

RAE-CEA 13P17

**RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O PROJETO:**

“Paisagem, território e história indígena na Aldeia Lalima, Miranda/MS”.

**Andressa Soreira**

**Victor Fossaluzza**

**São Paulo, Novembro de 2013**

**CENTRO DE ESTATÍSTICA APLICADA – CEA**  
**RELATÓRIO DE ANÁLISE ESTATÍSTICA**

**TÍTULO:** Relatório de análise estatística sobre o projeto: “Paisagem, território e história indígena na Aldeia Lalima, Miranda/MS”.

**PESQUISADORA:** Eduardo Bespalez

**ORIENTADORA:** Prof.<sup>a</sup> Dra. Fabíola Andréa Silva

**INSTITUIÇÃO:** Museu de Arqueologia e Etnologia da Universidade de São Paulo – MAE/USP

**FINALIDADE:** Tese de Doutorado

**RESPONSÁVEIS PELA ANÁLISE:** Andressa Soreira  
Victor Fossaluzza

**REFERÊNCIA DESTE TRABALHO:** FOSSALUZA, V. e SOREIRA, A. (2013). **Relatório de análise estatística sobre o projeto: “Paisagem, território e história indígena na Aldeia Lalima, Miranda/MS”.** São Paulo, IME-USP, 2013, (RAE-CEA 13P17).

## **FICHA TÉCNICA**

### **REFERÊNCIAS:**

BESPALEZ, E. (2009). **Levantamento Arqueológico e Etnoarqueologia na Aldeia Lalima, Mirando/MS.** Universidade de São Paulo.

BUSSAB, W. O., MORETTIN, P. A. (2013). **Estatística Básica.** 8<sup>a</sup> ed. Brasil. Saraiva.

GOODMAN, L. A., KRUSKAL, W. H. (1954). **Measures of association for cross classifications.** Journal of the American Statistical Association 49, no. 268, 732-764.

JOHNSON, R. A., & WICHERN, D. W. (2002). **Applied Multivariate Statistical Analysis.** Vol. 5. Upper Saddle River, NJ: Prentice Hall.

MARDIA, K. V., KENT, J. T., & BIBBY, J. M. (1979). **Multivariate Analysis (Probability and Mathematical Statistics).** London: Academic.

TORGO, L. (2006). **Introdução à Programação em R.** Portugal. Universidade do Porto.

### **PROGRAMAS COMPUTACIONAIS UTILIZADOS:**

Microsoft Excel for Windows (versão 2007)

Microsoft Word for Windows (versão 2007)

Software R (versão 2.14.2)

IBM SPSS Statistics (versão 20)

## **TÉCNICAS ESTATÍSTICAS UTILIZADAS:**

- Análise de Conglomerados (06:120)
- Análise Descritiva Multidimensional (03:020)
- Análise Descritiva Unidimensional (03:010)
- Associação e Dependência de Dados Qualitativos (06:020)
- Estimação Bayesiana (04:070)

## **ÁREA DE APLICAÇÃO:**

- Outros (14:990)

## Sumário

Resumo .....	6
1. Introdução.....	7
2. Descrição do estudo .....	8
3. Descrição das variáveis.....	9
4. Análise descritiva.....	9
5. Análise Inferencial .....	61
5.1 Análise de Conglomerados .....	61
5.2 Redes Bayesianas .....	75
6. Conclusões.....	81
Apêndice A.....	82
Apêndice B .....	83

## Resumo

Partindo do pressuposto de que a arqueologia possibilita a reconstrução das trajetórias históricas das populações até a atualidade, bem como a compreensão dos processos de continuidade e transformação cultural de tais trajetórias ao longo do tempo, o projeto tem o objetivo de analisar fragmentos cerâmicos encontrados em diferentes sítios arqueológicos localizados na Aldeia Lalima e, dessa forma, compreender um pouco melhor a histórica cultural da região. Com o intuito de distinguir e classificar os fragmentos arqueológicos foi aplicada a técnica multivariada Análise de Conglomerados. De maneira a melhor entender a relação entre as características dos fragmentos e, consequentemente, classificá-los quanto à sua origem, foi construída uma Rede Bayesiana.

## 1. Introdução

O estudo da arqueologia possibilita a reconstrução das trajetórias históricas das populações até a atualidade, bem como, a compreensão dos processos de continuidade e transformação cultural de tais trajetórias ao longo do tempo. Além disso, é de conhecimento que a presença de índios no território brasileiro é muito anterior ao processo de ocupação estabelecido pelos exploradores europeus.

Dentro desse contexto, encontra-se a Aldeia Lalima, uma terra indígena localizada na margem direita do rio Miranda, Pantanal do Mato Grosso do Sul, 255km a oeste da capital Campo Grande.

De acordo com pesquisas arqueológicas e etnoarqueológicas verificou-se que a Aldeia tem sido culturalmente construída e transformada pelas populações indígenas desde a pré-história.

É de conhecimento histórico que, apesar de habitar os mesmos lugares, cada população indígena possui seus próprios hábitos culturais. Tais hábitos são altamente influenciados pelas mudanças que ocorrem no território e na paisagem com o decorrer dos anos.

Dentro dessas mudanças destaca-se a variabilidade artefactual encontrada na Aldeia Lalima. Através das características de objetos cerâmicos e conhecimento histórico-cultural da região é possível identificar quais populações indígenas possivelmente habitaram a região desta aldeia.

Dessa forma, este trabalho tem como objetivo relacionar as características encontradas nos artefatos cerâmicos com as populações que ali habitaram, fornecendo insumo necessário para estudar as formações territoriais e a cronologia de ocupação na Aldeia Lalima.

## 2. Descrição do estudo

Foram realizadas, por meio de abordagens arqueológicas e com o auxílio de índios que vivem na região, coletas de materiais cerâmicos em diferentes locais da Aldeia Lalima – denominados sítios arqueológicos.

Foram coletados dados arqueológicos de 7 sítios em três temporadas de atividades em campo que ocorreram nos seguintes intervalos: Maio-Julho/2007, Maio-Julho/2011 e Abril-Julho/2012.

De maneira geral, procurou-se conduzir as coletas de modo que as variáveis relativas à distribuição espacial dos sítios, à cronologia de ocupação e às informações tecnológicas fossem integradas em função da compreensão da ocupação no território da Aldeia.

Cada sítio foi subdividido em sub-regiões de acordo com a localização e data da coleta, denominadas proveniência. Foram coletados artefatos na superfície terrestre, bem como, em níveis abaixo da superfície com profundidade de até 70 cm.

Após as coletas dos materiais algumas variáveis com relação à composição, forma e utilidade do fragmento arqueológico eram medidas e observadas com o intuito de caracterizar e, consequentemente, determinar a proveniência da peça.

Essa proveniência refere-se à possível Tradição tecnológica da peça, ou seja, sob qual forma de artesanato a peça foi submetida e qual grupo indígena produziu-a.

### Sítios arqueológicos estudados:

- Sítio 1 – Córrego Lalima
- Sítio 2 – José Rondon de Souza
- Sítio 3 – Tapera do Limpão
- Sítio 4 – Tapera do Gino
- Sítio 5 – Tapera da Mata do Urumbava
- Sítio 6 – Asa de Pote
- Sítio 8<sup>1</sup>

---

<sup>1</sup> O sítio 8 não possui um nome específico – apenas a nomenclatura Sítio 8.

### 3. Descrição das variáveis

Nos artefatos arqueológicos coletados foram medidas variáveis com relação à sua forma (peso, diâmetro e material), características visuais (cores, acabamento e formatos), localização e finalidade.

As duas únicas variáveis do estudo que assumem valores contínuos são as que representam o peso (medido em gramas) e a Espessura (medida em centímetros) dos fragmentos. As demais variáveis, incluindo aquelas que representam a altura dos fragmentos, já foram apresentadas categorizadas.

A principal variável abordada nesse estudo é a variável *Tradição tecnológica*, que representa a possível proveniência de cada peça e foi classificada pelo pesquisador através da análise das demais variáveis e de seu conhecimento histórico. Uma lista com as variáveis aferidas e a descrição de suas categorias encontra-se no apêndice A.

### 4. Análise descritiva

Tendo em vista que o objetivo principal deste trabalho é encontrar as variáveis significantes para determinar a proveniência das peças arqueológicas, a análise descritiva será feita tomando-se como base a variável *Tradição tecnológica*, ou seja, será analisada a distribuição dessa variável com relação a cada categoria das demais variáveis.

Inicialmente observou-se que apenas 766 dos 3.190 fragmentos disponíveis para análise foram classificados pelo pesquisador em um dos três grupos conhecidos; os demais ainda estão classificados como não identificados.

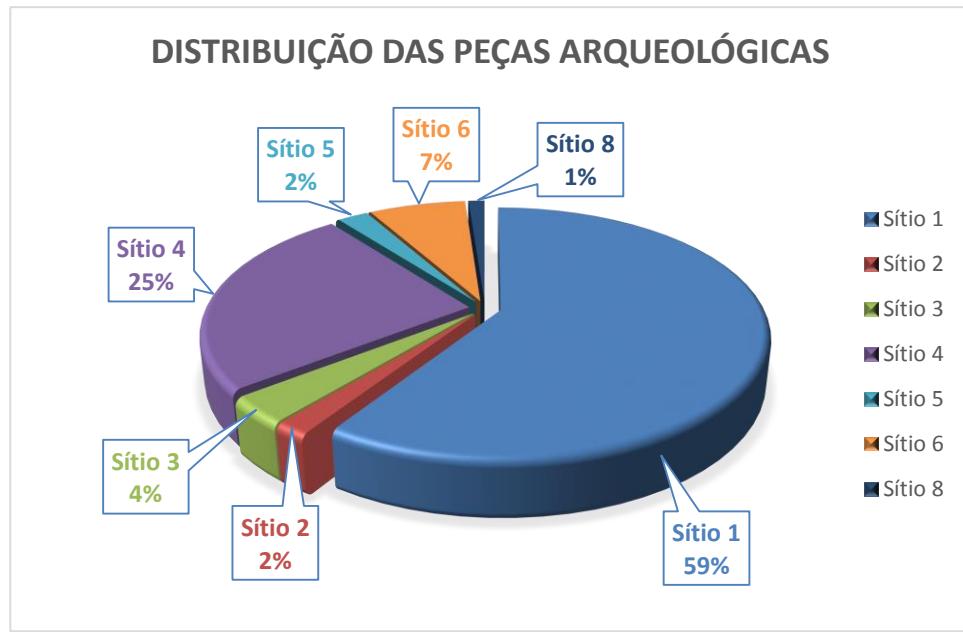
A análise descritiva da base identificou que a maioria das peças foram encontradas nos sítios 1 (59%) e 4 (25%) (Tabela 1). Com relação à classificação dos fragmentos, dos 766 classificados, 15,9% foram classificados no estilo Guarani, 5,3% no estilo Jacadigo, apenas 2,9% no estilo Kadiwéu (Tabela 2) e os 76% restantes ainda não foram classificados devido à falta de características marcantes.

A variável proveniência (Tabela 3) refere-se à uma subdivisão dos sítios arqueológicos, portanto, cada proveniência está relacionada unicamente a um determinado sítio. A classificação 99 da variável proveniência indica que o fragmento não teve sua proveniência especificada dentro do sítio de origem.

Já, para a distribuição dos fragmentos com relação ao nível de superfície em que foram encontrados, observamos que a maioria foi coletado na superfície (nível 0 – 81,9%), seguidos dos níveis 1 (até 10 cm de profundidade – 5,4%), nível 2 (até 20 cm de profundidade – 4,9%) e nível 3 (até 30 cm de profundidade – 3,7%). Os demais fragmentos foram encontrados nos níveis entre 30 – 70 cm e 28 fragmentos (0,9%) não tiveram seu nível identificado (Tabela 4).

**Tabela 1 - Distribuição dos fragmentos por Sítio arqueológico.**

Sítio	Frequência	%
1	1.892	59,3
2	57	1,8
3	111	3,5
4	801	25,1
5	71	2,2
6	225	7,1
8	33	1,0
Total	3.190	100



**Gráfico 1 – Distribuição dos fragmentos arqueológicos por Sítio arqueológico.**

**Tabela 2 - Distribuição dos fragmentos com relação à Tradição tecnológica.**

Tradição tecnológica	Frequência	%
Guarani	506	15,9
Jacadigo	168	5,3
Kadiwéy	92	2,9
Não identificado	2424	76,0
Total	3190	100

**Tabela 3 - Distribuição da proveniência por Sítio arqueológico.**

**Tabela 4 - Distribuição dos fragmentos por Nível de superfície.**

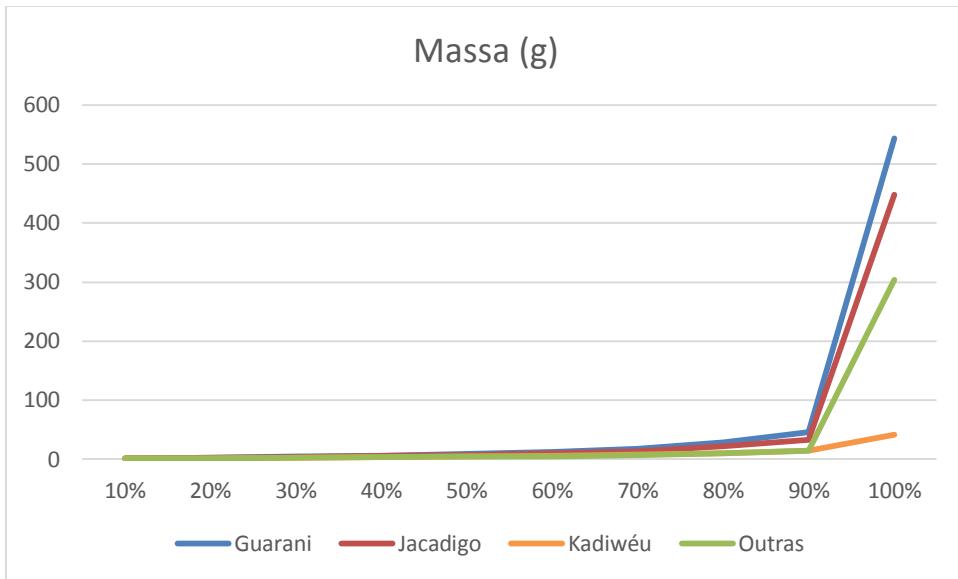
Nível	Frequência	%
0	2.608	81,8
1	173	5,4
2	157	4,9
3	118	3,7
4	73	2,3
5	22	0,7
6	8	0,3
7	3	0,1
99	28	0,9
Total	3.190	100

- Massa

A Massa<sup>2</sup> do fragmento arqueológico foi aferida em gramas. Observamos que os fragmentos com maior Massa foram encontrados nos Sítios 1 e 4 e, quando classificados, em sua maioria, eram da tradição Guarani (Gráficos 2 a 8 e Tabelas 5 a 10). A Tradição tecnológica Kadiwéu é a que apresenta, em sua maioria, os fragmentos com menor Massa. Para os fragmentos arqueológicos do Sítio 8, as variáveis Massa e Espessura ainda não haviam sido aferidas até o momento.

---

<sup>2</sup> Apesar de concluir que, por se tratar de fragmentos cerâmicos, seria melhor considerar o volume de cada fragmento para a análise, apresentamos aqui a distribuição encontrada para a variável massa apenas para questão descritiva.



**Gráfico 2 – Gráfico de percentis para a variável Massa (g). Análise geral, sem separação por sítio.**

**Tabela 5 – Estatísticas descritivas para a variável Massa (g) no Sítio 1.**

Tradição	Estatística					
	n	Mínimo	Máximo	Mediana	Média	E.P. Média
Guarani	429	1	543	9	21,5	2,0
Jacadigo	134	1	448	8,5	18,3	3,7
Outras	1329	1	135	5	8,7	0,4

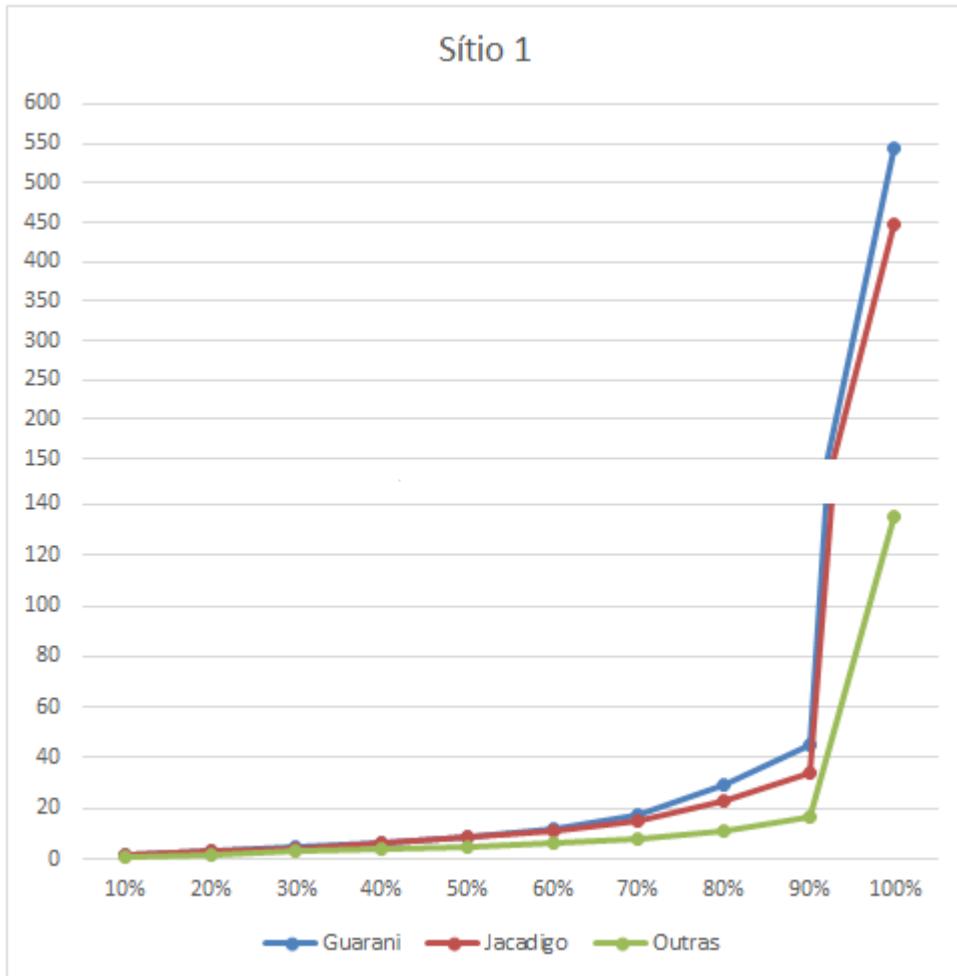


Gráfico 3 – Gráfico de percentis para a variável Massa (g) no Sítio 1.

Tabela 6 – Estatísticas descritivas para a variável Massa (g) no Sítio 2.

Tradição	n	Estatística				
		Mínimo	Máximo	Mediana	Média	E.P. Média
Kadiwéu	12	1	42	5	8,8	3,4
Outras	45	1	22	5	6,1	0,7

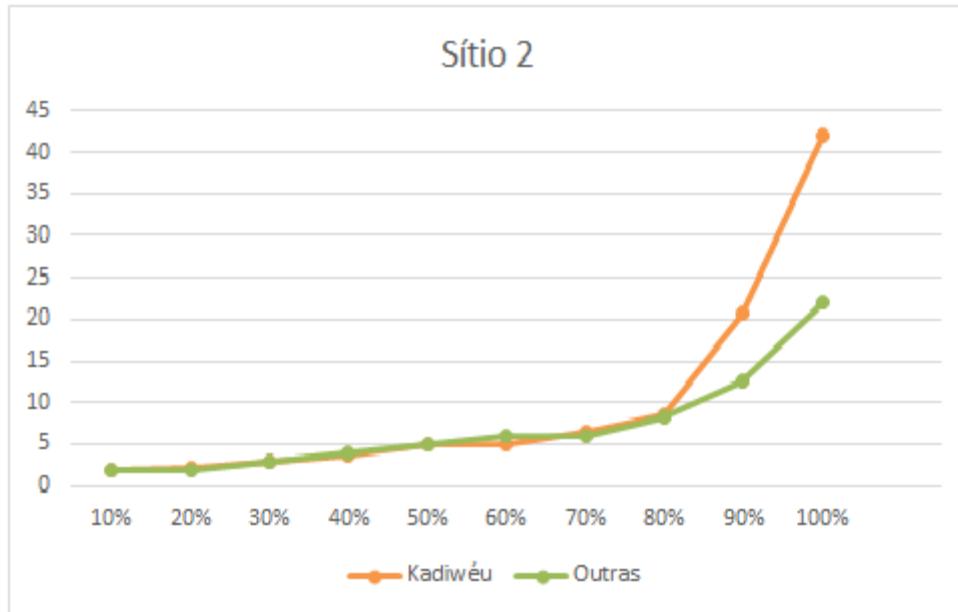


Gráfico 4 - Gráfico de percentis para a variável Massa (g) no Sítio 2.

Tabela 7 – Estatísticas descritivas para a variável Massa (g) no Sítio 3.

Tradição	n	Estatística				
		Mínimo	Máximo	Mediana	Média	E.P. Média
Guarani	1	1	1	1	1,0	1,0
Kadiwéu	3	5	30	10	15,0	7,6
Outras	107	1	116	5	8,5	1,2

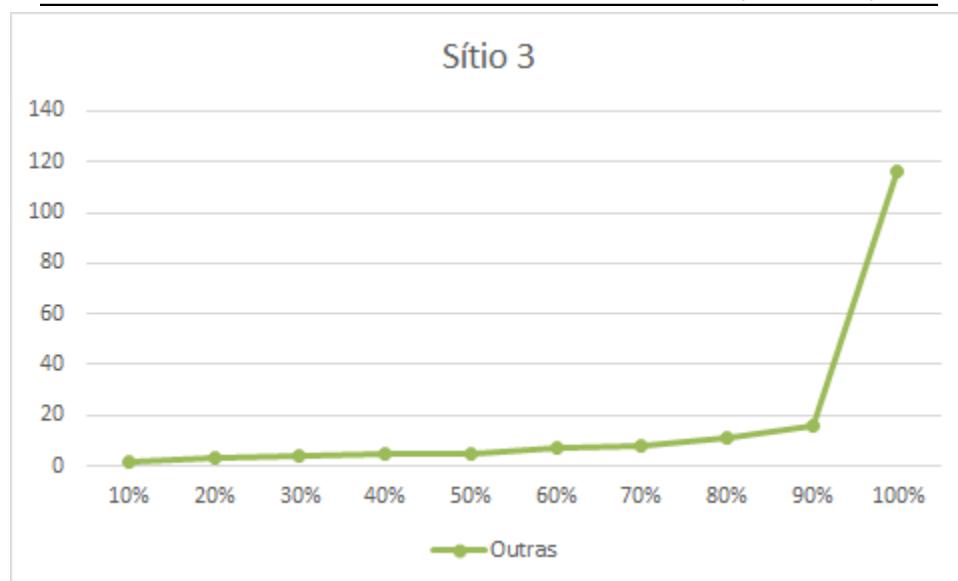
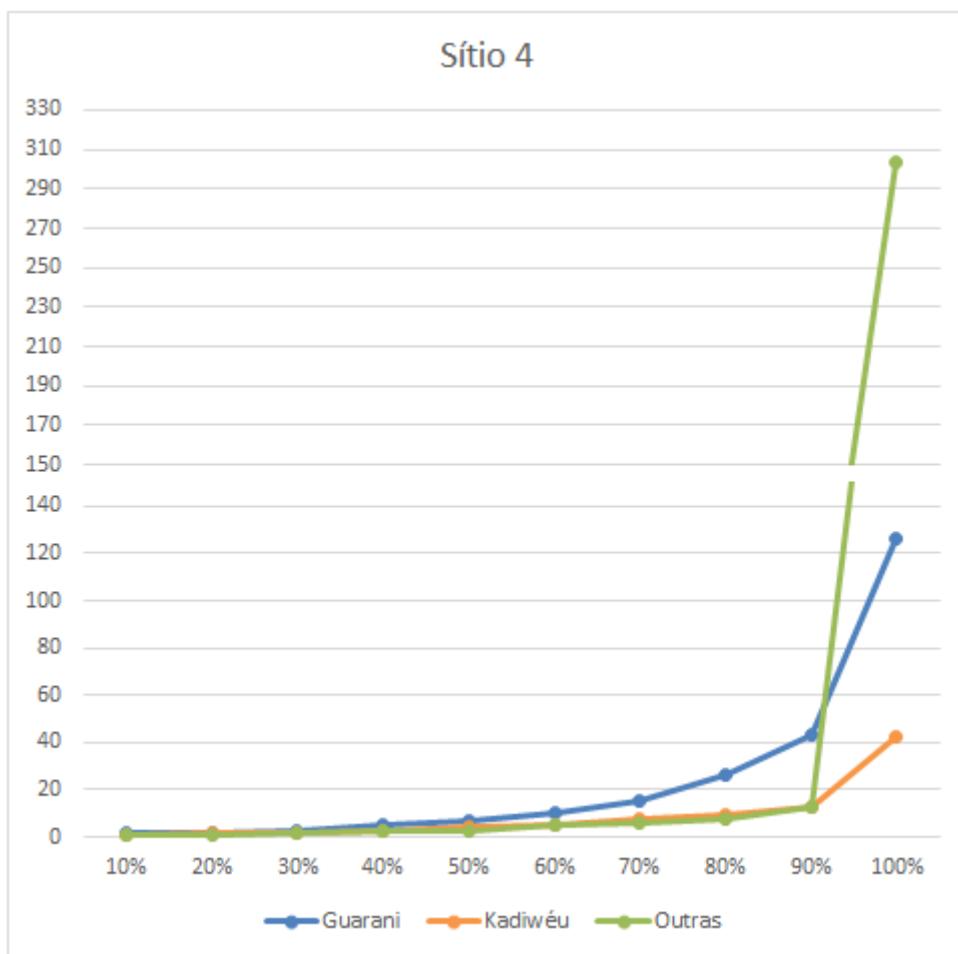


Gráfico 5 - Gráfico de percentis para a variável Massa (g) no Sítio 3.

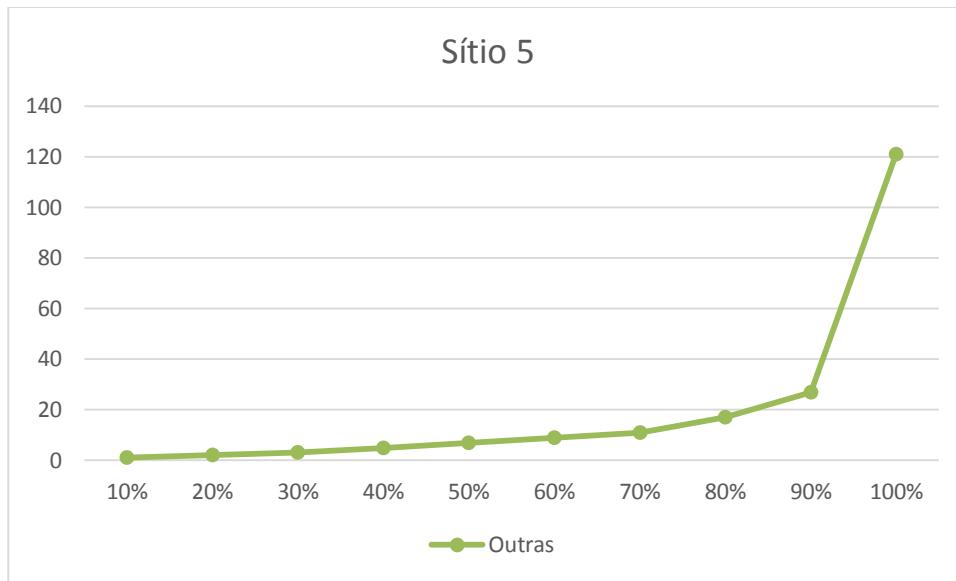
**Tabela 8 – Estatísticas descritivas para a variável Massa (g) no Sítio 4.**

Tradição	Estatística					
	n	Mínimo	Máximo	Mediana	Média	E.P. Média
Guarani	69	1	126	7	16,8	2,8
Kadiwéu	77	1	42	4	6,4	0,8
Outras	655	0	304	3	6,2	0,5

**Gráfico 6 - Gráfico de percentis para a variável Massa (g) no Sítio 4.**

**Tabela 9 - Estatísticas descritivas para a variável Massa (g) no Sítio 5.**

Tradição	Estatística					
	n	Mínimo	Máximo	Mediana	Média	E.P. Média
Outras	71	1	121	7	13,1	2,3

**Gráfico 7 - Gráfico de percentis para a variável Massa (g) no Sítio 5.****Tabela 10 - Estatísticas descritivas para a variável Massa (g) no Sítio 6.**

Tradição	Estatística					
	n	Mínimo	Máximo	Mediana	Média	E.P. Média
Guarani	7	5	121	45	48,6	15,8
Jacadigo	34	1	67	6,5	10,6	2,3
Outras	184	1	40	5	6,8	0,5

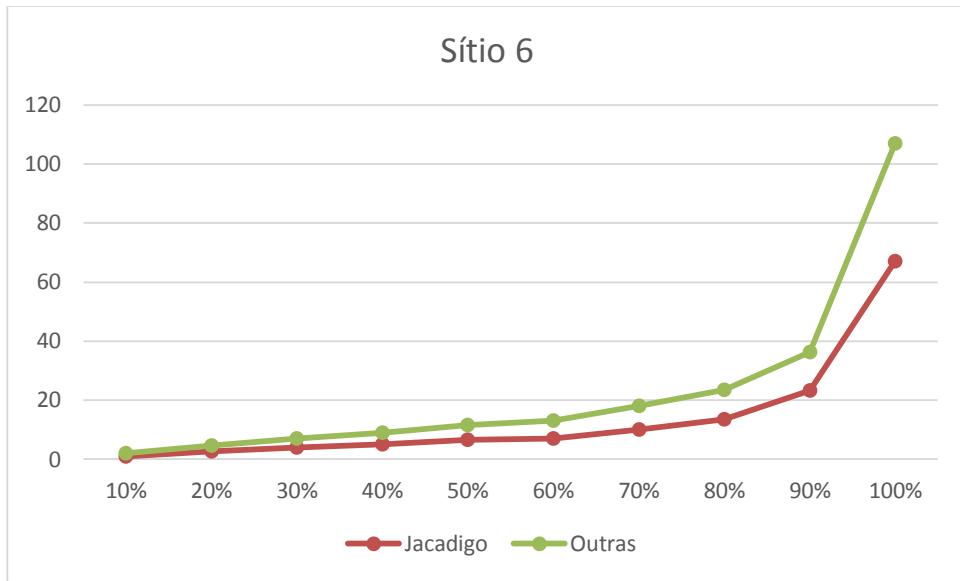


Gráfico 8 - Gráfico de percentis para a variável Massa (g) no Sítio 6.

- Espessura do fragmento (cm)

As Espessuras os fragmentos arqueológicos foram aferidas em milímetros. Nos sítio 1 e 4, a Tradição tecnológica Guarani é a que apresenta os fragmentos de maior diâmetro com relação às demais Tradições. No sítio 6, os fragmentos da Tradição tecnológica Jacadigo apresenta todos os percentis inferiores aos dos fragmentos não classificados (Gráficos 9 a 15 e Tabelas 11 a 16).

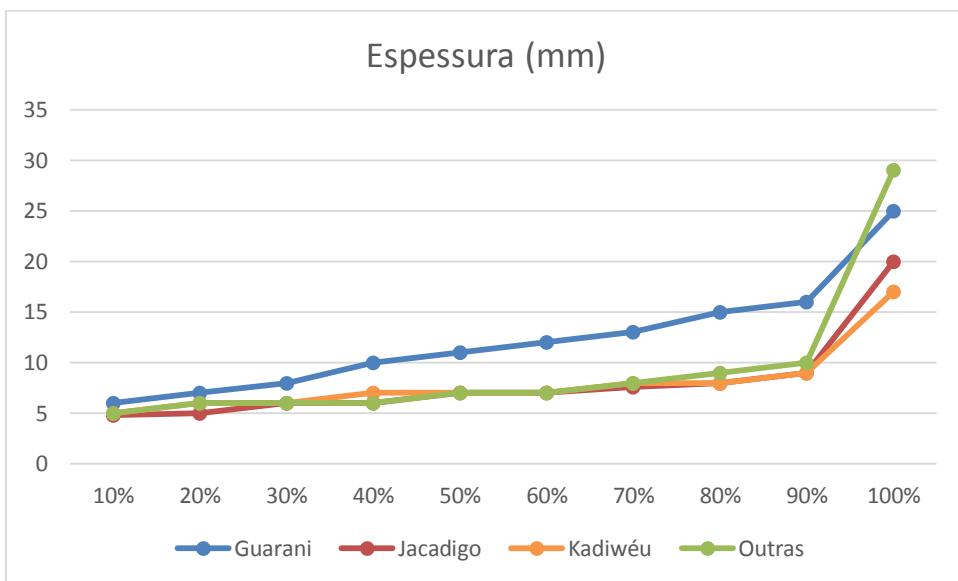
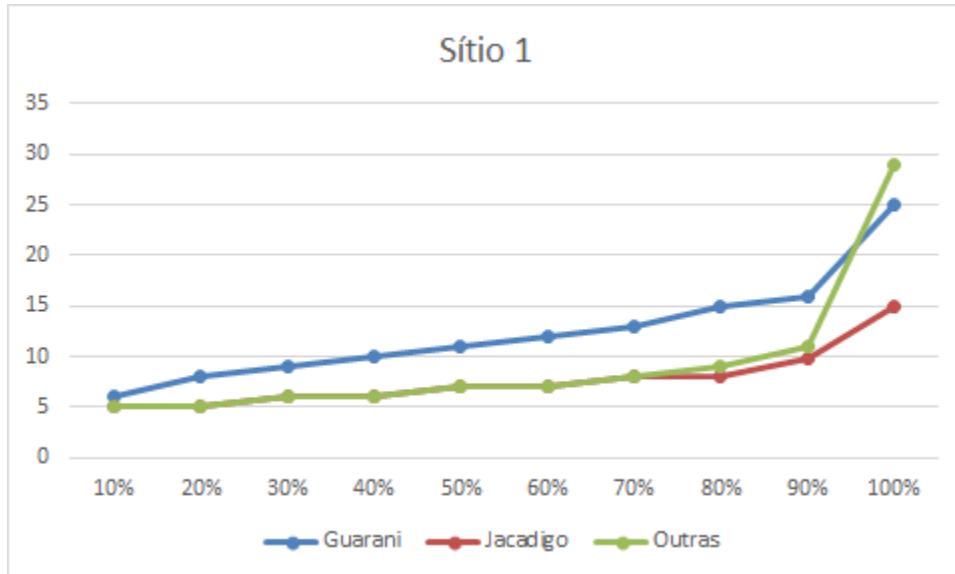


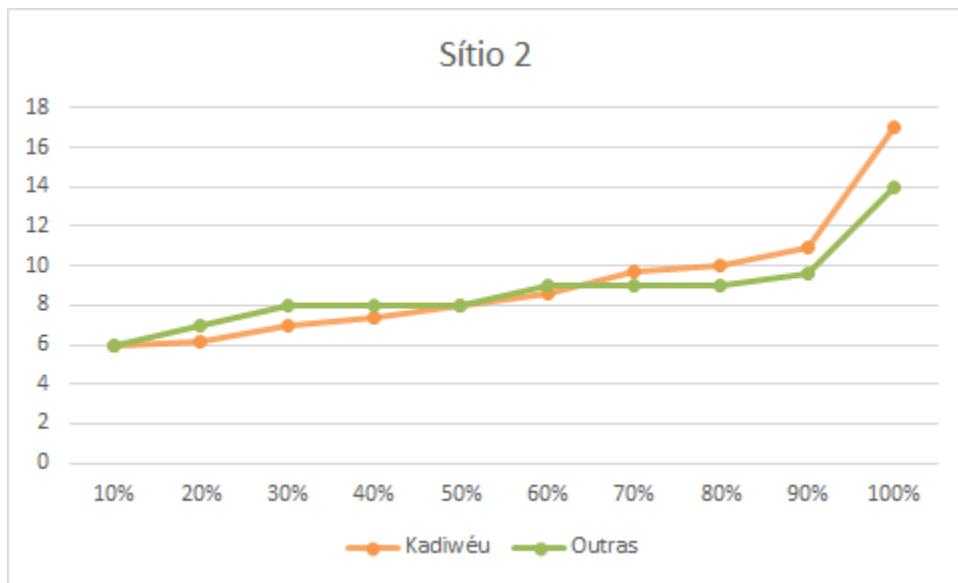
Gráfico 9 - Gráfico de percentis para a variável Espessura (cm). Análise geral – sem separação por sítio.

**Tabela 11 - Estatísticas descritivas para a variável Espessura (cm) no Sítio 1.**

Tradição	Estatística					
	n	Mínimo	Máximo	Mediana	Média	E.P. Média
Guarani	429	4	25	11	11,1	0,2
Jacadigo	134	3	15	7	7,0	0,2
Outras	1329	3	29	7	7,5	0,1

**Gráfico 10 - Gráfico de percentis para a variável Espessura (cm) no Sítio 1.****Tabela 12 - Estatísticas descritivas para a variável Espessura (cm) no Sítio 2.**

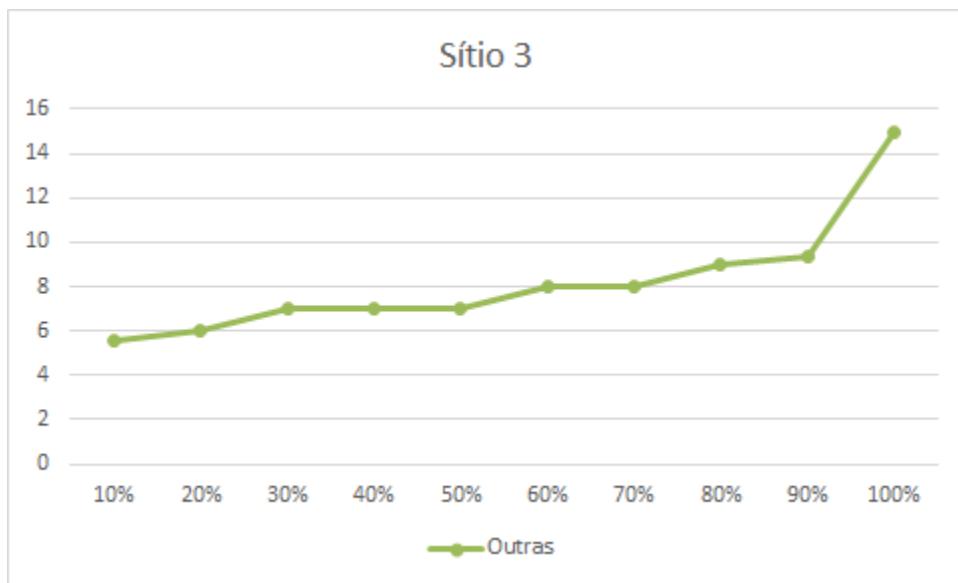
Tradição	Estatística					
	n	Mínimo	Máximo	Mediana	Média	E.P. Média
Kadiwéu	12	5	17	8	8,7	0,9
Outras	45	5	14	8	8,1	0,2



**Gráfico 11 - Gráfico de percentis para a variável Espessura (cm) no Sítio 2.**

**Tabela 13 - Estatísticas descritivas para a variável Espessura (cm) no Sítio 3.**

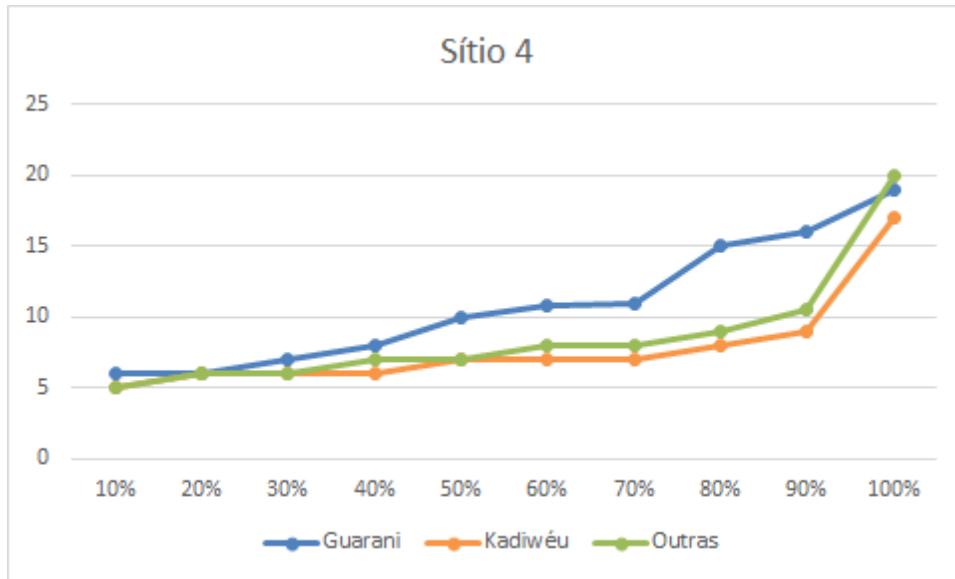
Tradição	n	Estatística				
		Mínimo	Máximo	Mediana	Média	E.P. Média
Guarani	1	6	6	6	6,0	6,0
Kadiwéu	3	6	9	8	7,7	0,9
Outras	107	4	15	7	7,6	0,2



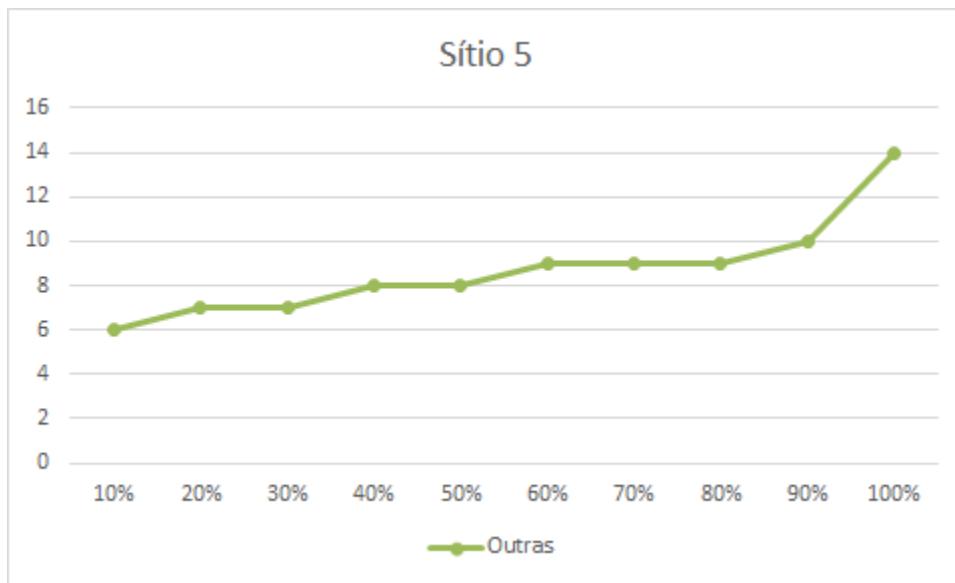
**Gráfico 12 - Gráfico de percentis para a variável Espessura (cm) no Sítio 3.**

**Tabela 14 - Estatísticas descritivas para a variável Espessura (cm) no Sítio 4.**

Tradição	Estatística					
	n	Mínimo	Máximo	Mediana	Média	E.P. Média
Guarani	69	4	19	10	10,1	0,5
Kadiwéu	77	4	17	7	7,0	0,2
Outras	655	3	20	7	7,6	0,1

**Gráfico 13 - Gráfico de percentis para a variável Espessura (cm) no Sítio 4.****Tabela 15 - Estatísticas descritivas para a variável Espessura (cm) no Sítio 5.**

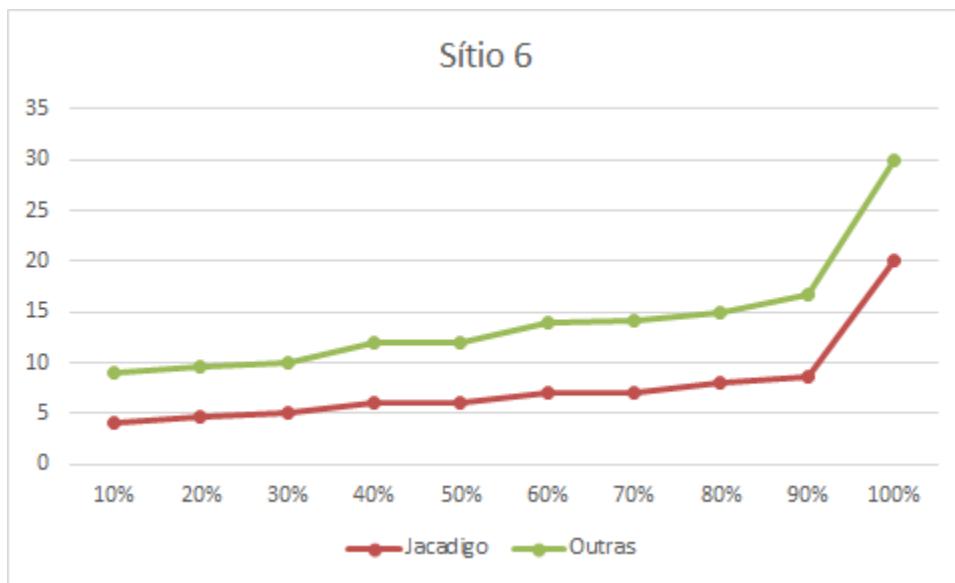
Tradição	Estatística					
	n	Mínimo	Máximo	Mediana	Média	E.P. Média
Outras	71	4	14	8	8,2	0,2



**Gráfico 14 - Gráfico de percentis para a variável Espessura (cm) no Sítio 5.**

**Tabela 16 - Estatísticas descritivas para a variável Espessura (cm) no Sítio 6.**

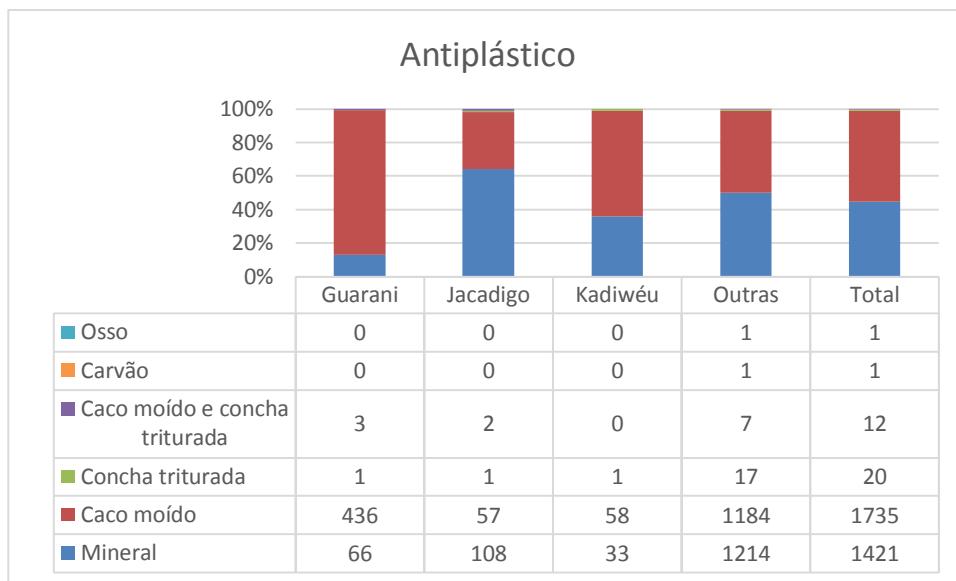
Tradição	Estatística					
	n	Mínimo	Máximo	Mediana	Média	E.P. Média
Guarani	7	7	17	10	11,7	1,4
Jacadigo	34	3	20	6	6,6	0,5
Outras	184	3	10	6	6,3	0,1



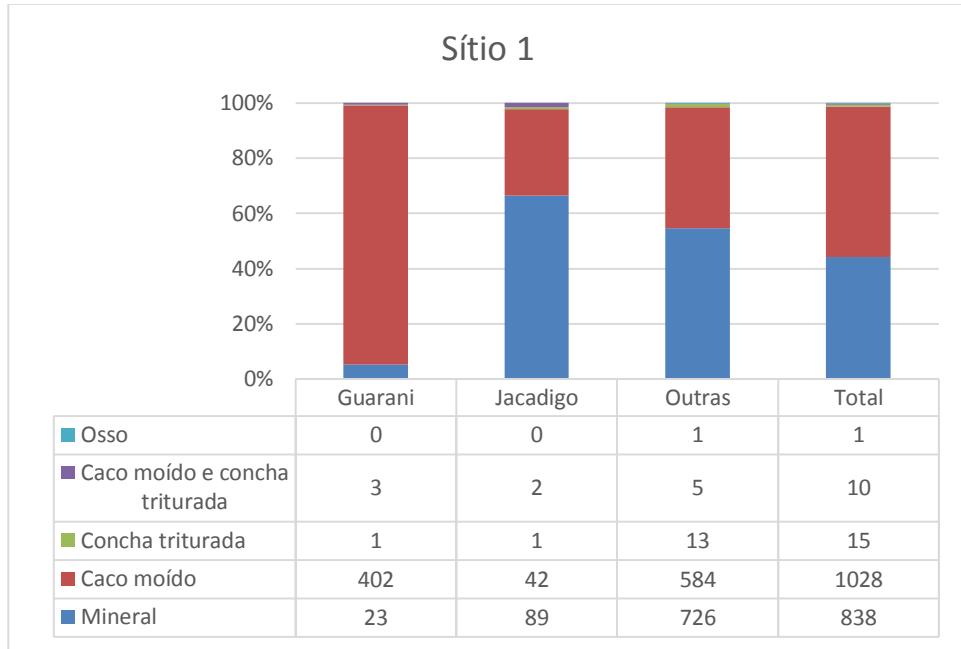
**Gráfico 15 - Gráfico de percentis para a variável Espessura (cm) no Sítio 6.**

- Antiplástico

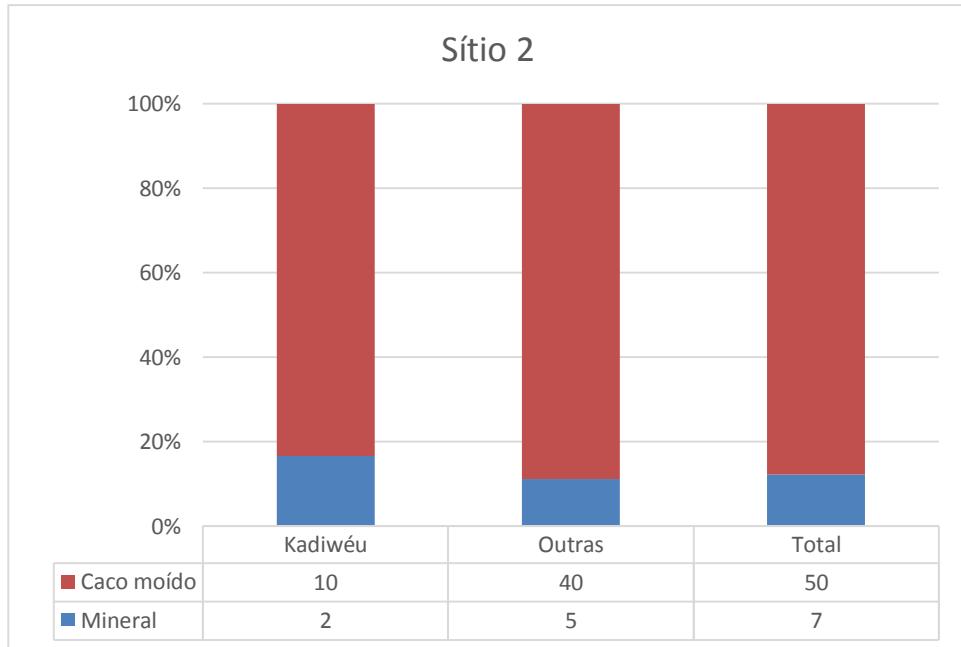
A variável Antiplástico refere-se ao tipo de material (mineral, caco, concha, osso, carvão) que constitui o fragmento arqueológico. Em todos os sítios analisados, as Tradições tecnológicas Guarani e Kadiwéu apresentam, em sua maioria, fragmentos constituídos de caco moído, já a Tradição tecnológica Jacadigo apresenta mais fragmentos constituídos de mineral (Figuras 1 a 7).



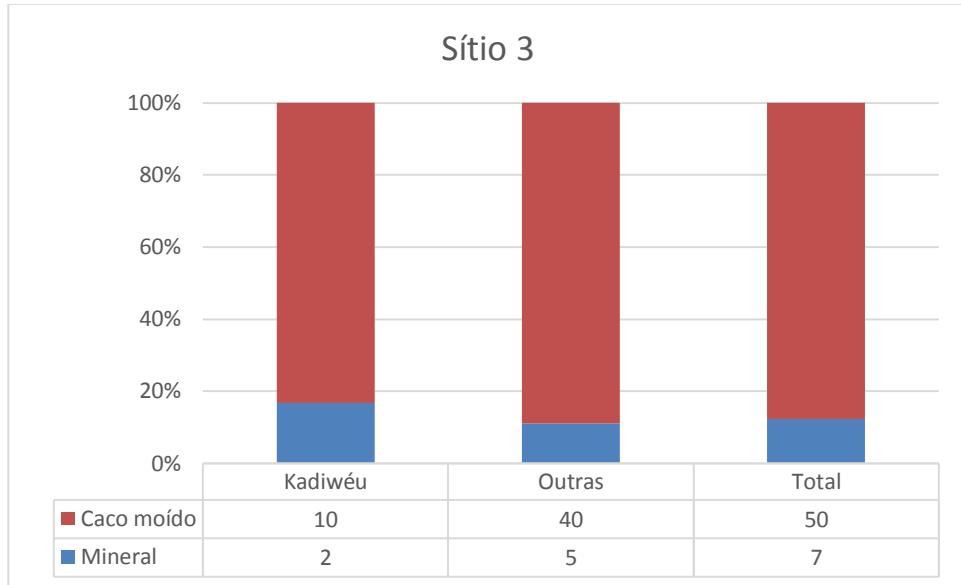
**Figura 1 - Distribuição da variável Antiplástico por Tradição tecnológica. Análise Geral – sem separação por sítio.**



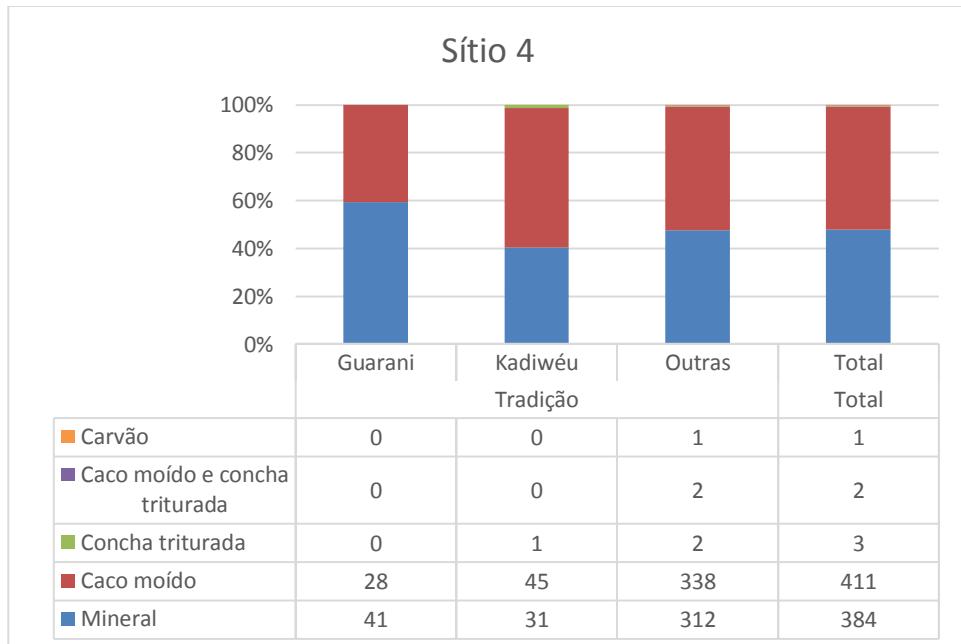
**Figura 2 – Distribuição da variável Antiplástico por Tradição tecnológica no Sítio 1.**



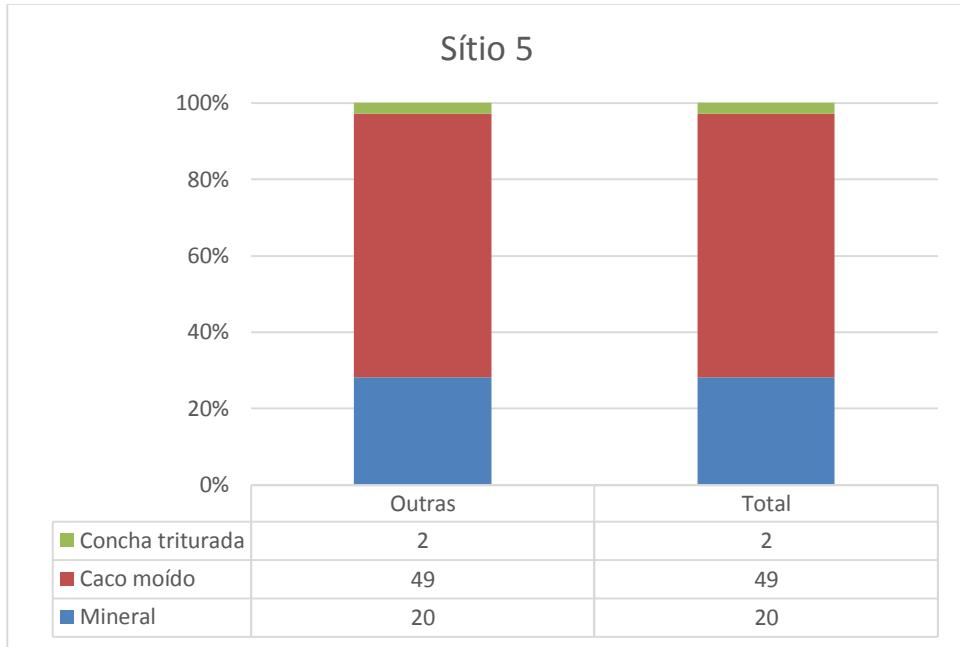
**Figura 3 – Distribuição da variável Antiplástico por Tradição tecnológica no Sítio 2.**



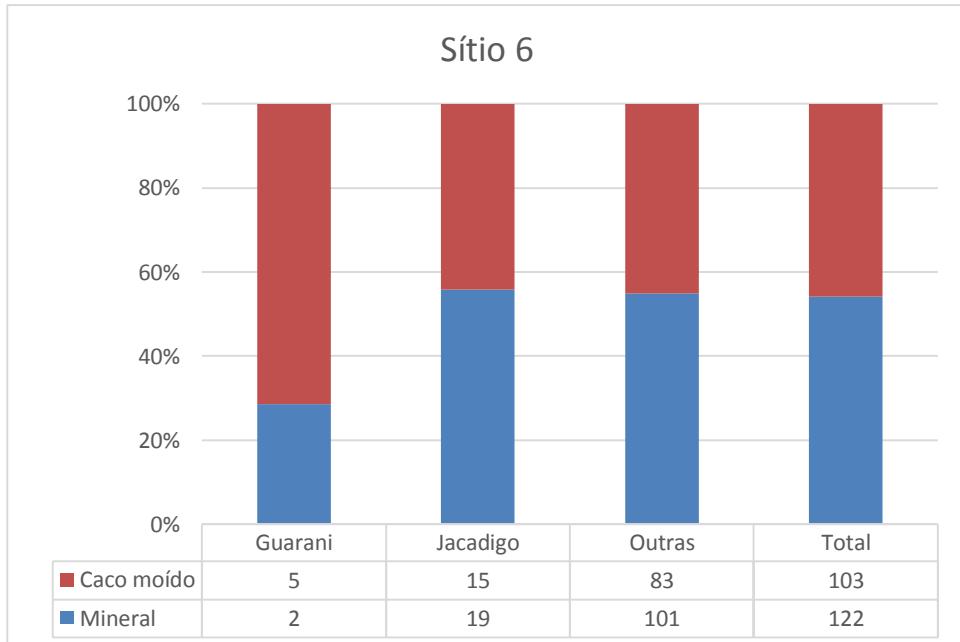
**Figura 4 – Distribuição da variável Antiplástico por Tradição tecnológica no Sítio 3.**



**Figura 5 – Distribuição da variável Antiplástico por Tradição tecnológica no Sítio 4.**



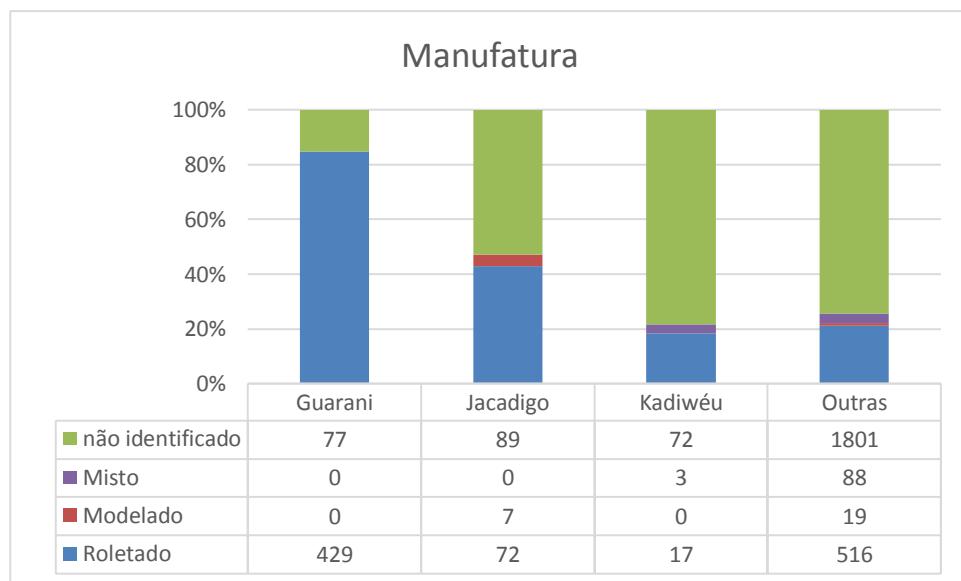
**Figura 6 – Distribuição da variável Antiplástico por Tradição tecnológica no Sítio 5.**



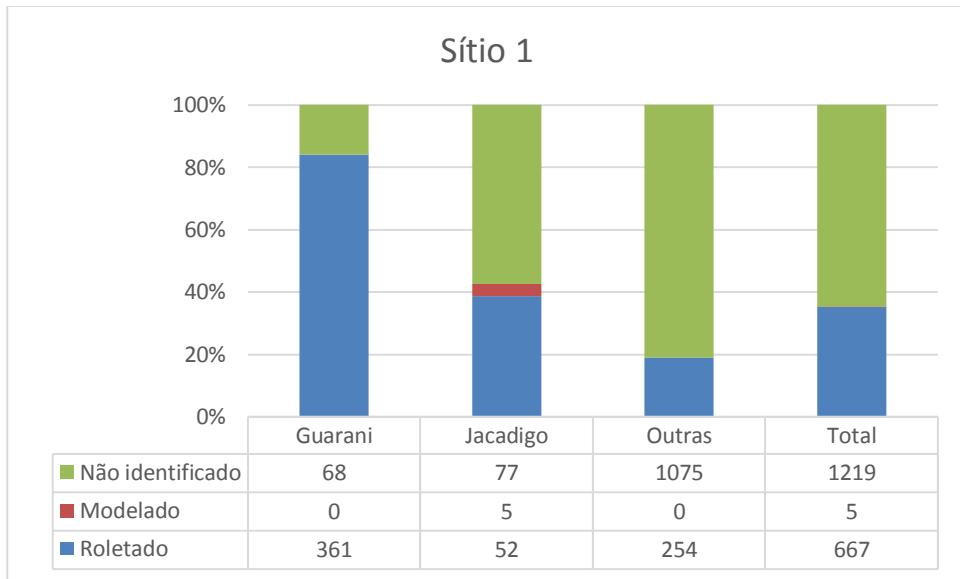
**Figura 7 – Distribuição da variável Antiplástico por Tradição tecnológica no Sítio 6.**

- Técnica de manufatura

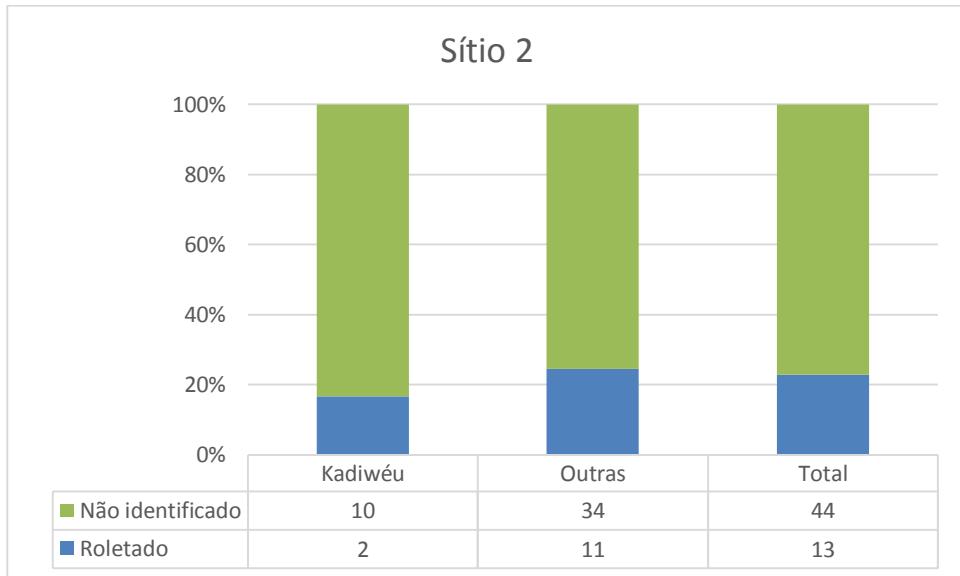
A variável Técnica de manufatura representa o tipo de manufatura aplicado à peça arqueológica. Em muitos dos fragmentos encontrados não foi possível identificar a técnica utilizada. A Tradição tecnológica Guarani é representada por fragmentos em que foi aplicada a técnica Roletado. A Técnica de manufatura Modelado foi observada em fragmentos ainda não classificados e em 5 fragmentos do sítio 1 que foram classificadas no grupo Jacadigo (Figuras 8 a 13).



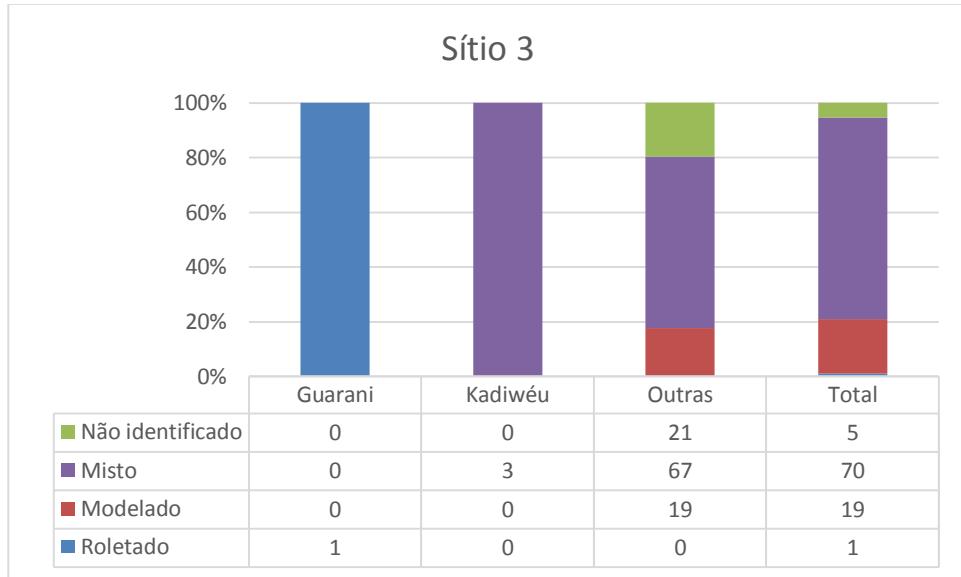
**Figura 8 - Distribuição da variável Manufatura por Tradição tecnológica. Análise Geral – sem separação por sítio.**



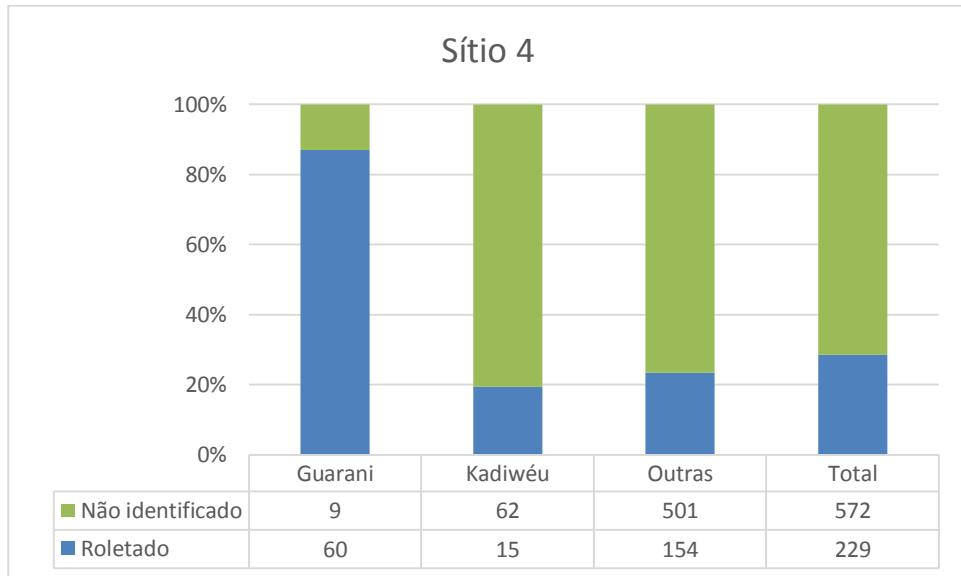
**Figura 9 - Distribuição da variável Manufatura por Tradição tecnológica no Sítio 1.**



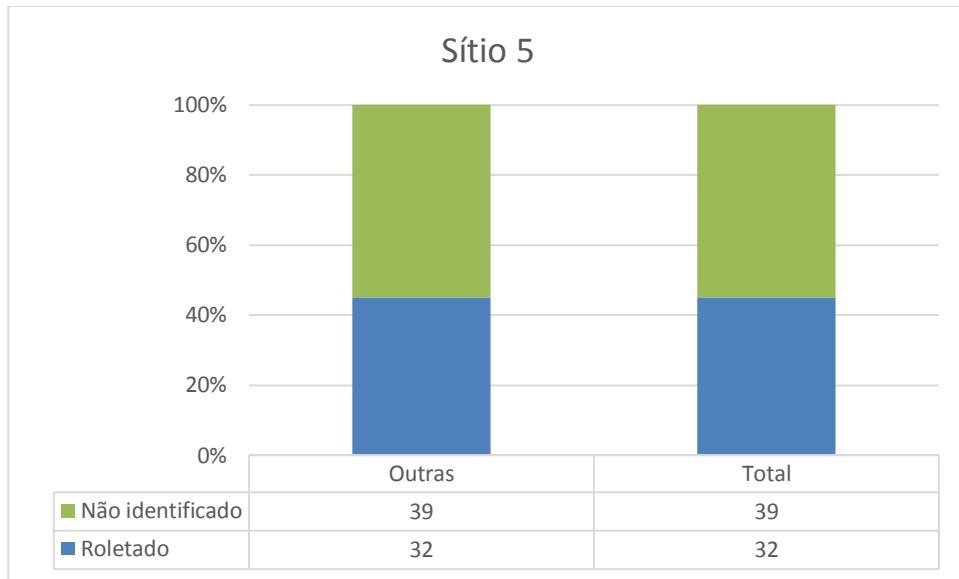
**Figura 10 - Distribuição da variável Manufatura por Tradição tecnológica no Sítio 2.**



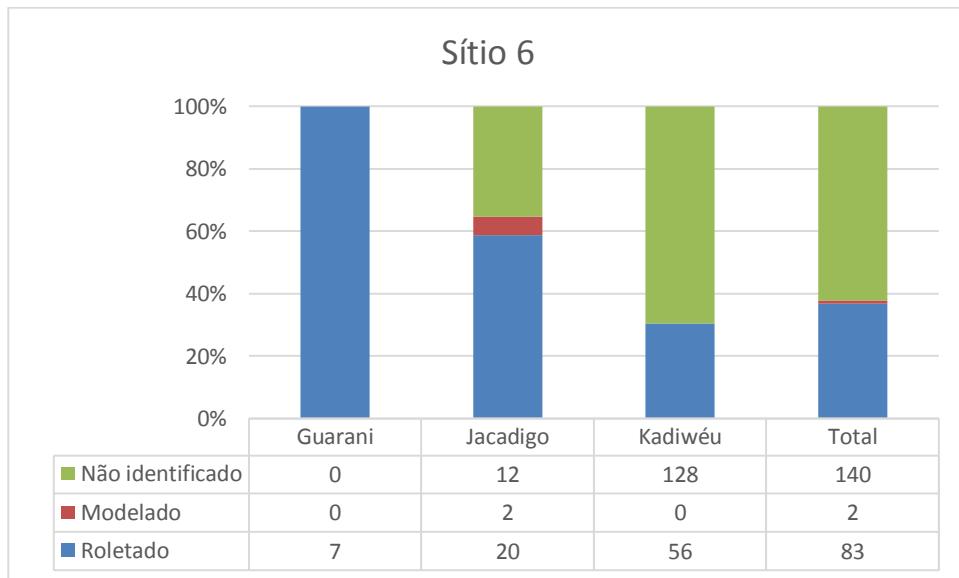
**Figura 11 - Distribuição da variável Manufatura por Tradição tecnológica no Sítio 3.**



**Figura 12 - Distribuição da variável Manufatura por Tradição tecnológica no Sítio 4.**



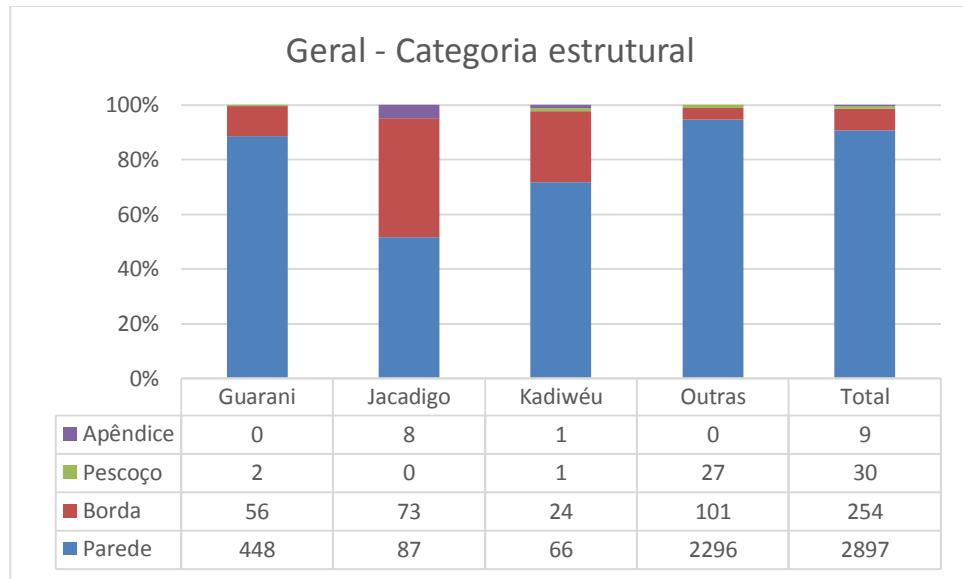
**Figura 13 - Distribuição da variável Manufatura por Tradição tecnológica no Sítio 5.**



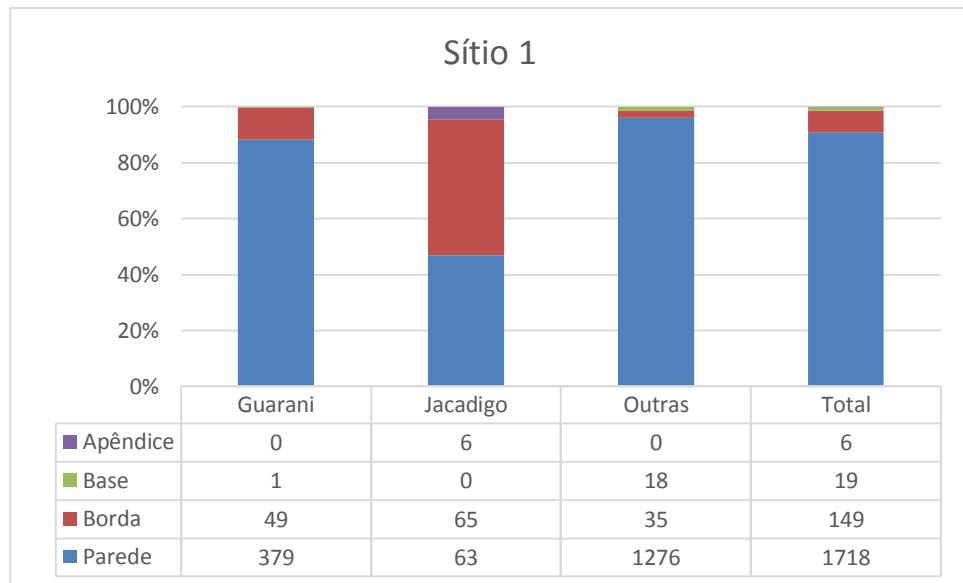
**Figura 14 - Distribuição da variável Manufatura por Tradição tecnológica no Sítio 6.**

- Categoria estrutural do fragmento

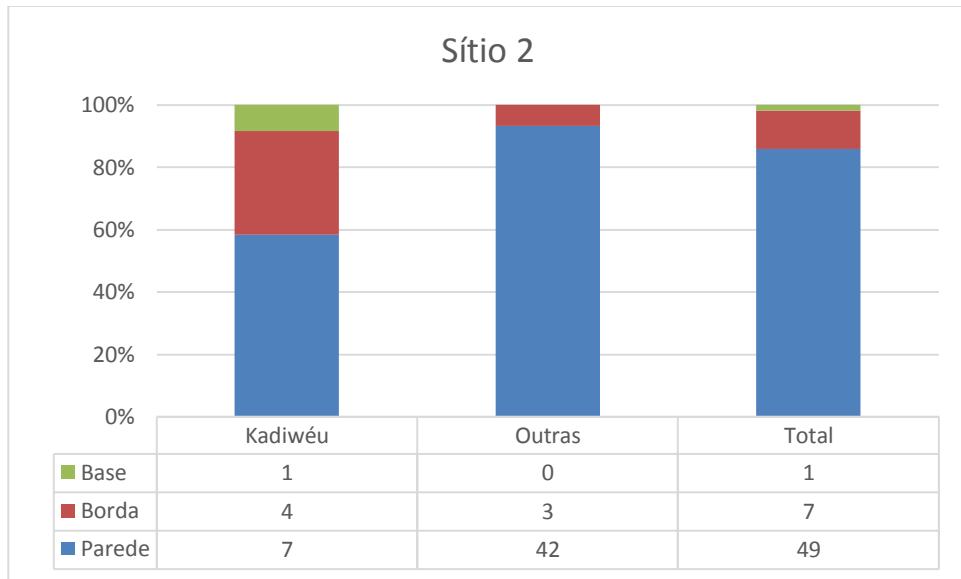
Essa variável refere-se à categoria estrutural que o fragmento fazia parte originalmente na peça. Observa-se, em todos os sítios, que a maior parte dos fragmentos são oriundas da parede da peça, seguidos de fragmentos que constituíam a borda da peça original (Figuras 14 a 19).



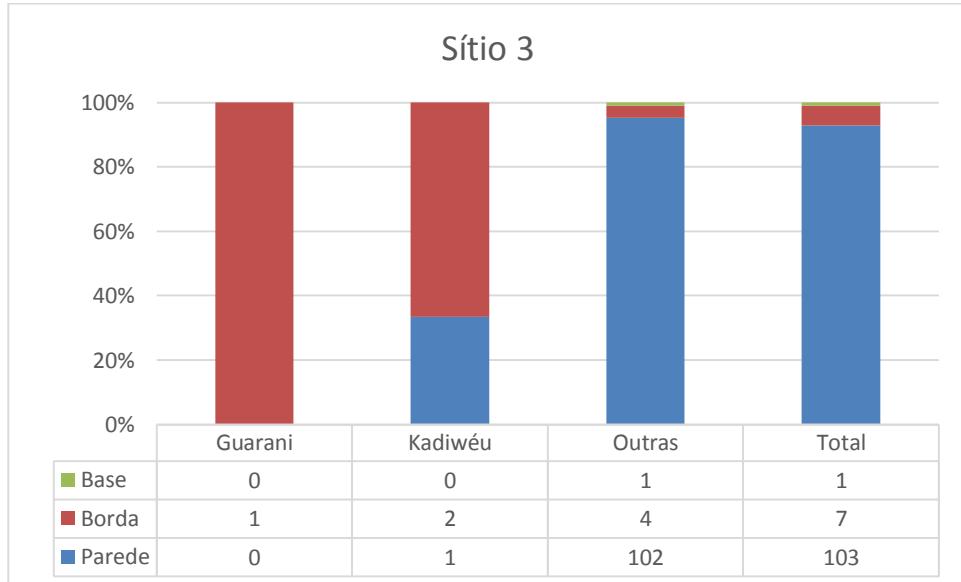
**Figura 15 - Distribuição da variável Categoria estrutural do fragmento por Tradição tecnológica. Análise geral – sem separação por sítio.**



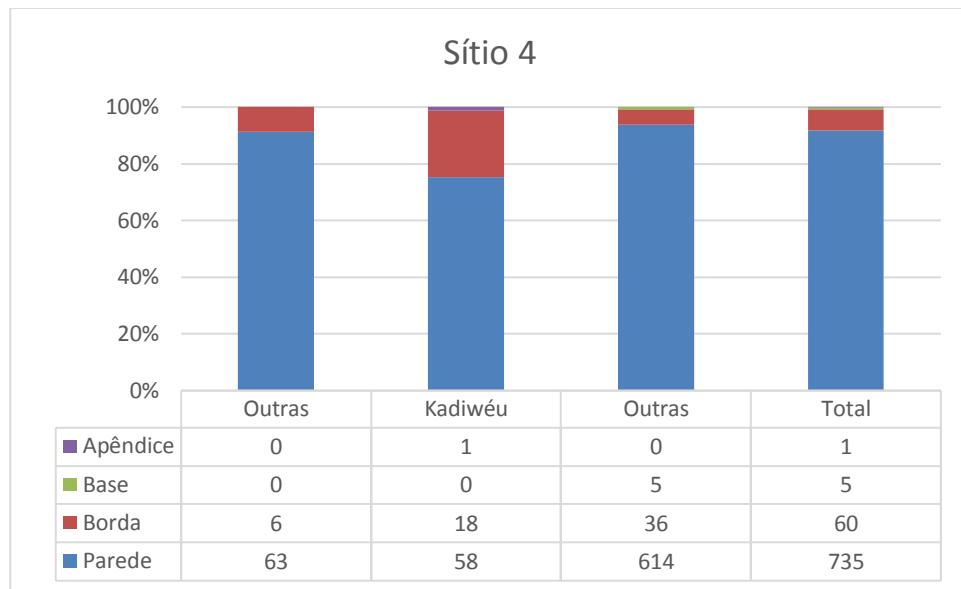
**Figura 16 - Distribuição da variável Categoria estrutural do fragmento por Tradição tecnológica no Sítio 1.**



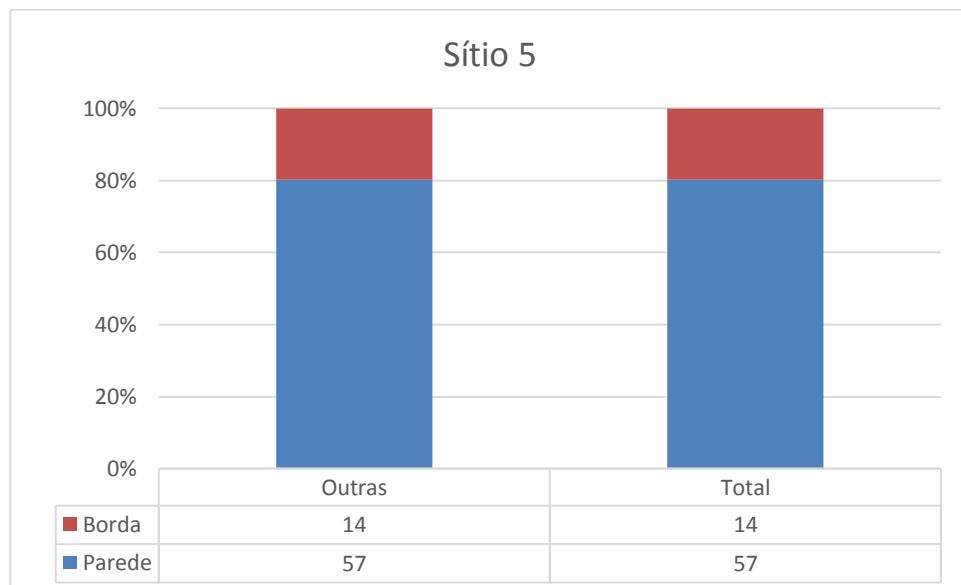
**Figura 17 - Distribuição da variável Categoria estrutural do fragmento por Tradição tecnológica no Sítio 2.**



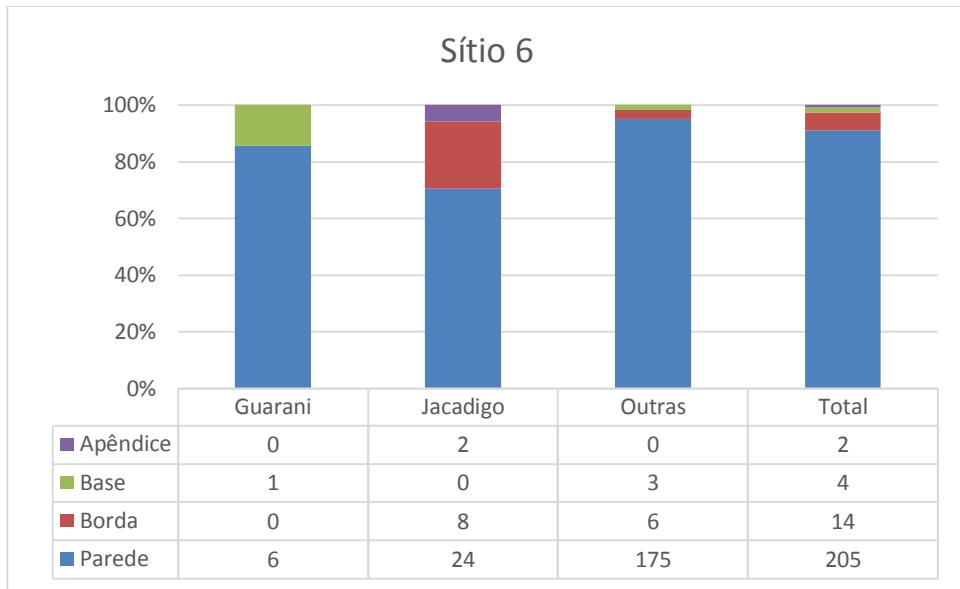
**Figura 18 - Distribuição da variável Categoria estrutural do fragmento por Tradição tecnológica no Sítio 3.**



**Figura 19 - Distribuição da variável Categoria estrutural do fragmento por Tradição tecnológica no Sítio 4.**



**Figura 20 - Distribuição da variável Categoria estrutural do fragmento por Tradição tecnológica no Sítio 5.**



**Figura 21 - Distribuição da variável Categoria estrutural do fragmento por Tradição tecnológica no Sítio 6.**

- Configuração do núcleo

Com relação à configuração do núcleo os fragmentos se apresentam, tanto por sítio, como em geral, bem distribuídos em todas as categorias disponíveis – exceto pela categoria 6, que só foi observado um fragmento, oriundo do sítio 4, com essa configuração (Figuras 21 a 27). Também não é possível observar claramente predominância de alguma configuração em determinada Tradição tecnológica.

1 = Oxidação completa, sem materiais orgânicos



2 = Oxidação incompleta, com materiais orgânicos, margens do core difusas



3 = Oxidação incompleta, com ou sem materiais orgânicos



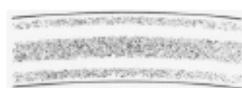
4 = Oxidação incompleta, com materiais orgânicos



5 = Oxidação incompleta, com materiais orgânicos

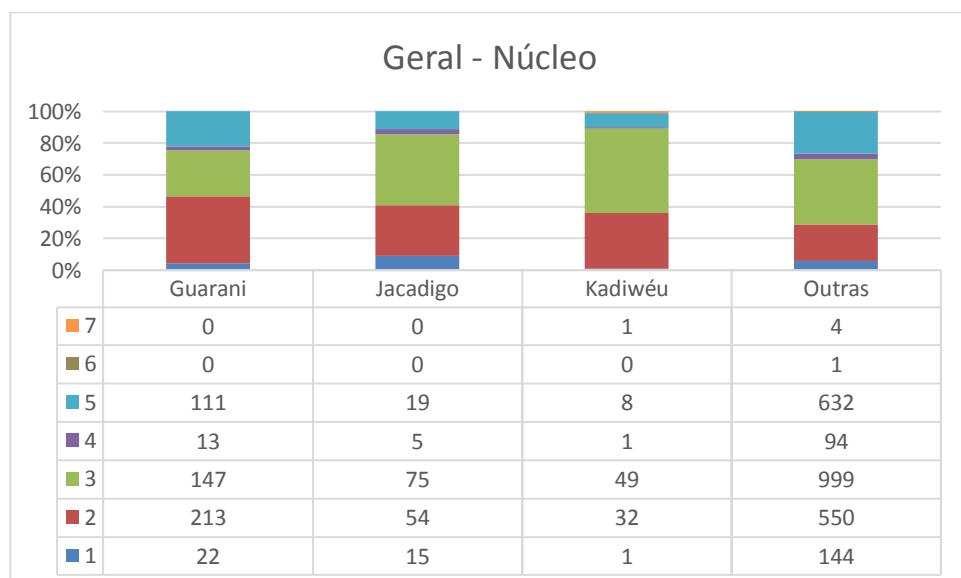


6 = Redução e resfriado rapidamente no ar (2 vezes), margens do core definidas, core duplo

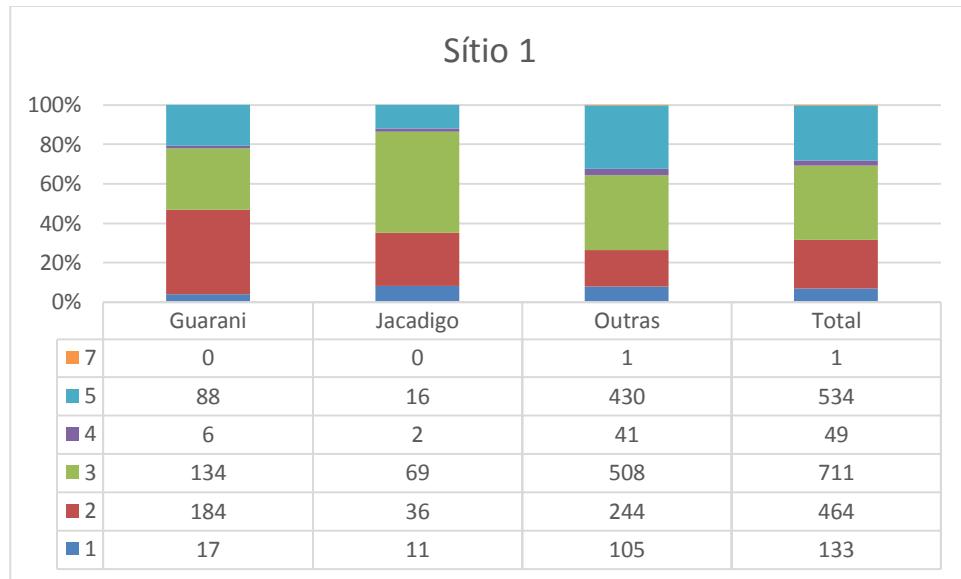


7 = Escuro, claro, escuro

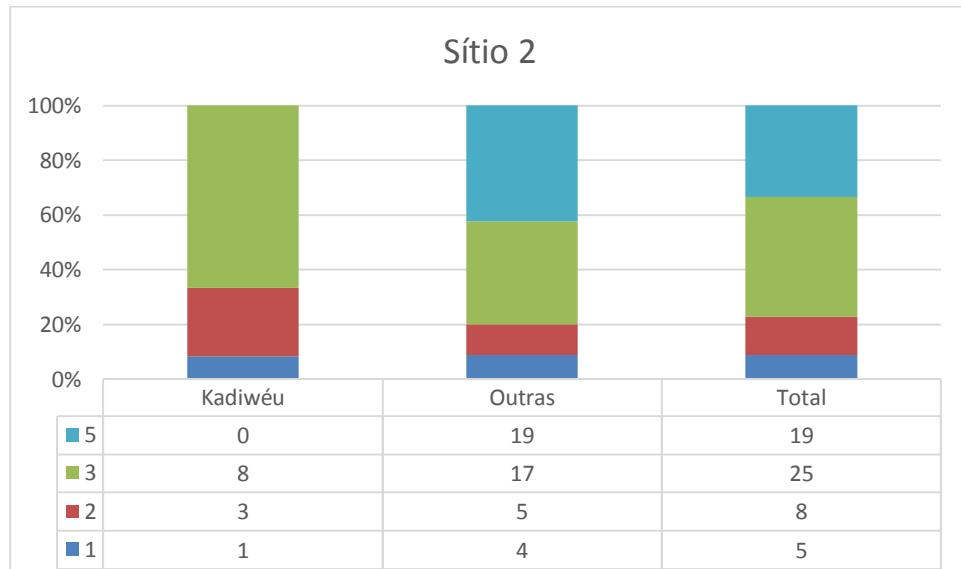
99 = Não identificado



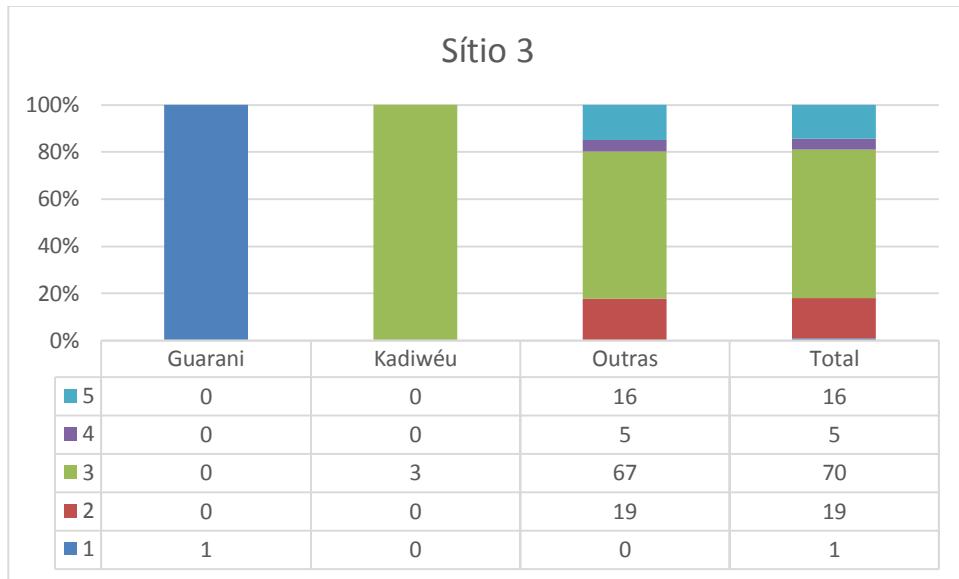
**Figura 22 - Distribuição da variável Núcleo do fragmento por Tradição tecnológica.**  
Análise geral – sem separação por sítio.



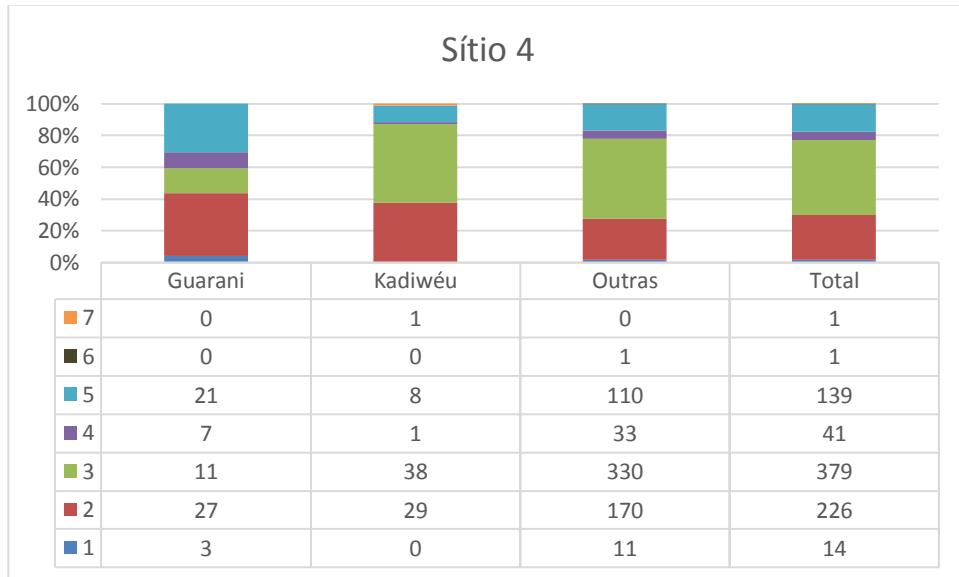
**Figura 23 - Distribuição da variável Núcleo do fragmento por Tradição tecnológica no Sítio 1.**



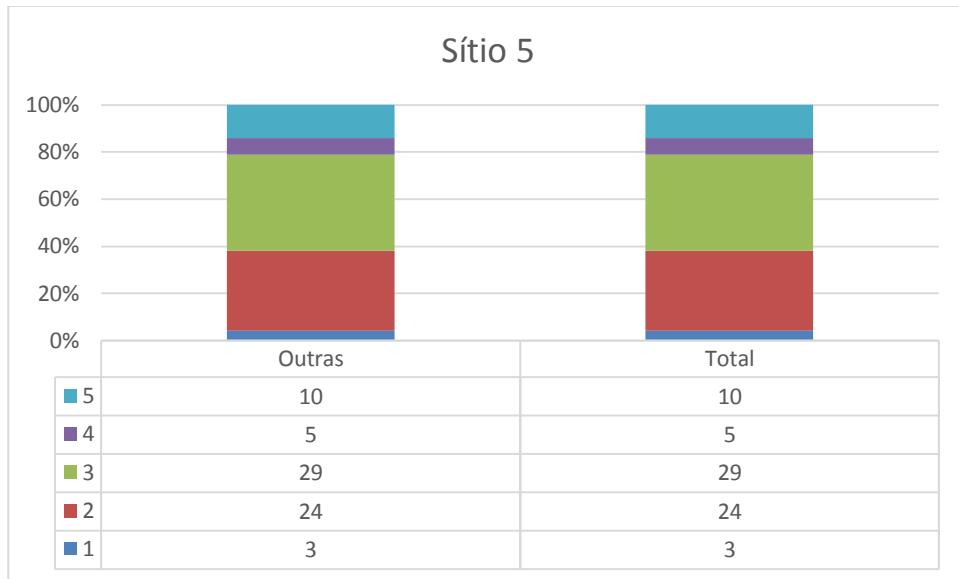
**Figura 24 - Distribuição da variável Núcleo do fragmento por Tradição tecnológica no Sítio 2.**



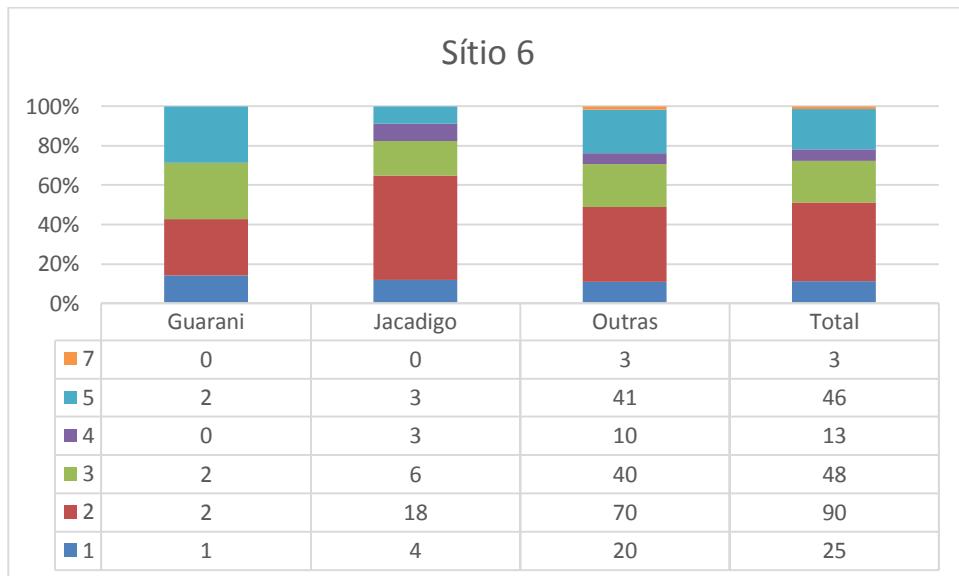
**Figura 25 - Distribuição da variável Núcleo do fragmento por Tradição tecnológica no Sítio 3.**



**Figura 26 - Distribuição da variável Núcleo do fragmento por Tradição tecnológica no Sítio 4.**



**Figura 27 - Distribuição da variável Núcleo do fragmento por Tradição tecnológica no Sítio 5.**



**Figura 28 - Distribuição da variável Núcleo do fragmento por Tradição tecnológica no Sítio 6.**

- Acabamento de superfície face externa

Com relação ao acabamento de superfície face externa, observa-se, em geral, que a maioria dos fragmentos, cerca de 70%, possuem o acabamento de Polimento (1) ou Alisamento (2). Também observa-se que um grande número de fragmentos possuem o

acabamento na face externa Corrugado (4) - cerca de 11% do total (Tabela 17), sendo esses, quase em sua totalidade, classificados na Tradição tecnológica Guarani.

**Tabela 17 – Distribuição da variável Acabamento face externa por Tradição tecnológica.**

Acabamento face externa	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
1	37	42	9	1152	1240	38,9
2	3	16	6	944	969	30,4
3	0	1	1	11	13	0,4
4	355	0	0	3	358	11,2
9	0	0	0	1	1	0,0
10	26	0	0	3	29	0,9
13	0	5	0	0	5	0,2
16	0	36	0	1	37	1,2
17	3	0	0	0	3	0,1
19	4	1	0	1	6	0,2
20	0	1	0	0	1	0,0
21	0	1	1	0	2	0,1
23	0	11	0	0	11	0,3
24	0	9	26	6	41	1,3
25	0	1	0	0	1	0,0
26	0	3	0	0	3	0,1
27	0	0	3	0	3	0,1
30	46	0	1	11	58	1,8
31	4	3	2	106	115	3,6
32	0	1	2	54	57	1,8
33	3	12	2	68	85	2,7
34	3	0	0	0	3	0,1
35	3	0	0	0	3	0,1
36	2	0	0	0	2	0,1
37	3	0	0	0	3	0,1
38	0	5	0	1	6	0,2
39	8	0	0	0	8	0,3
40	2	0	0	0	2	0,1
43	1	0	0	0	1	0,0
44	0	1	0	22	23	0,7
46	0	13	0	0	13	0,4
48	0	0	34	1	35	1,1
49	0	0	1	0	1	0,0
51	1	0	0	0	1	0,0
53	0	0	1	0	1	0,0

57	0	0	0	1	1	0,0
59	0	1	0	0	1	0,0
60	1	0	0	0	1	0,0
61	0	0	0	2	2	0,1
62	0	0	1	0	1	0,0
63	0	0	0	1	1	0,0
64	0	0	2	0	2	0,1
99	1	5	0	35	41	1,3
Total	506	168	92	2424	3190	100

- Acabamento de superfície face interna

Assim como observou-se no acabamento da face externa, o acabamento na face interna também é predominante as técnicas de Polimento (1) e Alisamento (2) da peça – cerca de 90% dos fragmentos encontrados (Tabela 18).

**Tabela 18 – Distribuição da variável Acabamento face interna por Tradição tecnológica.**

Acabamento face interna	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
0	0	0	0	2	2	0,1
1	298	74	31	1180	1583	49,6
2	166	73	7	1034	1280	40,1
3	0	6	1	19	26	0,8
19	0	0	0	1	1	0,0
27	0	0	1	0	1	0,0
30	14	0	0	2	16	0,5
31	7	3	33	71	114	3,6
32	3	0	2	41	46	1,4
33	3	9	13	32	57	1,8
37	1	0	0	0	1	0,0
38	0	0	0	1	1	0,0
39	1	0	0	0	1	0,0
41	1	0	0	0	1	0,0
44	3	1	1	26	31	1,0
47	0	0	1	0	1	0,0
52	0	1	0	0	1	0,0
55	0	0	1	0	1	0,0
63	0	0	1	0	1	0,0
99	9	1	0	15	25	0,8
Total	506	168	92	2424	3190	100

- Coloração externa

As variáveis de coloração, tanto interna quanto externa, possuem um número muito grande de categorias. Apesar disso, observa-se que não há concentração dos fragmentos em determinadas categorias. Porém, visando uma melhor forma de trabalhar com tais variáveis, as categorias serão agrupadas para obtenção de melhores resultados na análise inferencial (Tabelas 19 e 20).

**Tabela 19 – Distribuição da variável Coloração externa por Tradição tecnológica.**

Coloração externa	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
1	4	4	0	20	28	0,9
2	33	23	0	218	274	8,6
3	36	14	3	110	163	5,1
4	57	10	0	114	181	5,7
5	13	1	0	76	90	2,8
6	45	8	1	115	169	5,3
7	11	0	0	62	73	2,3
8	2	0	0	9	11	0,3
9	8	3	0	25	36	1,1
10	16	5	0	90	111	3,5
11	71	5	0	49	125	3,9
12	57	1	0	51	109	3,4
13	14	4	0	12	30	0,9
14	20	0	0	5	25	0,8
15	15	0	1	8	24	0,8
16	1	3	0	14	18	0,6
17	0	1	0	3	4	0,1
18	1	0	0	0	1	0,0
19	1	0	0	4	5	0,2
20	4	0	0	2	6	0,2
21	3	2	0	53	58	1,8
22	5	6	0	103	114	3,6
23	0	1	1	1	3	0,1
24	14	10	0	76	100	3,1
25	0	1	1	5	7	0,2
26	1	0	1	49	51	1,6
27	4	0	0	1	5	0,2
30	2	3	0	20	25	0,8
31	0	1	0	25	26	0,8

32	0	1	0	14	15	0,5
33	2	13	6	115	136	4,3
34	8	10	1	76	95	3,0
35	0	1	0	5	6	0,2
36	2	6	2	29	39	1,2
37	1	7	2	76	86	2,7
38	4	7	1	86	98	3,1
39	1	0	0	0	1	0,0
40	0	0	0	9	9	0,3
41	0	1	0	6	7	0,2
42	0	0	0	3	3	0,1
43	0	1	0	0	1	0,0
44	0	2	0	9	11	0,3
45	2	0	0	4	6	0,2
46	4	4	1	15	24	0,8
47	0	0	0	4	4	0,1
48	3	4	0	18	25	0,8
50	1	0	0	3	4	0,1
51	2	0	0	2	4	0,1
52	0	0	0	4	4	0,1
53	0	1	1	13	15	0,5
55	0	0	0	3	3	0,1
56	0	0	0	5	5	0,2
57	0	1	0	0	1	0,0
58	0	2	0	4	6	0,2
59	0	0	0	1	1	0,0
61	0	0	0	5	5	0,2
62	0	0	0	2	2	0,1
63	0	0	0	2	2	0,1
64	0	0	0	8	8	0,3
67	0	0	0	1	1	0,0
70	0	0	0	1	1	0,0
71	1	1	0	5	7	0,2
72	0	0	1	1	2	0,1
73	0	0	0	1	1	0,0
99	37	0	69	574	680	21,3
Total	506	168	92	2424	3190	100

- Coloração interna

**Tabela 20 – Distribuição da variável Coloração interna por Tradição tecnológica.**

Coloração interna	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
1	12	6	0	9	27	0,8
2	46	15	1	165	227	7,1
3	23	15	0	96	134	4,2
4	75	8	0	73	156	4,9
5	10	10	1	87	108	3,4
6	77	12	1	268	358	11,2
7	31	1	0	47	79	2,5
8	6	0	1	7	14	0,4
9	4	2	0	10	16	0,5
10	10	11	0	97	118	3,7
11	36	1	0	40	77	2,4
12	41	4	0	38	83	2,6
13	5	6	0	8	19	0,6
14	10	0	0	3	13	0,4
15	15	2	0	12	29	0,9
16	0	3	0	14	17	0,5
17	1	1	0	0	2	0,1
19	1	0	0	2	3	0,1
20	3	0	0	7	10	0,3
21	5	2	1	39	47	1,5
22	10	8	0	100	118	3,7
23	1	0	0	1	2	0,1
24	10	6	3	73	92	2,9
25	0	3	1	4	8	0,3
26	2	0	0	42	44	1,4
28	3	0	0	3	6	0,2
30	1	2	0	14	17	0,5
31	0	0	0	21	21	0,7
32	1	1	0	17	19	0,6
33	8	11	2	89	110	3,4
34	3	7	0	76	86	2,7
35	0	1	0	3	4	0,1
36	3	4	2	65	74	2,3
37	0	7	0	50	57	1,8
38	6	8	0	91	105	3,3
40	0	0	1	13	14	0,4
41	0	0	0	8	8	0,3

42	0	0	0	9	9	0,3
43	0	0	0	1	1	0,0
44	0	3	0	9	12	0,4
45	1	1	0	6	8	0,3
46	2	3	1	22	28	0,9
47	1	0	0	1	2	0,1
48	2	2	0	44	48	1,5
49	0	0	0	4	4	0,1
50	0	1	0	3	4	0,1
52	0	0	0	3	3	0,1
53	0	0	4	10	14	0,4
54	0	0	0	1	1	0,0
55	0	0	0	4	4	0,1
56	1	0	0	11	12	0,4
57	1	0	0	2	3	0,1
58	0	1	1	1	3	0,1
59	0	0	0	5	5	0,2
60	0	0	0	1	1	0,0
61	0	0	0	3	3	0,1
62	0	0	0	3	3	0,1
64	0	0	2	3	5	0,2
65	0	0	0	1	1	0,0
66	0	0	0	1	1	0,0
68	0	0	0	1	1	0,0
69	0	0	0	5	5	0,2
71	2	0	0	2	4	0,1
72	0	0	1	2	3	0,1
99	37	0	69	574	680	21,3
Total	506	168	92	2424	3190	100

- Acabamento de superfície no lábio

Sabendo que nem todas os fragmentos faziam parte da categoria estrutural *borda* da peça, para a variável acabamento de superfície no lábio, observa-se que a maioria dos fragmentos (92,3%) foram classificados na categoria não se aplica (0), já que não há como classificar os demais fragmentos com relação ao acabamento do lábio (Tabela 21).

**Tabela 21 – Distribuição da variável Acabamento do lábio por Tradição tecnológica.**

Acabamento lábio	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
0	450	95	68	2332	2945	92,3
1	35	16	1	46	98	3,1
2	14	22	1	32	69	2,2
3	0	0	0	2	2	0,1
5	0	0	15	0	15	0,5
10	0	1	0	0	1	0,0
11	0	0	1	0	1	0,0
12	0	1	0	0	1	0,0
13	0	3	0	0	3	0,1
23	0	19	0	1	20	0,6
31	0	1	0	2	3	0,1
32	0	0	0	2	2	0,1
33	1	2	2	4	9	0,3
34	5	2	0	1	8	0,3
38	0	5	0	0	5	0,2
45	0	0	2	0	2	0,1
50	0	0	1	0	1	0,0
54	0	1	0	0	1	0,0
56	0	0	1	0	1	0,0
99	1	0	0	2	3	0,1
Total	506	168	92	2424	3190	100

Assim como se observou para a variável Acabamento de superfície no lábio, alguns fragmentos, devido à sua categoria estrutural, não puderam ser classificados com relação à algumas características – além daqueles em que realmente não foi possível identificar.

Dessa forma, as variáveis a seguir não são tão relevantes para distinção entre os grupos, já que representam as características de determinados fragmentos em especial e, portanto, em sua maioria, os fragmentos foram classificados na categoria *não se aplica* (0) ou *não classificado* (99) (Tabelas 22 a 34).

**Tabela 22 – Distribuição da variável Parede por Tradição tecnológica.**

Parede	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
0	58	82	26	129	295	9,2
1	1	0	0	0	1	0,0
2	2	0	0	1	3	0,1
4	1	7	1	14	23	0,7
5	36	0	0	0	36	1,1
6	1	0	0	0	1	0,0
8	0	0	0	2	2	0,1
99	407	79	65	2278	2829	88,7
Total	506	168	92	2424	3190	100

**Tabela 23 – Distribuição da variável Borda por Tradição tecnológica.**

Borda	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
0	450	95	68	2322	2935	92,0
1	6	23	3	21	53	1,7
2	7	4	0	3	14	0,4
3	7	8	2	7	24	0,8
4	4	16	0	16	36	1,1
6	1	0	0	0	1	0,0
7	12	0	0	0	12	0,4
8	1	0	0	0	1	0,0
10	5	0	0	0	5	0,2
11	2	0	0	1	3	0,1
99	11	22	19	54	106	3,3
Total	506	168	92	2424	3190	100

**Tabela 24 – Distribuição da variável Lábio por Tradição tecnológica.**

Lábio	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
0	451	95	68	2325	2939	92,1
1	26	1	2	14	43	1,3
2	6	1	0	9	16	0,5
3	6	2	4	13	25	0,8
4	11	1	4	24	40	1,3
6	0	18	0	2	20	0,6
7	0	0	2	6	8	0,3
8	3	44	1	8	56	1,8
10	0	0	10	0	10	0,3
11	3	6	1	16	26	0,8
12	0	0	0	7	7	0,2
Total	506	168	92	2424	3190	100

**Tabela 25 – Distribuição da variável Diâmetro por Tradição tecnológica.**

Diâmetro	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
0	450	95	68	2326	2939	92,1
1	1	0	1	0	2	0,1
2	1	1	0	0	2	0,1
3	1	0	0	0	1	0,0
4	0	2	0	1	3	0,1
5	0	2	0	0	2	0,1
99	53	68	23	97	241	7,6
Total	506	168	92	2424	3190	100

**Tabela 26 – Distribuição da variável Base por Tradição tecnológica.**

Base	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
0	504	168	91	2395	3158	99,0
1	1	0	1	13	15	0,5
2	1	0	0	11	12	0,4
3	0	0	0	3	3	0,1
5	0	0	0	1	1	0,0
99	0	0	0	1	1	0,0
Total	506	168	92	2424	3190	100

**Tabela 27 – Distribuição da variável Apêndice por Tradição tecnológica.**

Apêndice	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
0	506	159	91	2424	3180	99,7
1	0	8	1	0	9	0,3
2	0	1	0	0	1	0,0
Total	506	168	92	2424	3190	100

**Tabela 28 – Distribuição da variável Classe por Tradição tecnológica.**

Classe	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
1	1	15	0	3	19	0,6
2	3	2	0	5	10	0,3
3	10	1	0	0	11	0,3
99	492	150	92	2416	3150	98,7
Total	506	168	92	2424	3190	100,0

**Tabela 29– Distribuição da variável Contorno por Tradição tecnológica.**

Contorno	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
1	1	14	0	3	18	0,6
2	1	3	0	3	7	0,2
3	2	0	0	0	2	0,1
4	7	0	0	0	7	0,2
99	495	151	92	2418	3156	98,9
Total	506	168	92	2424	3190	100,0

**Tabela 30 – Distribuição da variável Altura por Tradição tecnológica.**

Altura	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
1	2	2	0	1	5	0,2
2	0	3	0	0	3	0,1
99	504	163	92	2423	3182	99,7
Total	506	168	92	2424	3190	100,0

**Tabela 31 – Distribuição da variável Uso face externa por Tradição tecnológica.**

Uso face externa	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
1	4	0	0	7	11	0,3
2	25	5	4	246	280	8,8
4	2	0	0	3	5	0,2
6	0	1	0	0	1	0,0
8	1	0	0	8	9	0,3
9	0	0	0	1	1	0,0
99	474	162	88	2159	2883	90,4
Total	506	168	92	2424	3190	100,0

**Tabela 32 – Distribuição da variável Uso face interna por Tradição tecnológica.**

Uso face interna	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
1	5	0	0	3	8	0,3
3	18	5	0	109	132	4,1
4	1	0	0	3	4	0,1
5	20	3	0	25	48	1,5
8	1	0	0	9	10	0,3
9	0	0	0	1	1	0,0
99	461	160	92	2274	2987	93,6
Total	506	168	92	2424	3190	100,0

**Tabela 33 – Distribuição da variável Função primária por Tradição tecnológica.**

Função primária	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
1	14	0	0	0	14	0,4
2	1	0	0	0	1	0,0
3	29	0	0	0	29	0,9
4	11	0	0	0	11	0,3
5	0	3	0	0	3	0,1
6	0	0	0	1	1	0,0
7	39	8	4	341	392	12,3
8	15	1	0	17	33	1,0
9	0	2	0	1	3	0,1
12	3	0	0	0	3	0,1
99	394	154	88	2064	2700	84,6
Total	506	168	92	2424	3190	100,0

**Tabela 34 – Distribuição da variável função secundária por Tradição tecnológica.**

Função secundária	Tradição				Total	%
	Guarani	Jacadigo	Kadiwéu	Outras		
1	9	0	0	9	18	0,6
99	497	168	92	2415	3172	99,4
Total	506	168	92	2424	3190	100,0

➤ Análise de Associação

Com o intuito de verificar associações entre as variáveis categóricas utilizou-se o coeficiente  $\lambda$  de Goodman-Kruskal, obtido por:

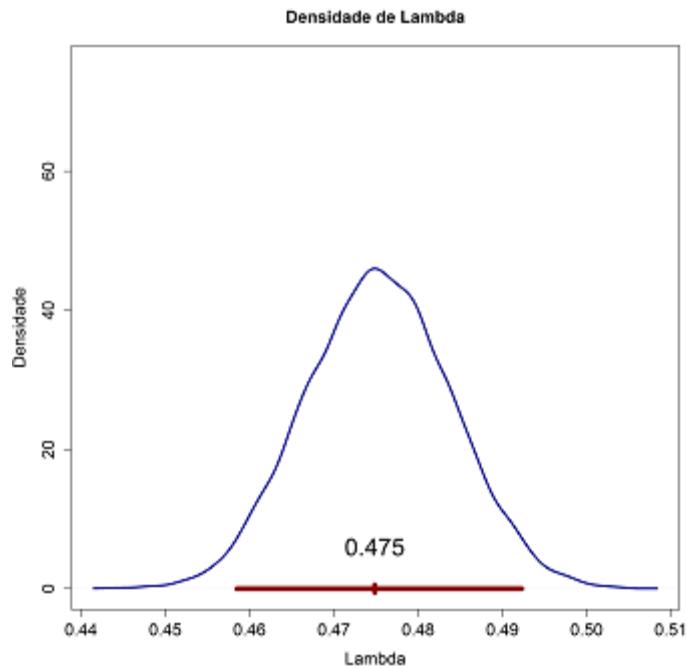
$$\lambda = \frac{\sum_{i \in L} \max_{j \in C} \theta_{ij} + \sum_{j \in C} \max_{i \in L} \theta_{ij} - \max_{i \in L} \theta_{i \cdot} - \max_{j \in C} \theta_{\cdot j}}{2 - \max_{i \in L} \theta_{i \cdot} - \max_{j \in C} \theta_{\cdot j}}$$

em que  $\theta_{ij}$  é a probabilidade do fragmento pertencer a i-ésima categoria da 1ª variável e a j-ésima categoria da 2ª variável.

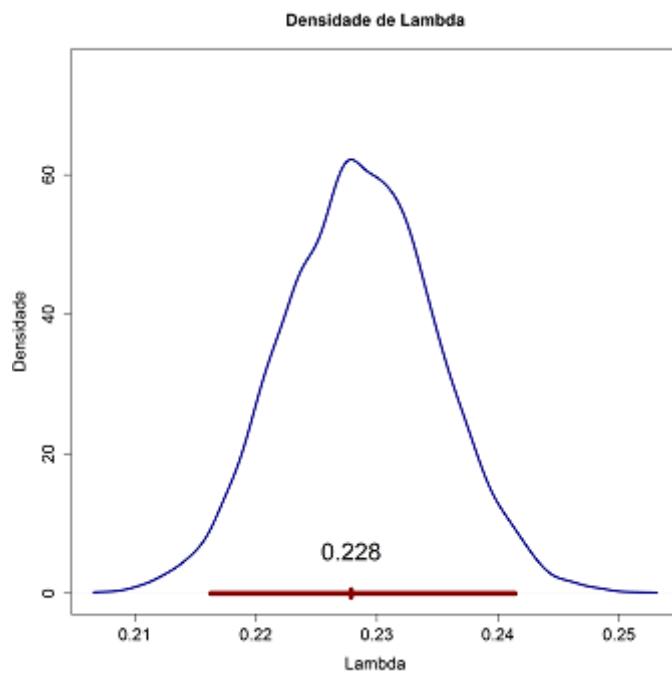
Esse coeficiente foi utilizado pois possui propriedade de simetria e é invariante quanto a mudança de ordem das linhas ou colunas - varia de 0, não associação, a 1, total associação. As estimativas pontuais e os intervalos de credibilidade (95%) para o valor de associação  $\lambda$  foram calculados sob abordagem Bayesiana.

A seguir são apresentados os gráficos com as distribuições de densidade de  $\lambda$ , bem como, as estimativas pontuais e intervalares encontradas para as associações estatisticamente significantes. Nesta primeira análise foram avaliadas as associações entre todas as variáveis contidas na base de dados; em um segundo momento, são avaliadas as associações entre a variável Tradição tecnológica e as demais variáveis categóricas.

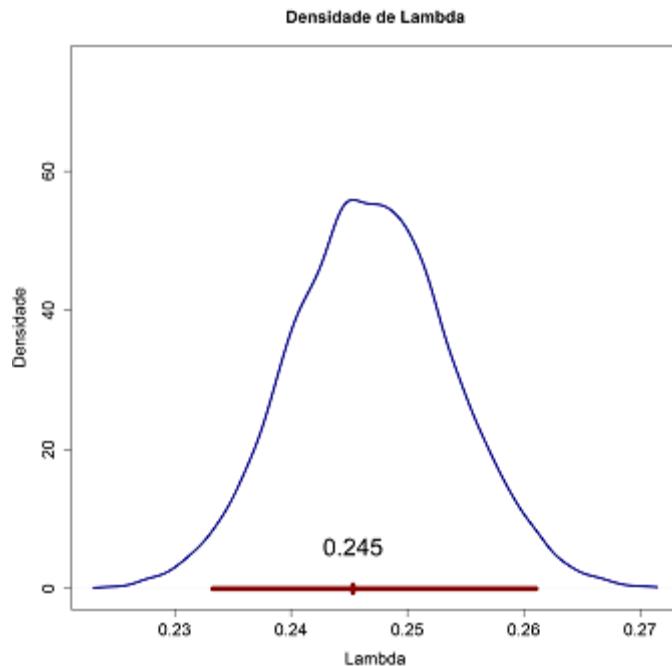
Além de algumas associações já esperadas, como sítio e proveniência, observou-se algumas outras associações consideráveis entre algumas variáveis – Antiplástico, acabamento face interna, manufatura, nível, categoria estrutural e Técnica de manufatura.



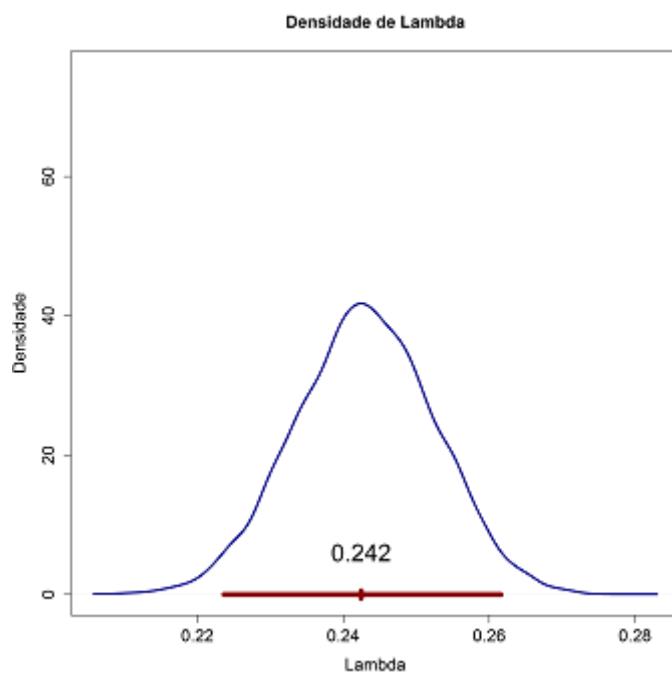
**Gráfico 16 – Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Sítio vs. Proveniência).**



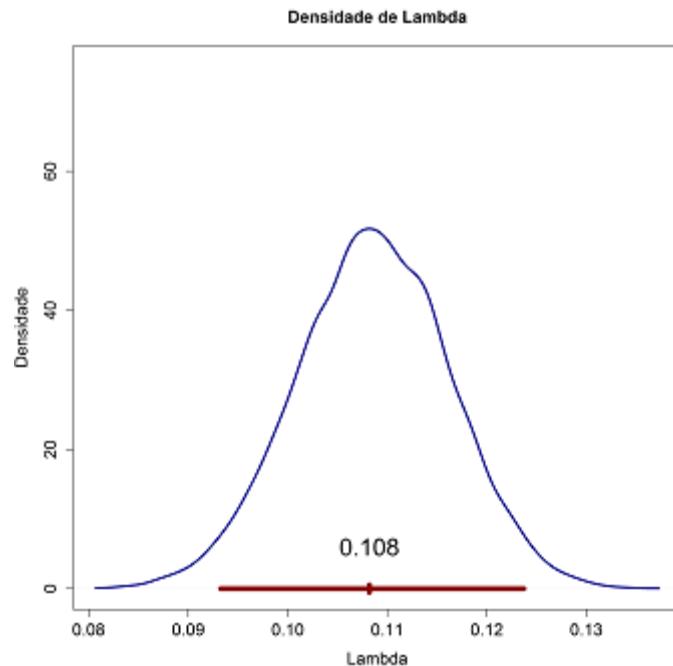
**Gráfico 17 – Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Sítio vs. Coloração externa).**



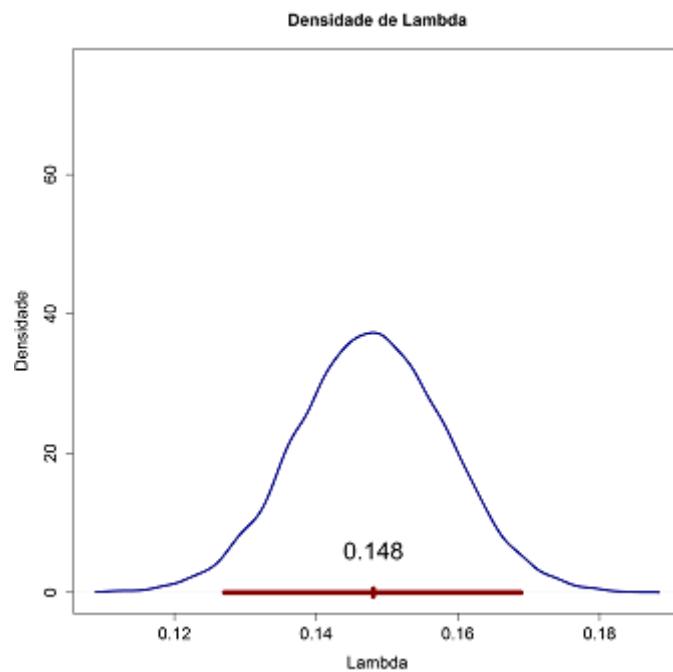
**Gráfico 18 – Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Sítio vs. Coloração interna).**



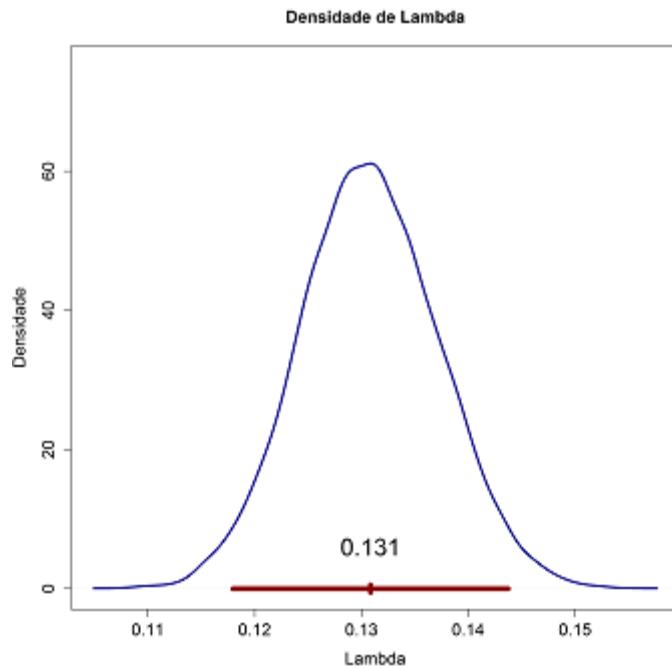
**Gráfico 19 – Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Proveniência vs. Antiplástico).**



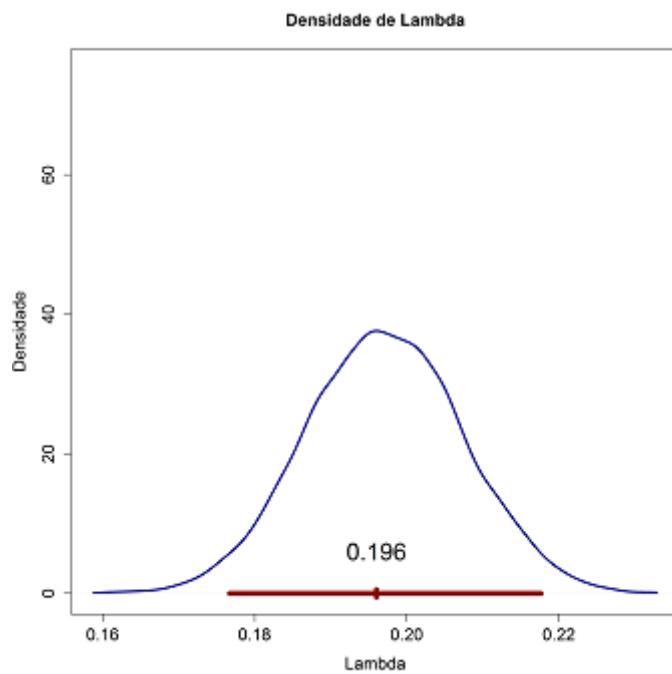
**Gráfico 20 – Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Proveniência vs. Acabamento face interna).**



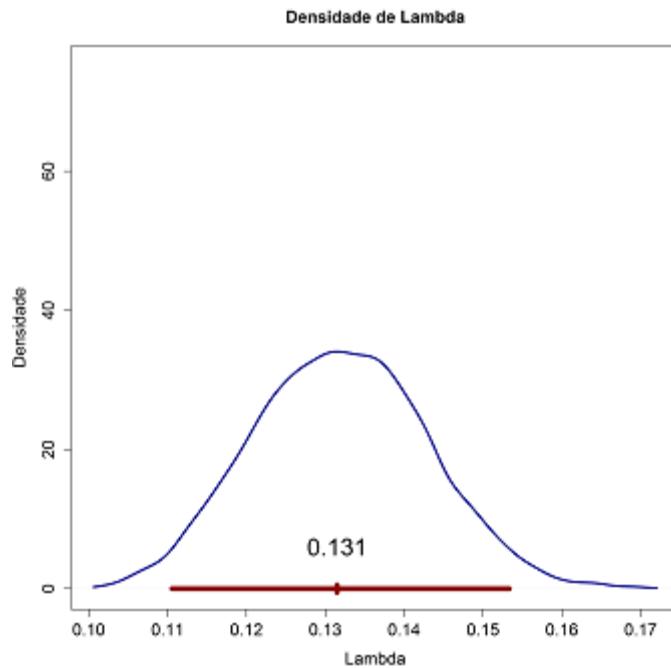
**Gráfico 21 – Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Proveniência vs. Manufatura).**



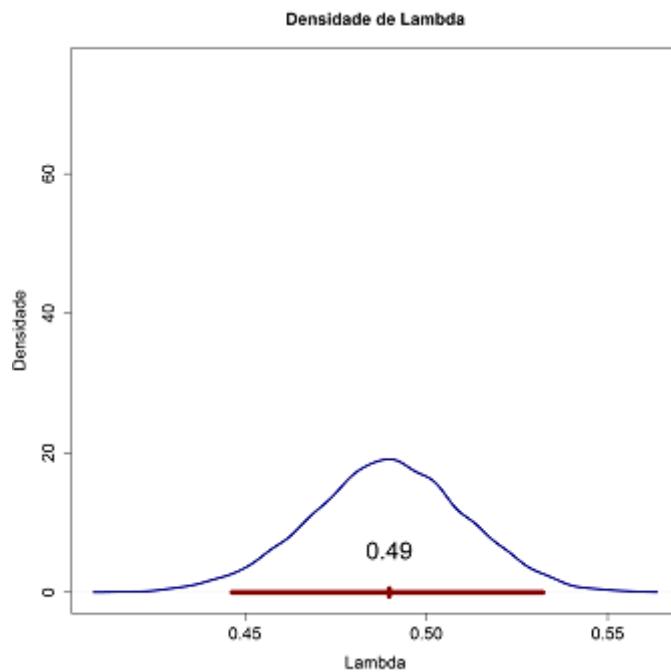
**Gráfico 22 – Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Proveniência vs. Nível).**



**Gráfico 23 – Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Acabamento face externa vs. Acabamento face interna).**



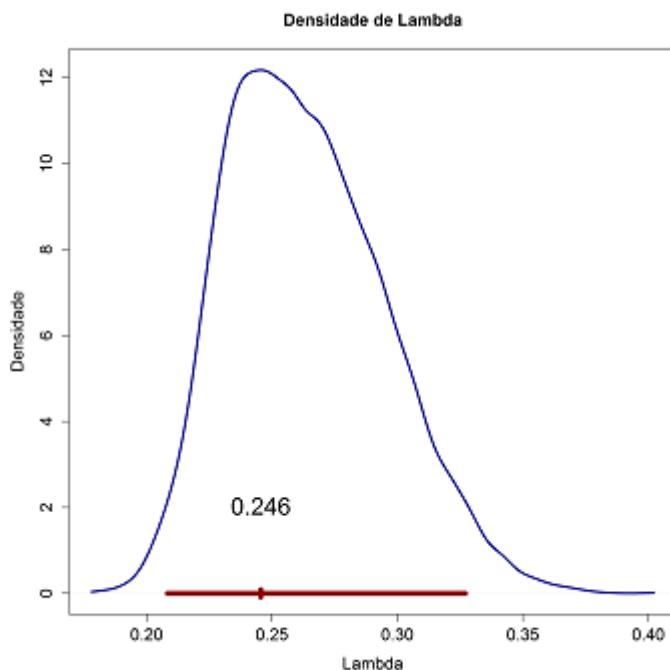
**Gráfico 24 – Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Acabamento face externa vs. Manufatura).**



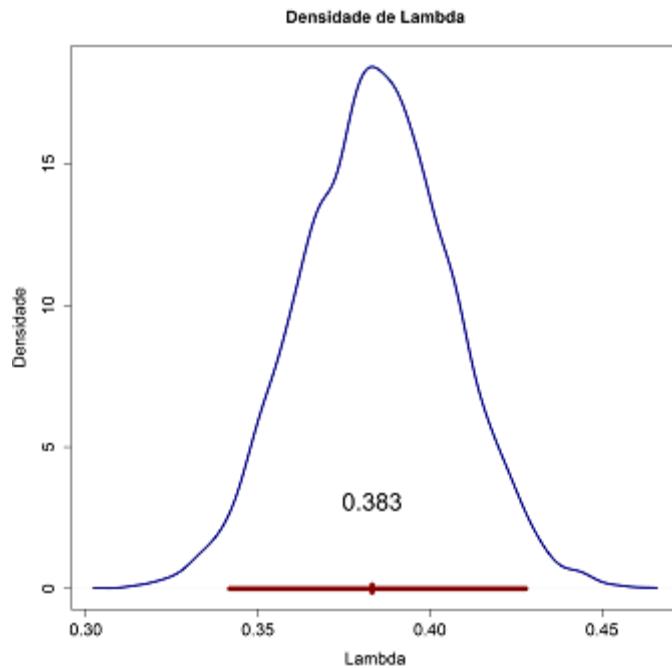
**Gráfico 25 - Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Acabamento lábio vs. Categoria).**

Com o objetivo de identificar as possíveis variáveis relacionadas à Tradição tecnológica, foram avaliadas as associações entre essa variável e a demais utilizando as informações apenas dos fragmentos já classificados pelo pesquisador. Dessa forma, as estimativas a seguir considerando as características dos 766 fragmentos classificados quanto a técnica tecnológica.

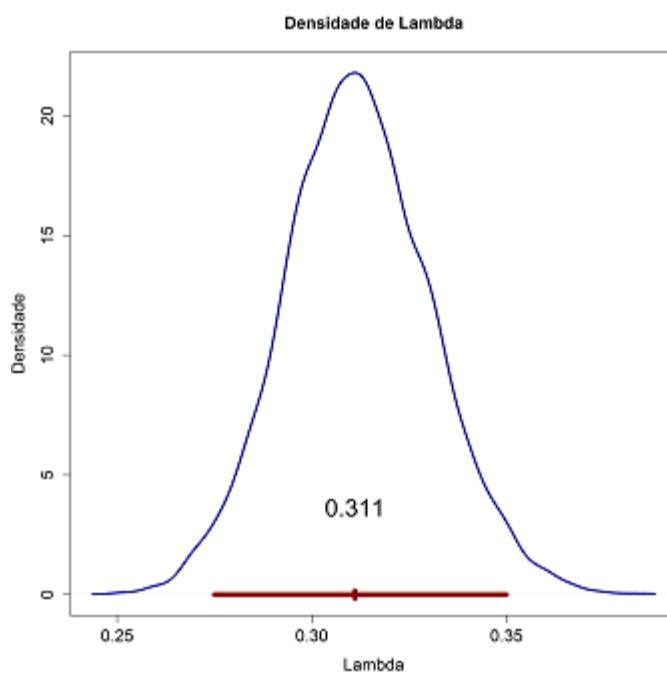
Analizando os valores encontrados para as estimativas de associação, observa-se que as variáveis sítio, proveniência, acabamento face externa, Antiplástico, manufatura e coloração (externa e interna) parecem ser relevantes para a distinção entre os grupos de Tradição tecnológica (Gráficos 26 a 32).



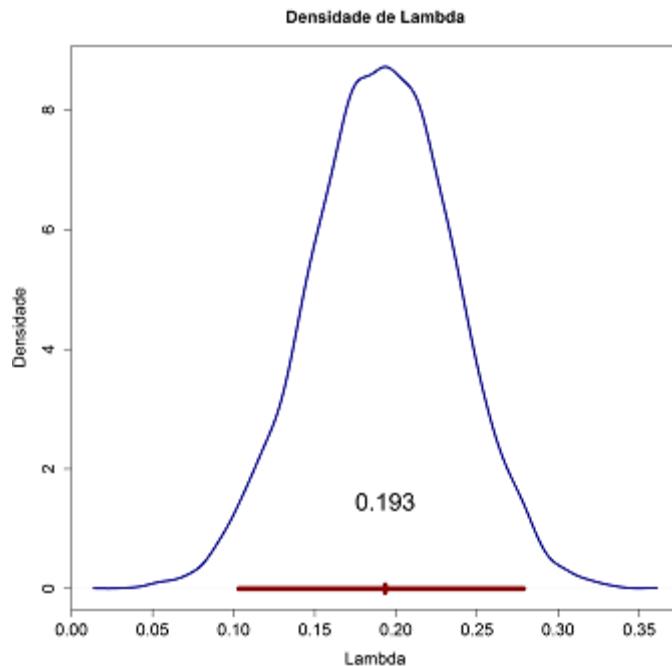
**Gráfico 26 - Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Tradição tecnológica vs. Sítio).**



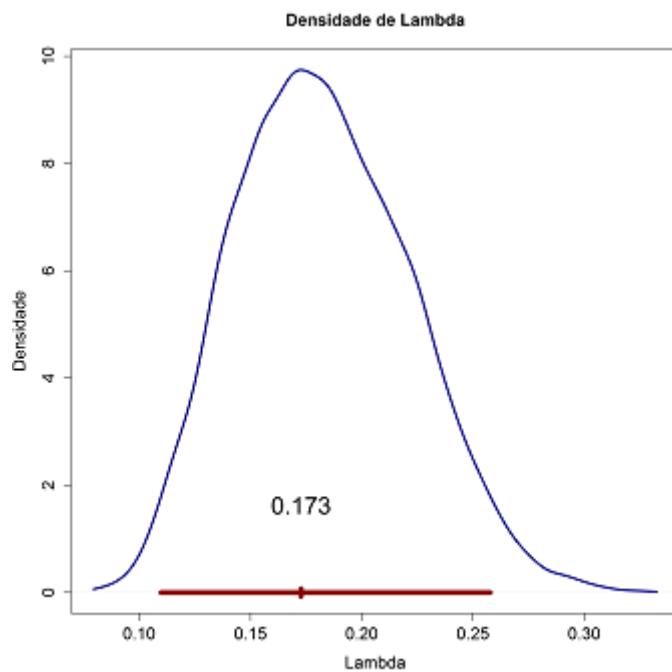
**Gráfico 27 - Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Tradição tecnológica vs. Proveniência).**



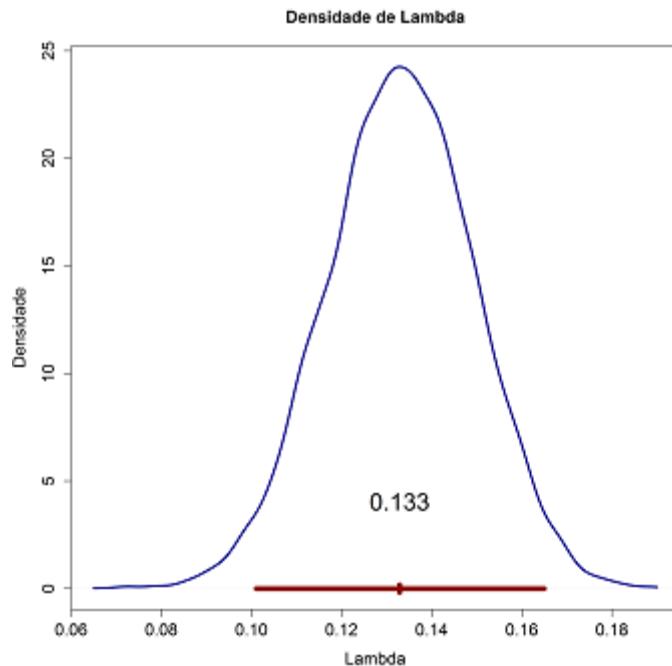
**Gráfico 28 - Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Tradição tecnológica vs. Acabamento face externa).**



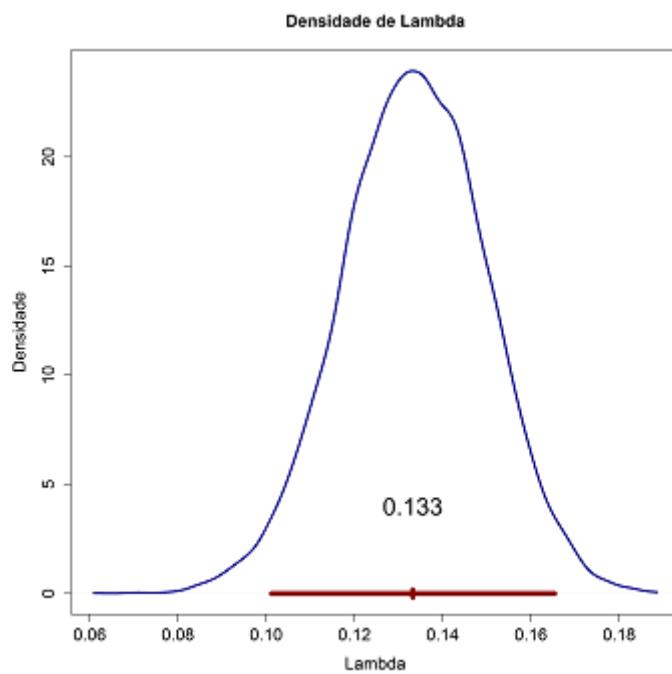
**Gráfico 29 - Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Tradição tecnológica vs. Antiplástico).**



**Gráfico 30 - Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Tradição tecnológica vs. Manufatura).**



**Gráfico 31 - Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Tradição tecnológica vs. Coloração interna).**



**Gráfico 32 - Densidade de probabilidade para o coeficiente de associação e estimativa pontual de  $\lambda$  (Tradição tecnológica vs. Coloração externa).**

## 5. Análise Inferencial

### 5.1 Análise de Conglomerados

Com o intuito de agrupar os fragmentos arqueológicos de acordo com suas características (físicas/composição e de utilização), foi aplicada a técnica de Análise de Conglomerados.

Por se tratar, em sua maioria, de variáveis categóricas, utilizou-se a métrica *Gower*<sup>3</sup> para obter a matriz de distâncias entre os fragmentos. Para a criação dos grupos, empregou-se o método de agrupamento de Ligação Completa (também conhecido como *técnica do vizinho mais distante*).

Levando-se em conta a análise de associação entre as variáveis feita na análise descritiva, as variáveis consideradas para o agrupamento foram as seguintes: Sítio, Nível, Antiplástico, Manufatura, Categoria estrutural, Acabamento da face externa, Coloração externa, Coloração interna, Núcleo, Função primária e Espessura.

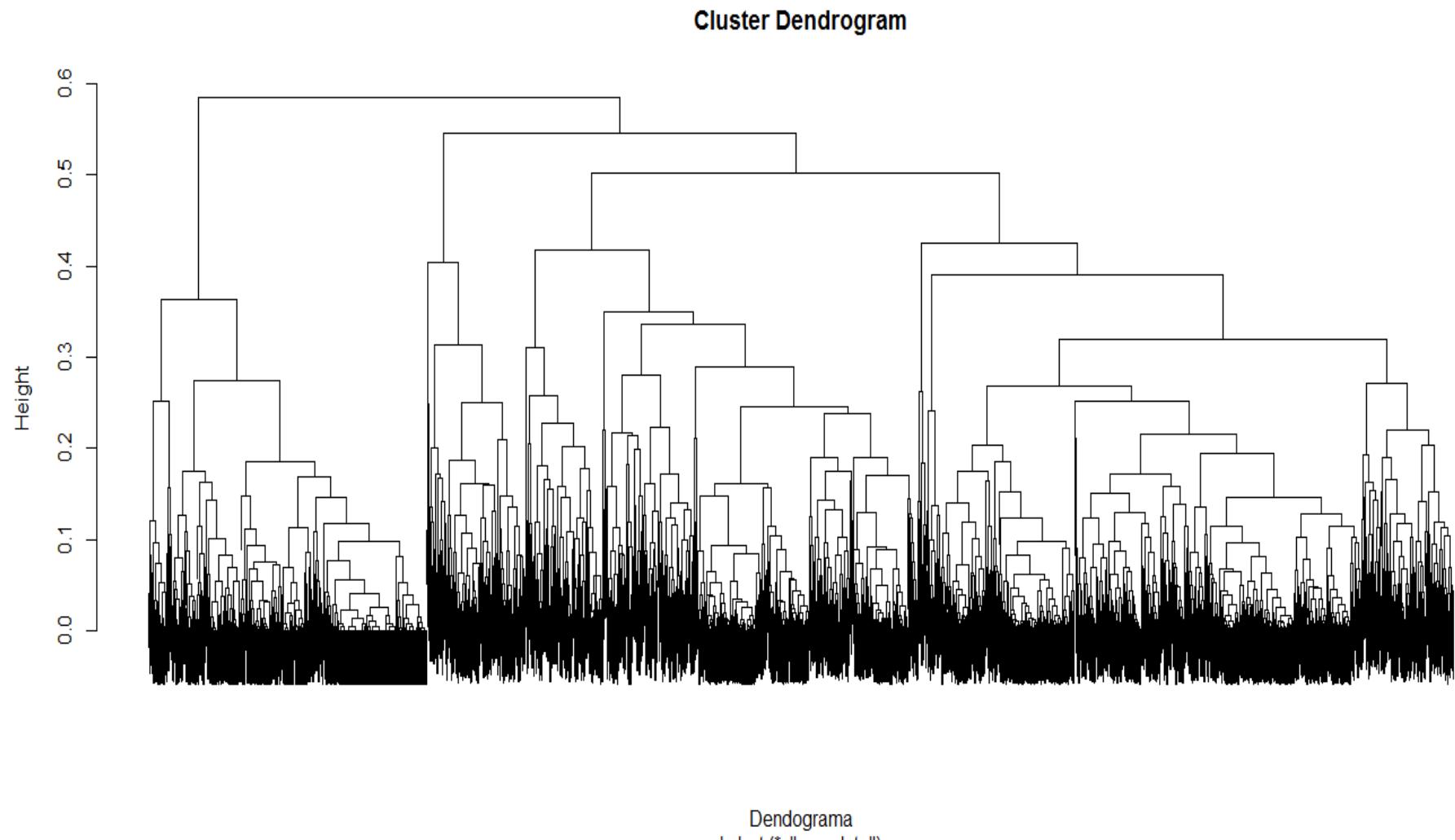
Apenas para ilustração do agrupamento, é apresentado no Gráfico 33 o dendograma dos fragmentos cerâmicos obtido na análise.

Visando maximizar a semelhança entre os fragmentos dentro de cada grupo formado e a dissimilaridade quando compara-se fragmentos entre grupos distintos, e de modo que o número de grupos fosse razoável nesse estudo<sup>4</sup>, optou-se pela formação de 4 grupos (Figura 29).

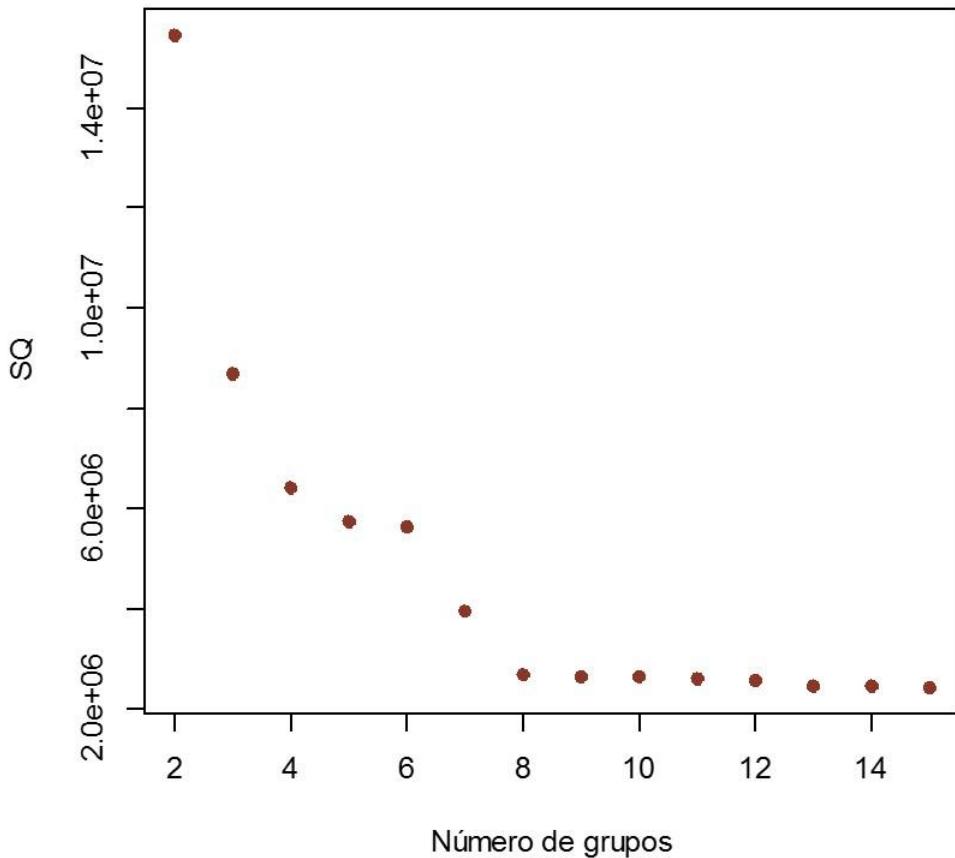
---

<sup>3</sup> Uma breve explicação da métrica *Gower*, bem como, da técnica de agrupamento utilizada pode ser vista no Apêndice B.

<sup>4</sup> É de conhecimento do pesquisador que há quatro possíveis tipos de tradições tecnológicas distintas. Dessa forma, optou-se por formar 4 grupos na Análise de Conglomerados.



**Gráfico 33 – Dendograma para os fragmentos cerâmicos obtido na Análise de Conglomerados.**



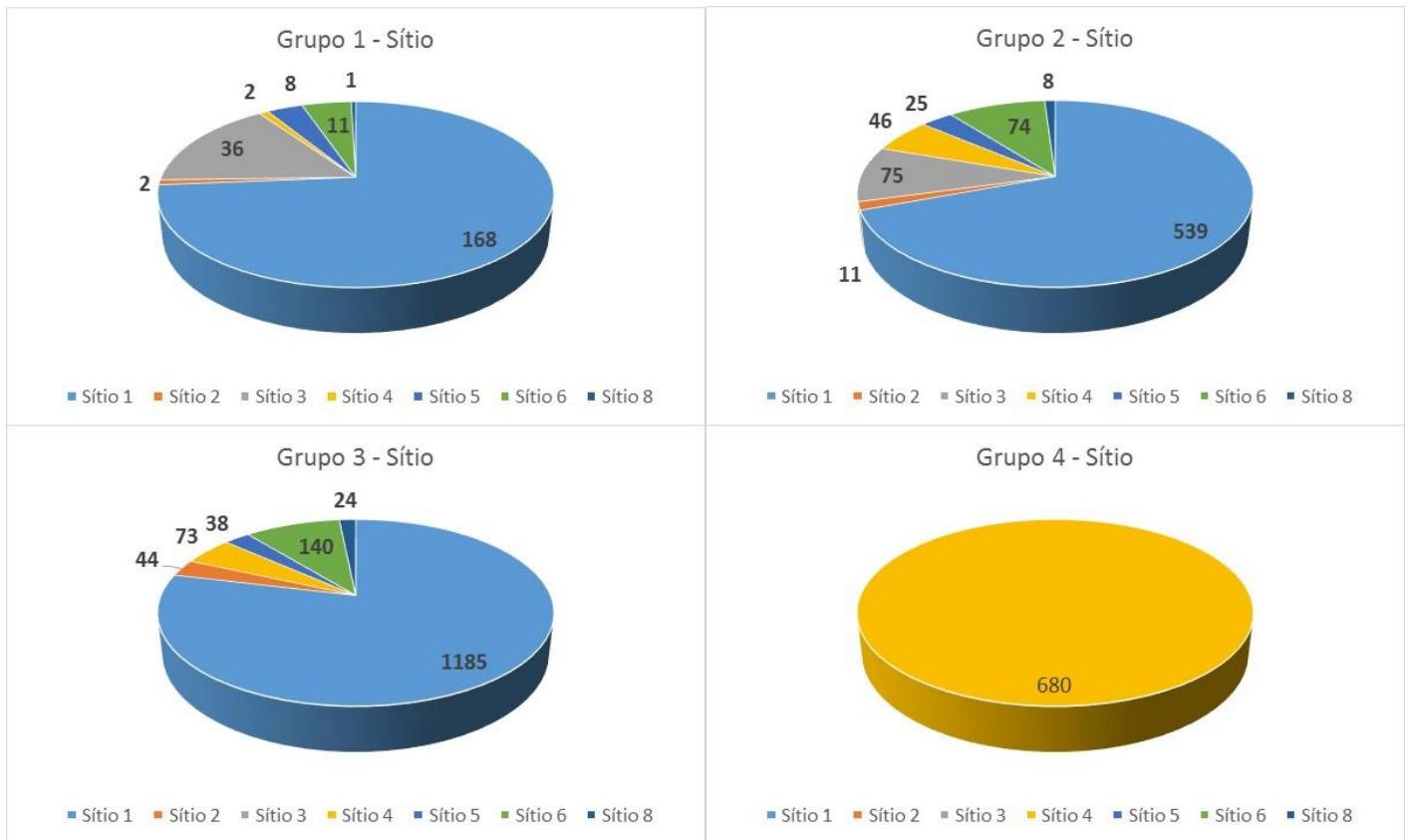
**Figura 29 – Gráfico da Soma de Quadrados na Análise de Conglomerados para determinar o número de grupos a serem formados.**

De forma a caracterizar cada grupo formado, apresenta-se a seguir a distribuição das variáveis em cada grupo.

**Tabela 35 – Distribuição dos grupos pela variável Tradição tecnológica.**

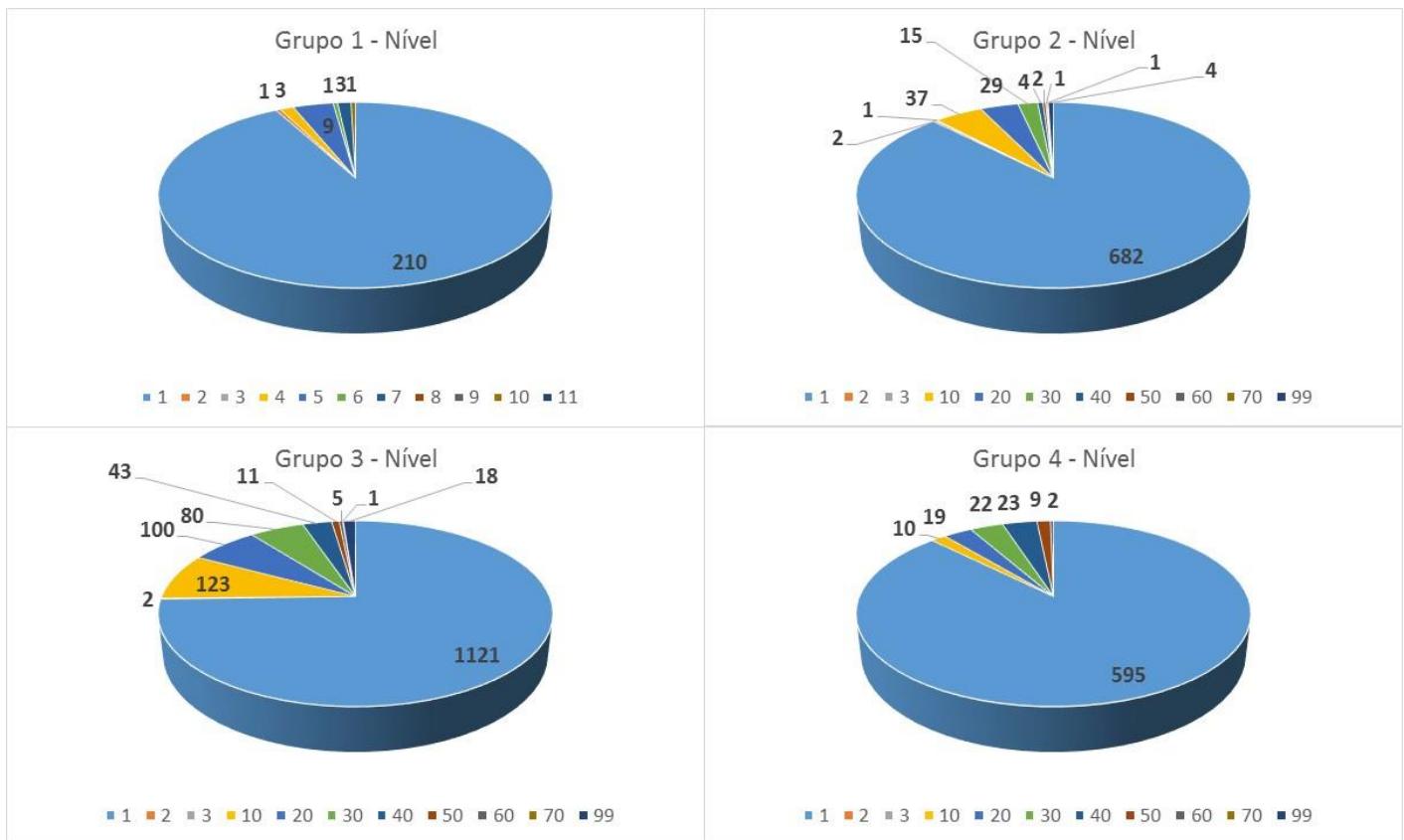
Grupo	Tradição			
	Guarani	Jacadigo	Kadiwéu	99
1	86 (17%)	12 (7%)	1 (1%)	129 (5%)
2	<b>315 (62%)</b>	<b>72 (43%)</b>	7 (8%)	384 (16%)
3	69 (14%)	<b>84 (50%)</b>	15 (18%)	1336 (55%)
4	36 (7%)	0	<b>69 (75%)</b>	575 (24%)

Na Tabela 35 observa-se que fragmentos classificados na tradição Guarani foram, em sua maioria, agrupados no Grupo 2. Na tradição Kadiwéu, há uma predominância no Grupo 4 e, na tradição Jacadigo, praticamente metade dos fragmentos foram classificados no Grupo 3 e a outra metade no Grupo 2. As peças não-classificadas foram alocadas, em sua maioria, no Grupo 3.



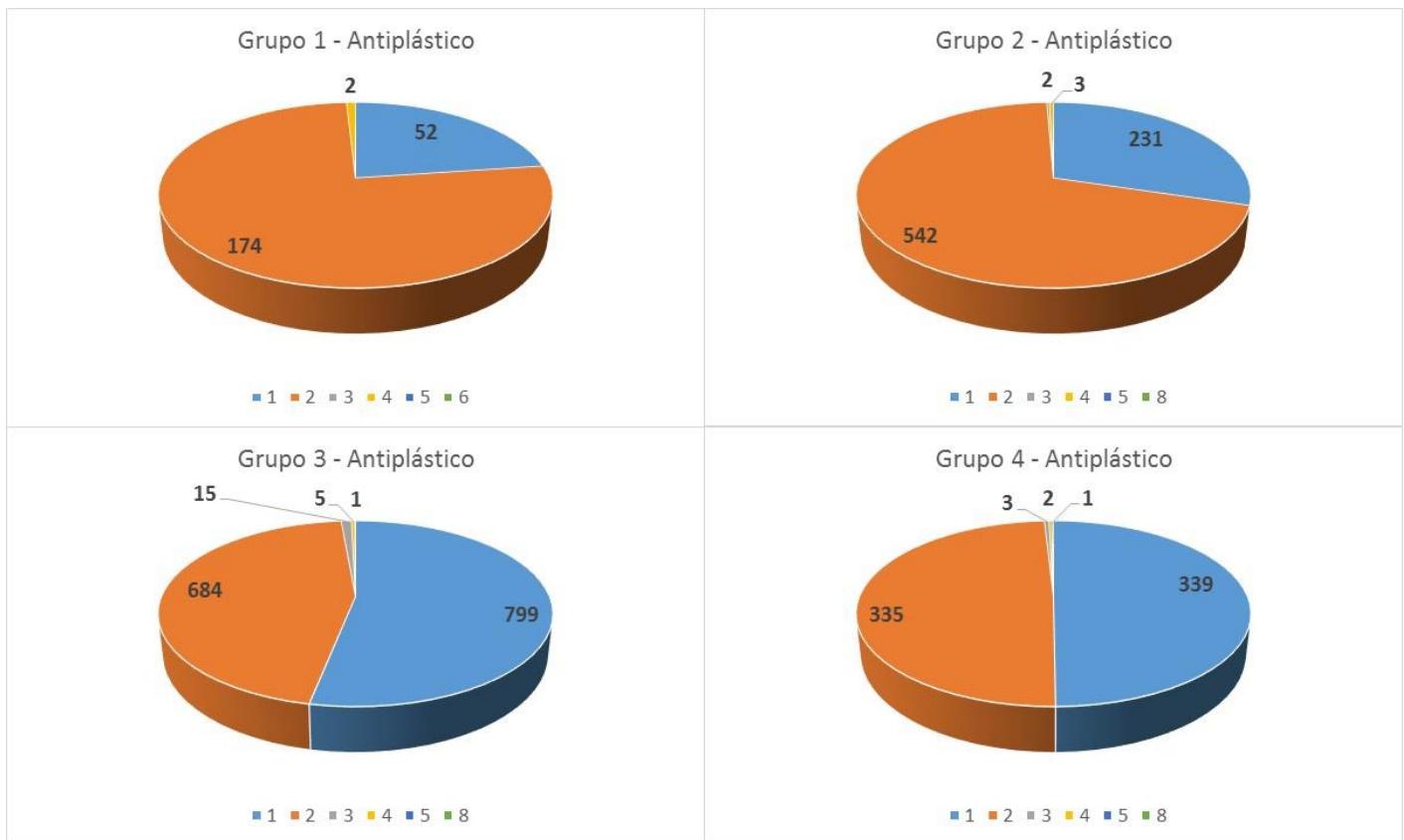
**Figura 30 – Distribuição da variável Sítio por grupo formado.**

Com relação à distribuição por Sítio (Figura 30), o grupo 4 é totalmente formado por fragmentos oriundos do Sítio 4; já os demais grupos tem peças, em sua maioria, do Sítio 1, sendo que nos grupos 1 e 2 há uma predominância secundária de peças do Sítio 3 e no grupo 3 essa predominância secundária se dá por fragmentos do Sítio 6.



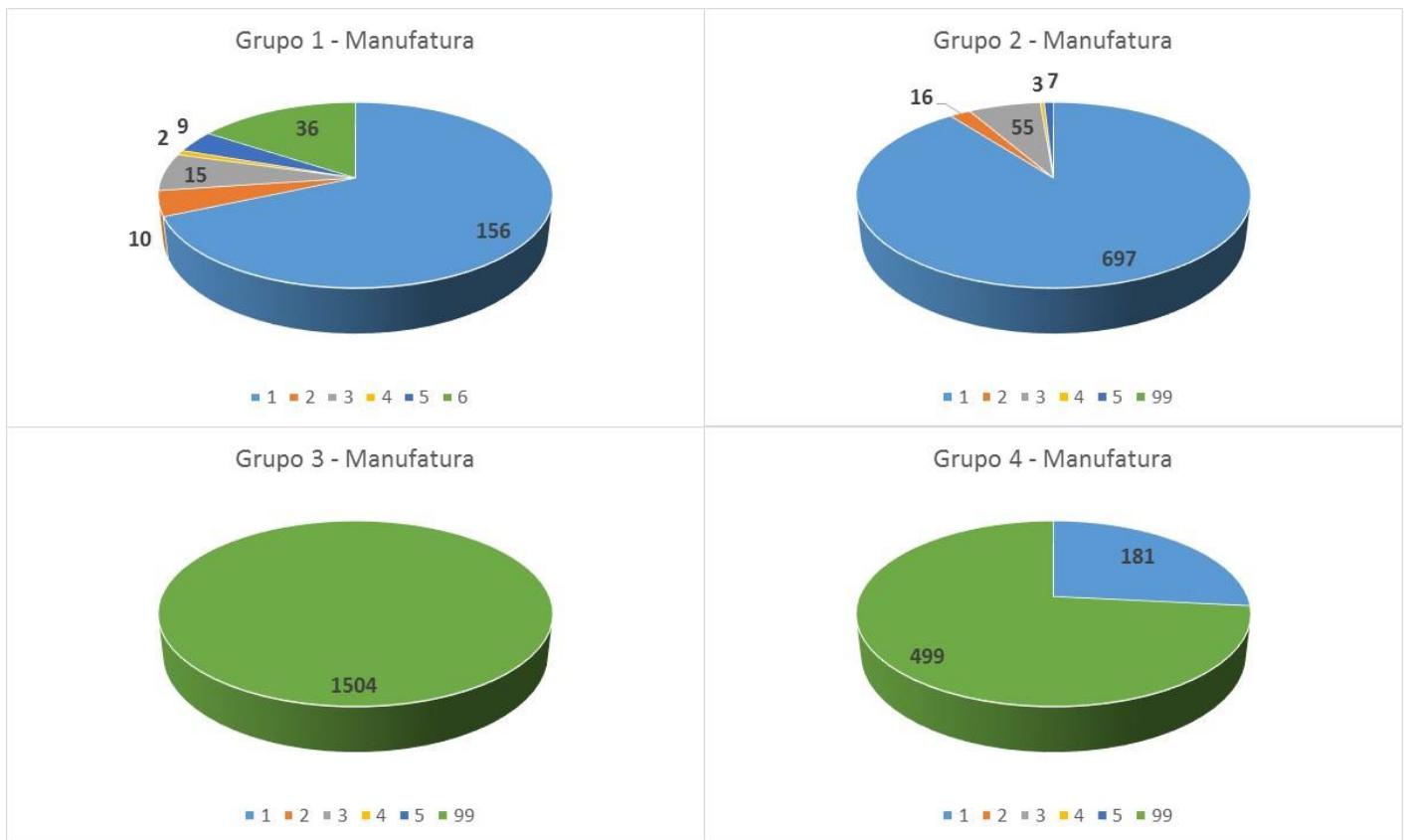
**Figura 31 – Distribuição da variável Nível de superfície por grupo formado.**

Para a variável Nível de superfície (Figura 31), sabe-se que a maioria dos fragmentos foram coletados no Nível 0 (superfície do solo), porém, destaca-se no grupo 3 uma frequência notável de fragmentos provenientes de níveis entre 10-30 cm abaixo da superfície (cerca de 20% no grupo).



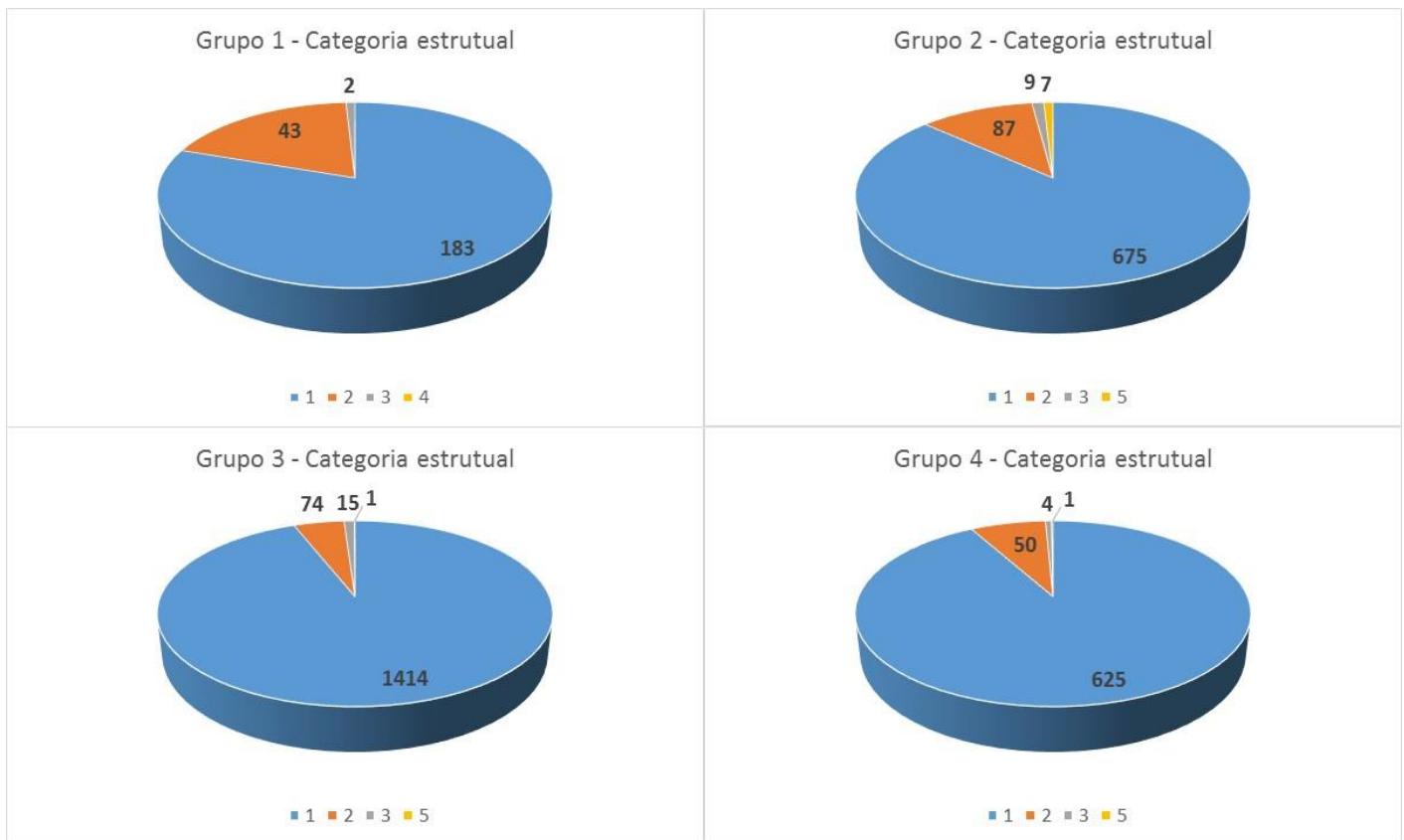
**Figura 32 – Distribuição da variável Antiplástico por grupo formado.**

Na distribuição da variável *Antiplástico* (Figura 32), vemos que nos grupos 1 e 2 há uma predominância da categoria 2 (caco moído); no grupo 3, porém, essa predominância encontra-se na categoria 1 (mineral) e, no grupo 4, as duas categorias se apresentam, praticamente, com a mesma distribuição de frequências.



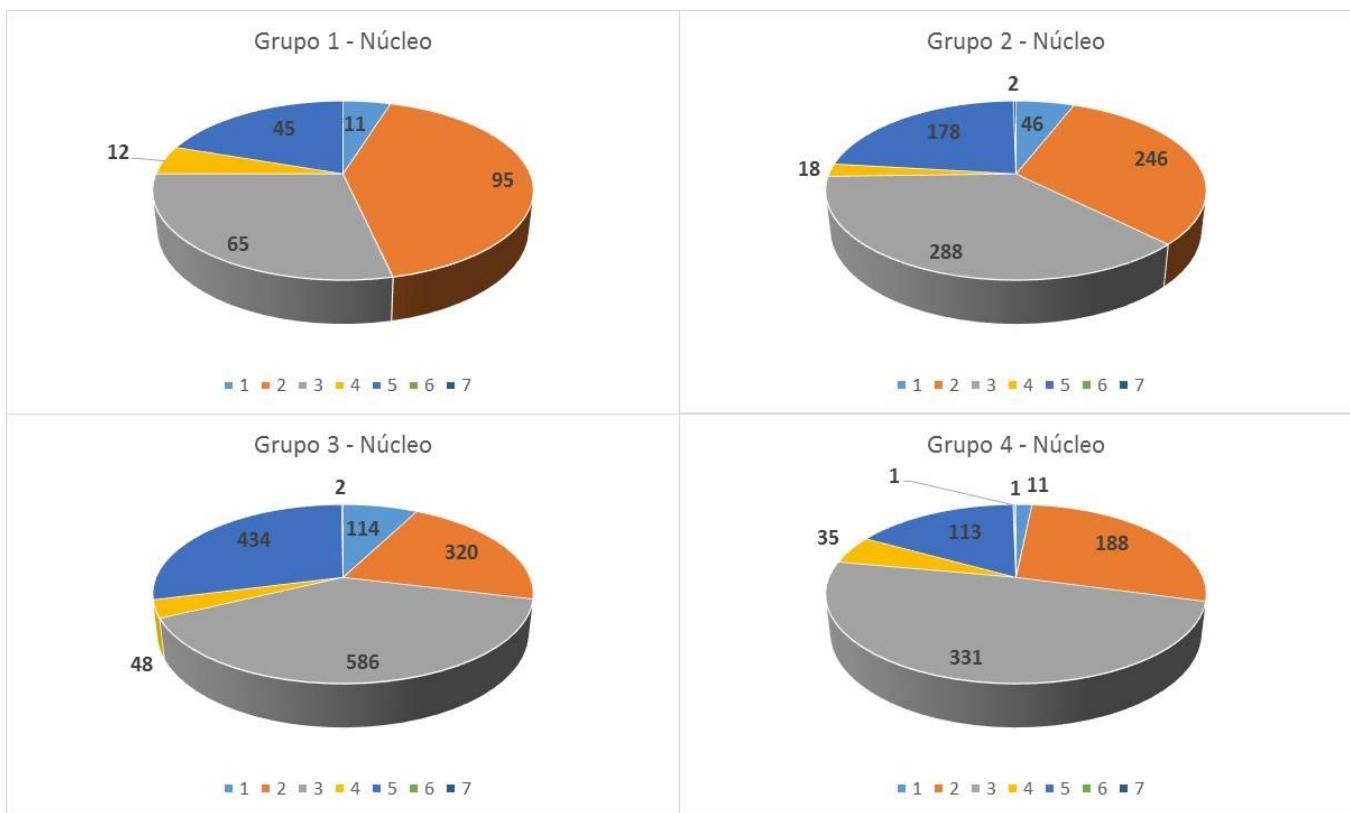
**Figura 33 – Distribuição da variável Manufatura por grupo formado.**

A variável *Manufatura* (Figura 33) é uma das que apresentam maior distinção na distribuição quando comparados os grupos. No grupo 1 as duas categorias que apresentam maiores frequências são *Roletado* e *não-identificado*, no grupo 2 *Roletado* e *Misto*, no grupo 3, por sua vez só há fragmentos não-identificados com relação à Técnica de manufatura e, por sua vez, no grupo 4, há apenas fragmentos pertencentes às categorias *Roletado* e *não-identificado*.

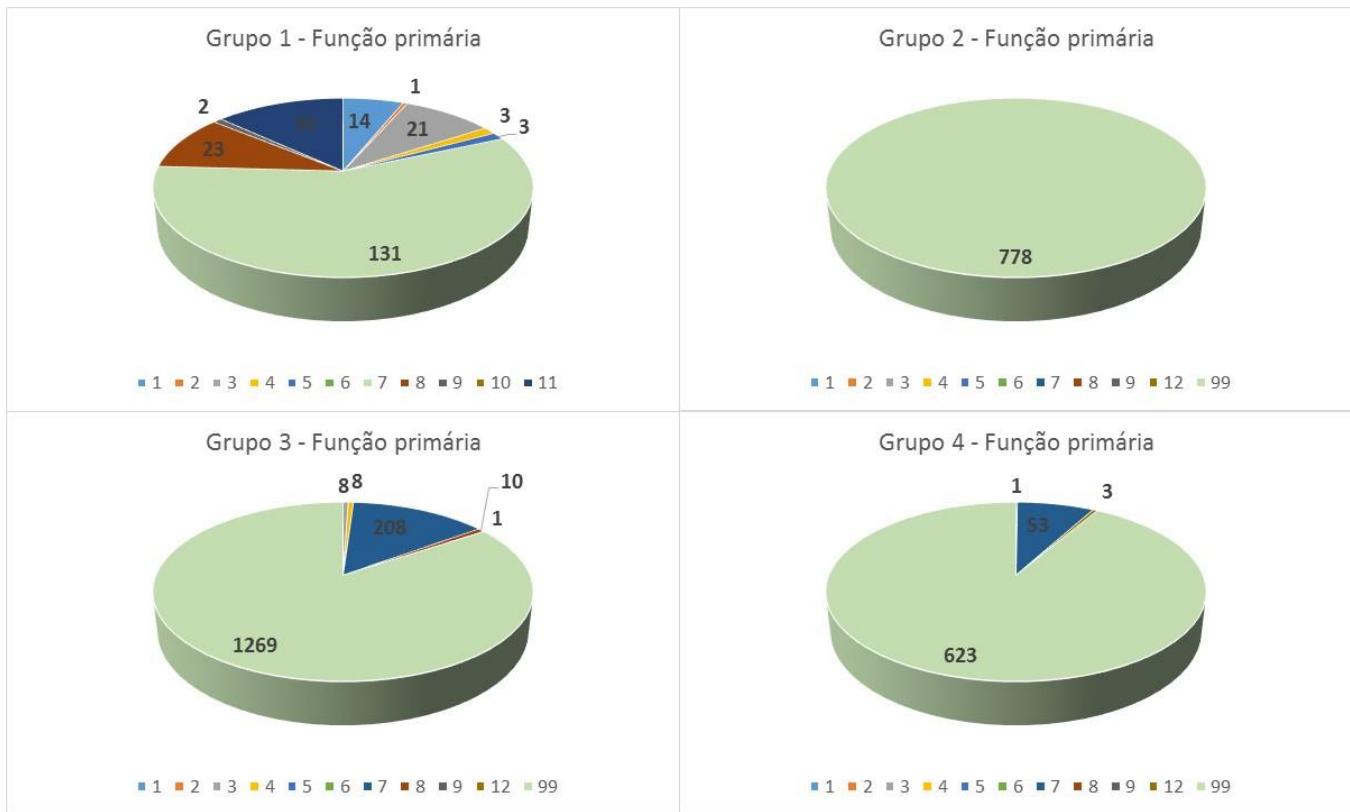


**Figura 34 – Distribuição da variável Categoria estrutural por grupo formado.**

Para a variável *Categoria estrutural* do fragmento (Figura 34) os quatro grupos formados são predominantemente formados por fragmentos de *parede* (cerca de 80%), seguido de peças de *borda*. Exceto no grupo 1 há, pelo menos, 1 fragmento de *apêndice* classificado nos demais grupos.

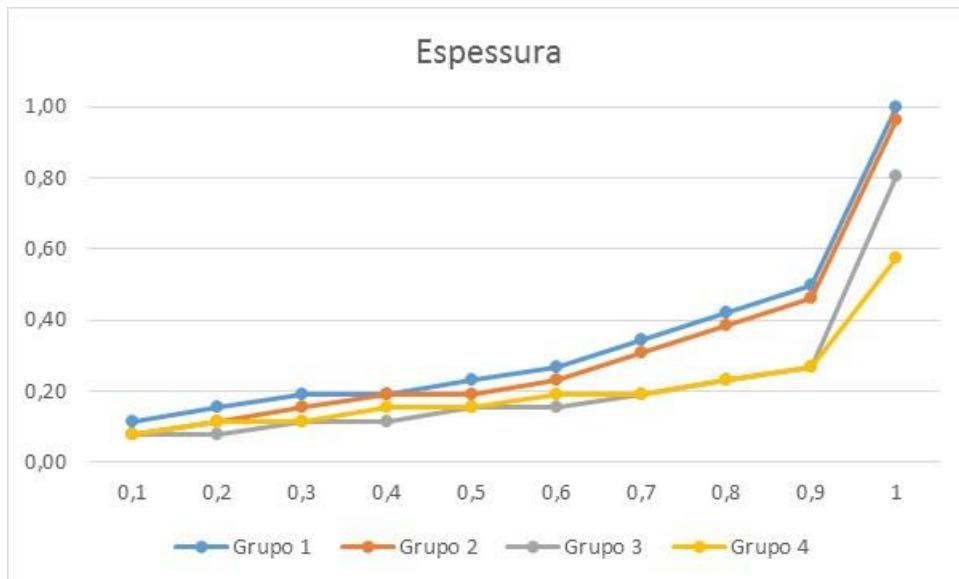


**Figura 35 – Distribuição da variável Núcleo por grupo formado.**



**Figura 36 – Distribuição da variável Função primária por grupo formado.**

A variável *Núcleo* (Figura 35) apresenta distribuições bem parecidas nos 4 grupos, sendo que as categorias 2 e 3 são predominantes. Observando a distribuição da variável *Função primária* (Figura 36), vemos que a maioria dos fragmentos não foram classificados com relação a esta característica, porém, grande parte dos que foram classificados em uma das demais categorias, foram alocados no grupo 1; alguns fragmentos classificados na categoria 7 foram alocados nos grupos 3 e 4 e, por sua vez, o grupo 2 apresenta, apenas, fragmentos não classificados com relação à função primária.



**Figura 37 – Gráfico de percentis para a variável Espessura por grupo formado.**

Com relação à *Espessura* dos fragmentos (Figura 37) observa-se que os grupos 1 e 2 são caracterizados por apresentaram peças mais espessas quando comparados com os outros dois grupos (3 e 4).

A variável *acabamento face externa* (Tabela 36) apresenta, nos grupos 1 e 2, predominância das categorias Alisamento (1), Polimento (2) e Corrugado – cerca de 22% do total por categoria. Já, nos grupos 3 e 4 destaca-se, apenas, as duas primeiras categorias – responsáveis por cerca de 65-80% do total.

Para as variáveis relacionados à coloração (Tabelas 37 e 38), como já observado na análise descritiva, sabe-se que há uma distribuição dos fragmentos por todas as categorias – fato esse observado também nos grupos 1, 2 e 3, porém, destaca-se que o grupo 4 é formado, exceto por 1 fragmento, apenas por peças não-classificados com relação à coloração (externa e interna).

**Tabela 36 – Distribuição da variável Acabamento face externa por grupo formado.**

Acabamento face externa	Grupo							
	1		2		3		4	
Categorias	Freq.	%	Freq.	%	Freq.	%	Freq.	%
1	48	21,0	220	28,3	630	41,9	342	50,4
2	56	24,5	149	19,2	660	43,9	104	15,3
3		0,0	1	0,1	7	0,5	5	0,7
4	53	23,1	278	35,7	6	0,4	21	3,1
9		0,0		0,0	1	0,1		0,0
10	3	1,3	9	1,2	12	0,8	5	0,7
13	1	0,4	1	0,1	3	0,2		0,0
16		0,0	8	1,0	29	1,9		0,0
17	1	0,4		0,0	1	0,1	1	0,1
19		0,0	6	0,8		0,0		0,0
20		0,0		0,0	1	0,1		0,0
21		0,0		0,0	2	0,1		0,0
23	1	0,4	3	0,4	7	0,5		0,0
24		0,0	4	0,5	14	0,9	23	3,4
25	1	0,4		0,0		0,0		0,0
26		0,0	1	0,1	2	0,1		0,0
27		0,0		0,0	2	0,1	1	0,1
30	8	3,5	17	2,2	28	1,9	5	0,7
31	1	0,4	16	2,1	37	2,5	61	9,0
32	10	4,4	15	1,9	11	0,7	21	3,1
33	1	0,4	11	1,4	29	1,9	44	6,5
34	1	0,4	2	0,3		0,0		0,0
35	1	0,4	1	0,1	1	0,1		0,0
36		0,0		0,0	2	0,1		0,0
37	1	0,4		0,0	2	0,1		0,0
38	1	0,4	2	0,3	3	0,2		0,0
39	3	1,3	3	0,4	2	0,1		0,0
40	2	0,9		0,0		0,0		0,0
43	1	0,4		0,0		0,0		0,0
44		0,0	7	0,9	8	0,5	8	1,2
46		0,0	13	1,7		0,0		0,0
48		0,0	3	0,4	2	0,1	30	4,4
49		0,0		0,0		0,0	1	0,1
51	1	0,4		0,0		0,0		0,0
53		0,0		0,0		0,0	1	0,1
57		0,0	1	0,1		0,0		0,0
59		0,0	1	0,1		0,0		0,0
60		0,0	1	0,1		0,0		0,0
61		0,0		0,0		0,0	2	0,3

62		0,0		0,0		0,0		1	0,1
63		0,0		0,0		0,0		1	0,1
64		0,0		0,0		0,0		2	0,3
99	34	14,8	5	0,6	2	0,1			0,0
Total	229	100	778	100	1504	100	679	100	

**Tabela 37 – Distribuição da variável Coloração externa por grupo formado.**

Coloração externa	Grupo							
	1		2		3		4	
Categorias	Freq.	%	Freq.	%	Freq.	%	Freq.	%
1	2	0,9	8	1,0	18	1,2		0
2	20	8,8	61	7,8	193	12,8		0
3	16	7,0	46	5,9	101	6,7		0
4	25	11,0	62	8,0	94	6,3		0
5	11	4,8	17	2,2	62	4,1		0
6	19	8,3	56	7,2	94	6,3		0
7	9	3,9	18	2,3	46	3,1		0
8	2	0,9	2	0,3	7	0,5		0
9	5	2,2	7	0,9	24	1,6		0
10	10	4,4	34	4,4	67	4,5		0
11	13	5,7	63	8,1	49	3,3		0
12	11	4,8	64	8,2	34	2,3		0
13	3	1,3	8	1,0	19	1,3		0
14	3	1,3	9	1,2	13	0,9		0
15	1	0,4	16	2,1	7	0,5		0
16	2	0,9	5	0,6	11	0,7		0
17	1	0,4		0,0	3	0,2		0
18	1	0,4		0,0		0,0		0
19	1	0,4		0,0	4	0,3		0
20	1	0,4	5	0,6		0,0		0
21	1	0,4	15	1,9	42	2,8		0
22	9	3,9	27	3,5	78	5,2		0
23		0,0		0,0	3	0,2		0
24	5	2,2	37	4,8	58	3,9		0
25	2	0,9		0,0	5	0,3		0
26	6	2,6	15	1,9	30	2,0		0
27		0,0	3	0,4	2	0,1		0
30	1	0,4	6	0,8	18	1,2		0
31	2	0,9	5	0,6	19	1,3		0
32		0,0	7	0,9	8	0,5		0
33	17	7,5	37	4,8	82	5,5		0
34	6	2,6	31	4,0	58	3,9		0

35	1	0,4		0,0	5	0,3	0
36	1	0,4	11	1,4	27	1,8	0
37	7	3,1	17	2,2	62	4,1	0
38	6	2,6	29	3,7	63	4,2	0
39		0,0		0,0	1	0,1	0
40		0,0	2	0,3	7	0,5	0
41		0,0	1	0,1	6	0,4	0
42		0,0	1	0,1	2	0,1	0
43		0,0	1	0,1		0,0	0
44	1	0,4	4	0,5	6	0,4	0
45		0,0	2	0,3	4	0,3	0
46		0,0	17	2,2	7	0,5	0
47		0,0		0,0	4	0,3	0
48	2	0,9	5	0,6	18	1,2	0
50	1	0,4	2	0,3	1	0,1	0
51		0,0	4	0,5		0,0	0
52		0,0		0,0	4	0,3	0
53		0,0	2	0,3	13	0,9	0
55	1	0,4	1	0,1	1	0,1	0
56		0,0	3	0,4	2	0,1	0
57		0,0	1	0,1		0,0	0
58		0,0	1	0,1	5	0,3	0
59		0,0		0,0	1	0,1	0
61	1	0,4	1	0,1	3	0,2	0
62		0,0	2	0,3		0,0	0
63	1	0,4		0,0	1	0,1	0
64		0,0	2	0,3	6	0,4	0
67		0,0		0,0	1	0,1	0
70		0,0	1	0,1		0,0	0
71		0,0	3	0,4	3	0,2	1 0,1
72		0,0	1	0,1	1	0,1	0
73		0,0		0,0	1	0,1	0
99	1	0,4		0,0		0,0	679 99,9
Total	228	100	778	100	1504	100	680 100

**Tabela 38 – Distribuição da variável Coloração interna por grupo formado.**

Coloração interna	Grupo							
	1		2		3		4	
Categorias	Freq.	%	Freq.	%	Freq.	%	Freq.	%
1	4	1,8	8	1,0	15	1,0		0
2	30	13,2	69	8,9	128	8,5		0
3	13	5,7	38	4,9	83	5,5		0

4	23	10,1	66	8,5	67	4,5	0
5	10	4,4	25	3,2	73	4,9	0
6	29	12,7	111	14,3	218	14,5	0
7	11	4,8	29	3,7	39	2,6	0
8	4	1,8	5	0,6	5	0,3	0
9	3	1,3	5	0,6	8	0,5	0
10	3	1,3	34	4,4	81	5,4	0
11	7	3,1	35	4,5	35	2,3	0
12	13	5,7	41	5,3	29	1,9	0
13	1	0,4	5	0,6	13	0,9	0
14	1	0,4	4	0,5	8	0,5	0
15	4	1,8	12	1,5	13	0,9	0
16	2	0,9	6	0,8	9	0,6	0
17		0,0		0,0	2	0,1	0
19	1	0,4	1	0,1	1	0,1	0
20	1	0,4	6	0,8	3	0,2	0
21	4	1,8	11	1,4	32	2,1	0
22	7	3,1	29	3,7	82	5,5	0
23		0,0		0,0	2	0,1	0
24	10	4,4	25	3,2	57	3,8	0
25	1	0,4	1	0,1	6	0,4	0
26	5	2,2	16	2,1	23	1,5	0
28	1	0,4	2	0,3	3	0,2	0
30	1	0,4	2	0,3	14	0,9	0
31	1	0,4	1	0,1	19	1,3	0
32	1	0,4	7	0,9	11	0,7	0
33	10	4,4	38	4,9	62	4,1	0
34	7	3,1	27	3,5	52	3,5	0
35		0,0		0,0	4	0,3	0
36	3	1,3	19	2,4	52	3,5	0
37		0,0	13	1,7	44	2,9	0
38	5	2,2	32	4,1	68	4,5	0
40		0,0	4	0,5	10	0,7	0
41		0,0	1	0,1	7	0,5	0
42	1	0,4	1	0,1	7	0,5	0
43		0,0	1	0,1		0,0	0
44		0,0	6	0,8	6	0,4	0
45	3	1,3	4	0,5	1	0,1	0
46	2	0,9	12	1,5	14	0,9	0
47		0,0		0,0	2	0,1	0
48	1	0,4	7	0,9	40	2,7	0
49		0,0	1	0,1	3	0,2	0
50	1	0,4	1	0,1	2	0,1	0

52		0,0		0,0	3	0,2		0
53		0,0	2	0,3	12	0,8		0
54		0,0		0,0	1	0,1		0
55		0,0	2	0,3	2	0,1		0
56	1	0,4	2	0,3	9	0,6		0
57		0,0	1	0,1	2	0,1		0
58	1	0,4		0,0	2	0,1		0
59		0,0	1	0,1	4	0,3		0
60		0,0		0,0	1	0,1		0
61		0,0	1	0,1	2	0,1		0
62	1	0,4	1	0,1	1	0,1		0
64		0,0	2	0,3	3	0,2		0
65		0,0	1	0,1		0,0		0
66		0,0	1	0,1		0,0		0
68		0,0		0,0	1	0,1		0
69		0,0		0,0	5	0,3		0
71	1	0,4	2	0,3		0,0	1	0,1
72		0,0	1	0,1	2	0,1		0
99		0,0		0,0	1	0,1	679	99,9
Total		228	100	778	100	1504	100	680
								100

## 5.2 Redes Bayesianas

Considerando que dos 3.190 fragmentos observados apenas 766 (24%) foram classificados quanto à Tradição tecnológica, deseja-se inferir o conhecimento dessa classificação para identificar em qual dos 3 grupos tecnológicos conhecidos cada fragmento pertence observadas as suas características.

Uma das técnicas estatísticas disponíveis para essa análise se utiliza de uma abordagem bayesiana e é conhecida como Redes Bayesianas<sup>5</sup>. Sucintamente, essa técnica considera as informações dos fragmentos já classificados para criar um *algoritmo (grafo) de classificação* de modo a classificar novos fragmentos em um dos três grupos possíveis (variável resposta do algoritmo – Tradição tecnológica). O resultado do algoritmo é, basicamente, um conjunto com as probabilidades de uma determinada peça pertencer a um dos 3 grupos possíveis. Dessa forma, pode-se determinar um ponto de corte na probabilidade de modo a classificar o novo fragmento no grupo em que a probabilidade de pertencer a este é maior que este ponto de corte.

As variáveis consideradas para a criação do grafo de classificação foram: Sítio, Nível, Antiplástico, Manufatura, Categoria, Núcleo e Espessura. Particularmente, por se

<sup>5</sup> Veja uma breve explicação da técnica no Apêndice B. Uma abordagem mais completa pode ser vista em (PEARL, 1988).

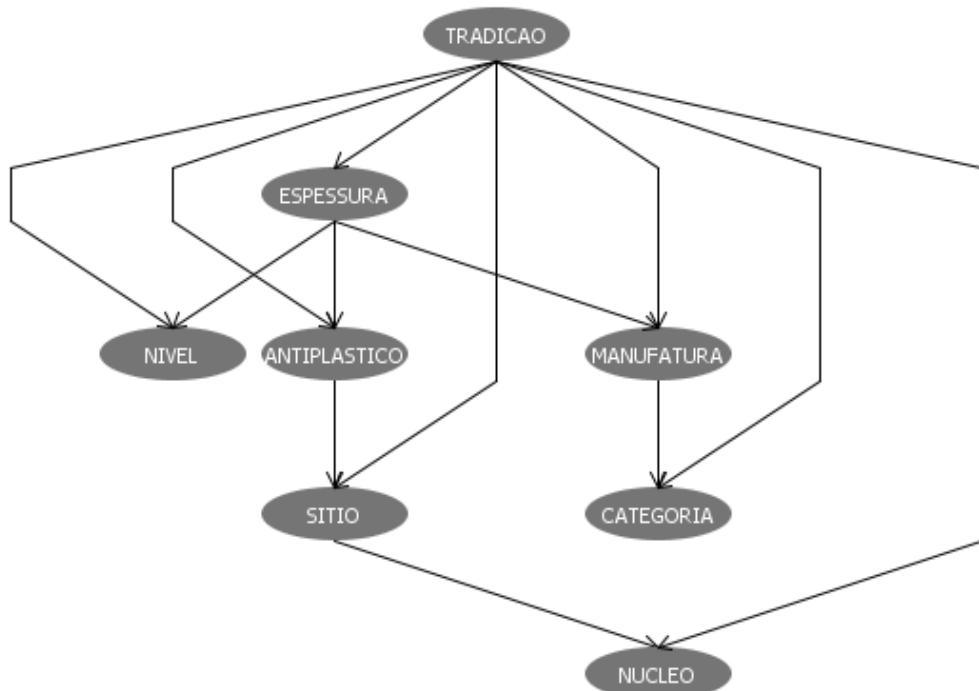
tratar de uma variável contínua, optou-se por aplicar, para cada valor assumido pela variável Espessura, a seguinte modificação, garantindo, dessa forma, valores entre 0 e 1:

$$e_{mi} = \frac{e_i - \min(e)}{\max(e) - \min(e)}$$

em que  $\min(e)$  e  $\max(e)$  são, respectivamente o mínimo e o máximo valor assumido pela variável Espessura e  $e_i$  é o valor original da variável Espessura para o i-ésimo fragmento.

Para a obtenção da Rede Bayesiana, portanto, considerou-se apenas os fragmentos já classificados e utilizou-se da técnica TAN para criar o algoritmo de classificação. O algoritmo TAN foi utilizado pois este permite considerar dependências/associações observadas entre as variáveis explicativas – associações essas que já vimos existir na análise de associações (parte descritiva). Essas associações estão relacionadas com a idéia de “pais e filhos” dentro de um Rede Bayesiana – isto quer dizer que, cada variável dependente de uma outra (pai) é chamada de filho. Em outras palavras, isto quer dizer que a distribuição de probabilidade da variável *filho* depende de qual categoria a variável *pai* pertence para um determinado fragmento.

A seguir apresenta-se a Rede Bayesiana gerada a partir do algoritmo:



**Figura 38 – Rede Bayesiana para classificação dos fragmentos arqueológicos.**

Tendo em mente a idéia de variáveis pai-filho, observamos na Rede Bayesiana gerada (Figura 38) que a variável *Espessura* torna-se pai das variáveis *Nível*, *Antiplástico* e *Manufatura*. A variável *Antiplástico*, por sua vez, torna-se pai da variável *Sítio* que é pai da variável *Núcleo*. Além disso, observa-se que a variável *Manufatura* é pai da variável *Categoria estrutural*.

Para geração do algoritmo, foram utilizados 66% dos fragmentos e 33% para validação. Dessa forma, dos 766 fragmentos disponíveis, 259 foram separados aleatoriamente para validar o algoritmo, sendo 177 da tradição Guarani, 53 da tradição Jacadigo e os 29 restantes da tradição Kadiwéu.

A seguir são apresentadas as tabelas de distribuições de probabilidade observadas para o conjunto de fragmentos utilizados na Rede Bayesiana.

**Tabela 39 – Distribuição de probabilidade na Rede Bayesiana da variável Espessura.**

TRADICAO	'(-inf-0.295455]'	'(0.295455-inf)'
1	0,393	0,607
2	0,9	0,1
3	0,909	0,091

**Tabela 40 – Distribuição de probabilidade na Rede Bayesiana da variável Nível**

TRADICAO	ESPESSURA	1	2	3	10	20	30	40	50	60	70
1	'(-inf-0.295455]	0,792	0,017	0,007	0,091	0,061	0,012	0,012	0,002	0,002	0,002
1	'(0.295455-inf)	0,954	0,005	0,002	0,011	0,018	0,002	0,005	0,002	0,002	0,002
2	'(-inf-0.295455]	0,722	0,003	0,003	0,049	0,036	0,095	0,056	0,016	0,01	0,01
2	'(0.295455-inf)	0,69	0,024	0,024	0,024	0,119	0,024	0,024	0,024	0,024	0,024
3	'(-inf-0.295455]	0,758	0,006	0,006	0,039	0,051	0,039	0,051	0,028	0,017	0,006
3	'(0.295455-inf)	0,654	0,038	0,038	0,038	0,038	0,038	0,038	0,038	0,038	0,038

**Tabela 41 – Distribuição de probabilidade na Rede Bayesiana da variável Antiplástico.**

TRADICAO	ESPESSURA	1	2	3	4
1	'(-inf-0.295455]	0,211	0,779	0,007	0,002
1	'(0.295455-inf)	0,079	0,908	0,002	0,011
2	'(-inf-0.295455]	0,637	0,337	0,01	0,017
2	'(0.295455-inf)	0,694	0,25	0,028	0,028
3	'(-inf-0.295455]	0,355	0,622	0,017	0,006
3	'(0.295455-inf)	0,35	0,55	0,05	0,05

**Tabela 42 – Distribuição de probabilidade na Rede Bayesiana da variável Manufatura.**

TRADICAO	ESPESSURA	1	2	3	99
1	'(-inf-0.295455]'	0,719	0,002	0,002	0,276
1	'(0.295455-inf)'	0,924	0,002	0,002	0,073
2	'(-inf-0.295455]'	0,43	0,03	0,003	0,537
2	'(0.295455-inf)'	0,361	0,194	0,028	0,417
3	'(-inf-0.295455]'	0,18	0,006	0,041	0,773
3	'(0.295455-inf)'	0,25	0,05	0,05	0,65

**Tabela 43 – Distribuição de probabilidade na Rede Bayesiana da variável Sítio.**

TRADICAO	ANTIPLASTICO	1	2	3	4	6
1	1	0,343	0,007	0,007	0,606	0,036
1	2	0,918	0,001	0,003	0,065	0,013
1	3	0,429	0,143	0,143	0,143	0,143
1	4	0,636	0,091	0,091	0,091	0,091
2	1	0,808	0,005	0,005	0,005	0,178
2	2	0,699	0,009	0,009	0,009	0,274
2	3	0,429	0,143	0,143	0,143	0,143
2	4	0,556	0,111	0,111	0,111	0,111
3	1	0,014	0,07	0,014	0,887	0,014
3	2	0,008	0,174	0,058	0,752	0,008
3	3	0,143	0,143	0,143	0,429	0,143
3	4	0,2	0,2	0,2	0,2	0,2

**Tabela 44 – Distribuição de probabilidade na Rede Bayesiana da variável Categoria estrutural.**

TRADICAO	MANUFATURA	1	2	3	5
1	1	0,892	0,103	0,003	0,001
1	2	0,25	0,25	0,25	0,25
1	3	0,25	0,25	0,25	0,25
1	99	0,816	0,158	0,019	0,006
2	1	0,326	0,66	0,007	0,007
2	2	0,056	0,056	0,056	0,833
2	3	0,25	0,25	0,25	0,25
2	99	0,725	0,253	0,006	0,017
3	1	0,711	0,237	0,026	0,026
3	2	0,25	0,25	0,25	0,25
3	3	0,3	0,5	0,1	0,1
3	99	0,709	0,25	0,02	0,02

**Tabela 45 – Distribuição de probabilidade na Rede Bayesiana da variável Núcleo.**

TRADICAO SITIO		1	2	3	4	5	7
1	1	0,041	0,427	0,311	0,015	0,205	0,001
1	2	0,167	0,167	0,167	0,167	0,167	0,167
1	3	0,375	0,125	0,125	0,125	0,125	0,125
1	4	0,049	0,382	0,16	0,104	0,299	0,007
1	6	0,15	0,25	0,25	0,05	0,25	0,05
2	1	0,086	0,267	0,508	0,011	0,124	0,004
2	2	0,167	0,167	0,167	0,167	0,167	0,167
2	3	0,167	0,167	0,167	0,167	0,167	0,167
2	4	0,167	0,167	0,167	0,167	0,167	0,167
2	6	0,122	0,5	0,176	0,095	0,095	0,014
3	1	0,167	0,167	0,167	0,167	0,167	0,167
3	2	0,1	0,233	0,567	0,033	0,033	0,033
3	3	0,083	0,083	0,583	0,083	0,083	0,083
3	4	0,006	0,369	0,481	0,019	0,106	0,019
3	6	0,167	0,167	0,167	0,167	0,167	0,167

Sumarização dos resultados da Rede Bayesiana

Casos classificados corretamente	233	89.9614 %
Casos classificados incorretamente	26	10.0386 %
Estatística Kappa	0,7931	
Número de casos	259	

**Tabela 46 – Tabela com a precisão do algoritmo de classificação.**

Taxa verdadeiros-positivos	Taxa falsos-positivos	Precisão	Curva ROC	Tradição tecnológica
0,92	0,13	0,94	0,96	Guarani
0,85	0,07	0,76	0,97	Jacadigo
0,90	0,00	0,96	1,00	Kadiwéu

## Matriz de Confusão

		Classificada		
		Original	Guarani	Jacadigo
	Guarani	162	14	1
		8	45	0
	Kadiwéu	3	0	26

Analisando os resultados acima, vemos que o desempenho do algoritmo de classificação é satisfatório, já que a precisão de acerto do algoritmo é de cerca de 89% (Tabela 46 e matriz de confusão).

Dessa forma, dado o resultado satisfatório do algoritmo para classificação dos fragmentos, deseja-se, portanto, extrapolar o resultado obtido para o grupo dos fragmentos ainda sem classificação com relação à Tradição tecnológica.

Analizando os fragmentos disponíveis e calculando, para um deles, as probabilidades de pertencer a cada um dos grupos, obteve-se a seguinte distribuição quanto a classificação por grupo:

**Tabela 47 – Distribuição dos fragmentos classificados na Rede Bayesiana segundo a Tradição tecnológica**

	Guarani	Jacadigo	Kadiwéu	Total
Freq.	421	1141	733	2295
%	18,3%	49,7%	31,9%	100%

Observamos que cerca da metade dos fragmentos foram classificados na tradição Jacadigo e 30%, aproximadamente, foram classificados na tradição Kadiwéu; já os fragmentos classificados na tradição Guarani representam cerca de 18% do total.

É válido ressaltar que dos 3.190 fragmentos disponíveis inicialmente, 766 já haviam sido classificados pelo pesquisador. Dos 2.424 restantes, foi possível classificar 2.295 (Tabela 47) e para os 129 que ainda faltaram não foi possível essa classificação devido ao fato desses fragmentos apresentarem características distintas daquelas dos fragmentos já classificados – características estas, portanto, não consideradas na Rede Bayesiana.

## 6. Conclusões

O estudo em questão permitiu avaliar quais características dos fragmentos são importantes para distingui-los com relação à Tradição tecnológica. Dessa forma, observou-se que as variáveis Antiplástico, Sítio, Acabamento da Face Externa, Espessura e Coloração são importantes para essa distinção. Além disso, utilizando-se do conhecimento do pesquisador e da distribuição probabilística das variáveis encontrada nos fragmentos foi possível criar um algoritmo de classificação com o intuito de orientá-lo em futuras classificações para novos fragmentos.

Destaca-se também o fato de que a maioria dos fragmentos já classificados pertenciam à tradição Guarani e os fragmentos classificados pela Rede Bayesiana foram, em sua maioria, classificados nas duas outras tradições: Jacadigo e Kadiwéu; assim, podemos inferir que os fragmentos oriundos da tradição Guarani possuem características físicas mais marcantes de serem analisadas individualmente e os das tradições Jacadigo e Kadiwéu necessitam de serem analisados levando em conta o conjunto de variáveis, sendo, portanto, mais fácil de serem classificados com o auxílio da Rede Bayesiana.

*Apêndice A*  
*Lista das variáveis*

*Apêndice B*  
*Conteúdo Teórico*

## Coeficiente de similaridade de Gower

O Coeficiente de Similaridade de Gower foi proposto por Gower em 1971 e é uma das medidas mais populares de similaridade para a análise simultânea de dados quantitativos e qualitativos.

O coeficiente de Similaridade de Gower  $s_{ij}$  compara dois casos  $i$  e  $j$  e é definido como:

$$s_{ij} = \frac{\sum_k w_{ijk} s_{ijk}}{\sum_k w_{ijk}}$$

em que:

$s_{ijk}$  denota a contribuição da  $k$ -ésima variável, assumindo valor 1 se os indivíduos  $i$  e  $j$  são semelhantes e 0, caso contrário.

$w_{ijk}$  é usualmente 1 ou 0 dependendo se ou não a comparação entre os indivíduos  $i$  e  $j$  é válida para a  $k$ -ésima variável (recebe 0 quando o valor da variável está ausente em um ou ambos indivíduos).

Nota-se que o efeito do denominador  $\sum_k w_{ijk}$  é dividir a soma de similaridades pelo número de variáveis.

## Clusters hierárquicos

Dado um conjunto de  $n$  indivíduos, o ponto de partida para os métodos de classificação hierárquicos será, em geral, uma matriz  $n \times n$  cujo elemento genérico  $(i, j)$  é uma medida de semelhança (ou dissimilaridade) entre o indivíduo  $i$  e o indivíduo  $j$ . Os critérios de semelhança (ou dissimilaridade) utilizados podem ser diversos, havendo alguns critérios específicos para dados de diversos tipos. Com grande frequência, existe uma matriz  $X \times p$  de observações multivariadas associadas aos indivíduos e que são usadas para o cálculo da matriz de semelhanças/dissimilaridade.

Podem considerar-se inicialmente os  $n$  indivíduos como constituindo  $n$  classes diferentes. Procedendo por etapas, vai-se fundindo um par de classes em cada etapa. O agrupamento a ser feito numa dada etapa é a junção dos dois subgrupos (classes) considerados mais “semelhantes”. A forma usual de representar graficamente as sucessivas uniões de subgrupos num método de classificação hierárquico é através de um dendrograma.

Um corte no dendrograma a qualquer nível de aglomeração produz uma classificação em k subgrupos ( $1 \leq k \leq n$ ). Note-se que um par de indivíduos que seja incluído numa mesma classe em qualquer etapa do processo não poderá mais ser separado em etapas posteriores, uma vez que estas consistem em junções de classes já existentes.

### Redes Bayesianas

As redes bayesianas foram desenvolvidas no início dos anos 1980 para facilitar a tarefa de predição em sistemas de inteligência artificial (AI) (Pearl, 2000).

Em resumo, redes bayesianas (RB) também conhecidas como redes de opinião, redes causais, gráficos de dependência probabilística, são modelos gráficos para raciocínio (conclusões) baseado na incerteza, onde os nós representam as variáveis (discreta ou contínua), e os arcos representam a conexão direta entre eles. Ela vem se tornando a metodologia padrão para a construção dos sistemas que se baseiam no conhecimento probabilístico e tem sido aplicada em uma variedade de atividades do mundo real.

Do ponto de vista aplicado, redes bayesianas constituem um modelo gráfico que representa de forma simples as relações de dependência das variáveis de um sistema.

As dependências em uma rede bayesiana são representadas na forma de um grafo direcionado acíclico e tem como uma das suas principais características a adaptabilidade, podendo, a partir de novas informações, gerar alterações na estrutura de dependência. Além disso, diferentemente dos métodos clássicos de inferência, dado um conjunto de informações, permite calcular a probabilidade de uma peça pertencer a cada um dos grupos formados.