Abstract

The Module 2 Case assignment will create dummy codes for categorical predictor variables and. check the assumptions of normality, homoscedasticity, and collinearity. It will also run multiple regression using three different methods including forced entry, stepwise, and hierarchical analysis. The purpose of this case assignment is to identify predictors for reading comprehension among children.

Part 1: This section will create dummy codes for categorical predictor variables.

Dummy Coding with Two Levels

The graphic table below indicates the dummy codes for categorical predictors of variables. These variables are dummy coded into two variables: 1: **LS visual** – where "1" indicates a visual learning style and "0" indicates not a visual learning style; and 2: **LS auditory** – where "1" indicates an auditory learning style and "0" indicates not an auditory learning style. The dummy coding is represented below:

Dummy Coded Variables

1. **LS Visual:** 0 = Not a visual learning style

    1 = visual learning style 1

2. **LS Auditory:** 0 = not an auditory learning style

    1 = auditory learning style

Frequency Tables 1:

The frequency table below indicates statistics of variables visual and auditory. As provided in the table, there are a total of 332 cases and non are missing. The mean for visual is .17 and for auditory is .42. The learning style value of 00 cuts off the 75 percentile (75% of cases fall at or below this value.

Statistics

		Visual	Auditory
N	Valid	332	332
	Missing	0	0
Mean		.17	.42
Std. Deviation		.372	.495
Percentile	25	.00	.00
	75	.00	.00

Frequency Tables 2:

The frequency table below lists the values of the variable visual and the frequency of occurrence of each.

Visual

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Not a visual learning style	277	83.4	83.4	83.4
	Visual Learning Style	55	16.6	16.6	100.0
	Total	332	100.0	100.0	

Frequency Tables 3:

The frequency table below includes the lists of values for variable auditory

Auditory

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Not auditory Learning Style	192	57.8	57.8	57.8
	Auditory Learning Style	140	42.2	42.2	100.0
	Total	332	100.0	100.0	

Part 2: This section will check the assumptions of normality, homoscedasticity, and colinearity. It will also describe and provide support regarding whether the assumptions were met (Include supporting tables/graphs).

Graph 1: Histogram SPSS Results:

Reporting result: Based on the details provided below, the assumption shows normal distribution as supported by the graph. The histogram shows some possible outliers.

Histogram

Dependent Variable: Reading

Mean = 2.08E-16
Std. Dev. = 0.990
N = 313
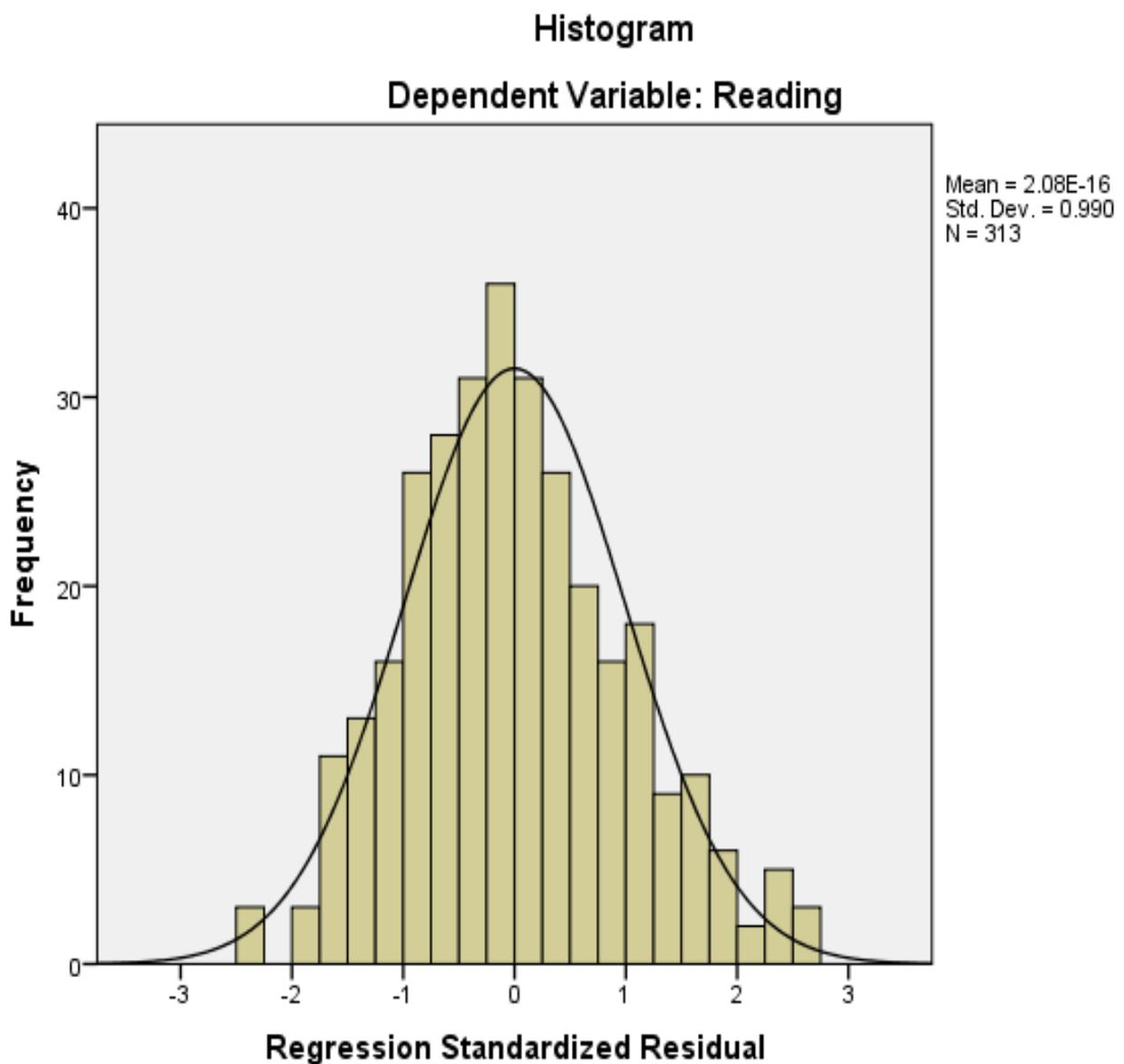
SPSS Results – P-P Plot:

Reporting result: According to the linear regression analysis, the assumptions indicate that the residuals are normally distributed. It is important to meet this assumption for the p-values for the t-tests to be valid.

Normal P-P Plot of Regression Standardized Residual

Dependent Variable: Reading
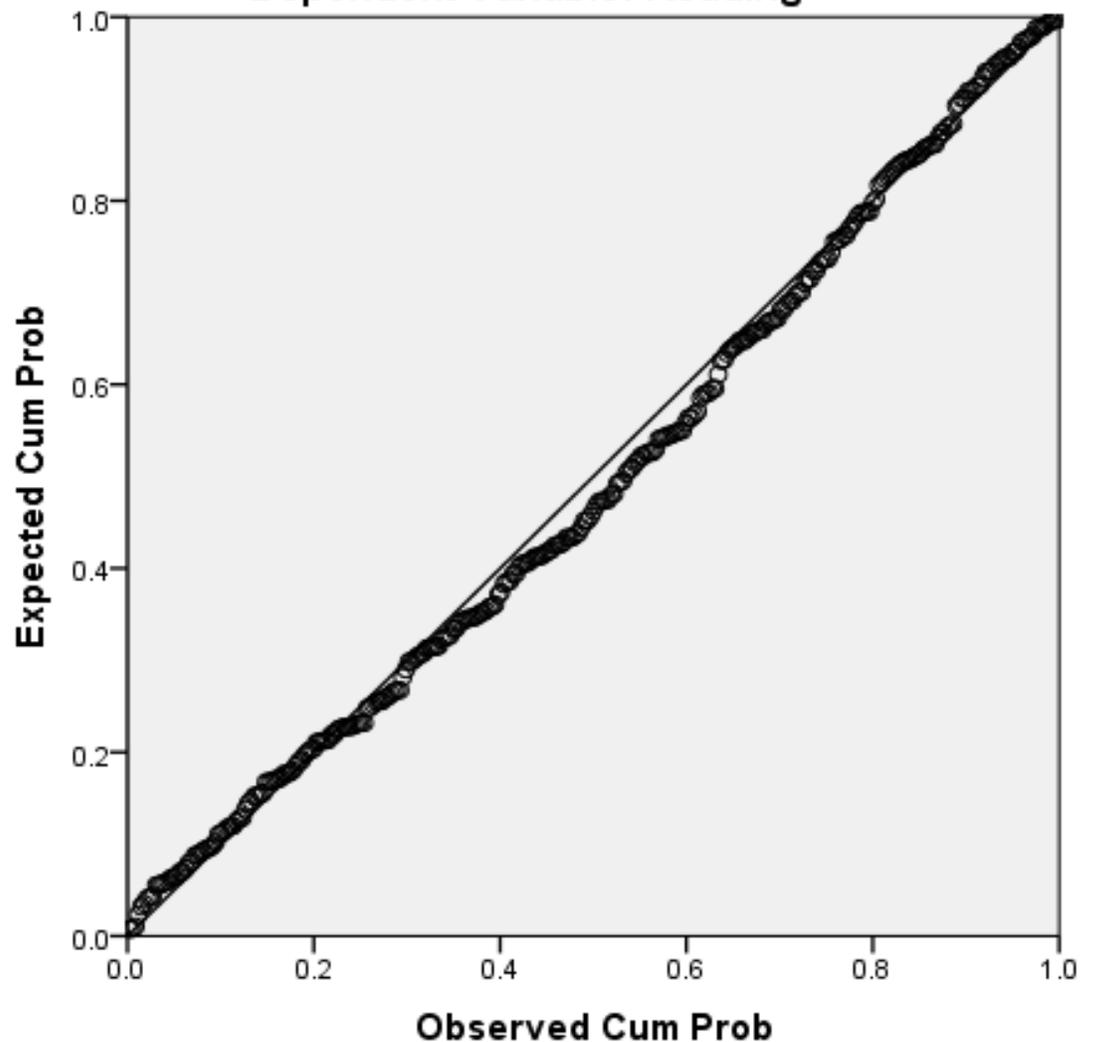
SPSS Results – Scatter plot

Reporting result: All the scatter plots suggest that the observation indicates no extra attention since all the points stand parallel to one another. The assumption of homoscedasticity indicates

that the residuals are approximately equal for all predicted dependent variable.

Scatterplot
Dependent Variable: Reading
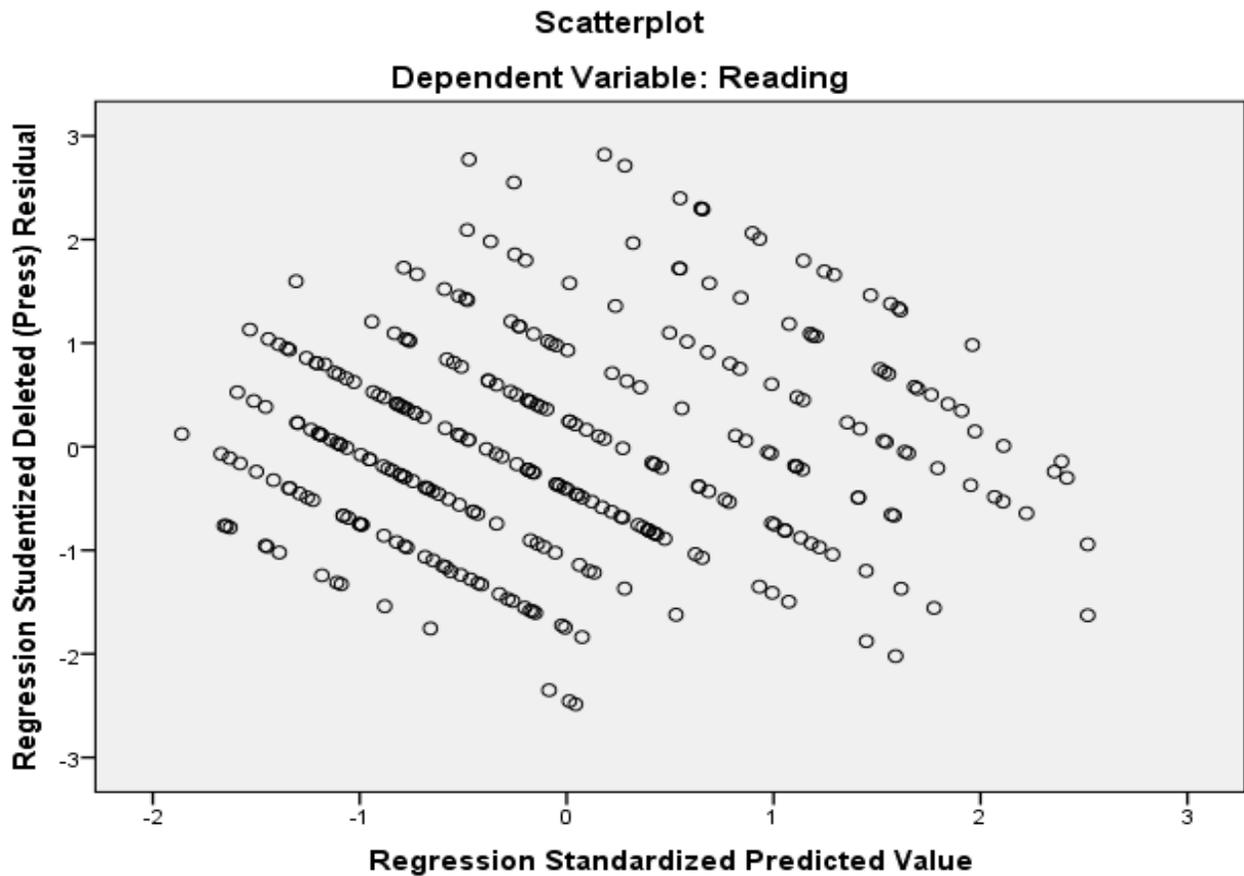
Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.710[a]	.503	.494	.18845

a. Predictors: (Constant), Learning Style, morpheme, id, Gender, visual, phoneme

b. Dependent Variable: Reading

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	11.019	6	1.837	51.716	.000[b]
	Residual	10.867	306	.036		
	Total	21.886	312			

a. Dependent Variable: Reading

b. Predictors: (Constant), Learning Style, morpheme, id, Gender, visual, phoneme

Reporting result: The regression model is statistically significant $F_{(6, 306)} =$ 51.716, p = .000, p<0.01

SPSS Results – Collinearity

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-.518	.124		-4.191	.000		
	Id	.000	.000	.197	4.669	.000	.910	1.099
	Phoneme	.649	.060	.510	10.891	.000	.740	1.352
	Visual	.174	.050	.162	3.457	.001	.740	1.351
	Morpheme	.089	.052	.073	1.723	.086	.905	1.105
	gender	-.024	.021	.021	-1.098	.273	.985	1.015
	Learning Style	-.009	.015	-.025	-.603	.547	.981	1.020

Reporting result: As indicated above, none of the independent variables is statistically significant. The VIF is above 5, which means that multicollinearity inflated the standard errors which lower the test below 2, which means that the significance level becomes above 0.05.

Part 3: This section will run multiple regression using three different methods:

1. **Forced Entry multiple regression:**

Table 1: SPSS Results

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.658[a]	.433	.424	.19941

a. Predictors: (Constant), Learning Style, morpheme, Gender, phoneme, Auditory

Table 2: SPSS Results

ANOVA[a]

Model	Sum of Squares	df	Mean Square	F	Sig.

1	Regression	9.496	5		1.899	47.760	.000[b]
	Residual	12.446	313		.040		
	Total	21.942	318				

a. Dependent Variable: Reading

b. Predictors: (Constant), Learning Style, morpheme, Gender, phoneme, Auditory

Table 3: SPSS Results

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.064	.063		1.015	.311
	Auditory	-.032	.024	-.060	-1.336	.183
	phoneme	.781	.056	.621	14.033	.000
	morpheme	.120	.054	.099	2.228	.027
	Gender	-.017	.023	-.032	-.748	.455
	Learning Style	-.014	.016	-.039	-.874	.383

a. Dependent Variable: Reading

The Forced Entry multiple regressions overall reporting results: The Forced Entry multiple regressions first tables above reports that the model accounted for 42.4% of the variance. The multiple R for the relationship between independent and dependent variables is 0.658. The overall relationship between the set of variables would is characterized as strong using the rule of thumb. The second table regression model is statistically significant F (5, 313) = 47.760, p= 00, p<0.01. We reject the null hypothesis that there is no relationship between the set of variables. There is a statistical significant relationship between the set of independent variables and the dependent variables. The third table regression equation for this model is Y = .064 − .032 Auditory + .781 Phoneme + .120 Morpheme − .017 Gender − .014 Learning style. The phoneme awareness: **βeta=.099, t=14.033, p<.01,** was the most influential predictor, followed by morpheme **βeta=.621, t=2.228, p<.05.**

2. Stepwise Multiple regressions:

Table 1: SPSS Results

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df 1	df2	Sig. F Change
1	.647[a]	.419	.417	.20052	.419	228.705	1	317	.000
2	.654[b]	.428	.424	.19936	.009	4.710	1	316	.031

. Predictors: (Constant), phoneme
. Predictors: (Constant), phoneme, morpheme

Table 2: SPSS Results

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9.196	1	9.196	228.705	.000[b]
	Residual	12.746	317	.040		
	Total	21.942	318			
2	Regression	9.383	2	4.691	118.046	.000[c]
	Residual	12.559	316	.040		
	Total	21.942	318			

a. Dependent Variable: Reading

b. Predictors: (Constant), phoneme

c. Predictors: (Constant), phoneme, morpheme

Table 3: SPSS Results

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.038	.030		1.295	.196
	phoneme	.815	.054	.647	15.123	.000
2	(Constant)	-.005	.035		-.137	.891
	phoneme	.782	.056	.622	14.073	.000
	morpheme	.116	.054	.096	2.170	.031

a. Dependent Variable: Reading

The Stepwise Multiple regressions reporting result: The Stepwise Multiple regressions first table above indicates that the phoneme alone accounts for 41.7% of the variance, while phoneme and morpheme accounted for 42.4% of the variance. The multiple R for the relationship between the subset of independent variables that predict the dependent variables are .647 and 0.654, which would be characterized as moderate using the rule of thumb. The stepwise regression second table indicates a significant model emerged that contained two variables **F (1, 317) = 228.705, p= 00, p<0.01** and **F (2, 316) = 118.046, p = .000, p < 0.01),** less than or equal to the level of significance of 0.05. We reject the null hypothesis that there is no relationship between the best subset of independent variables and the dependent variable. We support the research hypothesis that there is a statistical significance relationship between the set of independent variables and the dependent variable. The third table regression equation indicates that this model is **Y = - .005 + .782 Phoneme + .116 Morpheme.** The phoneme awareness: **βeta=.622, 14.073, p<0.01,** was the most influential predictor, followed by morpheme **βeta=.621, t=2.228, p=0.031, p<.05.**

3. **Hierarchical Multiple regression**

    Table 1: SPSS Results

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df 1	df2	Sig. F Change
					Change Statistics				
1	.658[a]	.433	.424	.19941	.433	47.760	5	313	.000

Predictors: (Constant), Learning Style, morpheme, Gender, phoneme, Auditory

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9.496	5	1.899	47.760	.000[b]
	Residual	12.446	313	.040		
	Total	21.942	318			

a. Dependent Variable: Reading

b.   Predictors: (Constant), Learning Style, morpheme, Gender, phoneme, Auditory

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.064	.063		1.015	.311
	Gender	-.017	.023	-.032	-.748	.455
	Auditory	-.032	.024	-.060	-1.336	.183
	phoneme	.781	.056	.621	14.033	.000
	morpheme	.120	.054	.099	2.228	.027
	Learning Style	-.014	.016	-.039	-.874	.383

a.   Dependent Variable: Reading

The Hierarchical multiple regression reporting results: The hierarchical multiple regressions

first table above indicates that the model accounted for 42.4% of the variance explained after the

influence of school, learning style, morpheme, gender, phoneme, auditory and word is removed.

The multiple R for the relationship between the subset of independent variables that predict the

dependent variables is .658, which is characterized as moderate using the rule of thumb. The

second table regression model is statistically significant **F (5, 303) = 47.760, p= .000, p<0.01**.

We reject the null hypothesis that there is no relationship between the set of variables. There is a

statistical significant relationship between the set of independent variables and the dependent

variables. The third table regression equation indicates that this model is **Y = .064 – .017 Gender

– .032 Auditory + .781 Phoneme + .120 Morpheme – .014 Learning style.** The phoneme

awareness: **βeta=.621, 14.073, p<0.01,** was the most influential predictor, followed by

morpheme **βeta=.099, t=2.228, p=0.027, p<.05.** For the independent variable, the probability of t

statistic is (1.015) for the b coefficient is .311, which is greater than the level of significance of

0.05. We do not reject the null hypothesis. Therefore, we conclude that this no statistical

significant relationship between independent variables and dependent variable.