



Análise de Agrupamento

Indicadores GeoSES & Indicadores Obstétricos

Wesley Pereira da Silva ¹

24 de janeiro de 2023

¹[Lattes](#)

Conteúdo

1	Conjunto de Dados	3
1.1	Listagem das Variáveis	3
1.2	Estrutura dos Dados	4
2	Análise de Agrupamento	16
	Siglas	18

Apresentação

1. *Contexto:* O presente relatório objetiva apresentar informações detalhadas sobre como as características socioeconômicas estão relacionadas a riscos gestacionais. Investigamos um conjunto de variáveis da base GeoSES¹ e buscamos traçar um paralelo junto a indicadores gerados a partir da base [Sistema de Informação sobre Nascidos Vivos \(SINASC\)](#)². Por meio de análise de agrupamento, identificamos grupos de Municípios do Estado de São Paulo com condições similares de desenvolvimento e variáveis de associação estatisticamente significativas. O conjunto de dados do experimento relatado neste relatório, foi reunido e organizado pelo [Observatório Obstétrico Brasileiro \(OObR\)](#)³.
2. *Métodos:* Esse foi um estudo ecológico baseado em população com análise de agrupamento de indicadores obstétricos associados a indicadores demográficos. Para verificar a correlação de características socioeconômicas e associado a risco gestacional, buscamos agrupar as variáveis pela técnica k-Means. Na sequência, listamos em árvore de decisão a ordem de variáveis que leva as instâncias aos seus respectivos grupos. Os dados de cada município de São Paulo foram tratados como vetores matemáticos e submetidos ao algoritmo citado. As diferentes magnitudes de medidas do conjunto de dados foram padronizadas para valores de uma distribuição normal, assim, $X \rightarrow X_N$, sendo $X_N \sim N(\mu = 0, \sigma = 1)$. Evidenciamos a associação de variáveis por meio do cálculo de correlação de Spearman e filtrando os valores de correlação superiores 0,4 e inferiores a -0,4.

¹<https://dados.gov.br/dataset/geoses>

²sinasc.saude.gov.br

³linktr.ee/observatorioobstetricobr

1

Conjunto de Dados

A fim de buscar uma abordagem compatível com o estudo de [1], que aponta grupos socioeconômicos nos quais há há riscos gestacionais associados a contaminação por COVID-19 no Brasil, selecionamos um conjunto de variáveis socioeconômicas e obstétricas. A seguir, seus nomes, descrições e estruturas.

1.1 Listagem das Variáveis

Indicadores Obstétricos

- Nascimentos - Quantidade de Nascimentos
- Porc.prematuros - Percentual de Prematuros
- Porc.cesareas - Percentual de Cesáreas
- Porc.grav_multipla - Percentual de Gravidez Múltiplas
- Porc.anomalias - Percentual de Anomalias
- Porc.nenhuma_consulta - Percentual de nenhuma consulta
- Porc.consulta7mais - Percentual de mais de 7 consultas
- Porc.feminino - Percentual do sexo feminino
- Porc.raca_mae_branca - Percentual de mãe com raça branca
- Porc.raca_mae_negra - Percentual de mãe com raça negra

- Porc_peso_menor_2500 - Percentual de peso menor de 2.5Kg
- Porc_apgar1_menor_7 - Percentual de APGAR1 menor 7
- Porc_apgar5_menor_7 - Percentual de APGAR5 menor 7

Indicadores GeoSES

- GeoSES
- renda média domiciliar
- educação
- mobilidade
- pobreza
- riqueza
- privação material
- segregação residencial (distanciamento residencial entre grupos econômico-raciais diferentes)

1.2 Estrutura dos Dados

Indicadores	n	μ	σ	min	25%	50%	75%	max
GeoSES	643	-0,23	0,26	-1,00	-0,39	-0,24	-0,09	1,00
educação	643	58,27	6,20	34,98	54,16	58,89	62,23	75,17
privacao_material	643	68,98	14,24	18,38	62,71	73,42	79,13	91,96
pobreza	643	21,97	9,19	5,57	15,81	20,17	26,17	62,65
riqueza	643	1,35	1,29	0,00	0,53	1,09	1,78	17,58
renda	643	2381,92	637,68	1132,09	1961,32	2283,12	2663,29	6707,76
segregação	643	-0,52	0,09	-0,74	-0,58	-0,54	-0,48	-0,11

Tabela 1.1: GeoSES - Descritivo das Variáveis para os Municípios de SP

Indicadores %	n	μ	σ	min	25%	50%	75%	max
Prematuros	643	0,08	0,04	0,00	0,06	0,08	0,10	0,25
Cesareas	643	0,65	0,15	0,15	0,54	0,66	0,77	0,97
Grav Múltipla	643	0,02	0,02	0,00	0,00	0,02	0,03	0,21

Anomalias	642	0,01	0,01	0,00	0,00	0,01	0,01	0,17
Zero Consulta	643	0,01	0,01	0,00	0,00	0,01	0,01	0,08
Consulta>=7	643	0,80	0,11	0,31	0,73	0,81	0,88	1,00
Feminino	643	0,49	0,05	0,25	0,46	0,49	0,51	0,86
Raça_mãe_branca	9	0,61	0,49	0,00	0,00	1,00	1,00	1,00
Raça_mãe_negra	9	0,39	0,49	0,00	0,00	0,00	1,00	1,00
Peso<2500	628	0,09	0,03	0,01	0,07	0,09	0,10	0,26
Apgar1<7	643	0,06	0,03	0,00	0,04	0,06	0,08	0,25
Apgar5<7	643	0,01	0,01	0,00	0,00	0,01	0,01	0,09

Tabela 1.2: Indicadores Obstétricos - Descritivo das Variáveis para os Municípios de SP

Análise de Agrupamento

Em função das diversas amplitudes das variáveis tratadas, foi realizado o procedimento de padronização, escalando os valores proporcionalmente para entre 0 e 1. A clusterização dos indicadores GeoSES para os 643 Municípios de São Paulo e os quantitativos ficaram da seguinte forma:

Grupo	Quantidade
Inferior	112
Intermediário	360
Superior	171

Grupo Inferior

Neste grupo, destaca-se os menores valores para o indicador GeoSES, sendo a média $\mu = -0,56$. A Tabela 1.5 lista a quantidade de elementos no Grupo, na coluna n. A seguir, tem-se o valor médio, o desvio padrão, o valor mínimo, o valor dos quantis 25, 50 e 75% e o valor máximo.

Indicadores %	n	μ	σ	min	25%	50%	75%	max
Prematuros	112	0,07	0,04	0,00	0,05	0,07	0,09	0,17
Cesareas	112	0,50	0,15	0,15	0,41	0,49	0,59	0,92
Grav Múltipla	112	0,02	0,02	0,00	0,00	0,01	0,02	0,10
Anomalias	112	0,01	0,01	0,00	0,00	0,00	0,01	0,05
Zero-Consulta	112	0,01	0,01	0,00	0,00	0,01	0,01	0,04
Consulta>=7	112	0,73	0,12	0,31	0,66	0,74	0,83	0,93
Feminino	112	0,49	0,06	0,36	0,46	0,49	0,52	0,70
Raca_mae_branca	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN

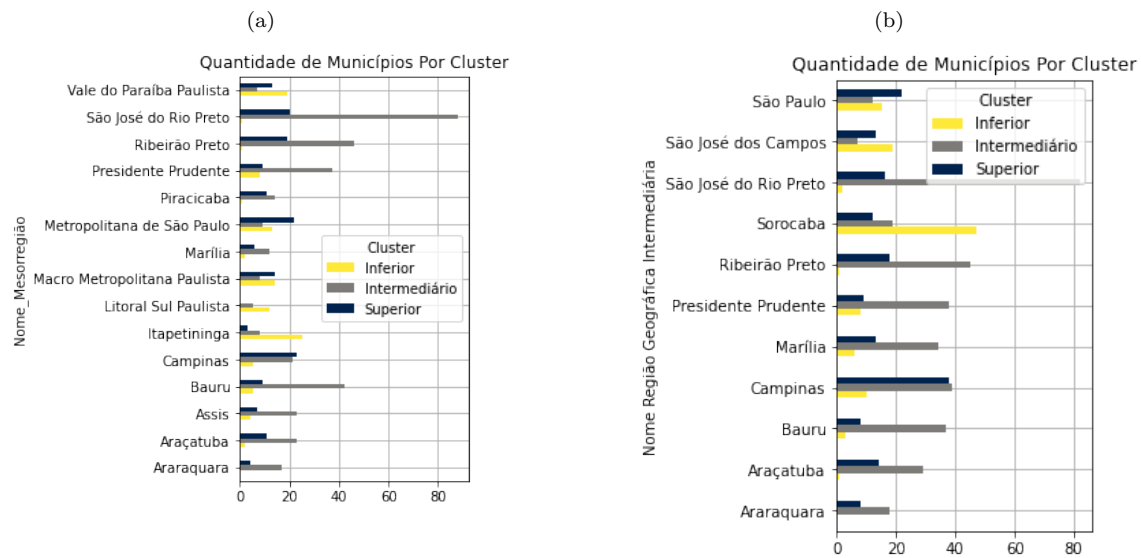
Raca_mae_negra	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Peso>2500	109	0,08	0,04	0,02	0,06	0,08	0,10	0,23
Apgar1<7	112	0,07	0,04	0,00	0,04	0,07	0,09	0,23
Apgar5<7	112	0,01	0,01	0,00	0,00	0,01	0,02	0,09
GeoSES	112	-0,56	0,17	-1,00	-0,66	-0,55	-0,45	-0,11
educação	112	64,04	4,72	53,57	61,29	64,34	67,03	75,17
privacao_material	112	45,02	11,73	18,38	37,44	46,25	54,10	68,20
pobreza	112	35,62	10,05	18,94	27,93	34,37	40,89	62,65
riqueza	112	0,83	0,76	0,00	0,35	0,57	1,16	4,63
renda	112	1824,43	356,51	1132,09	1584,87	1781,23	2057,18	2941,48
segregacao	112	-0,60	0,05	-0,72	-0,64	-0,60	-0,56	-0,48

Cluster Intermediário

Indicadores %	n	μ	σ	min	25%	50%	75%	max
Prematuros	360	0,08	0,04	0,00	0,06	0,08	0,10	0,25
Cesareas	360	0,68	0,14	0,25	0,57	0,70	0,79	0,97
Grav Múltipla	360	0,02	0,03	0,00	0,00	0,02	0,03	0,21
Anomalias	359	0,01	0,02	0,00	0,00	0,01	0,01	0,17
Zero_Consulta	360	0,01	0,01	0,00	0,00	0,00	0,01	0,08
Consulta>=7	360	0,81	0,11	0,40	0,75	0,82	0,89	1,00
Feminino	360	0,49	0,06	0,25	0,45	0,49	0,52	0,86
Raça_mãe_branca	4	0,50	0,58	0,00	0,00	0,50	1,00	1,00
Raça_mãe_negra	4	0,50	0,58	0,00	0,00	0,50	1,00	1,00
Peso>2500	348	0,09	0,04	0,01	0,07	0,09	0,11	0,26
Apgar1<7	360	0,06	0,04	0,00	0,04	0,06	0,08	0,25
Apgar5<7	360	0,01	0,01	0,00	0,00	0,01	0,02	0,08
GeoSES	360	-0,28	0,11	-0,64	-0,36	-0,28	-0,20	-0,02
educação	360	60,18	3,45	52,04	58,03	60,06	62,18	75,02
privacao_material	360	72,14	7,65	49,40	67,20	73,80	78,18	87,50
pobreza	360	20,67	5,65	6,90	16,60	20,41	24,29	38,29
riqueza	360	0,99	0,73	0,00	0,45	0,90	1,44	4,52
renda	360	2222,52	322,04	1516,46	1988,02	2191,23	2438,41	3431,94
segregacao	360	-0,55	0,04	-0,74	-0,58	-0,55	-0,52	-0,46

Tabela 1.4: Indicadores Obstétricos - Descritivo das Variáveis

Figura 1.1: Quantidade de Municípios nos Clusters Por Meso Região e Região Geográfica Intermediária



Cluster - Superior

Indicadores %	n	μ	σ	min	25%	50%	75%	max
Prematuros	171	0,08	0,03	0,00	0,07	0,09	0,10	0,25
Cesareas	171	0,69	0,11	0,42	0,60	0,69	0,78	0,97
Grav Múltipla	171	0,02	0,01	0,00	0,02	0,02	0,03	0,09
Anomalias	171	0,01	0,01	0,00	0,01	0,01	0,01	0,05
Zero Consulta	171	0,01	0,01	0,00	0,00	0,01	0,01	0,04
Consulta \geq 7	171	0,81	0,09	0,51	0,76	0,82	0,88	0,96
Feminino	171	0,48	0,03	0,36	0,48	0,49	0,50	0,55
Raça_mãe_branca	5	0,70	0,45	0,00	0,50	1,00	1,00	1,00
Raça_mãe_negra	5	0,30	0,45	0,00	0,00	0,00	0,50	1,00
Peso $>$ 2500	171	0,09	0,02	0,04	0,08	0,09	0,10	0,22
Apgar1 $<$ 7	171	0,06	0,02	0,00	0,04	0,06	0,07	0,13
Apgar5 $<$ 7	171	0,01	0,01	0,00	0,01	0,01	0,01	0,03
GeoSES	171	0,09	0,19	-0,18	-0,04	0,05	0,15	1,00
Educação	171	50,46	3,82	34,98	48,38	51,14	53,20	57,48
Privação_material	171	78,02	8,41	40,38	74,72	79,99	83,84	91,96
PPobreza	171	15,77	4,51	5,57	12,44	15,60	18,86	28,76

Riqueza	171	2,45	1,75	0,31	1,44	2,19	2,97	17,58
Renda	171	3082,63	693,67	2047,89	2665,17	2925,19	3316,17	6707,76
Segregação	171	-0,41	0,06	-0,51	-0,46	-0,43	-0,38	-0,11

Tabela 1.5: Indicadores Obstétricos - Descritivo das Variáveis

Figura 1.2: Quantidade de Municípios nos Clusters por Micro-Região e por Região Geográfica Imediata

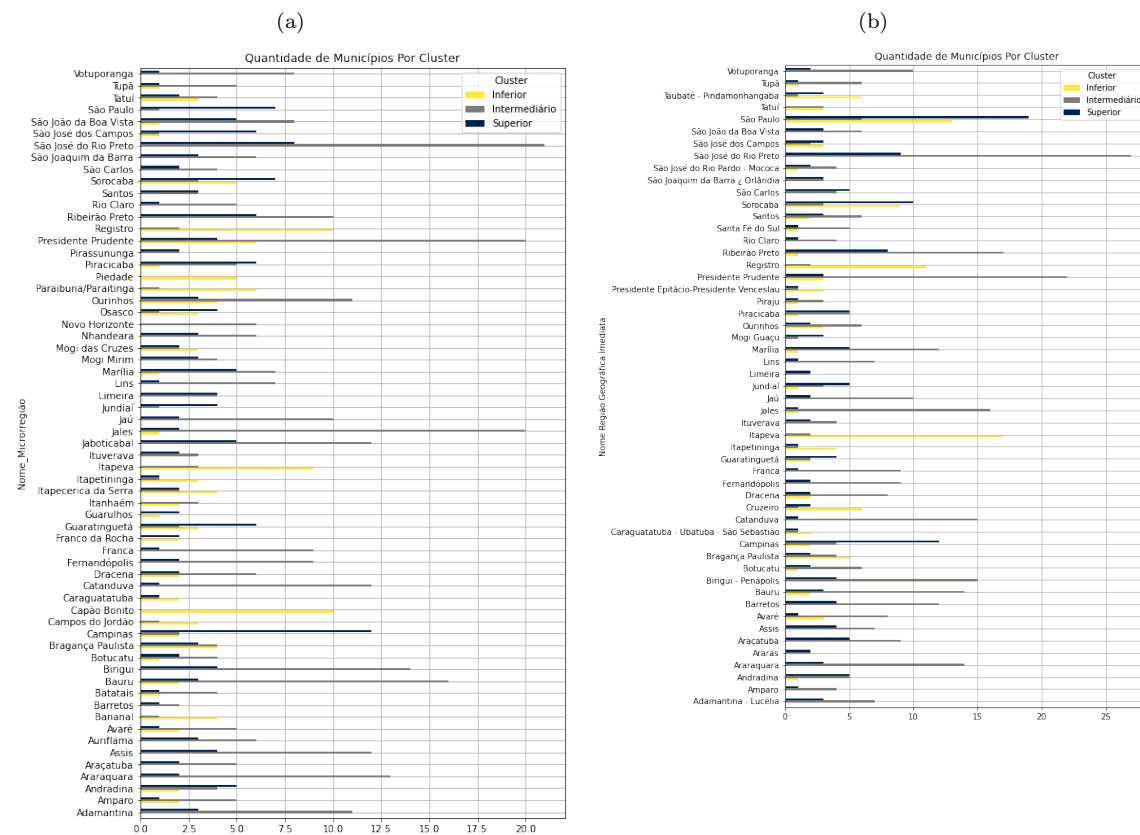


Figura 1.3: Distribuições Estatísticas do Grupo Inferior

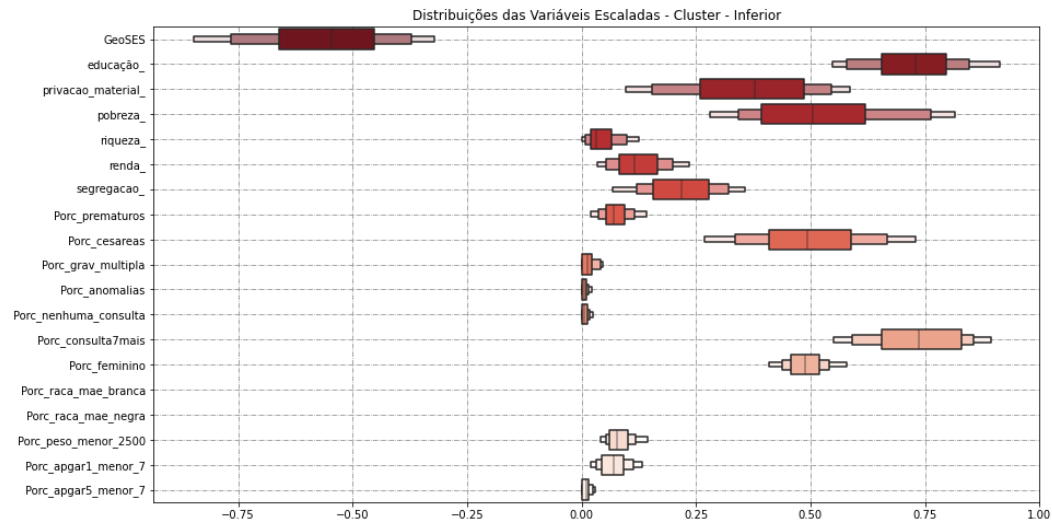


Figura 1.4: Correlação das Variáveis no Grupo Inferior

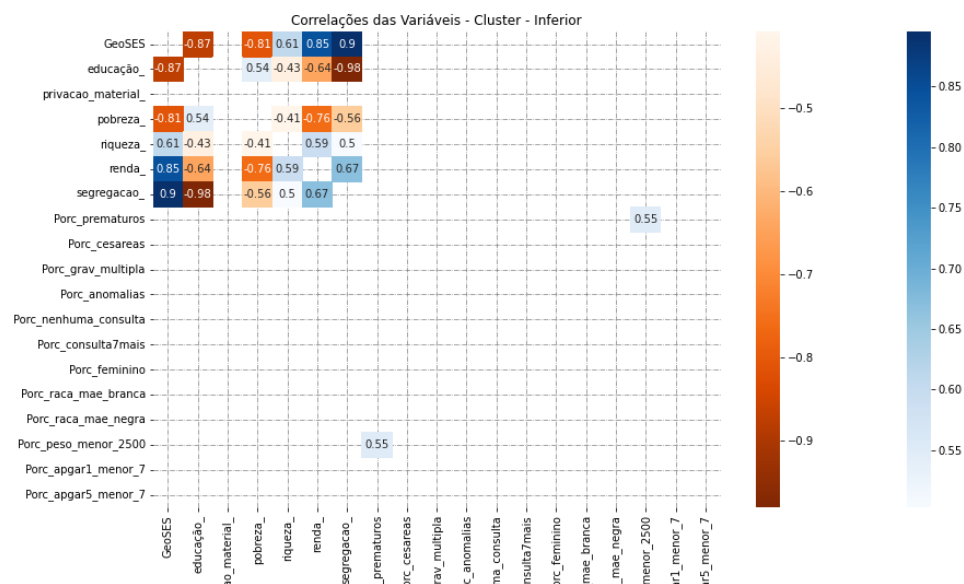


Figura 1.5: Mapa do Estado de Paulo - GeoSES nas Cidades do Grupo Inferior

Indicador GeoSES - Estado de São Paulo



Figura 1.6: Distribuições Estatísticas do Grupo Intermediário

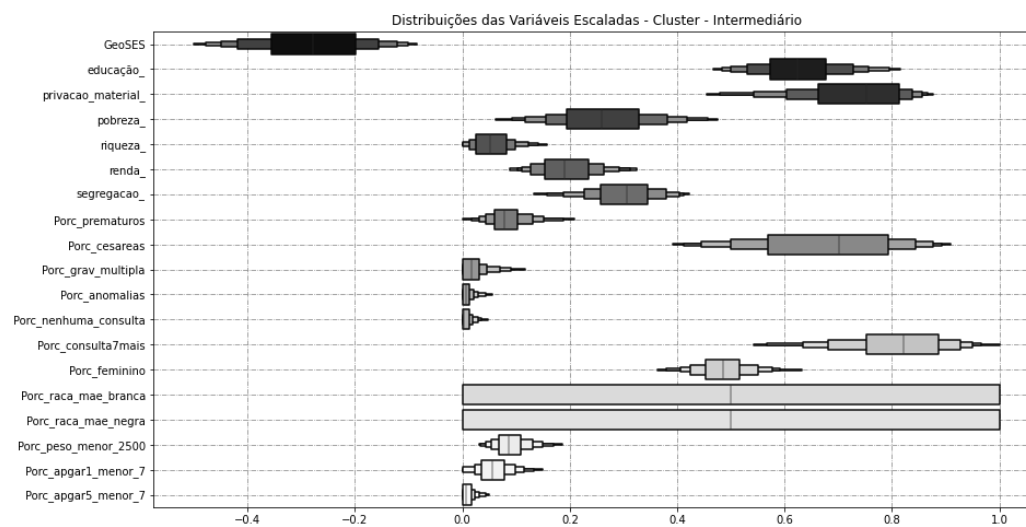


Figura 1.7: Correlação das Variáveis no Grupo Intermediário

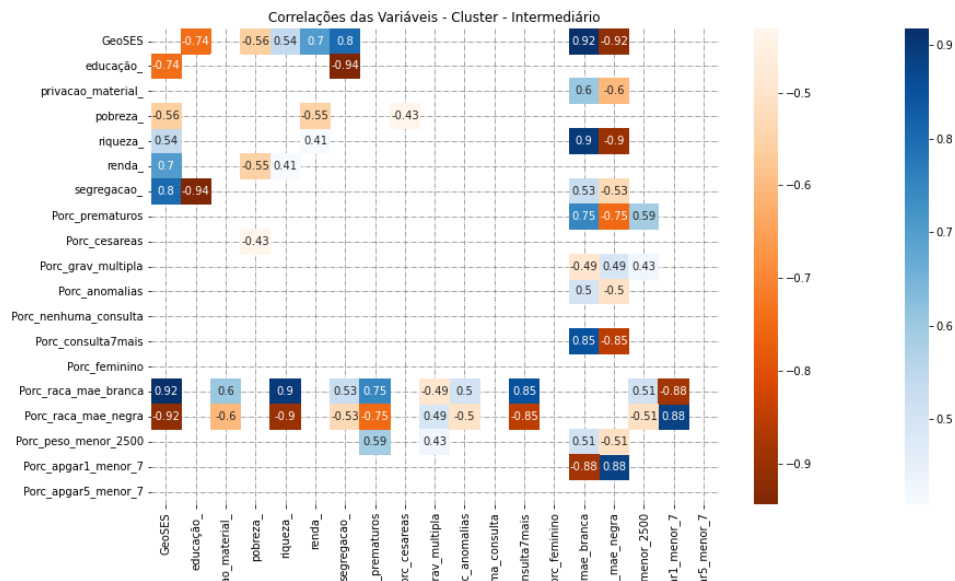


Figura 1.8: Mapa do Estado de Paulo - GeoSES nas Cidades do Grupo Intermediário

Indicador GeoSES - Estado de São Paulo

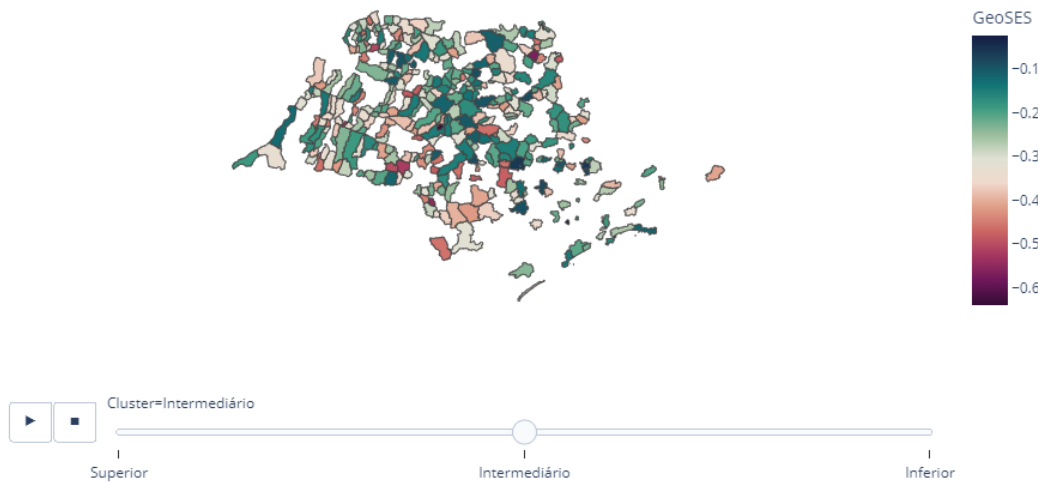


Figura 1.9: Distribuições Estatísticas do Grupo Superior

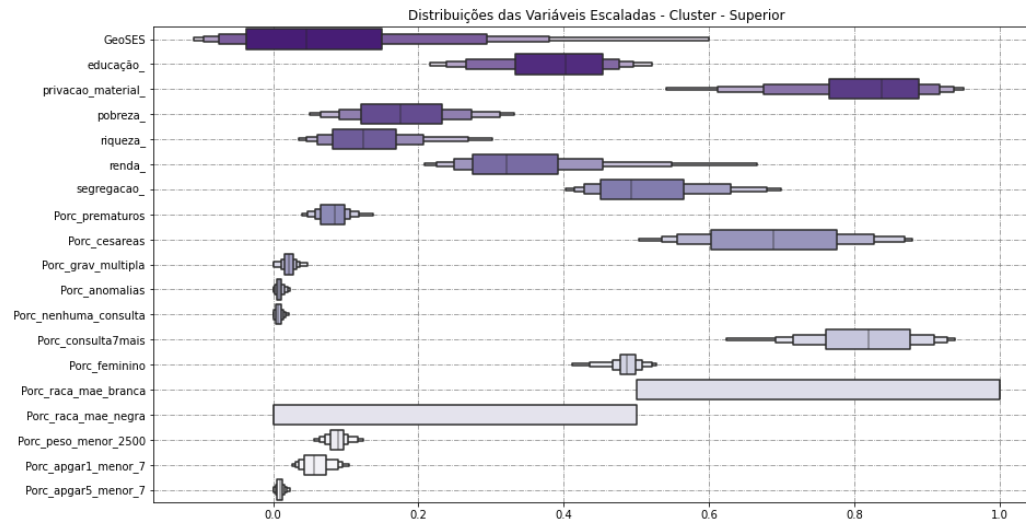
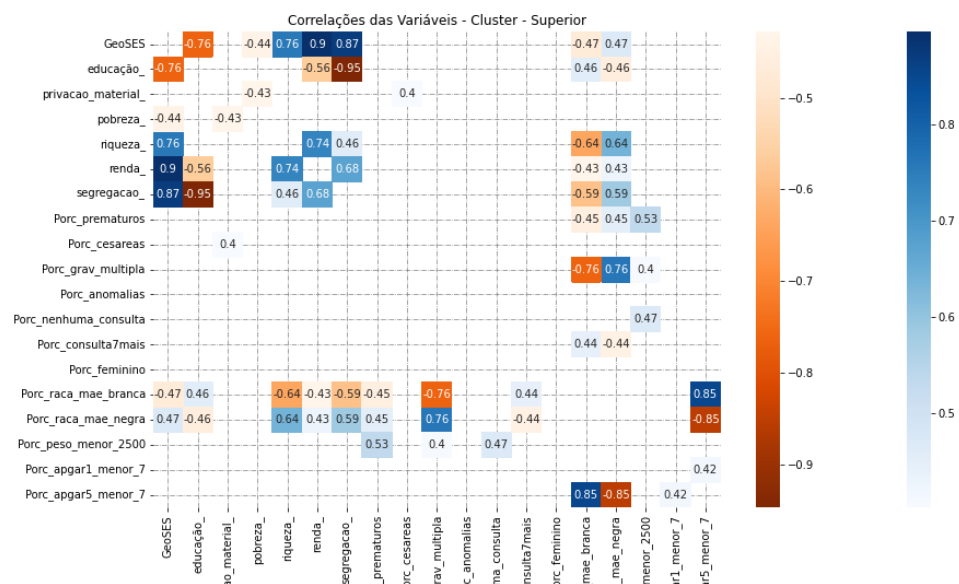


Figura 1.10: Correlação das Variáveis no Grupo Superior



Indicador GeoSES - Estado de São Paulo

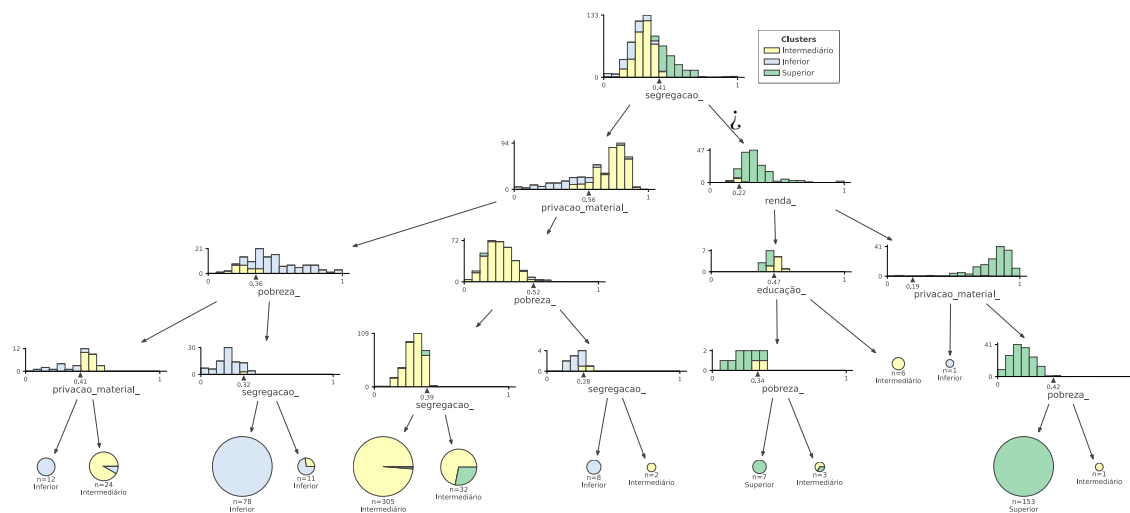


Figura 1.13: Árvore de Decisão do Agrupamento - Indicadores Obstétricos

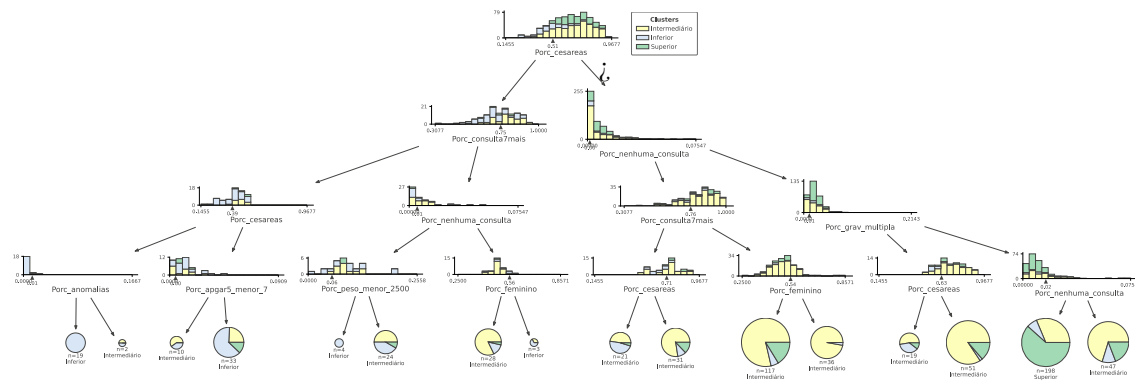
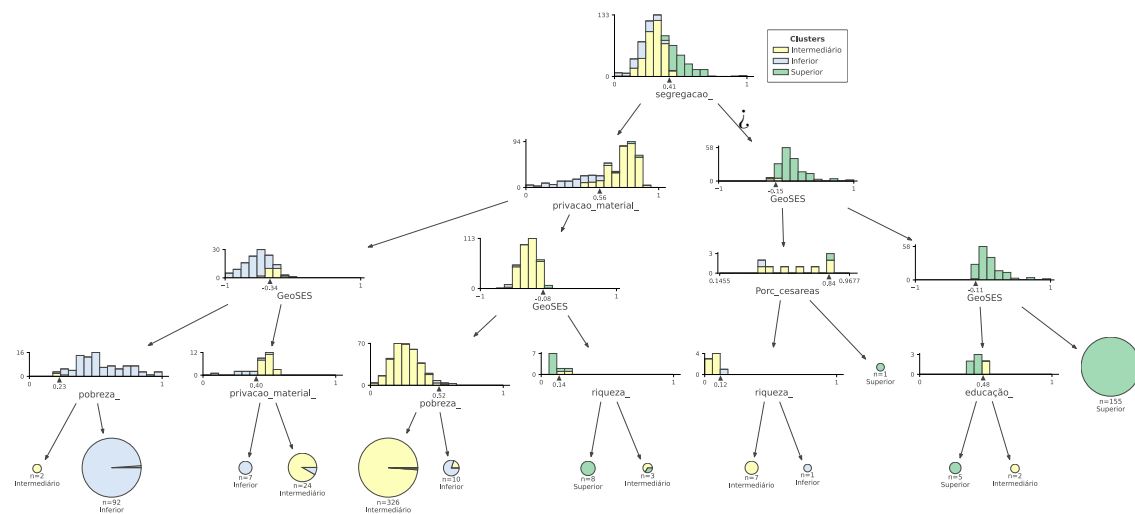


Figura 1.14: Árvore de Decisão do Agrupamento - Indicadores GeoSES e Obstétricos



2

Análise de Agrupamento

Bibliografia

- [1] Thayane Santos Siqueira, José Rodrigo Santos Silva, Mariana do Rosário Souza, Débora Cristina Fontes Leite, Thomas Edwards, Paulo Ricardo Martins-Filho, Ricardo Queiroz Gurgel, and Victor Santana Santos. Spatial clusters, social determinants of health and risk of maternal mortality by COVID-19 in Brazil: a national population-based ecological study. *The Lancet Regional Health - Americas*, 3:100076, 2021.

Siglas

OOBr Observatório Obstétrico Brasileiro [2](#)

SINASC Sistema de Informação sobre Nascidos Vivos [2](#)