Multivariate Analysis of Vehicle Safety

Leigh Cobbs
Texas A&M University
College Station, TX

Mary Cunningham
James Madison University
Harrisonburg, VA

Cheryl Gerde
Morehead State University
Morehead, KY

Abstract

Vehicle safety affects our lives daily. To measure safety, we took a large sample of popular vehicles and set out to create a vehicle safety rating system. To do this, we used two multivariate techniques, Principal Components Analysis and Discriminant Analysis. Principal Components Analysis reduced our set of variables to a smaller set of principal components. We then used Discriminant Analysis to classify vehicles by safety rating using principal components scores.

1. Introduction

People spend many hours of their lives on the road. They pack their loved ones into vehicles every day. According to the National Highway Traffic Safety Administration (NHTSA), in 2001 alone, 41,821 persons died in automobile accidents [5]. With so much responsibility riding on their means of transportation, it seems that people would want to choose their vehicles wisely. There is a vast sea of vehicle information out there, but what if consumers could just look at a vehicle and deduce its safety? Can we predict the safety of a car based only on measurements, and not on crash tests? What types and brands of vehicles are safer? These are some questions we would like to consider.

2. Objective

Our purpose is to collect data on a set of vehicles and to use this data to create a safety rating system. To analyze our data set, we utilize two methods of multivariate analysis, *Principal Components Analysis* (PCA) and *Discriminant Analysis* (DA). The main purpose of PCA is to reduce the dimensionality of the data set by creating principal components. These principal components are used as predictors in DA. Discriminant Analysis creates a classification rule so that we can assign a safety rating to vehicles not in our original data set. The statistical software package Minitab is used in executing PCA and DA, as well as in other manipulations of the data set.

3. Data Collection

3.1 Collection

Unfortunately we could not find an existing data set that contained information useful for analyzing car safety. Consequently, we set out to create a new, unique data set. Our first step was to decide on a collection of variables that we perceived applicable to vehicle safety. A large variety of vehicle measurements and specifications are found at www.edmunds.com, however, this site lacked a few specifications that we feel are important to safety, such as track and center of gravity measurements. These measurements are available at www.dms.dot.gov. Since these new measurements are recorded for only a select group of cars, including these variables greatly reduced possible vehicles for our data set. In addition, we chose to include vehicles of the same production year to reduce discrepancies, and since 2001 automobile data is most recent and complete, we only selected vehicles manufactured in 2001. The final data set contained p=17 variables on n=56 vehicles (Appendix Table A1). In addition to these 17 variables, we obtained safety performance ratings for each vehicle from www.crashtest.com, which are discussed in greater depth in later sections.

3.2 Variable Description

The following is a description of the 17 variables collected. Fourteen of the measurements were obtained from www.edmunds.com [4]. Two exterior measurements, *Length* and *Height* of the vehicle, and three interior measurements, *FrontHeadRoom, FrontShoulderRoom*, and *FrontLegRoom,* are measured in inches. *Weight* and *EngineSize* are measured in pounds and liters, respectively. *Price*, measured in U.S. dollars, is the manufacturer's suggested retail price (MSRP). The successive two variables are not as self- explanatory, so we give brief definitions. *Horsepower* is a measure of mechanical power determined by work and its relation to time. *Torque* is the maximum amount of force produced at a specific speed and is measured in feet per pound (ft/lb) [1]. The next three measurements are taken from www.dms.dot.gov [7]. *CoG Long* is the longitudinal center of gravity, measured in inches from the front axle and *CoG Lat* is the latitudinal center of gravity, also measured in inches, negative towards the driver's side. *Avg Track* is an average of two measurements, front track and rear track, which measure the distance between the left and right front wheels and the distance between the left and right rear wheels, respectively. Transformations using the natural logarithm function are used in the remaining variables for reasons to be discussed later. They include *LNFuelCap, LNWidth, LNTurnCircle*, and *LNWheelBase.* Fuel capacity is measured in gallons, while width and turning circle radius are measured in inches. Wheel base, measured in inches, is the distance from the center of the front wheels to the center of the rear wheels [1]. The variables are summarized in table form below.

Table 3.2

Interior	Exterior	Mechanical	Other
Fronthead	Length	LNFuelCap	Price
FrontShldr	LNWidth	Horsepower	
FrontLeg	Height	Torque	
	Weight		
	LNWheelBase		
	LNTurnCircle		
	Avg. Track		
	CoGLong		
	CoGLat		

3.3 Normality Test

To investigate the normality of each variable, we first construct a Quantile-Quantile plot, or Q-Q plot, to see if it exhibits a linear pattern. Most of the Q-Q plots appear to follow a linear pattern, however, we take a closer look with the Normality Test. We compute the sample correlation coefficient between the ordered sample values and their z-scores. If their correlation coefficient is sufficiently close to one, a perfect positive correlation, we fail to reject the null hypothesis, H_0: $\rho=1$, where ρ is the population correlation coefficient. Thus, the variable is considered to follow a normal distribution. With a sample size of 56 the test of H_0: $\rho=1$ versus H_1: $\rho<1$ will reject H_0 if $r_Q < .9695$ at $\alpha=0.01$, where r_Q is the sample correlation coefficient given by $r_Q=\text{cov}(\underline{x}_j,\underline{q}_j)/(\sqrt{\text{var}(x_j)}\sqrt{\text{var}(q_j)})$, and .9695 is the critical value [6]. For four of our 17 variables, fuel capacity, width, wheel base, and turn circle, we reject the null hypothesis. Since normality increases our confidence in our multivariate results, we elect to transform these four variables to try to achieve normality. We tried two Box-Cox transformations, square root and natural log functions, and found that the natural log produced higher correlation coefficients in all four cases. By taking the natural log, three more variables can be considered normally distributed at the $\alpha=0.01$ significance level. Width remains the only non-normally distributed variable. Since both multivariate techniques, PCA and DA, are robust to normality, we proceed with the analyses.

4. Principal Components Analysis

4.1 Theory

The first technique we apply is Principal Component Analysis. PCA is performed on *n* p-variate data vectors, given by $\underline{x} = (x_1, x_2, \ldots, x_p)'$. It is assumed that $\underline{x} \sim N(\underline{\mu},\Sigma)$, though PCA is robust to non-normality. The primary objective of PCA is to reduce the dimensionality of the data. By grouping highly correlated variables, linear combinations of the variables, called principal components, are generated. Each principal component is written in the form:

$$PC_i = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{ip}x_p,$$

where the a_{ij}'s are the principal component coefficients. The set of values for each principal component when calculated with a specific data vector is referred to as that vector's principal component scores.

The principal components are ordered so that each principal component accounts for a larger amount of variation than the next. Specifically, each PC_i is associated with an eigenvalue, λ_i, of Σ, where the λ_i's are ordered from largest to smallest. The coefficient vector, $\underline{a_i}$, for PC_i is the unit eigenvector paired with λ_i. It can be shown through Lagrangian maximization that the values of $\underline{a_i}$, the normalized eigenvector associated with λ_i, achieves the maximum possible variance of PC_i. In fact, $Var(PC_i) = \lambda_i$. Then the proportion of the variance in the original data explained by PC_i is given by $\lambda_i/(\lambda_1 + \lambda_2 + \ldots \lambda_p) = \lambda_i/tr(\Sigma) = \lambda_i/(\sigma_{11} + \sigma_{22} + \ldots \sigma_{pp}) = \lambda_i/(Var(x_1) + Var(x_2) + \ldots Var(x_p))$. Thus PC_1 is associated with the largest eigenvalue λ_1 and accounts for the largest amount of variation possible, PC_2 is associated with λ_2 and accounts for the next largest amount of variation, and so on.

A necessary assumption when applying PCA is the dependence of the variables. This assumption is necessary since it allows different variables that measure similar quantities to be grouped into a single principal component, which results in a few principal components being used to account for a large proportion of the variance of the original data. Thus the number of principal components is less than the number of original variables, *p*, and reduction of dimensionality is achieved. It is also important to note that the principal components are now uncorrelated, that is they are independent from each other. Thus, PCA essentially takes a set of correlated variables, and returns a new, smaller set of variables (in the from of the principal component scores) which are uncorrelated and account for most of the variance present in the original data [3].

4.2 Data

Before running PCA on the data, we conduct the Test for Independence, which tests the null hypothesis, H_0: $P = I$ versus H_1: $P \neq I$, where P is the correlation matrix. Our goal is to reject H_0, so that we may conclude dependence of the variables. We run the test, using a test statistic

$$\boldsymbol{\chi^2 = -(n - 1 - ((2p + 5) / 6) * \ln |R|},$$

with $p(p-1)/2$ degrees of freedom, where R is the maximum likelihood estimate of P and p=17 is the number of variables. In our analysis, the null hypothesis is rejected, since for $\alpha = 0.05$, $\chi^2 = 1183.44 > \chi^2_{136,0.05} = 164.2162$ and our p-value is less than zero. Therefore, the variables are in fact dependent, and we can proceed with the principal component analysis [6].

Using PCA, we set out to achieve 80% of the variance being explained by our principal components. Using Minitab, we can generate up to 17 principal components, their coefficients, and their scores for each data vector. Minitab also pairs each principal component with the corresponding eigenvalue and gives the amount of variance explained. The eigenvalues and variance explained by the first eight principal components is given Table 4.2 below.

Table 4.2

Eigenvalue	10.55	1.534	1.098	1.021	0.718	0.590	0.469	0.332
Proportion	0.609	0.090	0.065	0.060	0.042	0.035	0.028	0.020
Cumulative	0.609	0.699	0.764	0.824	0.866	0.901	0.929	0.948

Notice that the objective of 80% variance explained is achieved with the 4^{th} principal component. These four components actually account for 82.4% of the variation. Thus we are able to describe the 17 original variables with just four measurements, which can now be used for further analysis.

5. Discriminant Analysis

5.1 Theory

The second technique we apply is Discriminant Analysis (DA). The goal of DA is to classify a data vector, $\underline{x} \sim N(\underline{\mu},\Sigma)$, into one of k multivariate normal populations, $\Pi_1, \ldots, \Pi_k$. Samples are taken from each population to from the training sample. Using the training sample, a discriminant rule is found. This discriminant rule partitions each population and calculates the population mean. The partitioning is done in the manner that minimizes the total probability of misclassification. A new data unit can then be classified using the Mahalanobis squared distance (M-distance). The M-distance essentially measures how far the new data unit is from the population mean and is given by

$$D_i^2 = (\underline{x} - \underline{\mu_i})' \, \Sigma^{-1} \, (\underline{x} - \underline{\mu_i}),$$

where $\underline{\mu_i}$ is the mean vector for Π_i. The unit is then classified into the population for which the M-distance is minimum, that is into Π_i, where $D_i^2 = \min\{ D_1^2, \ldots, D_k^2 \}$. Two methods are available for Discriminant Analysis – Linear and Quadratic. Linear DA requires the assumption that each of the population covariance matrices are equal, while Quadratic DA does not. It should be noted that DA is also robust to non-normality [3].

5.2 Data

Populations

Regarding our data, we use Discriminant Analysis to group cars by their safety ratings. The ratings we obtained from www.crashtest.com were as follows: Excellent Safety, Good Safety, Acceptable But Better Choices Exist, Marginal Safety, and Poor Safety. Each rating is specific to the individual models, styles, and years of production. We assign each safety category a “star” rating, summarized in Table 5.2 below [2].

Table 5.2.1

Rating	Recommendation	Star Category
Excellent	Highly Recommended	5
Good	Recommended	4
Acceptable But Better Choices Exist	Not Recommended	3
Marginal	Unacceptable	2
Poor	Unacceptable	1

We failed to discover any cars with a safety rating of one star, and so each vehicle listed in our data set belongs to one of four groups, or populations: two, three, four, or five-stars. Using ratings from www.crashtest.com for each of the 56 cars, we place them into their corresponding populations. The first population, Π_1, corresponds to the five-star vehicles, and n_1=9 vehicles from the training sample belong to this population. Populations Π_2, Π_3, and Π_4 correspond to the four, three, and two-star populations, and their sample sizes are n_2=20, n_3=18, n_4=9 vehicles, respectively.

Sigma Test

With the training sample chosen, we perform the sigma test to see if Linear DA can be applied. We test H_0: $\Sigma_1=\Sigma_2=\Sigma_3=\Sigma_4=\Sigma$ versus H_1: at least one pair of Σ's are not equal, with α=0.05 and $p(p+1)(k-1)/2 = 17(18)(3)/2 = 459$ degrees of freedom. With the populations $\Pi_1, \ldots, \Pi_4$ defined as previously discussed, the test statistic $M/C \approx \chi^2_{p(p+1)(k-1)/2}$ is computed as follows:

S_i=Covariance matrix for Π_i

$$S_p = \text{pooled sigma matrix} = \sum_{i=1}^{4}((n_i-1)S_i)/(n_i-1)$$

$$M=[\sum_{i=1}^{4} n_i - 1]*[\ln |S_p|] - \sum_{i=1}^{4}(n_i-1)\ln|S_i| = 4761.093$$

$$1/C=1-[(2p^2+3p-1)/(6(p+1)]\{[\sum_{i=1}^{4} 1/(n_i-1)\,] - [1/(n_1+n_2+n_3+n_4-4)]\} = 0.336676$$

Thus M/C = 1602.947.

The critical value at a 5% level of significance level is $\chi^2_{459,.05}$ = 509.9475. Since $M/C > \chi^2_{459,.05}$ and the p-value is very close to zero, we reject H_0, and conclude that not all Σ_i's are equal and hence we cannot apply linear DA. We therefore must use quadratic DA [3].

Analysis

For the implementation of DA, we again use Minitab. We are unable to run the quadratic Discriminant Analysis directly on the 17 variables because they are too highly correlated. We therefore use the principal component scores, thus creating a set of uncorrelated data on which the DA can be run (Appendix Table A3). Apparent Error Rate refers to the percent of elements of in the training sample that is misclassified. It is obtained by subtracting the proportion correct from one. We want to achieve an AER of less than 10%. We begin by running the test on four principal component scores, since in PCA we found this number achieved the required 80% explained variance. This yields an AER of 48.2%. We then run the test on increasingly larger numbers of PC scores, which yield lower and lower AERs. Based on the 10% rule, eight PC scores adequately classify the data, since the resulting AER is 5.4%. The summary of the classification using these eight scores is shown in Table 5.2.2 below.

Table 5.2.2

Predicted Rating	Actual Rating			
	2-star	3-star	4-star	5-star
2-star	9	2	0	0
3-star	0	15	0	0
4-star	0	1	20	0
5-star	0	0	0	9
Total n_i	9	18	20	9
Proportion Correct	100%	83.3%	100%	100%

Notice that the two, four, and five-star vehicles are all classified correctly. Two of the three-star vehicles are misclassified as two-star vehicles, and a third is misclassified as a four-star vehicle. The total number of vehicles correctly classified is 53, or 94.6%.

With the classification rule established, we can now input new data on additional vehicles and classify them into one of the four populations. We collect data on five new vehicles from model years 2001 and 2002 (using the same method as before), standardize this data, and multiply the standardized variables by the PC coefficients to obtain their respective PC scores (Minitab requires the variables to be standardized to compute the scores). Once we obtain the classifications, we compare them to the actual classifications given on www.crashtest.com. The results are summarized in Table 5.2.3.

Table 5.2.3

Vehicle	$\mathbf{D_4^2}$	$\mathbf{D_3^2}$	$\mathbf{D_2^2}$	$\mathbf{D_1^2}$	**Predicted Rating**
2002 Toyota Highlander (4)	103.185	14.220	2.336	82.928	4-star
2002 Ford Taurus Wagon (4)	388.520	6.514	16.182	497.508	3-star
Jeep Cherokee 2DR 4x2 SE Utility	125.117	11.693	72.929	1449.547	3-star
Ford Escape 4x2 (3)	23.111	1.047	1.948	354.249	3-star
Chevy Silverado 4x4 (3)	20.811	14.670	6.446	2057.295	4-star

6. Discussion

The topic of car safety raises many questions. Among them, how do the best selling vehicles rate? We confront this question by examining the top ten best selling cars for the model year 2001. The best selling vehicles are, starting in descending order, Ford F-Series, Chevrolet Silverado, Ford Explorer, Honda Accord, Toyota Camry, Ford Taurus, Dodge Ram Pickup, Honda Civic, Ford Ranger, and Ford Focus [4]. Interestingly, eight out of ten of these top sellers are contained in our data set, and one, the Chevrolet Silverado, is one of our test vehicles. Of the nine vehicles, all but one have a four-star rating, meaning that these eight are recommended vehicles; only one, the Dodge Ram Pickup, in not recommended. This suggests good news for consumers; many popular cars are safe as well. Perhaps the car-buying public is concerned about safety and shop for safe cars, or perhaps, top car-makers are going to great lengths to ensure that their top sellers are safe. Ideally, both theories are correct. Examination of the data set offers some support for the second claim; popular car- makers such as Ford and Honda are consistently categorized into four- and five-star rating groups.

Finally, we recognize a limitation in our data set. Obviously, our data set contains a mixture of vehicle types, cars, vans, trucks, and sport utility vehicles (SUVs), however, due to limitations in data availability, the mix contains a large amount of trucks and SUVs. This can be justified by recognizing that half of the top ten best sellers, and further, the top three, are all trucks and SUVs, reflecting an increasing popularity for these vehicles. While this mixture may put smaller vehicles at a disadvantage in the discrimination, we feel that there is some value in comparing safety of all classes of vehicles. Clearly, there is room for further investigation.

7. Conclusions

Vehicle safety is an important issue that impacts many people. Further analysis could help ensure that automobiles are safer in the future. Car companies, for instance, want to produce safe vehicles to aid sales. Our analysis could be performed on a larger and broader sample to help companies establish if there are particular design features such as heavier vehicles with more horsepower that increase safety. The analysis could also be applied to more specific classes of variables – mid-size cars, extended cab trucks, small SUVs – to individualize the safety features for each class. Other variables than the ones that are discussed in the paper could be explored, such as vertical center of gravity and ground clearance. Consumers, too, should be aware of important features that contribute to the safety of the cars they drive. In the future, our analysis and others could be applied to later models in order to ascertain if the same variables are still important and to examine the effect of new features. In the end, statistical analysis of car safety could possibly save millions of dollars in damages and more importantly, save lives. Is it worth it? We think so.

8. Acknowledgements

As part of the 2002 SUMSRI experience, we would like to extend our gratitude to all of those that have helped us grow as mathematicians and statisticians during our time spent at SUMSRI. In particular, we would like to thank Dr. Vasant Waikar for all of the knowledge he imparted to us and for his support. We are grateful for our graduate assistant, Lindsay Kellam, who helped us in every step of the program, including help with topic selection, becoming familiar with Minitab, and paper revisions. Thank you to Dr. Tom Farmer, for his expertise throughout the writing process. Thanks also, to the many speakers and faculty members that took time out of their schedules to make us well rounded students. Finally, we extend our gratitude to the National Science Foundation, National Security Agency, and Miami University for sponsoring this opportunity for us to participate in SUMSRI.

9. References

[1] Cartrackers.com, Retrieved in July 2002 from www.cartrackers.com/auto_glossary.html

[2] Crashtest Results, *Safety Information: 2001*, Retrieved in July 2002 from www.crashtest.com/explanations/safety/ie.html

[3] Dallas E. Johnson, *Applied Multivariate Methods for Data Analysts*, Duxbury Press (1998)

[4] Edmunds.com, *Used Cars: 2001,* Retrieved in July 2002 from www.edmunds.com/used/index.html

[5] National Highway Traffic Safety Administration, *Accident Statistics,* Retrieved in July 2002 from www.nhtsa.dot.gov

[6] Richard A. Johnson and Dean W. Wichern, *Applied Multivariate Statistical Analysis (Fourth Edition),* Prentice Hall (1998)

[7] United States Department of Transportation, *Docket Management System, Simple Search-docket number 8298,* Retrieved in July 2002 from www.dms.dot.gov

Appendices

Table A1

Make & Model	Price	LNFuelCap	EngineSize	Horsepower	Torque	FrontHead	FrontShldr	FrontLeg	Length	LNWidth	Height
Acura MDX 4x4	37101	2.47654	3.5	240	245	38.7	61.2	41.5	188.5	4.17899	68.7
Lexus RX300 4WD Wagon 4Dr	31023	2.47654	3.0	220	222	39.5	57.7	40.7	180.3	4.18510	65.7
Toyota Sienna LE 4dr passenger van	21839	2.47654	3.0	210	220	40.6	60.4	41.9	194.1	4.19720	66.9
Volkswagen Jetta 4dr GLS Sedan	14856	2.52573	2.0	115	122	37.4	53.7	41.5	172.3	4.19720	56.9
Chevrolet Impala 4Dr LS Sedan	16944	2.58022	3.8	200	225	39.2	59.0	42.2	200.0	4.20020	57.3
Honda Odyssey 4Dr LX Psngr van	24212	2.58022	3.5	210	229	41.2	62.6	41.0	201.2	4.20020	68.5
Lincoln LS 4Dr V6 Sedan	23854	2.58022	3.0	210	205	40.4	57.7	42.8	193.9	4.20020	56.1
Ford Expedition 4x4 4Dr XLT Utility	27065	2.58022	4.6	215	290	39.8	63.9	40.9	204.6	4.20320	76.6
Volkswagen Jetta Wagon 4Dr GLS	16078	2.66026	2.0	115	122	38.6	53.7	41.5	173.6	4.20320	58.5
Subaru Legacy L wagon AWD	15488	2.67415	2.5	165	166	40.2	53.9	43.3	187.4	4.20618	59.6
Ford F-150 4X4 4door flsz crewcab	26557	2.67415	4.6	231	293	39.8	63.7	41.0	225.9	4.20916	73.9
Honda Civic 4DR DX	11112	2.68785	1.7	115	110	39.8	52.6	42.2	174.6	4.20916	56.7
Ford F- 150 4X2 4Dr XLT CrewCab SB	20769	2.69463	4.6	231	293	39.8	63.7	41.0	225.9	4.21065	73.9
Ford Windstar 4Dr SE	20975	2.70805	3.8	200	240	39.3	60.9	40.7	200.9	4.21213	68.2
Subaru Forester	16888	2.72785	2.5	165	166	40.2	53.5	43.0	175.6	4.21509	62.8
Ford Focus 4 Dr LX	9871	2.76632	2.0	110	125	39.3	53.7	43.1	174.9	4.22391	56.3
Ford Taurus 4Dr LX	12611	2.77259	3.0	155	185	40.0	57.3	42.2	197.6	4.22391	56.1
Honda Accord 4 Dr EX Sedan	17445	2.82731	2.3	150	152	38.5	56.9	42.1	189.4	4.22391	56.9
Chevrolet Suburban 4x4 4Dr K1500	30288	2.83321	5.3	285	325	40.7	65.2	41.3	219.3	4.22391	75.4
Chevrolet Tahoe 4DR 4x4 LS	29285	2.83908	5.3	285	325	40.7	65.2	41.3	196.9	4.22975	76.3
Ford Explorer 4x4 XLS Utility	20259	2.85647	4.0	210	240	39.9	56.7	42.4	190.7	4.23266	67.5
Mazda MPV 4Dr ES	19162	2.89037	2.5	160	165	41.0	59.8	40.8	187.0	4.23989	68.7

Ford Crown Victoria STD	14992	2.89037	4.6	220	265	39.4	60.8	42.5	212.0	4.24707	56.8
Toyota Camry LE	15276	2.89037	2.2	136	150	38.6	56.2	43.4	188.6	4.24992	55.1
Toyota Tundra Ext.Cab 4x2 Ltd. V8 SB	22179	2.89037	4.7	245	315	40.3	62.4	41.5	217.5	4.25135	70.9
Ford Ranger Ext.Cab 4x2 XL	12463	2.91777	3.0	150	185	39.2	53.8	42.2	202.9	4.25277	64.8
Dodge Stratus 4Dr ES	14232	2.91777	2.7	200	192	37.6	55.2	42.3	191.2	4.25419	54.9
Toyota Highlander 4x4 STD	22048	2.91777	2.4	155	163	40.0	57.9	40.7	184.4	4.25419	68.7
Ford Focus Street Wagon	12710	2.91777	2.0	130	135	39.3	53.7	43.1	178.2	4.25703	53.9
Dodge Durango 4X4 4dr Sport Wagon	21952	2.94444	4.7	235	295	39.8	57.3	41.9	193.3	4.26970	72.9
Dodge Ram Van Wagon 2Dr 1500	15395	2.94444	3.9	175	225	40.5	68.0	39.0	187.2	4.26970	79.5
Jeep Wrangler 4X4 2Dr Sport Utility	16327	2.95491	4.0	190	235	42.3	51.9	41.1	155.4	4.27528	71.1
Chevrolet Tracker 4Dr 4x4 LT	13810	2.97041	2.5	155	460	39.9	52.8	41.4	162.8	4.27805	66.3
Honda CRV FWD LX	17159	2.97041	2.0	146	133	40.5	53.3	41.5	177.6	4.28082	65.9
Jeep Grand Cherokee 4x4 Laredo	20654	2.98568	4.0	195	230	39.7	58.9	41.4	181.5	4.29046	69.4
Mitsubishi Montero Sport 3.5XS	19681	2.98568	3.5	197	223	38.9	55.5	42.8	181.1	4.29046	68.3
Chrysler PT Cruiser STD	16552	2.99573	2.4	150	162	40.4	54.6	40.6	168.8	4.29320	63.0
Mitsubishi Montero 4Dr Limited 4WD	26676	2.99573	3.5	200	235	41.4	58.5	42.7	188.9	4.29592	73.1
Toyota Corolla 4Dr CE	9578	2.99573	1.8	125	125	39.3	52.8	42.5	174.0	4.30000	53.7
Toyota Tacoma Ext.Cab 4x4 SB	16721	3.02042	3.4	190	220	38.7	53.9	42.8	203.0	4.30000	67.5
Toyota RAV4 4WD 4Dr STD	16170	3.03975	2.0	148	142	41.3	46.3	42.4	165.1	4.30271	65.3
Toyota Echo 2Dr STD Coupe	8315	3.04452	1.5	108	105	39.9	52.3	41.1	163.2	4.32015	57.9
Pontiac Aztek 2WD	16284	3.16969	3.4	185	210	39.7	58.7	40.5	182.1	4.33467	66.7
Pontiac Aztek 4x4	17953	3.21888	3.4	185	210	39.7	58.7	40.5	182.1	4.33467	66.7
Dodge Ram 1500 Quad cab 4x4	19824	3.21888	5.2	230	300	40.2	66.0	41.0	224.1	4.35028	74.6
Toyota Tacoma Xtracab 4x2	15684	3.21888	3.4	190	220	38.7	53.9	42.8	202.3	4.35927	67.5

hecrolet Tracker 2dr 4x4	11810	3.21888	2.0	127	134	40.9	53.0	41.4	151.8	4.36437	66.5
Chevrolet Astro van	16084	3.25810	4.3	190	250	39.2	64.0	41.6	189.8	4.36691	75.0
Toyota Prius 4dr	18358	3.25810	1.5	70	82	38.8	52.8	41.2	169.6	4.36818	57.6
Dodge Neon 4DR Highline ACR Sedan	10897	3.25810	2.0	150	135	38.4	53.4	42.2	174.4	4.37071	56.0
Dodge Ram Ext. Cab 4X4 4dr ST LB	21694	3.27336	5.9	245	345	40.2	66.0	41.0	244.1	4.37071	77.2
Hyundai Accent 2Dr GS Hatchback	6783	3.29584	1.6	105	106	38.9	52.8	42.6	166.7	4.37324	54.9
Isuzu Rodeo 4X4 4Dr LS Wagon	17806	3.40120	3.2	205	214	38.9	56.3	42.1	183.5	4.37324	69.2
Chev Cavalier	11133	3.43399	2.4	150	155	37.6	53.9	41.9	180.9	4.37324	53.0
Chev 5-10 4X2	26557	3.48124	4.6	231	293	39.8	63.7	41.0	225.9	4.37952	73.9

Table A1 (Continued)

Make & Model	Weight	LNWheelBase	LNTurnCircle	Avg Track	CoG Long	CoG Lat
Acura MDX 4x4	4387.0	4.46130	3.42751	66.65	46.12	-1.060
Lexus RX300 4WD Wagon 4Dr	3924.0	4.53582	3.44999	61.50	44.13	-0.700
Toyota Sienna LE 4dr passenger van	3932.0	4.53689	3.47197	62.55	47.95	-1.110
Volkswagen Jetta 4dr GLS Sedan	2908.0	4.56539	3.48124	59.20	41.43	-0.590
Chevrolet Impala 4Dr LS Sedan	3466.0	4.57471	3.49043	61.35	42.59	-0.290
Honda Odyssey 4Dr LX Psngr van	4248.0	4.58088	3.49043	66.25	50.75	-0.930
Lincoln LS 4Dr V6 Sedan	3593.0	4.58497	3.52930	60.50	54.73	-0.520
Ford Expedition 4x4 4Dr XLT Utility	5345.0	4.59411	3.53515	65.45	58.59	-1.000
Volkswagen Jetta Wagon 4Dr GLS	3079.0	4.59512	3.53515	59.05	43.04	-0.940
Subaru Legacy L wagon AWD	3345.0	4.59915	3.54674	57.55	47.30	-1.070
Ford F-150 4X4 4door flsz crewcab	4644.0	4.60916	3.54962	65.30	55.83	-1.090
Honda Civic 4DR DX	2421.0	4.63473	3.54962	57.95	41.82	-1.120
Ford F- 150 4X2 4Dr XLT CrewCab SB	4644.0	4.63473	3.56671	65.55	58.87	-0.970
Ford Windstar 4Dr SE	4223.0	4.63473	3.56671	63.90	47.42	-0.440
Subaru Forester	3140.0	4.63473	3.56953	57.75	45.01	-0.850
Ford Focus 4 Dr LX	2564.0	4.63570	3.57235	58.10	40.72	-0.740
Ford Taurus 4Dr LX	3354.0	4.63667	3.57795	61.90	39.58	-0.660
Honda Accord 4 Dr EX Sedan	3075.0	4.64535	3.57795	60.90	42.66	-0.760
Chevrolet Suburban 4x4 4Dr K1500	5447.0	4.64727	3.58629	65.60	61.99	-1.320
Chevrolet Tahoe 4DR 4x4 LS	5050.0	4.65396	3.59457	58.55	51.53	-1.470
Ford Explorer 4x4 XLS Utility	4045.0	4.65491	3.59457	60.65	46.68	-1.460
Mazda MPV 4Dr ES	3682.0	4.66250	3.59457	64.40	51.44	-0.680
Ford Crown Victoria STD	3917.0	4.66627	3.59731	60.40	41.32	-0.850
Toyota Camry LE	3097.0	4.66721	3.60550	65.20	55.09	-1.440

Toyota Tundra Ext.Cab 4x2 Ltd. V8 SB	4402.0	4.66721	3.61899	58.05	51.28	-1.380
Ford Ranger Ext.Cab 4x2 XL	3313.0	4.67189	3.62167	60.70	40.44	-1.150
Dodge Stratus 4Dr ES	3297.0	4.67189	3.62167	61.65	47.32	-0.930
Toyota Highlander 4x4 STD	3715.0	4.67563	3.62167	58.10	43.09	-0.740
Ford Focus Street Wagon	2717.0	4.68213	3.62167	57.50	46.72	-0.580
Dodge Durango 4X4 4dr Sport Wagon	4848.0	4.68491	3.62966	62.90	53.40	-0.890
Dodge Ram Van Wagon 2Dr 1500	4985.6	4.68491	3.63231	67.65	44.39	-1.350
Jeep Wrangler 4X4 2Dr Sport Utility	3316.0	4.68675	3.63759	58.40	49.69	-0.610
Chevrolet Tracker 4Dr 4x4 LT	2987.0	4.69592	3.64021	60.30	45.85	-0.660
Honda CRV FWD LX	3126.0	4.69684	3.64545	59.65	47.36	-0.600
Jeep Grand Cherokee 4x4 Laredo	3972.0	4.70502	3.64806	58.20	49.18	-1.350
Mitsubishi Montero Sport 3.5XS	4010.0	4.71133	3.64806	58.00	41.69	-0.720
Chrysler PT Cruiser STD	3123.0	4.71492	3.65584	63.40	48.23	-1.020
Mitsubishi Montero 4Dr Limited 4WD	4675.0	4.71671	3.66099	57.35	39.04	-1.200
Toyota Corolla 4Dr CE	2410.0	4.73795	3.67122	59.05	51.60	-1.240
Toyota Tacoma Ext.Cab 4x4 SB	3515.0	4.74057	3.68135	59.90	43.70	-0.540
Toyota RAV4 4WD 4Dr STD	2877.0	4.74232	3.68888	56.25	36.82	-1.130
Toyota Echo 2Dr STD Coupe	2035.0	4.75359	3.68888	63.60	51.27	-1.440
Pontiac Aztek 2WD	3779.0	4.75359	3.68888	63.20	41.89	-0.650
Pontiac Aztek 4x4	3779.0	4.77153	3.69635	63.15	44.96	-0.620
Dodge Ram 1500 Quad cab 4x4	5545.7	4.77996	3.70130	67.95	57.10	-1.390
Toyota Tacoma Xtracab 4x2	3355.0	4.79331	3.70130	57.25	52.52	-1.480
hecrolet Tracker 2dr 4x4	2811.0	4.80320	3.70130	57.45	41.50	-0.610
Chevrolet Astro van	4323.0	4.80320	3.72086	65.45	51.80	-0.900
Toyota Prius 4dr	2765.0	4.83390	3.74479	53.15	40.48	-1.750
Dodge Neon 4DR Highline ACR Sedan	2585.0	4.85437	3.75420	58.10	38.98	-0.078
Dodge Ram Ext. Cab 4X4 4dr ST LB	5600.0	4.86753	3.80444	67.95	57.90	-1.360
Hyundai Accent 2Dr GS Hatchback	2280.0	4.93087	3.82647	56.40	37.23	-1.410
Isuzu Rodeo 4X4 4Dr LS Wagon	4124.0	4.93087	3.82647	59.85	49.02	-0.900
Chev Cavalier	2617.0	4.93087	3.82647	57.25	37.57	-0.860
Chev 5-10 4X2	4644.0	4.93231	3.84374	54.70	48.62	-1.210

Table A2--PCA Coefficients

-0.221884	-0.141466	-0.138161	0.234683	0.132035	0.601457	0.507735	0.195582
-0.295755	-0.053484	-0.033877	0.042297	0.031394	-0.026629	-0.141466	-0.131524
-0.290794	0.041597	-0.088029	0.137289	0.002199	-0.057124	-0.210483	-0.163252
-0.273777	0.033209	-0.127491	0.300657	-0.109301	0.216493	-0.009016	-0.095253
-0.252409	-0.028591	-0.091250	0.257144	-0.118037	0.027090	-0.556213	-0.078627
-0.096650	-0.561852	-0.297525	0.091240	-0.276132	-0.559648	0.372590	-0.053522
-0.280890	-0.010114	0.184238	-0.120986	0.131989	0.047808	0.152136	-0.281049
0.138068	0.469511	-0.429691	0.244463	-0.387081	-0.002090	0.133847	-0.388761
-0.260020	0.384308	0.001516	-0.047672	0.027058	-0.215044	0.207409	0.137616
-0.283789	0.037722	0.126247	-0.052880	0.152497	-0.127891	0.217601	-0.299449
-0.253625	-0.322558	-0.125227	0.034270	0.129617	-0.045325	-0.273667	0.197873
-0.298972	-0.060850	-0.040797	0.059444	0.088673	0.099623	0.015021	-0.134938
-0.263362	0.323816	-0.004076	-0.052212	0.048026	-0.271273	0.109364	0.359853
-0.271754	0.264784	0.074358	0.014467	0.111174	-0.231379	-0.000339	0.208329
-0.193738	-0.071178	0.494759	-0.321223	-0.404174	0.100518	0.054662	-0.344243
-0.210264	-0.001590	-0.094390	-0.318868	-0.650704	0.215158	-0.055063	0.420761
0.077177	-0.018907	0.589751	0.683719	-0.249210	-0.114353	0.071298	0.187475

Table A3- PCA Scores

-2.21765	-0.37039	0.57602	0.39661	0.45449	2.93902	1.04455	-0.74045
-0.79442	-0.96107	0.77829	1.16521	0.80744	1.48559	0.65844	0.54315
-1.31011	-0.31322	-0.40877	-0.08215	-0.25072	-0.17194	0.75621	-0.28377
3.57424	0.88576	1.68917	-0.01903	0.89530	0.93655	-0.32872	0.38581
0.26805	1.15785	1.35105	1.64864	-0.29859	-0.27990	0.29155	-0.55067
-2.43859	-1.24184	0.59608	-0.32696	-0.67488	-0.20492	1.22460	-0.23427
-0.21946	0.65563	-0.14002	1.12745	-1.80447	0.24601	1.49957	0.31631
-4.72548	-0.73498	0.39912	-0.35450	-0.49059	0.92884	-0.16879	0.16419
3.15324	0.13255	0.62308	-0.59866	0.72260	0.52540	0.10470	0.32616
2.01871	0.55832	-1.49739	0.20222	-0.80239	-0.34189	0.52764	-0.30775
-5.15935	0.72935	0.37331	-0.44391	-0.11654	-0.05252	0.35564	0.83357
4.02445	0.10454	-0.40060	-0.80806	0.17073	-0.70900	0.44338	0.02770
-5.05011	0.85313	0.69952	-0.61639	-0.67528	-0.55477	-0.11936	0.88648
-2.14734	-0.08687	1.84676	0.73118	0.20384	0.00986	-0.07387	0.30495
2.43608	-0.18970	-1.02480	0.75962	-0.62922	-0.08779	0.37294	-0.27209
4.19227	0.89939	0.02385	0.17658	-0.26623	-0.70564	0.23410	-0.36876
1.02741	0.82118	0.92162	0.47384	0.03143	-1.37643	0.56969	-0.87166

Table A4- Eigenanalysis

Eigenvalue	10.355	1.534	1.098	1.021	0.718	0.59	0.469	0.332	0.257	0.216
Proportion	0.609	0.09	0.065	0.06	0.042	0.035	0.028	0.02	0.015	0.013
Cumulative	0.609	0.699	0.764	0.824	0.866	0.901	0.929	0.948	0.963	0.976

Eigenvalue	0.134	0.084	0.064	0.053	0.034	0.027	0.013
Proportion	0.008	0.005	0.004	0.003	0.002	0.002	0.001
Cumulative	0.984	0.989	0.992	0.996	0.998	0.999	1

Table A5- Car Ratings Table

Make & Model	Rank
Acura MDX 4x4	5
Lexus RX300 4WD Wagon 4Dr	5
Toyota Sienna LE 4dr passenger van	5
Volkswagen Jetta 4dr GLS Sedan	5
Chevrolet Impala 4Dr LS Sedan	5
Honda Odyssey 4Dr LX Psngr van	5
Lincoln LS 4Dr V6 Sedan	5
Ford Expedition 4x4 4Dr XLT Utility	5
Volkswagen Jetta Wagon 4Dr GLS	5
Subaru Legacy L wagon AWD	4
Ford F-150 4X4 4door flsz crewcab	4
Honda Civic 4DR DX	4
Ford F- 150 4X2 4Dr XLT CrewCab SB	4
Ford Windstar 4Dr SE	4
Subaru Forester	4
Ford Focus 4 Dr LX	4
Ford Taurus 4Dr LX	4
Honda Accord 4 Dr EX Sedan	4
Chevrolet Suburban 4x4 4Dr K1500	4
Chevrolet Tahoe 4DR 4x4 LS	4
Ford Explorer 4x4 XLS Utility	4
Mazda MPV 4Dr ES	4
Ford Crown Victoria STD	4
Toyota Camry LE	4
Toyota Tundra Ext.Cab 4x2 Ltd. V8 SB	4
Ford Ranger Ext.Cab 4x2 XL	4
Dodge Stratus 4Dr ES	4
Toyota Highlander 4x4 STD	4
Ford Focus Street Wagon	4
Dodge Durango 4X4 4dr Sport Wagon	3
Dodge Ram Van Wagon 2Dr 1500	3
Jeep Wrangler 4X4 2Dr Sport Utility	3
Chevrolet Tracker 4Dr 4x4 LT	3
Honda CRV FWD LX	3
Jeep Grand Cherokee 4x4 Laredo	3
Mitsubishi Montero Sport 3.5XS	3
Chrysler PT Cruiser STD	3
Mitsubishi Montero 4Dr Limited 4WD	3
Toyota Corolla 4Dr CE	3
Toyota Tacoma Ext.Cab 4x4 SB	3
Toyota RAV4 4WD 4Dr STD	3
Toyota Echo 2Dr STD Coupe	3
Pontiac Aztek 2WD	3
Pontiac Aztek 4x4	3
Dodge Ram 1500 Quad cab 4x4	3
Toyota Tacoma Xtracab 4x2	3
Checrolet Tracker 2dr 4x4	3
Chevrolet Astro van	2
Toyota Prius 4dr	2
Dodge Neon 4DR Highline ACR Sedan	2
Dodge Ram Ext. Cab 4X4 4dr ST LB	2
Hyundai Accent 2Dr GS Hatchback	2
Isuzu Rodeo 4X4 4Dr LS Wagon	2
Chev Cavalier	2
Chev 5-10 4X2	2

Table A6- Sample Correlation Coefficients Table

Variable	Correlation Coefficient
MSRP	.980
LNFuelCap	.988
EngineSize	.980
Horsepower	.991
Torque	.975
FrontHead	.992
FrontShldr	.972
FrontLeg	.982
Length	.982
LNWidth	.947
Height	.972
Weight	.991
LNWheelBase	.971
LNTurnCircle	.982
Avg Track (in)	.982
CoG Long (in from front axle)	.989
CoG Lat (neg twrds Driver Sd)	.990