



PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS
Instituto de Ciências Exatas e Informática
Prática sobre Tabelas *hash*

Cursos : *Engenharia de Computação*
Sistemas de Informação
Disciplina : *Algoritmos e Estruturas de Dados*
Professora : *Eveline Alonso Veloso*

Regras Básicas:

1. Estude bastante cada par de entrada/saída fornecido.
2. Todos os programas deverão ser desenvolvidos na linguagem de programação Java.
3. Essas práticas poderão ser desenvolvidas em grupos de, no máximo, três integrantes.
4. Cópias, se existirem, serão encaminhadas ao colegiado de coordenação didática do curso.
5. Fique atento ao *charset* dos arquivos de entrada e saída. Recomenda-se a utilização dos métodos da classe MyIO.java para **leitura de dados do teclado**. É necessário definir o *charset* a ser utilizado antes de começar a leitura de dados do teclado, da seguinte forma: `MyIO.setCharset("UTF-8")`.
6. As saídas esperadas cadastradas no VERDE pela professora foram geradas empregando-se: `System.out.println()`.
7. Para cada exercício, vocês devem submeter apenas um arquivo (.java) por grupo. Essa regra será necessária para a submissão de exercícios no VERDE e no identificador de plágios utilizado na disciplina.
8. A resolução (código) de cada exercício deverá ser submetida ao VERDE.
9. A execução do código submetido será realizada automaticamente pelo VERDE, mas o código será analisado e validado pela professora.

Base de Dados:

[Airbnb](#) é um serviço *online* comunitário que possibilita às pessoas anunciarem, encontrarem e reservarem acomodações e meios de hospedagem.



Foi fundado em Agosto de 2008 por Brian Chesky, Joe Gebbia e Nathan Blecharczyk em São Francisco, na Califórnia. O financiamento inicial foi obtido a partir da incubadora **Y Combinator**. Mais tarde, Greylock Partners, Sequoia Capital e Ashton Kutcher também investiram na empresa.

A plataforma permite aos indivíduos alugarem toda ou parte de sua própria casa, como uma

forma de acomodação extra. O *site* fornece um serviço de busca e reservas entre a pessoa que oferece a acomodação e o turista que busca pela locação. Abrange mais de 500 mil anúncios em mais de 35.000 cidades e 192 países.

No entanto, no mundo, centenas de cidades possuem restrições para locações de curto prazo. Na Europa, cidades como Barcelona, Amsterdã, Paris e Veneza, impuseram restrições à plataforma nesse sentido. Em Paris, por exemplo, os anfitriões não podem alugar seus imóveis por mais de 120 dias por ano e estes devem estar regulares perante a prefeitura e adimplentes com os tributos.

Nos Estados Unidos, cidades como Washington D.C., Los Angeles e Santa Monica, possuem restrições similares.

No Japão a situação não é diferente. A regulamentação nipônica exige registro dos anfitriões para que possam anunciar seus imóveis na plataforma. Ainda assim, limitando em 180 dias por ano o tempo que um imóvel pode ser alugado.

No Brasil ainda não há legislação que regulamente o Airbnb. No entanto, em abril de 2021, o Superior Tribunal de Justiça (STJ) decidiu que condomínios residenciais podem impedir o uso de imóveis para locação pela plataforma. Segundo a Quarta Turma do Tribunal, o sistema de reserva de imóveis pela plataforma digital é caracterizado como uma espécie de contrato atípico de hospedagem – distinto da locação por temporada e da hospedagem oferecida por empreendimentos hoteleiros, que possuem regulamentações próprias.

Exercício:

1. Tabela *hash* com endereçamento em separado

Crie uma tabela *hash* com endereçamento em separado. Em seguida, faça a inserção dos registros correspondentes a algumas acomodações conforme a entrada padrão. A chave de pesquisa será o atributo ***hostId*** (código do anfitrião) da acomodação. Não insira uma acomodação se sua chave de pesquisa já estiver na tabela *hash*. Por fim, pesquise se alguns registros estão armazenados na tabela *hash*, mostrando suas respectivas posições nessa tabela (endereços efetivos).

A função de transformação que deve ser aplicada é: ***(hostId mod tamanho da tabela hash)***, onde o tamanho da tabela *hash* é 7507.

Utilize a classe *Acomodacao* especificada e desenvolvida em prática anterior.

Seu programa deve ler um arquivo-texto chamado ***dados_airbnb.txt*** que, no VERDE, **localiza-se na pasta /tmp**. Você deve preencher um vetor de objetos da classe *Acomodacao* com os dados das diversas acomodações informadas nesse arquivo.

A **primeira linha** desse arquivo deve ser **descartada**, pois apenas informa os diversos **campos** presentes em cada uma das linhas seguintes. Cada uma das

linhas seguintes indica os dados de uma acomodação, separados por uma tabulação ('\t').

A **entrada padrão** é dividida em **duas partes**. A **primeira** contém, em cada linha, uma *string* indicando o **id da acomodação** que deve ser inserida na tabela *hash* de acomodações sobre a qual as pesquisas serão processadas. Após a palavra FIM, inicia-se a **segunda** parte da entrada padrão, que também é composta por várias linhas, sendo que cada uma possui uma *string* indicando o **código do anfitrião (hostId)** da acomodação que deve ser pesquisada na tabela *hash*. A última linha dessa parte também terá a palavra FIM.

A **saída padrão** será composta por várias linhas, uma para cada pesquisa realizada na tabela *hash*. Cada linha deve apresentar a posição de cada acomodação procurada na tabela *hash* (endereço efetivo); e, no final, as palavras **SIM** ou **NAO**, indicando se cada um dos códigos de anfitrião (*hostIds*) pesquisados existe ou não na tabela *hash*.

Além disso, crie um arquivo de *log* na pasta corrente com o nome matrícula_hashSeparado.txt com uma única linha contendo: seu número de matrícula, tempo de execução de seu algoritmo (em milissegundos) e número de comparações realizadas. Todas as informações desse arquivo de *log* devem ser separadas por uma tabulação '\t'.

2. Tabela *hash* com endereçamento aberto e *rehashing*

Crie uma tabela *hash* com endereçamento aberto e *rehashing*. Em seguida, faça a inserção dos registros correspondentes a algumas acomodações conforme a entrada padrão. A chave de pesquisa será o atributo **hostId** (código do anfitrião) da acomodação. Não insira uma acomodação se sua chave de pesquisa já estiver na tabela *hash*. Por fim, pesquise se alguns registros estão armazenados na tabela *hash*, mostrando suas respectivas posições nessa tabela (endereços efetivos).

A função de transformação que deve ser aplicada é: **(hostId mod tamanho da tabela hash + k * (hostId mod 3803)) mod tamanho da tabela hash**, onde o tamanho da tabela *hash* é 7919 e *k* indica o número de tentativas de se inserir o registro na tabela, começando com o valor 0.

Utilize a classe *Acomodacao* especificada e desenvolvida em prática anterior.

Seu programa deve ler um arquivo-texto chamado **dados_airbnb.txt** que, no VERDE, **localiza-se na pasta /tmp**. Você deve preencher um vetor de objetos da classe *Acomodacao* com os dados das diversas acomodações informadas nesse arquivo.

A **primeira linha** desse arquivo deve ser **descartada**, pois apenas informa os diversos **campos** presentes em cada uma das linhas seguintes. Cada uma das linhas seguintes indica os dados de uma acomodação, separados por uma tabulação ('\t').

A **entrada padrão** é dividida em **duas partes**. A **primeira** contém, em cada linha, uma *string* indicando o **id da acomodação** que deve ser inserida na tabela *hash* de acomodações sobre a qual as pesquisas serão processadas. Após a palavra FIM, inicia-se a **segunda** parte da entrada padrão, que também é composta por várias linhas, sendo que cada uma possui uma *string* indicando o **código do anfitrião** (*hostId*) da acomodação que deve ser pesquisada na tabela *hash*. A última linha dessa parte também terá a palavra FIM.

A **saída padrão** será composta por várias linhas, uma para cada pesquisa realizada na tabela *hash*. Cada linha deve apresentar a posição de cada acomodação procurada na tabela *hash* (endereço efetivo); e, no final, as palavras **SIM** ou **NAO**, indicando se cada um dos códigos de anfitrião (*hostIds*) pesquisados existe ou não na tabela *hash*.

Além disso, crie um arquivo de *log* na pasta corrente com o nome matrícula_hashRehashing.txt com uma única linha contendo: seu número de matrícula, tempo de execução de seu algoritmo (em milissegundos) e número de comparações realizadas. Todas as informações desse arquivo de *log* devem ser separadas por uma tabulação '\t'.