

CIA014 - Análise Exploratória de Dados e Visualização
Laboratório 4



Instituto de Educação Superior de Brasília - IESB
Ciência de Dados e Inteligência Artificial

Brasil e a Pandemia do Covid-19 - Novo Coronavírus

Rafhael de Oliveira Martins

29 de novembro de 2020

Conteúdo

1	Introdução	4
2	Justificativa/motivação para o estado	5
3	Descrição da base de dados, com dicionário de dados	6
3.1	Descrição da Base de Dados	6
3.2	Dicionário de Dados	6
3.2.1	COVID19.sas	6
3.2.2	COVID19_CADASTRO.sas	7
3.2.3	COVID19_SERIE.sas	8
3.2.4	COVID19_TAXAS_BRASIL.sas	9
3.2.5	COVID19_TAXAS_ESTADO.sas	10
3.2.6	COVID19_TAXAS_REGIAO.sas	10
4	Análise Exploratória e Visualização dos dados do Covid-19	12
4.1	Análise do Covid-19 no Brasil	12
4.1.1	Análise da População do Brasil segundo dados TCU 2019	12
4.1.2	Análise das Taxas da COVID-19 por Região do Brasil	14
4.1.3	Análise dos Casos Confirmados do COVID-19 por Região e População	15
4.1.4	Análise dos Óbitos Confirmados do COVID-19 por Região	19
4.2	Análise do Covid-19 no Estado do Paraná (PR)	20
4.2.1	Análise de Casos Novos Confirmados por Data	20
4.2.2	Análise de Óbitos Novos Confirmados por Data	20
4.2.3	Análise de Óbitos Novos e Casos Novos Confirmados por Semana Epidemiológica	22
5	Técnicas aplicadas para o desenvolvimento do trabalho	23
6	Conclusão	25
7	Referências Bibliográficas	26

Lista de Figuras

1	Casos de cononavírus no Brasil - Ministério da Saúde	4
2	Variáveis na base de dados COVID19.sas	7
3	Variáveis na base de dados COVID19_CADASTRO.sas	7
4	Variáveis na base de dados COVID19_SERIE.sas	9
5	Variáveis na base de dados COVID19_TAXAS_BRASIL.sas	9
6	Variáveis na base de dados COVID19_TAXAS_ESTADO.sas	10
7	Variáveis na base de dados COVID19_TAXAS_REGIAO.sas	11
8	Análise da COVID19_TAXAS_BRASIL.sas	12
9	Tabela da Poulção do Brasil por Região	12
10	Poulção do Brasil por Região	13
11	Poulção do Brasil por Estado	13
12	Tabela de taxas da COVID-19 por Região do Brasil	14
13	Casos Confirmados por Região do Brasil	14
14	Casos Confirmados por Estado do Brasil	15
15	População e Casos Confirmados na Região Centro-Oeste do Brasil	16
16	População e Casos Confirmados na Região Nordeste do Brasil	16
17	População e Casos Confirmados na Região Norte do Brasil	17
18	População e Casos Confirmados na Região Sudeste do Brasil	18
19	População e Casos Confirmados na Região Sul do Brasil	18
20	Óbitos Confirmados por Região do Brasil	19
21	Casos Novos Confirmados por data no Estado do Paraná (PR)	20
22	Óbitos Novos Confirmados por data no Estado do Paraná (PR)	21
23	Óbitos Novos e Casos Novos Confirmados por Semana Epidemiológica no Estado do Paraná (PR)	22
24	Dashboard Covid-19 por Região do Brasil	23
25	Dashboard Covid-19 por Estado do Brasil	24
26	Dashboard Covid-19 por Mês e Semana Epidemiológica no Estado do Paraná (PR)	24

1 Introdução

A doença do **coronavírus (COVID-19)** é uma doença infecciosa causada por um coronavírus recém-descoberto.

A maioria das pessoas que adoece em decorrência da COVID-19 apresentará sintomas leves a moderados e se recuperará sem tratamento especial.

O vírus que causa a COVID-19 é transmitido principalmente por meio de gotículas geradas quando uma pessoa infectada tosse, espirra ou exala. Essas gotículas são muito pesadas para permanecerem no ar e são rapidamente depositadas em pisos ou superfícies. Você pode ser infectado ao inalar o vírus se estiver próximo de alguém que tenha COVID-19 ou ao tocar em uma superfície contaminada e, em seguida, passar as mãos nos olhos, no nariz ou na boca.

O Brasil registrou o primeiro caso do novo coronavírus SARS-CoV-2, causador da doença covid-19, no dia 26 de fevereiro. Foi em São Paulo. Um homem de 61 anos, cuja identidade não foi revelada, que esteve na Itália de 9 a 21 de fevereiro, mais especificamente na região da Lombardia, um dos epicentros da crise naquele país. Desde então, a infecção se alastrou por todos os Estados por meio de um tipo de transmissão chamada de comunitária, que não permite se saber onde, exatamente, uma pessoa contraiu o vírus.

Brasil conta **158.456 mortes** e **5.468.270 casos de coronavírus**. Dados do Ministério da Saúde até **28 de outubro**. A figura 1, extraída do site "CORONAVÍRUS BRASIL" (<https://covid.saude.gov.br>) detalha esses números.

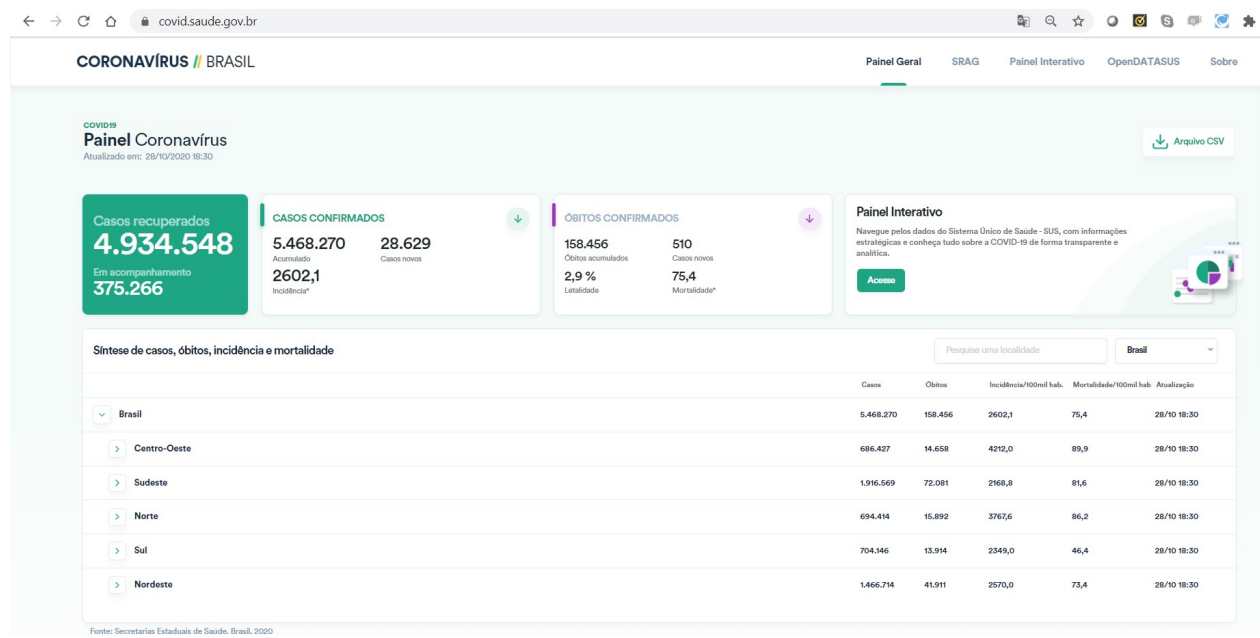


Figura 1: Casos de cononavírus no Brasil - Ministério da Saúde

Na figura 1, além dos totais de casos Confirmados, Recuperados e Óbitos, nos níveis de Brasil, Região e Estados você também encontra o Coeficiente de Incidência de COVID-19, Coeficiente de Mortalidade por COVID-19 e Mortalidade e a Taxa de Letalidade por COVID-19. A metodologia de cálculo desses coeficientes e da taxa está na aba "Sobre" do painel.

2 Justificativa/motivação para o estado

Brasil oficialmente República Federativa do Brasil, é o maior país da América do Sul e da região da América Latina, sendo o quinto maior do mundo em área territorial com 47,3% do território sul-americano contendo mais de 8 milhões de km² e sexto em população com mais de 210 milhões de habitantes. Único país na América onde se fala majoritariamente a língua portuguesa e o maior país lusófono do planeta, além de ser uma das nações mais multiculturais e etnicamente diversas, em decorrência da forte imigração oriunda de variados locais do mundo. Formado por 26 estados mais o Distrito Federal onde se localiza a sede do governo brasileiro e a capital do país a cidade de Brasília, totalizando 27 unidades federativas do Brasil (UF) com 5570 municípios. Sua moeda local é o Real representada pelo símbolo (R\$), sendo seu PIB (Produto Interno Bruto) em 2019 de R\$ 7,3 trilhões de reais, sendo o último trimestre divulgado (2º trimestre de 2020), o valor foi de R\$ 1 653,0 bilhões de reais. Mais informações do país, podem ser encontradas neste website (<https://www.gov.br/pt-br>).

O Estado do Paraná (PR) é uma das 27 Unidades da Federativas do Brasil, sendo localizado ao norte da região Sul do país. Possui 399 municípios, uma área de mais de 199 mil km² o décimo quinto do país, uma população estimada em mais de 11 milhões de habitantes, sendo sua capital é a cidade de Curitiba. Tendo o quinto maior PIB do Brasil com R\$ 421.375 bilhões de reais, é o estado onde se localiza uma das principais Usinas Hidrelétricas do Brasil a Usina de Itaipu Binacional sendo parte do Governo Brasileiro e parte do Governo Paraguaio, localizada na divisa entre os dois países, está no estado do PR do lado brasileiro, no Rio Paraná, na cidade de Foz do Iguaçu (Brasil) e cidade de Hernandarias (Paraguai). Mais informações do Estado, podem ser encontradas neste website (<http://www.parana.pr.gov.br/>).

O Estado do Paraná (PR) foi escolhido como proposta de estudo e análise exploratória dos dados da COVID-19 no Brasil e neste estado do PR conforme indicação do professor.

3 Descrição da base de dados, com dicionário de dados

3.1 Descrição da Base de Dados

Utilizando a base de dados do covid-19, citada anteriormente, foram gerados os arquivos relacionados a seguir, visando facilitar a análise exploratória e visualização dos dados:

1. covid19.sas7bdat (arquivo original no formato SAS)
2. covid19_cadastro.sas7bdat (dados de população)
3. covid19_serie.sas7bdat (série histórica até 28/10/2020)
4. covid19_taxas_brasil (coeficientes e taxa no nível Brasil)
5. covid19_taxas_regiao (coeficientes e taxa no nível Região)
6. covid19_taxas_estado (coeficientes e taxa no nível Estado)

Os arquivos estão disponíveis na pasta CIA (X:/Dados/Dados_2019/Covid-19) nos laboratórios do IESB e disponíveis para download no link

<https://drive.google.com/drive/folders/14Z3UWwdxyz-YZB7u5kU3GF3LsdwJYF6t?usp=sharing>

3.2 Dicionário de Dados

O dicionário de dados visa explicar as variáveis presentes na base de dados abordada neste trabalho. Portanto, as bases de dados detêm as seguintes variáveis:

3.2.1 COVID19.sas

- **Região:** Nome da região do País.
- **Estado:** Sigla da Unidade da Federação (UF) de cada estado do Brasil.
- **Município:** Nome do município ao qual foi computado os dados.
- **Coduf:** Código da Unidade da Federação (UF) representada numericamente.
- **Codmun:** Código do município representada numericamente.
- **CodRegiaoSaude:** Código da região saúde, representada numericamente.
- **NomeRegiaoSaude:** Nome da regioao saúde.
- **Data:** Data representada com YYYY-MM-DD (Ano-Mês-Dia).
- **SemanaEpi:** Número da semana.
- **PopulacaoTCU2019:** Tamanho da população no ano de 2019 segunda base do Tribunal de Contas da União (TCU).
- **CasosAcumulados:** Número de casos Acumulados, soma do número de casos por dia, tornando o número acumulado, sabendo-se o total.
- **CasosNovos:** Número de casos novos, informa o número de casos novos no dia em questão.
- **ObitosAcumulados:** Número de obitos acumulados, soma do número de obitos por dia, tornando o número acumulado, sabendo-se o total.

- **ObitosNovos:** Número de obitos novos, informa o número de obitos novos no dia em questão.
- **RecuperadosNovos:** Número de recuperados novos, soma destes números no dia em questão.
- **EmAcompanhamentoNovos:** Número de casos novos em acompanhamento, soma destes números, sabendo-se o total.
- **Interior/Metropolitana:** Número da região se é Interior ou Metropolitana.

Variables in Creation Order					
#	Variable	Type	Len	Format	Informat
1	regiao	Char	12	\$CHAR12.	\$CHAR12.
2	estado	Char	2	\$CHAR2.	\$CHAR2.
3	municipio	Char	33	\$CHAR33.	\$CHAR33.
4	coduf	Num	8	BEST2.	BEST2.
5	codmun	Num	8	BEST6.	BEST6.
6	codRegiaoSaude	Num	8	BEST5.	BEST5.
7	nomeRegiaoSaude	Char	42	\$CHAR42.	\$CHAR42.
8	data	Num	8	YYMMDD10.	YYMMDD10.
9	semanaEpi	Num	8	BEST2.	BEST2.
10	populacaoTCU2019	Num	8	BEST9.	BEST9.
11	casosAcumulado	Num	8	BEST7.	BEST7.
12	casosNovos	Num	8	BEST5.	BEST5.
13	obitosAcumulado	Num	8	BEST6.	BEST6.
14	obitosNovos	Num	8	BEST4.	BEST4.
15	Recuperadosnovos	Num	8	BEST7.	BEST7.
16	emAcompanhamentoNovos	Num	8	BEST6.	BEST6.
17	interior/metropolitana	Num	8	BEST1.	BEST1.

Figura 2: Variáveis na base de dados COVID19.sas

A base de dados COVID19.sas tem um tamanho total de 232MB contendo 1.214.572 milhões de observações ou linhas e 17 Variáveis ou colunas.

3.2.2 COVID19.CADASTRO.sas

- **Região:** Nome da região do País.
- **Estado:** Sigla da Unidade da Federação (UF) de cada estado do Brasil.
- **Coduf:** Código da Unidade da Federação (UF) representada numericamente.
- **PopulacaoTCU2019:** Tamanho da população no ano de 2019 segunda base do Tribunal de Contas da União (TCU).

Variables in Creation Order					
#	Variable	Type	Len	Format	Informat
1	regiao	Char	12	\$CHAR12.	\$CHAR12.
2	estado	Char	2	\$CHAR2.	\$CHAR2.
3	coduf	Num	8	BEST2.	BEST2.
4	populacaoTCU2019	Num	8	BEST9.	BEST9.

Figura 3: Variáveis na base de dados COVID19.CADASTRO.sas

A base de dados COVID19_CADASTRO.sas tem um tamanho total de 128KB contendo 27 observações ou linhas e 4 Variáveis ou colunas.

3.2.3 COVID19_SERIE.sas

- **Região:** Nome da região do País.
- **Estado:** Sigla da Unidade da Federação (UF) de cada estado do Brasil.
- **Município:** Nome do município ao qual foi computado os dados.
- **Coduf:** Código da Unidade da Federação (UF) representada numericamente.
- **Codmun:** Código do município representada numericamente.
- **CodRegiaoSaude:** Código da região saúde, representada numericamente.
- **NomeRegiaoSaude:** Nome da regio saúde.
- **Data:** Data representada com YYYY-MM-DD (Ano-Mês-Dia).
- **SemanaEpi:** Número da semana.
- **CasosAcumulados:** Número de casos Acumulados, soma do número de casos por dia, tornando o número acumulado, sabendo-se o total.
- **CasosNovos:** Número de casos novos, informa o número de casos novos no dia em questão.
- **ObitosAcumulados:** Número de obitos acumulados, soma do número de obitos por dia, tornando o número acumulado, sabendo-se o total.
- **ObitosNovos:** Número de obitos novos, informa o número de obitos novos no dia em questão.
- **RecuperadosNovos:** Número de recuperados novos, soma destes números no dia em questão.
- **EmAcompanhamentoNovos:** Número de casos novos em acompanhamento, soma destes números, sabendo-se o total.
- **Interior/Metropolitana:** Número da região se é Interior ou Metropolitana.

A base de dados COVID19_SERIE.sas tem um tamanho total de 1MB contendo 6.669 mil observações ou linhas e 16 Variáveis ou colunas.

Variables in Creation Order					
#	Variable	Type	Len	Format	Informat
1	regiao	Char	12	\$CHAR12.	\$CHAR12.
2	estado	Char	2	\$CHAR2.	\$CHAR2.
3	municipio	Char	33	\$CHAR33.	\$CHAR33.
4	coduf	Num	8	BEST2.	BEST2.
5	codmun	Num	8	BEST6.	BEST6.
6	codRegiaoSaude	Num	8	BEST5.	BEST5.
7	nomeRegiaoSaude	Char	42	\$CHAR42.	\$CHAR42.
8	data	Num	8	YYMMDD10.	YYMMDD10.
9	semanaEpi	Num	8	BEST2.	BEST2.
10	casosAcumulado	Num	8	BEST7.	BEST7.
11	casosNovos	Num	8	BEST5.	BEST5.
12	obitosAcumulado	Num	8	BEST6.	BEST6.
13	obitosNovos	Num	8	BEST4.	BEST4.
14	Recuperadosnovos	Num	8	BEST7.	BEST7.
15	emAcompanhamentoNovos	Num	8	BEST6.	BEST6.
16	interior/metropolitana	Num	8	BEST1.	BEST1.

Figura 4: Variáveis na base de dados COVID19_SERIE.sas

3.2.4 COVID19_TAXAS_BRASIL.sas

- **Populacao_Brasil:** Número com o valor total da população brasileira.
- **Região:** Nome da região do País, neste caso, o Brasil como região.
- **Estado:** Sigla da Unidade da Federação (UF) de cada estado do Brasil, neste caso o Brasil como estado.
- **Casos_Confirmados_Brasil:** Número de casos confirmados no Brasil com a COVID-19.
- **Obitos_Confirmado_Brasil:** Número de obitos confirmados no Brasil devido a COVID-19.
- **Brasil_Incidencia:** Número total representando o valor da taxa de incidência no Brasil pela COVID-19.
- **Brasil_Mortalidade:** Número total representando o valor de mortes pela COVID-19 no Brasil.
- **Brasil_Letalidade:** Número total representando o valor da letalidade pela COVID-19 no Brasil.

Variables in Creation Order					
#	Variable	Type	Len	Format	Informat
1	Populacao_Brasil	Num	8	BEST9.	BEST9.
2	Regiao	Char	6		
3	Estado	Char	2		
4	Casos_Confirmados_Brasil	Num	8	BEST5.	BEST5.
5	Obitos_Confirmados_Brasil	Num	8	BEST4.	BEST4.
6	Brasil_Incidencia	Num	8		
7	Brasil_Mortalidade	Num	8		
8	Brasil_Letalidade	Num	8		

Figura 5: Variáveis na base de dados COVID19_TAXAS_BRASIL.sas

A base de dados COVID19.TAXAS.BRASIL.sas tem um tamanho total de 128KB contendo 1 observação ou linha e 8 Variáveis ou colunas.

3.2.5 COVID19_TAXAS_ESTADO.sas

- **Região:** Nome da região do País, neste caso, o Brasil como região.
- **Estado:** Sigla da Unidade da Federação (UF) de cada estado do Brasil, neste caso o Brasil como estado.
- **Coduf:** Código da Unidade da Federação (UF) representada numericamente.
- **PopulacaoTCU2019:** Tamanho da população no ano de 2019 segunda base do Tribunal de Contas da União (TCU).
- **Casos_Confirmados_Estado:** Número de casos confirmados nos Estados com a COVID-19.
- **Obitos_Confirmado_Estados:** Número de obitos confirmados nos Estados devido a COVID-19.
- **Estado_Incidencia:** Número total representando o valor da taxa de incidência nos Estados pela COVID-19.
- **Estados_Mortalidade:** Número total representando o valor de mortes pela COVID-19 nos Estados.
- **Estados_Letalidade:** Número total representando o valor da letalidade pela COVID-19 nos Estados.

Variables in Creation Order					
#	Variable	Type	Len	Format	Informat
1	regiao	Char	12	\$CHAR12.	\$CHAR12.
2	estado	Char	2	\$CHAR2.	\$CHAR2.
3	coduf	Num	8	BEST2.	BEST2.
4	populacaoTCU2019	Num	8	BEST9.	BEST9.
5	Casos_Confirmados_Estado	Num	8	BEST5.	BEST5.
6	Obitos_Confirmados_Estado	Num	8	BEST4.	BEST4.
7	Estado_Incidencia	Num	8		
8	Estado_Mortalidade	Num	8		
9	Estado_Letalidade	Num	8		

Figura 6: Variáveis na base de dados COVID19_TAXAS_ESTADO.sas

A base de dados COVID19.TAXAS_ESTADO.sas tem um tamanho total de 128KB contendo 27 observações ou linhas e 9 Variáveis ou colunas.

3.2.6 COVID19_TAXAS_REGIAO.sas

- **Região:** Nome da região do País, neste caso, o Brasil como região.
- **Populacao_Regiao:** Número com o valor total da população por região do Brasil.
- **Casos_Confirmados_Regiao:** Número de casos confirmados por Regiao com a COVID-19.
- **Obitos_Confirmado_Regiao:** Número de obitos confirmados por Região devido a COVID-19.
- **Regiao_Incidencia:** Número total representando o valor da taxa de incidência nas Região dos Estados brasileiros pela COVID-19.

- **Regiao_Mortalidade:** Número total representando o valor de mortes pela COVID-19 por Região.
- **Regiao_Letalidade:** Número total representando o valor da letalidade pela COVID-19 por Região.

Variables in Creation Order					
#	Variable	Type	Len	Format	Informat
1	regiao	Char	12	\$CHAR12.	\$CHAR12.
2	Populacao_Regiao	Num	8	BEST9.	BEST9.
3	Casos_Confirmados_Regiao	Num	8	BEST5.	BEST5.
4	Obitos_Confirmados_Regiao	Num	8	BEST4.	BEST4.
5	Regiao_Incidencia	Num	8		
6	Regiao_Mortalidade	Num	8		
7	Regiao_Letalidade	Num	8		

Figura 7: Variáveis na base de dados COVID19_TAXAS_REGIAO.sas

A base de dados COVID19_TAXAS_REGIAO.sas tem um tamanho total de 128KB contendo 5 observações ou linhas e 7 Variáveis ou colunas.

As Figuras 2, 3, 4, 5, 6 e 7 mostram as variáveis da base de dados com seu nome da variável ou coluna *Variable*, como também seu tipo, Numérica (Num) ou Categórica (Char) mostrado na coluna *Type*, tamanho da variável mostrado na coluna *Len* e seu formato disposto na base, demonstrado nas colunas *Format* e *Informat*.

4 Análise Exploratória e Visualização dos dados do Covid-19

Através das ferramentas SAS Enterprise Guide 7.1 e Power BI Desktop, será apresentado neste trabalho informações da COVID-19 no Brasil e no estado do Paraná (PR).

4.1 Análise do Covid-19 no Brasil

Com cerca de um pouco mais de 210 milhões de habitantes, o Brasil vem tendo um grande número de casos de COVID-19. Como em outros países em que sua população é grande em quantidade como em alguns também em território, a COVID-19 aumenta rapidamente e em diversas localidades do território nacional, no Brasil este reflexo vem acontecendo da mesma forma, vemos isso na tabela abaixo na *figura 8*, em que nos mostra as taxas e o avanço da COVID-19 no Brasil até a data de 28/10/2020, última data registrada na base de dados coletada neste dia para a análise deste trabalho.

Populacao_Brasil	Regiao	Estado	Casos_Confirmados_Brasil	Obitos_Confirmados_Brasil	Brasil_Incidencia	Brasil_Mortalidade	Brasil_Letalidade
210147125	Brasil	BR	547E4	16E4	2602.1150658	75.402411525	2.8977354812

Figura 8: Análise da COVID19.TAXAS.BRASIL.sas

Em 28/10/2020 o Brasil registrou uma marca de 5.468.270 milhões de CASOS CONFIRMADOS de COVID-19 e 158.456 mil OBITOS CONFIRMADOS pela COVID-19, uma perda grande se comparado com a TAXA DE INCIDÊNCIA DE CASOS que girava em torno de 2602.11 casos por 100 mil habitantes no Brasil diariamente, sendo a TAXA DE MORTALIDADE em 75.40 mortes por 100 mil habitantes e uma TAXA DE LETALIDADE de 2.89%.

4.1.1 Análise da População do Brasil segundo dados TCU 2019

Analysis Variable : populacaoTCU2019							
regiao	N Obs	Mean	Std Dev	Minimum	Maximum	Sum	N
Centro-Oeste	4	4074268.50	1984502.06	2778986.00	7018354.00	16297074.00	4
Nordeste	9	6341294.89	4178364.19	2298696.00	14873064.00	57071654.00	9
Norte	7	2632997.14	2891098.57	605761.00	8602865.00	18430980.00	7
Sudeste	4	22092858.25	17497876.40	4018650.00	45919049.00	88371433.00	4
Sul	3	9991994.67	2448597.02	7164788.00	11433957.00	29975984.00	3

Figura 9: Tabela da Poulção do Brasil por Região

A Região Sudeste possui a maior população do país com 88.371.433 milhões de habitantes, seguido da Região Nordeste com 57.057.654 milhões de habitantes e em terceiro lugar com maior população vem a Região Sul do país com 29.975.984 milhões de habitantes, de acordo com a tabela da *figura 9*. Vemos ainda que a Região Norte vem em quarto lugar com 18.430.980 milhões de habitantes e em quinto lugar a Região Centro-Oeste com 16.297.074 milhões de habitantes. Informações estas registradas pelo TCU em 2019 disposta nessa base de dados da COVID-19 do DATASUS, ainda podemos ver está distribuição da população brasileira por região na *figura 10* ficando bem claro a região mais populosa até a de menor população.

Uma curiosidade desta tabela da *figura 9* é a informação de Mínimo e Máximo onde podemos ver por região qual detém a menor e maior população do brasil por região, mas só por está tabela não sabemos qual estado da região dispõe do valor informado, sendo necessário analisar a distribuição por Estado, vemos isto na *figura 11*.

Na **Região Centro-Oeste** o Estado de Mato Grosso do Sul (MS) é o de menor população com 2.778.986 milhões de habitantes seguido do Estado de Goiás (GO) com a maior população de 7.018.354 milhões de habitantes. Na **Região Nordeste** o Estado de Sergipe (SE) é o de menor população com 2.298.696 milhões de habitantes seguido do Estado da Bahia (BA) com a maior população de 14.873.064 milhões de habitantes

sendo também o quarto Estado mais populoso do país. Na **Região Norte** o Estado de Roraima (RR) é o de menor população com 605.761 mil habitantes seguido do Estado do Pará (PA) com a maior população de 8.602.865 milhões de habitantes.

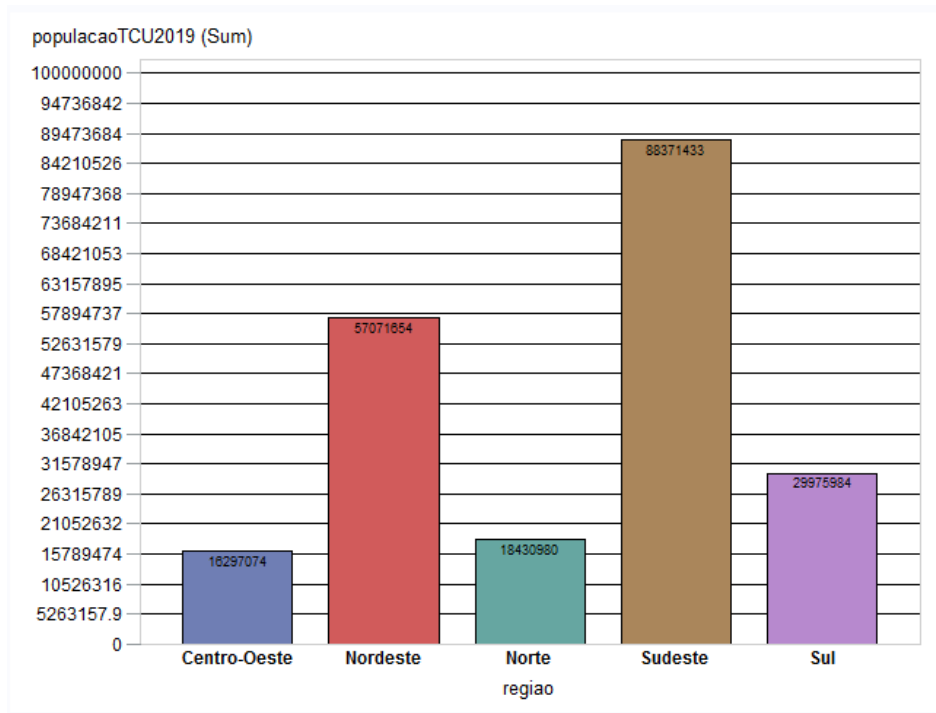


Figura 10: Poulção do Brasil por Região

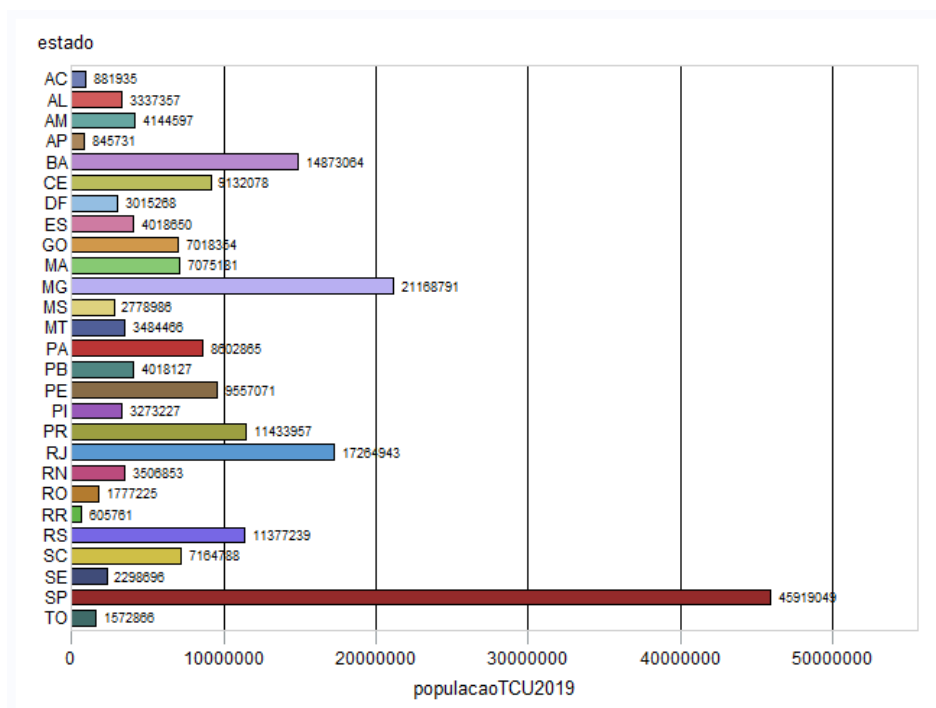


Figura 11: Poulção do Brasil por Estado

Na **Região Sudeste** o Estado do Espírito Santo (ES) é o de menor população com 4.018.650 milhões de

habitantes seguido do Estado de São Paulo (SP) com a maior população de 45.919.049 milhões de habitantes sendo também a mais populosa do país. Na **Região Sul** o Estado de Santa Catarina (SC) é o de menor população com 7.164.788 milhões de habitantes seguido do Estado do Paraná (PR) com a maior população de 11.433.957 milhões de habitantes sendo também o terceiro Estado mais populoso do país.

4.1.2 Análise das Taxas da COVID-19 por Região do Brasil

regiao	Populacao_Regiao	Casos_Confirmados_Regiao	Obitos_Confirmados_Regiao	Regiao_Incidencia	Regiao_Mortalidade	Regiao_Letalidade
Centro-Oeste	16297074	686E3	15E3	4211.9646754	89.942525879	2.1354055129
Nordeste	57071654	147E4	42E3	2569.9518013	73.43575499	2.8574759633
Norte	18430980	694E3	16E3	3767.645562	86.224389588	2.28854833
Sudeste	88371433	192E4	72E3	2168.7653294	81.565951296	3.7609394705
Sul	29975984	704E3	14E3	2349.0338132	46.417158483	1.9760106569

Figura 12: Tabela de taxas da COVID-19 por Região do Brasil

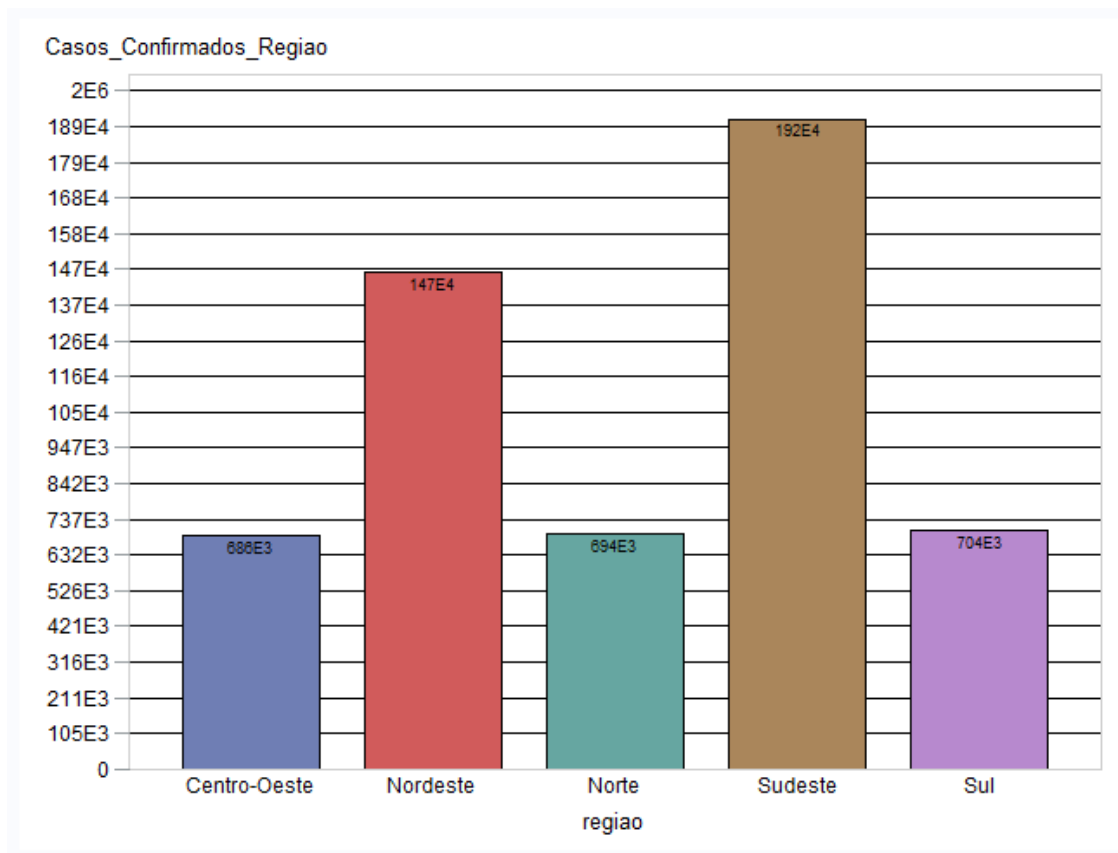


Figura 13: Casos Confirmados por Região do Brasil

A **Região Sudeste** do país possui o maior volume de casos confirmados com a COVID-19 com 1.916.569 milhões de casos confirmados como mostrado na *figura 13*, seguido da **Região Nordeste** do país com 1.466.714 milhões de casos confirmados, puxado principalmente pelos estados de *São Paulo (SP)* e o estado da *Bahia (BA)*, respectivamente a cada região, como mostrado os casos confirmados por estado na *figura 14*. Até a data do dia 28/10/2020 que foi registrada pela base de dados, a **Região Centro-Oeste** foi a que teve o menor número de casos confirmados com o novo coronavírus com 686.427 mil casos confirmados.

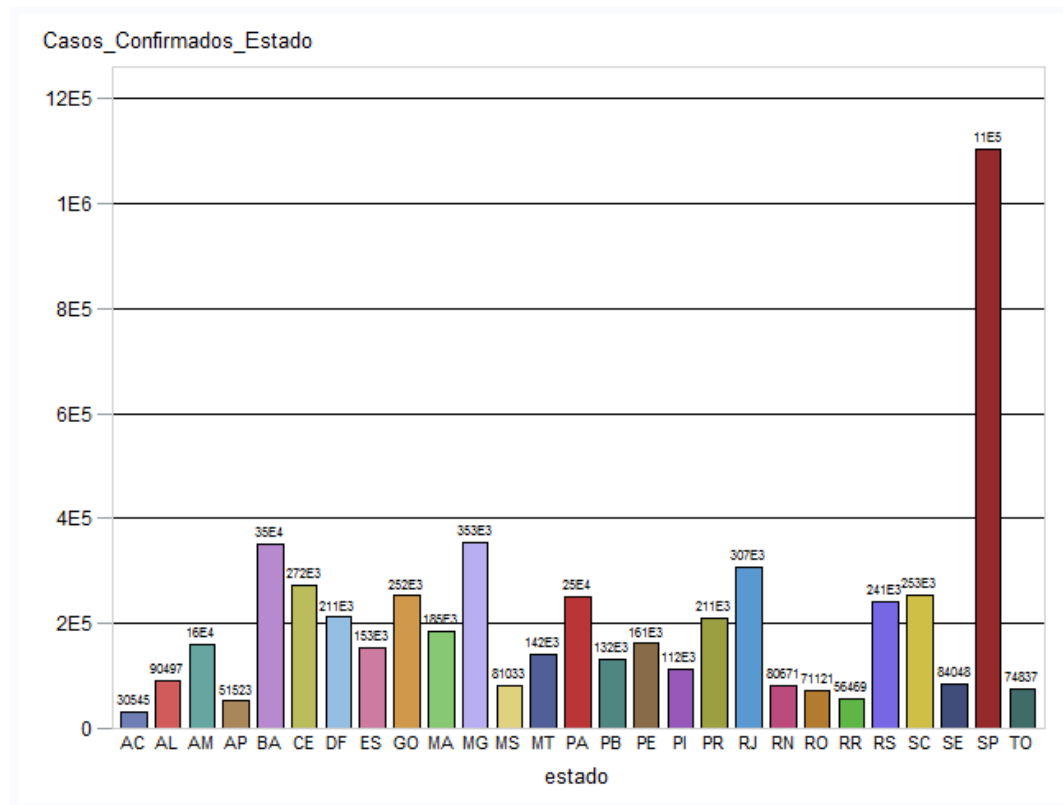


Figura 14: Casos Confirmados por Estado do Brasil

Fazendo uma ligação entre tamanho populacional e casos confirmados com a COVID-19, pode-se perceber que as regiões com os maiores números de casos confirmados são as mais populosas e as com menores números de casos confirmados, as regiões de menor número populacional, como vimos nas *figuras 10, 11, 12, 13 e 14*.

4.1.3 Análise dos Casos Confirmados do COVID-19 por Região e População

Observando as *figuras 15, 16, 17, 18 e 19*, é nítido a distribuição de casos confirmados com o novo coronavírus por região e estados nestas regiões. Vemos também que nos estados onde a população é maior conforme mostrado na *figura 11*, os casos confirmados são maiores também.

Na **Região Centro-Oeste** o estado de *Goiás (GO)* com uma população de 7.018.354 milhões de habitantes, teve 252.333 mil casos confirmados com o COVID-19, e o estado do *Mato Grosso do Sul (MS)* com 2.778.986 milhões de habitantes, teve até 28/10/2020 cerca de 81.033 mil casos confirmados com a COVID-19, sendo o menos afetado na região, conforme mostrado na *figura 15*.

Na **Região Nordeste** vemos o estado da *Bahia (BA)* com 14.873.064 milhões de habitantes e o estado do *Ceara (CE)* com 9.132.078 milhões de habitantes, estes com 349.711 mil casos confirmados e 272.104 mil casos confirmados respectivamente, sendo os maiores em população como também em casos confirmados com o COVID-19 da região. O estado do *Rio Grande do Norte (RN)* é o estado menos afetado da região com uma população de 3.506.853 milhões de habitantes e 80.671 mil casos confirmados, como mostrado na *figura 16*.

Na **Região Norte** o estado do *Pará (PA)* com uma população de 8.602.865 milhões de habitantes e o estado do *Amazonas (AM)* com uma população 4.144.597 milhões de habitantes estão com o maior volume de casos confirmados na região com 250.336 mil e 159.583 mil respectivamente.

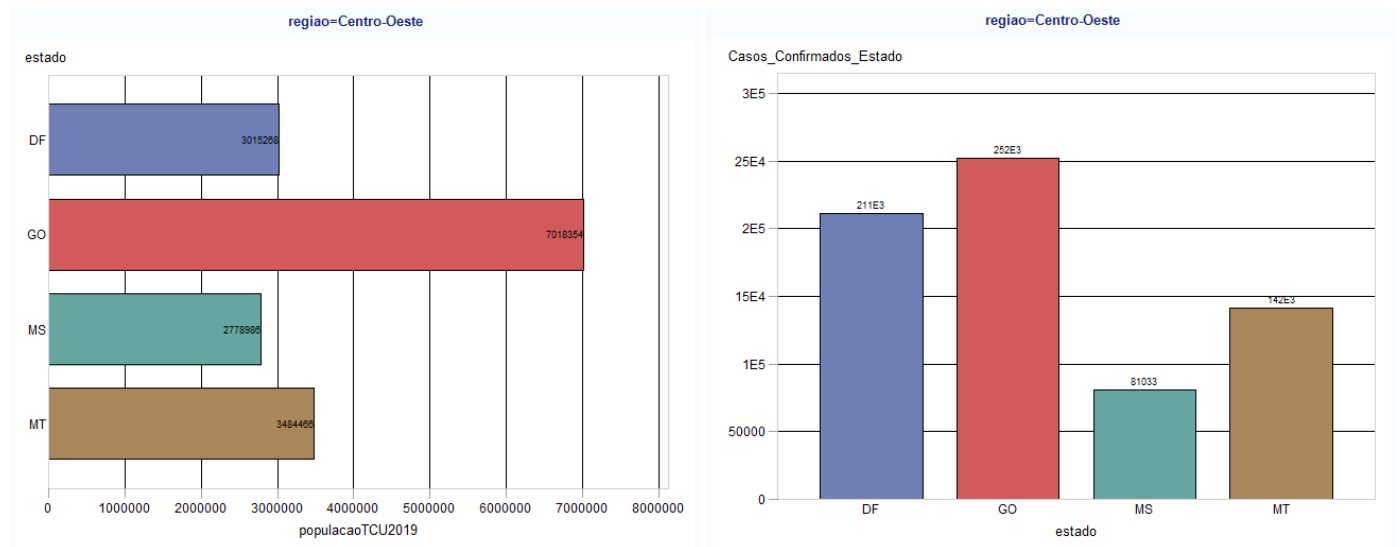


Figura 15: População e Casos Confirmados na Região Centro-Oeste do Brasil

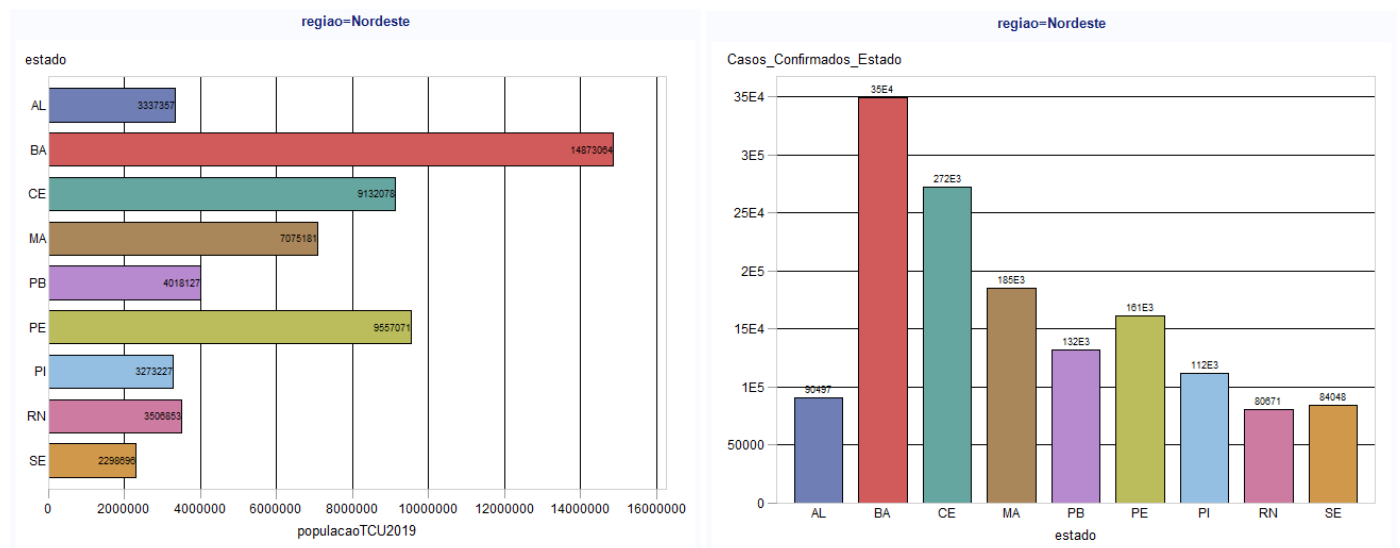


Figura 16: População e Casos Confirmados na Região Nordeste do Brasil

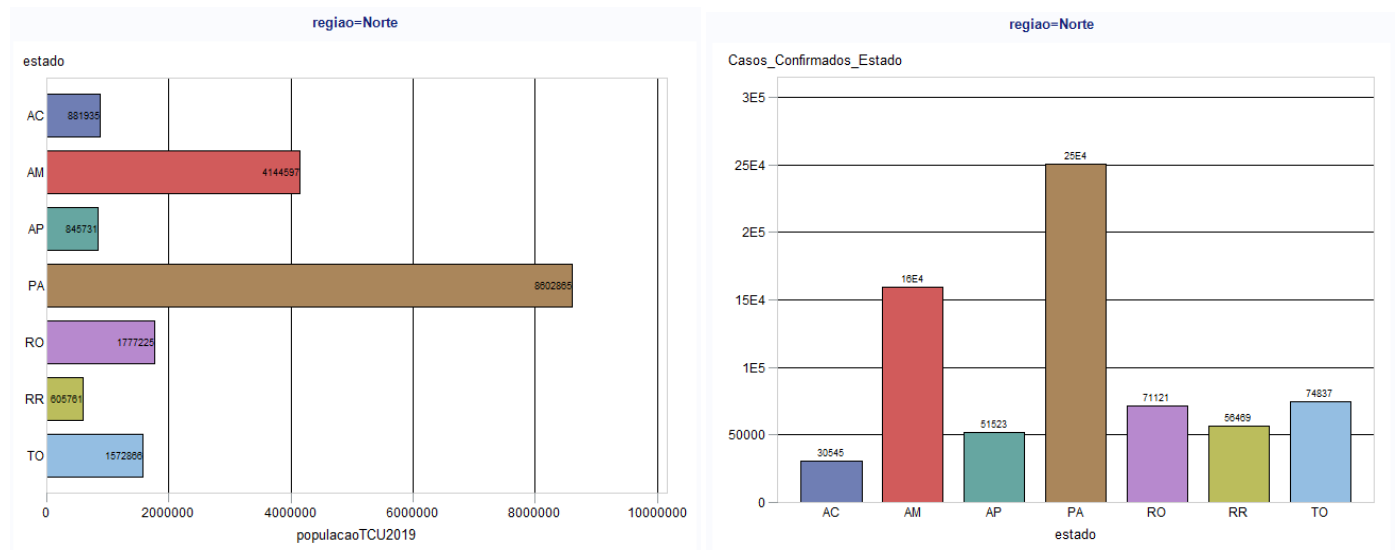


Figura 17: População e Casos Confirmados na Região Norte do Brasil

O estado menos afetado na região foi o *Acre (AC)* com uma população de 881.935 mil habitantes e 30.545 mil casos confirmados pelo COVID-19, como mostrado na *figura 17*.

Este resultado já altera um pouco a lógica de casos confirmados e população, pois vemos que o estado do *Acre (AC)* não é o menor em tamanho populacional, mas foi o menos afetado até a data de 28/10/2020, última do registro da base de dados analisada.

Na **Região Sudeste** os estados de *São Paulo (SP)* e *Minas Gerais (MG)* são os mais populosos com 45.919.049 milhões e 21.168.791 milhões de habitantes e tiveram 1.103.582 milhões e 353.311 mil casos confirmados respectivamente, sendo os mais afetados da região. O estado menos afetado foi o *Espírito Santo (ES)* com 4.018.650 milhões de habitantes e 153.004 mil casos confirmados, como mostrado na *figura 18*.

Na **Região Sul** a diferença entre os três estados da região entre população e casos confirmados é pouca, mas podemos ver que o estado de *Santa Catarina (SC)* possui a menor população com 7.164.788 milhões de habitantes e a menor população entre os estados da região mas com o maior volume de casos confirmados com 252.551 mil casos confirmados, sendo o mais afetado pelo COVID-19. O estado do *Paraná (PR)* é o menos afetado pelo COVID-19 até a data de 28/10/2020 com a maior população da região de 11.433.957 milhões de habitantes e 210.901 casos confirmados, como mostrado na *figura 19*.

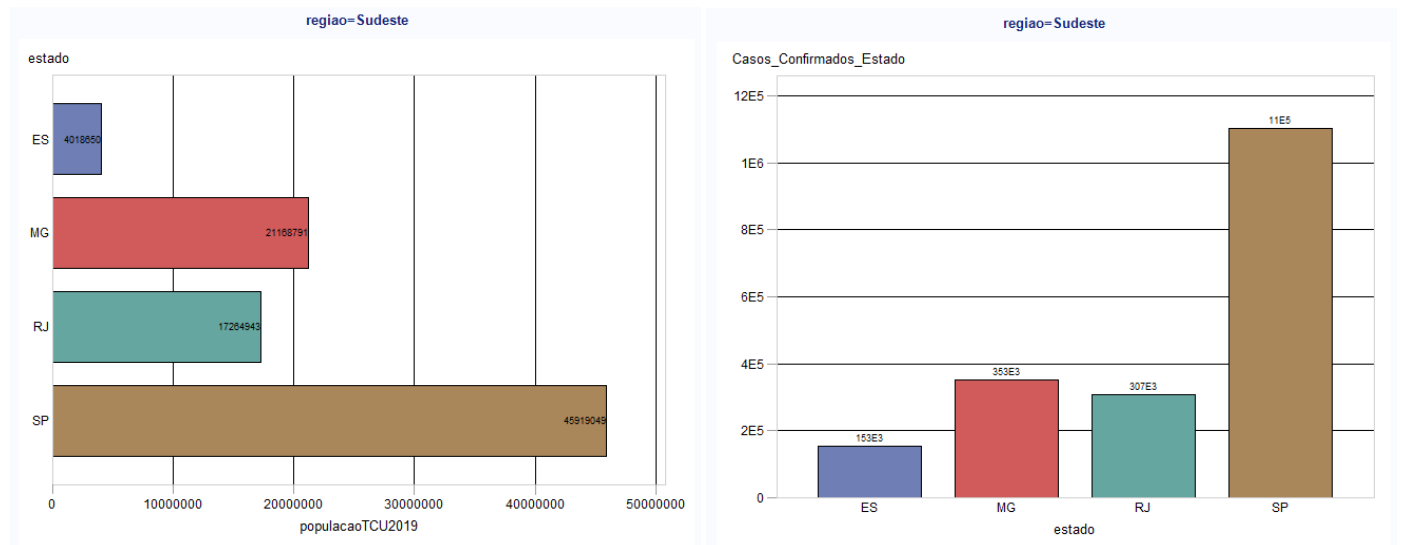


Figura 18: População e Casos Confirmados na Região Sudeste do Brasil

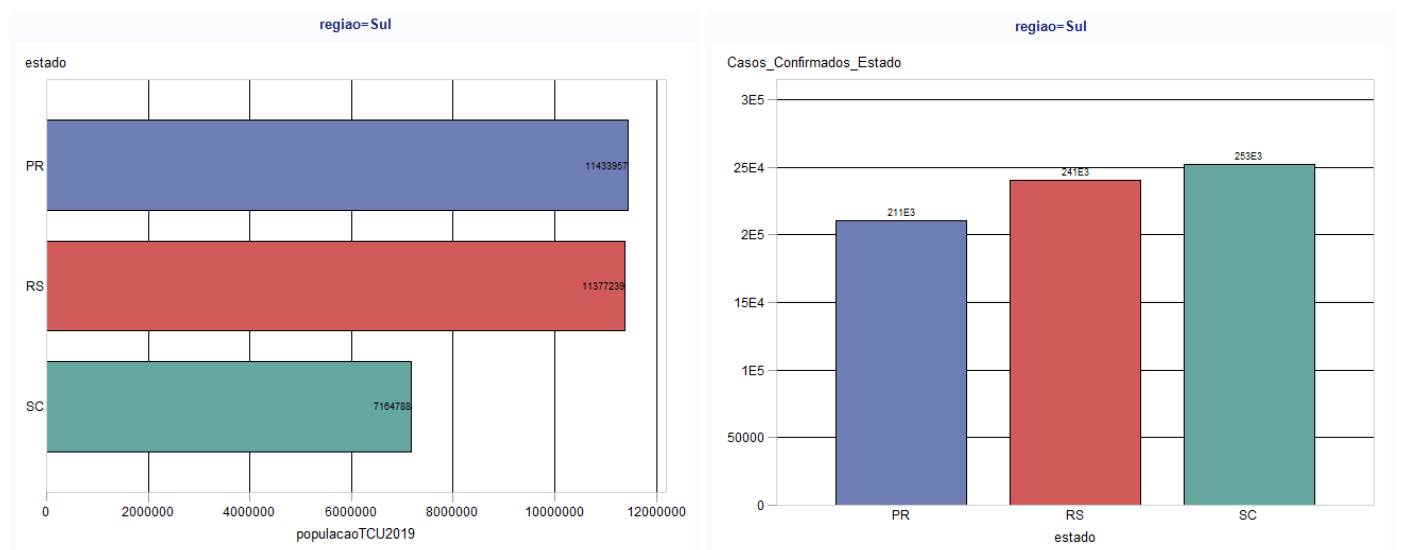


Figura 19: População e Casos Confirmados na Região Sul do Brasil

4.1.4 Análise dos Obitos Confirmados do COVID-19 por Região

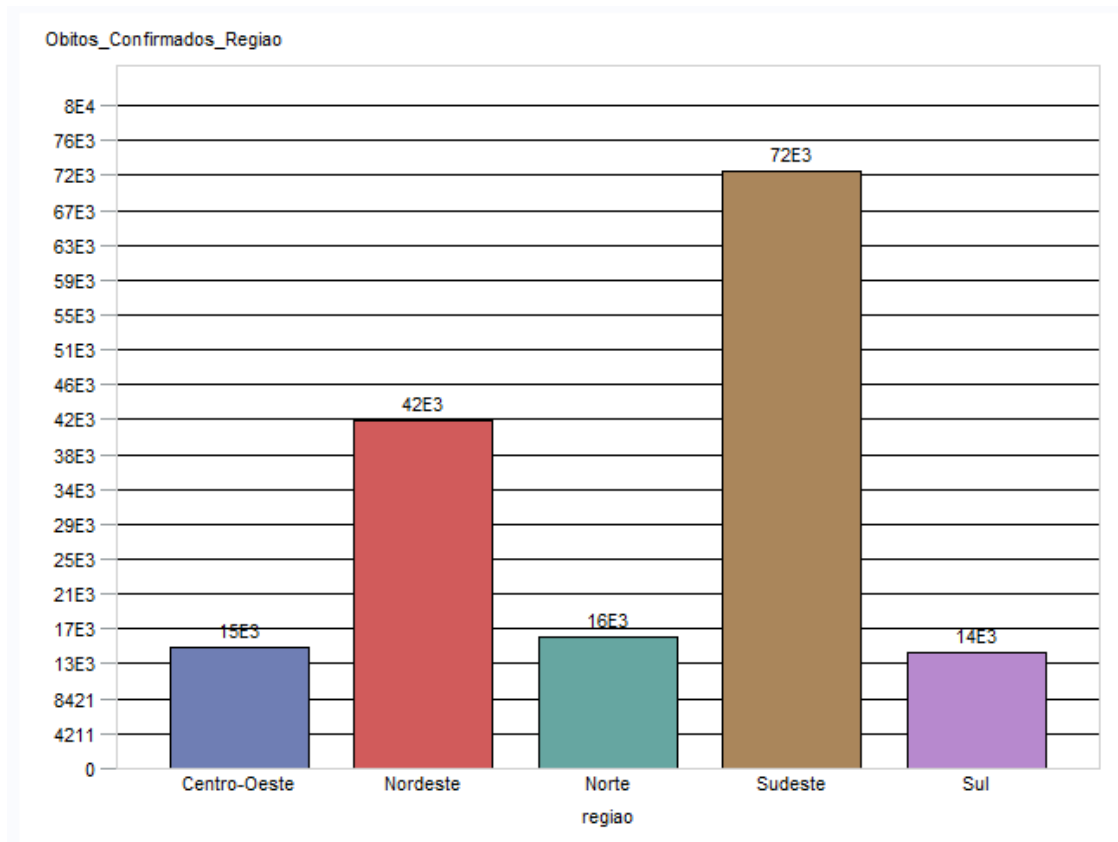


Figura 20: Obitos Confirmados por Região do Brasil

Como foi possível ver nas análises anteriores, vimos uma relação entre tamanho populacional e casos confirmados, quanto maiores, maior era os casos confirmados. Mas também vimos que em alguns estados isto não se seguia, e tiveram estados com populações menores e volume de casos confirmados maiores.

Por sua vez vemos os obitos por região na *figura 20* onde vemos claramente a **Região Sudeste** seguido da **Região Nordeste** com os maiores volumes de obitos confirmados pelo COVID-19 no Brasil, estes com 72.081 mil e 41.911 mil obitos confirmados, sendo estas as mais afetada do país.

Vemos também que a **Região Sul** do país foi a menos afetada com 13.914 mil obitos confirmados, como mostrado na *tabela da figura 12 e figura 20*.

4.2 Análise do Covid-19 no Estado do Paraná (PR)

Analisando os dados do COVID-19 no Estado do *Paraná (PR)* como proposta de estudo, encontra-se uma curiosidade e também um dado a se refletir, o estado do *Paraná (PR)* é o maior estado da *Região Sul do Brasil*, tendo a maior população da região, mas obteve o menor número de casos confirmados da região.

Para esta análise será utilizado a amostra dos dados do COVID-19, sendo a base de dados COVID19_SERIE.sas, está abordará as informações que precisasse para obter informação do COVID-19 no estado, por se tratar de uma amostra, não representa os valores totais das taxas para o estado, ou seja, não são os dados da **População (conjunto de todos os elementos ou resultados sob investigação), mas uma Amostra (qualquer subconjunto da população)** (MORETTIN; BUSSAB, 2017), neste caso os dados obtidos na base foram coletados até a data de 28/10/2020.

4.2.1 Análise de Casos Novos Confirmados por Data

Vê-se na *figura 21* os casos confirmados do COVID-19 no estado do *Paraná (PR)*, interessante que nos mostra um alto número de casos confirmados entre 29/06/2020 a 27/09/2020, consequência do período de pico de casos da pandemia do novo Coronavírus.

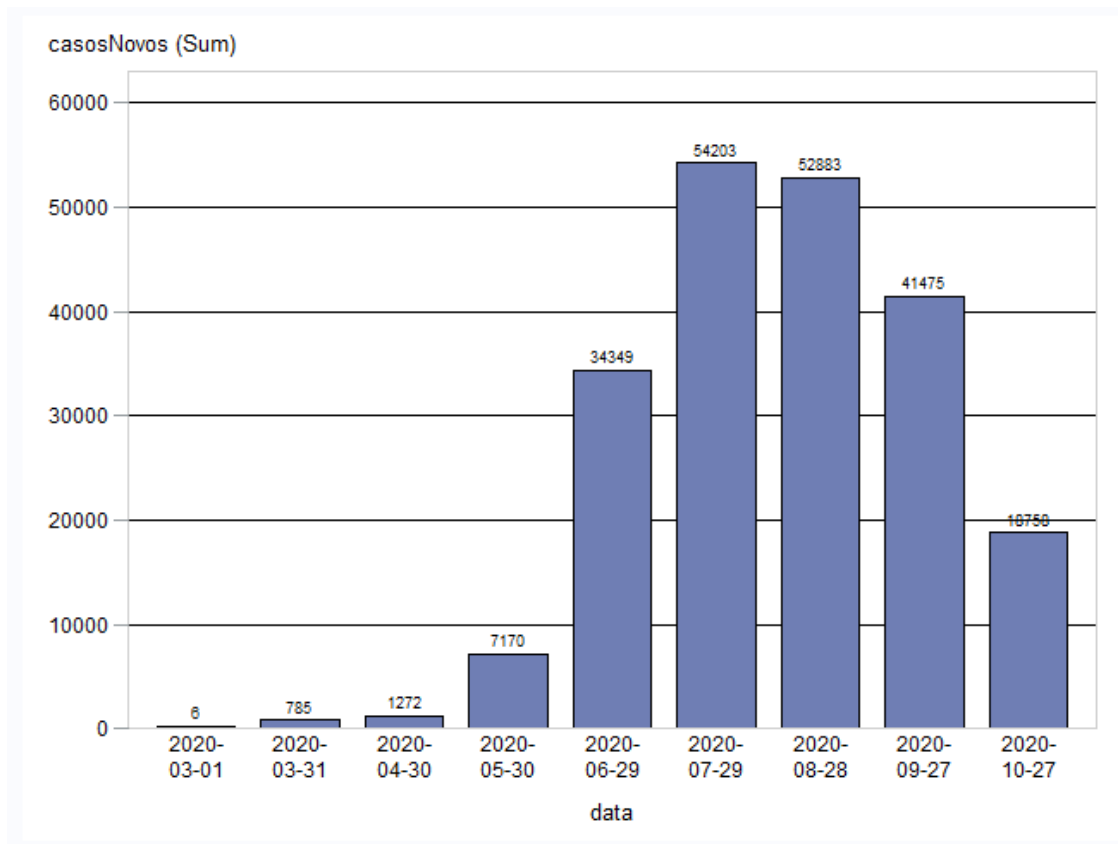


Figura 21: Casos Novos Confirmados por data no Estado do Paraná (PR)

4.2.2 Análise de Óbitos Novos Confirmados por Data

No estado do *Paraná (PR)* vemos o começo do aumento de casos entre 30/04/2020 e 30/05/2020 demonstrado na *figura 21*.

Em 29/07/2020, ou seja, no mês de Julho o estado teve 54.203 mil casos confirmados o pico da pandemia, já

em 28/08/2020, ou seja, mês de agosto teve uma pequena queda 52.883 mil casos confirmados, uma queda de 1.320 mil casos confirmados.

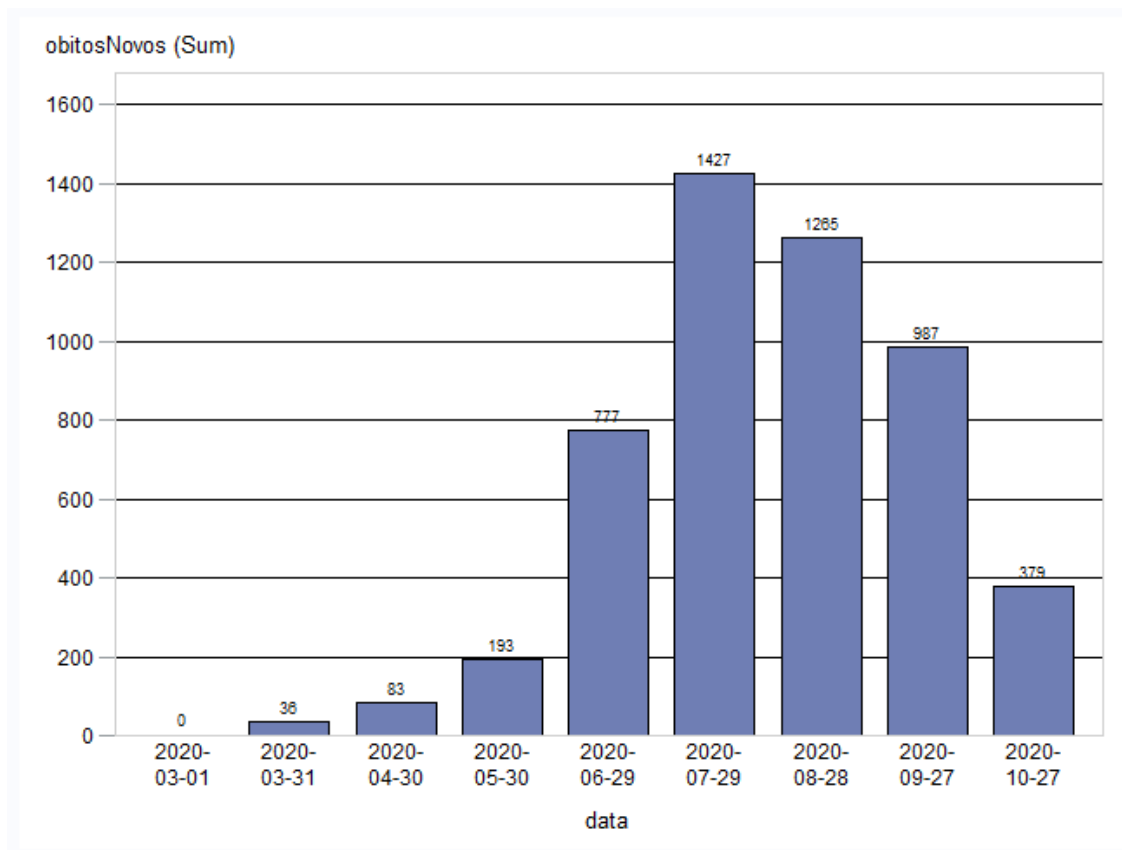


Figura 22: Óbitos Novos Confirmados por data no Estado do Paraná (PR)

Pico de obitos ocorreu em meados de 29/07/2020 no estado, crescimento dos obitos vem desde o mês de Maio como mostrado na *figura 22*, ou seja, meados de 30/05/2020 em diante. O estado obteve uma maior queda no número de obitos pelo COVID-19 entre 27/09/2020 e 27/10/2020 com queda de cerca de 608 óbitos a menos registrado pela base de dados, até 28/10/2020.

Uma ótima notícia desde que o maior número de mortes no *Paraná (PR)* foi de 1.427 mil óbitos confirmados pelo novo Coronavírus. A queda no número de mortes começou a se desenvolver entre 29/07/2020 e 28/08/2020.

4.2.3 Análise de Óbitos Novos e Casos Novos Confirmados por Semana Epidemiológica

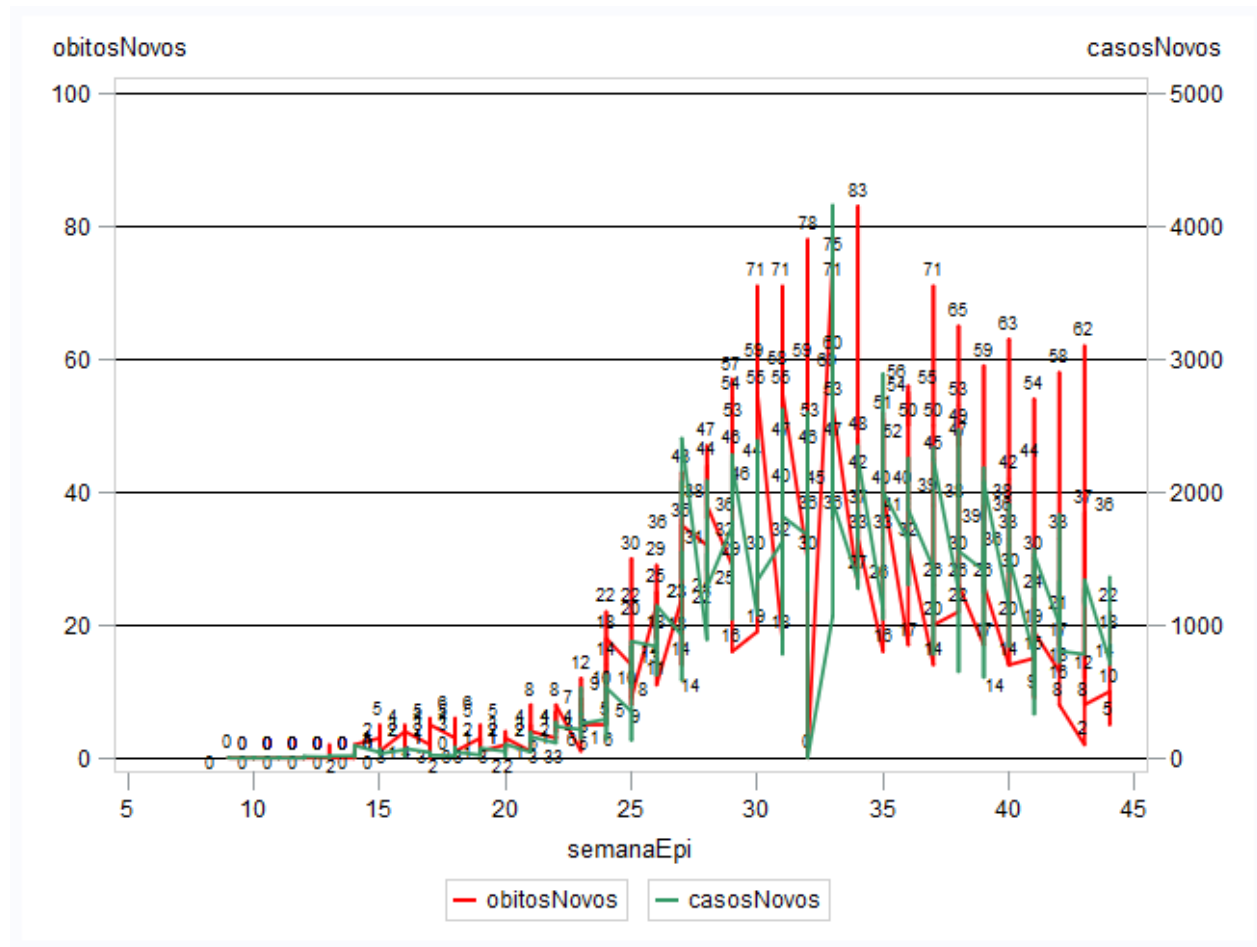


Figura 23: Óbitos Novos e Casos Novos Confirmados por Semana Epidemiológica no Estado do Paraná (PR)

Semana Epidemiológica entre 25 e 40 foram as que mais teve impacto em número de casos novos e óbitos novos por COVID-19 no Estado do *Paraná (PR)*, devido ao aumento pode-se perceber no Gráfico de linhas da *figura 23* que entre as semanas 30 e 35 houve o pico de casos novos e óbitos novos, mas também vemos nestas mesmas semanas que teve uma queda bruta em casos novos e óbitos novos, chegando a 0 casos novos e 0 óbitos novos.

O interessante deste gráfico de linhas que é possível ver claramente que o número de casos novos e óbitos novos começaram a ocorrer entre a semana 10 e 15, tendo números crescente destes a partir da semana 15 em diante.

Fazendo uma ligação na análise dos gráficos de barras das *figuras 21 e 22*, e o gráfico de linhas da *figura 23*, podemos perceber que entre as semanas de maiores números de casos novos e óbitos novos como também a maior queda nestes, dentro do período de pico da pandemia do novo Coronavírus no estado do *Paraná (PR)* ocorreu justamente entre os meses de Julho e Agosto, semanas entre 30 e 35, chegando o níveis 0 de casos novos e 0 de óbitos novos, como por sua vez a 75 de casos novos e 83 de óbitos novos pelo COVID-19.

5 Técnicas aplicadas para o desenvolvimento do trabalho

Utilizando a Estatística descritiva através do Software SAS Enterprise Guide 7.1, neste trabalho pode ser realizado com cálculos de várias métricas e gráficos desenvolvidos para geração de informações contidas nos dados brutos (rawdata) assim, foi possível gerar informação (valor) com os dados.

Utilizei para este trabalho uma linha de demonstração e calculo de métricas por região do Brasil, mostrando desde a população total, como também população por região do país. Casos Novos Confirmados por região e um comparativo de população por casos novos confirmados. Também utilizei informações dos dados de Óbitos Novos Confirmados e foi realizado breve um comparativo entre casos novos e óbitos novos confirmados.

Ao decorrer do trabalho utilizei gráficos de barras e de linhas que demonstram melhor a distribuição de casos, óbitos e demais variáveis por se tratar de dados como uma série temporal, ou seja, os dados ocorrer e são registrados ao decorrer de um tempo, neste casos por data (dia/mês/ano) e por semana (número da semana registrada).

Através da base de dados e está análise no Software SAS Enterprise Guide 7.1, foi possível realizar este Dashboard no Power BI Desktop.

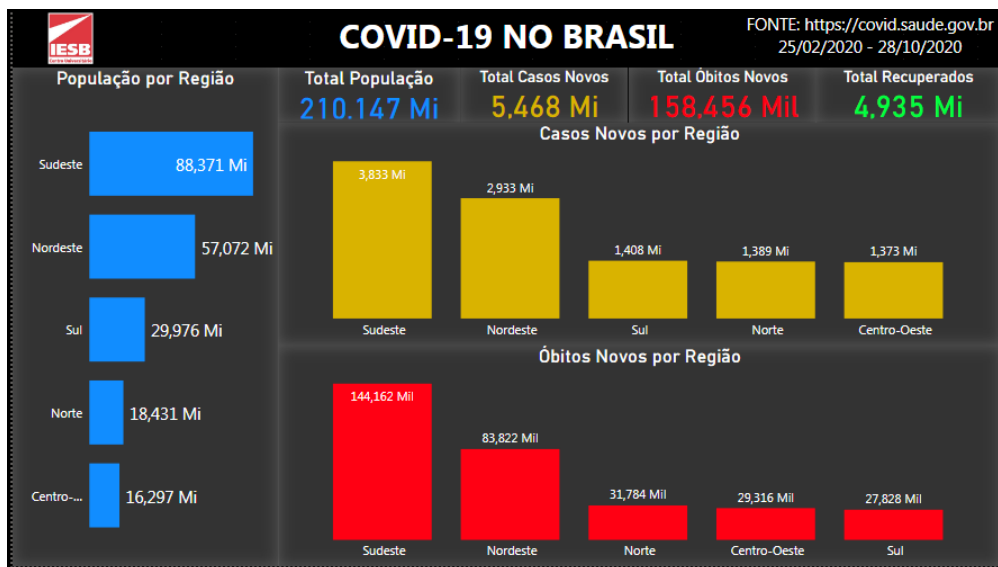


Figura 24: Dashboard Covid-19 por Região do Brasil

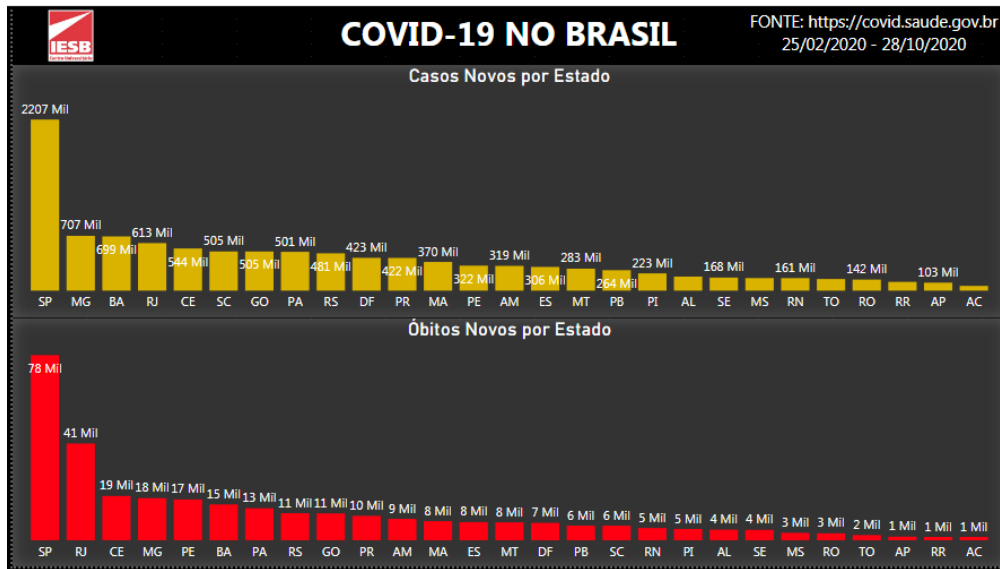


Figura 25: Dashboard Covid-19 por Estado do Brasil

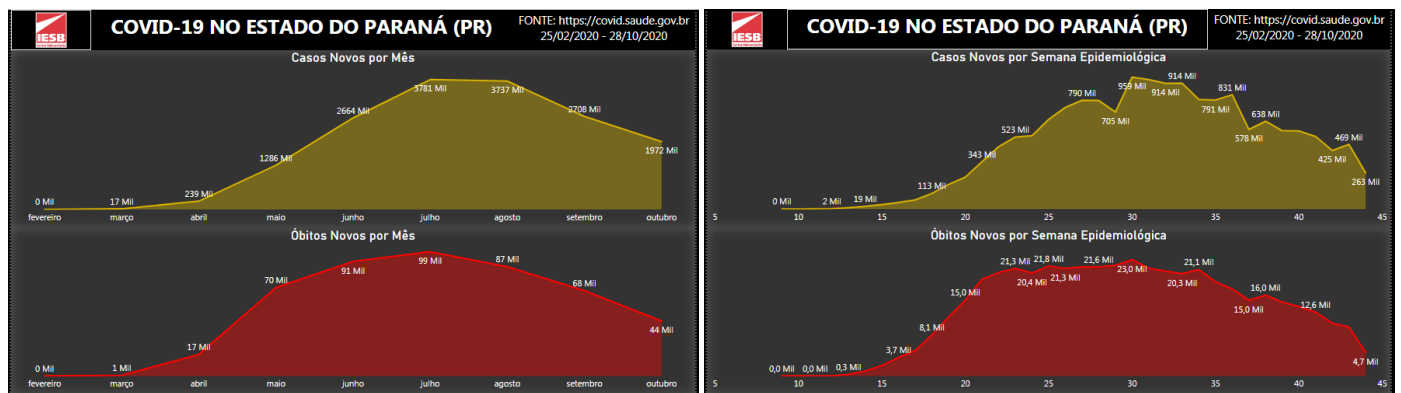


Figura 26: Dashboard Covid-19 por Mês e Semana Epidemiológica no Estado do Paraná (PR)

6 Conclusão

Através de vários cálculos e demonstrações gráficas por gráficos de barras e linhas como também tabelas demonstrativas, com isso foi possível descobrir um certo padrão nos dados da COVID-19 em que quanto maior a população maior o número de casos novos confirmados. Por se tratar de uma Pandemia, o Novo Coronavírus ainda é um mistério de como erradicar este de municípios, cidades, países, ou seja, do mundo em si, por sua vez a população como os governos foram descobrindo ao decorrer da Pandemia que determinadas medidas preventivas como: Isolamento Social, Uso de Máscaras, Uso de Álcool em Gel e Líquido, Limpeza frequente de mãos e partes em contato direto com saliva, e secreções do corpo, que são um transmissores do COVID-19.

Por sua vez, este novo Coronavírus foi observado por diversos países que quanto maior o número de pessoas (aglomeração) em determinados locais ou espaços públicos, a contaminação foi maior, e por isso muitos adotaram medida de isolamento social, como o Lockdown (Confinamento), em que a população não pode sair de suas casas a não ser se fosse casos extremos de necessidade alimentar ou de saúde. Com o padrão de contaminação que ocorre com aglomeração de pessoas, foi possível descobrir que quanto maior a população maior o número de casos, mas foi possível verificar também que em alguns estados, este padrão não se segue, como no estado do Paraná (PR) na região sul do país, em que a população do estado é a maior da região mas o número de casos foi o menor da região, podemos pensar que tal fato se dê por medidas restritivas de locomoção de pessoas, isolamento social, evitando ao máximo aglomeração de pessoas em espaços públicos. Vemos também medidas para o estado do Paraná (PR) em que verificamos sua população, casos novos confirmados, óbitos novos confirmados, como comparação entre os dois indicadores do COVID-19 no Brasil e no estado do Paraná (PR).

Conclui-se neste trabalho um entendimento que através de medidas de tendência central, entre estas estatísticas básicas e de resumo, com o auxílio de ferramentas como a utilizada o Software de Análise de Dados o SAS Enterprise Guide 7.1, foi possível analisar e encontrar medidas, padrões e distribuição de acontecimentos do Novo Coronavírus no Brasil e no Estado do Paraná (PR).

7 Referências Bibliográficas

COLE NUSSBAUMER KNAFLIE. Storytelling com dados: Um guia sobre visualização de dados para profissionais de negócios. Alta Books, 2018.

GOVERNO DO BRASIL. Governo do Brasil. 2020. Disponível em: <https://www.gov.br/pt-br>.

MORETTIN, P. A.; BUSSAB, W. O. Estatística Básica 9ª Edição. Saraiva Educação SA, 2017.

PARANÁ. Governo do Estado do Paraná. 2020. Disponível em: <http://www.parana.pr.gov.br/>.

WIKIPÉDIA. Brasil. 2020. Disponível em: <https://pt.wikipedia.org/wiki/Brasil>.

WIKIPÉDIA. Paraná. 2020. Disponível em: <https://pt.wikipedia.org/wiki/Paran%C3%A1>.