

# Plano Analítico para Classificação estatística de deputados federais como bancada evangélica

DOCUMENTO: SAP-2021-012-JG-v01

De: Felipe Figueiredo Para: Josir Gomes

Data: 2021-10-12

## SUMÁRIO

1	LISTA DE ABREVIATURAS.....	2
2	INTRODUÇÃO.....	2
2.1	Contexto.....	2
2.2	Objetivos.....	2
2.3	Hipóteses.....	2
3	DADOS.....	3
3.1	Dados brutos.....	3
3.2	Tabela de dados analíticos.....	3
4	VARIÁVEIS DO ESTUDO.....	3
4.1	Desfechos primário e secundários.....	3
4.2	Covariáveis.....	4
5	MÉTODOS ESTATÍSTICOS.....	4
5.1	Análises estatísticas.....	4
5.1.1	Análise descritiva.....	4
5.1.2	Análise inferencial.....	4
5.1.3	Modelagem estatística.....	4
5.2	Significância e Intervalos de Confiança.....	4
5.3	Tamanho da amostra e Poder.....	5
5.4	Softwares utilizados.....	5
6	EXCEÇÕES E OBSERVAÇÕES.....	5

---

7 REFERÊNCIAS.....	5
8 APÊNDICE.....	5
8.1 Disponibilidade.....	5
8.2 Análise exploratória de dados.....	6

---

## Histórico do documento

Versão	Alterações
01	Versão inicial

## 1 LISTA DE ABREVIATURAS

- AIC: Criterio de informação de Akaike (do inglês, *Akaike's Information Criterion*)
- DP: Desvio padrão

## 2 INTRODUÇÃO

### 2.1 Contexto

Aplicar método de seleção estatística ou redução dimensional de variáveis para selecionar as variáveis que têm maior contribuição na explicação da variância observada nos dados, e ajustar modelo de regressão logística para quantificar os efeitos das variáveis incluídas no modelo final.

### 2.2 Objetivos

Quantificação do efeito da receita recebida durante a campanha, número de votos e outras características de deputados federais eleitos em 2018 na classificação como pertencendo à bancada evangélica.

### 2.3 Hipóteses

Deputados federais da bancada evangélica que foram eleitos em 2018 podem ser identificados com base nas doações recebidas durante a campanha eleitoral, número de votos recebidos e outras características.

## 3 DADOS

### 3.1 Dados brutos

Base de dados recebida contendo características dos deputados federais eleitos em 2018.

### 3.2 Tabela de dados analíticos

Todas as variáveis da tabela de dados analíticos foram identificadas de acordo com as descrições das variáveis, e os valores foram identificados de acordo com o dicionário de dados providenciado. Estas identificações possibilitarão a criação de tabelas de resultados com qualidade de produção final.

Depois dos procedimentos de limpeza e seleção 12 variáveis foram incluídas na análise com 514 observações. A Tabela 1 mostra a estrutura dos dados analíticos.

**Tabela 1** Estrutura da tabela de dados analíticos após seleção e limpeza das variáveis.

id	partido	uf	capilaridade	primeira	sexo	evangelico	num_votos	posicao	decil_filiados	decil_deputados	total_receita
1											
2											
3											
...											
514											

A tabela de dados analíticos serão disponibilizados na versão privada do relatório, e serão omitidas da versão pública do relatório.

## 4 VARIÁVEIS DO ESTUDO

### 4.1 Desfechos primário e secundários

O desfecho primário está definido como a classificação entre deputados da bancada evangélica e outros deputados explicada pela receita total recebida.

### 4.2 Covariáveis

As seguintes características dos deputados federais serão consideradas para inclusão na análise: Número de votos recebidos, posicionamento político, capilaridade, a unidade da federação, o partido (sigla), o sexo e se é o primeiro mandato. As seguintes

características dos partidos serão consideradas para inclusão na análise: decil do número de deputados eleitos e decil do número de filiados.

As receitas discriminadas em suas diversas origens não serão consideradas na análise, devido à baixa representatividade de valores em suas distribuições (figura A1).

## 5 MÉTODOS ESTATÍSTICOS

### 5.1 Análises estatísticas

#### 5.1.1 Análise descritiva

As características dos deputados serão descritas, por estado, como média (DP) ou frequência e proporção (%), conforme apropriado. As distribuições serão sumarizadas em tabelas e visualizadas em gráficos exploratórios

#### 5.1.2 Análise inferencial

As inferências serão baseados no modelo de regressão binomial (cf. seção 5.1.3).

#### 5.1.3 Modelagem estatística

A classificação dos deputados como pertencendo ou não à bancada evangélica foi estabelecida por métodos extrâneos aos dados. Como o desfecho primário é uma variável binária será criado um modelo de regressão logística para quantificar a contribuição das variáveis disponíveis para explicar a associação com a classificação. Não serão incluídos efeitos de interação entre as variáveis.

A técnica de seleção estatística de variáveis “para a frente” (*stepwise forward-selection*) para reduzir o número de variáveis que contribuam para a análise. Será usado o critério de  $p < 0,20$  para inclusão no modelo multivariado. Os modelos intermediários criados serão comparados com o AIC.

### 5.2 Significância e Intervalos de Confiança

Todas as análises serão realizadas ao nível de significância de 5%. Todos os testes de hipóteses e intervalos de confiança calculados serão bicaudais.

### 5.3 Tamanho da amostra e Poder

N/A

### 5.4 Softwares utilizados

Esta análise será realizada utilizando-se o software R versão 4.1.1.

## 6 EXCEÇÕES E OBSERVAÇÕES

## 7 REFERÊNCIAS

- **SAR-2021-012-JG-v01** – Quantificação do efeito do número de votos, verba de doação e outras características na explicação da classificação de bancada evangélica vs outros.

## 8 APÊNDICE

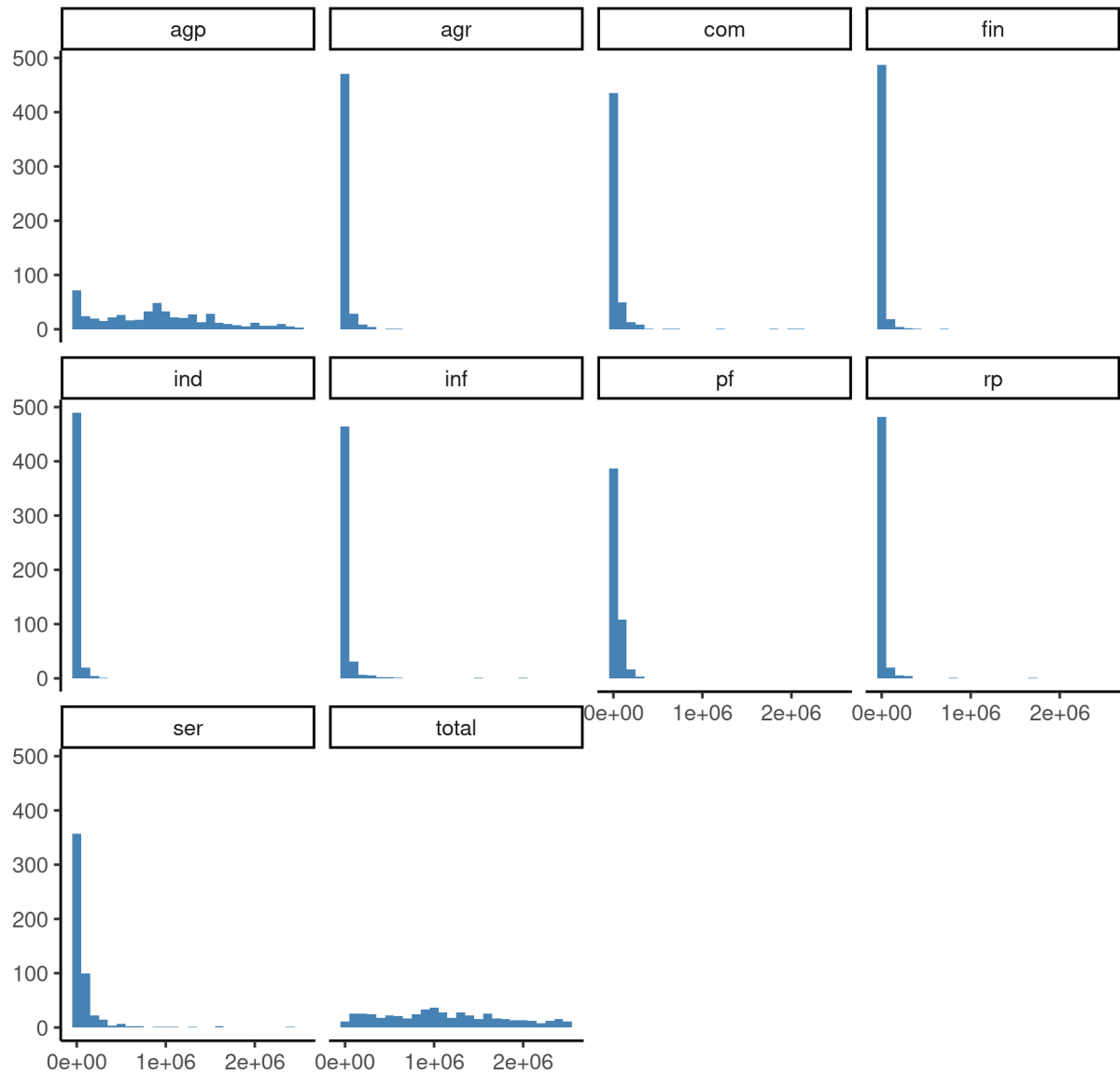
### 8.1 Disponibilidade

Tanto este plano analítico como o relatório correspondente (**SAR-2021-012-JG-v01**) podem ser obtidos no seguinte endereço:

<https://github.com/philsf-biostat/SAR-2021-012-JG/>

## 8.2 Análise exploratória de dados

Distribuição das receitas, por origem



**Figura A1** Distribuição das receitas de deputados federais, por origem (agp = receita que veio do Partido ao invés de apoiadores privados (empresariais ou não); agr = setor agrícola; com = setor do comércio; fin = setor específico dos bancos e outras instituições financeiras e imobiliárias; ind = setores da indústria; inf = setor de infra-estrutura; pf = pessoa física; rp = recursos próprios; ser = setor de serviços).