

Results

ANOVA

Introduction

An analysis of variance (ANOVA) was conducted to determine whether there were significant differences in MathScores by FavoriteColor.

Assumptions

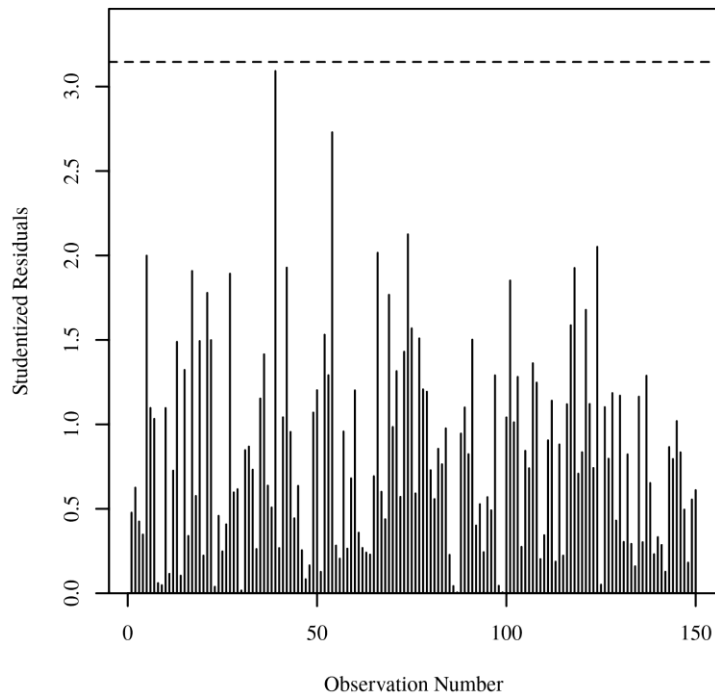
Normality. A Shapiro-Wilk test was conducted to determine whether the model residuals could have been produced by a normal distribution (Razali & Wah, 2011). The results of the Shapiro-Wilk test were not significant based on an alpha value of .05, $W = 1.00$, $p = .944$. This result suggests the possibility that the residuals of the model were produced by a normal distribution cannot be ruled out, indicating the normality assumption is met.

Homogeneity of Variance. A Levene's test was conducted to determine whether the model residuals have similar variances between the groups of the independent variables. The results of the Levene's test were not significant based on an alpha value of .05, $F(2, 147) = 0.11$, $p = .892$. This result suggests the residuals of the model have similar variances, indicating the homogeneity of variance assumption is met.

Outliers. To identify influential points, Studentized residuals were calculated and the absolute values were plotted against the observation numbers (Field, 2017; Pituch & Stevens, 2015). Studentized residuals are calculated by dividing the model residuals by the estimated residual standard deviation. An observation with a Studentized residual greater than 3.15 in absolute value, the 0.999 quantile of a t distribution with 149 degrees of freedom, was considered to have significant influence on the results of the model. Figure 1 presents the Studentized residuals plot of the observations. Observation numbers are specified next to each point with a Studentized residual greater than 3.15.

Figure 1

Studentized residuals plot for outlier detection



Results

The ANOVA was examined based on an alpha value of .05. The results of the ANOVA were significant, $F(2, 147) = 7.63, p < .001$, indicating there were significant differences in MathScores among the levels of FavoriteColor (Table 1). The eta squared was 0.09 indicating FavoriteColor explains approximately 9% of the variance in MathScores. The means and standard deviations are presented in Table 2.

Table 1

Analysis of Variance Table for MathScores by FavoriteColor

Term	SS	df	F	p	η_p^2
FavoriteColor	809.97	2	7.63	< .001	0.09
Residuals	7,804.14	147			

Figure 2

Means of MathScores by FavoriteColor with 95.00% CI Error Bars

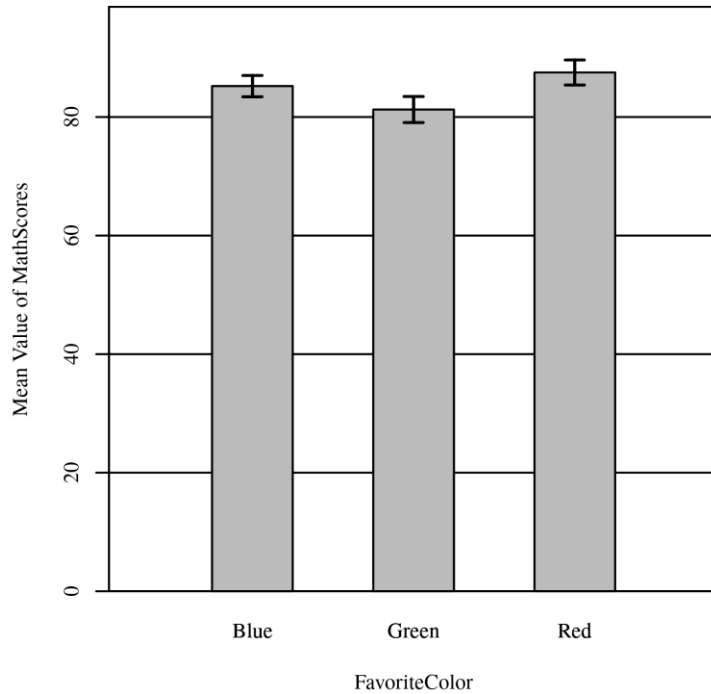


Table 2

Mean, Standard Deviation, and Sample Size for MathScores by FavoriteColor

Combination	<i>M</i>	<i>SD</i>	<i>n</i>
Blue	85.21	7.57	68
Green	81.26	6.91	38
Red	87.51	7.15	44

Note. A '-' indicates the sample size was too small for the statistic to be calculated.

Post-hoc

A *t*-test was calculated between each group combination to further examine the differences among the variables based on an alpha of .05. The Tukey HSD *p*-value adjustment was used to correct for the effect of multiple comparisons on the family-wise error rate. For the main effect of FavoriteColor, the mean of MathScores for Blue ($M = 85.21$, $SD = 7.57$) was significantly larger than for Green ($M = 81.26$, $SD = 6.91$), $p = .022$. For the main effect of

FavoriteColor, the mean of MathScores for Green ($M = 81.26$, $SD = 6.91$) was significantly smaller than for Red ($M = 87.51$, $SD = 7.15$), $p < .001$. No other significant effects were found.

References

- Field, A. (2017). *Discovering statistics using IBM SPSS statistics: North American edition*. Sage Publications
- Intellectus Statistics [Online computer software]. (2023). Intellectus Statistics.
<https://statistics.intellectus360.com>
- Pituch, K. A., & Stevens, J. P. (2015). *Applied multivariate statistics for the social sciences* (6th ed.). Routledge Academic. <https://doi.org/10.4324/9781315814919>
- Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21-33.

Glossaries

ANOVA (Analysis of Variance)

The ANOVA is used to assess differences in a dependent variable by the given independent variable(s). If the independent variable(s) have more than two levels, and statistical significance is found in the ANOVA, pairwise comparisons (called post-hoc tests) are conducted to determine the paired differences.

Degrees of Freedom (df): Used with the F to determine the p -value. The two df s are derived from the number of groups and sample size.

F Ratio (F): Used with the two df values to determine the p value, calculated by dividing the between subjects MS by the residuals MS .

Mean Square (MS): Used to determine the F ratio; calculated by dividing the SS by df .

Normality: Refers to the distribution of the data. The assumption is that the data follows the bell-shaped curve. If the data are not normally distributed, the results may not be reliable.

Outlier: A data point that is abnormally distant from a set of observations.

p -value: The probability that the null hypothesis - no difference in the dependent variable by the independent variable - is true.

Residuals: Refers to the difference between the predicted value for the dependent variable and the actual value of the dependent variable.

Studentized Residuals: Residuals that are scaled by dividing the each residual by the estimated standard deviation of the residuals.

Sum of Squares (SS): Used with the df to determine the MS .

Raw Output

Analysis of Variance Table for MathScores by FavoriteColor

Included Variables:

MathScores and FavoriteColor

Sample Size (Complete Cases):

N = 150

Shapiro-Wilk Test:

W = 0.996, p = 0.944

Levene's Test:

F(2, 147) = 0.115, p = 0.892

ANOVA Results:

Term	SS	df	F	p	η_p^2
FavoriteColor	809.975	2	7.628	7.045×10^{-04}	0.0940
Residuals	7,804.137	147			

Means Table:

Combination	M	SD	n
Blue	85.212	7.567	68
Green	81.261	6.914	38
Red	87.512	7.152	44

Note. A '-' indicates the sample size was too small for the statistic to be calculated.

Post-hoc Comparisons with Tukey's Honest Significant Difference Test:

Comparison	M	Lwr Limit	Upr Limit	p
Green-Blue	-3.951	-7.445	-0.457	0.0224
Red-Blue	2.300	-1.037	5.638	0.236
Red-Green	6.251	2.431	10.071	4.698×10^{-04}

Note. M is calculated on the differences between the groups in each comparison.

Upper and lower limits for the means are calculated using a 95% confidence interval.