**SCM 517- DOE Project:**

**LEGO Car**



**Team members:**

Juqin Liu

Ishani Shah

Aurelien Misteli

Anurag Kumar

**Exclusive summary**

Design of Experiment (DOE) is a method that can help us find the relationship between factors or interaction of factors affecting a process. In our project, we designed LEGO race cars and used Minitab to run and analyze our experiments. There are lots of factors to be considered and we determined four factors that may impact our design. The four factors are length, aero dynamism, weight and size of back tires. In our report we’ll illustrate in detail about the effects of four factors and the interactions between them. In the project, we tried to find the input factors that truly contribute to achieve our goal. DOE method helps us to optimize our design of LEGO cars.

**Objective and goals**

In our case project, we have been asked to design a car from multiple LEGO parts. Our goal is to maximize the distance of our race car and try to minimize our cost at the same time if possible. Our ultimate goal is to win the Grand Prix and get 1% bonus points.

To optimize our car, it's essential to reply to these questions:

- What are the variables (intern and external factors) that influence the response

- Define the nominal requirement for the output

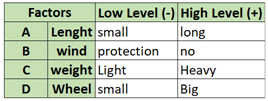
- Which factors influence most of our output?

- How to limit the variation of the uncontrollable factor with the controllable input?

After defining our goal, we determine the process output to achieve it. How to maximize distance? The distance of an object results from the combination of multiple forces, such as speed, environment, weight, energy, and endurance. In our case, with our material, we have to focus on how to collect them, especially the speed and endurance of the car. Then, establish how to control and/or improve it. That leads us to our second steps

To determine all the factors, with the team, we realized a brainstorming and fishbone diagram. One by one, on a backdoor, we wrote all our different ideas and discussed them. If a process was written three or four times, it was automatically implemented in our design. For a variable, which appears one or two times, the teammate had to explain why with the pros and the cons.

In the end, we concluded that four factors have to be performed in the experiment with a low (-) and high (+) level.



Length (A): The size of the car defines stability.

Aero Dynamism (B): To control one external factor, the wind, we look at the spoiler to control our stiction and our speed. The shape of the front spoiler may influence the speed of the car by limiting the strength of the wind. Also, the combination with a back spoiler we might increase our stability. So, we try a different model with (+) and without (-) spoilers)

Weight (C): The weight of a car has as many advantages as an inconvenience. A heavy car (+) will increase your speed but a long process reduces your stability. It makes it harder to control its path, which means the car can easily derive from its line. On the contrary, a light car can get his optimal speed in a limited time

Back Tire (D): In our kit, we dispose of 6wheel with two distinct sizes, small (-) and big (+). We test different wheels to observe the impact on the floor and the variation of the speed. The team decided that putting the big wheels in the front was not necessary because, with the new shape of the car, it will reduce our split streaming.

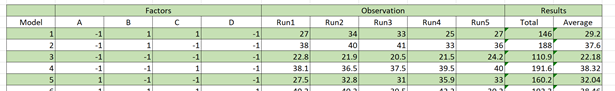
After knowing all the controllable and uncontrollable variables, we proceed to design of the experiments:

The car is composed of 4 factors, so we have to design 2^4 = 16 models to test all the possible factors level combinations. For each model, we did 5 runs in the experiment. We have 4 mains factors (A, B, C D), six two-way interactions(A\*B,A\*C,A\*D,B\*C,B\*D,C\*D), four three-way interaction (A\*B\*C, A\*B\*D, A\*C\*D, B\*C\*D) and a four-way interaction(A\*B\*C\*D)

Because of the four dimensions, we cannot illustrate the relation with a shape in the cube plot graph.

Run the experiments: After assembling everything, we run the car in a condition similar to the car circuit. For the ramp, we use a dashboard and draw a starting line. Each model started at the same distance. With the help of tools (trash and two books), one side of the ramp was lifted to reach an angle of 30°.We use the Law of cosines to get the approximate of the angle. Also, the prototype was placed in an inside room, far from 3 meters of the wall to avoid any crash. Finally, we placed a roll at the end of the ramp to measure our distance in inches.

After setting the environment and design the models, we collected our data:



Model: refers to different models

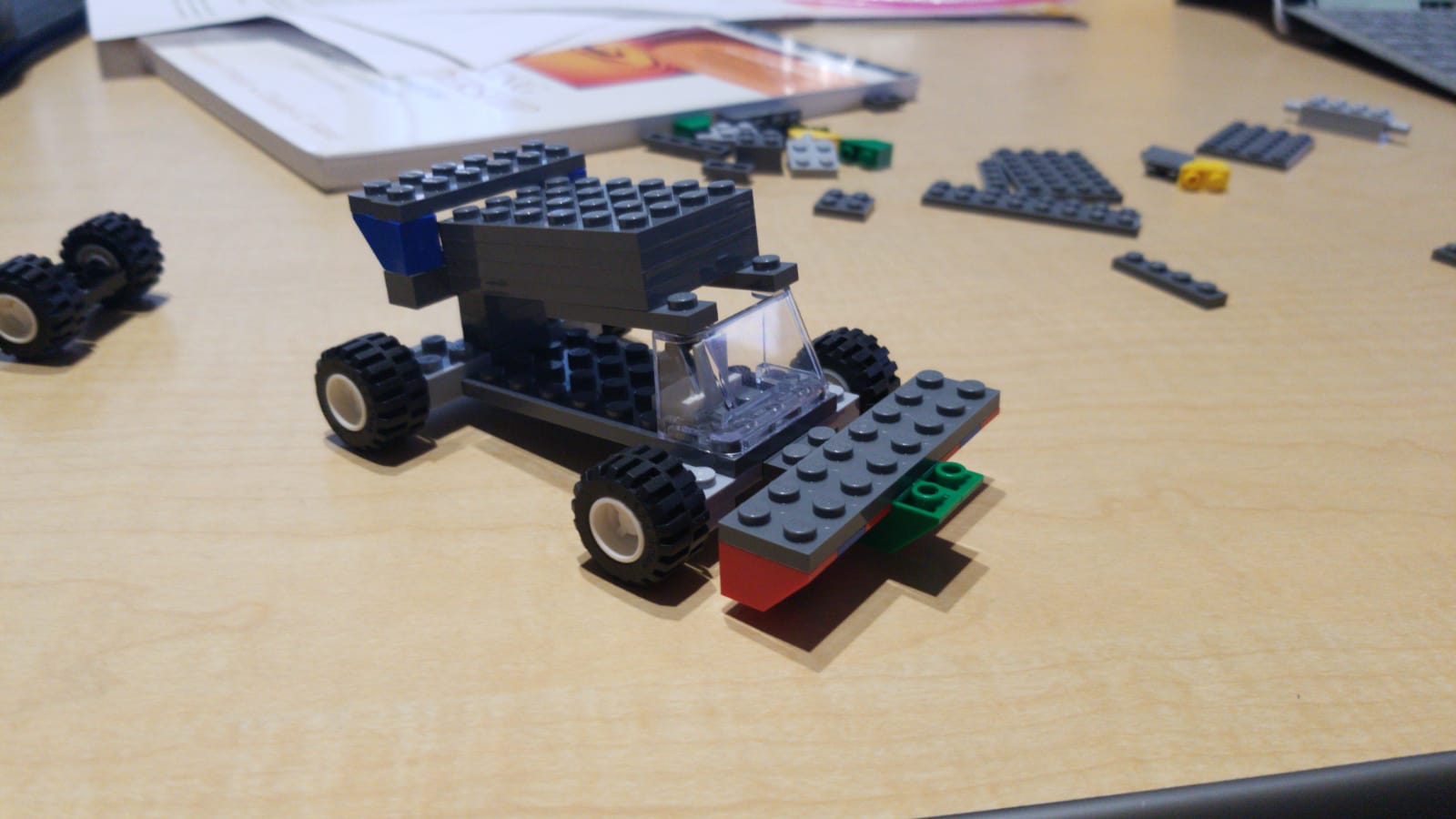
Factors: refers to which variable is used with low and high level as a reference

Observation: refers to a specific attempt of the cars (which is our response in Minitab)

Results: we compute the total and the average distance of each model

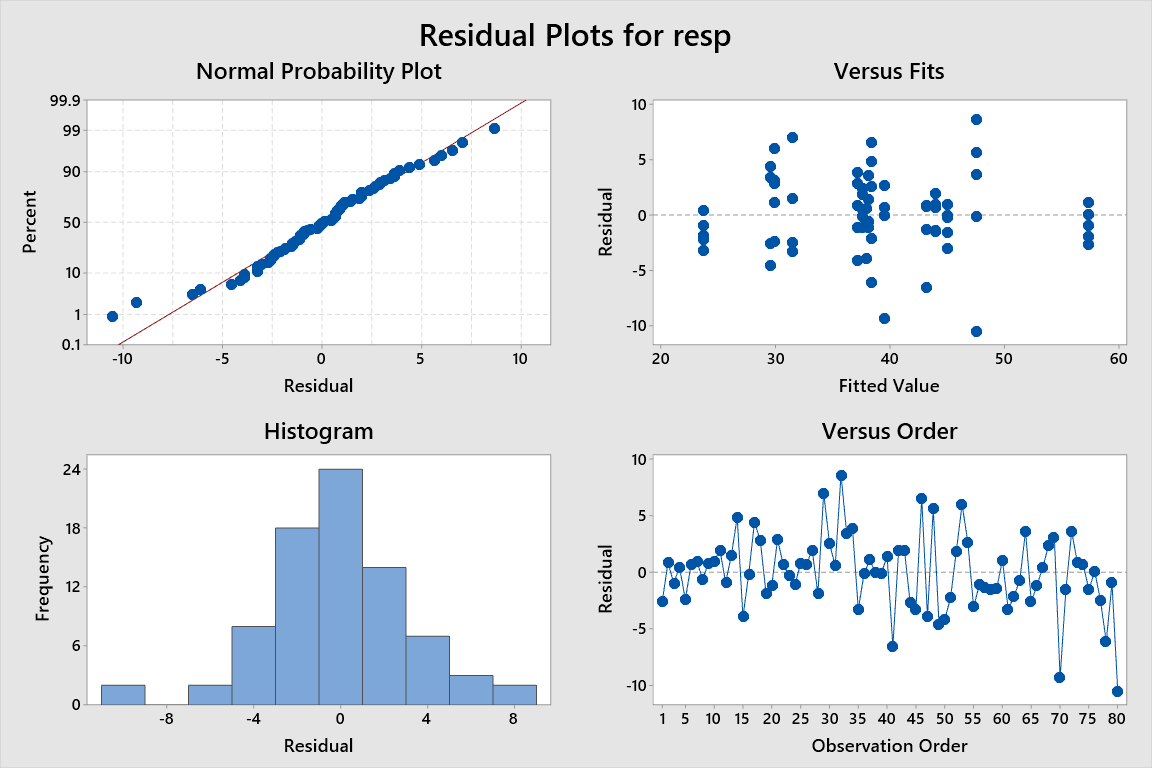
For example, car number 4 is composed of a small chassis, a spoiler as protection, a heavy structure and small wheels in the back. For its run2, it performed 36.5 inches. In the end, it has a total of 191.6 inches with an average of 38.32inches per attempt.

Picture of the car number 4

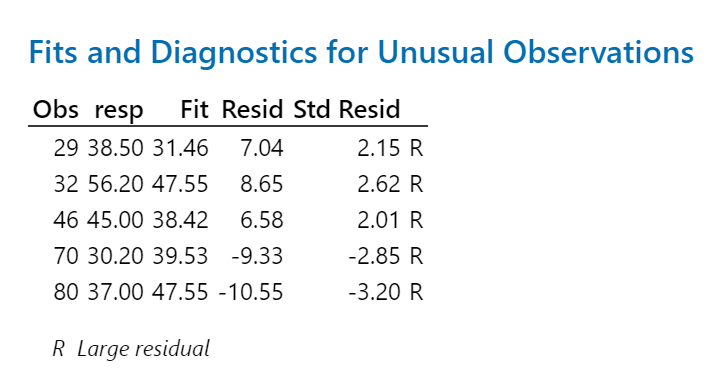


We created a full factorial design in Minitab and analyzed data using ANOVA. Since we have a 2-level design, it’s necessary to have more than one replicates to get the different calculation results for residual errors. If we only do 1 run, the degree freedom will be 0 and we are not able to get the F-stat (as well as p-value) in our ANOVA report. We did 5 runs under the stick control of our experiment environment and process to get a more accurate result.

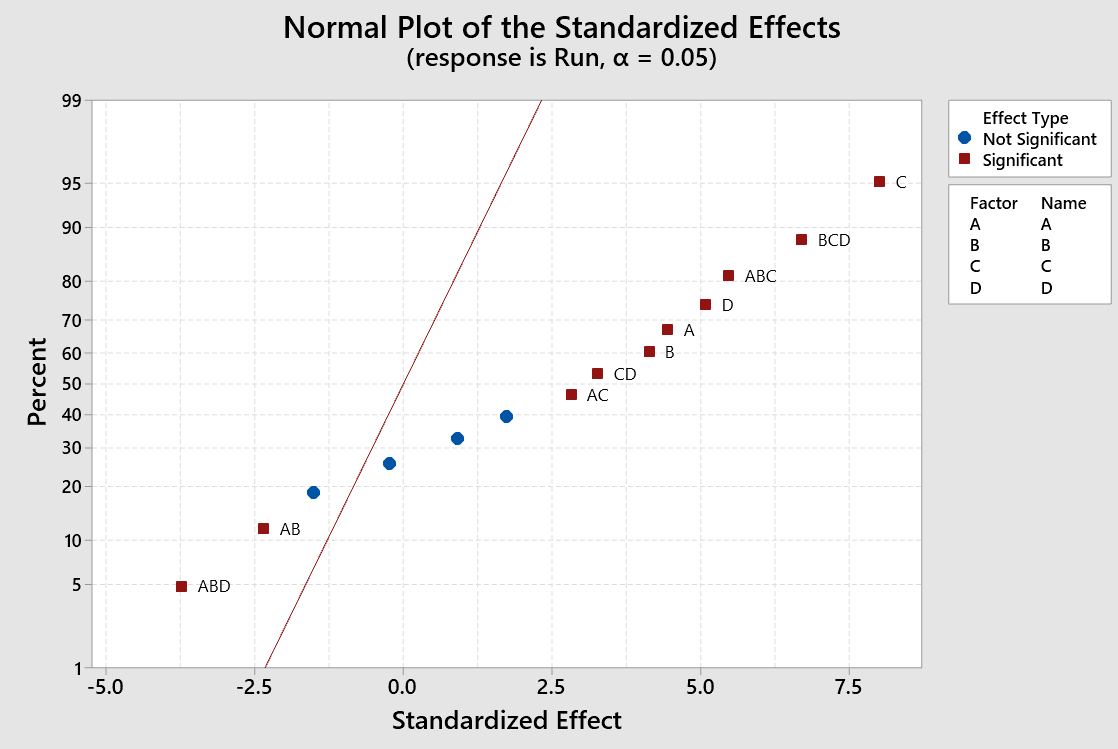
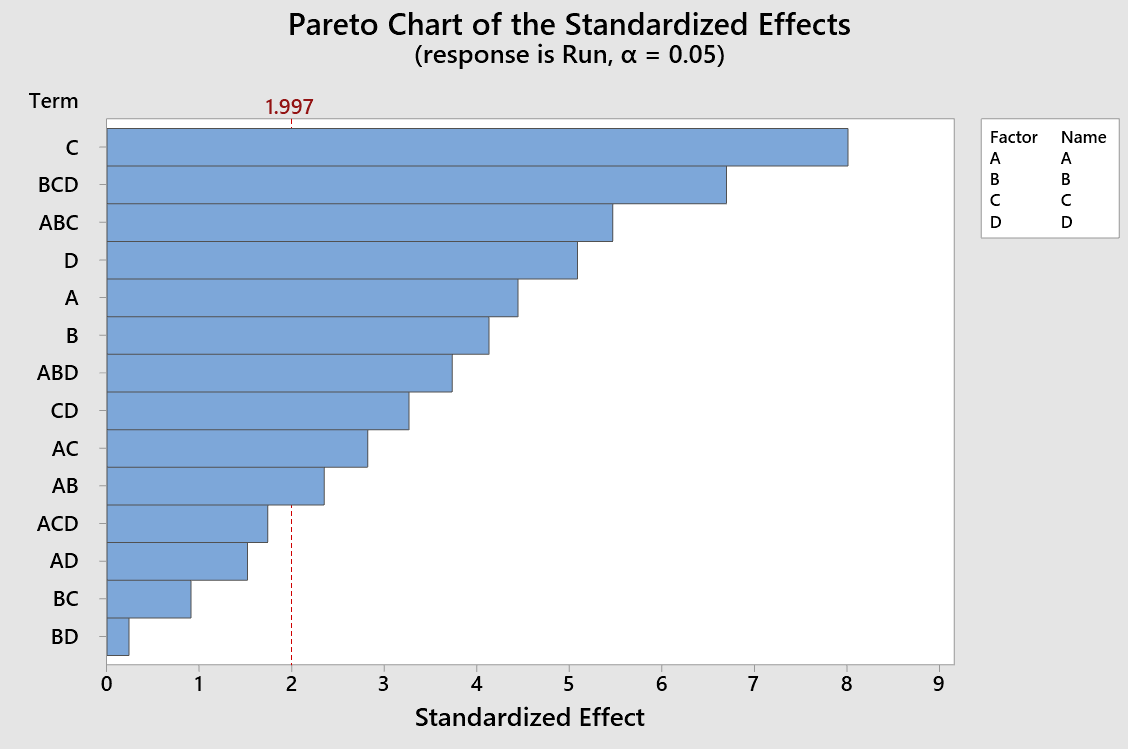
In order to control all other potential variables during the experiment, the same person did the five runs, measured and collected data of the experiment in the same place. We created a DOE project using Minitab to get a randomized order to do the tests and record validated distances. Our graphical analysis and ANOVA report analysis are as follows:



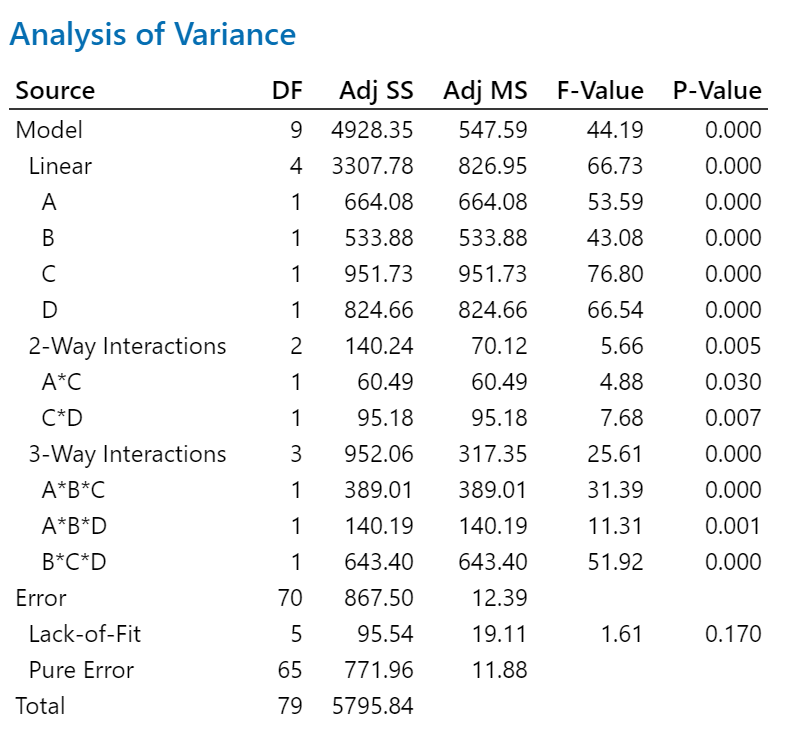
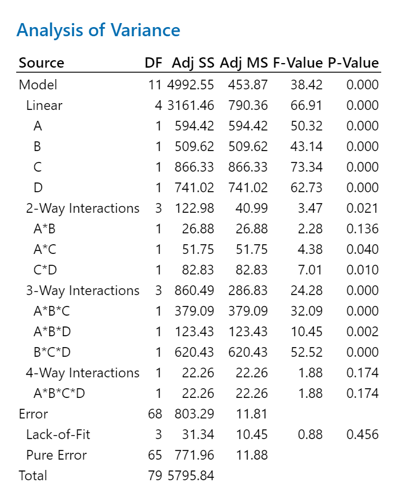
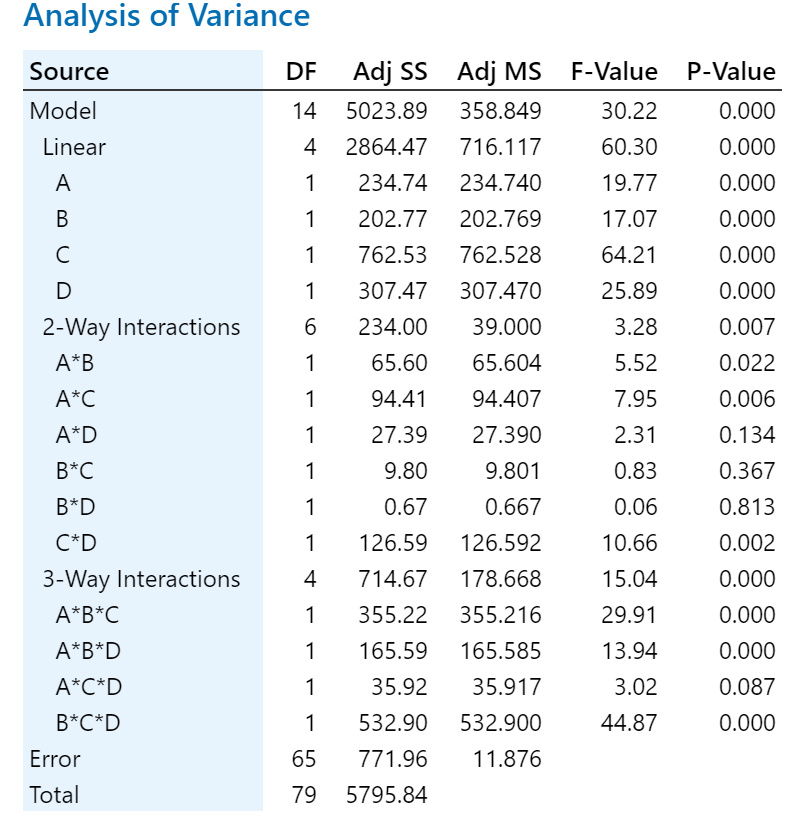
Residual plots helped us to determine whether our model adequate or not. In the histogram, there is no skewness and our data are normally distributed. Then we use the normal probability plot to verify the assumption of our residuals are normally distributed. We can see there is no any curve which indicates the skewness in our graph. There are probably some outliers that may influence the accuracy of our p-value. In the residual versus fits graph, we can see the points are randomly distributed on both sides and no patterns. The residuals versus order graph shows the data we collected in order. There are no trends and patterns on the graph and the residuals falls randomly around the center line.



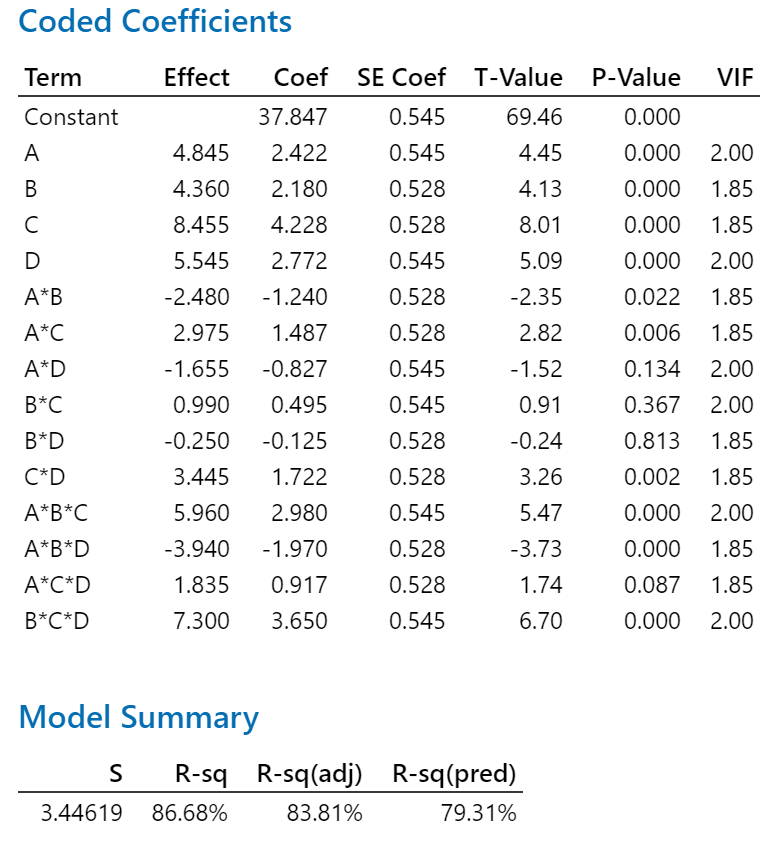
The unusual observation is one of the most important tables for use to diagnose our measurements because it can mislead our result. There are 5 unusual observations in our experiment. The “R” tells us that all of 5 unusual observations are outliers. We tried to remove the unusual observations. The results in the *Model Summary* table like R^2 increased a bit. The non-significant in ANOVA table still the same, which are the interactions of A\*D, B\*C, B\*D and A\*C\*D. Since it's hard for us to find out the causes of unusual observations and the changes are not significant, we keep our original data.



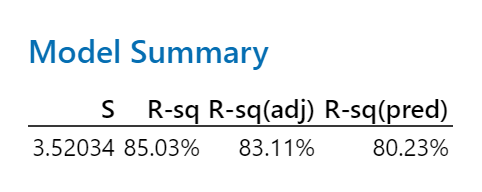
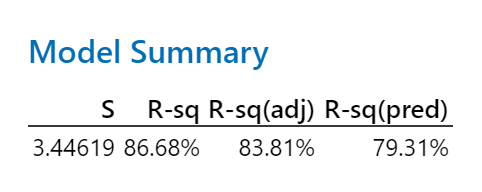
We use a Pareto chart to see whether the main and interactions effects are significant or not. As shows in the Pareto chart, factor C (Weight) has the largest effect and interactions B \* D (aero dynamism \* back tire size) has the smallest effect. In the normal plot graph, we can also see the four insignificant points.



We use the default 95% confidence interval (α = 0.05) to compare our p-value with the significant level. If the p-value is <= 0.05, there is a significant association between distance (response value) and factors. On the contract, if p-value > 0.05, there is not a significant association between the distance and factors. As we can see from the report, the 2-way interactions including A\*D, B\*C and B\*D and 3-way interactions of A\*C\*D are not significant, which means the relationship between distance and those interactions of factors may not depend on the value of other factors (input variables). After remodeling by removing A\*D, B\*C, B\*D, A\*C\*D, A \* B, and A \* B \* C \* D, all factors and interactions are significant.



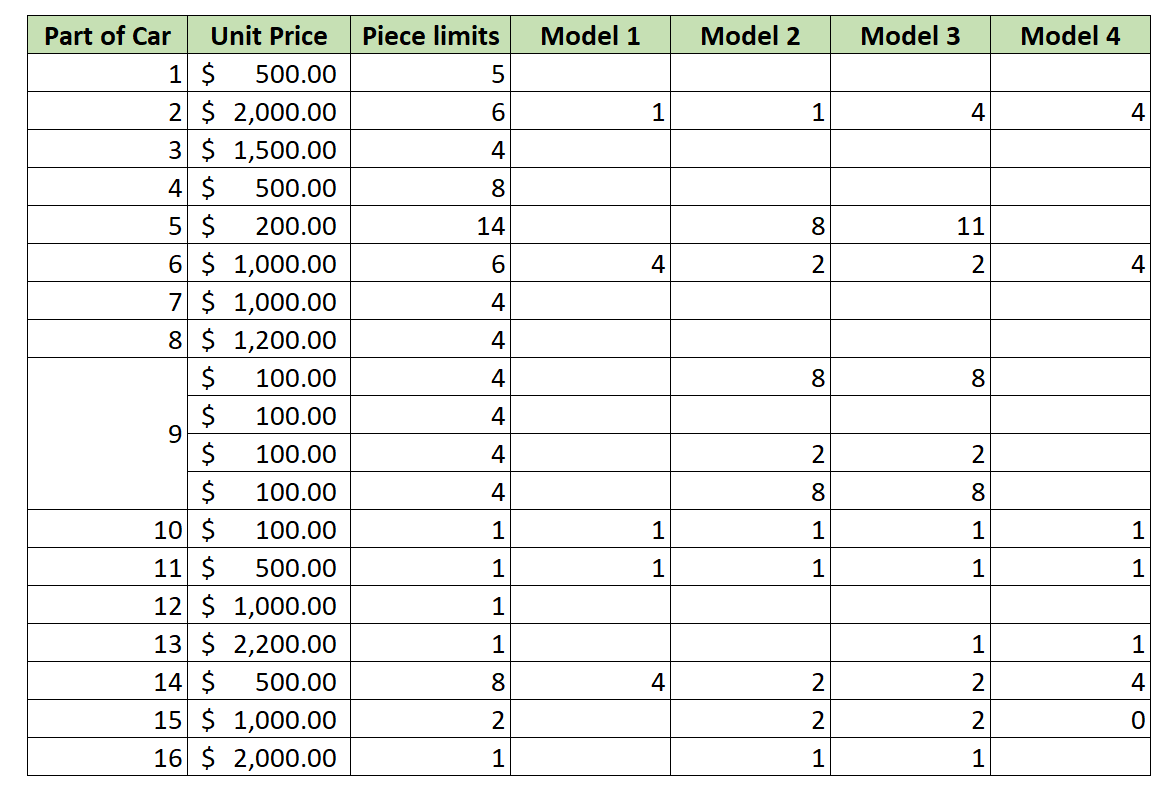
From the coded coefficients table, we can see whether a factor or interactions is a significant predictor of the response. and we can easily get the effect of different factors or interactions using coefficient \* 2.

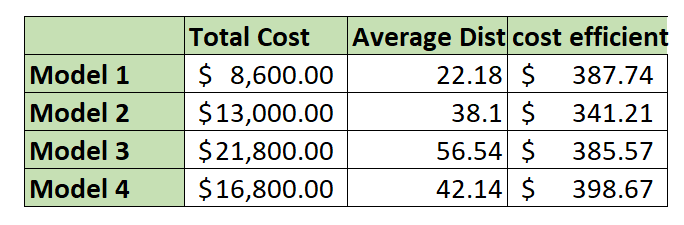


In this table, we can see the goodness-of-fit statistics, which tells us how well the model fits our collected data. The lower standard error (S) value, the better our model describes the response. After remodeling, the S decreased a little bit, but the predicted R-squared value increased. Besides, the gap between pred R-sq and R-sq becomes smaller at the same time. When pred R-sq is less than R-sq, it means our model is overfit, which is because we have non-important terms of effects in the population. The model becomes better after remodel. Our regression equation is as follows:

Regression Equation = 38.485 - 3.060A+ 2.744 B - 3.664 C - 3.410 D + 0.924 A\*C+ 1.159 C\*D - 2.342 A\*B\*C + 1.406 A\*B\*D - 3.012 B\*C\*D

Cost matrix for sample models





Here is some sample cost-efficient analysis of our models. To calculate the efficiency, we simply use the total cost of the model divides the average distance of that model. As you can see the details of model pieces and the costs in the table. If we take cost into consideration, the sample model 2 will be our choice. We chose model 3 the one that runs the furthest distance.

In all, our team had fun during this project. We created lots of models before starting the experiment for fun. DOE project helps us have a better understanding about the design and implementation of projects.

In our experiment, we found out the weight is one of the main effects. The heavier weight, the further distance. However, the cost will probably be more expensive since we use LEGO. LEGO pieces usually weigh more with increasing pieces and it will increase the cost.