882 | .039893   | 7.78  | 0.000 | .2318173 .3885591     |
| female    | -1.421982| .4618985  | -3.08 | 0.002 | -2.329394 -.5145693   |
| fexper    | -.0572396| .0212425  | -2.69 | 0.007 | -.0989711 -.0155081   |
| expersq   | -.0058552| .0008455  | -6.93 | 0.000 | -.0075162 -.0041942   |
| _cons     | 4.521433 | .4037018  | 11.20 | 0.000 | 3.728349 5.314516     |
(i) How much of the variation in wages is explained by the model.
   \[ R^2 = .2024 \]

(ii) Interpret your estimated coefficients. Are your estimated coefficients significant?

(iii) Show your regression equations graphically in a scatter diagram. Describe your graph.

(d) Construct a variable that allows for differing rates of diminishing returns between men and women. Write out your new regression model and explain and interpret the variables.

Adding \( X_5 = X_1^2 \cdot X_2 \) to the model to account for differing non-linear returns to experience. The model becomes,

\[
Wage = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + U.
\]

(e) (i) What are the expected signs of the coefficients?

(ii) Write out the regression equation for men.

\[
Wage = \beta_0 + \beta_1 X_1 + \beta_2 X_4 + U.
\]

(iii) Write out the regression equation for women.

\[
Wage = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) X_1 + (\beta_4 + \beta_5) X_4 + U.
\]
(f) Estimate the above equation:

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs =526</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1617.89444</td>
<td>5</td>
<td>323.578889</td>
<td>F(  5,   520) = 30.36</td>
</tr>
<tr>
<td>Residual</td>
<td>5542.51985</td>
<td>520</td>
<td>10.658692</td>
<td>Prob &gt; F = 0.0000</td>
</tr>
<tr>
<td>Total</td>
<td>7160.41429</td>
<td>525</td>
<td>13.6388844</td>
<td>R-squared = 0.2259</td>
</tr>
</tbody>
</table>

|                       |            |     |            | Adj R-squared= 0.2185 |
|                       |            |     |            | Root MSE = 3.2648 |

\[
\begin{array}{|c|c|c|c|c|c|c|}
\hline
\text{wage} & \text{exper} & \text{female} & \text{fexper} & \text{expersq} & \text{femexp2} & \text{cons} \\
\hline
\text{Coef.} & 0.4279455 & 0.1184071 & -0.3077607 & -0.0085424 & 0.0057427 & 3.770214 \\
\text{Std. Err.} & 0.0523676 & 0.6414639 & 0.076134 & 0.0011474 & 0.0016773 & 0.4558869 \\
\text{t} & 8.17 & 0.18 & -4.04 & -7.45 & 3.42 & 8.27 \\
\text{P>t} & 0.000 & 0.854 & 0.000 & 0.000 & 0.001 & 0.000 \\
\text{[95% Conf. Interval]} & 0.3250675 & -1.141772 & -0.457328 & -0.0107965 & 0.0024475 & 2.874608 \\
\text{Root MSE} & 3.2648 \\
\hline
\end{array}
\]

(i) How much of the variation in wages is explained by the model?
\[ R^2 = .2185 \]

(ii) Interpret your estimated coefficients. Are your estimated coefficients significant?

(iii) Show your regression equations graphically in a scatter diagram. Describe your graph.