



Mercado de telefonia móvel no Brasil

Análise de dados do mercado regulado brasileiro
de 2009 a 2023

Beatriz Pierri
Fernanda Castro
Graciele Rodrigues
Jane Adriana
Lilian Nakayama
Risla Miranda

Bootcamp Análise de Dados ENAP
Outubro 2023

SUMÁRIO

| | |
|------------------|---|
| Introdução | 3 |
|------------------|---|

PARTE I

| | |
|-------------|---|
| Dados | 5 |
|-------------|---|

| | |
|-----------------|---|
| Objetivos | 6 |
|-----------------|---|

| | |
|-------------------|---|
| Metodologia | 7 |
|-------------------|---|

PARTE II

| | |
|--------------------|----|
| Documentação | 12 |
|--------------------|----|

PARTE III

| | |
|-------------------------------------|----|
| Destaques da análise de dados | 17 |
|-------------------------------------|----|

| | |
|----------------|----|
| Produtos | 26 |
|----------------|----|

INTRODUÇÃO

Este relatório tem como objetivo organizar e concentrar a documentação e destaques das análises de dados realizadas a partir de dados brutos da Agência Nacional de Telecomunicações - Anatel, referentes ao mercado de telefonia móvel no período de 2009 a 2022.

A análise de dados sobre o tema foi resultado de um projeto escolhido no Bootcamp de Análise de Dados - turma exclusiva para mulheres - da Escola Nacional de Administração Pública (ENAP).

O relatório está dividido em três partes: a primeira discute a estrutura e a documentação do projeto; a segunda traz destaques encontrados na análise dos dados; e a terceira resume os produtos - além deste relatório.

PARTE I

DADOS

Os dados utilizados no projeto são os dados brutos disponibilizados pela Anatel através do painel de acessos de telefonia móvel¹. Os dados brutos contêm dados de 2005 a 2023. Sendo que o período de 2005 a 2008 não possui os dados agregados por UF, portanto, escolhemos não utilizar nas análises os anos de 2005 a 2008.

As planilhas dos dados brutos disponibilizadas pela Anatel trazem a quantidade dos acessos por empresa, por porte do cliente, por modalidade e tecnologia de acesso, para todos os meses de determinado ano e/ou semestre. Acessos são os chips ativos naquele mês de referência. É importante mencionar que nem todas as empresas têm dezembro do ano preenchido e que o valor vai se acumulando a cada mês, ou seja, se em janeiro aquele município ou estado tem x acessos, em fevereiro - supondo que tenham esses meses preenchidos - terá $x+y$, sendo y a quantidade de acessos de fevereiro, podendo variar para mais ou para menos (isto é, a empresa pode ter tido um acréscimo ou decréscimo da sua base de clientes no mês seguinte).

As diferenças entre as planilhas dos dados de 2009 a 2018 e de 2019 a 2023 se resumem a: i. divisão entre semestres e ano: as planilhas de 2019 a 2023 estão segregadas em dois semestres para cada ano, além de organizarem os dados mensais em duas planilhas diferentes, sendo uma em colunas e outra em linhas, e ii. ausência das colunas referentes ao código nacional, nome do município, modalidade de cobrança e tipo de produto para dados de 2009 a 2018.

Como as diferenças entre os períodos 2009-2018 e 2019-2023 não impactam as análises propostas, nós optamos pela série temporal de 2009 a

¹ Nós utilizamos os dados brutos atualizados até 27/09/2023.

2023, utilizando-se dos acessos por UF e por código nacional e das tecnologias por ano.

OBJETIVOS

1. Analisar, no período de 2009 a 2023, a concentração de mercado de telefonia móvel por UF e por DDD a partir da parcela de mercado por grupo econômico e do indicador de HHI (Índice Herfindahl-Hirschman).
2. Analisar, no período de 2009 a 2023, a evolução da tecnologia de geração no Brasil.

METODOLOGIA

Após a conversa inicial sobre o projeto, os objetivos foram acordados e nivelados, levando em conta o tempo e as tecnologias aprendidas durante o Bootcamp. Também acordamos quais produtos queríamos no final das análises, tanto para documentar, quanto para publicizar as informações.

Para a análise e visualização dos dados usamos Python e as bibliotecas pandas, numpy, plotly, matplotlib, folium e seaborn. O PowerBI foi utilizado para construir uma visualização dos dados de concentração do mercado por UF. O Streamlit foi utilizado para construir uma página web para concentrar os produtos advindos das análises.

Os dados de 2009 a 2018 estavam em uma única planilha, os demais anos tinham os dados separados por semestre. Além dessa separação por semestre, as planilhas também eram separadas de acordo com a estruturação da quantidade de acesso por mês: a Anatel colocou um sufixo nas planilhas (_colunas) identificando que estas planilhas separavam o mês por colunas, nas outras a separação é feita por linhas.

Para dinamizar e não incorrer em problemas de memória - as planilhas com a estruturação dos meses por linhas tinham uma quantidade significativa de linhas, chegando, em uma delas a 13 milhões, o que poderia sobrecarregar os nossos computadores - decidimos utilizar as planilhas com o sufixo _colunas.

É possível observar a seguir como as planilhas com sufixo _colunas são estruturadas:

df_colunas.head(10)

| | CNPJ | Código Nacional | Município | UF | Modalidade de Cobrança | Tecnologia | Tecnologia Geração | Empresa | Porte da Prestadora | Tipo de Pessoa | Tipo de Produto | Código IBGE Município | Grupo Econômico | 2019-01 | 2019-02 | 2019-03 | 2019-04 | 2019-05 | 2019-06 |
|---|---------------|-----------------|-------------|----|------------------------|------------|--------------------|------------------|---------------------|-----------------|-----------------|-----------------------|-----------------|---------|---------|---------|---------|---------|---------|
| 0 | 1778972000174 | 11 | Santo André | SP | Pós-pago | LTE | 4G | America Net S.a. | Pequeno Porte | Pessoa Física | VOZ+DADOS | 3547809 | OUTROS | NaN | NaN | NaN | NaN | NaN | 1.0 |
| 1 | 1778972000174 | 11 | Santo André | SP | Pós-pago | LTE | 4G | America Net S.a. | Pequeno Porte | Pessoa Jurídica | VOZ+DADOS | 3547809 | OUTROS | NaN | NaN | 102.0 | 3.0 | 33.0 | NaN |
| 2 | 1778972000174 | 11 | Arujá | SP | Pós-pago | LTE | 4G | America Net S.a. | Pequeno Porte | Pessoa Física | VOZ+DADOS | 3503901 | OUTROS | 10.0 | 300.0 | 779.0 | 1187.0 | 1115.0 | 1110.0 |
| 3 | 1778972000174 | 11 | Arujá | SP | Pós-pago | LTE | 4G | America Net S.a. | Pequeno Porte | Pessoa Jurídica | VOZ+DADOS | 3503901 | OUTROS | 1.0 | 2.0 | 1.0 | 2.0 | 6.0 | 4002.0 |
| 4 | 1778972000174 | 11 | Barueri | SP | Pós-pago | LTE | 4G | America Net S.a. | Pequeno Porte | Pessoa Física | VOZ+DADOS | 3505708 | OUTROS | 1.0 | NaN | NaN | NaN | 1001.0 | NaN |
| 5 | 1778972000174 | 11 | Barueri | SP | Pós-pago | LTE | 4G | America Net S.a. | Pequeno Porte | Pessoa Jurídica | VOZ+DADOS | 3505708 | OUTROS | 4.0 | 10.0 | 22.0 | 27.0 | 32.0 | 3005.0 |
| 6 | 1778972000174 | 11 | Cajamar | SP | Pós-pago | LTE | 4G | America Net S.a. | Pequeno Porte | Pessoa Física | VOZ+DADOS | 3509205 | OUTROS | NaN | NaN | 10.0 | NaN | NaN | NaN |
| 7 | 1778972000174 | 11 | Cotia | SP | Pós-pago | LTE | 4G | America Net S.a. | Pequeno Porte | Pessoa Física | VOZ+DADOS | 3513009 | OUTROS | 3.0 | 2.0 | 2.0 | 1001.0 | 4.0 | NaN |
| 8 | 1778972000174 | 11 | Cotia | SP | Pós-pago | LTE | 4G | America Net S.a. | Pequeno Porte | Pessoa Jurídica | VOZ+DADOS | 3513009 | OUTROS | 1.0 | NaN | NaN | 6.0 | 31.0 | 32.0 |
| 9 | 1778972000174 | 11 | Diadema | SP | Pós-pago | LTE | 4G | America Net S.a. | Pequeno Porte | Pessoa Física | VOZ+DADOS | 3513801 | OUTROS | NaN | NaN | NaN | NaN | NaN | 2.0 |

Figura 1 - Dataframe do ano de 2019

Dividimos entre nós para que cada integrante ficasse com um ano para proceder o tratamento e às análises dos dados - apenas a Risla ficou com a planilha com o compilado dos anos de 2009 a 2018. As tabelas de 2009 a 2018 com os acessos por coluna não tinham os dados de janeiro de 2009. Esses dados estavam em outra tabela, de 2005 a 2008. Como as tabelas tinham tamanhos iguais, pegamos apenas a coluna de janeiro de 2009 e adicionamos à tabela de 2009.

Para a tabela de tecnologia, não foi possível fazer o passo de transpor janeiro de 2009 conforme descrito no parágrafo anterior. Após análise do arquivo de tecnologia de 2005 a jan/2009, observou-se que as duas planilhas não possuem o mesmo tamanho e nem as mesmas colunas, portanto, para dinamizar a análise e evitar ruído nos dados, decidimos não acrescentar a coluna de jan/2009 para a análise de tecnologia. Assim, as análises de tecnologia foram realizadas a partir de fev/2009 até dez/2018.

1. Concentração de mercado

Para a análise da concentração de mercado de telefonia móvel, dividimos a análise em dois blocos: um por UF e outro por DDD. A parcela de mercado foi calculada a partir da média da soma de acessos por grupo econômico.

Como visto, os dados brutos contidos nas planilhas disponibilizadas pela Anatel continham o número de acessos (chips operacionais) vigentes em

todos os meses do ano, conforme demais especificações por tipo de acesso (tipo de cliente, de tecnologia, de tecnologia de geração, de modalidade de acesso, etc.). Escolhemos utilizar a média de acessos anual para afastar distorções que poderiam aparecer, como, por exemplo, a empresa não ter preenchido o mês de dezembro - e, como consequência, aparecer como zero acessos no ano sob análise.

Escolhemos desagregar os dados em UF e em DDD para termos uma análise interessante sobre regionalização da concentração de mercado, uma vez que as análises acadêmicas costumam partir dos acessos em nível nacional.

Com a média de acessos por ano e por grupo econômico, conseguimos calcular a parcela de mercado de cada grupo econômico (*market share*). *Market share* é a participação de mercado (em %) que cada grupo econômico apresenta em determinado delimitador geográfico: no caso do presente trabalho, utilizamos duas delimitações geográficas de referência: i) Unidades da Federação (UF) e ii) áreas abrangidas pelo mesmo código DDD de telefonia móvel.

E, a partir da parcela de mercado, calculamos o indicador de HHI. Segundo o Guia de Análise de Atos de Concentração Horizontal do Cade ("Guia H")², "o Índice Herfindahl-Hirschman (HHI) é uma metodologia de medição do grau de concentração dos mercados. O HHI é calculado com base no somatório do quadrado das participações de mercado de todas as empresas de um dado mercado. O HHI pode chegar até 10.000 pontos, valor no qual há um monopólio, ou seja, em que uma única empresa possua 100% do mercado".

Assim temos a participação de mercado e o HHI por UF e por DDD.

² Disponível em <https://cdn.cade.gov.br/Portal/centrais-de-conteudo/publicacoes/guias-do-cade/guia-para-analise-de-atos-de-concentracao-horizontal.pdf>. Acesso em 31/10/2023. p. 24.

Como é possível observar na figura 1, há NaN³ distribuídos nas planilhas, principalmente nas colunas de mês que indicam a quantidade de acessos. Rodamos as análises de parcelas de mercado com os dados NaN presentes e também zerando os NaN utilizando `fillna(0)`.

Percebemos que as análises de concentração de mercado com os NaNs superestimavam as empresas que não tinham tantas entradas de dados de acesso. Como fizemos o cálculo da média de acessos do ano, caso uma empresa tenha enviado apenas o valor de, por exemplo, três meses, a média não levaria em conta doze meses do ano e sim três, o que superestima a média dessa empresa.

Diante disso, após os testes metodológicos, decidimos tratar o NaN como zero através do método `fillna(0)`. Assim temos uma escolha metodológica que leva em conta a variável ano como doze meses e trata igualmente o valor indefinido.

Ainda, a média dos acessos do ano foi calculada a partir da soma de acessos naquele ano de cada linha da tabela. Para o cálculo da parcela de mercado, que por convenção chamamos de *market share*, pegamos a média de acessos por grupo econômico, tanto por UF, quanto por DDD.

Com o valor do *market share*, calculamos o HHI elevando os valores do *market share* de cada empresa (agrupando por UF e depois por DDD) ao quadrado e depois somando os resultados. Para analisar o grau de concentração de mercado a partir do HHI utilizamos os critérios da *Federal Trade Commission* (FTC): abaixo de 1500 temos um mercado pouco concentrado; entre 1500 e 2500 temos uma concentração moderada de mercado; e acima de 2500 é um mercado altamente concentrado.

³ NaN significa Not a Number e remete a um valor indefinido ou irrepresentável.

2. Tecnologia

Há duas colunas referentes às tecnologias: tecnologia e tecnologia geração. Nos cadernos Jupyter utilizados para a análise dos dados, trazemos conceitos e entendimentos sobre as tecnologias vigentes em cada ano. Fizemos uma análise inicial das tecnologias, entretanto, os tipos de tecnologia exigem um conhecimento mais aprofundado e, em nível comercial, o que é oferecido é a tecnologia geração, a saber: M2M⁴, 1G, 2G, 3G, 4G e 5G.

A partir das tecnologias geração temos um retrato da qualidade de acesso à internet através do celular. Portanto, escolhemos focar nas tecnologias de geração para analisar a distribuição da qualidade de telefonia no Brasil.

As análises levaram em conta a distribuição das tecnologias de geração por UF e também por empresa/grupo econômico. Além disso, especificamente sobre a 5G - a tecnologia de ponta hoje -, analisamos a distribuição de 5G por estado nos anos em que ela começou a ser comercializada, e fizemos um ranking dos Top 10 municípios brasileiros com maior representatividade nacional (em número de acessos) de tecnologia 5G em 2023⁵ além de mapa com o percentual de acesso 5g em cada UF.

⁴ M2M refere-se a comunicação entre máquinas. O M2M padrão, geralmente, refere-se às máquinas de cartão.

⁵ Conforme dados disponibilizados pela Anatel até setembro de 2023.

PARTE II

DOCUMENTAÇÃO

É possível acessar os cadernos Jupyter e demais códigos utilizados no tratamento, análise, organização e visualização dos dados no repositório no GitHub dos projetos⁶. Para evitar ruídos e possíveis perdas durante o projeto, decidimos criar um repositório específico para trabalhar durante projeto⁷. Os cadernos Jupyter trazem orientações de como rodar o código, além de comentários documentando as variáveis e funções criadas e as análises realizadas.

1. Tratamento dos dados

Para facilitar a análise dos dados, as integrantes que ficaram com os anos de 2019 a 2023 utilizaram o merge para ter em uma única planilha os dados do primeiro e do segundo semestres. Utilizando o pandas, lemos os arquivos e salvamos em dataframes.

Para verificar a quantidade de NaNs, fizemos uma função para checar a proporção de NaNs em relação ao tamanho da tabela. Os NaNs foram checados também por maiores empresas - o que vimos que não eram tão comuns. Conforme decisão metodológica, utilizamos o `fillna(0)` para transformar os NaNs em zeros.

Fizemos também função e checagem de grafia dos grupos econômicos. Em apenas duas tabelas tivemos divergência com a palavra Telefônica e consertamos com uma função:

```
def corrigir_nome(nome):  
    if nome == 'TELEFÔNICA':  
        return 'TELEFONICA'  
    else:
```

⁶ Por meio deste link é possível ter acesso aos outros projetos realizados no âmbito deste Bootcamp.

⁷ Para acessar o repositório utilizado durante o processo do projeto [clique aqui](#).

```
return nome
```

2. Padronização de nomes

Para facilitar nossa comunicação e a leitura dos códigos, assim como padronizar o entendimento na equipe, padronizamos os nomes das variáveis dos dataframes (df_ano_uf, df_ano_ddd, df_ano_tec), das novas colunas (media_ano, market_share, hhi_empresa, HHI) com os cálculos de concentração, dos notebooks (acessos_ano, tecnologia_ano) e também dos arquivos csv (cm_ano_uf, cm_ano_ddd, hhi_ano_uf, hhi_ano_ddd, tec_ano_5g, tec_ano_geracao, tec_ano_tipo) exportados para visualização dos dados.

3. Análises dos dados

3.1. Concentração de mercado

Para fazermos os cálculos referentes à concentração de mercado, adicionamos uma coluna chamada media_ano em cada dataframe: `round(df_ano[lista_colunas].mean(axis=1), 2)`. A variável lista_colunas tem uma lista com as colunas de cada mês.

Fizemos dois dataframes dos grupos econômicos para cada agrupamento - UF e DDD. Para isso, utilizamos o método `groupby`:

- Para UF: `round(pd.DataFrame(df_ano.groupby(['UF', 'Grupo Econômico']))['media_ano'].sum()), 2)`
- Para DDD: `round(pd.DataFrame(df_ano.groupby(['Código Nacional', 'Grupo Econômico']))['media_ano'].sum()), 2)`

Para o cálculo do *market share*, criamos uma coluna chamada *market_share*, pegando a *media_ano* e dividindo pela soma da *media_ano* do agrupamento por UF ou DDD.

Para o cálculo do HHI, primeiro calculamos o valor do *market share* de cada empresa ao quadrado, adicionando este valor a uma coluna chamada

hhi_empresa. Como decidimos não colocar em porcentagem o valor do market_share para facilitar a visualização no PowerBI, o cálculo do hhi_empresa foi multiplicado por 10.000 para termos a base correta para interpretação do HHI. Adicionamos uma coluna chamada HHI ao dataframe UF e DDD, somando os valores hhi_empresa agrupado por UF ou DDD, de acordo com o dataframe em análise.

Criamos uma função chamada analise_hhi que itera os valores de HHI por estado e retorna qual critério está o valor:

```
def analise_hhi(valor_hhi):  
    if valor_hhi <= 1500:  
        return "pouco concentrado"  
    if valor_hhi > 1500 and valor_hhi <= 2500:  
        return "moderado"  
    else:  
        return "altamente concentrado"
```

Essa função foi aplicada aos dataframes UF e DDD em uma nova coluna chamada analise_mercado_hhi.

3.2. Tecnologia

Para a análise das tecnologias, utilizamos o arquivo exportado csv que já possuía as colunas media_ano e também já estava com os dois semestres. Entretanto, as tabelas de 2009 a 2018 tratadas para a análise de mercado não continham a coluna sobre tecnologia, assim, o roteiro de tratamento foi utilizado para a tabela contendo as tecnologias de 2009 a 2018⁸.

Como os arquivos csv lidos para dataframe estavam com vírgula nos valores de média do ano, fizemos a conversão para tipo número float:
`df_tec['media_ano'] = df_tec['media_ano'].str.replace(',', '
'.', regex=True).astype(float).`

⁸ Lembrando que, conforme discutido na metodologia (p. 8), o ano de 2009 na tabela começa a partir de fevereiro.

Fizemos uma análise exploratória para entender as tecnologias e as tecnologias de geração nas tabelas, calculando o percentual de cada uma para o ano. Para isso, adicionamos uma coluna chamada percentual aos dataframes que agruparam a tecnologia por tipo (usando a coluna Tecnologia) e a tecnologia por geração (usando a coluna Tecnologia Geração).

A partir de 2021 pudemos fazer uma análise específica de 5G a partir da média de acessos por UF, criando um dataframe específico com uma coluna chamada percentual para listar a porcentagem de 5G por estado.

4. Organização dos dados

Como transformamos os agrupamentos em dataframes, fizemos uma exportação em csv tanto para ter a tabela com os valores calculados, mas também para facilitar a visualização dos dados no PowerBI.

Para tanto, padronizamos os nomes dos arquivos e, como na análise de dados de concentração de mercado não adicionamos a coluna de ano, fizemos um script em Python (cod_add_coluna.ipynb) que lê a pasta com os arquivos, abre o csv e cria uma coluna chamada ano.

Também fizemos um script em Python (merge df tecnologia.ipynb) para concatenar todos os arquivos referentes às análises de tecnologia, facilitando para a visualização da evolução tecnológica ano a ano.

5. Visualização dos dados

Para a visualização de dados, além da ferramenta PowerBI, utilizamos o matplotlib, seaborn, folium e plotly para a evolução das tecnologias de 2009 a 2023. Para a evolução das tecnologias de geração utilizamos um gráfico de linhas múltiplas a partir da porcentagem de acessos das tecnologias de geração por ano.

PARTE III

DESTAQUES DAS ANÁLISES DE DADOS

1. Concentração de mercado

Conforme estudos anteriores, confirmamos que a concentração do mercado de telefonia é alta e utilizar os dados desagregados em estados foi interessante para perceber que, mesmo com estados com diferença entre o *market share* das empresas, ainda temos sempre, como dominantes no mercado de telefonia móvel, os mesmos quatro grupos econômicos - até a compra da Oi que foi finalizada no primeiro semestre de 2022: Telefônica (VIVO), Telecom Italia (TIM), TELECOM AMERICAS (CLARO) e Oi, conforme a figura a seguir:



| Grupo Econômico | Média de Acessos |
|--------------------------|----------------------|
| TELEFONICA (VIVO) | 1.127.088.513 |
| TELECOM AMERICAS (CLARO) | 942.441.635 |
| TELECOM ITALIA (TIM) | 894.828.115 |
| OI | 562.790.189 |
| OUTROS | 38.574.201 |
| NEXTEL | 18.509.353 |
| Total | 3.584.232.006 |

Figura 2 - Média de acessos por grupo econômico

Fonte: Painel BI Mercado Regulado de Telefonia Móvel

De 2015 para 2016 observamos uma queda da média de acessos, indicando que houve desativação de quantidade significativa de chips das operadoras. E, conforme o mapa em bolhas da Figura 3 adiante, no qual quanto maior a bolha, maior a média de acessos, a maior média de acessos durante todos os anos analisados se concentra na região sudeste, especificamente nos estados de São Paulo, Rio de Janeiro e Minas Gerais.

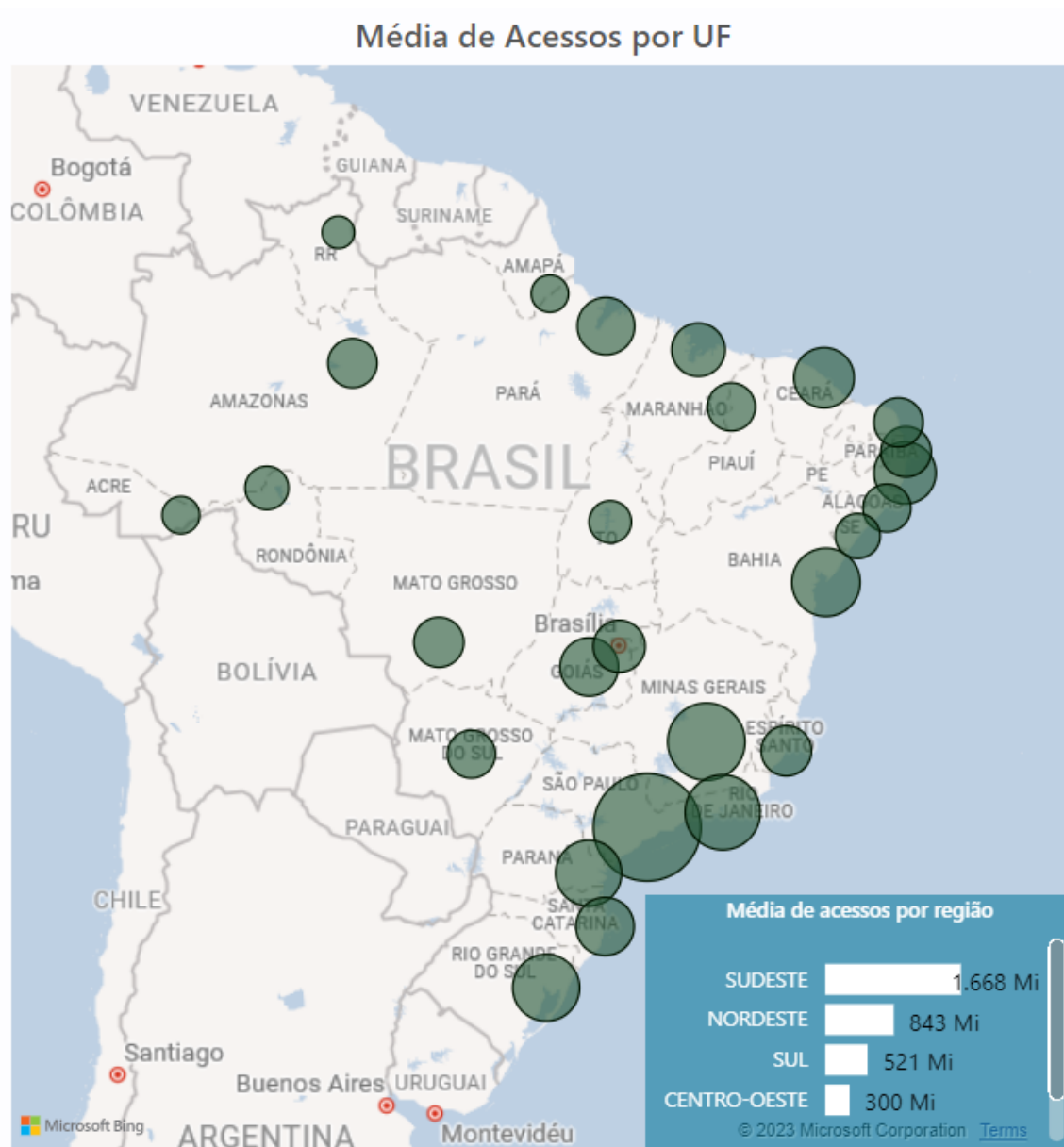


Figura 2 - Média de acessos por grupo econômico

Fonte: Painel BI Mercado Regulado de Telefonia Móvel

Com a finalização do acesso da Oi no primeiro semestre de 2022, já podemos perceber um aumento nos índices de HHI em todos os estados, inclusive apresentando um aumento de mais de 1000 em seis estados - Espírito Santo, Maranhão, Minas Gerais, Rondônia, Sergipe e Tocantins - que já apresentavam um mercado altamente concentrado pelo menos nos últimos cinco anos.

| | UF | HHI_2022 | analise_mercado_hhi_2022 | HHI_2023 | analise_mercado_hhi_2023 | diff_hhi |
|----|----|-----------|--------------------------|-----------|--------------------------|-----------|
| 7 | ES | 4338.6335 | altamente concentrado | 5982.8489 | altamente concentrado | 1644.2154 |
| 9 | MA | 2505.3774 | altamente concentrado | 3610.9166 | altamente concentrado | 1105.5392 |
| 10 | MG | 2643.6564 | altamente concentrado | 3662.8476 | altamente concentrado | 1019.1912 |
| 20 | RO | 3110.5866 | altamente concentrado | 4491.7232 | altamente concentrado | 1381.1366 |
| 24 | SE | 3477.2562 | altamente concentrado | 5360.0210 | altamente concentrado | 1882.7648 |
| 26 | TO | 3130.2870 | altamente concentrado | 4813.1052 | altamente concentrado | 1682.8182 |

Figura 3 - Dataframe com os estados que tiveram um aumento de mais de mil no índice de HHI

Ainda sobre o HHI, temos apenas dois estados que apresentaram o mercado moderadamente concentrado, a saber: Minas Gerais em 2015 e em 2016; e Rio de Janeiro de 2018 a 2020. O mercado esteve moderadamente concentrado no Rio de Janeiro uma vez que a Nextel aumentou sua parcela de mercado nesses anos, porém sua base de clientes foi integrada em 2021 após ser comprada pela Claro em 2019.

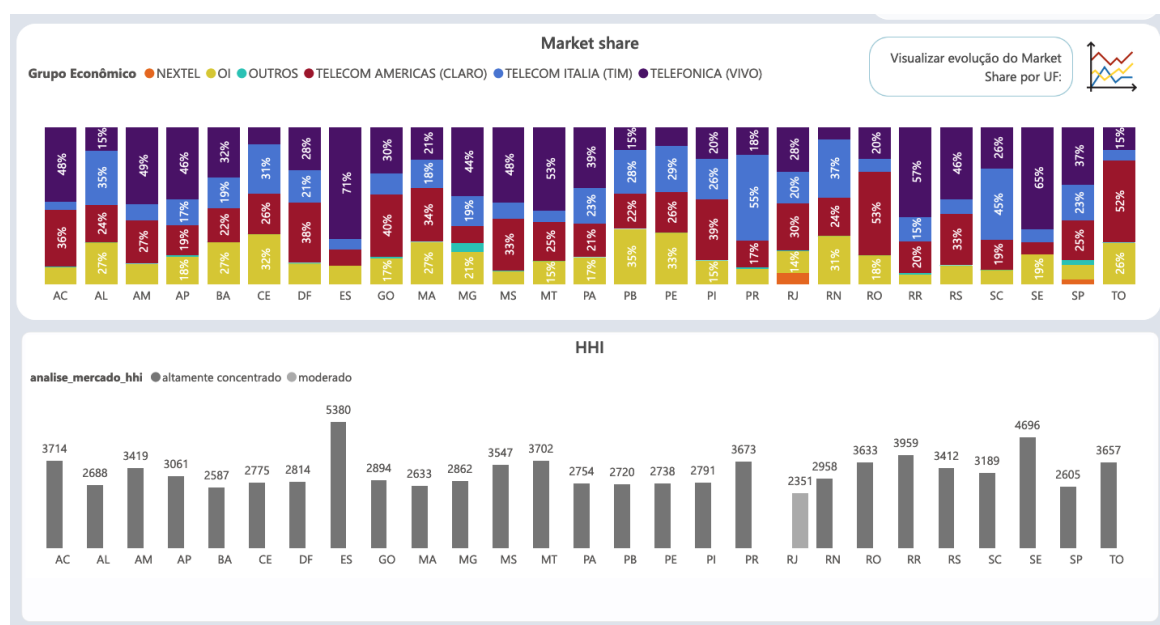


Gráfico 1 - Comparativos de *market share* e HHI do ano de 2020

Fonte: Painel BI Mercado Regulado de Telefonia Móvel

A partir do gráfico acima é possível enxergar como o *market share* de cada grupo econômico reflete no HHI. Espírito Santo, por exemplo, tem a maior concentração de mercado no Brasil no ano de 2020, tendo a Telefônica 71% da parcela de mercado no estado. No caso do Rio de Janeiro, que apresentou no ano de 2020 um mercado moderadamente concentrado, temos a figura significativa de um quarto *player*, a Nextel.

Além disso, utilizando a base histórica de dados (2009 a 2023) que possuímos, foi possível realizar a análise da evolução do *market share* de cada grupo econômico em cada estado no decorrer dos anos. Como exemplo, pode-se observar que a operadora Oi em alguns estados da região nordeste (CE, PB, PE e RN) detinha participação de mercado relevante (acima de 30%) desde 2009 e, a partir de 2022, com a sua interrupção, a operadora Vivo, principalmente, ganhou significativa parcela de mercado nessa região.

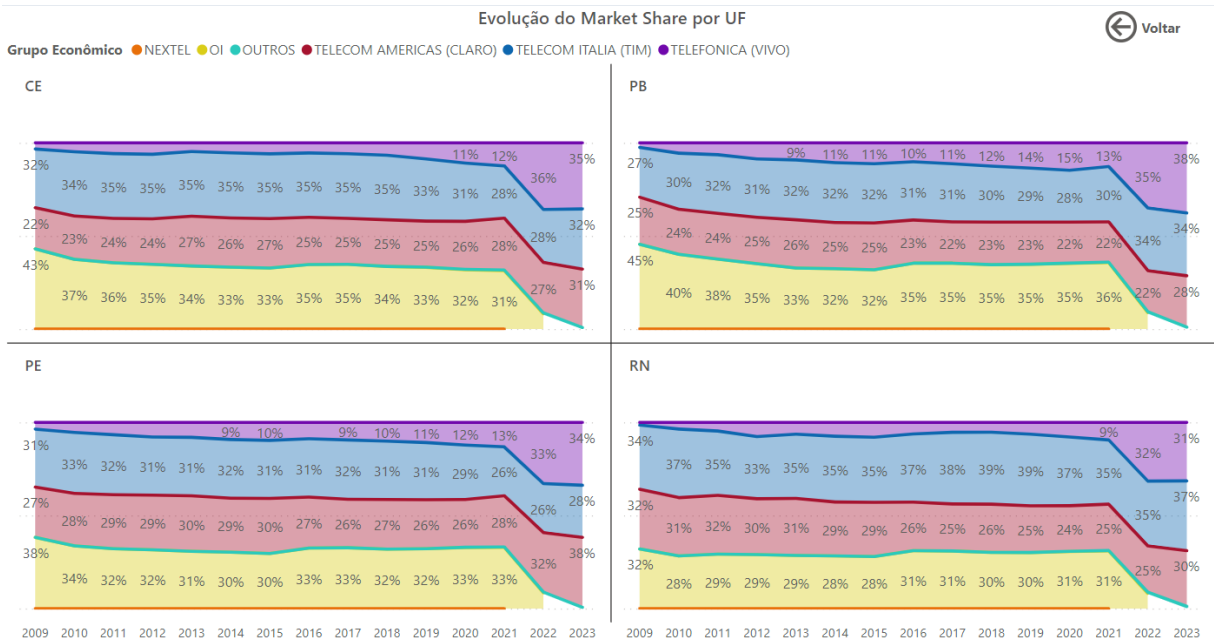


Figura 4 - Gráficos da evolução de market share no Ceará, Paraíba, Pernambuco e Rio Grande do Norte

Fonte: Painel BI Mercado Regulado de Telefonia Móvel

2. Tecnologia

Quanto à tecnologia, percebemos a entrada da 5G em 2021 e, apesar de nova, há uma ascensão pequena, sendo a 4G a tecnologia com a maior média de acesso em 2023.

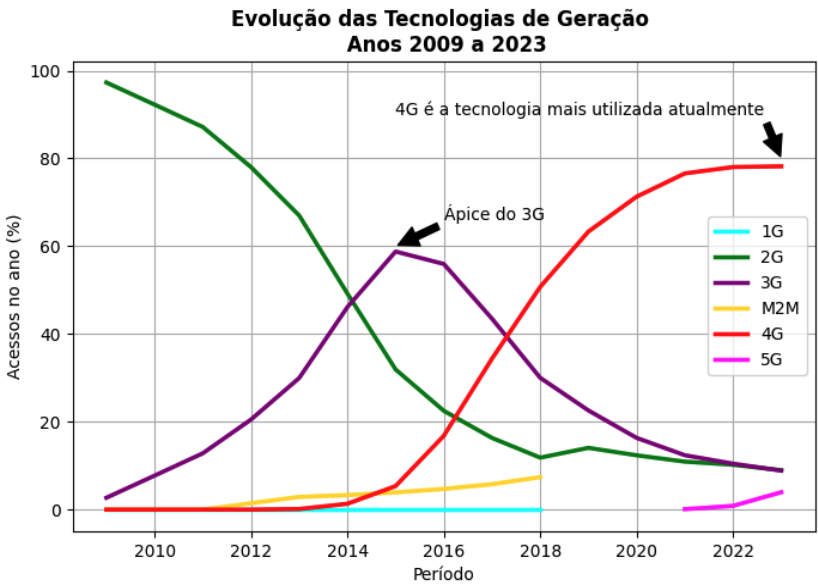


Gráfico 2 - Evolução das tecnologias de geração de 2009 a 2023

Tendo o ano de 2023 como base, temos que a maior parte de acessos de 5G é feita a partir de cobrança do tipo pós-pago, já a 4G, a tecnologia de geração mais utilizada atualmente, temos uma divisão equilibrada entre pós e pré-pago, conforme gráfico a seguir:

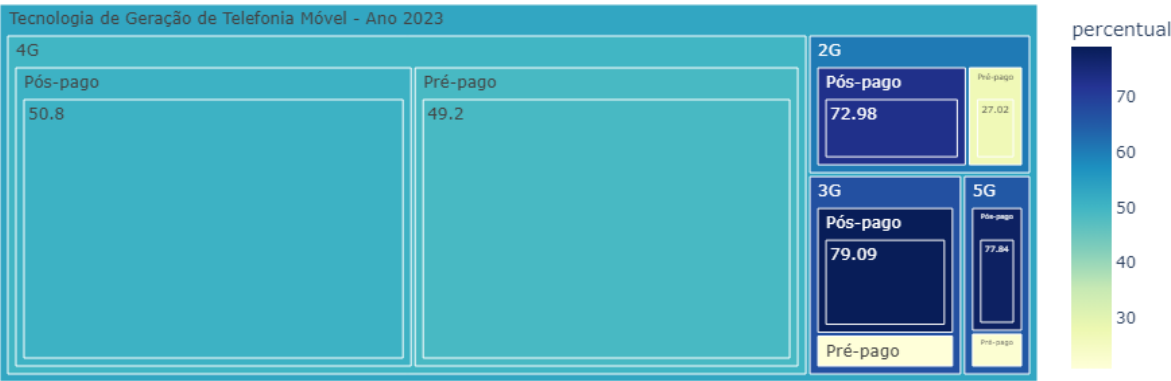


Gráfico 3 - Modalidade de cobrança por tecnologia de geração em 2023

Percentual de acesso 5G por UF

Fonte: <https://bootcamp-mercado-telefonos.streamlit.app/Mapas>

| Tecnologia 5G - Top 10 Municípios (2023) | | |
|--|----------------|------------|
| | Municípios | percentual |
| 1 | São Paulo | 19,52 |
| 2 | Rio de Janeiro | 9,73 |
| 3 | Brasília | 4,68 |
| 4 | Belo Horizonte | 3,85 |
| 5 | Curitiba | 3,43 |
| 6 | Fortaleza | 2,78 |
| 7 | Goiânia | 2,62 |
| 8 | Salvador | 2,57 |
| 9 | Manaus | 2,3 |
| 10 | Porto Alegre | 2,08 |

Figura 6 - Dataframe em formato tabular com os 10 municípios com o maior percentual de tecnologia 5G

Ainda utilizamos ferramentas como mapa *choroplético* para construção de mapas dinâmicos utilizando uma base de dados de média de acessos por DDD no período de 2009 a 2023, para isso fizemos uso de bibliotecas como plotly.express e plotly.io. Os mapas apresentados nas figuras 5 a 8 mostram o retrato da concentração de acessos no retrato de 2009 e 2023 (ou 2022, no caso da da operadora Oi), quanto mais escuro, na escala de cores, estiver determinada região, maior a concentração em número de acessos da operadora.

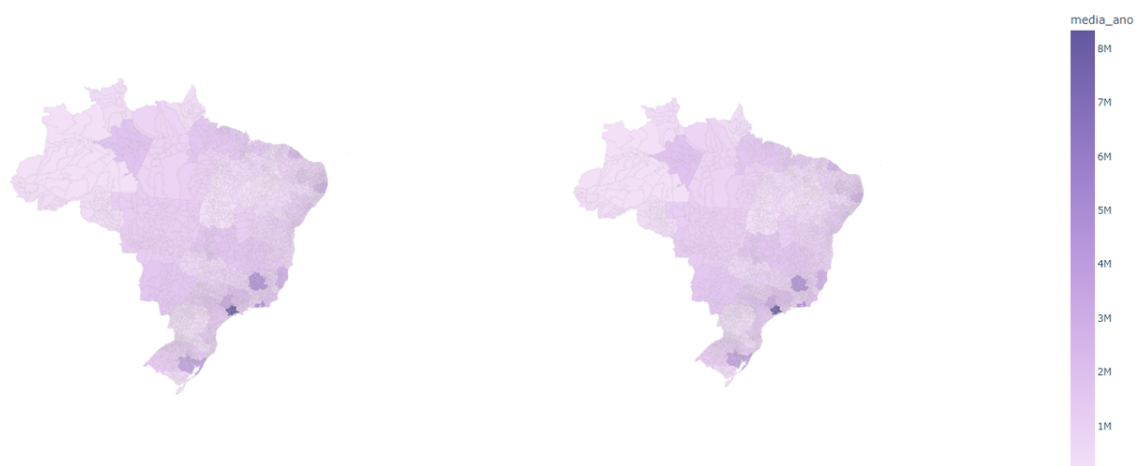


Figura 5 - Mapa do grupo econômico Telefônica (Principal operadora Vivo) - 2009 e 2023
(escala de 1 milhão a 8 milhões de acessos)

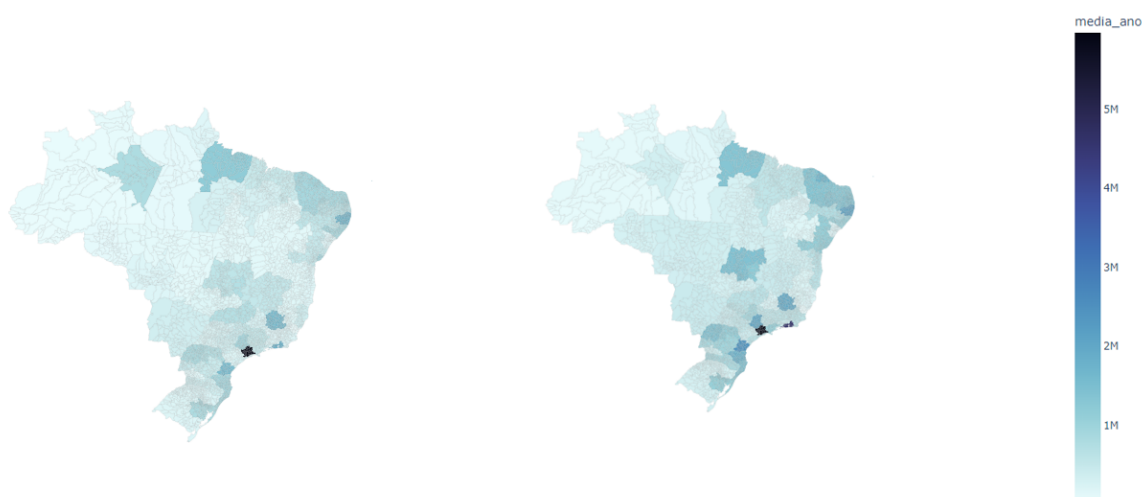


Figura 6 - Mapa do grupo econômico Telecom Itália (Principal operadora TIM) - 2009 e 2023
(escala de 1 milhão a 5 milhões de acessos)

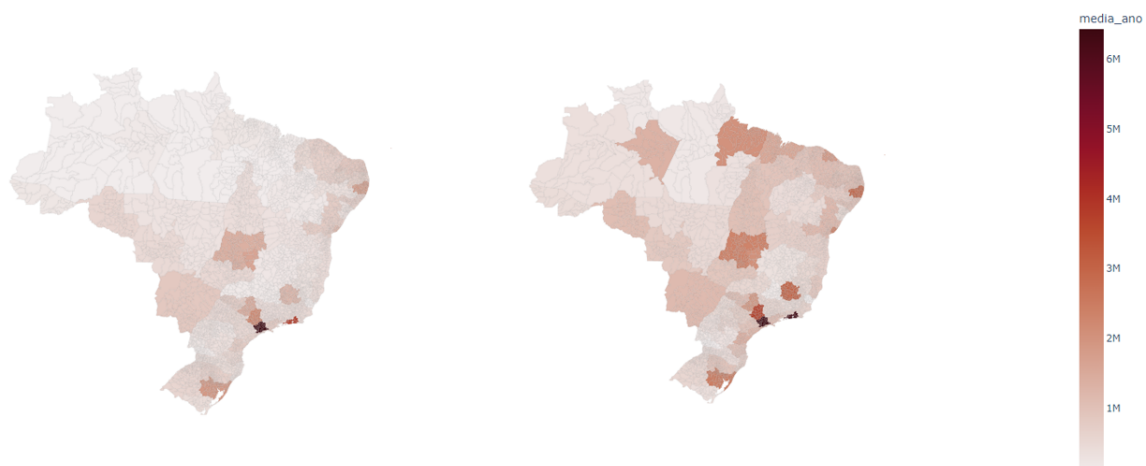


Figura 7 - Mapa do grupo econômico Telecom América (Principal operadora Claro) - 2009 e 2023 (escala de 1 milhão a 6 milhões de acessos)

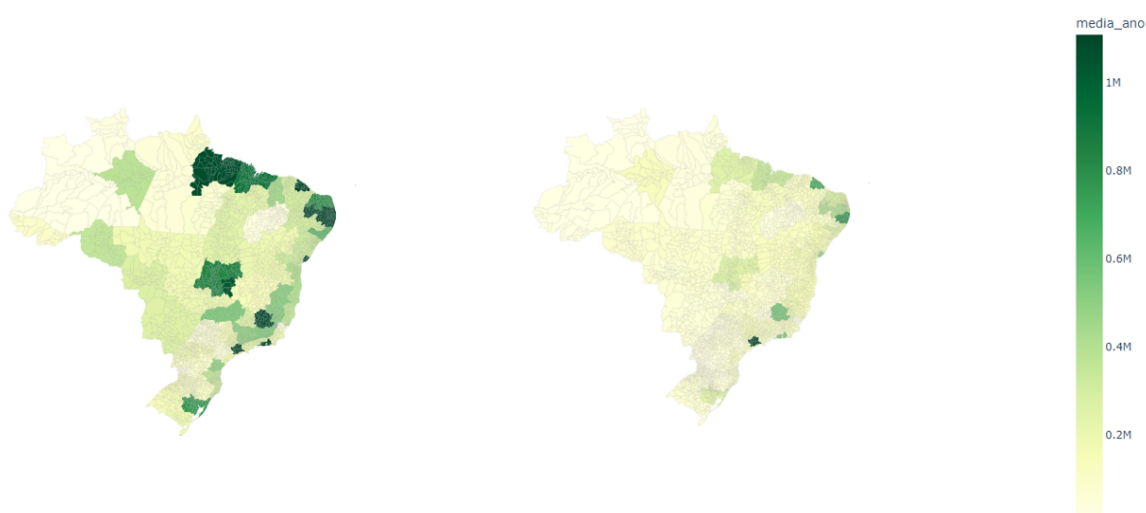


Figura 8 - Mapa da operadora Oi - 2009 e 2022 (escala de 0,2 milhão a 1 milhão de acessos)

Observamos que a operadora Oi tinha uma grande concentração no mercado norte e nordeste e também no centro-oeste, em 2009. Como foi vendida, a partir do segundo semestre de 2022 esse mercado foi absorvido. Observa-se também que o grupo econômico Telecom Itália (TIM), teve perda de mercado ao longo dos anos na região norte e um aumento no centro-oeste. O grupo econômico Telefônica (Vivo) manteve-se mais estático ao longo do período analisado.

PRODUTOS

Além deste Relatório Final, temos como produtos da análise de dados do projeto um painel feito em PowerBI e uma página na web feita com Streamlit.

O painel de PowerBI traz visualizações das análises relacionadas à concentração do mercado de telefonia entre 2009 e 2023. Há três opções de visualização: 1) distribuição geográfica de acessos; 2) comparativo de *market share* e HHI; e 3) tabela dos dados com diversas possibilidades de filtros.

Na primeira opção temos o mapa do Brasil no qual é possível visualizar a média de acessos por região e por UF. Também há um gráfico de barras da média de acesso por grupo econômico. Há como filtrar por grupo econômico e também visualizar os dados de cada ano.

Na segunda opção temos os dados de *market share* e de HHI. Com essa visualização entendemos o porquê da alta concentração e de como a divisão do mercado se dá. Nesta opção também é possível visualizar a evolução do *market share* por UF ao clicar no ícone no canto superior direito.

Na terceira opção podemos visualizar os dados de concentração de mercado em forma tabular, aplicando os filtros desejados, possibilitando uma análise estática dos índices utilizados para analisar a concentração do mercado.

A página web em Streamlit está disponível no endereço <https://bootcamp-mercado-telefonia.streamlit.app>. Nela é possível encontrar um resumo do projeto, a equipe técnica, gráficos estáticos visualizando os dados sobre tecnologia no mercado de telefonia e os links para os outros produtos - painel PowerBI e este relatório final. A aplicação web objetiva publicizar e concentrar os produtos da análise, além de facilitar o acesso ao resumo e à visualização de dados do projeto, especialmente a de tecnologia, uma vez que criamos gráficos estáticos e dinâmicos.