

Escolha e baixe um arquivo csv a partir do link e use-o nas questões a seguir. Na resolução desta questão, não enviar os arquivos envolvidos, mas somente os resultados pedidos em arquivo PDF.

Requisitos:

- Gere as somas md5 e sha1 do arquivo encriptado.
- Compacte via linha de comando o arquivo csv original usando compressões zip, gzip, bzip2, rar e 7zip. Depois disso compare os tempos e tamanhos dos arquivos gerados. Dica: para obter o tempo de execução, use o comando time. Exemplo: `time zip iris.zip iris.csv`

Sobre a minha resolução

- Para reproduzir meus testes (caso queira) deixei um script sh onde é baixado um arquivo csv, e nele são realizados:
 - Testes de hashing com md5sum e sha1sum do arquivo data.csv
 - Testes dos tempos de compactação entre cada desses formatos:
 - Zip
 - Gzip
 - Bzip2
 - 7zip
 - Rar
- Ainda quando cada arquivo compactado é criado é medido o tempo e o arquivo compactado é deletado.
- Da mesma maneira o data.csv, caso ele não esteja na sua máquina será baixado e no final do script será excluído.
- Como resultado será gerado dois arquivos com extensão `.txt` com os seguintes resultados:
 - solve-hash.txt: Arquivo contendo as chaves hasheadas que foram geradas em cada um dos scripts (sha1sum e md5sum)
 - solve-compression.txt: Sobre o resultado dos tempos para cada tipo de compactação
- Para executar o scripts rode:

```
bash script.sh
```

Resultados

- Ao abrir o arquivo `solve-hash.txt` você verá os seguintes resultados:

```
--- Hashing data.csv ---
```

```
md5sum: 75c9e7b5dd6548ebf9d879937905931c  data.csv
```

```
sha1sum: aa0cca517cc4261cb28a8b0829e6e75e8ecd589d  data.csv
```

- Ao abrir o arquivo `solve-compression.txt` verá os seguintes resultados:

```
-- Time between each type of compression --
```

```
zip:
```

```
real    0m13,420s
```

```
user    0m13,247s
```

```
sys     0m0,124s
```

```
gzip:
```

```
real    0m11,810s
```

```
user    0m11,716s
```

```
sys     0m0,088s
```

```
bzip2:
```

```
real    0m17,056s
```

```
user    0m16,809s
```

```
sys     0m0,144s
```

```
7zip:
```

```
real    1m38,461s
```

```
user    3m57,986s
```

```
sys     0m1,564s
```

```
rar:
```

```
real    0m32,957s
```

```
user    1m43,964s
```

```
sys     0m0,531s
```

- Com o resultado acima podemos fazer uma tabela para melhor visualização dos dados

| type compression | total time (s) |
|------------------|----------------|
| zip | 13.420 |
| gzip | 11.810 |
| bzip2 | 17.056 |
| 7zip | 98.461 |
| rar | 32.957s |

- Pela tabela acima é perceptível que o melhor tempo de compressão foi o `gzip` com apenas 11.810 segundos e o pior foi o `7zip` com 98.461 segundos
- Acredito que o tempo possa variar de máquina para máquina, mas estes foram meus resultados.