

BOSTON UNIVERSITY  
GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

**PRODUCING STONE AND STATE:  
THE INTERSECTION OF DOMESTIC AND INSTITUTIONAL  
ECONOMIES IN CLASSIC MAYA SOCIETY**

by

**MARY E. CLARKE**

B.F.A., Savannah College of Art and Design, 2013

Submitted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

2020

© Copyright by  
MARY E. CLARKE  
2020

## Appendix F

In collaboration with consultants from Boston University's Master of Science in Statistical Practice (MSSP) program, an exploratory data analysis and a statistical data analysis were produced. These analyses and the tests therein aimed to measure the diversity of consumed low, mid, and high value resources by households included in this study.

### *Exploratory Data Analysis*

An exploratory data analysis (EDA) was used to visualize the data and household comparison (FIGURE F.1). The aim of this analysis was to view the number of goods each household consumed at each designated value. In the plot, the red, green, and blue lines represent the number of low, mid, and high value resource varieties. Each node represents a household, which is labeled along the x-axis. The dashed lines represent 2 standard deviations away from the mean and the thin solid lines indicate the mean. The standard deviation lines provide an estimation of the variation of the data and gives an idea of how far each household is away from the mean. As illustrated in FIGURE F.1, all points lie within 2 standard deviations from the mean, suggesting that they may all be similar to each other and that one household is not significantly different than the others. Overall, there appears to be a greater similarity among the low value resources, whereas the high value resources present higher variability.

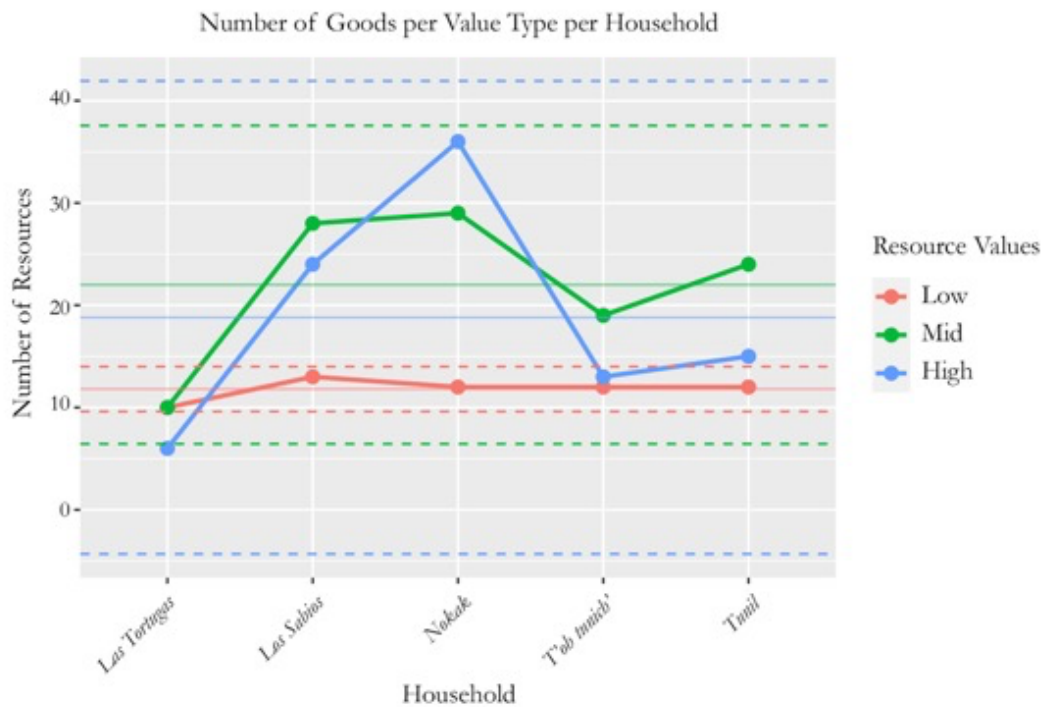


FIGURE F.1. The distribution of consumed resources by value and household.

The same plots are presented with the independent and limited networks excluded (FIGURE F.2). All plots lie within the 2 standard deviations, which suggests that they may all be similar to each other. Akin to the results from the first plot (FIGURE F.1), low value resources are consumed in a similar diversity, whereas there is more variability in the amounts of high value resource varieties consumed by each household. Additionally, the curves in this second plot appear smoother than the first plot, suggesting that when independent and limited networks are excluded, household consumption patterns are generally more similar.

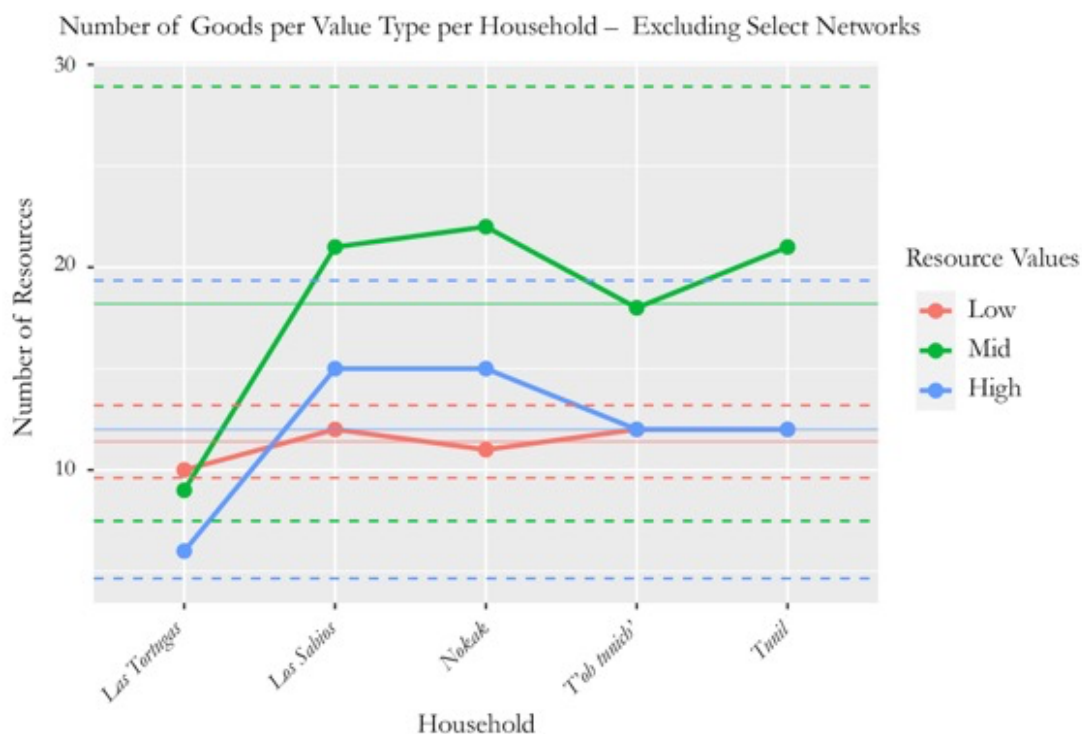


FIGURE F.2. The distribution of consumed resources by value and household when independent and limited exchange networks are excluded.

Following methods employed by Garraty (2009), the Brainerd-Robinson (BR) coefficient or the H-score, was considered. However, because each household get one h-score, no statistical comparisons could be made, and conclusions are exploratory. With this analysis, the aim is to compare the individual household units with the aggregate (or “expected”) distribution for all households combined. Larger values of the index indicate a greater distance between the unit and the aggregate distribution, which would then indicate a heterogeneous sample. Smaller values indicate more uniformity and therefore a homogeneous sample. In the case of the present data, the proportional of each resource variety for each household was compared to the overall proportion of the corresponding resource variety within value categories. That is, if we are looking at the low valued goods,

$$P_{value=low,household=Nokak} = \frac{\text{number of low value goods in Nokak}}{\text{number of goods in Nokak}} \times 100$$

and the expected proportion of low valued goods is

$$\bar{p}_{value=low} = \frac{\text{number of low value goods in all households}}{\text{number of goods in total}} \times 100$$

Using these proportions, the scaled BR coefficient for a household is calculated by summing over the differences for each value type of good and dividing by 200:

$$BR_{household} = \frac{\sum_{value} (|\bar{p}_{value} - P_{value,household}|)}{200}.$$

The scaled BR coefficient is illustrated (FIGURE F.3) alongside the 2 standard deviation lines away from the mean. Every household's assemblage lies within the 2 standard deviations, suggesting that there may not be a significant difference between them. It is also noted that the *Los Sabios* household has the most homogenous assemblage in terms of its consumption patterns, relative to the expected consumption pattern. Conversely, the *Las Tortugas* household is the most heterogenous. However, looking at the magnitudes of the scaled BR coefficients, there does not seem to be much difference between these households.

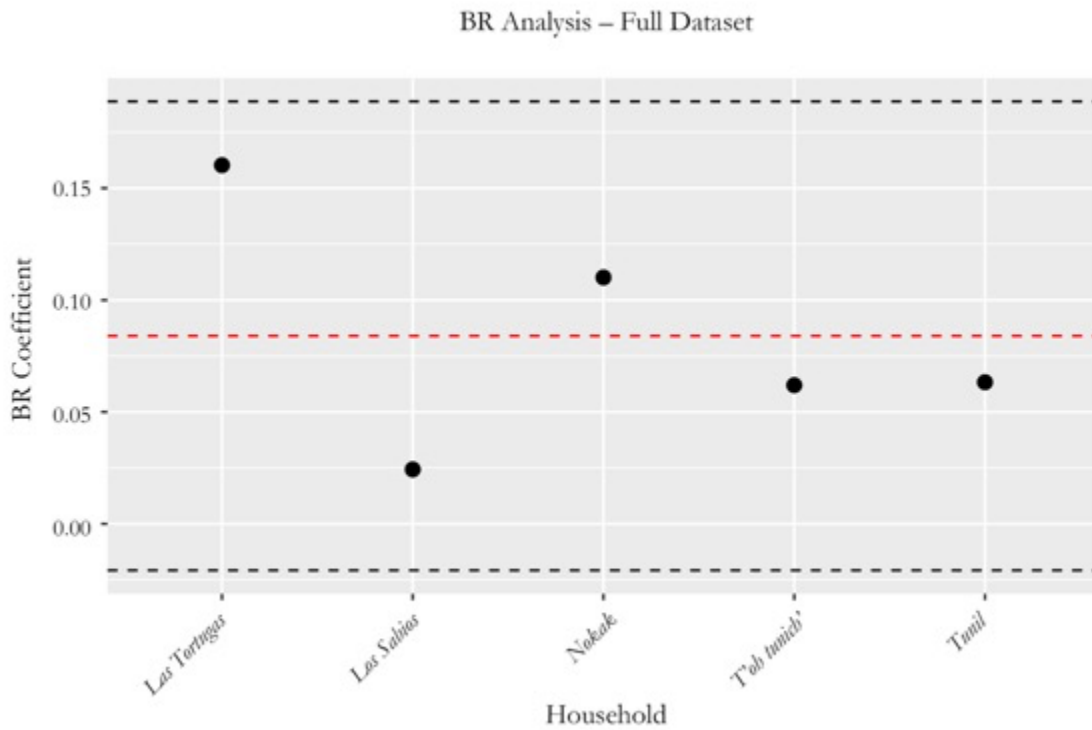


FIGURE F.3. The BR coefficient distributions for households.

When the independent and limited exchange networks are excluded (FIGURE F.4), the coefficients decrease, which suggests they become more homogenous. All nodes are within two standard deviations from the mean, but the *Las Tortugas* household presents the highest scale BR coefficient and is closer to the 2 standard deviation bar. This may suggest that *Las Tortugas* could be different from the other households, but the data are not sufficient to conclude this. However, the scaled BR coefficients for the households are more similar to each other in this plot (FIGURE F.4) than the previous (FIGURE F.3), which may suggest that there is no difference between the households when the independent and limited networks are excluded.

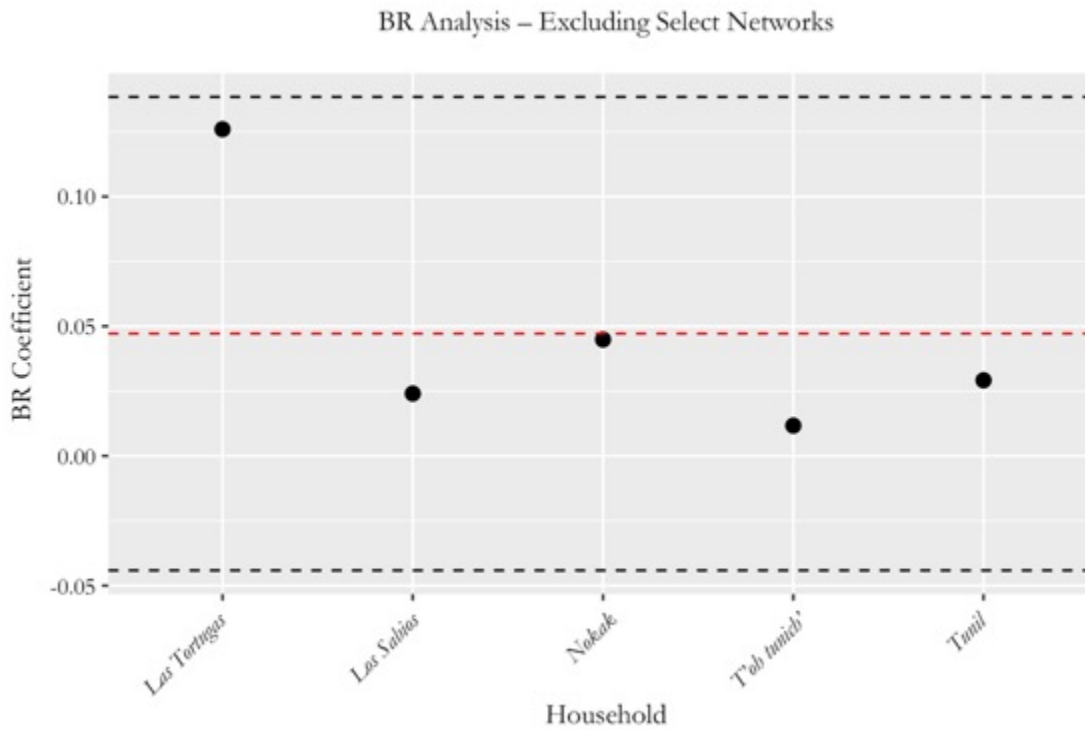


FIGURE F.4. The BR coefficient distributions for households with the independent and limited exchange networks excluded.

A paired dendrogram and heatmap was also produced for this dataset (FIGURE F.5). This plot illustrates two main components: a dendrogram on what clusters are formed from the data and a heatmap detailing how the clusters were determined. To generate the dendrogram located at the top of top of the plot, a clustering algorithm was used. Initially, each household was assigned to its own cluster, and then the algorithm proceeds iteratively, where, at each step, the two most similar clusters were grouped together until there is just a single cluster with all the households. The length of the lines suggests the distance or similarities of the clusters. The heatmap provides additional details on the sequence of the clustering as the rows and columns with similar colors are joined together first. Here, these colors represent the number of resource varieties consumed by each household within each value category.



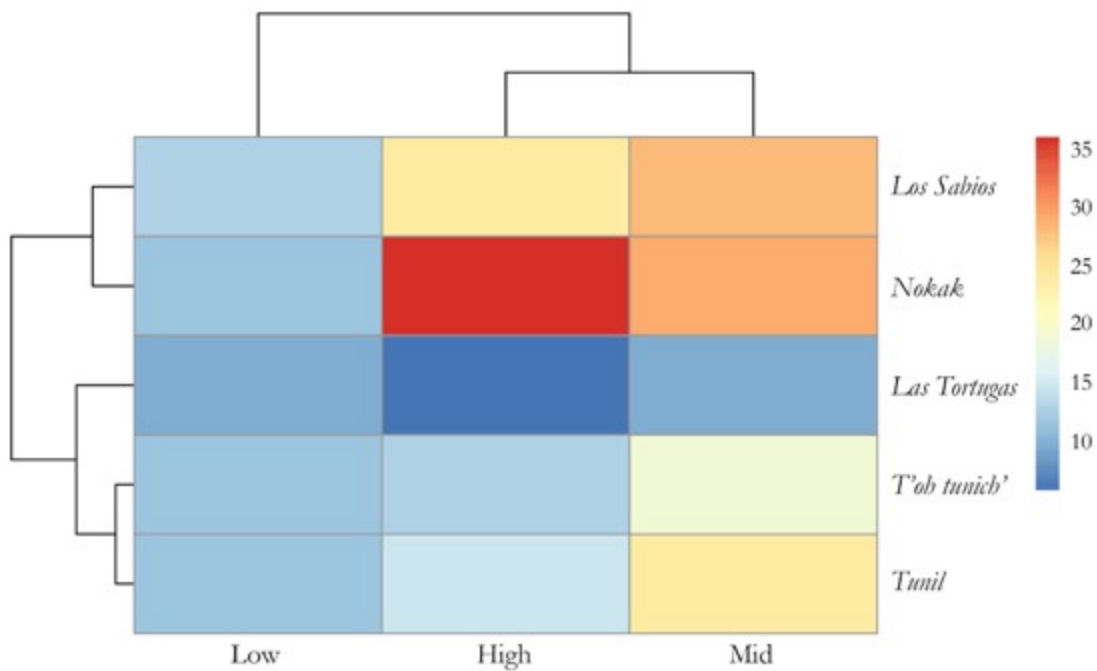


FIGURE F.5. A heatmap of the overall dataset.

From this plot, it is possible to observe that the *T'oh tunich'* and *Tunil* households are the most similar as they were joined first and because the lengths of their lines are the shortest. Additionally, *Las Tortugas* appears to be more alike these two households than the *Los Sabios* and *Nokak* households, which are grouped together first, which indicates that they are most similar to each other. Viewing vertical labels, the mid and high value groups are clustered together first, indicating that they are more similar to each other than the low value category.

### *Statistical Analysis*

Analyses of these data were conducted using multinomial tests that use log likelihood ratios. These tests aim to compare the observed distribution of goods to an expected distribution based on the grouping of the households. A pairwise test was first completed between each household by comparing the observed probabilities of goods

to the expected probability. If the compared households were from the same distribution (i.e., they share a similar purchasing power), then they were grouped together and present as essentially the same in terms of their consumption patterns related to low, mid, and high value categories.

More specifically, it was assumed that the number of different valued goods for each household come from a multinomial distribution with probability  $\pi_{household,value}$ , where *household* and *value* are the given household and its consumed value. For example,  $\pi_{Nokak,low}$  is the probability of the *Nokak* household consuming a low value good. As households are being paired and then compared, the null hypothesis is that the probabilities of the type of good consumed by the two households is equal:

$\pi_{household_1,value} = \pi_{household_2,value} = \pi_{value}$ , where  $\pi$  is the expected probability of the level of valued goods if the two households shared the same purchasing power. For each test, a p-value was computed, which is the probability of seeing the observed data given the assumption that the null hypothesis is true. If the p-value is low (usually below  $\alpha = 0.05$ ), then the null hypothesis may be eliminated because it is not likely to obtain the observed data if the null hypothesis is true. If the null hypothesis is rejected, then it is possible to conclude that the two households do not share the same purchasing power the same and are, thus, distinct from each other. However, similarity is concluded if the null hypothesis is not rejected.

Using the p-values obtained from the pairwise tests, a cluster dendrogram was generated (FIGURE F.6). This dendrogram was produced in a similar manner as the previous example (FIGURE F.5). However, it is important to note that log p-values are used here because they are generally more stable, and because the log of decimals is

negative, the negative log p-values are used so that positive numbers may be employed in the analysis.

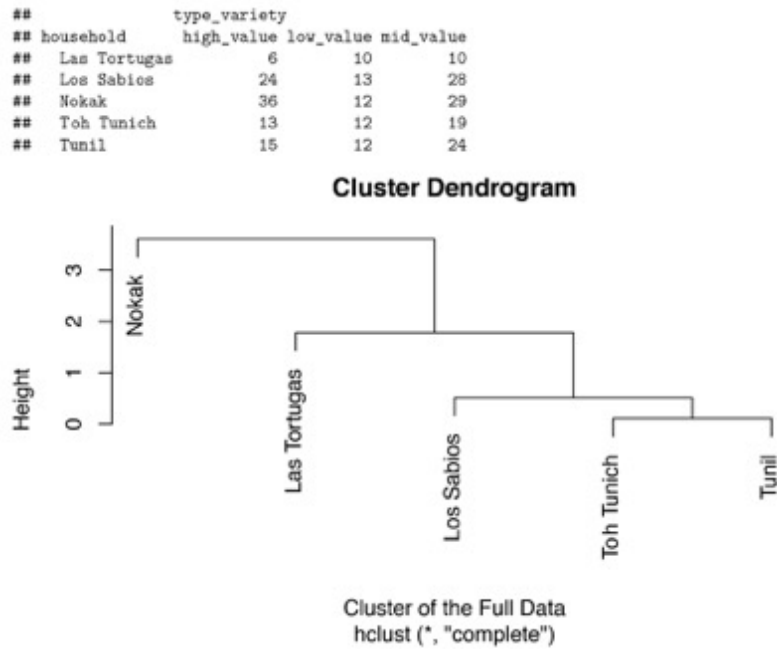


FIGURE F.6. The cluster dendrogram of the full dataset.

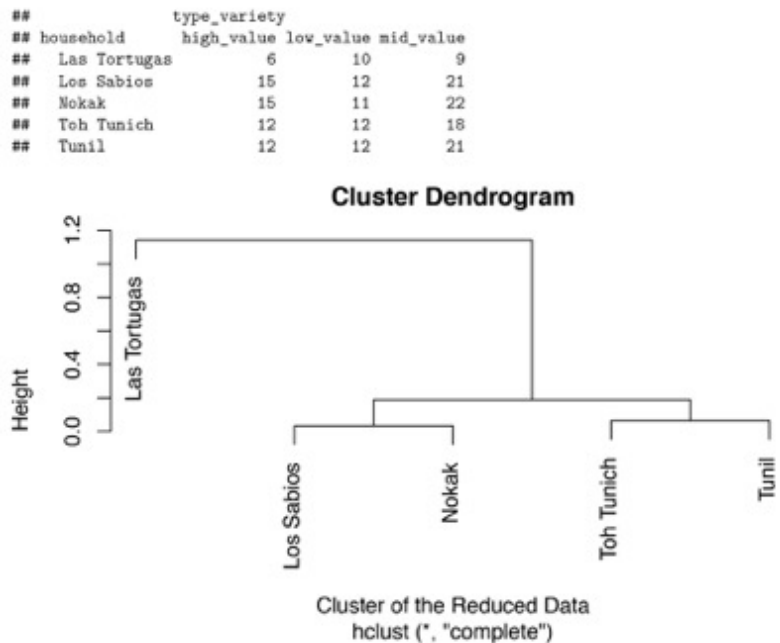


FIGURE F.7. The cluster dendrogram of the reduced dataset where the limited and independent networks were excluded.

In the two dendrograms (FIGURES F.6 AND F.7), the y-axis (the height) is a measure of closeness of either individual data points or clusters. Looking at the dendrograms based on the full dataset (FIGURE F.6), it appears to suggest that the *Nokak* and *Las Tortugas* households are different from the other three households, with *Nokak* being the most different. Looking at the reduced dataset (FIGURE F.7), distributions suggest that *Las Tortugas* is the main outlier. The height of this dendrogram indicates that *Las Tortugas* is quite different from the other clusters, but also that the range of the heights is smaller, which suggests that the five households appear to be more similar when independent and limited networks are excluded.

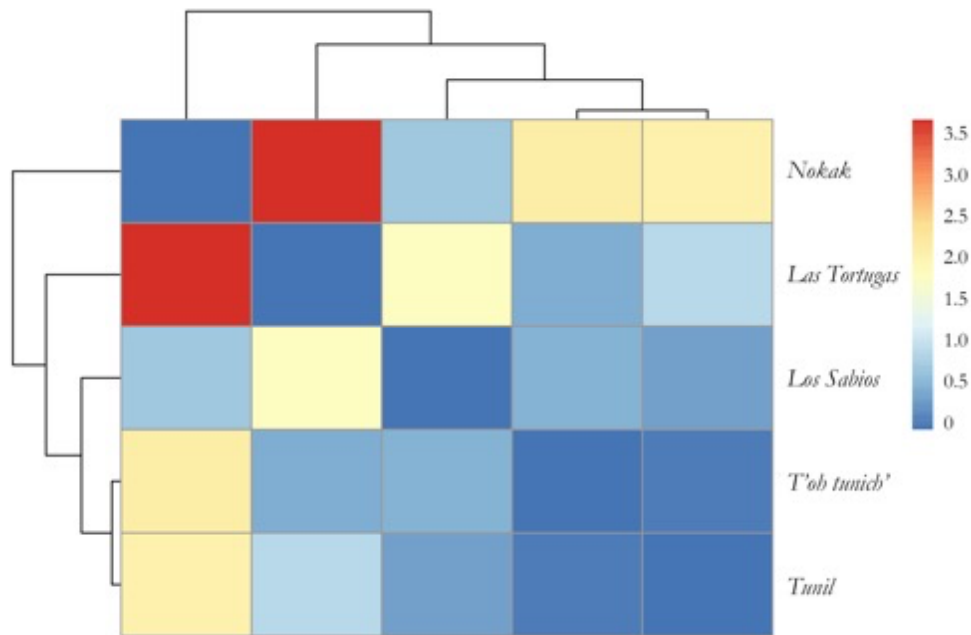


FIGURE F.8. The Heatmap of the overall dataset using p-values.

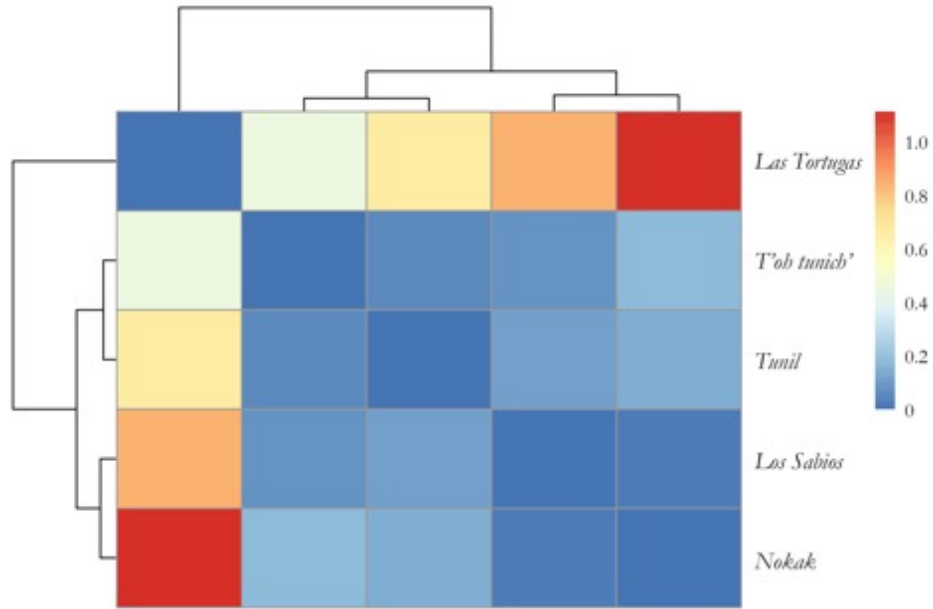


FIGURE F.9. The Heatmap of the reduced dataset using p-values.

The heatmaps provide additional details regarding how the dendrograms made the observed clusters (FIGURES F.8 AND F.9). Each color represents the negative log p-value for each pairwise test. The higher the log p-value, the redder the color will be. As negative log p-values were considered, a higher number means a smaller decimal point that indicates increased difference among the compared households according to the p-values. The full dataset (FIGURE F.8) illustrates that the *Nokak* and *Las Tortugas* households are the most different from the other three households. The reduced dataset (FIGURE F.9) illustrates that the *Las Tortugas* household has the most variety of colors indicating that it is different than the other households which all appear to be quite similar. These observations are consistent with the cluster dendrograms (FIGURES F.6 AND F.7).

The above dendrograms only provide an idea of how the households would be clustered. In order to make statistical claims, it needs to be determined whether the p-values obtained from the tests are statistically significant. To do this, the p-values will

be compared against a threshold. However, a family-wise corrected threshold is used to account for the multiple hypothesis approach to this analysis. Without this correction, there may be false positives. This is due to the definition of the p-value threshold, or significance level, which is the probability of rejecting the null hypothesis when the null hypothesis is true. When more hypotheses are tested, it becomes more likely that the null hypothesis will be rejected, which may lead to false positives. Here, a family-wise corrected threshold of  $-\log(0.05/n)$  is used, where 0.05 is the typical threshold used and  $n$  is the number of tests conducted, which is 10 in this case. Note that the negative log of this threshold is used so that it is on the same scale as the p-values. Additionally, the null hypothesis is rejected if the results are greater than the threshold because negative log p-values are employed.

Using the family-wise corrected p-values, graphs were produced to visualize household similarities and differences (FIGURE F.10). If there is an edge between the nodes, then the p-value is not statistically significant (i.e., the negative log p-value is greater than the threshold), which means that these households are similar. For both graphs generated using the results of the multinomial tests, they suggest that all five households are similar with each other because they are all connected. However, the thickness of the lines represents the negative log p-values for the tests between each pair. The thicker the line, the greater the log p-value, which means the difference between the households connected by this line is greater than if they were connected by a thinner line.

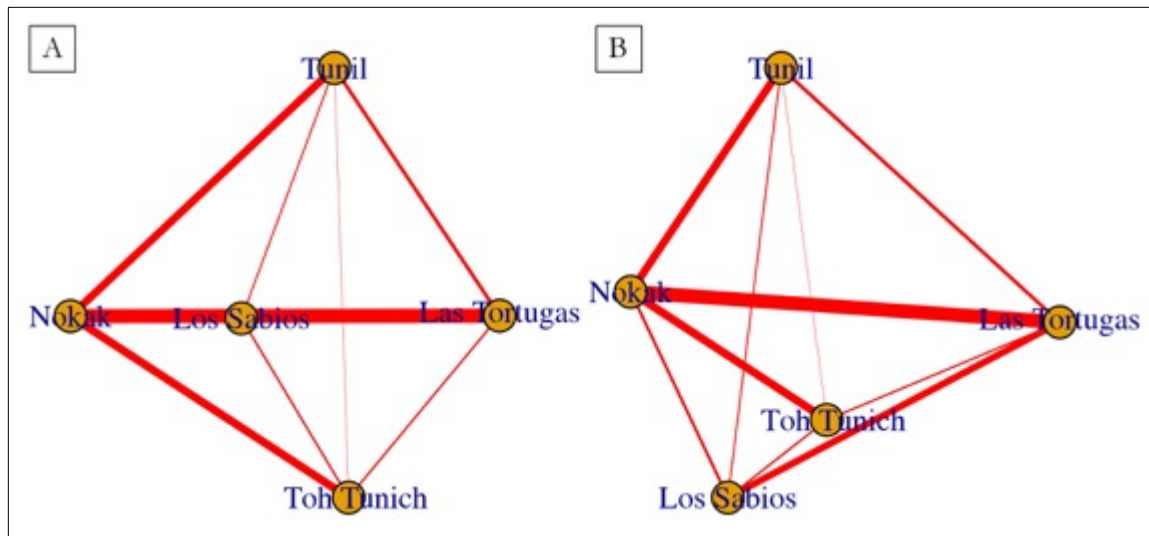


FIGURE F.10. The family-wise corrected p-values for the A) complete and B) reduced datasets.

One final multinomial hypothesis test was conducted using the groupings obtained from the pairwise tests and graph (FIGURE F.10). However, instead of a pairwise test, this test evaluates if all households were provisioning themselves through the same network, a marketplace exchange, as suggested by the grouping:  $\pi^{Nokak, value} = \pi^{Los Sabios, value} = \pi^{Las Tortugas, value} = \pi^{Toh tunich', value} = \pi^{Tunil, value} = \pi^{value}$ . Note, a normal p-value is used, this the null hypothesis is rejected if the p-value is below  $\alpha = 0.05$ . The results suggest that both the p-values for the full and reduced datasets are not statistically significant, which indicates that the households have a shared heterogenous pattern of consumption. The reduced dataset has a much higher p-value (0.94) than the p-value for the full dataset (0.23). Although both of the p-values are not statistically significant, these findings could support the EDA and the hypothesis that, once the independent and limited exchange networks are removed, the remainder of the resources consumed by households were acquired within a shared exchange network, such as a marketplace.

From this analysis, it is possible to conclude that the households were provisioning themselves primarily through a shared marketplace exchange network. However, there are a few limitations to the analysis. First, there is a lack of data points in which we are using to compare each household. If more attributes were included, greater differences between households could be identified. Second, the similarities in the low and mid value consumption patterns may be overshadowing the differences in high value resource consumption, which led to the conclusion that all households are similar. Although statistical claims to differences between households cannot be made, the EDA suggests that there may be some slight differences between these households.