Mert Dogan
Z1910844

1. A commonly discussed issue is why people are successful (for example, http://freakonomics.com/podcast/sports-ep-3/. One perspective is that success depends on effort/hard work, whereas another perspective is that luck has a major part in determining success. What can you add to the conversation, based on the topic discussed in this module?

Some athletes emphasize the value of hard work, while others highlight the role of luck. Most seem to agree that both hard work and luck impact success as separate independent factors. We can add depth to the conversation by introducing moderation analysis and interaction effects and how they differ from separately accounting luck and hard work. Moderation analysis examines whether a third variable influences the relationship between two variables. To do this, we can state the dependent variable, success, and select either hard work or luck as an independent variable while designating the other as the moderator. If we were to select hard work as the independent variable, we can then test whether the impact of hard work on success changes depending on the level of luck. If we can determine that there is an interaction effect, we can then look at the coefficient and if positive, we can conclude that an increase in hard work and luck have synergistic effects on increasing success. In other words, the level of luck influences how much hard work contributes to success. Likewise, differences in luck can lead to different levels of success for athletes who put in the same amount of hard work.

2. Use the Cars dataset from SASHELP directory to explore if Drive Train impacts the relationship between MSRP(y) and Horsepower (x). Conduct a thorough analysis and state your conclusions.

In order to test the interaction effect between Horsepower and Drive Train on MSRP, we need to conduct linear regression analysis. By running the analysis, we can see that the model is significant (Appendix A.1). Parameter estimates table (Appendix A.2) gives us the effect of the interaction term between horsepower and drive train. There are three different drivetrains (all, front, and rear), and we can observe that the comparison group is the drivetrain rear. As compared to the drivetrain rear, drivetrain all and front have a significant interaction effect with p-values less than <.0001. Therefore, we can conclude that there is an interaction effect between horsepower and drivetrain on MSRP. Thus, Drive Train impacts the relationship between MSRP and Horspower. We can observe negative coefficients on both all and front drivetrains which suggests as horsepower increases, all and front drivetrain MSRP tends to decrease more compared to the rear drivetrain.

3. Use the dataset named “FinalQ2” posted in the datasets folder to solve this question. Examine if Grade category and total amount interact with each other to impact VAR7 Project Approval (0 = not approved, 1 = approved).

In order to examine whether grade and total amount interact with each other to impact VAR7, we need to conduct binary logistic regression. By running the analysis, we can see that the model is significant by observing that the p-values of Likelihood Ratio, Score, and Wald are all less than <.0001 (Appendix B.1). By looking at the Analysis of Effect table (Appendix B.2) we can see that the p-value of the interaction effect term, between Grade category and total amount is significant with a p-value of 0.0432. Therefore, we can conclude that there is a significant interaction effect between grade category and total amount on whether the project gets approved or not.
Also, the Analysis of Maximum Likelihood Estimates table (Appendix B.3) shows that Elementary and High Grades do not individually show a significant interaction effect when compared to the Middle grade.

APPENDIX A

A.1

Least Squares Model (No Selection)

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	1.169209E11	23384186059	222.70	<.0001
Error	422	44310688407	105001631		
Corrected Total	427	1.612316E11			

A.2

Parameter Estimates					
Parameter	DF	Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-27020	3413.380075	-7.92	<.0001
Horsepower	1	278.450623	12.455849	22.36	<.0001
DriveTrain All	1	21090	5664.556531	3.72	0.0002
DriveTrain Front	1	21787	4120.815561	5.29	<.0001
DriveTrain Rear	0	0	.	.	.
Horsepowe*DriveTrain All	1	-98.044295	22.455315	-4.37	<.0001
Horsepowe*DriveTrain Front	1	-116.505251	17.227412	-6.76	<.0001
Horsepowe*DriveTrain Rear	0	0	.	.	.

APPENDIX B

B.1

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	82.8416	5	<.0001
Score	90.9145	5	<.0001
Wald	87.8670	5	<.0001

B.2

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
Grade_Category	2	14.6466	0.0007
TotalAmount	1	38.2570	<.0001
TotalAmou*Grade_Cate	2	6.2827	0.0432

B.3

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	1.7774	0.0403	1943.6384	<.0001
Grade_Category	Elementary	1	0.0796	0.0448	3.1511	0.0759
Grade_Category	High	1	-0.1099	0.0629	3.0539	0.0805
Grade_Category	Middle	0	0	.	.	.
TotalAmount		1	-0.00011	0.000041	7.3194	0.0068
TotalAmou*Grade_Cate	Elementary	1	-0.00009	0.000048	3.4125	0.0647
TotalAmou*Grade_Cate	High	1	0.000018	0.000060	0.0963	0.7563
TotalAmou*Grade_Cate	Middle	0	0	.	.	.