Glass Identification

Blaise Chidlow, Edith Garcia, Steven Maxwell, Cait Moran

DePaul University 2018

* **Powerpoint -** <https://maildepaul-my.sharepoint.com/:p:/g/personal/cmoran21_mail_depaul_edu/ETnT3ksKl2hGqTVE8SzEpkEBw3KoiLg5ciVvvFFVRZHCyA?e=eZdlRm>

# Abstract

Using a data set with the refractive indices and weight percent of eight oxides of 214 glass shards, this study sought to find out which oxides were more prevalent between different categories of glass as well as whether the chemical composition and refractive index determined the different types of glass. A MANOVA found that some oxides were good indicators for some glass types but not for others. Additionally, MANOVA contrast tests found a significant difference between the groups of: windows vs non-window glass, building vs vehicle glass, and float vs non-float glass. Discriminant analysis found that the success rates for classifying windows vs non-windows, as well as for classifying building vs vehicle glass, were higher than the success rate for classifying float vs non-float glass.

# Background

When a crime occurs, and glass is shattered, being able to match the glass at the crime scene to remnants left on a person or to a vehicle is an incredibly vital component in any crime scene investigation. Usually, a shard of glass is being compared to glass whose origin is already known. To determine whether glass evidence is a match, the shards must share some properties that have been measured in past studies including refractive index, manufacturing process, or weight percent in oxides, among others. Since the aforementioned properties are relevant to our dataset, these will be the focus.

Glass is defined to be a state of matter that has some properties of both liquids and crystalline solids. Crystalline solids are known for having a pattern in their atomic structure whereas for glass it is random. This state is known as being amorphous, or between solid and liquid. Since the structure is random, there isn’t one specific composition that defines all glass. However, at the molecular level, there are enough similarities ("Chemistry of Glass," 2011). There are a number of elements that most glass types are made of. These include Si, Na, Ca, Pb, and K ("What Elements are Used to Make Glass?," 2018). The composition is usually described in terms of oxide weight percent.

Oxides are compounds of oxygen and another element. Their weight percent is the ratio of the weight of the oxide over the entire weight of the specimen, in this case, the glass shard (Francis, n.d.). The combinations of different oxides result in different physical properties in the finished product. Having some amount or none of specific oxides, changes the thermal resistance, durability, flexibility, and more physical properties of the glass ("What Elements are Used to Make Glass?,"2018). This allows for glass to meet different needs such as whether it will be used in labs or whether a window needs to be strong enough to resist extreme weather changes.

Since glass composition will be different for glass that needs to be exposed to the environment, such as windows and vehicle headlamps, as opposed to glass that will be used mostly indoors, such as containers and tableware, it is expected these categories will have or lack specific oxides that will help classify them into their respective categories. Glass for both windows and vehicles are usually created by a process known as the float process. Containers are usually created by the glassblowing process.

The float process is where all the elements, as raw materials, are put into a furnace that eventually produces the glass product floating on top of melted tin. Ribbons of glass are what come out once the product has been cooled ("Glass Manufacturing Process," n.d.). The glassblowing process is one where a glob of molten glass is blown into and shaped into a container. It can also be blown inside a mold to maintain uniform shape, such as a when making long-necked bottles (“The Story of Glass,” n.d.). One thing to keep in mind is that even though glass may be manufactured from the same factory, there will be variation in the concentration of the elements due to issues in the manufacturing process. Those issues include having imperfect measurements and improper mixing of raw materials, the age of the machinery, and variations in the heating (Gamble & Kirk, 1943). Therefore, it is expected that those glass shards from the dataset will show a difference between those that were float processed to those that were not. One thing to note is that in this dataset, there is no specification as to what process was used for the non-float glass shards. However, despite the method of processing not being clear for some of the glass in our dataset, what is known for every glass shard is its refractive index.

Refractive index is a ratio of velocity of light in a vacuum to the velocity of the light in the medium, in this case, glass (Bottrell, 2009). Refractive index is easy to measure on fragments of glass and has had good discrimination potential in other studies (Bottrell, 2009). In a more recent study from 2015, using glass obtained from a hit and run case, the densities, in grams per cubic centimeter, and the refractive indices of both samples of windshield glass were higher than those of headlight glass (Theogene & Rajiv, 2015). Therefore, refractive index has been shown to be a reliable source of discrimination among glass.

Bottrell mentioned that one of the takeaways in their study was that in a single source of glass variations do exist, but they will be largest between glass of different types (2009). Therefore, variations in refractive index, oxide composition, and perhaps method of processing are what this project sought to find. Specifically, our research objectives were to find which chemical combinations are more prevalent in different glass types and whether the chemical composition and refractive index determined the different types of glass in our dataset.

# The Data

The data for this study were accessed from the UCI Machine Learning Repository on May 20th , 2018. The sample size is n = 214 observations of 10 variables, with 9 being quantitative and 1 being categorical. Note that the original data contained 11 variables; however, for this study, we disregarded the numeric ID variable that was included in the original dataset. The source of the data is the Central Research Home Office Forensic Science Service, from September 1987. In the original study, the researchers' interest in the classification of glass types was motivated by crime scene investigation.

For this study, a total of ten variables are utilized. They are described in the following table:

Table 1: List of Variables and Responses:

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Type | Description and Response | How the Variable Will Be Referenced in the Report |
| RI | Quantitative | Refractive Index | 'RI' |
| Sodium | Quantitative | Weight percent of sodium oxide | 'Sodium' |
| Magnesium | Quantitative | Weight percent of magnesium oxide | 'Magnesium' |
| Aluminum | Quantitative | Weight percent of aluminum oxide | 'Aluminum' |
| Silicon | Quantitative | Weight percent of silicon oxide | 'Silicon' |
| Potassium | Quantitative | Weight percent of potassium oxide | **'**Potassium' |
| Calcium | Quantitative | Weight percent of calcium oxide | 'Calcium' |
| Barium | Quantitative | Weight percent of barium oxide | 'Barium' |
| Iron | Quantitative | Weight percent of iron oxide | 'Iron' |
| Type of Glass | Categorical | Contains seven levels, denoting seven different types of glass. | 'Glass Type' |

Note that, for the variable 'Glass Type', the meaning of each of the seven levels is as follows:

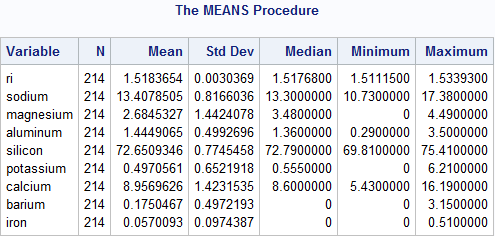
|  |  |
| --- | --- |
| Glass Type Value | Meaning |
| 1 | Building windows that are float processed |
| 2 | Building windows that are not float processed |
| 3 | Vehicle windows that are float processed |
| 4 | Vehicle windows that are not float processed (Not present in this dataset) |
| 5 | Containers |
| 6 | Tableware |
| 7 | Headlamps |

For this study, the dataset contains 214 observations, each with a value for the above variables (I.e. there are no missing values).

## Descriptive Statistics

Using SAS, we present mean, standard deviation, median, min and max of the 9 quantitative variables in the following table:

Table 2: Descriptive Statistics of the Quantitative Variables

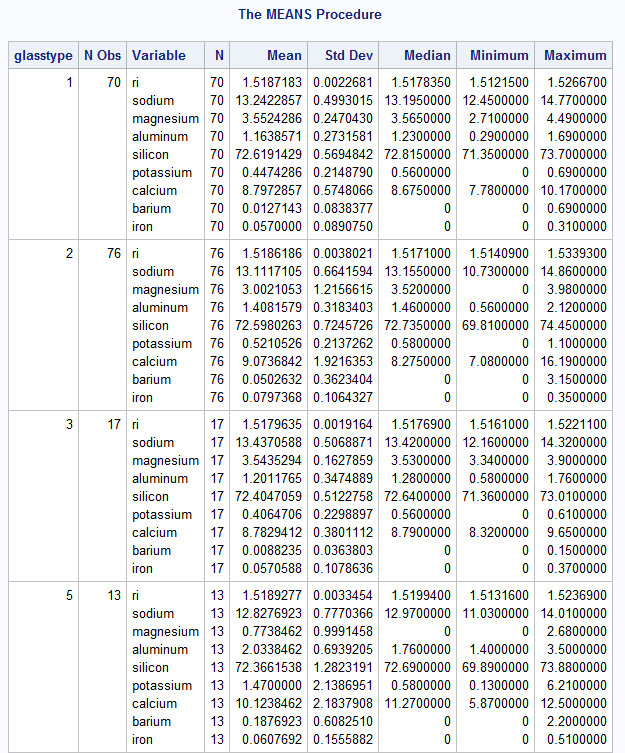


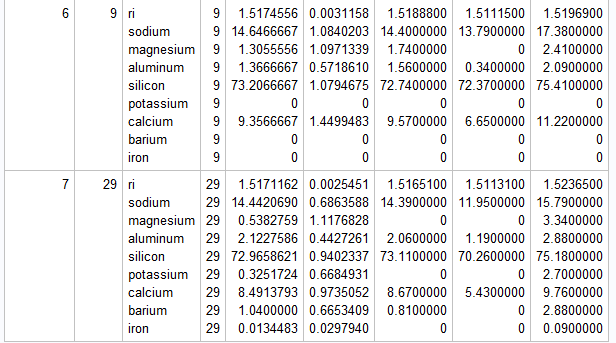
In examining the output from Table 2, observe that 'RI' has a mean value of about 1.518 with a standard deviation of about 0.003. The variable 'Sodium' has a mean value of about 13.408 with a standard deviation of about 0.817. The variable 'Magnesium' has a mean value of about 2.685 with a standard deviation of about 1.442. The variable 'Aluminum' has a mean value of about 1.445 with a standard deviation of about 0.499. The variable 'Silicon' has a mean value of about 72.651 with a standard deviation of about 0.775. The variable 'Potassium' has a mean value of about 0.497 with a standard deviation of about 0.652. The variable 'Calcium' has a mean value of about 8.957 with a standard deviation of about 1.423. The variable 'Barium' has a mean value of about 0.175 with a standard deviation of about 0.175. The variable 'Iron' has a mean value of about 0.057 with a standard deviation of about 0.097.

### Descriptive Statistics by Glass Type

In the following table, the summary statistics of each level of 'Glass Type' for the 9 quantitative variables

Table 3: Descriptive Statistics for the 9 Quantitative Variables by 'Glass Type'

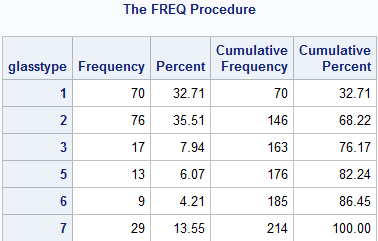




In Table 3, the descriptive statistics for the 9 quantitative variables for each level of the variable 'Glass Type' is shown. Observe that the differences among the variables do not vary too much among the six groups.

For the categorical variable 'Glass Type', the frequency table is below:

Table 4: Frequencies of the variable 'Glass Type'



From Table 4, observe that 70 (32.71%) of the observations belong to glass type 1, 76 (35.51%) of the observations belong to glass type 2, 17 (7.94%) of the observations belong to glass type 3, 13 (6.07%) of the observations belong to glass type 5, 9 (4.21%) of the observations belong to glass type 6, and 29 (13.55%) of the observations belong to glass type 7. Observe that none of the observations in the dataset belong to glass type 4. I.e., there are no observations that are classified as "Vehicle windows that are not float processed."

In commenting on the frequencies in Table 3, that glass types 2, 1, and 7 have the most counts, with 76, 70, and 29 respectively. Glass types 6, 5, and 3 have the fewest counts, with 9, 13, and 17 respectively.

# Analysis

## Multivariate Analysis of Variance (MANOVA)

In an attempt to answer the research question, *"What chemicals are more prevalent in certain types of glass?,* a MANOVA test may be used to start the exploration of this topic. MANOVA is the multivariate form of the univariate ANOVA test. MANOVA tests the difference in two or more vectors of means when there are several dependent variables. In the case of the glass data being evaluated, the GlassType variable is the dependent variable with 7 different levels or 7 different dependent outcomes.

There are a few assumptions the MANOVA test should account for including multivariate normality, equality of covariance matrices, near absence of multivariate outliers, linearity, absence of multicollinearity. For the purposes of the study, the multivariate normality equality of covariance matrices will be evaluated.

### Testing Assumptions- Multivariate Normality

Q-Q (quantile – quantile) plots give an excellent representation of the multivariate normality. The independent variables are evaluated for each dependent glass type. See the results below. Ideal normality should be a nearly straight positively linear plots of the points. This can also be reinforced by explicit Shapiro-Wilks tests, if desired.

|  |  |
| --- | --- |
| Glass Type 1: Float Processed Building Windows | Glass Type 2: Non-Float Processed Building Windows |
|  |  |

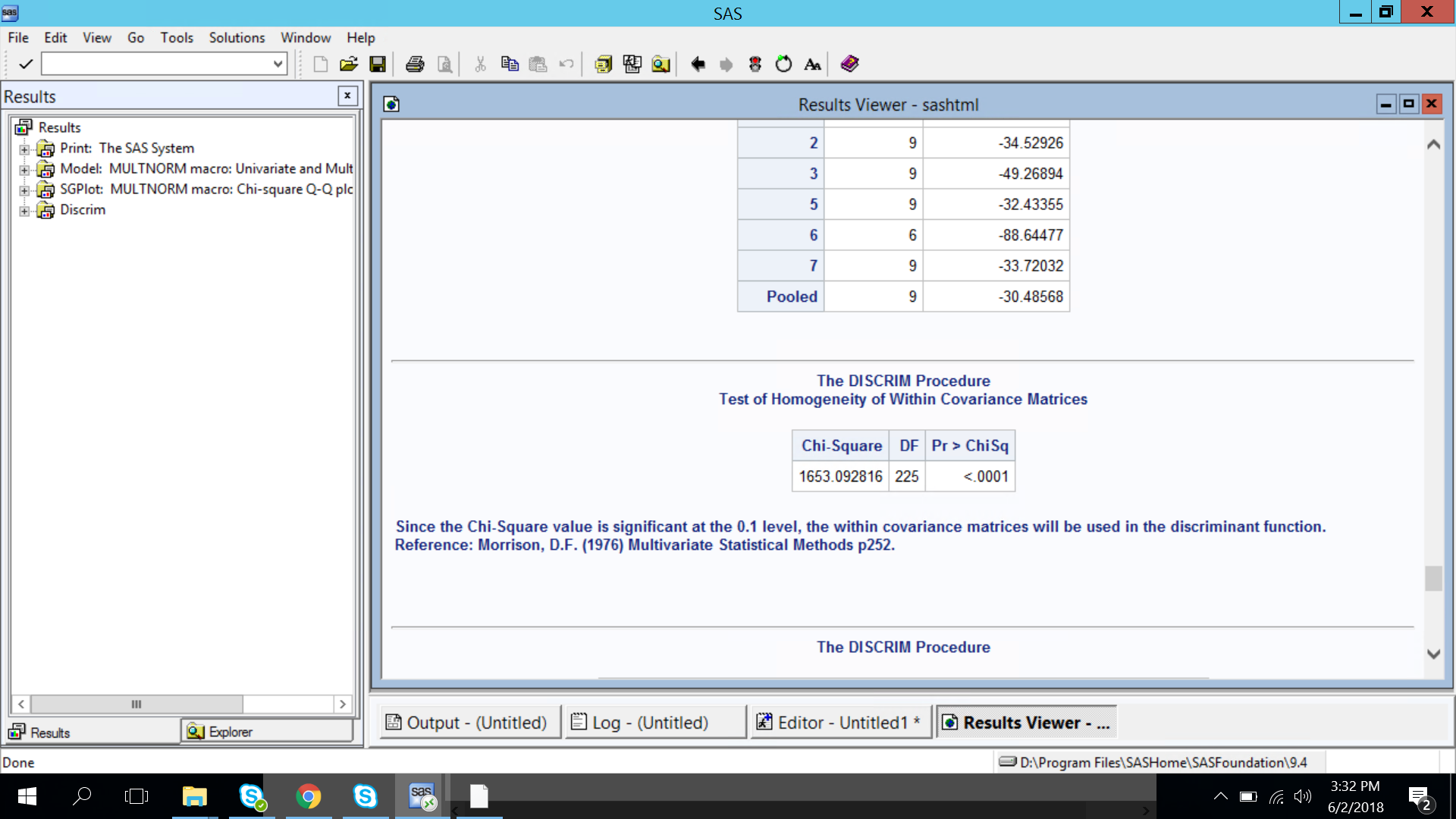
|  |  |
| --- | --- |
| Glass Type 3: Float Processed Vehicle Windows | Glass Type 4: Non-Float Processed Vehicle Windows |
|  | *N/A* |

|  |  |
| --- | --- |
| Glass Type 5: Containers | Glass Type 6: Tableware |
|  | MULTNORM: Removing observations with missing values...  MULTNORM: Checking for singularity of covariance matrix...  WARNING: The variable Potassium is constant.  WARNING: The variable Barium is constant.  WARNING: The variable Iron is constant.  WARNING: The Correlation Matrix is not positive definite.  ERROR: Covariance matrix is singular. |

|  |
| --- |
| Glass Type 7: Headlamps |
|  |

According to the multivariate normality tests, the variables are not considered normal in any group and do not meet the assumptions.

### Testing Assumptions- Equality of Covariances



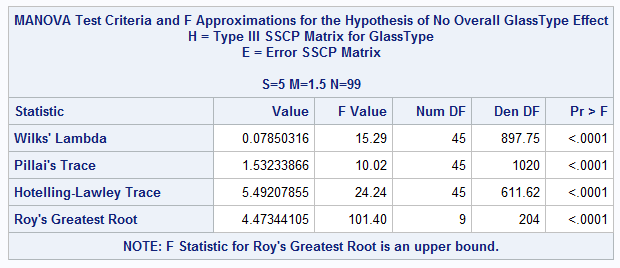
Because the p-value is less than 0.05, the Box M-test says that there is a significant difference in the covariance matrices for each glass type. The covariances are not equal.

### Assumptions Discussion

Since both the multivariate normality and equality of covariance tests are violated in this data set, running a MANOVA may not give us completely accurate results.

To better prepare the data, a multivariate Box-Cox transformation could be done to make things more normal. Testing for and handling outliers would also give some improvements, as well as increasing the sample size in order for the Central Limit Theorem to take effect. As it is rare that all assumptions are satisfied and MANOVA can still be robust for parametric tests, for the purposes of this study, it will be acknowledged that the MANOVA assumptions may need some additional work can be thought of as a limitation. A MANOVA will still be run to see what the data might show and/or explain.

### One Way MANOVA using Ref\_Index Sodium Magnesium Aluminum Silicon Potassium Calcium Barium Iron



There is a significant difference in the mean vectors for the 7 types of glass. Since the One-Way MANOVA shows this significant difference, individual Tukey tests can be run on significant variables to see how each of them impact the different Glass Types.

#### Refractive Index

|  |  |  |
| --- | --- | --- |
| Test | Output | Interpretation |
| F (Significance) |  | The p-value of F for Refractive Index is above 0.05 so it is not considered significant. |

#### Sodium

|  |  |  |
| --- | --- | --- |
| Test | Output | Interpretation |
| F (Significance) |  | The p-value of F for Sodium is less than 0.05, so it is considered a significant variable and a Tukey test is performed. |
| Tukey |  | Significant difference between:   * Glass Type 3 and 6 * Glass Type 1 and 6 * Glass Type 2 and 6 * Glass Type 5 and 6 * Glass Type 3 and 7 * Glass Type 1 and 7 * Glass Type 2 and 7 * Glass Type 5 and 7 |

This shows Sodium levels may be a good indicator when trying to classify as Tableware or Headlamps.

#### Magnesium

|  |  |  |
| --- | --- | --- |
| Test | Output | Interpretation |
| F (Significance) |  | The p-value of F for Magnesium is less than 0.05, so it is considered a significant variable and a Tukey test is performed. |
| Tukey |  | Significant difference between:   * Glass Type 1 and 2 * Glass Type 1 and 6 * Glass Type 1 and 5 * Glass Type 1 and 7 * Glass Type 2 and 6 * Glass Type 2 and 5 * Glass Type 2 and 7 * Glass Type 3 and 6 * Glass Type 3 and 5 * Glass Type 3 and 7 |

This shows Magnesium levels may be a good indicator when trying to classify between float processed building windows, non-float processed building windows, and float processed vehicle windows. Since we know there were not any non-float processed vehicle windows in the data, Magnesium may be considered a key differentiator between building and vehicle windows in general.

#### Aluminum

|  |  |  |
| --- | --- | --- |
| Test | Output | Interpretation |
| F (Significance) |  | The p-value of F for Aluminum is less than 0.05, so it is considered a significant variable and a Tukey test is performed. |
| Tukey |  | Significant difference between:   * Glass Type 1 and 7 * Glass Type 2 and 7 * Glass Type 3 and 7 * Glass Type 6 and 7 * Glass Type 1 and 5 * Glass Type 2 and 5 * Glass Type 3 and 5 * Glass Type 6 and 5 * Glass Type 1 and 2 |

This shows Aluminum levels may be a good indicator when trying to classify as Containers or Headlamps. It is also a good way to tell the difference between float and non-float processed window glass.

#### Silicon

|  |  |  |
| --- | --- | --- |
| Test | Output | Interpretation |
| F (Significance) |  | The p-value of F for Silicon is less than 0.05, so it is considered a significant variable and a Tukey test is performed. |
| Tukey |  | Significant difference between:   * None |

Even though Silicon is a significant variable, it is not a good indicator alone to differentiate glass type. This is most likely a side of correlation with another element.

#### Potassium

|  |  |  |
| --- | --- | --- |
| Test | Output | Interpretation |
| F (Significance) |  | The p-value of F for Potassium is less than 0.05, so it is considered a significant variable and a Tukey test is performed. |
| Tukey |  | Significant difference between:   * Glass Type 1 and 5 * Glass Type 2 and 5 * Glass Type 3 and 5 * Glass Type 6 and 5 * Glass Type 7 and 5 |

This shows Potassium levels may be a good indicator to identify Container glass.

#### Calcium

|  |  |  |
| --- | --- | --- |
| Test | Output | Interpretation |
| F (Significance) |  | The p-value of F for Calcium is less than 0.05, so it is considered a significant variable and a Tukey test is performed. |
| Tukey |  | Significant difference between:   * Glass Type 1 and 5 * Glass Type 7 and 5 |

This shows Calcium levels may be a good indicator when trying to classify Container glass compared to float processed building windows or headlamps.

#### Barium

|  |  |  |
| --- | --- | --- |
| Test | Output | Interpretation |
| F (Significance) |  | The p-value of F for Barium is less than 0.05, so it is considered a significant variable and a Tukey test is performed. |
| Tukey |  | Significant difference between:   * Glass Type 1 and 7 * Glass Type 2 and 7 * Glass Type 3 and 7 * Glass Type 5 and 7 * Glass Type 6 and 7 |

This shows Barium levels may be a good indicator when trying to classify Headlamps.

#### Iron

|  |  |  |
| --- | --- | --- |
| Test | Output | Interpretation |
| F (Significance) |  | The p-value of F for Iron is less than 0.05, so it is considered a significant variable and a Tukey test is performed. |
| Tukey |  | Significant difference between:   * Glass Type 2 and 7 |

This shows iron levels may be a good indicator when trying to tell the difference between non-float processed building windows and headlamps.

#### Contrasts

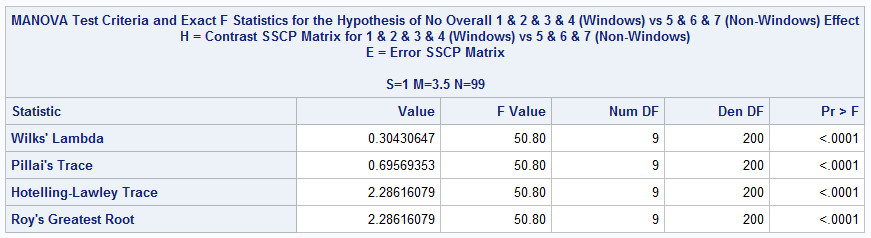
Taking a closer look at the levels of the data, deeper comparisons can be made in order to uncover more information. An effective way to do this is by using contrasts in conjunction with MANOVA analysis. The organization of information below sets up the proper logic to draw conclusions of how to create subsets of the data to further evaluate differences in mean vectors.

* Window glass (building windows and vehicle windows)
  + float processed
    - building windows
    - vehicle windows
  + non-float processed
    - building windows
    - vehicle windows
* Non-window glass
  + containers
  + tableware
  + headlamps

For contrasts, according to the data organization above, the following comparisons fit well with the class distributions and logic:

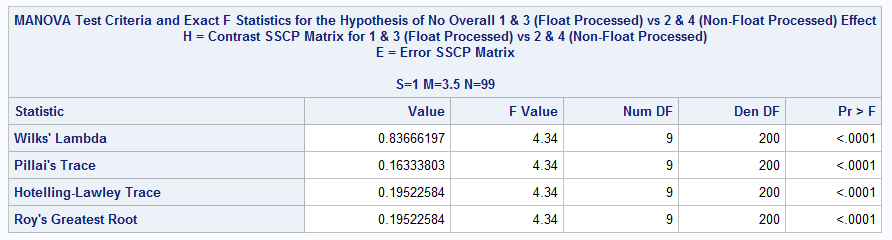
* window vs non-window
* float processed vs non-float processed
* building vs vehicle

##### Window vs. Non-Window: Is there a significant difference when comparing windows (either building or vehicle) and non-windows?



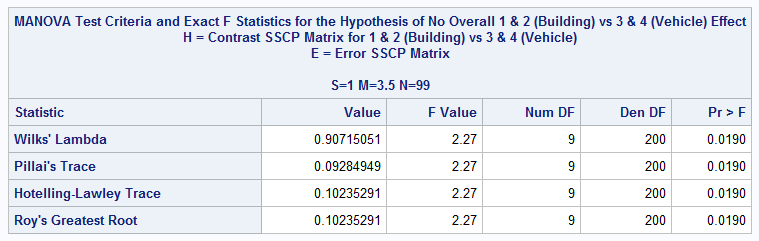
Because the p-values for F of all four listed statistics are less than 0.05, yes, there is a significant difference in the mean vectors between 1 & 2 & 3 & 4 vs 5 & 6 & 7. There is a significant difference in the mean vectors for building windows versus non-windows.

##### Float Processed vs. Non-Float Processed: Is there a significant difference when comparing float processed vs non-float processed glass?



Because the p-values for F of all four listed statistics are less than 0.05, yes, there is a significant difference in the mean vectors between 1 & 3 vs 2 & 4. There is a significant difference in the mean vectors for float processed vs non-float processed glass.

##### Building vs. Vehicle: Is there a significant difference when comparing building vs vehicle glass?



Because the p-values for F of all four listed statistics are less than 0.05, yes, there is a significant difference in the mean vectors between 1 & 2 vs 3 & 4. There is a significant difference in the mean vectors for building versus vehicle glass. Since these p-values all show as 0.0190 as opposed to <0.0001 for the previous two comparisons, it may be noted that there is very slightly less of a difference between building vs vehicle when compared to the <0.0001 output of window vs non-window and float-processed vs non-float processed.

## Discriminant Analysis

A discriminant analysis was performed on the data to determine how well new observations could be classified according to their classification groups as well as individual glass types. Discriminant analysis is a method of statistical analysis used to classify unknown observations into known groups based on measured characteristics, in the case of the data, these characteristics are the elemental compounds and the refractive index.

### Assumptions Discussion

Similar to MANOVA, discriminant analysis follows the assumptions that the data are multivariate normal, the variables are independent, there are equal variance and covariance matrices, and that the population means are significantly different. Also, like MANOVA, the assumptions of equal variance and covariance matrices, as well as, the normality of the data are violated. However, quantitative discriminant analysis is robust to non-normality according to a paper by R. Clarke, William, A. Lachenbruch, Peter, and Broffitt, Barabara titled “Communications in Statistics - Theory and Methods: How normality affects the quadratic discriminant function.” In the case of unequal variances, the use of Bartlett’s test determines whether or not linear or discriminant analysis is appropriate.

### Method

In running discriminant analysis, first there needs to be a testing set as well as a training set. Then the prior probabilities must be determined, then based on the frequencies of each type of glass. To determine whether or not the data for each glass type can be classified according to their glass type, the use of linear or quadratic analysis must be determined by means of the Bartlett test. This test is used to determine whether or not the variance-covariance matrices or homogenous for all populations involved. When testing for homogenous variance-covariance matrices, we found the p-values to be less than 0.001 and rejected the null hypothesis for all analyses, meaning that quadratic discriminant analysis was appropriate for all test.

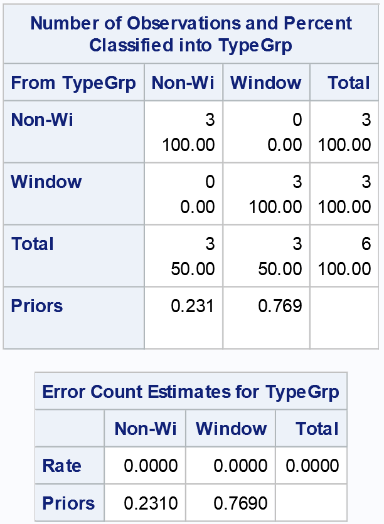
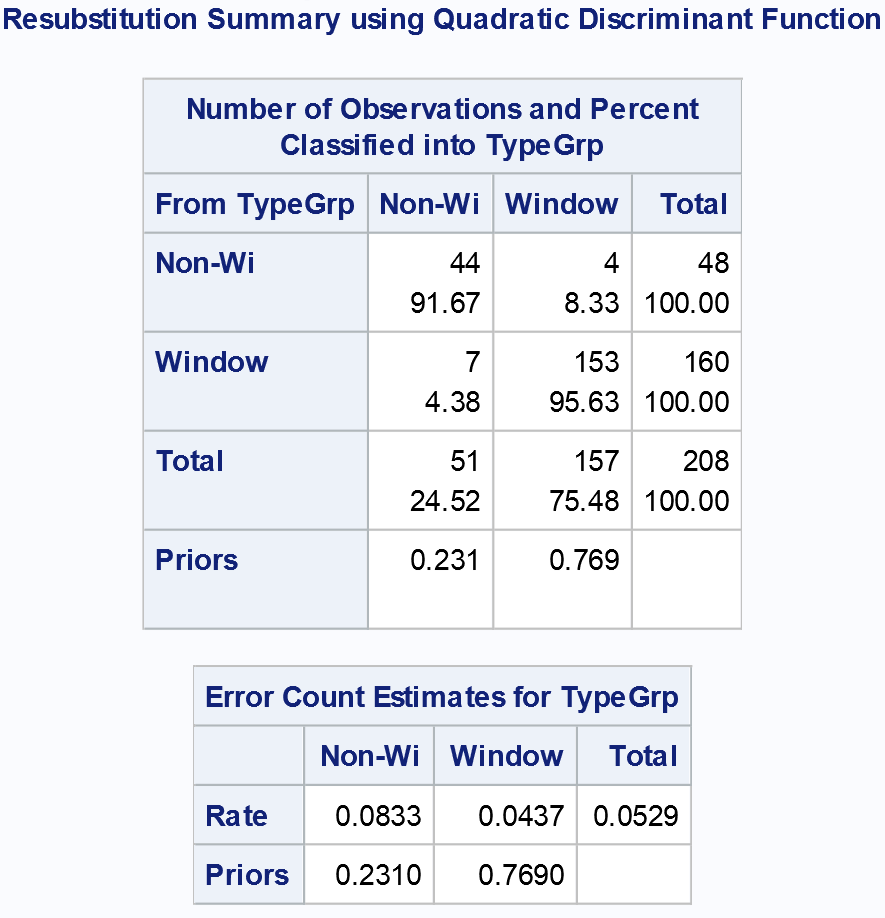
A testing dataset was created by removing one observation from each glass type for a total of 6 observations. Once the quadratic discriminant analysis functions were generated, the functions were evaluated via the resubstitution method. Unfortunately, with quadratic discriminant analysis, SAS does not produce functions for evaluation. Instead, the calibration data of the analysis was used to classify the test data. This testing set was then used to evaluate the accuracy of the quadratic functions.

### Results

The results of the QDA for the group classifications can be seen in Figures 4, 5, and 6 on the following page.

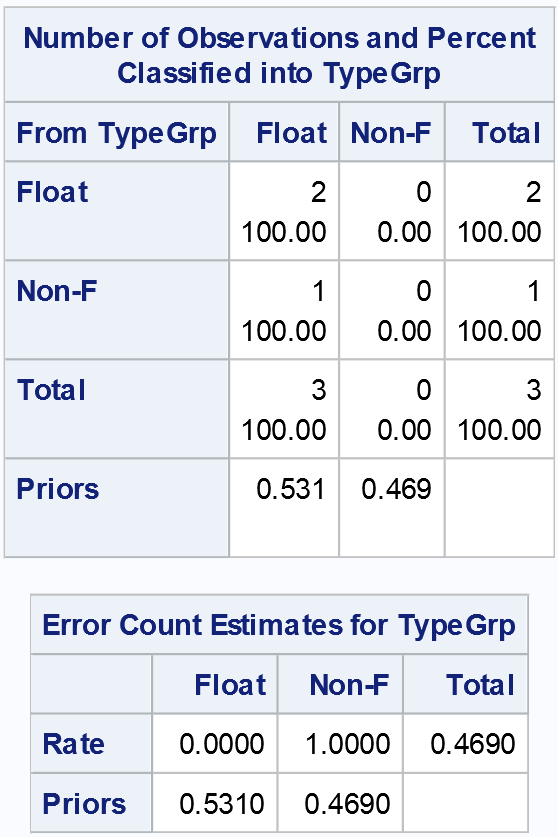
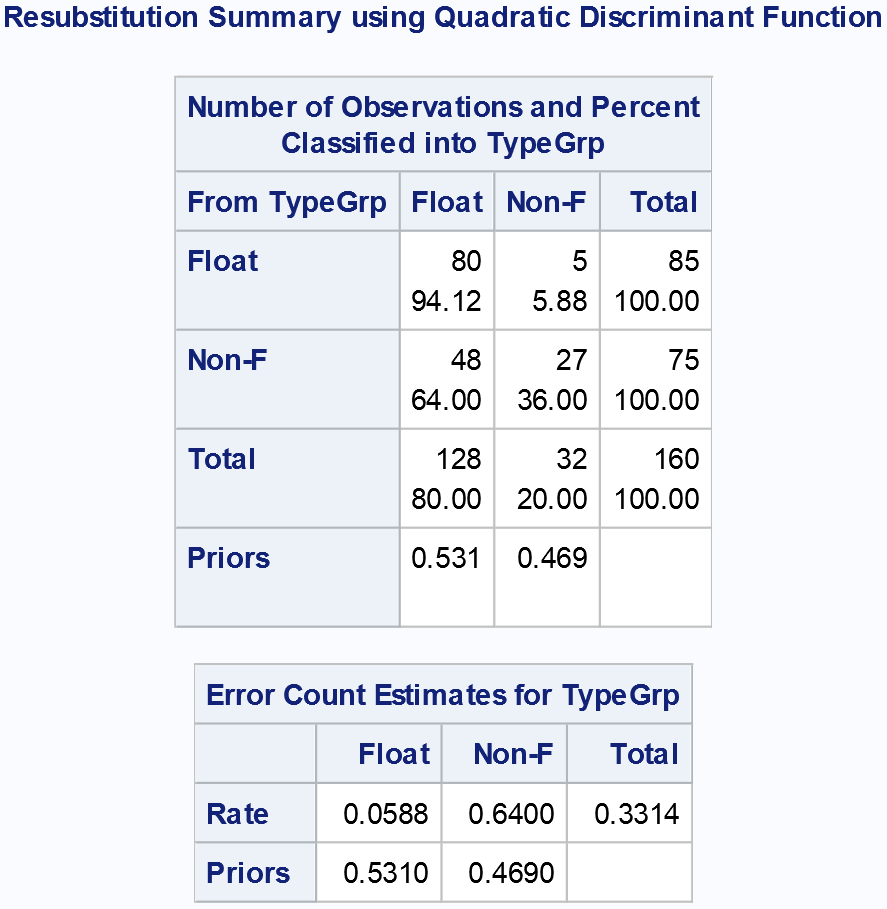
Window vs. Non-Window

Figure 4. Classification Summary of QDA of Window and Non-Window glass types for the training and testing data sets.



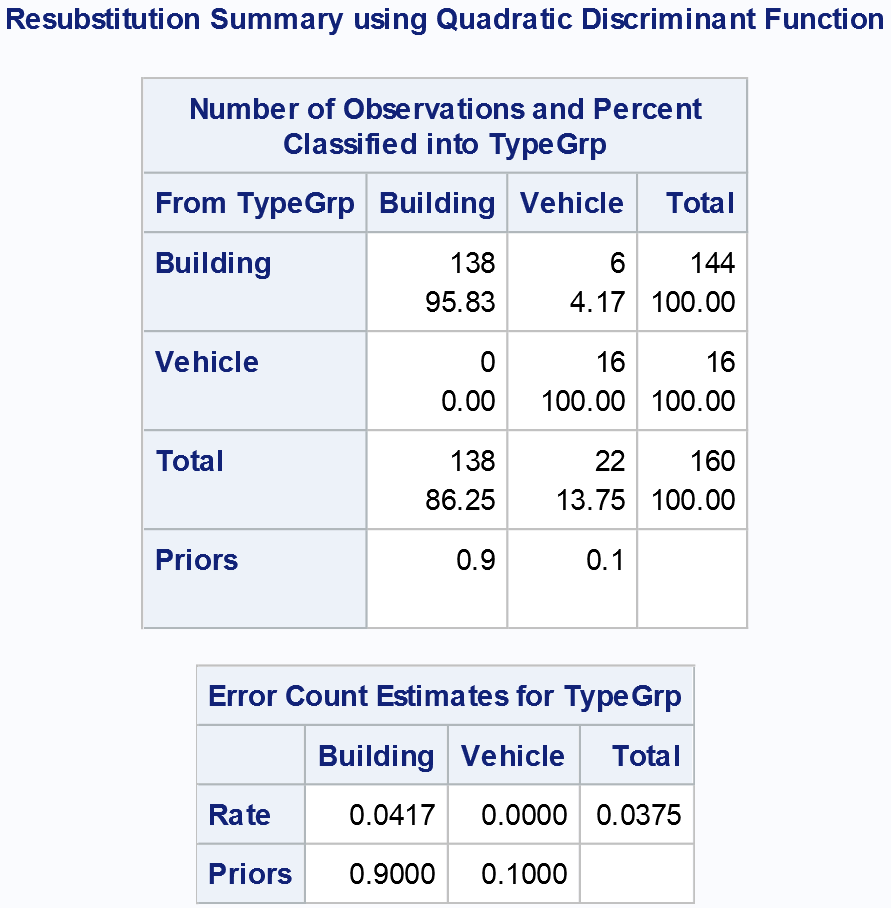
Float Processed vs. Non-Float Processed

Figure 5. Classification Summary of QDA of Float Processed and Non-Float Processed glass types for the training and testing data sets.



Building vs. Vehicle

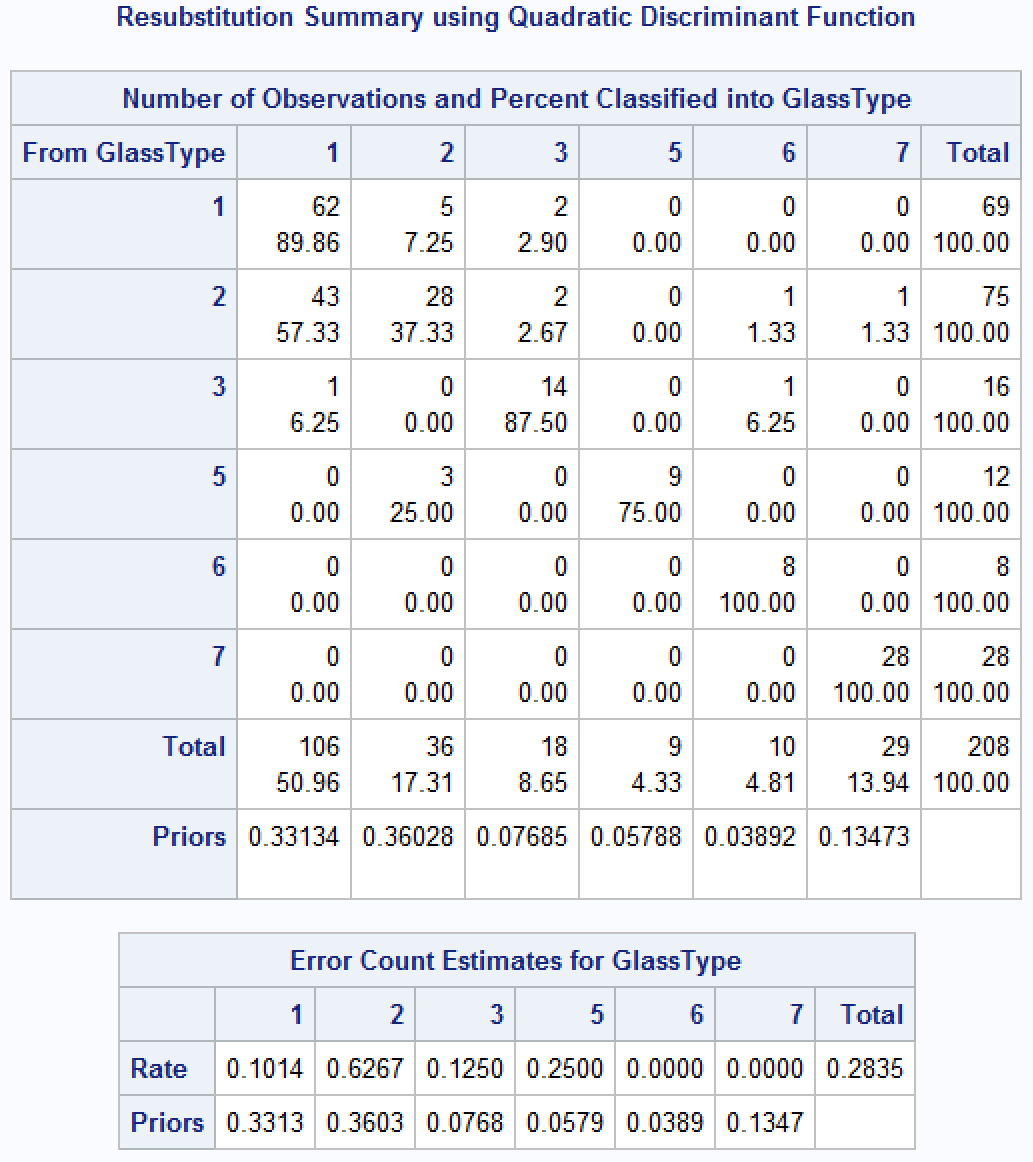
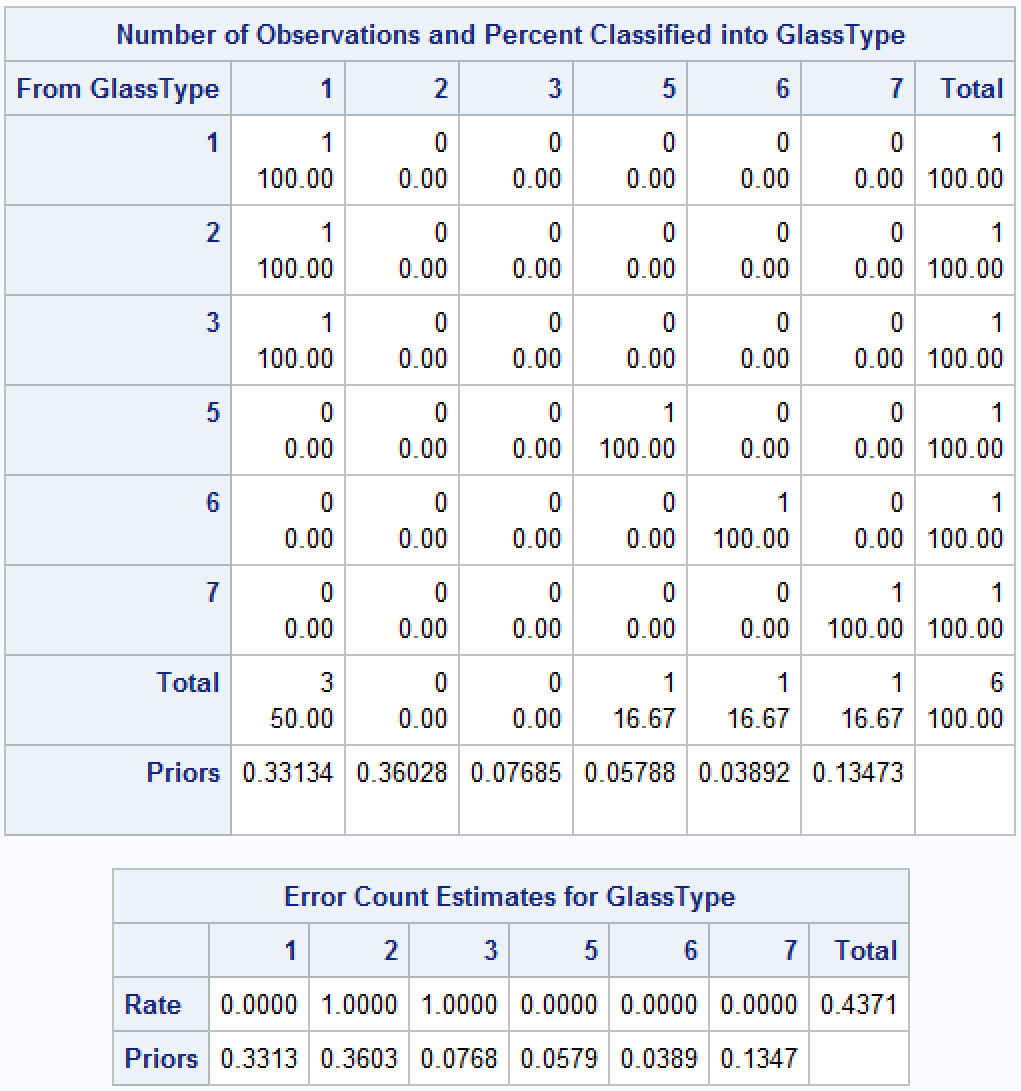
Figure 6. Classification summary of QDA of Building and Vehicle glass types for the training and testing data sets.

Differentiation between Window vs. Non-Window as well as Building vs. Vehicle proved to be highly successful with overall error rates of 5.3% and 3.8% respectively while Float vs. Non-Float Processed had a much higher error rate of 33.1%. When applied to the testing dataset, the classifications were consistent with 0% and 10% for Window vs. Non-Window and Building vs. Vehicle and 46.9% for Float vs. Non-Float.

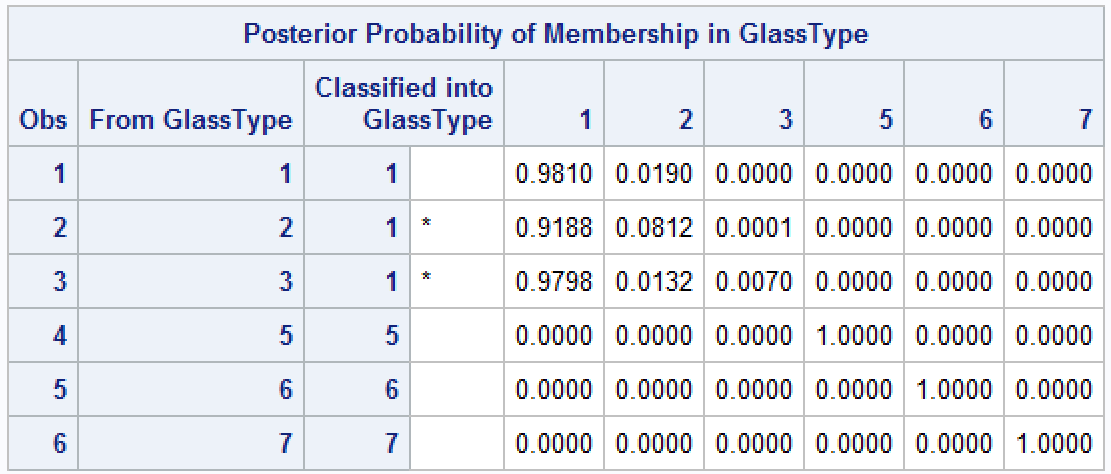
QDA was then applied to the data in its entirety, the results of the classification can be seen in Figure 7.

Figure 7. Classification summary of QDA of Building and Vehicle glass types for the training and testing data sets.

With an error rate of 28%, the analysis classified tableware and headlamp glass with 100% accuracy, building and vehicle glass that is float processed with 89.9% and 87.5% accuracy, respectively. Container glass and building windows that are not float processed weren’t classified as well with 75% and 37.3% accuracy respectively.

Figure 8. Classification summary of QDA of Building and Vehicle glass types for the training and testing data sets.



Classification of the new observations can be seen in Figure 8. After classifying the new observations, it was found that the float processed building window glass, containers, tableware, and headlamps were all classified correctly while the non-float processed window glass and float processed vehicle glass were classified incorrectly as float processed building window glass.

## Conclusions

To conclude, our study sought to find which chemical combinations are more prevalent in different glass types and whether the chemical composition and refractive index determined the different types of glass in our dataset. A MANOVA found that there were significant differences in glass type by all eight elements in our dataset but not by refractive index. The summary of results by glass type can be found in Table 5 and answers which chemical elements are more prevalent, which can reveal themselves as different oxides.

Regarding sodium oxide, two of the glass types under the non-window category, tableware and headlamps, were significantly different from all vehicular and building window glass types. The other glass type under the non-window category, containers, was only significantly different from tableware and headlamps but not any of the window glass types (vehicular and building).

For magnesium oxide, all the non-window glass types were significantly different from the window glass types. Additionally, within the window glass types, non-float processed building glass was significantly different from float processed building glass.

For aluminum oxide, two of the non-window glass types, containers and headlamps, were significantly different from all window glass types as well as the non-window tableware glass type. In addition, within the window glass types, like for magnesium oxide, the non-float processed building glass was significantly different from float processed building glass.

For potassium oxide, all glass types were significantly different from containers. Therefore, potassium seems to be a good indicator for container glass types. For calcium oxide, container glass types were significantly different from headlamp glass and float-processed building window glass types. For barium oxide, all glass types were significantly different than headlamp glass. Therefore, barium oxide appears to be a good indicator for headlamp glass. Lastly, for iron oxide, the only significant difference appeared between non-float processed building window glass and headlamp glass.

Table 5: Summary of results by glass type

|  |  |  |  |
| --- | --- | --- | --- |
| *Glass Type* | *Prevalent oxides* | *Glass Type* | *Prevalent oxides* |
| *1. Float building window* | Sodium can distinguish from tableware and headlamps  Magnesium can distinguish from non-window and non-float building glass  Aluminum can distinguish from container and headlamps  Potassium can distinguish from containers  Calcium can distinguish from containers  Barium can distinguish from headlamp glass | ***5. Containers*** | Sodium can distinguish from tableware and headlamps  Magnesium can distinguish from all window types  Aluminum can distinguish from window and tableware  Potassium can distinguish from all window, tableware, and headlamp glass  Calcium can distinguish from headlamp and float building window glass  Barium can distinguish from headlamp glass |
| *2. Non-float building window* | Sodium can distinguish from tableware and headlamps  Magnesium can distinguish from non-window and float building glass  Aluminum can distinguish from container and headlamps  Potassium can distinguish from containers  Barium can distinguish from headlamp glass  Iron can distinguish from headlamp glass | ***6. Tableware*** | Sodium can distinguish from all window and container glass  Magnesium can distinguish from all window glass  Aluminum can distinguish from container and headlamps  Potassium can distinguish from containers  Calcium can distinguish from containers  Barium can distinguish from headlamp glass |
| *3. Float vehicle window* | Sodium can distinguish from tableware and headlamp  Magnesium can distinguish from non-window glass  Aluminum can distinguish from container and headlamps  Potassium can distinguish from containers  Barium can distinguish from headlamp glass | ***7. Headlamps*** | Sodium can distinguish from all window and container glass  Magnesium can distinguish from all window glass  Aluminum can distinguish from window glass and tableware  Potassium can distinguish from containers  Calcium can distinguish from containers  Barium can distinguish from all window, tableware and container glass  Iron can distinguish from non-float building window glass |

Our results seem to suggest that in answering what determines different types of glass, only chemical composition is needed. In one study, refractive index alone was sufficient to distinguish between the samples of headlight and windshield glass (Gamble & Kirk, 1943). The reason for that could be that even when looking at the refractive indices of the samples in that study, the differences in values between those two types of glass were very small. The authors, Gamble and Kirk, mentioned that the differences in values were "not larger than might occur at times as local variations due to manufacturing imperfections."(1943). They also used different techniques in addition to examining refractive index. Therefore, in our study there might have not been any significant differences found in refractive index because the values in our dataset ranged from 1.51 to 1.53. Also, other studies that used refractive index only looked at 2-3 different glass types and so any variation could have been sufficient when looking at fewer glass types (Terry, Van, & Lynch, 1983; Bridge, Powell, Steele, & Sigman, 2007).

To answer whether the chemical composition and refractive index determined the different types of glass in our dataset, a quadratic discriminant analysis was performed. Interestingly, this analysis found high success rates of classification between window vs non-window glass and between building vs vehicle glass but a lower success rate for classifying float vs non-float processed glass. This could be because the processing method is more of an observable property whereas chemical composition and refractive index are measurable properties (Bottrell, 2009). In this study we were hoping the measures of refractive index and weight percent of oxides would suffice in finding a difference. Plus, in our dataset, we had glass, i.e., tableware and headlamp glass, that had no indication regarding by what process it was created. Non-float processed vehicle glass was not included in the dataset at all which could also contribute to this.

## Limitations (if any) and Other Variables You Wish You Wish You Would Have Included in the Study

Regarding the limitations of this study, one of the main limitations was the sample size. In particular, when conducting MANOVA and discriminant analysis, the small number of cases for some of the levels of the variable 'Glass Type' hindered the analysis. Furthermore, there were no observations for level 4 of 'Glass Type' (I.e., there were no observations classified as "Vehicle windows that are not float processed"). In checking the assumptions of MANOVA and discriminant analysis, it was found that the multivariate normality assumption was violated; however, we add as a comment that the tests used in this study are robust in the presence of violations of the multivariate normality assumption.

Finally, the scope of this study is limited to analyses based on 6 distinct types of glass. Future studies might benefit from both larger samples and a more diverse collection of glass types.

## References Section

Bottrell, M. C. (2009). Forensic Glass Comparison: Background Information Used in Data Interpretation. *Forensic Science Communications, 11*(2). Retrieved May 19, 2018, from https://archives.fbi.gov/archives/about-us/lab/forensic-science-communications/fsc/april2009/review

Chemistry of glass. (2011, December 1). Retrieved from https://www.cmog.org/article/chemistry-glass

Finch, H. (2005). Comparison of the Performance of Nonparametric and Parametric MANOVA Test Statistics when Assumptions Are Violated. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 1*(1), 27-38.

Francis, Eden. (n.d.). Weight Percent. Retrieved from http://dl.clackamas.edu/ch105-04/weight.htm

Gamble, L., Burd, D. Q., & Kirk, P. L. (January 01, 1943). Glass Fragments as Evidence. A Comparative Study of Physical Properties. *Journal of Criminal Law and Criminology (1931-1951), 33,* 5, 416.

Glass manufacturing process. (n.d.) Retrieved from https://www.aisglass.com/glass-manufacturing-process

Terry, K. W., Van, R. A., Lynch, B. F., Government Chemical Laboratories (W.A.), Criminology Research Council (Australia), & Western Australian Institute of Technology. (1983). *Identification of small glass fragments for forensic purposes*. Perth, W.A.: The authors.

The story of glass. (n.d.) Retrieved from https://www.packagingsa.co.za/info-library/packaging-types/glass/the-story-of-glass/

Theogene, N., & Rajiv, A. (2015). Identification of Headlight and Windshield glass of Car. *Research Journal of Recent Sciences,4*((IYSC-2015)), 166-170. Retrieved May 19, 2018, from http://isca.in/rjrs/archive/v4/iIYSC-2015/28.ISCA-IYSC-2015-9FMS-01.php

UC Irvine Machine Learning Repository. (1987). *Glass Identification Database.* [Data file]. Retrieved from https://archive.ics.uci.edu/ml/machine-learning-databases/glass/

What elements are used to make glass? (2018, April 21). Retrieved from https://libanswers.cmog.org/faq/143934

Zetlein, M. B. (January 01, 1979). Glass classification by elemental composition: Numerical evaluation. *Forensic Science International, 13,* 55-64.

## Appendix

Multivariate Normality

%inc [\\tsclient\C\Users\cmoran\Desktop\multnorm.sas](file:///\\tsclient\C\Users\cmoran\Desktop\multnorm.sas);

data glass7;

set glass;

if GlassType=7;

run;

%multnorm(data=glass7,var=Ref\_Index Sodium Magnesium Aluminum Silicon Potassium Calcium Barium Iron,plot=mult)

Equality of Covariance Matrices

proc discrim pool=test data=glass;

class GlassType;

var Ref\_Index Sodium Magnesium Aluminum Silicon Potassium Calcium Barium Iron;

run;

One Way MANOVA – All vars included

proc glm data=glass;

class GlassType;

model Ref\_Index Sodium Magnesium Aluminum Silicon Potassium Calcium Barium Iron = GlassType;

manova h=GlassType / printe printh;

run;

One Way MANOVA with Contrasts for Windows vs Non-Windows – 1 & 2 & 3 & 4 vs 4 & 6 & 7

proc glm data=glass;

class GlassType;

model Ref\_Index Sodium Magnesium Aluminum Silicon Potassium Calcium Barium Iron = GlassType;

contrast '1 & 2 & 3 & 4 (Windows) vs 5 & 6 & 7 (Non-Windows)' GlassType 0.33 0.33 0.33 -0.33 -0.33 -0.33;

manova h=GlassType / printe printh;

run;

One Way MANOVA with Contrasts for Float Processed vs Non-Float Processed– 1 & 3 vs 2 & 4

proc glm data=glass;

class GlassType;

model Ref\_Index Sodium Magnesium Aluminum Silicon Potassium Calcium Barium Iron = GlassType;

contrast '1 & 3 (Float Processed) vs 2 & 4 (Non-Float Processed)' GlassType 0.5 -1.0 0.5;

manova h=GlassType / printe printh;

run;

One Way MANOVA with Contrasts for Building vs Vehicle– 1 & 2 vs 3 & 4

proc glm data=glass;

class GlassType;

model Ref\_Index Sodium Magnesium Aluminum Silicon Potassium Calcium Barium Iron = GlassType;

contrast '1 & 2 (Building) vs 3 & 4 (Vehicle)' GlassType 0.5 0.5 -1.0;

manova h=GlassType / printe printh;

run;

Discriminant Analysis

/\*Create Window and Non-Window Groups in new dataset glasswind\*/

**data** glasswind;

set glass;

if Type=**1** then TypeGrp='Window';

if Type=**2** then TypeGrp='Window';

if Type=**3** then TypeGrp='Window';

if Type=**5** then TypeGrp='Non-Window';

if Type=**6** then TypeGrp='Non-Window';

if Type=**7** then TypeGrp='Non-Window';

**run**;

/\*Obtain frequencies for priors\*/

**proc** **freq** data=glasswind;

tables TypeGrp;

**run**;

/\*Discriminant Analysis for Window vs. Non-Window\*/

**proc** **discrim** data=glasswind pool=test;

class Typegrp;

var Ri Na Mg Al Si K Ca Ba Fe;

priors 'Non-Wi'=**0.238** 'Window'=**0.762**;

**run**;

/\*Remove Non-Window data in new dataset glass2\*/

**data** glass2;

set glass;

if Type ne **5**;

if Type ne **6**;

if Type ne **7**;

**run**;

/\*Create Float and Non-Float Groups in new dataset glassfnf\*/

**data** glassfnf;

set glass2;

if Type=**1** then TypeGrp='Float';

if Type=**3** then TypeGrp='Float';

if Type=**2** then TypeGrp='Non-Float';

**run**;

/\*Obtain frequencies for priors\*/

**proc** **freq** data=glassfnf;

tables TypeGrp;

**run**;

/\*Discriminant Analysis for Float vs. Non-Float\*/

**proc** **discrim** data=glassfnf pool=test;

class TypeGrp;

var Ri Na Mg Al Si K Ca Ba Fe;

priors 'Non-F'=**0.466** 'Float'=**0.534**;

**run**;

/\*Create Building and Vehicle Groups in new dataset glassbv\*/

**data** glassbv;

set glass2;

if Type=**1** then TypeGrp='Building';

if Type=**2** then TypeGrp='Building';

if Type=**3** then TypeGrp='Vehicle';

**run**;

/\*Obtain frequencies for priors\*/

**proc** **freq** data=glassbv;

tables TypeGrp;

**run**;

/\*Discriminant Analysis for Building vs. Vehicle\*/

**proc** **discrim** data=glassbv pool=test;

class Typegrp;

var Ri Na Mg Al Si K Ca Ba Fe;

priors 'Vehicle'=**0.104** 'Building'=**0.896**;

**run**;

/\*Obtain frequencies for prior probabilities\*/

**proc** **freq** data=glass;

tables GlassType;

**run**;

/\*Discriminant Analysis with test for equal covariance matrices using prior probabilities and the resubstitution method. This also creates a data set of the calibration information to be used for classifying additional observations\*/

**proc** **discrim** data=glass method=normal outstat=glassstat pool=test;

class GlassType;

priors '1'=**0.332** '2'=**0.361** '3'=**0.077** '5'=**0.058** '6'=**0.039** '7'=**0.135**;

var Ref\_Index Sodium Magnesium Aluminum Silicon Potassium Calcium Barium Iron;

**run**;

/\*Classification of new observations\*/

**proc** **discrim** data=glassstat testdata=glasstest testout=tout2 testlist;

class GlassType;

var Ref\_Index Sodium Magnesium Aluminum Silicon Potassium Calcium Barium Iron;

**run**;

* Assign responsibilities
  + Abstract – Edith
  + Background and purpose – Edith
  + Sample Collection Techniques, Methodology, & Design – Blaise
  + List of Variables, Responses, and Codebook of Variables in the Analysis (if applicable) - Blaise
  + Descriptive Statistics (Added by Blaise on 5/26)
  + The Body of the Report
    - Discriminant Analysis - Steven
    - MANOVA – Cait
  + Summary/Conclusions & Recommendations – Edith
  + Limitations (if any) and Other Variables You Wish You Wish You Would Have Included in the Study - Blaise
  + References – Edith
  + Appendix – Steven and Cait
  + Powerpoint - Cait, Blaise, Edith, Steven
  + Handout – Blaise, Edith
  + Presentation - All
* **Powerpoint -** <https://maildepaul-my.sharepoint.com/:p:/g/personal/cmoran21_mail_depaul_edu/ETnT3ksKl2hGqTVE8SzEpkEBw3KoiLg5ciVvvFFVRZHCyA?e=eZdlRm>

**Dataset:**

* <https://archive.ics.uci.edu/ml/datasets/Glass+Identification>