

MBA em Ciência de Dados

Trabalho Final

Disciplina de
Inteligência de Negócios

Alunos:

Gabriel Costa Neves
Rodrigo Monteiro dos Santos
Thiago Severo Garbin

Professor:

MSc. Eduardo Arruda

Sumário

1	Contextualização	3
2	Questões de Negócio	4
3	Fonte de Dados	5
3.1	Estrutura	5
4	Modelagem Multidimensional	6
5	Implementação	7
5.1	Plataforma Tableau	7
5.1.1	Rotina de Extração, Transformação e Carga	7
5.1.2	Visualizações	7
5.2	Plataforma Microsoft Power BI	7
5.2.1	Rotina de Extração, Transformação e Carga	7
5.2.2	Visualizações	7
6	Conclusão	8

1 Contextualização

Os dados escolhidos se tratam daqueles referentes às pessoas que se alistaram no serviço militar entre os anos de 2015 e 2019 e que estão disponíveis no Portal Brasileiro de Dados Abertos, os quais contêm informações relativas a medidas corporais, localidades de nascimento, alistamento e residência, além de informações sociais como escolaridade, estado civil e religião.

No Brasil, para indivíduos do sexo masculino, o alistamento no serviço militar é obrigatório, devendo este ser realizado no ano em que se adquire a maioridade, ou seja, 18 anos. Desta forma, é possível afirmar que o contexto de negócios neste caso se trata da recepção destes indivíduos por parte das Forças Armadas em todos os anos, a qual, conforme suas disponibilidades e interesses, colherá as devidas informações, realizará testes e decidirá quanto ao aproveitamento e ingresso no serviço. De acordo com o decreto 57.654 de 20/01/1966, os brasileiros chamados para o alistamento obrigatório, são denominados conscritos.

2 Questões de Negócio

Antes de apresentar as perguntas, é importante mencionar que se decidiu por fazer um dashboard para respondê-las, de maneira que foram escolhidos cinco filtros, os quais quando alterados se aplicam a todas as visualizações. Desta forma, os filtros escolhidos foram: ano de alistamento, status em relação a dispensa, religião, estado e região.

Questões:

- 1 - Qual a distribuição do Índice de Massa Corporal (IMC) para os conscritos?
- 2 - Qual o percentual das religiões entre os conscritos?
- 3 - Qual a distribuição do nível de escolaridade entre os conscritos?
- 4 - Qual a média e mediana de altura dos conscritos?
- 5 - Qual a porcentagem de conscritos que se alistaram ou não no ano correto?
- 6 - Quais as nacionalidades, fora a brasileira, que mais aparecem?

3 Fonte de Dados

A fonte é o Portal Brasileiro de Dados Abertos. Os dados são fornecidos pelo Exército Brasileiro, conforme o link a seguir:

<https://dados.gov.br/dataset/servico-militar>

3.1 Estrutura

Os dados estão dispostos em arquivos do tipo csv, separados por cada ano de alistamento. Cada linha representa um indivíduo alistado e contém as seguintes informações separadas por colunas:

ANO_NASCIMENTO, inteiro, ano de nascimento do cidadão;

PESO, inteiro, peso em quilos do cidadão;

ALTURA, inteiro, altura em centímetros do cidadão;

CABECA, inteiro, tamanho da cabeça em centímetros do cidadão;

CALCADO, inteiro, número do calçado do cidadão;

CINTURA, texto, tamanho da cintura em centímetros do cidadão;

RELIGIAO, texto, religião do cidadão;

MUN_NASCIMENTO, texto, município de nascimento do cidadão;

UF_NASCIMENTO, texto, UF de nascimento do cidadão (KK - UF no exterior);

PAIS_NASCIMENTO, texto, país de nascimento do cidadão;

ESTADO_CIVIL, texto, estado civil (Solteiro, Casado, Desquitado, Divorciado, Viúvo, Separado Judicialmente, Outros) do cidadão;

SEXO, texto, sexo (M - Masculino, F - Feminino) do cidadão;

ESCOLARIDADE, texto, escolaridade do cidadão;

VINCULACAO_ANO, inteiro, ano de alistamento do cidadão;

DISPENSA, texto, se o cidadão foi dispensado ou não (Sem dispensa, Com dispensa);

MUN_RESIDENCIA, texto, município de residência do cidadão;

UF_RESIDENCIA, texto, UF de residência do cidadão (KK - UF no exterior);

PAIS_RESIDENCIA, texto, país de residência do cidadão;

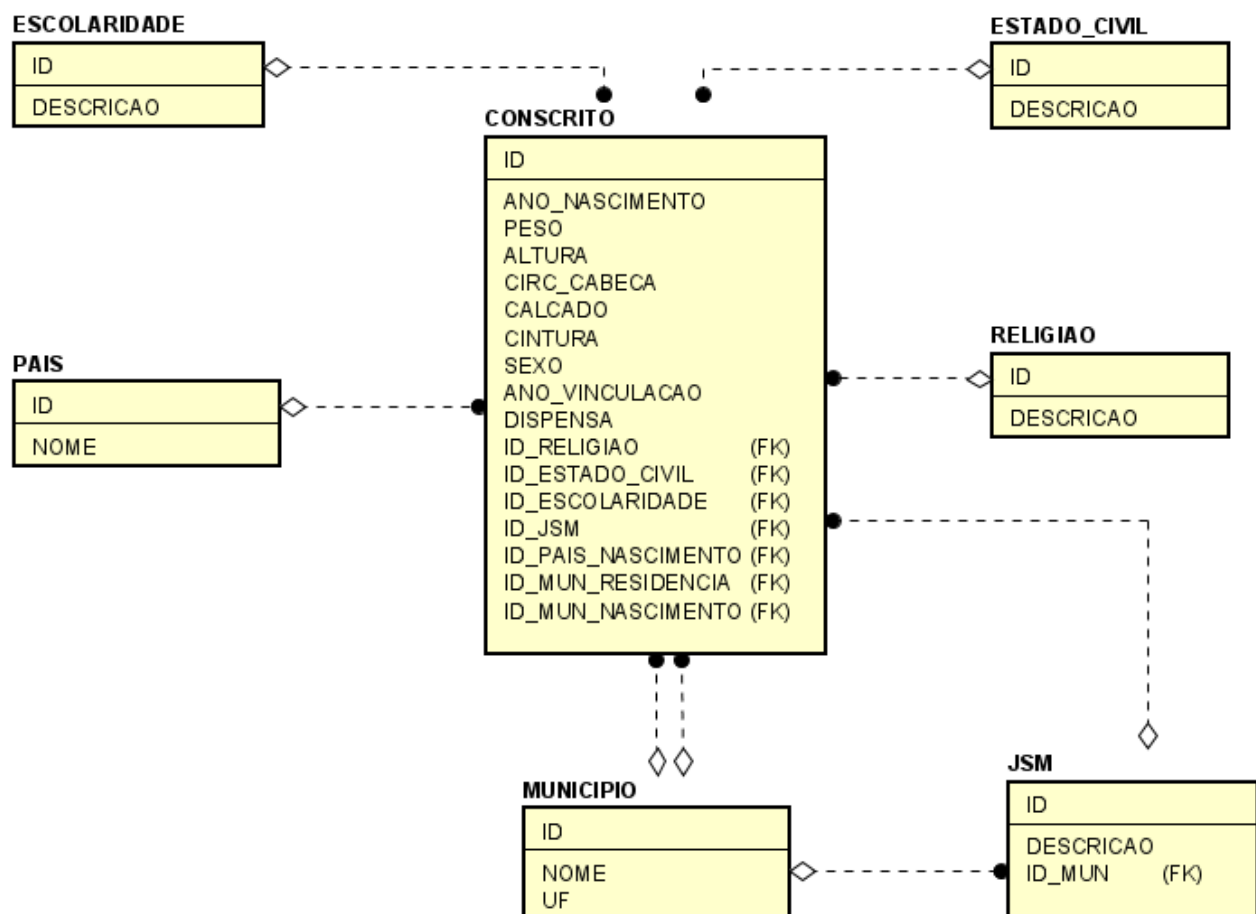
JSM, texto, Junta de Serviço Militar em que o cidadão realizou o alistamento;

MUN_JSM, texto, município da Junta de Serviço Militar;

UF_JSM, texto, UF da Junta de Serviço Militar (KK - UF no exterior).

4 Modelagem Multidimensional

A modelagem do cubo foi realizada no software Astah. Procurou-se normalizar os dados e diminuir a redundância entre as tabelas. As chaves primárias das tabelas são chaves do tipo *surrogate* e não representam as chaves naturais dos dados da aplicação. Tal escolha foi feita para facilitar a implementação dos *joins* com a tabela fato Conscrito.

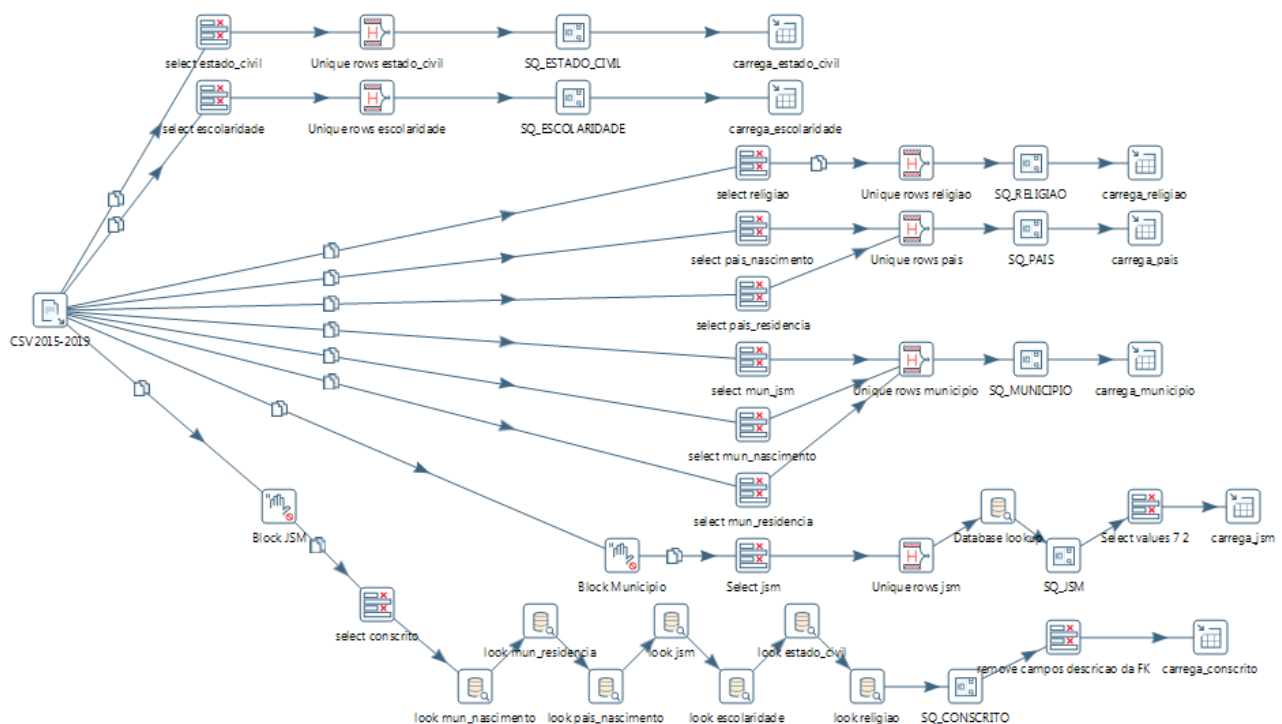


5 Implementação

5.1 Plataforma Tableau

5.1.1 Rotina de Extração, Transformação e Carga

Como forma de pré-processamento, as tabelas CSV originais dos cinco anos analisados foram reunidas em um único arquivo CSV para que então fosse realizada a remoção das linhas que continham valores nulos por meio da linguagem Python. Posteriormente, utilizou-se o arquivo resultante no software Pentaho para a realização das transformações e inserção das tabelas normalizadas no SGBD Oracle.

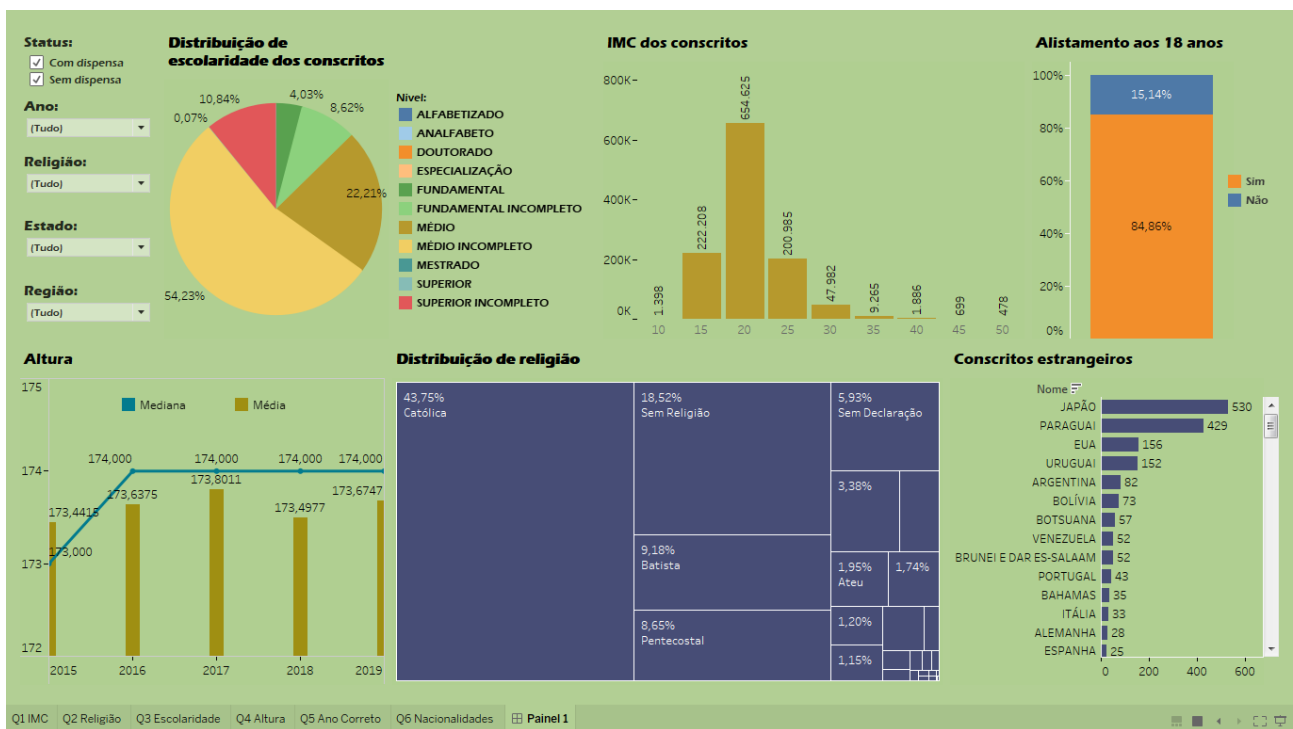


5.1.2 Visualizações

Dashboard Tableau:

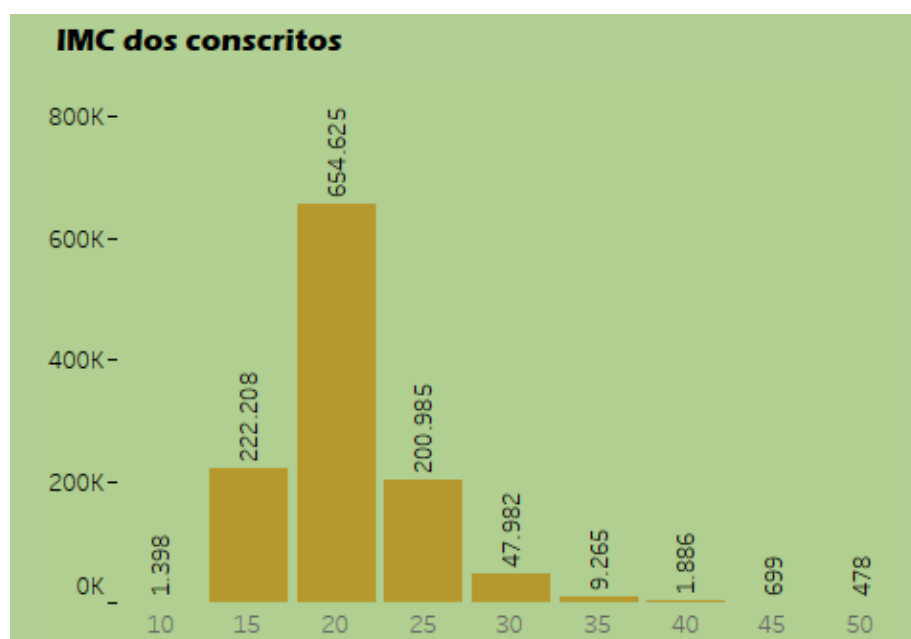
Link para o vídeo do dashboard Tableau em funcionamento:

https://www.youtube.com/watch?v=vx6hZpdMa_I



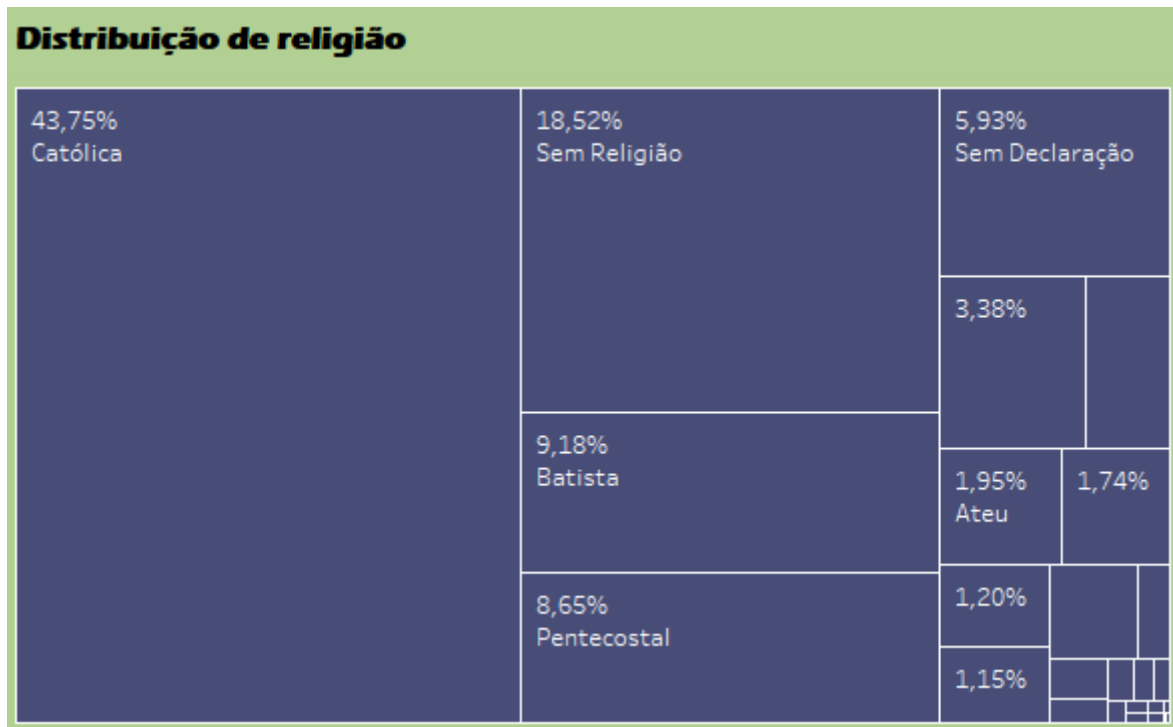
Questão 1 - Tableau:

Qual a distribuição do Índice de Massa Corporal (IMC) para os conscritos?



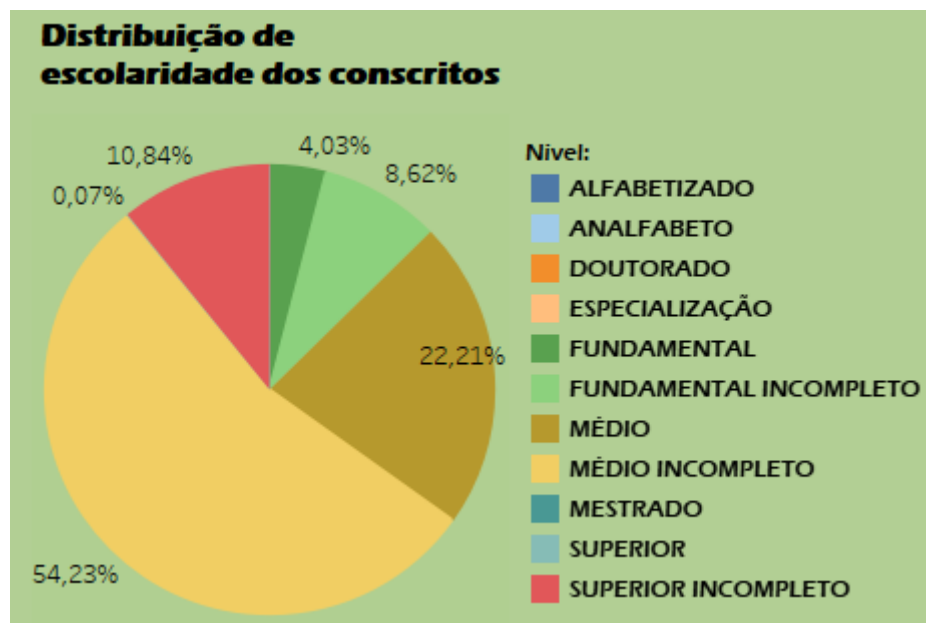
Questão 2 - Tableau:

Qual o percentual das religiões entre os conscritos?



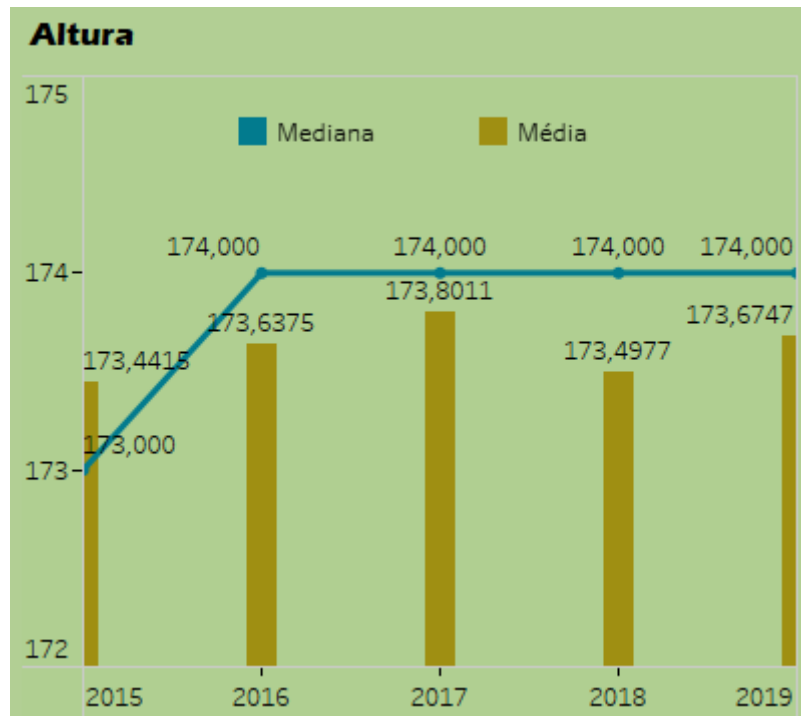
Questão 3 - Tableau:

Qual a distribuição do nível de escolaridade entre os conscritos?



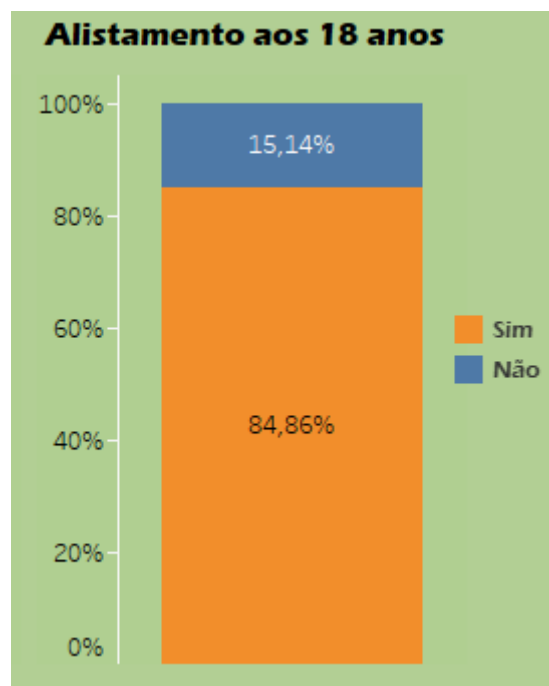
Questão 4 - Tableau:

Qual a média e mediana de altura dos conscritos?



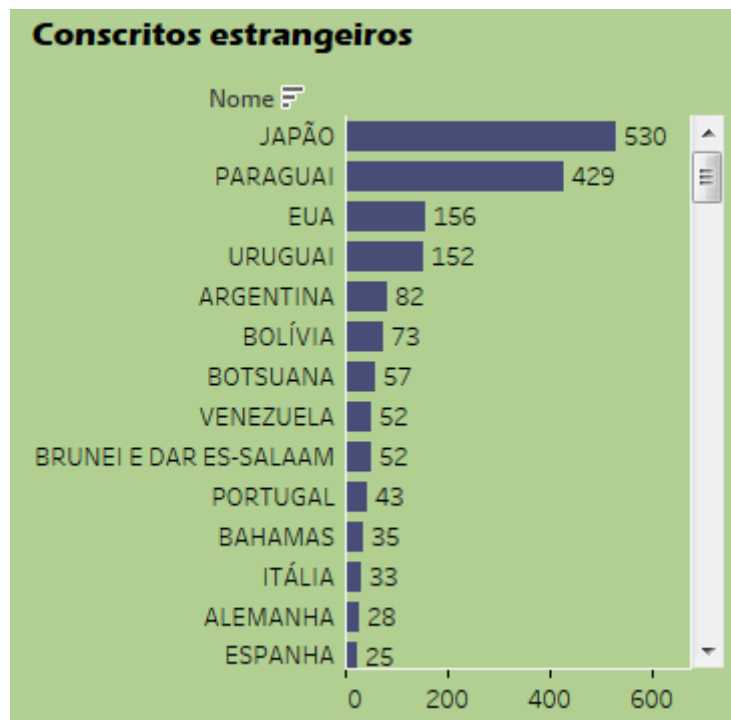
Questão 5 - Tableau:

Qual a porcentagem de conscritos que se alistaram ou não no ano correto?



Questão 6 - Tableau:

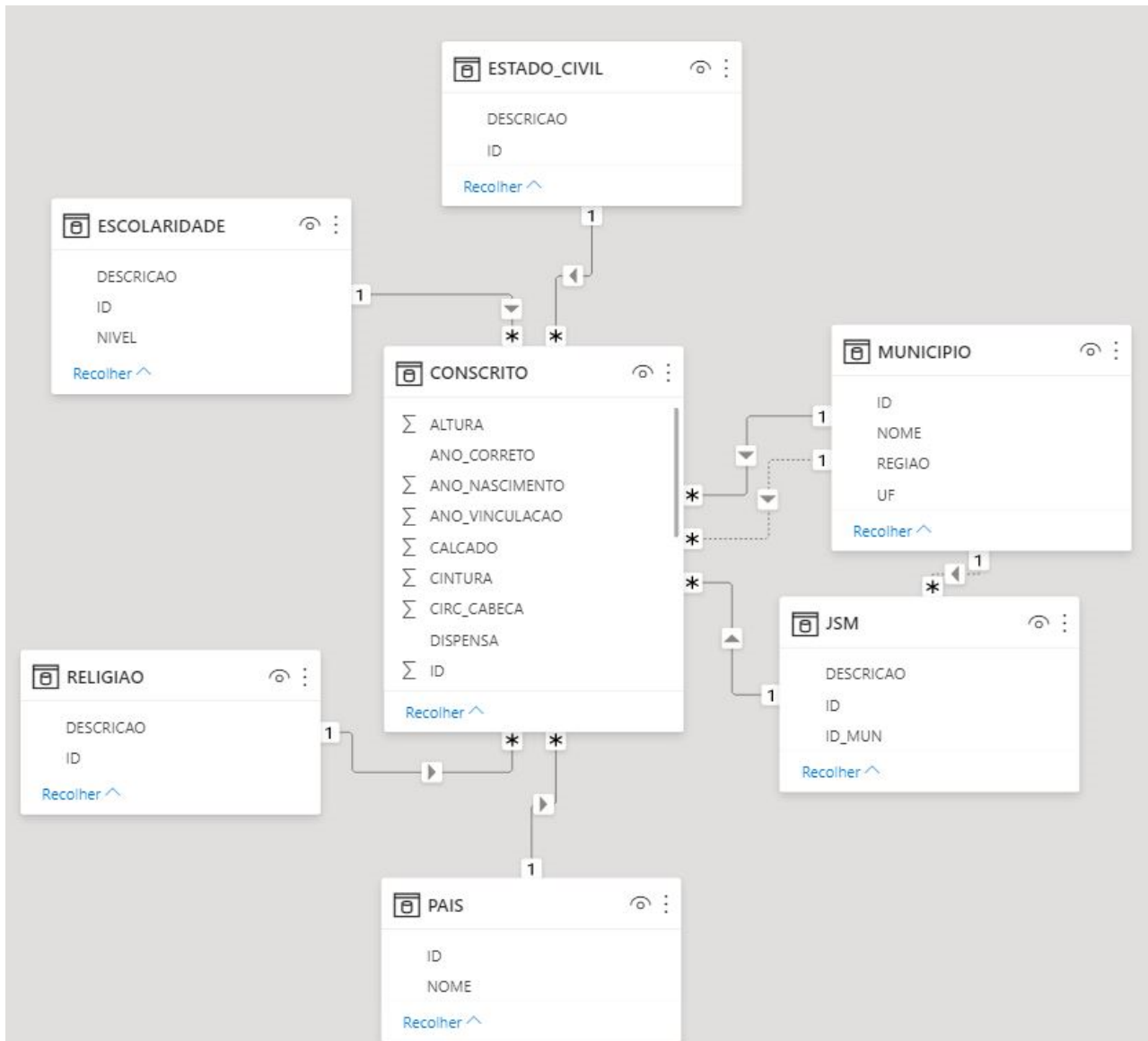
Quais as nacionalidades, fora a brasileira, que mais aparecem?



5.2 Plataforma Microsoft Power BI

5.2.1 Rotina de Extração, Transformação e Carga

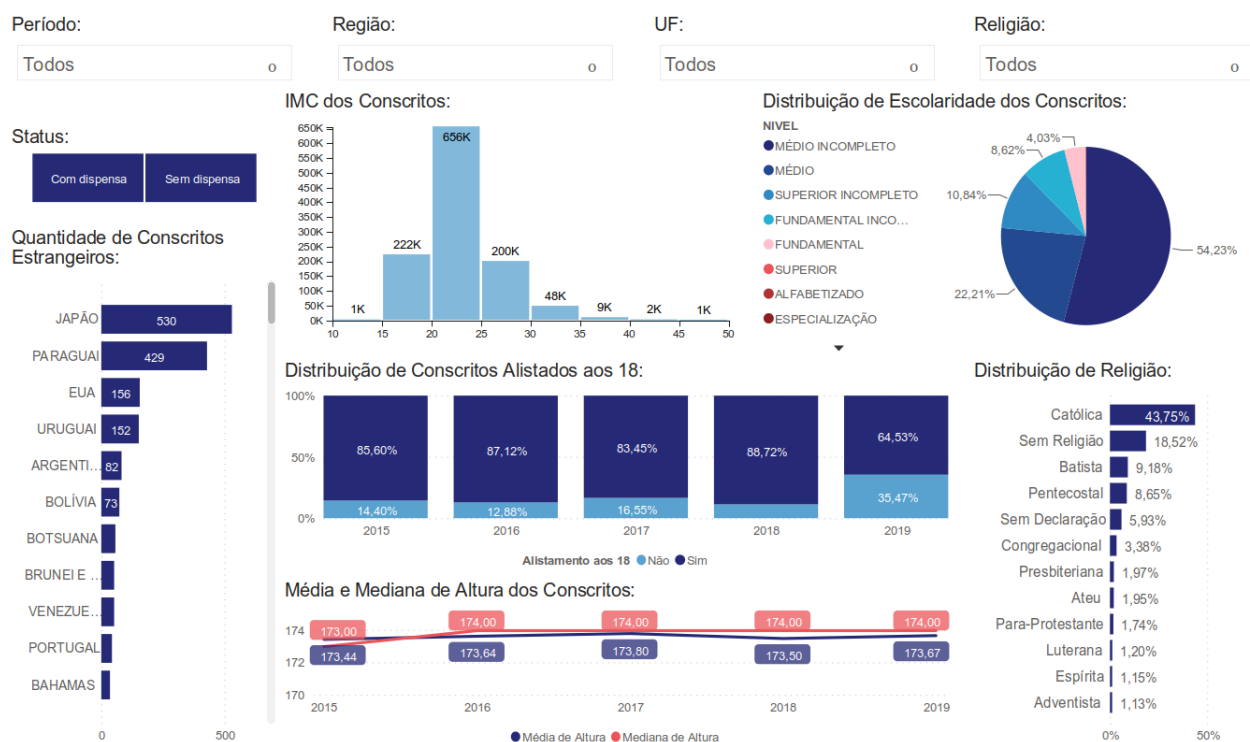
Considerando que não foi encontrada ferramenta com a nomenclatura solicitada, foi utilizada a parte de fluxo de dados do próprio Power BI desktop para a geração de relacionamentos conforme a imagem a seguir.



5.2.2 Visualizações

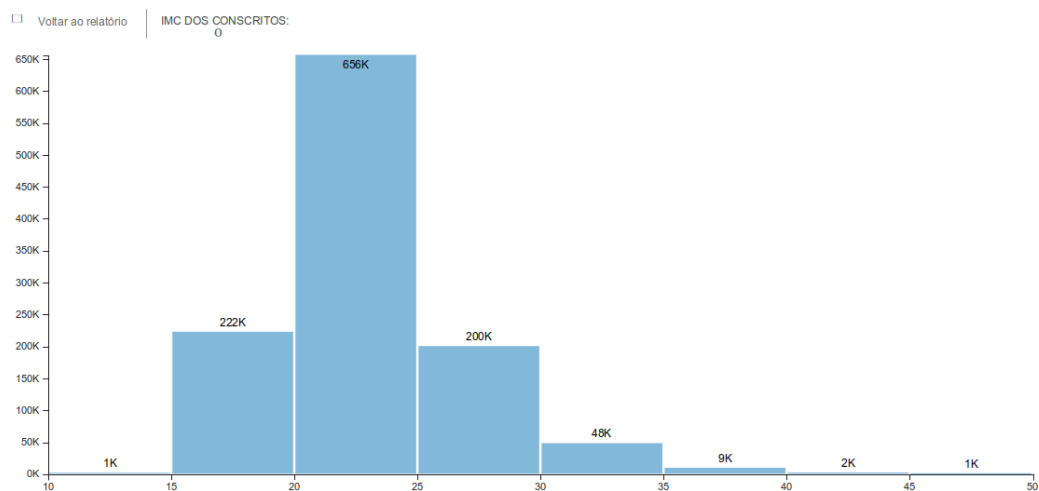
Dashboard Power BI:

Link para o vídeo do dashboard Power BI em funcionamento:

<https://www.youtube.com/watch?v=EMD1LT8sBZA>

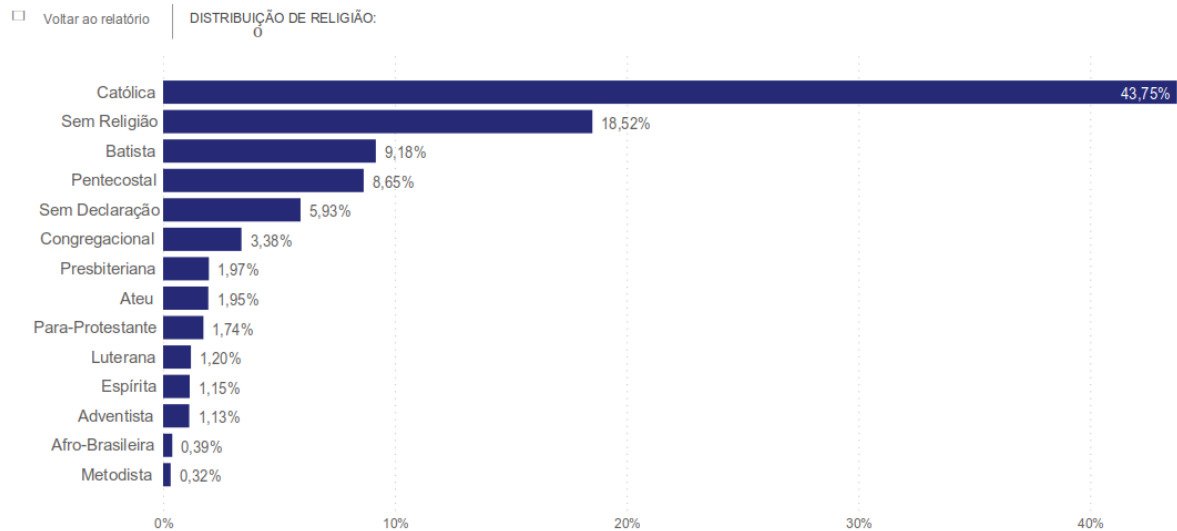
Questão 1 - Power BI:

Qual a distribuição do Índice de Massa Corporal (IMC) para os conscritos?



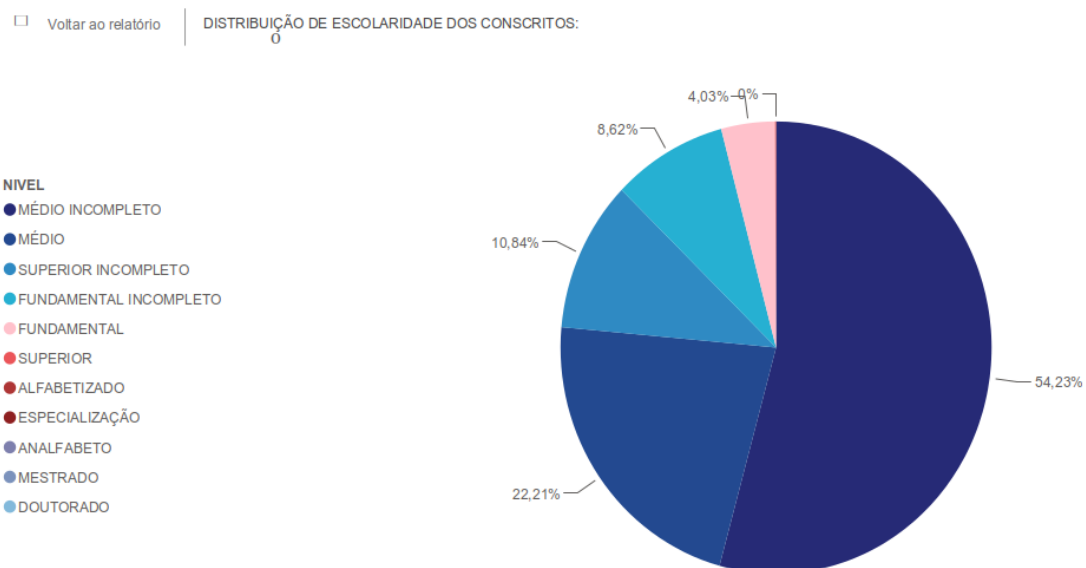
Questão 2 - Power BI:

Qual o percentual das religiões entre os conscritos?



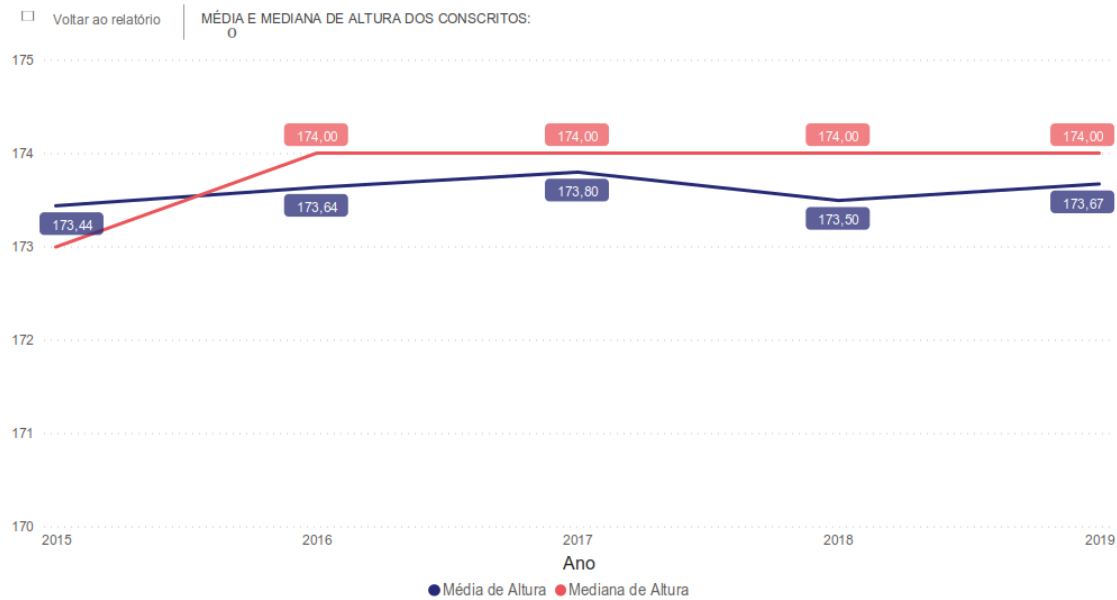
Questão 3 - Power BI:

Qual a distribuição do nível de escolaridade entre os conscritos?



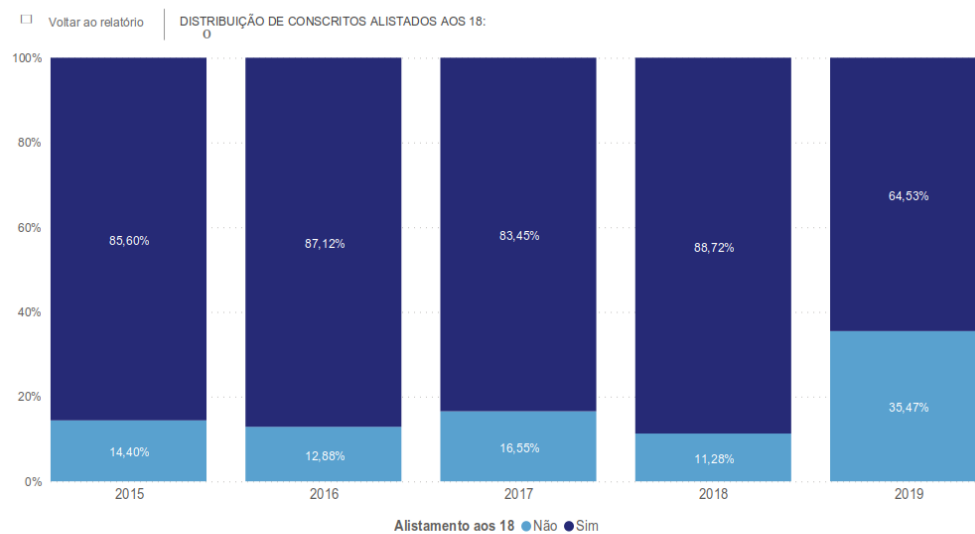
Questão 4 - Power BI:

Qual a média e mediana de altura dos conscritos?



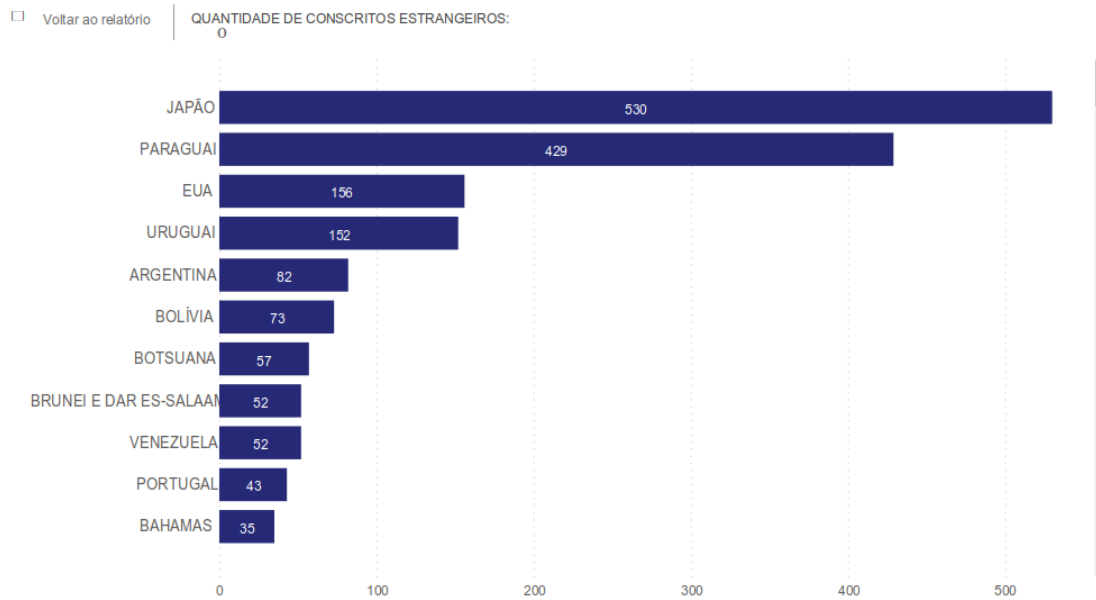
Questão 5 - Power BI:

Qual a porcentagem de conscritos que se alistaram ou não no ano correto?



Questão 6 - Power BI:

Quais as nacionalidades, fora a brasileira, que mais aparecem?



6 Conclusão

A análise dos dados é uma ferramenta fundamental, seja qual for o ramo de atuação. Conhecer e analisar o mercado é tarefa indispensável para o sucesso dos negócios nos dias de hoje.

Neste trabalho, demonstramos o uso de recursos de inteligência de negócios para analisar dados do alistamento militar Brasileiro. O negócio em questão é atípico, pois trata-se de uma função exclusiva do Estado e desta forma não há preocupação com variáveis de mercado e dados de concorrentes. As principais vantagens obtidas, em casos como este, são da coletividade, uma vez que a boa prestação de serviços por parte das forças armadas pode beneficiar de diversas formas a sociedade brasileira.

Desde o passo de extração dos dados, que foram acessados em um portal de dados públicos, passando pela modelagem e carga dos dados no SGBD, percebemos a necessidade do conhecimento de diversas áreas de atuação. A compreensão de áreas, como banco de dados, ferramentas de ETL e principalmente do domínio de negócio explorado, são essenciais para a execução do trabalho de inteligência de negócios.

Por fim, a análise em ferramentas, como o Tableau e o Power BI, permitiu ao grupo uma boa avaliação e síntese dos dados disponibilizados. Concluímos que se torna evidente que o uso de informações geradas para análise de inteligência de negócios são cruciais para a criação de estratégias, sobretudo nas públicas em que os interesses coletivos devem ser priorizados.