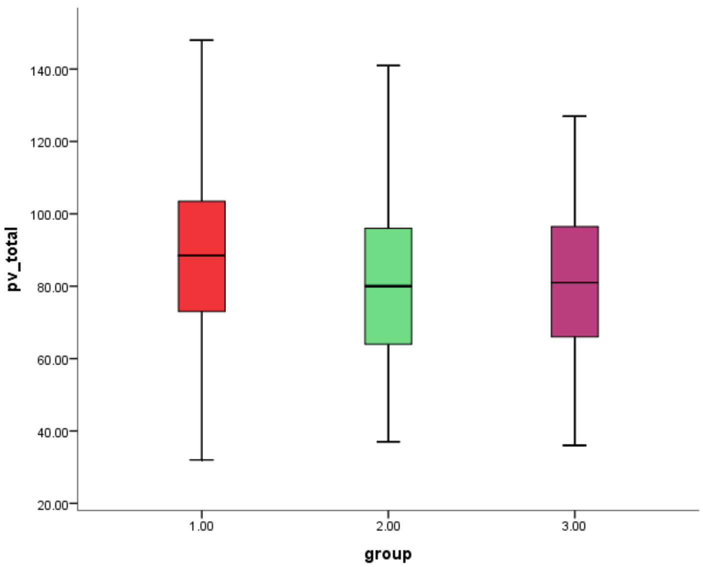
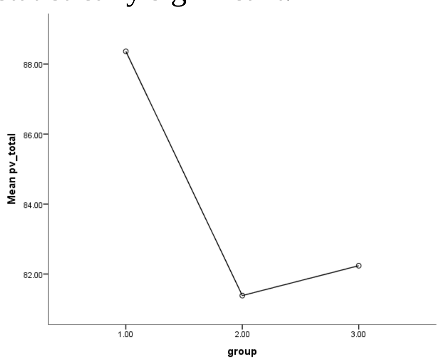
PART 1: One-Way ANOVA

1. The order of means: group 1 > group 3 > group 2. The group that had lost one of their parents have a larger mean score of perceived vulnerability than the other 2 group. The group with 2 living married parents have the lowest mean vulnerability score of all. However, the 3 groups have overlapping variances (error bars), which makes it difficult to conclude significant statistical differences between them.



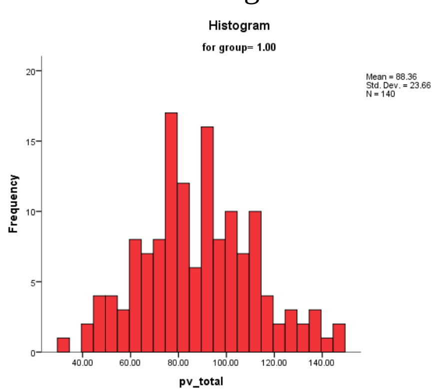
2. $H_0: \mu_1 = \mu_2 = \mu_3$
3. $H_1: At\ leas\ one\ \mu_j\ is\ not\ equal\ to\ others.$
4. The sample means are different, which is not compatible with the null hypothesis. But we can't conclude that we reject the null, because we don't know if the difference is statistically significant.

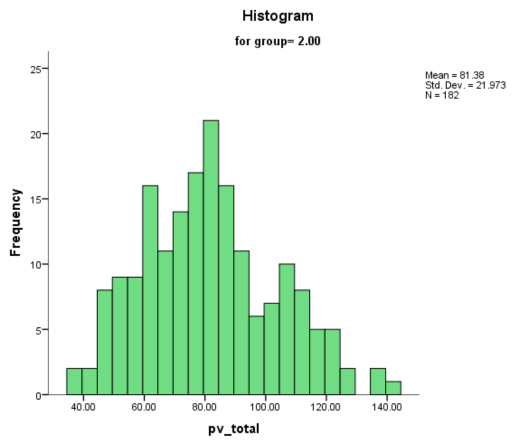


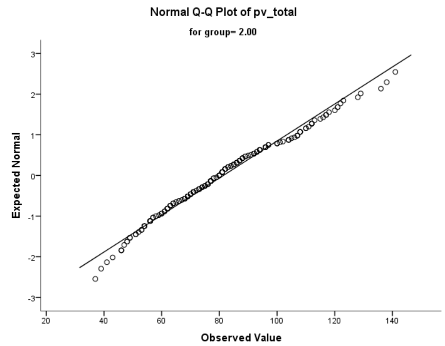
5. Although the group means are different, the variability within groups is so high and overlapping that it seems unlikely that the groups differ significantly in terms of pv scores.
6. $SS_{between} = 140(88.3643 - 84.0814)^2 + 182(81.3846 - 84.0814)^2 + 59(82.2373 - 84.0814)^2 = 4092.3310$

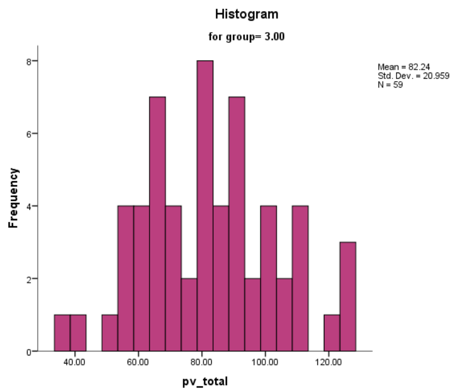
$$MS_{between} = \frac{4092.3310}{2} = 2046.1655$$

7. It indicates the average variation among sample means, and shows the part of the total variation explained by the group differences.
8. No. We also need to know about the average variation among individual observations within groups and compare these two in order to determine whether the observed variation among the sample means is due largely to chance, or if there are real differences between them.
9. $R^2 = \eta^2 = \frac{SS_{between}}{SS_{toatal}} = \frac{SS_{between}}{SS_{toatal}} = \frac{4092.3310}{194772.4776} = 0.021$
10. 2% of the variance in the perceived vulnerability score is explained by group membership (being in one of the 3 groups in our sample).
11. $F_{critical}(2, 378) = 3.019$
12. The F value is 4.056, which exceeds the critical F value. Hence, we reject the null hypothesis.
13. At least one of group's population mean (of pv score) is different from the others.
14. The p-value is .018, which indicates if the null hypothesis were true, there would only be a probability of 1.8% that we could get the results (the F-value) we got from the F-test.
15. Yes. Since this is an omnibus test, we can't be sure where exactly this difference is coming from, and which group pair are statistically significantly different. Is it that group 1 is different from the other 2? Is it that both groups 1 and 3 are different from group 2? Or other scenarios?
16. Conducting a Tukey post-hoc test, reveals that groups 1 and 2 are the only pair that differ significantly (p= .016).
17. Normality tests (K-S test, Q-Q plot and the histograms) show that group 2 is not normally distributed and is right-skewed. The other groups are normally distributed.









The Levene's test of equality of variances at the other hand, turns out not to be significant, which indicates equality of variances.

Levene's Test of Equality of Error Variances[a]

Dependent Variable: pv_total

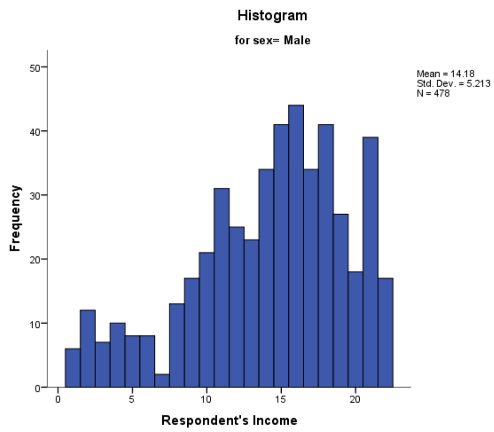
F	df1	df2	Sig.
.601	2	378	.549

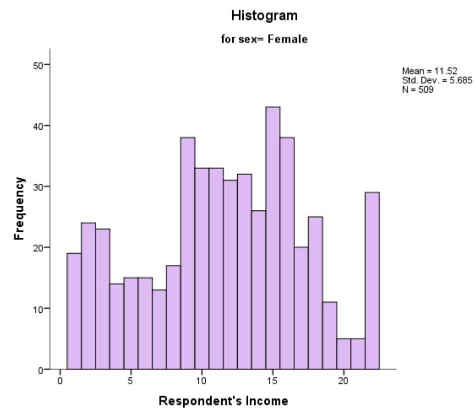
PART 2: Analysis of Variance (No covariate)

1. There are 478 participants in the male group and 509 participants in the female group.

Income in each group [Mean(SD)]: M: 14.18 (5.21) , F: 11.52(5.68)

Neither of the groups is normally distributed. Male group's income is left-skewed, while females' income looks bimodal to me:





2. The income is significantly different between male and female groups.

Tests of Between-Subjects Effects

Dependent Variable: Respondent's Income

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	1735.363[a]	1	1735.363	58.178	.000	.056	58.178	1.000
Intercept	162845.320	1	162845.320	5459.432	.000	.847	5459.432	1.000
sex	1735.363	1	1735.363	58.178	.000	.056	58.178	1.000
Error	29380.828	985	29.828					
Total	193067.000	987						
Corrected Total	31116.190	986						

a. R Squared = .056 (Adjusted R Squared = .055)
b. Computed using alpha = .05

3. A one-way analysis of variance was used to compare the mean income of men (M = 14.18, SD = 5.21 and women (M = 11.52, SD = 5.68. Using an alpha level of 0.05, this test was found to be statistically significant, $F(1, 985) = 58.178, \text{p} < .001, {\eta_p}^2 = 0.056$. The observed power was 1, which indicates type 1 error is very unlikely. The mean income for men was therefore significantly higher than the mean income for women.
4. Normality is violated as explained above, but since samples are large enough can be ignored.

PART 3: Analysis of Covariance

1. $Y_{ij} = \mu_Y + \alpha_j + \beta_w(X_{ij} - \mu_X) + \varepsilon_{ij}$
2. When adjusted for education level, the two groups still show statistically significant differences in terms of income, $F(1, 984) = 63.903, p < .001$.

Tests of Between-Subjects Effects

Dependent Variable: Respondent's Income

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	5341.220[a]	2	2670.610	101.955	.000	.172	203.909	1.000
Intercept	481.889	1	481.889	18.397	.000	.018	18.397	.990
educ	3605.857	1	3605.857	137.659	.000	.123	137.659	1.000
sex	1673.892	1	1673.892	63.903	.000	.061	63.903	1.000
Error	25774.971	984	26.194					
Total	193067.000	987						
Corrected Total	31116.190	986						

a. R Squared = .172 (Adjusted R Squared = .170)
b. Computed using alpha = .05

3. H_0: $\mu'_{Y1} = \mu'_{Y2}$, H_1: $\mu'_{Y1} \neq \mu'_{Y2}$

H_0: $\beta_w = 0$, H_1: $\beta_w \neq 0$

4. No. Having the education included in the analysis, the F-test doesn't change notably. Therefore, it seems to be useless.
5. M: 14.15, F: 11.54. They are only slightly different from the original means (M a little decreased and F a little increased).

Estimates

Dependent Variable: Respondent's Income

Respondent's Sex	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Male	14.153[a]	.234	13.694	14.613
Female	11.547[a]	.227	11.102	11.993

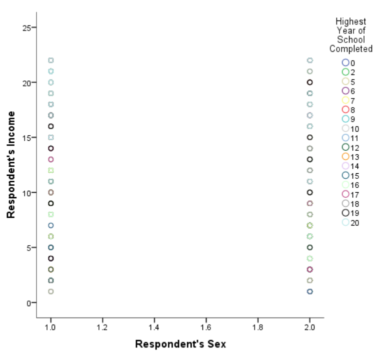
a. Covariates appearing in the model are evaluated at the following values:
Highest Year of School Completed = 13.71.

6. Both from the Coefficinets table (no significant interaction term) and the scatter plot (parallel), it seems that the homogeneity assumption is met.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	16.831	.555		30.317	.000
	Respondent's Sex	-2.653	.348	-.236	-7.627	.000
2	(Constant)	7.401	.957		7.730	.000
	Respondent's Sex	-2.606	.326	-.232	-7.994	.000
	Highest Year of School Completed	.682	.058	.340	11.733	.000
3	(Constant)	8.737	2.497		3.499	.000
	Respondent's Sex	-3.540	1.644	-.315	-2.153	.032
	Highest Year of School Completed	.653	.077	.326	8.478	.000
	sex_edu_int	.068	.118	.086	.579	.563

a. Dependent Variable: Respondent's Income



7. ANOVA model. Because by running the ANCOVA we lose one degree of freedom, which decreases the power, while the results remain statistically unchanged and our exploration revealed the ANCOVA is useless here.

PART 4: Analysis of Covariance through Multiple Regression

1. Intercept is the mean income for the reference (male) group. Unstandardized slope indicates the difference between female group's income and the male group.
2. For the intercept: $H_0: \mu_1 = 0$, For the slope: $H_0: \mu_1 = \mu_2 = 0$. The null is rejected, hence the mean income of females is different than the mean income for men in the population.

3. $R^2 = .056$, which means 5.6% of the income difference in the population can be explained by sex.
4. No. The interaction term is not significant.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	8.737	2.497		3.499	.000
	Respondent's Sex	-3.540	1.644	-.315	-2.153	.032
	Highest Year of School Completed	.653	.077	.326	8.478	.000
	sex_edu_int	.068	.118	.086	.579	.563

a. Dependent Variable: Respondent's Income

5. The intercept indicates the adjusted (for education level) mean of income in men (7.4). The slope for sex indicates the difference between the adjusted mean of income in the women group and the men group (-2.6).
6. Model: $Y' = 8.737 - 3.54(sex) + .653(education) + .068(sex * education)$

   6 years: $Y' = 8.737 - 3.54(1.52) + .653(6) + .068(7.0547) = 7.75$

   20 years: $Y' = 8.737 - 3.54(1.52) + .653(20) + .068(7.0547) = 16.89$

Descriptive Statistics

	Mean	Std. Deviation	N
Respondent's Income	12.81	5.618	987
Respondent's Sex	1.52	.500	987
Highest Year of School Completed	13.71	2.802	987
sex_edu_int	7.0547	7.08232	987

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	8.737	2.497		3.499	.000
	Respondent's Sex	-3.540	1.644	-.315	-2.153	.032
	Highest Year of School Completed	.653	.077	.326	8.478	.000
	sex_edu_int	.068	.118	.086	.579	.563

a. Dependent Variable: Respondent's Income