

Cássio Henrique Pereira Coelho - 201307175

Felipe de Oliveira Rios - 201307182

Thiago da Silva Fagundes - 201307206

Estruturas de Dados 2 - Tarefa 3

Introdução

Segundo proposto, foram implementados algoritmos de pesquisa sequencial e binária na linguagem Python (versão 3), os quais foram executados sobre os conjuntos de dados fornecidos. Ambos código e saída gerada estão anexados na última seção do relatório.

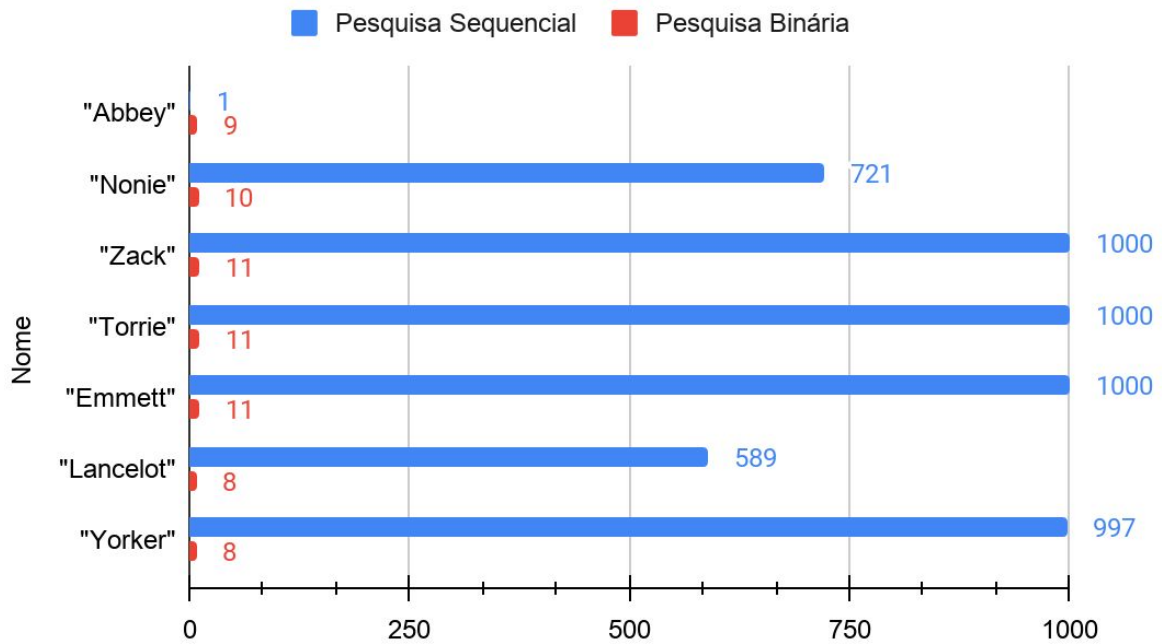
Resultados Obtidos

Como esperado, para todos os arquivos, que possuíam um grande número de entradas, o algoritmo de busca binária obteve resultados consideravelmente superiores. Apesar de que, no melhor caso (quando o elemento buscado está na primeira posição), o desempenho do algoritmo sequencial tenha se mostrado superior, no pior caso (quando o elemento buscado não está na lista) e no caso médio (o elemento está no meio da lista), o algoritmo binário mostrou resultados muito superiores aos resultados da busca sequencial. Essa diferença entre os desempenhos se mostrou cada vez maior para números cada vez maiores de entradas. Abaixo, seguem os resultados obtidos em tabelas e seus respectivos gráficos comparativos.

Arquivo Contendo 1000 Entradas

Nome	Comparações (Sequencial)	Comparações (Binária)
"Abbey"	1	9
"Nonie"	721	10
"Zack"	1000	11
"Torrie"	1000	11
"Emmett"	1000	11
"Lancelot"	589	8
"Yorker"	997	8

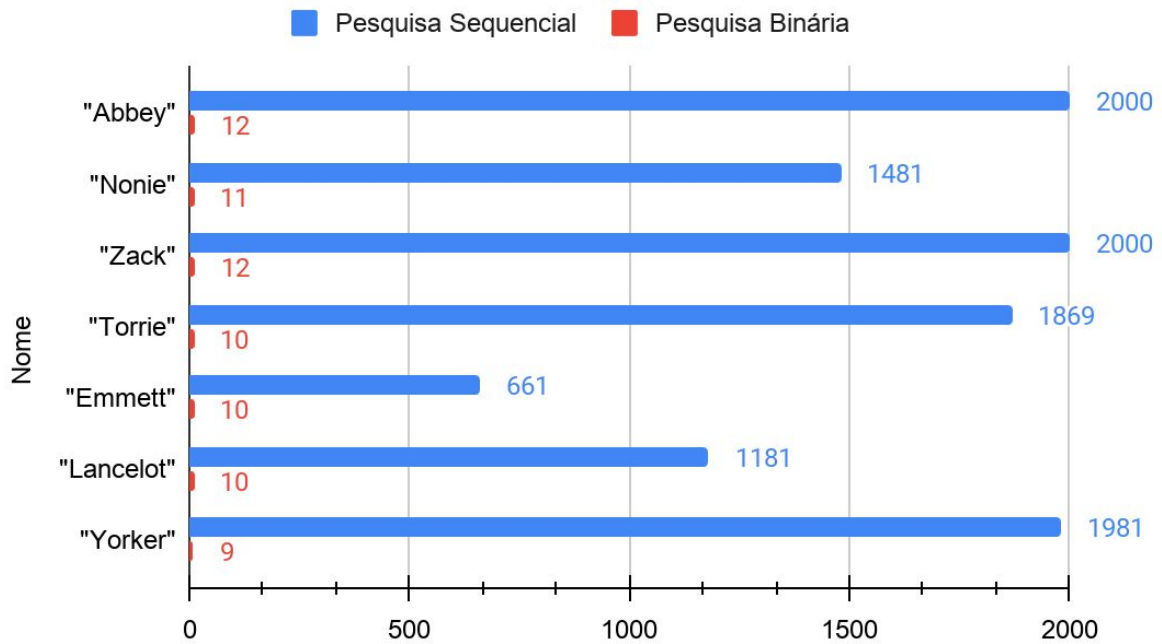
Número de Comparações dos Algoritmos de Busca Sequencial e Binária em um arquivo de 1000 entradas



Arquivo Contendo 2000 Entradas

Nome	Comparações (Sequencial)	Comparações (Binária)
"Abbey"	2000	12
"Nonie"	1481	11
"Zack"	2000	12
"Torrie"	1869	10
"Emmett"	661	10
"Lancelot"	1181	10
"Yorker"	1981	9

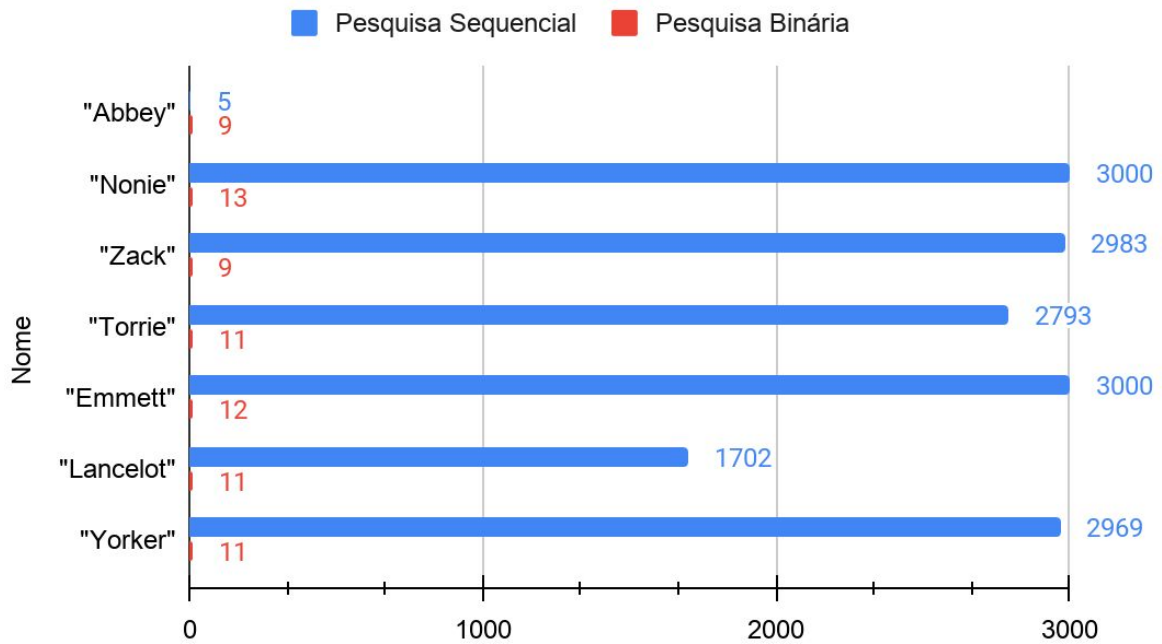
Número de Comparações dos Algoritmos de Busca Sequencial e Binária em um arquivo de 2000 entradas



Arquivo Contendo 3000 Entradas

Nome	Comparações (Sequencial)	Comparações (Binária)
"Abbey"	5	9
"Nonie"	3000	13
"Zack"	2983	9
"Torrie"	2793	11
"Emmett"	3000	12
"Lancelot"	1702	11
"Yorker"	2969	11

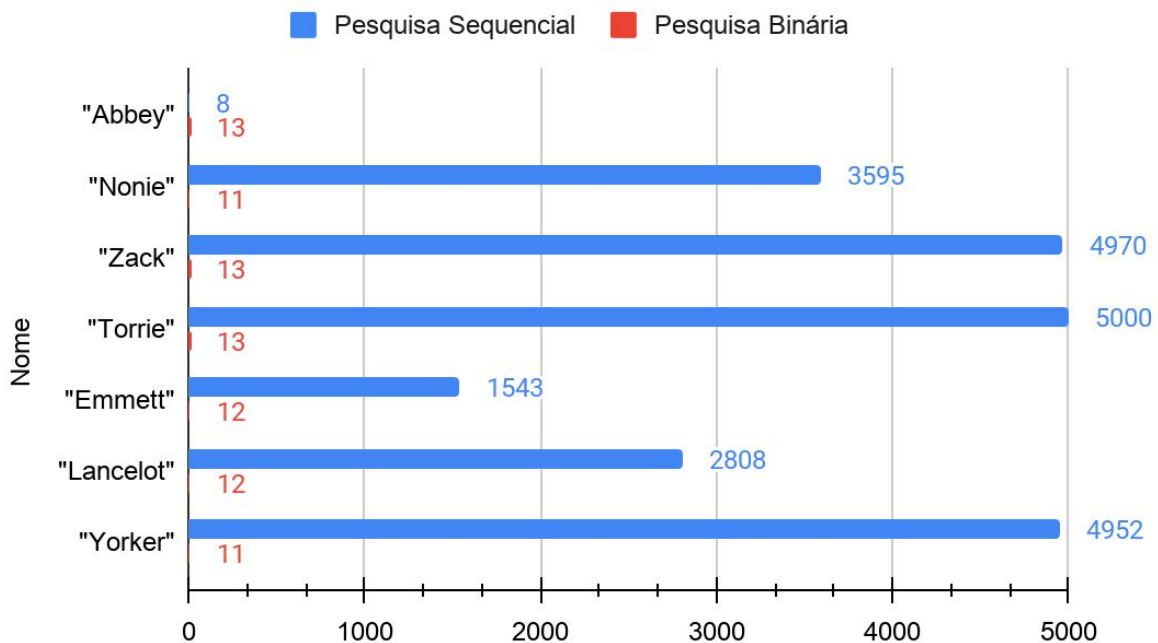
Número de Comparações dos Algoritmos de Busca Sequencial e Binária em um arquivo de 3000 entradas



Arquivo Contendo 5000 Entradas

Nome	Comparações (Sequencial)	Comparações (Binária)
"Abbey"	8	13
"Nonie"	3595	11
"Zack"	4970	13
"Torrie"	5000	13
"Emmett"	1543	12
"Lancelot"	2808	12
"Yorker"	4952	11

Número de Comparações dos Algoritmos de Busca Sequencial e Binária em um arquivo de 5000 entradas



Conclusão

Dados os resultados e sua discrepância, é indiscutível a escolha do algoritmo de pesquisa sequencial para pesquisas por elementos em conjuntos de dados ordenados, e portanto, para o sistema proposto; todavia, este método requer que o conjunto de dados esteja previamente ordenado, e seria incapaz de retornar os resultados pretendidos caso este requisito não seja cumprido. Assim, para conjuntos de dados não-ordenados, entre o método de busca sequencial e o método de busca binária, é preferível ordenar os dados usando um método de ordenação de boa performance, e usar o método de busca binária. Caso a ordenação não seja possível, não se tem escolha a não ser adotar o método de busca sequencial. Em síntese, a baixa performance do método de busca sequencial faz com que não seja aplicável em sistemas de larga escala, e seu uso é conveniente apenas em protótipos iniciais, dada sua simplicidade e facilidade de implementação.

Apêndice

Código Executável

pesquisa_sequencial.py

```
def pesquisa_sequencial(lista, nome):
    tamanho = len(lista)
    compara = 0
    for i in range(0, tamanho):
        compara += 1
        if(nome == lista[i]["first_name"]):
            print("\t\tEntrada encontrada após {}
comparações.".format(str(compara)) )
            return i
        print("\t\tEntrada não encontrada após {}
comparações.".format(str(compara)) )
    return -1
```

pesquisa_binaria.py

```
def pesquisa_binaria(lista, nome, inicio = 0, fim = None, compara = 0):
    if fim is None:
        fim = len(lista) - 1

    meio = int((inicio + fim)/2)

    auxiliar = lista[meio]["first_name"]

    compara += 1

    if inicio > fim:
        print("\t\tEntrada não encontrada após {}
comparações.".format(str(compara)) )
        return -1
    elif(nome == auxiliar):
        print("\t\tEntrada encontrada após {}
comparações.".format(str(compara)) )
        return meio
    else:
        if nome > auxiliar:
            return pesquisa_binaria(lista, nome, meio + 1, fim, compara)
        else:
            return pesquisa_binaria(lista, nome, inicio, meio - 1,
```

compara)

main.py

```
# coding: utf-8

from csv import DictReader
import json

import pesquisa_binaria
import pesquisa_sequencial
import sys

def main():

#=====
# Constantes e pré-carregamento de arquivos
#=====

    comprimento_linhas = 80

    # Lista com os caminhos dos arquivos nos quais será efetuada a
    # busca.
    lista_caminhos = [
        "datasets_prontos/1000/crescente.csv",
        "datasets_prontos/2000/crescente.csv",
        "datasets_prontos/3000/crescente.csv",
        "datasets_prontos/5000/crescente.csv"
    ]

    lista_nomes = [] # Lista que será preenchida com os nomes

    # Preenche a lista com os nomes do arquivo.
    with open("nomes_para_busca.txt", "r") as arquivo_nomes:
        for nome in arquivo_nomes.readlines():
            lista_nomes.append(nome.strip())

#=====
# Parte da pesquisa sequencial
```

```

#=====
=====
    print("="*comprimento_linhas)
    print("\tPESQUISA SEQUENCIAL")
    print("="*comprimento_linhas)

    for caminho in lista_caminhos:

        print("\tRealizando pesquisas no arquivo
        \{}".format(caminho))
        print("-"*comprimento_linhas)

        # Loop que itera sobre a lista de caminhos de arquivos nos quais
        # será realizada a pesquisa.
        with open(caminho, "r") as arquivo:
            dados = list(DictReader(arquivo))

            # Loop que itera sobre os nomes a serem buscados.
            for nome in lista_nomes:
                print("\tProcurando por \{}".format(nome))
                resultado =
                pesquisa_sequencial.pesquisa_sequencial(dados, nome)
                if resultado >= 0:
                    print(dictToString(dados[resultado], indent=3))
                    print("\t\t0 elemento estava na posição
                    {}".format(resultado))
                    print("")

                print("\n{}".format("-"*comprimento_linhas))

#=====
=====

    # Parte da pesquisa binária

#=====
=====

    print("="*comprimento_linhas)
    print("\tPESQUISA BINÁRIA")
    print("="*comprimento_linhas)

    # Loop que itera sobre a lista de caminhos de arquivos nos quais
    # será realizada a pesquisa.

```



```

    for caminho in lista_caminhos:

        print("\tRealizando pesquisas no arquivo
        \{}".format(caminho))
        print("-"*comprimento_linhas)

        with open(caminho, "r") as arquivo:
            dados = list(DictReader(arquivo))

            # Loop que itera sobre os nomes a serem buscados.
            for nome in lista_nomes:
                print("\tProcurando por \{} no
                arquivo...".format(nome))
                resultado = pesquisa_binaria.pesquisa_binaria(dados,
                nome)

                if resultado >= 0:
                    print(dictToString(dados[resultado], indent=3))
                    print("\t\t0 elemento estava na posição
                    {}".format(resultado))
                    print("")

                print("\n{}".format("-"*comprimento_linhas))

def dictToString(dictionary, indent=0):
    leading = "\t"*indent
    string = ""

    for key, value in dictionary.items():
        string += "{}{}: \{}".format(leading, key, value)

        if key != list(dictionary.keys())[-1]:
            string += "\n"

    return string

if __name__ == "__main__":
    main()

```

O projeto completo pode ser encontrado no repositório do github, em <https://github.com/cassioozn/Tarefa-3-Metodos-de-Pesquisa>.

Saída Gerada pela Execução

```
=====
=====
PESQUISA SEQUENCIAL
=====
=====
Realizando pesquisas no arquivo
"datasets_prontos/1000/crescente.csv"...
-----
-----
Procurando por "Abbey" no arquivo...
Entrada encontrada após 1 comparações:
    first_name: "Abbey"
    last_name: "Atmore"
    email: "aatmore1x@wufoo.com"
    gender: "Male"
    ip_address: "62.7.21.108"
0 elemento estava na posição 0.

Procurando por "Nonie" no arquivo...
Entrada encontrada após 721 comparações:
    first_name: "Nonie"
    last_name: "Purrington"
    email: "npurringtonlo@about.com"
    gender: "Female"
    ip_address: "108.158.117.48"
0 elemento estava na posição 720.

Procurando por "Zack" no arquivo...
Entrada não encontrada após 1000 comparações.

Procurando por "Torrie" no arquivo...
Entrada não encontrada após 1000 comparações.

Procurando por "Emmett" no arquivo...
Entrada não encontrada após 1000 comparações.

Procurando por "Lancelot" no arquivo...
Entrada encontrada após 589 comparações:
    first_name: "Lancelot"
    last_name: "Blindmann"
    email: "lblindmannu@oakley.com"
    gender: "Male"
    ip_address: "234.250.12.69"
0 elemento estava na posição 588.

Procurando por "Yorker" no arquivo...
```

Entrada encontrada após 997 comparações:
first_name: "Yorker"
last_name: "Pavlishchev"
email: "ypavlishchevny@nydailynews.com"
gender: "Male"
ip_address: "105.177.75.82"
O elemento estava na posição 996.

Realizando pesquisas no arquivo
"datasets_prontos/2000/crescente.csv"...

Procurando por "Abbey" no arquivo...
Entrada não encontrada após 2000 comparações.

Procurando por "Nonie" no arquivo...
Entrada encontrada após 1481 comparações:
first_name: "Nonie"
last_name: "Glavias"
email: "nglavias9b@techcrunch.com"
gender: "Female"
ip_address: "157.59.102.50"
O elemento estava na posição 1480.

Procurando por "Zack" no arquivo...
Entrada não encontrada após 2000 comparações.

Procurando por "Torrie" no arquivo...
Entrada encontrada após 1869 comparações:
first_name: "Torrie"
last_name: "Filippo"
email: "tfilippoem@imdb.com"
gender: "Female"
ip_address: "69.114.61.72"
O elemento estava na posição 1868.

Procurando por "Emmett" no arquivo...
Entrada encontrada após 661 comparações:
first_name: "Emmett"
last_name: "Szapiro"
email: "eszapiropq@techcrunch.com"
gender: "Male"
ip_address: "235.77.99.43"

O elemento estava na posição 660.

Procurando por "Lancelot" no arquivo...

Entrada encontrada após 1181 comparações:

first_name: "Lancelot"
last_name: "Eusden"
email: "leusdeny@gmpg.org"
gender: "Male"
ip_address: "7.145.235.42"

O elemento estava na posição 1180.

Procurando por "Yorker" no arquivo...

Entrada encontrada após 1981 comparações:

first_name: "Yorker"
last_name: "Browell"
email: "ybrowellqs@google.com"
gender: "Male"
ip_address: "14.108.228.39"

O elemento estava na posição 1980.

Realizando pesquisas no arquivo

"datasets_prontos/3000/crescente.csv"...

Procurando por "Abbey" no arquivo...

Entrada encontrada após 5 comparações:

first_name: "Abbey"
last_name: "Wellwood"
email: "awellwoodlm@ed.gov"
gender: "Male"
ip_address: "210.58.44.93"

O elemento estava na posição 4.

Procurando por "Nonie" no arquivo...

Entrada não encontrada após 3000 comparações.

Procurando por "Zack" no arquivo...

Entrada encontrada após 2983 comparações:

first_name: "Zack"
last_name: "Ferron"
email: "zferronin@webmd.com"
gender: "Male"
ip_address: "243.35.238.94"

O elemento estava na posição 2982.

Procurando por "Torrie" no arquivo...

Entrada encontrada após 2793 comparações:

first_name: "Torrie"
last_name: "Hughesdon"
email: "thughesdonby@skype.com"
gender: "Female"
ip_address: "158.9.226.100"

O elemento estava na posição 2792.

Procurando por "Emmett" no arquivo...

Entrada não encontrada após 3000 comparações.

Procurando por "Lancelot" no arquivo...

Entrada encontrada após 1702 comparações:

first_name: "Lancelot"
last_name: "Crotty"
email: "lcrottylt@narod.ru"
gender: "Male"
ip_address: "142.233.200.20"

O elemento estava na posição 1701.

Procurando por "Yorker" no arquivo...

Entrada encontrada após 2969 comparações:

first_name: "Yorker"
last_name: "Franklin"
email: "yfranklin82@wisc.edu"
gender: "Male"
ip_address: "183.177.220.180"

O elemento estava na posição 2968.

Realizando pesquisas no arquivo
"datasets_prontos/5000/crescente.csv"...

Procurando por "Abbey" no arquivo...

Entrada encontrada após 8 comparações:

first_name: "Abbey"
last_name: "Atmore"
email: "aatmore1x@wufoo.com"
gender: "Male"
ip_address: "62.7.21.108"

O elemento estava na posição 7.

Procurando por "Nonie" no arquivo...

Entrada encontrada após 3595 comparações:

first_name: "Nonie"
last_name: "Purrington"
email: "npurringtonlo@about.com"
gender: "Female"
ip_address: "108.158.117.48"

O elemento estava na posição 3594.

Procurando por "Zack" no arquivo...

Entrada encontrada após 4970 comparações:

first_name: "Zack"
last_name: "Nelmes"
email: "znelmeskp@imgur.com"
gender: "Male"
ip_address: "133.79.25.94"

O elemento estava na posição 4969.

Procurando por "Torrie" no arquivo...

Entrada não encontrada após 5000 comparações.

Procurando por "Emmett" no arquivo...

Entrada encontrada após 1543 comparações:

first_name: "Emmett"
last_name: "Dowdall"
email: "edowdall13k@yahoo.co.jp"
gender: "Male"
ip_address: "78.54.106.138"

O elemento estava na posição 1542.

Procurando por "Lancelot" no arquivo...

Entrada encontrada após 2808 comparações:

first_name: "Lancelot"
last_name: "Blindmann"
email: "lblindmannu@oakley.com"
gender: "Male"
ip_address: "234.250.12.69"

O elemento estava na posição 2807.

Procurando por "Yorker" no arquivo...

Entrada encontrada após 4953 comparações:

first_name: "Yorker"
last_name: "Pavlishchev"
email: "ypavlishchevny@nydailynews.com"

```
gender: "Male"
ip_address: "105.177.75.82"
O elemento estava na posição 4952.
```

```
-----
-----
=====
=====
```

PESQUISA BINÁRIA

```
=====
=====
```

```
Realizando pesquisas no arquivo
"datasets_prontos/1000/crescente.csv"...
```

```
-----
-----
```

```
Procurando por "Abbey" no arquivo...
Entrada encontrada após 9 comparações:
  first_name: "Abbey"
  last_name: "Atmore"
  email: "aatmore1x@wufoo.com"
  gender: "Male"
  ip_address: "62.7.21.108"
O elemento estava na posição 0.
```

```
Procurando por "Nonie" no arquivo...
Entrada encontrada após 10 comparações:
  first_name: "Nonie"
  last_name: "Purrington"
  email: "npurringtonlo@about.com"
  gender: "Female"
  ip_address: "108.158.117.48"
O elemento estava na posição 720.
```

```
Procurando por "Zack" no arquivo...
Entrada não encontrada após 11 comparações.
```

```
Procurando por "Torrie" no arquivo...
Entrada não encontrada após 11 comparações.
```

```
Procurando por "Emmett" no arquivo...
Entrada não encontrada após 11 comparações.
```

```
Procurando por "Lancelot" no arquivo...
Entrada encontrada após 8 comparações:
  first_name: "Lancelot"
```

```
last_name: "Blindmann"
email: "lblindmannu@oakley.com"
gender: "Male"
ip_address: "234.250.12.69"
O elemento estava na posição 588.
```

Procurando por "Yorker" no arquivo...

Entrada encontrada após 8 comparações:

```
first_name: "Yorker"
last_name: "Pavlishchev"
email: "ypavlishchevny@nydailynews.com"
gender: "Male"
ip_address: "105.177.75.82"
O elemento estava na posição 996.
```

Realizando pesquisas no arquivo
"datasets_prontos/2000/crescente.csv"...

Procurando por "Abbey" no arquivo...

Entrada não encontrada após 12 comparações.

Procurando por "Nonie" no arquivo...

Entrada encontrada após 11 comparações:

```
first_name: "Nonie"
last_name: "Glavias"
email: "nglavias9b@techcrunch.com"
gender: "Female"
ip_address: "157.59.102.50"
O elemento estava na posição 1480.
```

Procurando por "Zack" no arquivo...

Entrada não encontrada após 12 comparações.

Procurando por "Torrie" no arquivo...

Entrada encontrada após 10 comparações:

```
first_name: "Torrie"
last_name: "Filippo"
email: "tfilippoem@imdb.com"
gender: "Female"
ip_address: "69.114.61.72"
O elemento estava na posição 1868.
```


Procurando por "Emmett" no arquivo...

Entrada encontrada após 10 comparações:

first_name: "Emmett"

last_name: "Szapiro"

email: "eszapiropq@techcrunch.com"

gender: "Male"

ip_address: "235.77.99.43"

0 elemento estava na posição 660.

Procurando por "Lancelot" no arquivo...

Entrada encontrada após 10 comparações:

first_name: "Lancelot"

last_name: "Eusden"

email: "leusdeny@gmpg.org"

gender: "Male"

ip_address: "7.145.235.42"

0 elemento estava na posição 1180.

Procurando por "Yorker" no arquivo...

Entrada encontrada após 9 comparações:

first_name: "Yorker"

last_name: "Browell"

email: "ybrowellqs@google.com"

gender: "Male"

ip_address: "14.108.228.39"

0 elemento estava na posição 1980.

Realizando pesquisas no arquivo

"datasets_prontos/3000/crescente.csv"...

Procurando por "Abbey" no arquivo...

Entrada encontrada após 9 comparações:

first_name: "Abbey"

last_name: "Wellwood"

email: "awellwoodlm@ed.gov"

gender: "Male"

ip_address: "210.58.44.93"

0 elemento estava na posição 4.

Procurando por "Nonie" no arquivo...

Entrada não encontrada após 13 comparações.

Procurando por "Zack" no arquivo...

Entrada encontrada após 9 comparações:

first_name: "Zack"
last_name: "Ferron"
email: "zferronin@webmd.com"
gender: "Male"
ip_address: "243.35.238.94"

0 elemento estava na posição 2982.

Procurando por "Torrie" no arquivo...

Entrada encontrada após 11 comparações:

first_name: "Torrie"
last_name: "Hughesdon"
email: "thughesdonby@skype.com"
gender: "Female"
ip_address: "158.9.226.100"

0 elemento estava na posição 2792.

Procurando por "Emmett" no arquivo...

Entrada não encontrada após 12 comparações.

Procurando por "Lancelot" no arquivo...

Entrada encontrada após 11 comparações:

first_name: "Lancelot"
last_name: "Crotty"
email: "lcrottylt@narod.ru"
gender: "Male"
ip_address: "142.233.200.20"

0 elemento estava na posição 1701.

Procurando por "Yorker" no arquivo...

Entrada encontrada após 11 comparações:

first_name: "Yorker"
last_name: "Franklin"
email: "yfranklin82@wisc.edu"
gender: "Male"
ip_address: "183.177.220.180"

0 elemento estava na posição 2968.

Realizando pesquisas no arquivo
"datasets_prontos/5000/crescente.csv"...

Procurando por "Abbey" no arquivo...

Entrada encontrada após 13 comparações:

first_name: "Abbey"
last_name: "Atmore"
email: "aatmore1x@wufoo.com"
gender: "Male"
ip_address: "62.7.21.108"

0 elemento estava na posição 7.

Procurando por "Nonie" no arquivo...

Entrada encontrada após 11 comparações:

first_name: "Nonie"
last_name: "Purrington"
email: "npurringtonlo@about.com"
gender: "Female"
ip_address: "108.158.117.48"

0 elemento estava na posição 3594.

Procurando por "Zack" no arquivo...

Entrada encontrada após 13 comparações:

first_name: "Zack"
last_name: "Nelmes"
email: "znelmeskp@imgur.com"
gender: "Male"
ip_address: "133.79.25.94"

0 elemento estava na posição 4969.

Procurando por "Torrie" no arquivo...

Entrada não encontrada após 13 comparações.

Procurando por "Emmett" no arquivo...

Entrada encontrada após 12 comparações:

first_name: "Emmett"
last_name: "Dowdall"
email: "edowdall13k@yahoo.co.jp"
gender: "Male"
ip_address: "78.54.106.138"

0 elemento estava na posição 1542.

Procurando por "Lancelot" no arquivo...

Entrada encontrada após 12 comparações:

first_name: "Lancelot"
last_name: "Blindmann"
email: "lblindmannu@oakley.com"
gender: "Male"
ip_address: "234.250.12.69"

O elemento estava na posição 2807.

Procurando por "Yorker" no arquivo...

Entrada encontrada após 11 comparações:

first_name: "Yorker"

last_name: "Pavlishchev"

email: "ypavlishchevny@nydailynews.com"

gender: "Male"

ip_address: "105.177.75.82"

O elemento estava na posição 4952.

O arquivo de log também está incluso no repositório do projeto no github.