For this milestone assignment, focus on your first research question pertaining to identifying **differences between the store locations and days of the week**:

1. **Provide the names of all group members.**

Miza Syafiqah Mohamad Shanudin and Phichchaya Sutaporn

2. **State the research question.**

How do differences between the store locations and days of the week affect the average units sold for three subway stores in downtown state college?

How do football and none football days affect the average units sold on Friday and Saturday for three subway stores in downtown state college?

3. **Using the data set your client provided, run an ANOVA, including the main effects and two-way interaction terms, to address the research question. Provide your ANOVA table.**

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Store	2	352228	176114	33.80	0.000
DayOfWk	6	675455	112576	21.61	0.000
Football	1	273461	273461	52.49	0.000
Store*DayOfWk	12	823863	68655	13.18	0.000
Store*Football	2	8586	4293	0.82	0.439
Error	414	2156868	5210		
Lack-of-Fit	3	1982	661	0.13	0.945
Pure Error	411	2154886	5243		
Total	437	6330195			

Table 1: Analysis of Variance table for the full model.

4. **If necessary, reduce the model. Provide your ANOVA table for the reduced model. Clearly describe what model reduction technique you used (e.g. backward elimination with an 0.05 alpha level, or AIC, etc.).**

We get our reduced model using the backward elimination technique with an 0.05 alpha level.

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Store	2	2083833	1041916	200.16	0.000
DayOfWk	6	675455	112576	21.63	0.000
Football	1	273461	273461	52.53	0.000
Store*DayOfWk	12	928832	77403	14.87	0.000
Error	416	2165454	5205		
Lack-of-Fit	5	10568	2114	0.40	0.847
Pure Error	411	2154886	5243		
Total	437	6330195			

Table 2: Analysis of Variance table for the reduced model.

5. **Create residual plots for your final, reduced model. Are the assumptions satisfied? Are there any outliers (e.g. standard residuals beyond ±2)? If there are outliers, what dates do they correspond to?**

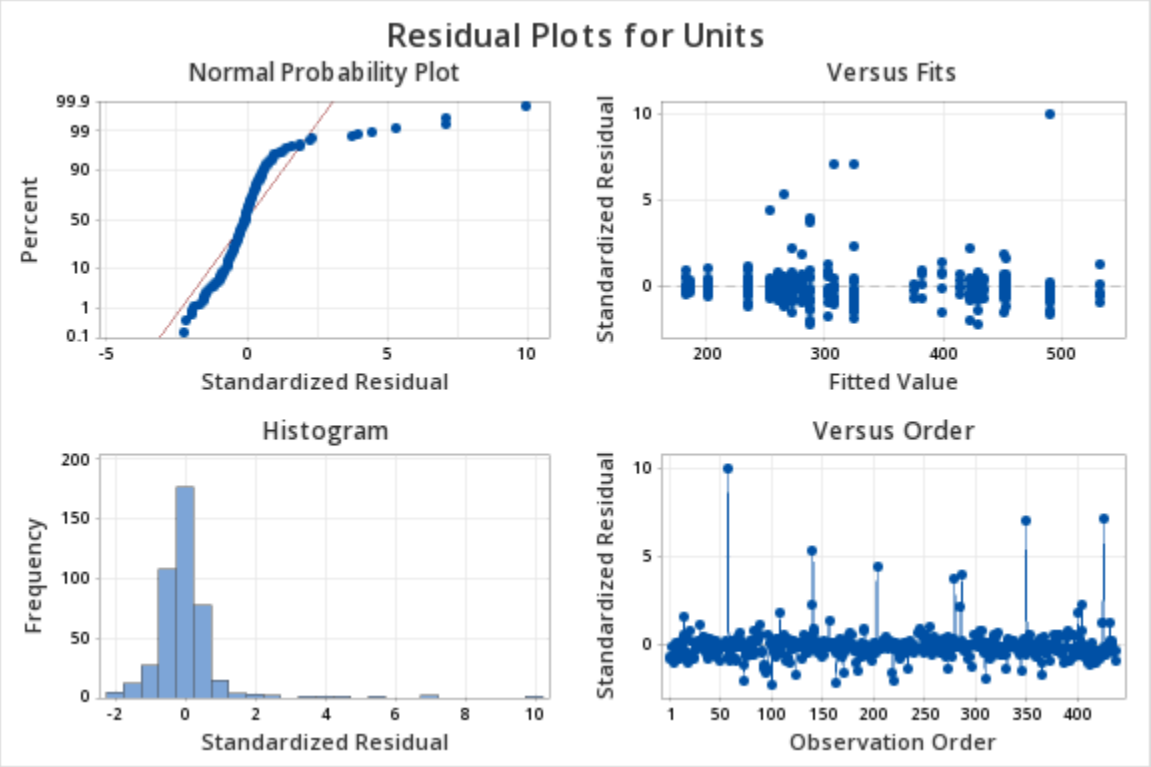

Figure1: Residual Plots for Units of the reduced model

The data is not normally distributed because there are outliers in the normal probability plot; the points do not follow the red line. Looking at the histogram, the graph is skewed to the right justifying that the data is not normally distributed. Looking at the residual vs fit plot, there is no clear cone pattern meaning that the assumption of equal variance is met. However, there are 12 outliers whose standard residuals are beyond +2. However, 3 of them have the value of standardized residuals really closed to +- 2 therefore we will only focus on the other 9 outliers. These outliers will be omitted during the study as they does not reflect the number of units sold on a regular basis. Below is the table showing the list of those outliers:

Date	Day of week	Location	Event
2/18/2017	Saturday	Pugh street	THON
2/25/2017	Saturday	Pugh street	St Patty's
11/3/2016	Thursday	Pugh street	National Sandwhich Day
2/25/2017	Saturday	Burrows street	St Patty's
11/3/2016	Thursday	East College Ave.	National Sandwhich Day
2/18/2017	Saturday	East College Ave.	THON
11/3/2016	Thursday	Burrows street	National Sandwhich Day
2/24/2017	Friday	Burrows street	St Patty's
2/24/2017	Friday	Pugh street	St Patty's

Table 3: Outliers table showing dates, days of week and locations they correspond to.