

Relatório do 1º Trabalho de Estrutura de Dados

Gabriel Silva Malta e Leonardo Perassoli*

Maio de 2020

1 Motivação e Introdução

O seguinte trabalho consiste na leitura, processamento e manipulação de dados lidos a partir de um arquivo csv. O arquivo covid19_clean_complete.csv é disponibilizado pelo site kaggle.com e reúne dados de aproximadamente 200 países sobre o andamento da doença COVID-19.

Dadas as especificações, buscamos aprender a processar dados, aprofundar o conhecimento em métodos de ordenação e criar uma ferramenta que, devido a seriedade da pandemia e atualidade do problema, oferecesse praticidade para obter informação sobre o cenário de 2020.

A plataforma nos apresenta dados em relação ao número de casos, mortos e recuperados da doença, bem como fichas técnicas sobre cada país registrado, nos ajudando a compreender o comportamento da doença em diferentes intervalos de tempo, desde o final de Janeiro até o dia 20 de Maio.

2 Os Registros

Foi necessário a criação de 3 structs para a execução do trabalho:

- Dado

```
typedef struct dado{
    char state[TAMs], country[TAMc];
    float lat, longe;
    int conf, death, recov, dia, mes, ano;
}Dado;
```

Figure 1: Registro de Dado

- Bucket

```
typedef struct bucket{
    Dado chave;
    int virgem;
}Bucket;
```

Figure 2: Registro de Bucket

*10262401 e 10376026, respectivamente

- País

```
typedef struct pais{
    char pais[TAMc];
    float mortes,confirmados,recuperados;
    float tx_recov,tx_mort;
    float var_med;
}Pais;
```

Figure 3: Registro de País

3 Os TADs

Para a manipulação dos dados, criamos uma rotina de TADs:

- Calcular a quantidade de linhas do arquivo csv;
- Criar e escrever dados tratados em um arquivo txt;
- Ordenar com MergeSort;
- Ordenar com InsertionSort;
- Iniciar e preencher uma tabela hash;
- Buscar um país na tabela hash;
- Pedir e guardar datas para o usuário;
- Buscar país com maior número de casos em um intervalo;
- Retornar a taxa de recuperados e mortos de um país.

4 Rodando o programa

Segue abaixo um exemplo de como o programa funciona:

4.1 Início do programa

A primeira escolha a ser feita no algoritmo é o método de ordenação. Após a escolha do método, existem três opções para o critério de ordenação:

- Nome do país;
- Latitude;
- Longitude.

```
#####
##### Acompanhamento COVID-19 #####
#####

Realizado pelos Alunos: Leonardo Perassoli
                        Gabriel Silva Malta

Os dados lidos por esse programa podem ser encontrados na url: https://www.kaggle.com/imdevskp/corona-virus-report
O intuito do programa e ser uma plataforma de acompanhamento dos dados a respeito da doenca COVID-19. Os metodos de
ordenacao e estruturacao dos dados sao para fins de avaliacao da materia SCC0224 Estrutura de Dados II do ICMC - USP.

***** Digite com qual metodo deseja ordenar seus dados *****

a)Insertion-Sort
b)Mergesort

Opcao:
```

Figure 4: Primeira tela do programa

```
Os dados lidos por esse programa podem ser encontrados na url: https://www.kaggle.com/imdevskp/corona-virus-report
O intuito do programa e ser uma plataforma de acompanhamento dos dados a respeito da doenca COVID-19. Os metodos de
ordenacao e estruturacao dos dados sao para fins de avaliacao da materia SCC0224 Estrutura de Dados II do ICMC - USP.

***** Digite com qual metodo deseja ordenar seus dados *****

a)Insertion-Sort
b)Mergesort

Opcao: b

***** Digite em qual dimensao deseja ordenar *****

1)Ordenar por nome do pais
2)Ordenar por latitude (Sul -> Norte)
3)Ordenar por longitude (Oeste -> Leste)

Opcao: 2

Tempo de execucao da ordenacao foi de 0.0480 segundos
Dados ordenados e processados!
Montando a tabela hash...
```

Figure 5: Escolha do critério

Dependendo do critério escolhido, o documento criado terá diferentes tipos de organização.

Após a criação do txt, será exibido na tela o tempo demorado para a execução do método escolhido. Além disso, a tela mostra que a tabela hash foi inicializada e preenchida.

```

País: United Kingdom
Estado: Falkland Islands (Malvinas)
Data: 20/5/20
Lat/Long: -51.80S/-59.520
Número de Confirmados: 13
Número de Mortes: 0
Número de Recuperados: 13

País: United Kingdom
Estado: Falkland Islands (Malvinas)
Data: 21/5/20
Lat/Long: -51.80S/-59.520
Número de Confirmados: 13
Número de Mortes: 0
Número de Recuperados: 13

País: Australia
Estado: Tasmania
Data: 22/1/20
Lat/Long: -41.45S/145.97L
Número de Confirmados: 0
Número de Mortes: 0
Número de Recuperados: 0

País: Australia
Estado: Tasmania
Data: 23/1/20
Lat/Long: -41.45S/145.97L
Número de Confirmados: 0
Número de Mortes: 0
Número de Recuperados: 0

```

Figure 6: Exemplo do arquivo de saída

```

Os dados lidos por esse programa podem ser encontrados na url: https://www.kaggle.com/imdevskp/corona-virus-report
O intuito do programa e ser uma plataforma de acompanhamento dos dados a respeito da doenca COVID-19. Os metodos de
ordenacao e estruturacao dos dados sao para fins de avaliacao da materia SCC0224 Estrutura de Dados II do ICMC - USP.

***** Digite com qual metodo deseja ordenar seus dados *****

a)Insertion-Sort
b)Mergesort

Opcao: b

***** Digite em qual dimensao deseja ordenar *****

1)Ordenar por nome do pais
2)Ordenar por latitude (Sul -> Norte)
3)Ordenar por longitude (Oeste -> Leste)

Opcao: 2

Tempo de execucao da ordenacao foi de 0.0480 segundos
Dados ordenados e processados!
Montando a tabela hash...

```

Figure 7: Criação da tabela hash

4.2 Menu

No menu principal, 7 opções estarão disponíveis:

Escolhendo a opção número 1, podemos pesquisar as fichas dos países que possuem dados registrados.

```

***** Menu de Opcoes *****

1)Pesquisar ficha de um pais (dado uma data)
2)Pais com maior numero de mortos (dado um periodo)
3)Pais com maior numero de confirmados (dado um periodo)
4)Pais com maior taxa de mortalidade (dado um periodo)
5)Pais com maior taxa de recuperacao (dado um periodo)
6)Variacao media de confirmados, mortos e recuperados de um pais(dado um periodo)
7)Sair do programa

Opcao: 1

Digite o nome do pais para pesquisar:

```

Figure 8: Menu principal

4.3 Simulando uma pesquisa de ficha

Após entrar com a opção desejada, será requisitado o nome de um país. No exemplo, foi digitado Brasil com "s", mas no arquivo csv, Brasil está escrito com "z". Logo, um erro foi acusado:

```

***** Menu de Opcoes *****

1)Pesquisar ficha de um pais (dado uma data)
2)Pais com maior numero de mortos (dado um periodo)
3)Pais com maior numero de confirmados (dado um periodo)
4)Pais com maior taxa de mortalidade (dado um periodo)
5)Pais com maior taxa de recuperacao (dado um periodo)
6)Variacao media de confirmados, mortos e recuperados de um pais(dado um periodo)
7)Sair do programa

Opcao: 1

Digite o nome do pais para pesquisar:Brasil
Nao foi possivel encontrar esse Pais. Algumas sugestoes que podem se assemelhar com o que procura:

Brazil
Brunei

Digite o nome do pais para pesquisar:

```

Figure 9: Escolha errada do país

A função sugestoes() indicará alguns nomes de países que se assemelham com o que foi digitado.

Entrando Brazil, teremos que digitar uma data para obtermos a ficha:

```
***** Menu de Opcoes *****
1)Pesquisar ficha de um pais (dado uma data)
2)Pais com maior numero de mortos (dado um periodo)
3)Pais com maior numero de confirmados (dado um periodo)
4)Pais com maior taxa de mortalidade (dado um periodo)
5)Pais com maior taxa de recuperacao (dado um periodo)
6)Variacao media de confirmados, mortos e recuperados de um pais(dado um periodo)
7)Sair do programa

Opcao: 1

Digite o nome do pais para pesquisar:Brasil
Nao foi possivel encontrar esse Pais. Algumas sugestoes que podem se assemelhar com o que procura:

Brazil
Brunei

Digite o nome do pais para pesquisar:Brazil

Digite um mes entre Janeiro (1) e Maio (5): 4
Digite o dia: 20
Digite o ano correspondente (somente os dois numeros finais): 20

Pesquisando por: Brazil na data 20/4/2020...

Ficha do pais:
Indice na tabela hash: 5711
Nome: Brazil
Numero de Casos:40743
Numero de Curas:22130
Numero de Mortos:2587
```

Figure 10: Tela da simulação da pesquisa

Podemos ver o número de mortos, casos curados e casos totais.

```
Ficha do pais:
Indice na tabela hash: 5711
Nome: Brazil
Numero de Casos:40743
Numero de Curas:22130
Numero de Mortos:2587
```

Figure 11: Ficha do pais Brasil

Ao terminar o processo, a tela do menu principal será exibida novamente possibilitando a procura por outra ficha de país e/ou outras informações.

4.4 Simulando uma pesquisa de quantidade de mortos

A quantidade de mortes é calculada por períodos. Portanto, será requisitado duas datas diferentes. Com essas datas, o algoritmo calcula a variação de mortos das datas de cada país e nos retorna qual foi o país com o maior número de mortes.

```
***** Menu de Opcoes *****
1)Pesquisar ficha de um pais (dado uma data)
2)Pais com maior numero de mortos (dado um periodo)
3)Pais com maior numero de confirmados (dado um periodo)
4)Pais com maior taxa de mortalidade (dado um periodo)
5)Pais com maior taxa de recuperacao (dado um periodo)
6)Variacao media de confirmados, mortos e recuperados de um pais(dado um periodo)
7)Sair do programa

Opcao: 2

Primeira Data:

Digite um mes entre Janeiro (1) e Maio (5): 1
Digite o dia: 22
Digite o ano correspondente (somente os dois numeros finais): 20

Segunda Data:

Digite um mes entre Janeiro (1) e Maio (5): 20
Digite o numero do mes entre Janeiro (1) e Maio (5): 5
Digite o dia: 20
Digite o ano correspondente (somente os dois numeros finais): 20
O pais com maior numero de mortes nesse periodo e:

US : 93439
```

Figure 12: Escolhendo as datas para a pesquisa

No exemplo, pesquisamos qual país tem a maior quantidade de mortos entre a data 22/1/2020 e a data 20/5/20. Podemos concluir que, desde o começo da pandemia, em Janeiro, até um dos ultimos registros do arquivo, no final de Maio, os Estados Unidos sofreram com 93439 mortes.

4.5 Simulando uma pesquisa de maior número de casos

Assim como a quantidade de mortes, é possível, dado um período, pesquisar qual país possui a maior quantidade de casos entre duas datas escolhidas pelo usuário.

```
***** Menu de Opcoes *****
1)Pesquisar ficha de um pais (dado uma data)
2)Pais com maior numero de mortos (dado um periodo)
3)Pais com maior numero de confirmados (dado um periodo)
4)Pais com maior taxa de mortalidade (dado um periodo)
5)Pais com maior taxa de recuperacao (dado um periodo)
6)Variacao media de confirmados, mortos e recuperados de um pais(dado um periodo)
7)Sair do programa

Opcao: 3

Primeira Data:

Digite um mes entre Janeiro (1) e Maio (5): 1
Digite o dia: 22
Digite o ano correspondente (somente os dois numeros finais): 20

Segunda Data:

Digite um mes entre Janeiro (1) e Maio (5): 5
Digite o dia: 20
Digite o ano correspondente (somente os dois numeros finais): 20
O pais com maior numero de confirmados nesse periodo e:

US : 1551852
```

Figure 13: Escolhendo as datas para a nova pesquisa

Logo, entre 22/1/20 e 20/5/20, foram registrados nos Estados Unidos 1.551.852 casos.

4.6 Taxas

Foram implementados os cálculos das seguintes taxas:

- Taxa de Mortalidade por período;
- Taxa de Recuperação por período;
- Variação média.

4.6.1 Simulando a pesquisa das taxas por período

```
Primeira Data:
Digite um mes entre Janeiro (1) e Maio (5): 1
Digite o dia: 22
Digite o ano correspondente (somente os dois numeros finais): 20

Segunda Data:
Digite um mes entre Janeiro (1) e Maio (5): 5
Digite o dia: 20
Digite o ano correspondente (somente os dois numeros finais): 20
O pais com maior taxa de mortalidade nesse periodo e:
US : 4.01 por cento      Conf.: 1939603 & Mortes: 93439
```

Figure 14: Taxa de Mortalidade

```
Opcao: 5
Primeira Data:
Digite um mes entre Janeiro (1) e Maio (5): 1
Digite o dia: 22
Digite o ano correspondente (somente os dois numeros finais): 20

Segunda Data:
Digite um mes entre Janeiro (1) e Maio (5): 5
Digite o dia: 20
Digite o ano correspondente (somente os dois numeros finais): 20
O pais com maior taxa de recuperapão nesse periodo e:

Western Sahara : 50.00 por cento Total de Casos: 18 & Recup.: 6
Seychelles : 50.00 por cento      Total de Casos: 33 & Recup.: 11
Saint Lucia : 50.00 por cento     Total de Casos: 54 & Recup.: 18
Saint Kitts and Nevis : 50.00 por cento Total de Casos: 45 & Recup.: 15
Papua New Guinea : 50.00 por cento      Total de Casos: 24 & Recup.: 8
Eritrea : 50.00 por cento          Total de Casos: 117 & Recup.: 39
Dominica : 50.00 por cento         Total de Casos: 48 & Recup.: 16
Cambodia : 50.00 por cento         Total de Casos: 366 & Recup.: 122
```

Figure 15: Taxa de Recuperação

```

Opcao: 6
Digite o nome do pais para pesquisar:Brazil

Primeira Data:

Digite um mes entre Janeiro (1) e Maio (5): 5
Digite o dia: 20
Digite o ano correspondente (somente os dois numeros finais): 20

Segunda Data:

Digite um mes entre Janeiro (1) e Maio (5): 1
Digite o dia: 22
Digite o ano correspondente (somente os dois numeros finais): 20

Brazil: entre 20/5/2020 e 22/1/2020

Taxa de Mortalidade : 3.35 por cento      Conf.: 562663.0 & Mortos: 18859.0
Taxa de Recuperacao : 20.74 por cento    Conf.: 562663.0 & Rec.: 116683.0
Variacao do Numero de Casos : 100.000000 por cento

```

Figure 16: Variação do país

5 Encerrar o programa

```

***** Menu de Opcoes *****

1)Pesquisar ficha de um pais (dado uma data)
2)Pais com maior numero de mortos (dado um periodo)
3)Pais com maior numero de confirmados (dado um periodo)
4)Pais com maior taxa de mortalidade (dado um periodo)
5)Pais com maior taxa de recuperacao (dado um periodo)
6)Variacao media de confirmados, mortos e recuperados de um pais(dado um periodo)
7)Sair do programa

Opcao: 7
Obrigado e Ate Logo...

Process returned 0 (0x0)   execution time : 375.976 s
Press any key to continue.

```

Figure 17: Encerramento do programa

6 Fluxo do Programa

Nessa seção, é possível observar um resumo do comportamento do programa, o que acontece desde a leitura do arquivo csv até o cálculo das taxas.

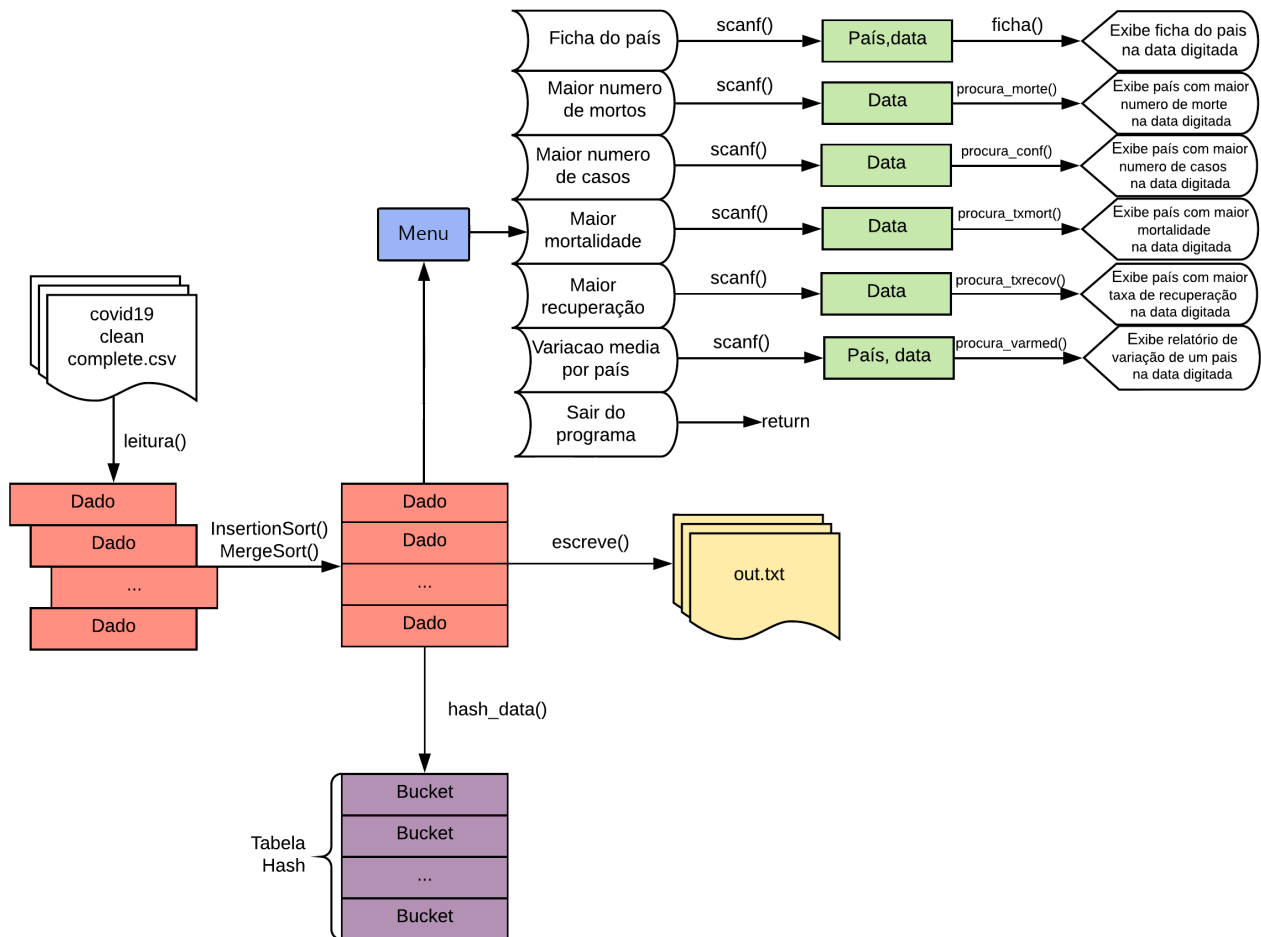


Figure 18: Visualização da Estrutura do Programa

7 Conclusão

Após nosso primeiro contato com um programa de processamento de dados, foi possível perceber que, mesmo com uma quantidade pequena de informações e com um certo tratamento, podemos ter compreensão de uma situação que parece ser muito mais intrincada do que realmente é.

Além disso, se interpretados corretamente, esses dados podem nos mostrar que países que estão usando métodos diferentes para tratar a pandemia estão obtendo resultados muito diferentes. Governos preparados, como o chinês ou o alemão, lidaram de maneira correta ao não negar a ciência, observando aspectos e levando em consideração materiais com informações de surtos anteriores, conseguindo minimizar as potenciais baixas e danos à economia.