Multiple Regression Model

A study was carried out to explore the relationship between Aggression and several potential predicting factors in 666 children who had an older sibling. Variables measured were Parenting Style (high score = bad parenting practices), Computer Games (high score = more time spent playing computer games), Television (high score = more time spent watching television), Diet (high score = the child has a good diet low in harmful additives), and Sibling Aggression (high score = more aggression seen in their older sibling). Past research indicated that parenting style and sibling aggression were good predictors of the level of aggression in the younger child. All other variables were treated in an exploratory fashion.

1. **State the underlying assumptions for the statistical test.**

**Linearity** - the relationships between the predictors and the outcome variable should be linear

**Normality** - the residuals should be normally distributed

**Homogeneity of variance (homoscedasticity)** - the residual variance should be constant

**Independence** - the residuals associated with one observation are not correlated with the errors of any other observation

1. **State whether the assumptions have been met**. If the assumptions were not met (either in actuality or hypothetically), state what alternatives you have available to you.

From the correlation matrix it is clear that there is a significant correlation existing between parenting style and television. Hence there might be a chance of multicollinearity. Possible alternatives are:

If appropriate combine these two variables or Remove one of the variables from the model.

From the histogram it is clear that the distribution of residuals is approximately normally distributed. Hence the normality assumption is satisfied.

Since there is no specific pattern for the points, the points on the plot of residuals against the fitted value are at random. Hence we can conclude that the errors are independent and the residual variances are constant. Hence the homogeneity of variance and independence assumption are also satisfied.

1. **State the null and alternative (research) hypotheses.**

**Null hypothesis:**

H0: The estimated regression model is not significant in predicting the child aggression.

**Alternative hypothesis:**

Ha: The estimated regression model is significant in predicting the child aggression.

**This is fine but really this needs to be more overarching—no differences among the variables**

1. **Copy your syntax file and paste it into your MS Word Document.**

**Syntax:**

GET

FILE='C:\Users\Dell\AppData\Local\Temp\Child Aggression-1.sav'.

DATASET NAME DataSet1 WINDOW=FRONT.

REGRESSION

/MISSING LISTWISE

/STATISTICS COEFF OUTS R ANOVA

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT Aggression

/METHOD=ENTER Television Computer\_Games Sibling\_Aggression Diet Parenting\_Style

/SCATTERPLOT=(\*ZRESID ,\*ZPRED)

/RESIDUALS HISTOGRAM (ZRESID) NORMPROB (ZRESID)

/SAVE PRED ZPRED ADJPRED MAHAL COOK LEVER ZRESID DRESID SDRESID SDBETA SDFIT COVRATIO.

1. **For your output file: Select all 🡪 Copy all objects 🡪 Paste into your MS word document.**

| **Variables Entered/Removedb** | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | | Variables Entered | | | Variables Removed | | | | Method | | | |
| 1 | | Parenting Style, Sibling Aggression, Use of Computer Games., Good Diet, Time spent watching television. | | | . | | | | Enter | | | |
| a. All requested variables entered.  b. Dependent Variable: Aggression | | | | | | | | | | | | |
| **Model Summaryb** | | | | | | | | | | | | | |
| Model | | R | R Square | | | Adjusted R Square | | | | Std. Error of the Estimate | | | |
| 1 | | .287a | .083 | | | .076 | | | | .30709 | | | |
| a. Predictors: (Constant), Parenting Style, Sibling Aggression, Use of Computer Games., Good Diet, Time spent watching television.  b. Dependent Variable: Aggression | | | | | | | | | | | | | |
| **ANOVAb** | | | | | | | | | | | | | | | | | | |
| Model | | | | Sum of Squares | | | | df | | | Mean Square | | | | | F | Sig. | |
| 1 | Regression | | | 5.602 | | | | 5 | | | 1.120 | | | | | 11.882 | .000a | |
| Residual | | | 62.240 | | | | 660 | | | .094 | | | | |  |  | |
| Total | | | 67.842 | | | | 665 | | |  | | | | |  |  | |
| a. Predictors: (Constant), Parenting Style, Sibling Aggression, Use of Computer Games., Good Diet, Time spent watching television.  b. Dependent Variable: Aggression | | | | | | | | | | | | | | | | | | |
| **Coefficientsa** | | | | | | | | | | | | | | | | | | | | |
| Model | | | | | | | Unstandardized Coefficients | | | | | | | | Standardized Coefficients | | | t | | Sig. |
| B | | | | | Std. Error | | | Beta | | |
| 1 | (Constant) | | | | | | -.005 | | | | | .012 | | |  | | | -.416 | | .677 |
| Time spent watching television. | | | | | | .033 | | | | | .046 | | | .032 | | | .715 | | .475 |
| Use of Computer Games. | | | | | | .142 | | | | | .037 | | | .152 | | | 3.851 | | .000 |
| Sibling Aggression | | | | | | .082 | | | | | .039 | | | .084 | | | 2.106 | | .036 |
| Good Diet | | | | | | -.109 | | | | | .038 | | | -.115 | | | -2.864 | | .004 |
| Parenting Style | | | | | | .057 | | | | | .015 | | | .177 | | | 3.891 | | .000 |
| a. Dependent Variable: Aggression | | | | | | | | | | | | | | | | | | | | |









| **Correlations** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | Aggression | Time spent watching television. | Use of Computer Games. | Sibling Aggression | Good Diet | Parenting Style |
| Aggression | Pearson Correlation | 1 | .159\*\* | .186\*\* | .129\*\* | -.009 | .211\*\* |
| Sig. (2-tailed) |  | .000 | .000 | .001 | .825 | .000 |
| N | 666 | 666 | 666 | 666 | 666 | 666 |
| Time spent watching television. | Pearson Correlation | .159\*\* | 1 | .166\*\* | .228\*\* | .096\* | .526\*\* |
| Sig. (2-tailed) | .000 |  | .000 | .000 | .013 | .000 |
| N | 666 | 666 | 666 | 666 | 666 | 666 |
| Use of Computer Games. | Pearson Correlation | .186\*\* | .166\*\* | 1 | .225\*\* | .242\*\* | .215\*\* |
| Sig. (2-tailed) | .000 | .000 |  | .000 | .000 | .000 |
| N | 666 | 666 | 666 | 666 | 666 | 666 |
| Sibling Aggression | Pearson Correlation | .129\*\* | .228\*\* | .225\*\* | 1 | .231\*\* | .174\*\* |
| Sig. (2-tailed) | .001 | .000 | .000 |  | .000 | .000 |
| N | 666 | 666 | 666 | 666 | 666 | 666 |
| Good Diet | Pearson Correlation | -.009 | .096\* | .242\*\* | .231\*\* | 1 | .266\*\* |
| Sig. (2-tailed) | .825 | .013 | .000 | .000 |  | .000 |
| N | 666 | 666 | 666 | 666 | 666 | 666 |
| Parenting Style | Pearson Correlation | .211\*\* | .526\*\* | .215\*\* | .174\*\* | .266\*\* | 1 |
| Sig. (2-tailed) | .000 | .000 | .000 | .000 | .000 |  |
| N | 666 | 666 | 666 | 666 | 666 | 666 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed).  \*. Correlation is significant at the 0.05 level (2-tailed). | | | | | | | |

1. For Multiple and Logistic Regression models, ensure that you include appropriate measures of model fit as well as the specific procedure used (e.g., Hierarchical, Enter, Stepwise, Forward, Backward).

The estimated regression equation is given by,

Aggression = -0.005 + 0.033 \* Television + 0.142 \* Computer Games + 0.082 \* Sibling aggression – 0.109 \* Good Diet + 0.057 \* Parenting Style

The regression coefficients can be interpreted as follows:

**You were only supposed to consider sibling aggression and parenting style—see assignment 1 explanation on blackboard**

For a unit increase in television hours, the child aggression increases by 0.033 units.

For a unit increase in computer game hours, the child aggression increases by 0.142 units.

For a unit increase in Sibling aggression, the child aggression increases by 0.082 units.

For a unit increase in Good Diet, the child aggression decreases by 0.109 units.

For a unit increase in Parenting Style, the child aggression increases by 0.057 units.

The significance of regression coefficients are tested using Student’s t test.

Computer games, Sibling aggression, Good Diet and Parenting Style are significant in predicting the child aggression at the 0.05 significance level, since the respective p-values are less than 0.05.

Television is not significant in predicting child aggression, since the corresponding p-value is greater than 0.05.

The significance of the regression model is tested using F-test.

Here the F-statistic = 11.882

P-value = 0.000

Hence we can conclude that the regression model is significant in predicting the child aggression, since the F-statistic is significant with p-value less than 0.05.

The model adequacy is measured using the R2 value. Here R2 = 0.083. Thus only 8.3% variability in the child regression can be explained by the regression model.

Mediating vs. moderating variable?

Reference:

Lewis-Beck MS, (1993). Regression Analysis, Beverley Hills, CA: Sage.

Wayne DW, (1995). Biostatistics, 6th ed, New York: John Wiley & Sons.

**Not APA I KNOW I SHOULD FIX THAT**

**Grading rubric:**

**Assumptions (5): 5**

**Assumptions met? (5): 5**

**Null and Alt Hyp (5): 5**

**Syntax (5): 3—didn’t follow assignment description to only include two variables**

**Output (5): 3—didn’t follow assignment**

**Results (10): 5—didn’t follow assignment**

**APA (10): 5—references not in APA—no citations**

**Power and effect size (5): 0—not mentioned**

**Total (50): 31**