

# Trabalho de Estrutura de Dados

## Etapa 2

### Ordenação Externa

Nome do Aluno: Arthur Augusto Magalhães ----- Matricula: 202110830

Nome do Aluno: Aaron Martins Leão Ferreira ----- Matricula: 202120496

Nome do Aluno: Wildes Augusto Sousa-----Matricula:202020602

Para realização da segunda parte do trabalho, o grupo teve como estratégia adotar a intercalação polifásica, uma vez que, a mesma utiliza apenas uma fita livre diferente de intercalação de balanceamento de vários caminho, tenho como desempenho a redução de cópias de arquivos adicionais. Sendo assim, o presente trabalho consiste de 3 partes (**Convert.cpp**, **auxiliar.hpp**, **exrenal\_merge\_sort.cpp**), que trabalhando em conjunto permitem a ordenação do arquivo **property-transfer-statistics-march-2022-quarter-csv.csv**. Por se tratar de um arquivo muito grande teve-se início na ideia pensada anteriormente na primeira etapa, a conversão para binário.

A **Convert.cpp** é responsável por ler o arquivo ‘,csv’ e convertê-lo para o formato binário, trabalhando em conjunto com o **auxiliar.hpp** ele separa cada um dos campos desejados, que são separados por vírgula, e transfere para a estrutura (struct ) encontrada no arquivo **auxiliar.hpp**

O arquivo **auxiliar.hpp** é responsável por guardar os atributos e valores que serão utilizados para ordenar o arquivo, nele temos 14 atributos que são:

1. Series\_reference
2. Period

3. Data\_value
4. Status
5. Units
6. Magnitude
7. Subject
8. Periodicity
9. Group,
- 10.Series\_title\_1
- 11.Series\_title\_2
- 12.Series\_title\_3
- 13.Series\_title\_4
- 14.Series\_title\_5

Na etapa da ordenação será usado apenas 2 desses atributos, que serão o **Series\_reference** e o **Period**, mas por enquanto armazenamos cada um desses dados ao ler as strings, então chega a função sem retorno **rw\_dados**, que irá pegar os vetor de caracter (strings) e adaptá-las para os tipos desejados, como por exemplo o atributo "Period ", que irá ser utilizado na forma float.

Já o arquivo `external_merge_sort.cpp` é o coração da máquina e o responsável pela ordenação do arquivo binário, o método escolhido foi o Merge Sort Externo, que usa a ideia de separar para conquistar, o arquivo é separado em várias partes(blocos) do tamanho da memória interna disponível, sendo a etapa mais importante a ordenação por intercalação.

Os blocos ordenados são distribuídos de forma desigual entre as fitas disponíveis, Obs: apenas uma fita é deixada livre. Com isso, por meio de comparações, ordenar o arquivo de forma crescente.

Abrindo o arquivo de entrada o programa cria pedaços que serão armazenados em um vetor e mais tarde utilizados para a ordenação, após isso se

insere esses pedaços em um vetor e os ordena, após a ordenação o arquivo binário ordenado é criado e o programa é encerrado.