

TRABALHO FINAL

INFORMAÇÕES GERAIS:

- Esta atividade versa sobre as *Collections* da linguagem Python: *Container datatypes*. Python possui *containers* embutidos: *dict*, *list*, *set* e *tuple*. Python 3 implementa *containers* especializados: *namedtuple*, *deque*, *ChainMap*, *Counter*, *OrderedDict*, *defaultDict*, *UserDict*, *UserList*, *UserString*.
- O trabalho compreende 4 (quatro) questões, que devem ser respondidas através de um texto argumentativo com a justificativa da resposta.
- Nas questões que exigem implementação, descrever os algoritmos utilizados e explicar como executar a aplicação.
- Escrever um relatório técnico (formato pdf, máx. 3 pág) com a solução das questões.
- A apresentação do trabalho deverá ser feita em sala de aula.

QUESTÕES:

1. Descrever as principais características das classes, incluindo vantagens e desvantagens.
2. Fazer uma aplicação Python para criar uma tabela de símbolos usando cada uma dessas classes para armazenar as palavras do arquivo “leipzig100k.txt” que possui cerca de 140.000 palavras. As palavras normalmente são separadas com o caractere `<space>`, `<tab>` e `<newline>`, através do método `split()`. No entanto, podem ser utilizados outros delimitadores como: dígitos e sinais de pontuação. Qual o tempo total gasto para incluir os nomes em cada uma das classes implementadas, SEM considerar o tratamento de palavras repetidas e sem fazer a conversão para letras minúsculas ou maiúsculas? Construir um gráfico com o tempo de execução das aplicações criadas na questão anterior, para avaliar a operação de inclusão nas diversas estruturas.
3. Implementar um trecho de código para consultar 10 palavras, a saber: Lisbon, NASA, Kyunghee, Konkuk, Sogang, momentarily, rubella, vaccinations, government, Authorities. Fazer um gráfico com o tempo médio da execução das 10 consultas nas classes estudadas, para avaliar a operação de busca de dados nessas estruturas.
4. Implementar um trecho de código para excluir 10 palavras, a saber: Lisbon, NASA, Kyunghee, Konkuk, Sogang, momentarily, rubella, vaccinations, government, Authorities. Fazer um gráfico com o tempo médio de execução das 10 exclusões nas classes estudadas, para avaliar a operação de remoção de dados nessas estruturas.

OBS:

- Nas avaliações temporais, tentar ser justo e usar sempre o mesmo computador e, se possível, considerar o tempo médio de pelo menos cinco execuções;
- Não é preciso implementar todas as classes, mas tentar contemplar as mais usadas.