Homework 5

Rucha Rane

1. Before the Interaction, adjusted R-squared= 0.555. So, 55.5% of the Variance in Income is explained by Education and PctWomen.
After the interaction terms were introduced in the model, adjusted R-squared=0.577. So, 57.7% of the variance in Income was now explained by the independent variables.
Also, significance F change was =0.016, which is statistically significant (@ <0.005). This implies that adding the interaction terms to the model increased the model fit significantly (increase in model fit due to interaction effects was found to be statistically significant).

Regression Line: 6938.857 -882.297 Educ + 64.91 PctWomen – 9.463(Educ * PctWomen)

```
NEW FILE.
DATASET NAME DataSet1 WINDOW=FRONT.
NEW FILE.
DATASET NAME DataSet2 WINDOW=FRONT.

GET DATA
  /TYPE=XLS
  /FILE='C:\Users\rrane1\Desktop\occupation - homework 5.xls'
  /SHEET=name 'occupation'
  /CELLRANGE=full
  /READNAMES=on
  /ASSUMEDSTRWIDTH=32767.
EXECUTE.
DATASET NAME DataSet3 WINDOW=FRONT.
AGGREGATE
  /OUTFILE=* MODE=ADDVARIABLES
  /BREAK=
  /education_mean=MEAN(education)
  /PctWomen_mean=MEAN(PctWomen).
COMPUTE deveducation=education_mean - education.
EXECUTE.
COMPUTE devPctWomen=PctWomen_mean - PctWomen.
EXECUTE.
COMPUTE educationPctWomendevprod=devPctWomen * deveducation.
EXECUTE.
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT income
  /METHOD=ENTER devPctWomen deveducation
  /METHOD=ENTER educationPctWomendevprod
  /RESIDUALS DURBIN
  /CASEWISE PLOT(ZRESID) OUTLIERS(3).
```

Regression

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	deveducation, devPctWomen[b]	.	Enter
2	educationPctWomendevprod[b]	.	Enter

a. Dependent Variable: income

b. All requested variables entered.

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.751[a]	.564	.555	2820.911	.564	61.458	2	95	.000	
2	.768[b]	.590	.577	2749.817	.026	5.976	1	94	.016	1.792

a. Predictors: (Constant), deveducation, devPctWomen

b. Predictors: (Constant), deveducation, devPctWomen, educationPctWomendevprod

c. Dependent Variable: income

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	978112712.197	2	489056356.099	61.458	.000[b]
	Residual	755966449.803	95	7957541.577		
	Total	1734079162.000	97			
2	Regression	1023298865.742	3	341099621.914	45.110	.000[c]
	Residual	710780296.258	94	7561492.513		
	Total	1734079162.000	97			

a. Dependent Variable: income

b. Predictors: (Constant), deveducation, devPctWomen

c. Predictors: (Constant), deveducation, devPctWomen, educationPctWomendevprod

Coefficients[a]

Model	Unstandardized Coefficients	Standardized Coefficients	t	Sig.	Collinearity Statistics

		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	6938.857	284.955		24.351	.000		
	devPctWomen	65.364	9.144	.485	7.148	.000	.996	1.004
	deveducation	-929.227	104.393	-.604	-8.901	.000	.996	1.004
2	(Constant)	6988.895	278.527		25.092	.000		
	devPctWomen	64.910	8.916	.482	7.280	.000	.996	1.004
	deveducation	-882.297	103.557	-.574	-8.520	.000	.962	1.040
	educationPctWomendevprod	-9.463	3.871	-.164	-2.445	.016	.966	1.036

a. Dependent Variable: income

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics		
						Tolerance	VIF	Minimum Tolerance
1	educationPctWomendevprod	-.164[b]	-2.445	.016	-.244	.966	1.036	.962

a. Dependent Variable: income

b. Predictors in the Model: (Constant), deveducation, devPctWomen

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	devPctWomen	deveducation	educationPctWomendevprod
1	1	1.062	1.000	.00	.47	.47	
	2	1.000	1.030	1.00	.00	.00	
	3	.938	1.064	.00	.53	.53	
2	1	1.204	1.000	.05	.02	.36	.37
	2	1.024	1.084	.37	.50	.06	.04
	3	.981	1.108	.51	.41	.05	.04
	4	.791	1.233	.07	.06	.53	.55

a. Dependent Variable: income

Casewise Diagnostics[a]

Case Number	Std. Residual	income	Predicted Value	Residual
2	5.684	25879	10248.00	15631.003
24	4.242	25308	13642.49	11665.510

a. Dependent Variable: income

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1666.82	14330.21	6938.86	3247.995	98
Residual	-8055.571	15631.003	.000	2706.960	98
Std. Predicted Value	-1.623	2.276	.000	1.000	98
Std. Residual	-2.929	5.684	.000	.984	98

a. Dependent Variable: income

2. After including the centered interaction terms, 57.7% of the variance in Income is accounted for by this Multiple Regression Model.

Both the F test and the t-test show that the regression results are significant at 0.000 for both Centered_Education and Centered_Pct Women variables.

Also, the interaction between the two centered variables is found to be significant at 0.016.

Regression Line: `6988.895 + 882.297 Educ - 64.91 PctWomen – 9.463(Educ * PctWomen)`

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.768[a]	.590	.577	2749.817

a. Predictors: (Constant), Educ_PctWomen_Centered, PctWomen_Centered, Educ_Centered

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1023298865.742	3	341099621.914	45.110	.000[b]
	Residual	710780296.258	94	7561492.513		
	Total	1734079162.000	97			

a. Dependent Variable: income

b. Predictors: (Constant), Educ_PctWomen_Centered, PctWomen_Centered, Educ_Centered

2.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	6988.895	278.527		25.092	.000
	Educ_Centered	882.297	103.557	.574	8.520	.000
	PctWomen_Centered	-64.910	8.916	-.482	-7.280	.000
	Educ_PctWomen_Centered	-9.463	3.871	-.164	-2.445	.016

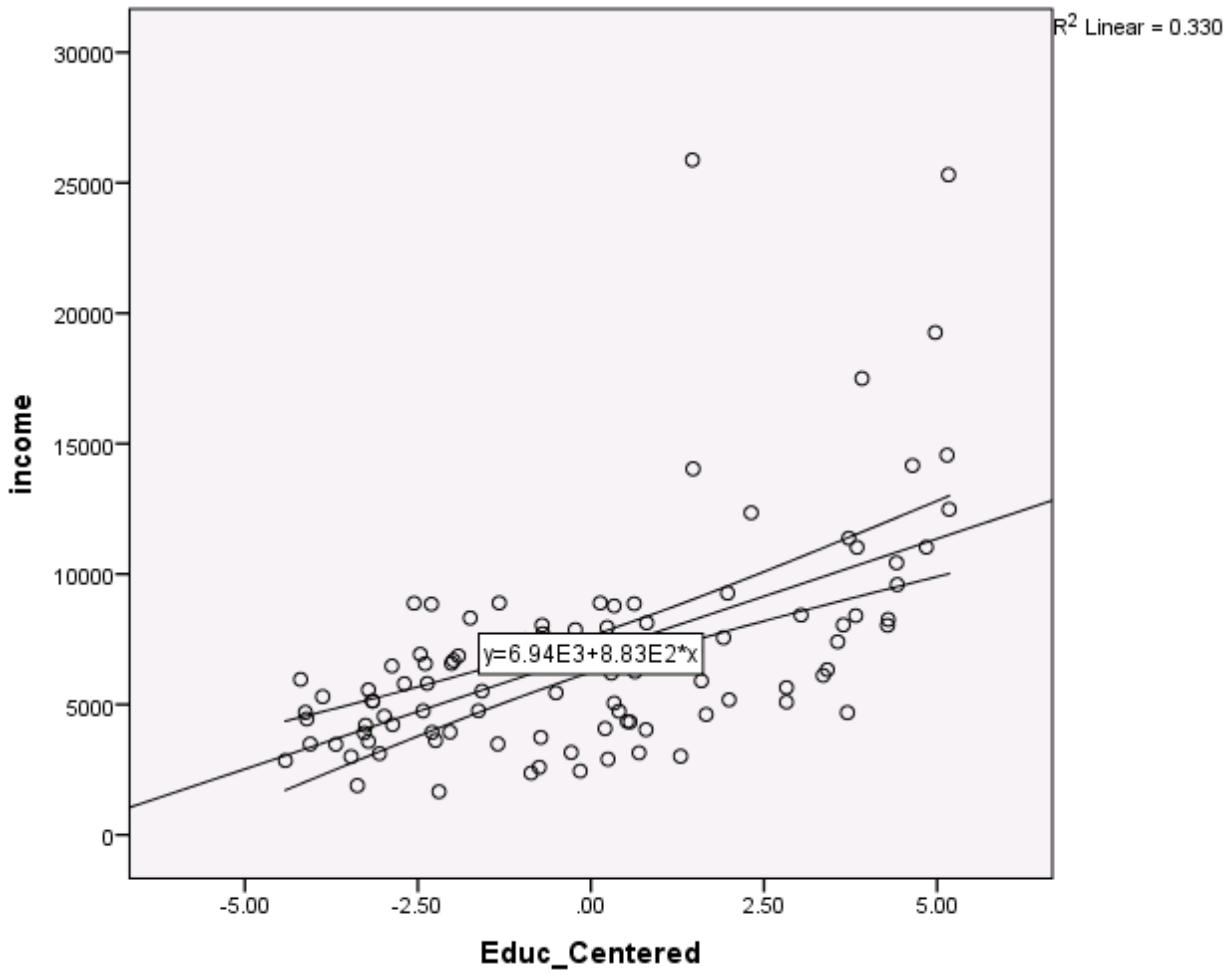
a. Dependent Variable: income

Also, Beta1=882.297. This implies that the expected change in the Income is 882.297 units for an incremental change in Education_Centered, when Education_Centered=0, holding Education_PctWomen_Centered (Interaction Term) constant.

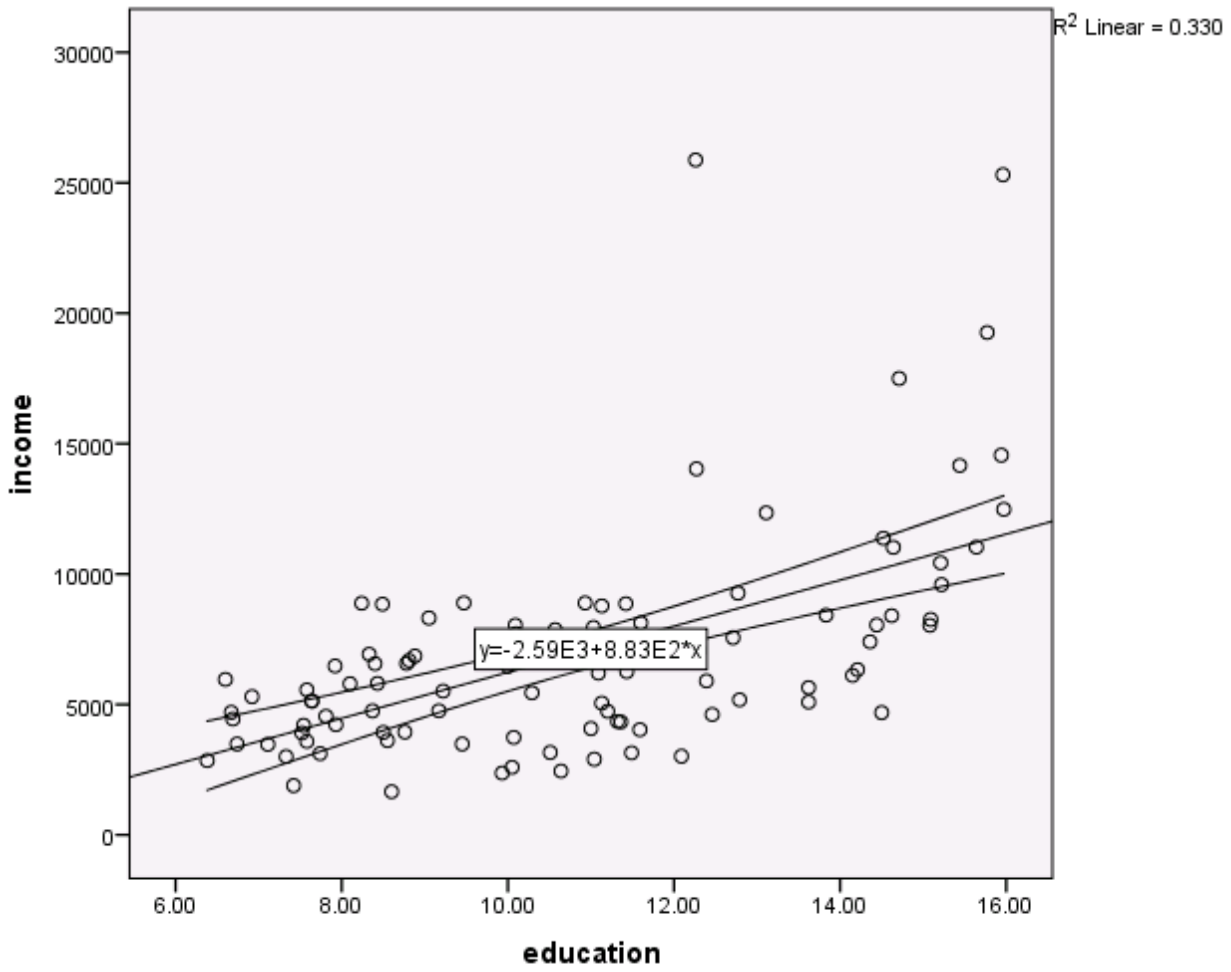
Beta 2= -64.910. This implies that the measure of the incremental change in the slope of X/Y line (change in Income with respect to Educ_Centered and PctWomen_Centered) is 64.91 units less for an incremental change in Education Centered, holding Education_PctWomen_Centered (Interaction Term) constant.

3. Centering the data did not change R-squared, which remained at 57.7%. Beta for the highest order coefficient is the same = 882.297. There seems to be no interation, because centering did not chance anything except Beta 0, which changed from 6938.857 to 6988.895

4. PctWomen = mean -SD, mean, mean + SD, according to question 1;



4. **PctWomen = mean -SD, mean, mean + SD, according to question 1;**



6. Hypothesis Test :