



Trabalho 1 – Análise de dados do covid-19 no Brasil usando métodos de ordenação

Prof. Lilian Berton

São José dos Campos, 2020

Trabalho 1

Passo 1. Selecionar um conjunto de dados sobre covid-19 (de preferência no Brasil)

Ex: Distribuição de Respiradores

URL: http://sage.saude.gov.br/dados/repositorio/distribuicao_respiradores.csv

Ex: Distribuição de Anestésicos e Bloqueadores Musculares

<https://dados.gov.br/dataset/distribuicao-de-anestesicos-e-bloqueadores-musculares1>

Ex: Notificação do coronavírus

URL: <https://covid.saude.gov.br/>
https://dados.gov.br/dataset/notificacao_covid

Ex: Distribuição de Equipamentos de Proteção Individual e Insumos– Covid- 19

<https://dados.gov.br/dataset/distribuicao-de-equipamentos-de-protecao-individual-e-insumos-covid-191>

Também é possível utilizar outros conjuntos de dados nessa temática que você julgar interessante.

Trabalho 1

Pode ser feito
em qualquer
linguagem!

Passo 2. Fazer um programa que lê esses dados de uma planilha .csv ou documento .txt e armazena em uma struct / vários arrays com as informações que você julgar importantes.

Você não precisa armazenar todas as informações da planilha csv, mas aquelas que julgar relevantes para o desenvolvimento de seu trabalho.

Por exemplo: no caso de distribuição de respiradores eu poderia armazenar a data de entrega, o estado, a quantidade e o valor.

DATA	FORNECEDOR	DESTINO	ESTADO/MUNICIPIO	TIPO	QUANTID	VALOR	DESTINAT	UF	DATA DE ENTREGA
19/04/2020	VYAIRE	RIO DE JANEIRO	ESTADO	UTI	40	2400000	ALMOXAR	RJ	19/04/2020
22/04/2020	VYAIRE	RIO DE JANEIRO	ESTADO	UTI	20	1200000	LABORAT	RJ	23/04/2020
08/05/2020	MAGNAMED	RIO DE JANEIRO	ESTADO	UTI	25	1500000	ALMOXAR	RJ	09/05/2020
08/05/2020	KTK	RIO DE JANEIRO	ESTADO	UTI	15	900000	ALMOXAR	RJ	09/05/2020
09/05/2020	LEISTUNG	RIO DE JANEIRO	ESTADO	UTI	6	360000	ALMOXAR	RJ	09/05/2020
09/05/2020	VYAIRE	RIO DE JANEIRO	ESTADO	UTI	1	60000	ALMOXAR	RJ	09/05/2020

Trabalho 1

Passo 3. Ordenar os arrays visando identificar diferentes padrões.

Utilizar três algoritmos de ordenação com diferentes complexidades $O(n^2)$, $O(n \log n)$, $O(n)$. Pode escolher qualquer algoritmo de ordenação.

Exemplo de padrões que poderiam ser analisados no dataset distribuição de respiradores:

- a) Quais foram as maiores/menores quantidades de respiradores distribuídos?
- b) Quais foram as datas em que houve mais/menos entregas de respiradores?
- c) Considerando os valores recebidos por estados, qual estado recebeu o maior montante?
- d) Qual fornecedor fez mais entregas?
- e) Existe alguma anomalia em meus dados?

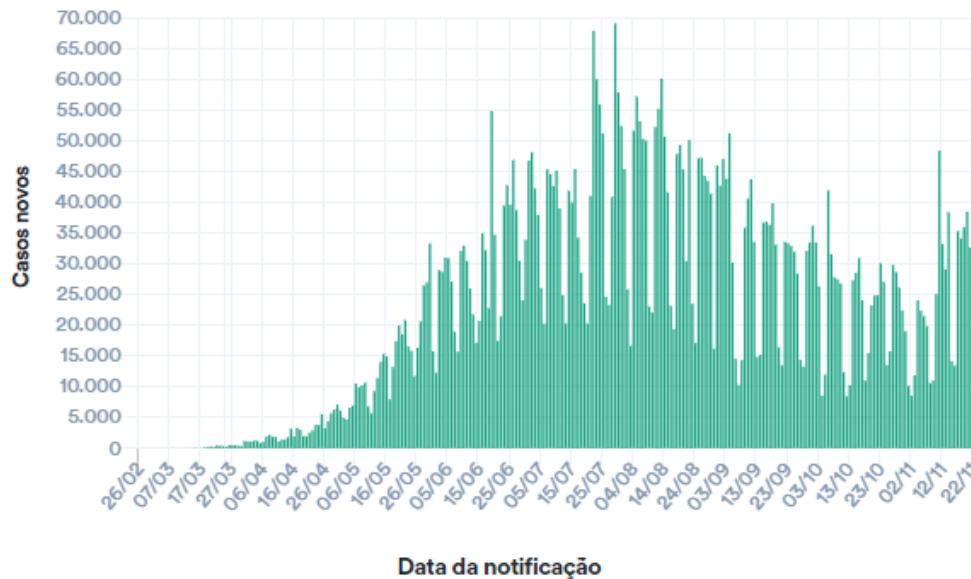
Você pode escolher outras questões e padrões interessantes para analisar em seus dados.

Trabalho 1

Passo 4. Comparar os métodos de ordenação. Calcular o tempo que cada método levou para ordenar cada array e apresentar uma tabela comparativa.

Passo 5. Salvar os vetores ordenados e plotar gráficos a fim de facilitar a visualização dos maiores / menores valores / anomalias em seus dados.

Casos novos de COVID-19 por data de notificação



Pode ser feito
em qualquer
linguagem!

Entrega

- Fazer upload no classroom de uma pasta zip/rar contendo:
- 1 arquivo csv com a base de dados usada: nome_sobrenome_base
- 1 arquivo com os algoritmos implementados: nome_sobrenome_algoritmo
- 1 arquivo .txt com o link no youtube contendo um vídeo de 5 a 15 min explicando como você abordou o problema + compilar e explicar seu código + mostrar os resultados: nome_sobrenome_video
(não será avaliada a qualidade do vídeo, o importante é permitir a visualização de seu código + compilação)
- 1 relatório de ao menos 3 páginas usando o template dupla coluna da IEEE, nome_sobrenome_relatorio
<https://www.ieee.org/conferences/publishing/templates.html>

No relatório comentar sobre os dados que você escolheu, incluir as questões e padrões que você analisou em seus dados, os gráficos gerados, a tabela de tempo comparando a execução dos algoritmos, o hardware utilizado, as referências bibliográficas

Sugestão como escrever um relatório:

Abstract—Resumo do artigo.

I. INTRODUÇÃO

Contexto, problema, citar trabalhos relacionados, explicar sua proposta para resolver o problema, a contribuição do seu trabalho (o que ele tem de diferente dos demais e por que é melhor), organização do trabalho.

II. TRABALHOS RELACIONADOS

Citar e explicar ao menos 5 trabalhos relacionados. Escrever pelo menos 4 linhas para cada trabalho. Um trabalho relacionado é usa exatamente a mesma técnica ou uma técnica similar para resolver exatamente o mesmo problema ou problemas parecidos.

Por exemplo, se seu trabalho é usar redes neurais para identificar rostos (classificação), você pode citar trabalhos que usam redes neurais para identificar placas de carro ou outros objetos. Além disso, você pode citar trabalhos que usam outras técnicas para identificar rostos. Porém, seguindo esse contexto, não comente trabalhos que usem técnicas para classificação de dados que não sejam imagens ou mesmo para a tarefa de regressão.

Fulano [?] propôs ...

Em [?], os autores apresentaram ...

O algoritmo ABC, proposto em [?], ...

III. METODOLOGIA

Aqui, explique sua proposta de trabalho, incluindo pré-processamento dos dados, fluxogramas, algoritmos, diferenças da sua modificação em alguma técnica em relação à técnica original.

Justifique as modificações explicando por que sua proposta deve ser melhor do que os trabalhos relacionados, em termos de qualidade da solução ou custo computacional (memória e tempo).

IV. ANÁLISE EXPERIMENTAL

A. Conjunto de dados

TABLE I

LEGENDA DA TABELA FICA EM CIMA.

B. Configuração do algoritmo e do ambiente computacional

Ambiente (hardware e software): cpu, memória, clock, sistema operacional e versão, compilador e versão, pacotes e versões

C. Critérios de análise

Executar as técnicas ao menos 30x com sementes diferentes para o gerador de números aleatórios, garantindo que os 30 resultados serão diferentes.

Medidas de qualidade e desempenho que serão utilizadas.

D. Resultados e discussão

Apresentar os resultados da sua proposta.

Comparar com outras técnicas que você tenha executado ou com resultados da literatura. Fazer gráficos dos resultados como histogramas, linhas, boxplots, etc.

No caso de tabelas, exibir média, desvio-padrão, mediana. Realizar testes de hipótese para verificar se as diferenças são estatisticamente significativas.

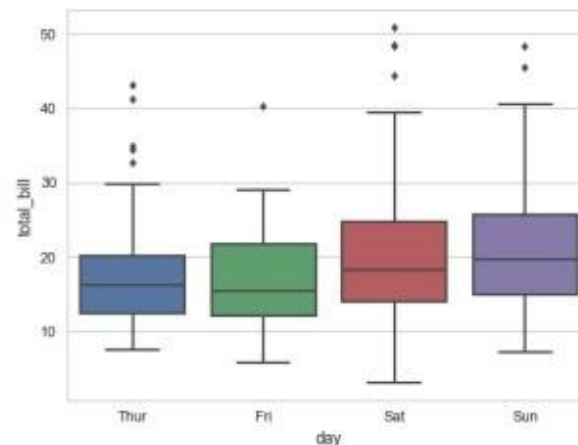


Fig. 1. Legenda da figura fica embaixo.

V. CONCLUSIONS

Contexto, problema, proposta, resultados, trabalhos futuros que podem melhorar a sua metodologia.

Não esqueça de incluir as referências bibliográficas que vc utilizou!

Avaliação

- Data de entrega: 21/12/2020 as 23:59
- Notas
 - (4,0) corretude dos algoritmos (processamento de arquivos, uso de vetores/structs/listas, dos 3 algoritmos de ordenação, bibliotecas, etc).
 - (3,0) quantidade de dados/arrays utilizados, análises dos dados e gráficos gerados.
 - (1,0) vídeo com apresentação do trabalho.
 - (2,0) relatório descrevendo os resultados.

