

UNIVERSIDADE PRESBITERIANA MACKENZIE

CARLOS SUETIRO KAERIYAMA

JENIFFER HELLEN VICENTE PIRES

JOÃO FRANCISCO NEVES

JONATHAN HENRIQUE GOMES DA PAZ

**ANÁLISE DO CONSUMO DE ENERGIA ELÉTRICA DO MUNICÍPIO DO RIO DE
JANEIRO**

SÃO PAULO

2022

Sumário

LISTA DE ILUSTRAÇÕES	3
LISTA DE TABELAS	4
Introdução	5
Elaboradores do estudo	5
Apresentação da Empresa	5
Contexto do Estudo	6
Metadados	7
Referências de aquisição do dataset	7
Descrição da origem	7
Descrição do dataset	8
Problema do estudo	8
Análise exploratória de dados	13
Tipos de dados	13
Medidas de posição e dispersão	14
Relações entre os atributos	15
Evolução	16
Valores perdidos ou incorretos	18
Conclusão	18
Apresentação do projeto	19
Referências	19

LISTA DE ILUSTRAÇÕES

Gráfico 01	Boxplot com o consumo MWh per capita.....	10
Gráfico 02	Histograma do consumo kWh por habitante.....	11
Gráfico 03	Gráfico de correlação entre os atributos do dataset.....	16
Gráfico 04	Consumo de energia elétrica médio mensal (kWh) per capita anual (1980 a 2019).....	17
Gráfico 05	Crescimento populacional no período de 1980 à 2019.....	17
Gráfico 06	Consumo de energia anual da população no período de 1980 à 2019.....	18

LISTA DE TABELAS

Tabela 01	Informações dos elaboradores do estudo.....	05
Tabela 02	Cronograma de entregas do estudo.....	07
Tabela 03	Anos com consumo de energia (kWh) per capita no intervalo de 150 a 180 kWh.....	12
Tabela 04	Anos com consumo de energia (kWh) per capita no intervalo de 190 a 200 kWh.....	12
Tabela 05	Anos com consumo de energia (kWh) per capita no intervalo de 220 a 230 kWh.....	13

Introdução

O município do Rio de Janeiro é o segundo mais populoso do Brasil, possuindo uma população estimada de 6.718.903 habitantes segundo dados do DATA RIO. A distribuição de energia elétrica foi concedida à empresa Light Serviços de Eletricidade S.A.

Devido a sua relevância, tanto populacional, quanto política, é necessário realizar uma análise do consumo de energia elétrica do município do Rio de Janeiro para identificar e corrigir problemas em relação a demanda de energia elétrica, bem como projetar as demandas futuras com a finalidade de atender as necessidades da população carioca.

A análise realizada neste trabalho tem o objetivo de descrever a situação atual do consumo de energia elétrica do município do Rio de Janeiro. Foram aplicadas técnicas de estatística descritiva para realizar a Análise Descritiva. A finalidade da análise é identificar o que aconteceu, em relação ao consumo de energia elétrica. No final do estudo são apresentados os códigos utilizados para a Análise Descritiva .

Elaboradores do estudo

O estudo do consumo de energia elétrica do Município do Rio de Janeiro é realizado pelos discentes do curso de graduação em Tecnologia em Ciência de Dados da Universidade Presbiteriana Mackenzie. Os nomes dos alunos e seus registros acadêmicos são apresentados na tabela 01.

Tabela 01 – Informações dos elaboradores do estudo

Nome	Registro Acadêmico
Carlos Suetiro Kaeriyama	10922003274
Jeniffer Hellen Vicente Pires	10922008462
João Francisco Neves	10922007911
Jonathan Henrique Gomes da Paz	10922008659

Apresentação da Empresa

A Light Serviços de Eletricidade S. A. (Light SESA) é a empresa responsável pelo fornecimento de energia elétrica para o Município do Rio de Janeiro. A Light SESA é uma das cinco empresas do Grupo Light – formado também pela Light Energia, a Light Soluções, a Light Esco e a Light Com. A empresa é a quarta maior distribuidora de energia do Brasil, em

número de clientes, e a quinta maior em quantidade de energia distribuída, segundo o Anuário Estatístico de Energia Elétrica de 2010, publicado pela Empresa de Pesquisa Energética (EPE), vinculada ao Ministério de Minas e Energia.

A Light tem como Missão: *“Prover energia e serviços com qualidade e de forma sustentável, contribuindo para o bem-estar e o desenvolvimento da sociedade”*, sua visão é: *“Ser a melhor empresa do setor elétrico, reconhecida pela rentabilidade, eficácia na gestão e qualidade dos serviços”* e seus valores são: *“Valorização da vida, Ética, Senso de dono, Simplicidade e Meritocracia”*.

Conforme o Relatório Anual Integrado de Sustentabilidade de 2021, a força de trabalho da Light é composta por 5.223 colaboradores próprios – sendo 154 pessoas com deficiência, 8.656 terceirizados e 13 estagiários.

Para aproveitar as oportunidades na área de Data Science e se tornar uma empresa data driven, a área de Tecnologia da Light preparou as fundações de analytics e gestão de dados. Isso permitiu que os usuários das áreas, na ponta, cientistas de dados, construíssem seus modelos paramétricos estatísticos, que dão suporte a processos decisórios com mais precisão que contribuem para a eficiência operacional. O resultado desse trabalho foi o projeto de Sistema de detecção de fraude por meio de Random Forest, no qual o Software suportado por algoritmo de múltiplas árvores de decisões (Random Forest) é capaz de analisar os dados históricos e criar regras de maneira automática, com o objetivo de identificar potenciais consumidores fraudadores ou com anomalias de consumo que indiquem irregularidades. Outro projeto implementado, foi o Desenvolvimento de modelos analíticos avançados, com base em aprendizado de máquina para detecção de padrões de fraude ou anomalias sobre base de clientes telemedidos. Onde a ferramenta de aprendizado computacional e metodologia para aumentar o acerto das seleções de clientes potenciais fraudadores do segmento de medição indireta (BTI) é de baixa tensão (200A).

Contexto do Estudo

Estudo do consumo de energia elétrica dos habitantes do Município do Rio de Janeiro. Todos os documentos necessários para a análise serão armazenados na nuvem, na pasta do Google Drive acessada pelo link <https://bityli.com/iKNWAAB>. O dataset será explorado com auxílio do aplicativo colab.to/R.

Tabela 02 – Cronograma de entregas do estudo

Cronograma	Período
Montagem de grupo, definição de projeto, escolha do dataset, escolha da ferramenta para análise do dataset.	15/08/2022 - 06/09/2022
Elaboração da proposta analítica, análise exploratória dos dados.	07/09/2022 - 09/10/2022
Construção de apresentação gráfica dos resultados analíticos e elaboração do data storytelling.	10/10/2022 - 02/11/2022
Ajustes do relatório final e elaboração de um vídeo de apresentação do projeto.	03/11/2022 - 28/11/2022

Metadados

Referências de aquisição do dataset

O dataset de consumo de energia elétrica do Município do Rio de Janeiro foi coletado no site DATA.RIO no formato.xlsx. Conforme informação do próprio site, qualquer pessoa tem acesso ao seu conteúdo e sua licença é CC BY. Os dados foram publicados em 26 de setembro de 2018 e sua última atualização ocorreu em 9 de fevereiro de 2021.

Os dados de consumo de energia elétrica do dataset foram obtidos através da empresa Light Serviços de Eletricidade S.A e os dados da população através do Instituto Brasileiro de Geografia e Estatística (IBGE) e do Instituto Pereira Passos (IPP).

Descrição da origem

O Instituto Pereira Passos é uma autarquia da Prefeitura do Rio de Janeiro e sua responsabilidade é o desenvolvimento de informações estatísticas, mapas, estudos e pesquisas com foco na Cidade do Rio de Janeiro. Sob sua gestão está o projeto Armazém dos Dados, que atualmente se chama DATA.RIO.

O DATA.RIO é o modelo utilizado para o planejamento, integração, gestão e disseminação da informação da Prefeitura do Rio de Janeiro.

A Light Serviços de Eletricidade S.A, responsável pelos dados de consumo de energia elétrica, é a empresa responsável pela distribuição de energia elétrica no Município do Rio de Janeiro e faz parte do Grupo Light.

O Instituto Brasileiro de Geografia e Estatística, responsável pelos dados populacionais, é uma entidade da administração pública federal vinculada ao Ministério do Planejamento e se constitui como principal provedor de dados e informações do país.

Descrição do dataset

O conjunto de dados (dataset) contém a série histórica da evolução do consumo total (MWh), médio anual (em MWh/ hab), mensal (em KWh/ hab) e diário (em MWh/ hab) de energia elétrica no Município do Rio de Janeiro entre os anos de 1980 a 2019. No registro do consumo total foram incluídas as categorias de consumo de energia elétrica residencial, industrial, comercial, rural, poder público, iluminação pública, serviço público e consumo próprio. O dataset possui 40 observações e 6 atributos.

Para os cálculos dos dados de consumo médio anual, mensal e diário por habitante, foram utilizados os dados da população do Município do Rio de Janeiro. Para a obtenção dos dados populacionais foram utilizadas diversas fontes e técnicas, conforme descrito abaixo:

- Para os anos de 1980, 1991, 2000 e 2010 - Censos Demográficos / IBGE;
- Intervalos 1981 a 1990 e 1992 a 1999 - cálculos IPP;
- Ano de 2007 - Contagem IBGE;
- Intervalo 2001 a 2006, 2008, 2009 e intervalo 2011 a 2019 - Estimativas de população, tabela 6579 do SIDRA/IBGE, com data de referência em janeiro de 2021.
- Os valores para o período 2001 a 2009 foram revistos em junho de 2010 a partir dos cálculos de população por interpolação para o intervalo intercensitário 2000 – 2010.

Problema do estudo

Considerando o consumo total anual de energia em MWh, no ano de 1980 foi registrado o menor consumo de energia no valor de 8 871 979 MWh. Ademais, no ano de 2014 foi verificado o maior consumo de energia do dataset, 18 096 224 MWh.


```
[40] #Anos em que houveram menor e maior consumo de energia
data[data$Total.MWh. == min(data$Total.MWh.)|data$Total.MWh. == max(data$Total.MWh.),c(1,2)]
```

A data.frame: 2 x 2

	Ano	Total.MWh.
	<int>	<dbl>
1	1980	8871979
35	2014	18096224

Em seguida, foi analisado o consumo per capita por ano, adicionando o atributo ao dataset, de acordo com as análises abaixo.

```
[41] #Consumo anual de Megawatt por habitante
MWh.Hab = c(0)
MWh.Hab = data$Total.MWh./data$Populacao.hab.
df = data.frame(data$Ano, data$Populacao.hab.,data$Total.MWh.,MWh.Hab)
colnames(df)[3] = "Total.MWh"
colnames(df)[1] = "Ano"
colnames(df)[2] = "Populacao"
head(df)
```

A data.frame: 6 x 4

	Ano	Populacao	Total.MWh	MWh.Hab
	<int>	<int>	<dbl>	<dbl>
1	1980	5090790	8871979	1.742751
2	1981	5125066	8884447	1.733528
3	1982	5159573	9594291	1.859513
4	1983	5194312	10902884	2.099004
5	1984	5229285	11622014	2.222486
6	1985	5264493	11765155	2.234813

O consumo de energia per capita anual no período considerado, possui valor médio de 2.457 MWh/Hab e mediana de 2.480 MWh/Hab. O atributo apresenta desvio padrão de 0.27, o que é significativo, indicando alta variabilidade nos dados.

```
[43] #variabilidade do consumo de energia por habitante anual
summary(df$MWh.Hab)
sd(df$MWh.Hab)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.734	2.346	2.480	2.457	2.675	2.829

0.268594006601974

Entretanto, todos os dados estão dentro do intervalo de três desvios padrões, como pode ser observado na imagem abaixo.

```
[22] 1 #Todos os valores estão dentro do intervalo de 3 desvios padrões.  
2 k = mean(df$MWh.Hab)+3*sd(df$MWh.Hab)  
3 p = mean(df$MWh.Hab)-3*sd(df$MWh.Hab)  
4 print(nrow(df[df$MWh.Hab < k & df$MWh.Hab > p,])/nrow(df))  
  
[1] 3.262864  
[1] 1.6513  
[1] 1
```

Analisando o gráfico boxplot, note que os registros de consumo de energia (MWh/Hab) per capita apresentam dois outliers.

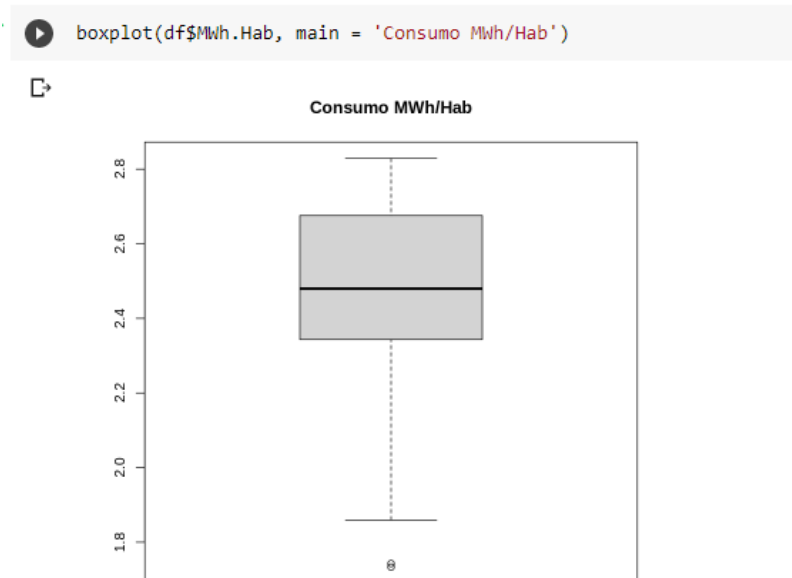


Gráfico 01: Boxplot com o consumo MWh per capita.

Para identificar esses outliers, utilizamos o método de variação interquartil (IQR). No caso, nos anos de 1980 e 1981, os consumos de energia registrados foram considerados como pontos fora da curva, com valores, respectivamente, iguais à 1.7428 MWh/Hab e 1.7335 MWh/Hab.

```
#Outliers MWh/Hab no ano
iqr = IQR(df$MWh.Hab)
Q1 = quantile(df$MWh.Hab, probs = c(.25))
Q3 = quantile(df$MWh.Hab, probs = c(.75))
df[df$MWh.Hab > Q3+1.5*iqr | df$MWh.Hab < Q1-1.5*iqr,]
```



A data.frame: 2 × 4

	Ano	Populacao	Total.MWh	MWh.Hab
	<int>	<int>	<dbl>	<dbl>
1	1980	5090790	8871979	1.742751
2	1981	5125066	8884447	1.733528

No gráfico 2, temos o histograma com o consumo de energia (kWh) per capita. Note que a distribuição é bimodal, possuindo moda nos intervalos de 190 kWh à 200 kWh, e 220 kWh à 230 kWh. O menor consumo de energia (kWh) foi dado no intervalo de variação de 150 kWh a 180 kWh, tendo ocorrido nos anos 1982 e 1983, conforme visto na tabela 03. Observa nas tabelas 4 e 5, os anos em que aconteceram consumo de energia no intervalo de variação modal supracitado.

```
#Consumo mensal de KWh/Hab
hist(1000*df$MWh.Hab/12, main = 'Consumo KWh/Hab', col = 'blue', xlab = 'KWh/Hab')
```

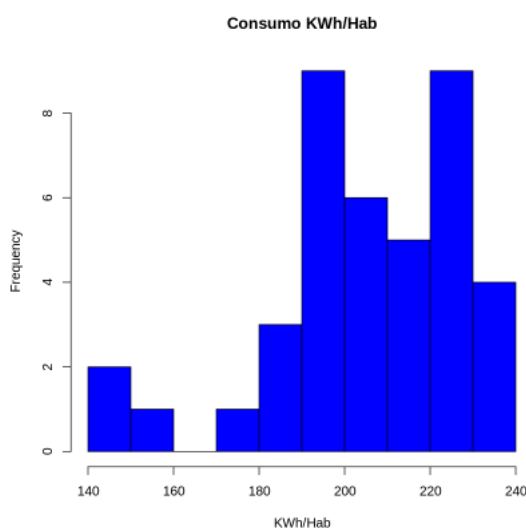


Gráfico 02: Histograma do consumo kWh per capita.

Tabela 03 - Anos com consumo de energia (kWh) per capita no intervalo de 150 a 180 kWh.

```
[ ] 1 #Anos com consumo de energia (kWh/Hab) entre [150;180].
    2 df[1000*df$MWh.Hab/12>=150 & 1000*df$MWh.Hab/12 <= 180,]
```

A data.frame: 2 x 4

	Ano	Populacao	Total.MWh	MWh.Hab
	<int>	<int>	<dbl>	<dbl>
3	1982	5159573	9594291	1.859513
4	1983	5194312	10902884	2.099004

Tabela 04 - Anos com consumo de energia (kWh) per capita no intervalo de 190 a 200 kWh.

```
[ ] 1 df[1000*df$MWh.Hab/12>=190 & 1000*df$MWh.Hab/12 <=200,]
```

A data.frame: 9 x 4

	Ano	Populacao	Total.MWh	MWh.Hab
	<int>	<int>	<dbl>	<dbl>
7	1986	5299939	12446331	2.348391
8	1987	5335623	12475527	2.338158
9	1988	5371547	12652650	2.355495
13	1992	5521452	13124963	2.377085
14	1993	5562429	13106850	2.356318
15	1994	5603709	13386636	2.388889
23	2002	5937253	13739063	2.314044
24	2003	5974081	13868520	2.321448
26	2005	6094183	14341920	2.353379

Tabela 05 - Anos com consumo de energia (kWh) per capita no intervalo de 220 a 230 kWh.

```
[ ] 1 df[1000*df$MWh.Hab/12>=220 & 1000*df$MWh.Hab/12 <=230,]
```

A data.frame: 9 × 4

	Ano	Populacao	Total.MWh	MWh.Hab
	<int>	<int>	<dbl>	<dbl>
18	1997	5729397	15319686	2.673874
28	2007	6132342	16520491	2.693994
29	2008	6161047	16596539	2.693786
31	2010	6320446	16935391	2.679461
32	2011	6355949	16901704	2.659194
33	2012	6390290	17108243	2.677225
34	2013	6429923	17476913	2.718060
36	2015	6476631	17661298	2.726927
37	2016	6498837	17341175	2.668350

Análise exploratória de dados

A análise exploratória de dados abaixo está disponível no diretório “Analise_exploratoria”. O data set no formato csv encontra-se no diretório “Dataset”, assim como sua versão original no formato xlsx.

Tipos de dados

Como dito anteriormente, o dataset possui 6 atributos e 40 registros. Sendo Ano e População do tipo inteiros, e os outros atributos tipo numérico.

```
[50] str(data)
```

```
'data.frame': 40 obs. of 6 variables:
 $ Ano      : int  1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 ...
 $ Total.MWh: num  8871979 8884447 9594291 10902884 11622014 ...
 $ MedioAnual.MWh.hab.: num  1.74 1.73 1.86 2.1 2.22 ...
 $ MedioMensal.KWh.hab.: num  145 144 155 175 185 ...
 $ MedioDiario.KWh.hab.: num  4.84 4.82 5.17 5.83 6.17 ...
 $ Populacao.hab.: int  5090790 5125066 5159573 5194312 5229285 5264493 5299939 5335623 5371547 5407713 ...
```

Medidas de posição e dispersão

Total.MWh.

Medida de posição:

```
[51] print(paste("Média: ", mean(data$Total.MWh)))  
      print(paste("Mediana: ", median(data$Total.MWh)))  
      print("Separatrizes:")  
      print(quantile(data$Total.MWh))
```

```
[1] "Média: 14461747.20405"  
[1] "Mediana: 14442423.185"  
[1] "Separatrizes:"  
      0%      25%      50%      75%     100%  
8871979 13105396 14442423 16630536 18096224
```

Medida de dispersão:

```
[52] print(paste("Variância: ", var(data$Total.MWh)))  
      print(paste("Desvio padrão: ", sd(data$Total.MWh)))  
      print(paste("Amplitude: ", max(data$Total.MWh)-min(data$Total.MWh)))
```

```
[1] "Variância: 6085238278875.45"  
[1] "Desvio padrão: 2466827.57380313"  
[1] "Amplitude: 9224244.834"
```

Populacao.hab.

Medida de posição:

```
[33] print(paste("Média: ", mean(data$Populacao.hab.)))  
      print(paste("Mediana: ", median(data$Populacao.hab.)))  
      print("Separatrizes:")  
      print(quantile(data$Populacao.hab.))
```

```
[1] "Média: 5850557.3"  
[1] "Mediana: 5836327.5"  
[1] "Separatrizes:"  
    0%    25%    50%    75%   100%  
5090790 5435020 5836328 6220144 6718903
```

Medida de dispersão:

```
[34] print(paste("Variância: ", var(data$Populacao.hab.)))  
      print(paste("Desvio padrão: ", sd(data$Populacao.hab.)))  
      print(paste("Amplitude: ", max(data$Populacao.hab.)-min(data$Populacao.hab.)))
```

```
[1] "Variância: 232311993361.087"  
[1] "Desvio padrão: 481987.544819456"  
[1] "Amplitude: 1628113"
```

Relações entre os atributos

Existe forte correlação entre os atributos Ano e População, consumo de energia (MWh) total anual e População, consumo de energia (MWh) per capita e consumo de energia (MWh) total anual, por último consumo de energia (MWh) total anual e Ano, conforme vemos no gráfico 3 que apresenta a correlação entre os atributos supra mencionados.

```

1 #O mapa de calor indica correlação entre os seguintes atributos: (Ano, População),
2 #(Total MWh, População), (MWh/Hab, Total MWh), (Total MWh, Ano).
3 f = signif(cor(df),2)
4 heatmap(f, cexRow= 1.1, cexCol = 1.1)

```

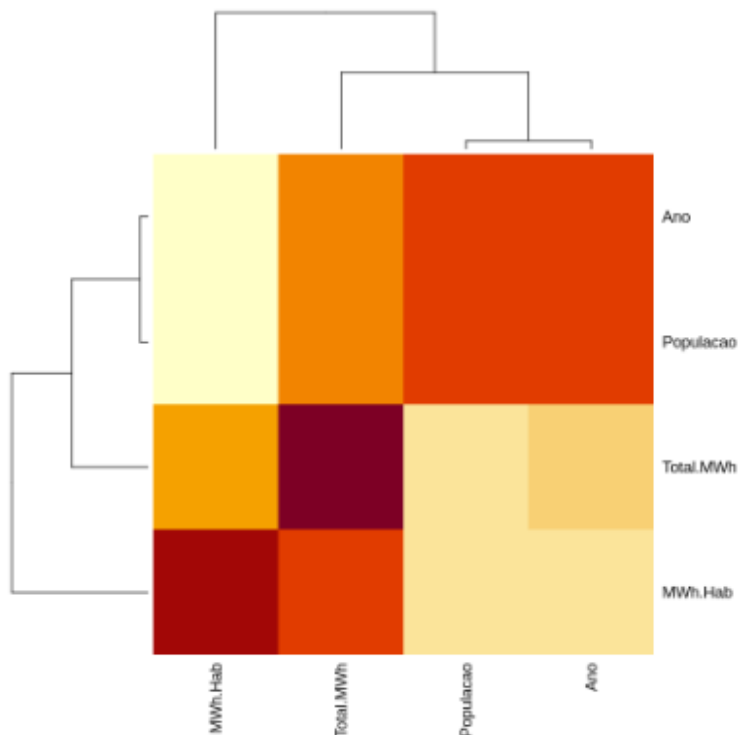


Gráfico 03: Gráfico de correlação entre os atributos do dataset.

Evolução

No período considerado no gráfico 4, o consumo de energia médio mensal (kWh) per capita está crescendo. Sendo que, nos anos de 2000 foi registrado o maior consumo, sendo seguido de uma queda de consumo de energia. No ano de 2002, ocorreu a retomada de crescimento no consumo de energia.


```
[26] 1 #gráfico de linha entre a população e consumo médio mensal kWh/hab entre 1980 à 2019.
      2 plot(df$Ano, 1000*df$MWh.Hab/12, main = 'Consumo de energia médio mensal por habitante entre 1980 e 2019',
      3 xlab = 'Ano', ylab = 'kWh/Hab', col = 'darkblue', type = 'l')
```

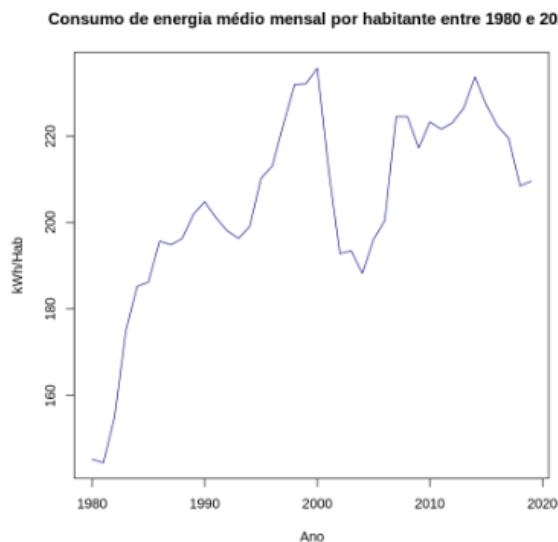


Gráfico 04: Consumo de energia elétrica médio mensal (kWh) per capita anual (1980 a 2019).

No gráfico 5, vemos que a população na cidade do Rio de Janeiro está aumentando no decorrer dos anos. Significando que será demandado cada vez mais geração de energia para atender as necessidades dos cariocas.

```
1 #Crescimento populacional nos anos de 1980 à 2019
2 plot(df$Ano, df$Populacao/1000, main = 'Crescimento populacional entre 1980 à 2019',
3 xlab = 'Ano', ylab = 'População (milhares de habitantes)',
4 type = 'l', col = 'red')
```

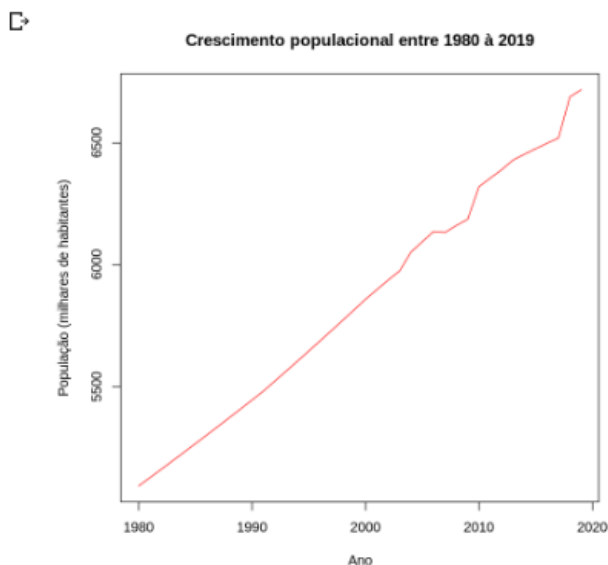


Gráfico 05: Crescimento populacional no período de 1980 à 2019.

No gráfico 6, vemos a tendência de crescimento do consumo de energia anual, devido ao aumento populacional.

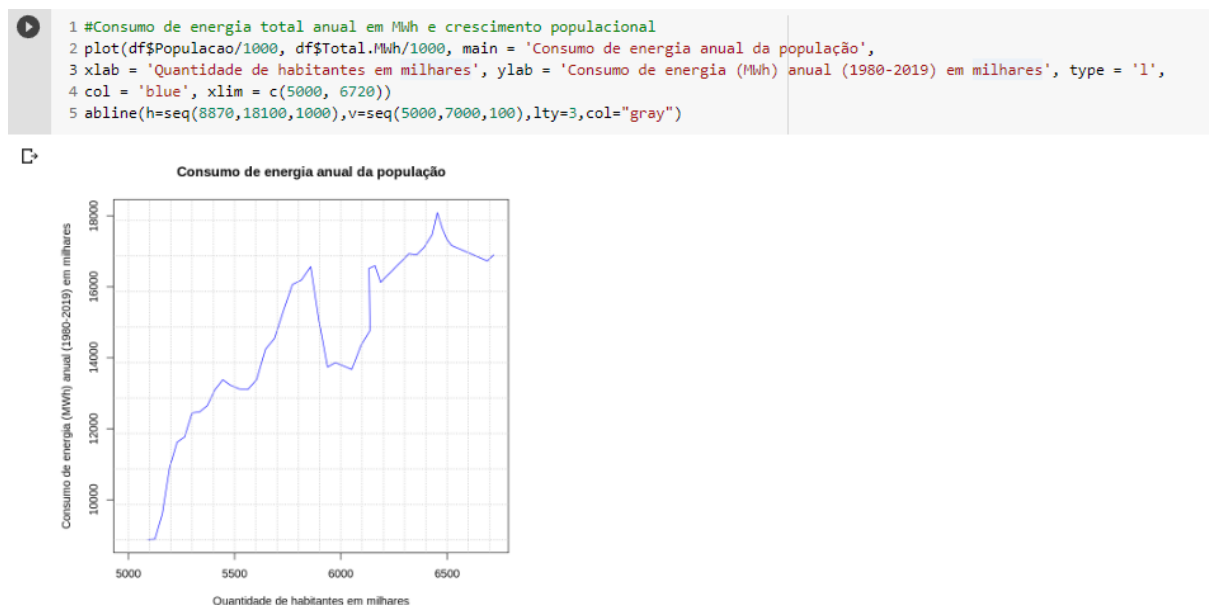


Gráfico 06: Consumo de energia anual da população no período de 1980 à 2019.

Valores perdidos ou incorretos

```
[57] sapply(data, function(x) sum(is.na(x)))
```

Ano: 0 Total.MWh.: 0 MedioAnual.MWh.hab.: 0 MedioMensal.KWh.hab.: 0 MedioDiario.KWh.hab.: 0 Populacao.hab.: 0

Conclusão

Foi feito o estudo do dataset disponível no domínio DATA RIO da Prefeitura Municipal do Rio de Janeiro, onde observamos o consumo de energia no período de 1980 à 2019. Verificamos que o aumento da população do município ao longo dos anos redundou numa demanda crescente por energia. Logo é importante desenvolver novas fontes de energia, preferencialmente ecológicas, visando minimizar a agressão ao meio ambiente, garantindo assim que a demanda por energia dos cariocas seja atendida. Evitando que venha a ocorrer períodos de crise energética, com eventuais recessões, sendo necessário adotar uma postura de austeridade energética no município.

A posteriori se faz mister realizar uma análise diagnóstica sobre os fatos levantados na análise descritiva elaborada neste documento. O intuito da análise diagnóstica será identificar as causas que levaram a um consumo de energia (KWh) anual per capita (entre 150 e 180 KWh) reduzido no período de 1982 à 1983. De modo semelhante, entender porque no período de 1997, 2007 - 2008, 2010 - 2013 e 2015 - 2016 o consumo de energia (KWh)

anual per capita foi de 220 a 230 KWh. Quais fatores influenciaram para que a demanda por energia fosse tão alta em relação aos anos anteriores. Verificar a possibilidade de mitigar esse consumo sem perda de produtividade econômica dos cariocas.

Por último, através da análise prescritiva poderemos projetar o consumo de energia da população da capital do estado do Rio de Janeiro. Com essa projeção estabeleceremos um plano de geração de energia de forma sustentável que atenda a todos os habitantes. Deste modo, reduziríamos a probabilidade de crise energética, com consequente crise econômica.

Apresentação do projeto

A apresentação do projeto foi postada no You Tube. O link para acesso é:

<https://youtu.be/zJkmcigNQno>

Referências

CONSUMO TOTAL, MÉDIO ANUAL, MENSAL E DIÁRIO DE ENERGIA ELÉTRICA POR HABITANTE NO MUNICÍPIO DO RIO DE JANEIRO ENTRE 1980 - 2019. DATA.RIO, 2017. Disponível em: <https://bityli.com/McErymS>. Acesso em: 03 de set. de 2022.

DATA RIO. Data.Rio Informações sobre a cidade, c. 2017. Página inicial. Disponível em: <https://www.data.rio/>. Acesso em: 03 de set. de 2022.

LIGHT. Informações sobre a empresa Light. Páginas consultadas. Disponível em:

<http://www.light.com.br/grupo-light/Empresas-do-Grupo/light-servicos-de-eletricidade.aspx>

<http://ri.light.com.br/sustentabilidade/relatorios/> . Acesso em: 06 de out. de 2022.