

Instituto de Pesquisa Econômica Aplicada - IPEA

CHAMADA PÚBLICA IPEA/PNPD Nº 088/2020

Relatório de Execução do Teste Aplicado

Candidato: **Bruno Miranda**

brunomiranda.work@gmail.com

<https://github.com/bmsmiranda/myportfolio/tree/main/IPEA>

61 9.9933 4416

Entrega do teste: 09/03/2020

Brasília, 2020.

Sumário

CONSIDERAÇÕES INICIAIS.....	3
CONFIGURAÇÕES PARA REALIZAÇÃO DO TESTE APLICADO	4
RESPOSTA DAS PERGUNTAS.....	5

CONSIDERAÇÕES INICIAIS

No link do meu perfil do GitHub indicado, foram disponibilizados além do material do teste aplicado, os relatórios finais do projeto em que atuei no Ipea, bem como dois artigos acadêmicos publicados por mim como primeiro autor.

Cabe a ressalva que todo o material foi disponibilizado para fins de comprovação de experiência e suas informações contidas não possuem caráter de divulgação de informação.

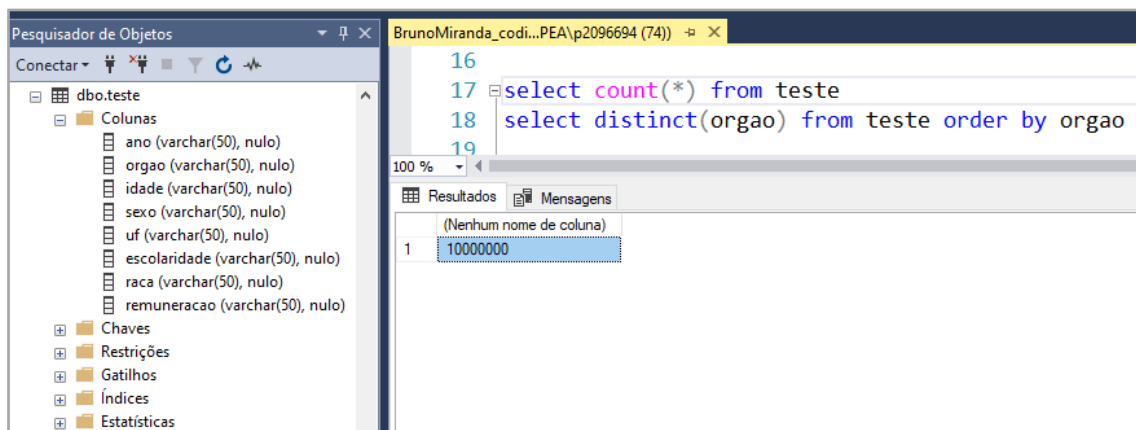
A seguinte estrutura foi disponibilizada no portfólio do Git:

- **PNPD-088-2020**
 - TESTE APLICADO
 - O código escrito em SQL para o teste foi disponibilizado para avaliação.
 - O código escrito em R para o teste foi disponibilizado para avaliação.
 - EXEMPLO DE ARTIGOS ACADÊMICOS PUBLICADOS
 - Constam dois artigos escritos, 1 publicado e 1 aceito para publicação.
- **POMUC-GASTOS-2019-2020**
 - Constam os documentos entregues ao final do Projeto Política sobre Mudanças do Clima (PoMuC) - Identificação dos Gastos Orçamentários Públicos Brasileiros com a Política Nacional sobre Mudança do Clima, realizado no Ipea.
 - No documento “POMUC-Gastos - relatório parcial de banco de dados (vrs. revisão)” consta a arquitetura da solução construída para banco de dados e análises feitas.
 - No documento “Relatório Identificação Gastos PoMuC” constam as análises estatísticas realizadas.
 - GASTOS_CARGA_SIOP
 - Consta o código fonte de uma solução de ETL codificada em Python 3.7 para o projeto executado.

CONFIGURAÇÕES PARA REALIZAÇÃO DO TESTE APLICADO

Primeiramente, dado a quantidade de registros disponibilizados (10 milhões), se optou por criar um banco de dados para fazer a importação dos registros.

- Se utilizou a versão *standalone* do SQL Server 2016
- A carga dos dados e criação da tabela de registro, foram realizadas via utilitário de importação, dado o caráter de simplificação do teste.



RESPOSTA DAS PERGUNTAS

Para a pergunta 1 não houve opção de resposta equivalente encontrada, uma vez que das opções disponíveis para resposta (3036, 3604, 3867, 3472 e 3336), nenhuma possui correspondência aos dados do *dataset* disponibilizado para o teste. Desse modo, o candidato preferiu não marcar nenhuma das opções disponíveis.

```
20 -- Pergunta 1
21 -- Resposta: opção não encontrada para seleção
22 select
23     count(*)
24 from
25     teste
26 where
27     1 = 1
28     and sexo = 'M'
29     and idade > 50
30     and uf = 'BA'
31     and ano = '2019'
32
```

00 %

Resultados Mensagens

(Nenhum nome de coluna)	
1	3108

Para a pergunta 2, foi realizada uma query buscando na base os funcionários de gênero feminino, agrupados por ano e órgão de atuação profissional. Uma vez ordenado os resultados, a resposta encontra foi:

- Ano de 2017, órgão ABC com um quantitativo de 882 funcionárias.

```

34  -- Pergunta 2
35  -- Resposta: 2017 - ABC
36  select
37      ano, orgao, count(*) as qtd
38  from
39      teste
40  where
41      1 = 1
42      and sexo = 'F'
43  group by
44      ano
45      , orgao
46  order by
47      qtd desc
48

```

Resultados

Mensagens

ano	orgao	qtd
2017	ABC	882
2018	ABC	852
2017	SBF	852
2017	ANA	850
2018	SBF	847
2017	SBB	835
2018	SBB	827
2016	ANA	784
2013	ANA	779

Para a pergunta 3, não houve resposta equivalente numérica dentre as opções disponíveis. Cabe ressaltar que no comando da questão não levou em consideração que na base de dados, para o ano de 2015 no órgão Ipea, existem registros com a escolaridade não preenchida, o que impacta no resultado final do cálculo, uma vez que o percentual será encontrado via a fórmula: $(\text{Qtd.Escolaridade 4} / \text{Qtd.Funcionario Total Ipea 2015}) * 100$.

- Desse modo, a opção “Nenhuma das alternativas foi selecionada”.

```

50 -- pergunta 3
51 -- Resposta: nenhuma das alternativas
52 declare @count decimal(8, 2);
53 set @count = cast((select count(*) from teste where orgao = 'Ipea' and ano = '2015'
54
55 select
56     escolaridade
57     , count(*) as qtd_esc
58     , @count as qtd_total
59     , ((CONVERT(decimal, (count(*))) / @count) * 100) as perc
60 from teste
61 where
62     1 = 1
63     and orgao = 'Ipea'
64     and escolaridade = 4
65     and ano = '2015'
66     and (escolaridade is not null and escolaridade != '')
67 group by
68     escolaridade
69 order by

```

escolaridade	qtd_esc	qtd_total	perc
4	470	1042.00	45.105566200

Para a pergunta 4, aplicou-se o filtro indicado para a consulta e se concluiu que dentre as mulheres para o ano de 2010 e órgão UFSJ, a raça Amarelo foi a menos predominante.

```

73 -- Pergunta 4
74 -- Resposta: Amarelo com 30 registros
75 select
76     raca, count(*) as qtd
77 from teste
78 where
79     1 = 1
80     and sexo = 'F'
81     and ano = '2010'
82     and orgao = 'UFSJ'
83     and (raca is not null and raca != '')
84 group by
85     raca
86 order by

```

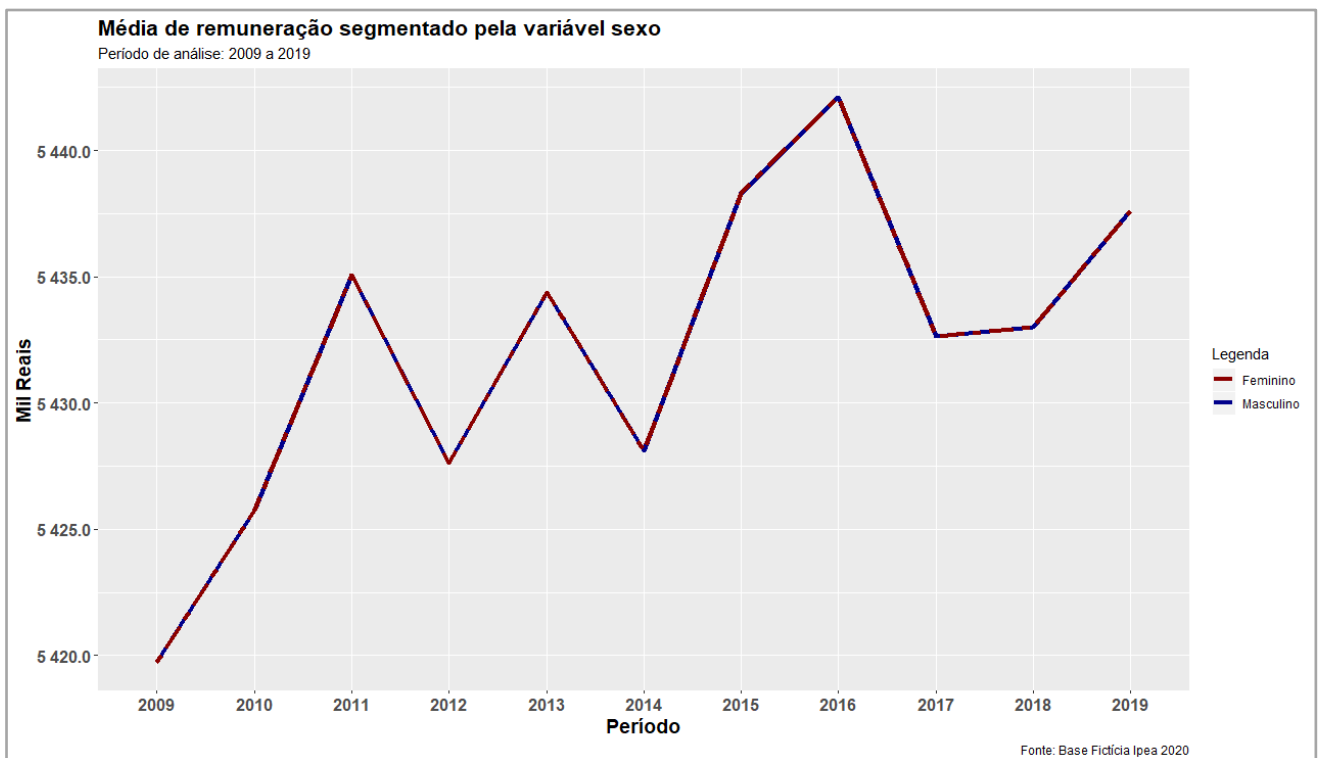
raca	qtd
Amarelo	30
Indigena	31
Preto	57
Pardo	87
Branco	102

Para a pergunta 5, optou-se por codificar em R um total de 3 análises:

i. MÉDIA DE REMUNERAÇÃO SEGMENTADO POR MEIO DA VARIÁVEL SEXO

O objetivo dessa análise foi de maneira simplória e não conclusiva, tentar captar o fenômeno da discriminação salarial devido ao gênero do funcionário. Desse modo, se obteve para os anos 2009-2019 a média salarial dos funcionários ano a ano.

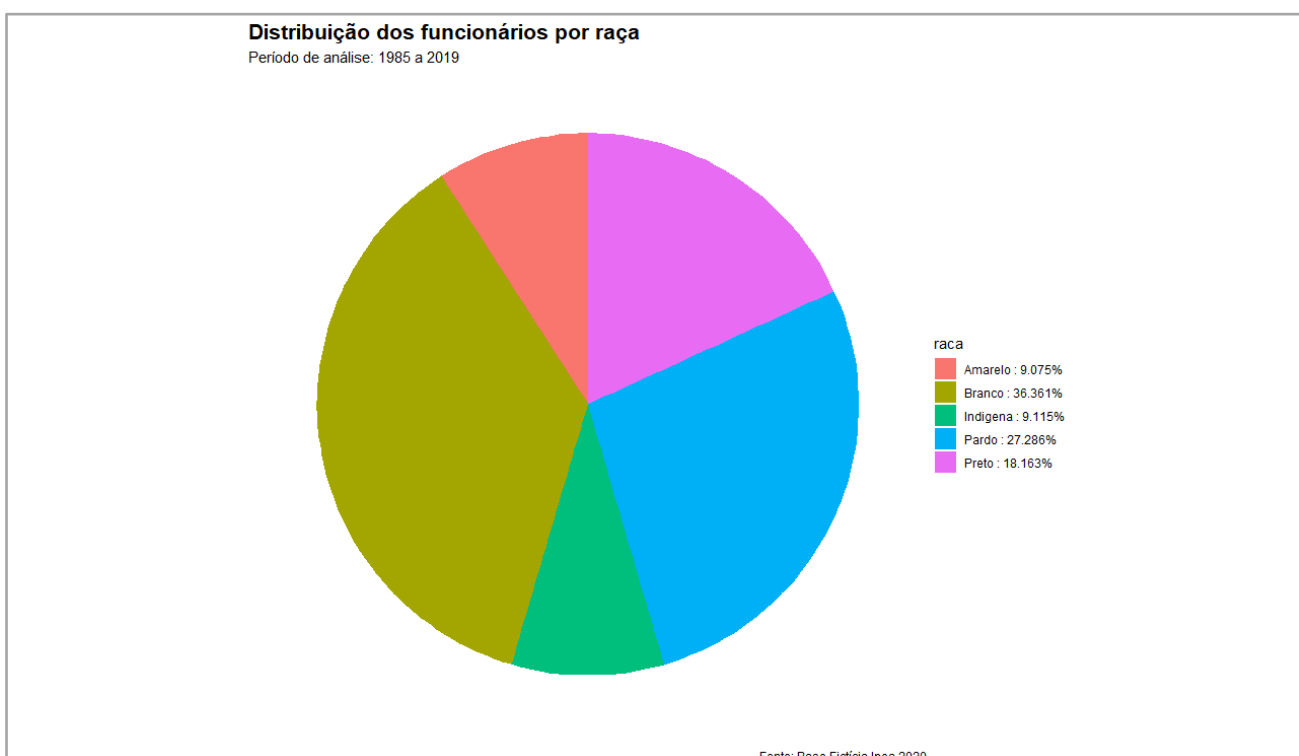
- Observou-se o comportamento de duas séries temporais independentes e sobrepostas, mostrando quase equivalência nos valores calculados para ambos os sexos, feminino e masculino.
- Por simplicidade, não se levou em conta a escolaridade dos funcionários como critério de segmentação.



ii. Distribuição dos funcionários por raça

O objetivo dessa análise foi de buscar a distribuição do fenômeno raça dentre os funcionários cadastrados. Desse modo, se obteve para os anos 1985-2019 a variável raça dos funcionários, excluindo os registros sem o preenchimento dessa variável.

- Como resultado da análise, observa-se, a prevalência da raça Branco e Pardo, com 36,36% e 27,28, respectivamente.



iii. Quantitativo de funcionários categorizado por sexo e segmentado por ano

O objetivo dessa última análise foi de buscar a evolução de um padrão de distribuição do fenômeno gênero dentre os funcionários cadastrados. Desse modo, se obteve para os anos 1985-2019 a variável sexo e o quantitativo dos funcionários, ano a ano.

- Como resultado da análise, pode-se observar a dominância do sexo Masculino na série histórica, a qual apresentar a partir do ano de 2017, um movimento de queda frente ao quantitativo computado para o sexo Feminino.

