

Contagem de frequência de Bytes em diversos tipos de arquivo

CEFET/RJ– Centro Federal de Educação Tecnológica Celso Suckow da Fonseca

Disciplina de Organização e Estrutura de Arquivos

Natália Nunes Vieira

Resumo: Este estudo tem como objetivo realizar uma análise sobre métodos de compressão de arquivos, onde serão analisadas as frequências de bytes ocorrentes em um arquivo antes e depois dos mesmos serem compactados como *.zip* ou *.rar*.

Introdução

O trabalho consiste em elaborar um código que seja capaz de contar a frequência de bytes de um arquivo, depois compactar estes arquivos e realizar uma nova contagem de frequência. Identificar o grau de compressão de cada um deles e redigir uma documentação com o que foi analisado. Para isso, o desenvolvimento foi dividido em três fases:

Fase 1 (um):

Desenvolver um algoritmo que seja capaz de ler cada Byte de um arquivo e contar a frequência em que eles ocorrem.

O algoritmo precisa:

- Ler uma pasta com diversos tipos de arquivo;
- Ler cada arquivo da pasta individualmente;
- Ler cada Byte do arquivo;
- Identificar qual Byte foi lido;
- Incrementar uma lista com a frequência que cada Byte aparece no arquivo;
- Escrever a lista na tela.

Linguagem: Python

Fase 2 (dois):

Compactar os arquivos utilizando formatos já conhecidos.

Será preciso:

- Escolher pelo menos dois formatos de compactação;
- Compactar cada um dos arquivos;
- Identificar o tamanho anterior e posterior à compactação.

Formato dos arquivos: *.txt*, *.pdf*, *.jpeg*, *.png*, *.mp3*, *.mp4*.

Formatos para compactação: *.rar* e *.zip*

Fase 3 (três):

Comparar a frequência de Bytes dos arquivos antes e depois de serem compactados.

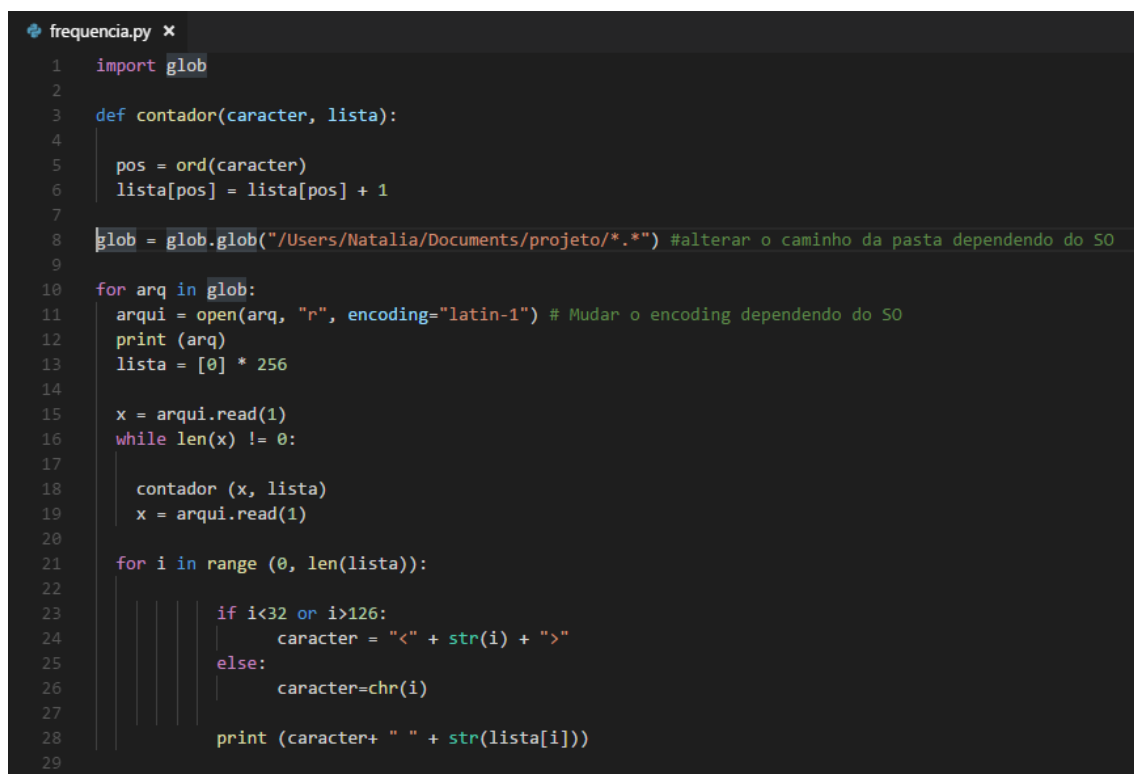
Será preciso:

- Realizar a contagem da frequência de Bytes do arquivo antes da compactação;
- Realizar a contagem da frequência de Bytes do arquivo após a compactação;
- Identificar qual a diferença da contagem entre os dois testes;
- Redigir um documento com o que for identificado entre os arquivos.

Desenvolvimento

Fase 1 (um) - Implementação:

O algoritmo foi desenvolvido na linguagem de programação Python, utilizando a IDE do *Visual Studio Code*. O código encontra-se disponível no *Github* descrito no rodapé do relatório.



```
1  import glob
2
3  def contador(caracter, lista):
4
5      pos = ord(caracter)
6      lista[pos] = lista[pos] + 1
7
8  glob = glob.glob("/Users/Natalia/Documents/projeto/*.") #alterar o caminho da pasta dependendo do SO
9
10 for arq in glob:
11     arqui = open(arq, "r", encoding="latin-1") # Mudar o encoding dependendo do SO
12     print (arq)
13     lista = [0] * 256
14
15     x = arqui.read(1)
16     while len(x) != 0:
17
18         contador (x, lista)
19         x = arqui.read(1)
20
21     for i in range (0, len(lista)):
22
23         if i<32 or i>126:
24             caracter = "<" + str(i) + ">"
25         else:
26             caracter=chr(i)
27
28         print (caracter+ " " + str(lista[i]))
29
```

Figura 1: Contador de frequência de Bytes

Para o código da Figura 1: Contador de frequência de Bytes é importante ressaltar que, para que os testes ocorram com precisão, é necessário alterar o caminho da pasta onde estarão os arquivos para teste, e também verificar o *encoding* utilizado por cada sistema operacional.

```

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto\imagem-png.png
<0> 220
<1> 120
<2> 27
<3> 28
<4> 59
<5> 27
<6> 21
<7> 19
<8> 86
<9> 26
<10> 31
<11> 33
<12> 39
<13> 0
<14> 19
<15> 20
<16> 80
<17> 55

```

Figura 2: Exemplo de saída no console

Para que fosse gerada uma identificação para cada caractere, utilizamos a tabela ASC II, que pode ser encontrada no GitHub informado no rodapé deste documento.

Fase 2 (dois) - Formatos:

Para a realização dos testes, foram escolhidos os formatos de arquivo *.txt*, *.pdf*, *.jpeg*, *.png*, *.mp3*, *.mp4*. E os formatos de compactação *.zip* e *.rar*.







Biblioteca Documentos		
projeto		
Nome	Tipo	Tamanho
 Super Mario Bross - papelão	Vídeo MP4	62.845 KB
 degradê	Imagem JPEG	40 KB
 cores	Imagem PNG	7 KB
 Contos para a infância - Guerra Junqueiro	Documento de Texto	182 KB
 Apenas mais uma de amor - Lulu Santos	Som no Formato MP3	3.406 KB
 A descoberta do mundo - Clarice Lispector	Adobe Acrobat Document	1.718 KB

Figura 3: Arquivos nos formatos selecionados

Para cada arquivo foram utilizados os padrões de compactação abaixo descritos na Figura 4: Especificações de compactação *.rar* e na Figura 5: Especificações de compactação *.zip*.

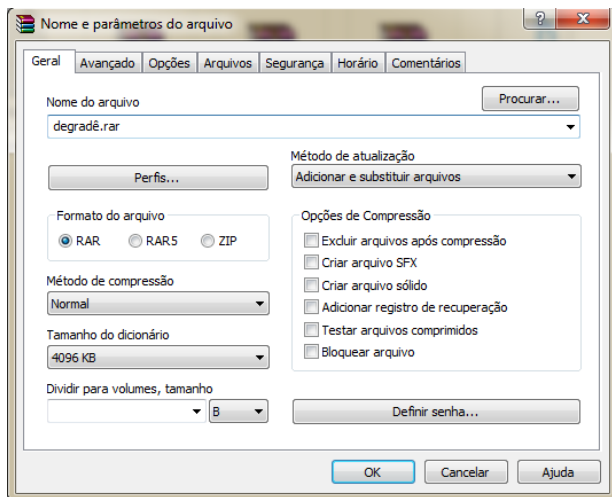


Figura 4: Especificações de compactação .rar

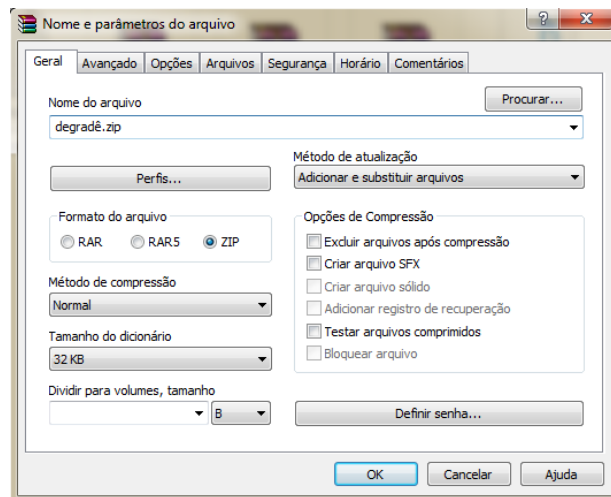


Figura 5: Especificações de compactação .zip

Todos os arquivos foram compactados, conforme a Figura 6: Arquivos originais e compactados.

Biblioteca Documentos		
projeto		
Nome	Tipo	Tamanho
Super Mario Bross - papelão	Arquivo ZIP do WinRAR	62.732 KB
Super Mario Bross - papelão	Arquivo do WinRAR	62.846 KB
Super Mario Bross - papelão	Vídeo MP4	62.845 KB
degradê	Arquivo ZIP do WinRAR	39 KB
degradê	Arquivo do WinRAR	39 KB
degradê	Imagem JPEG	40 KB
cores	Arquivo ZIP do WinRAR	7 KB
cores	Arquivo do WinRAR	7 KB
cores	Imagem PNG	7 KB
Contos para a infância - Guerra Junqueiro	Arquivo ZIP do WinRAR	71 KB
Contos para a infância - Guerra Junqueiro	Documento de Texto	182 KB
Contos para a infância - Guerra Junqueiro	Arquivo do WinRAR	67 KB
Apenas mais uma de amor - Lulu Santos	Arquivo ZIP do WinRAR	3.384 KB
Apenas mais uma de amor - Lulu Santos	Arquivo do WinRAR	3.378 KB
Apenas mais uma de amor - Lulu Santos	Som no Formato MP3	3.406 KB
A descoberta do mundo - Clarice Lispector	Arquivo ZIP do WinRAR	1.621 KB
A descoberta do mundo - Clarice Lispector	Arquivo do WinRAR	1.614 KB
A descoberta do mundo - Clarice Lispector	Adobe Acrobat Document	1.718 KB

Figura 6: Arquivos originais e compactados

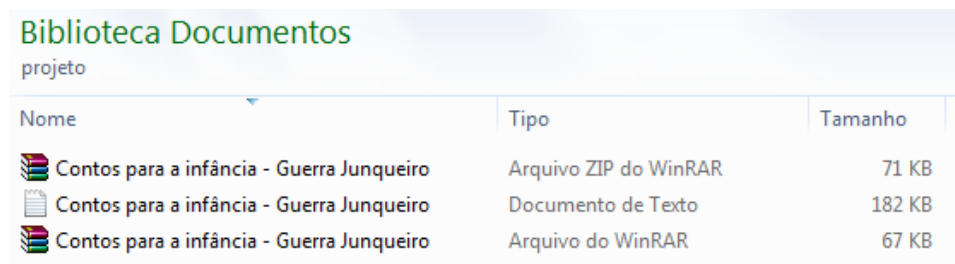
Fase 3 (três) – Testes:

Para a realização dos testes foi utilizada uma máquina com a seguinte configuração: processador *Core i5*, 8GB de memória RAM e um SSD de 120 GB. Os arquivos utilizados para execução dos testes encontram-se disponíveis na pasta do drive informada no rodapé deste relatório.

Abaixo, os resultados identificados em cada teste.

✓ .txt:

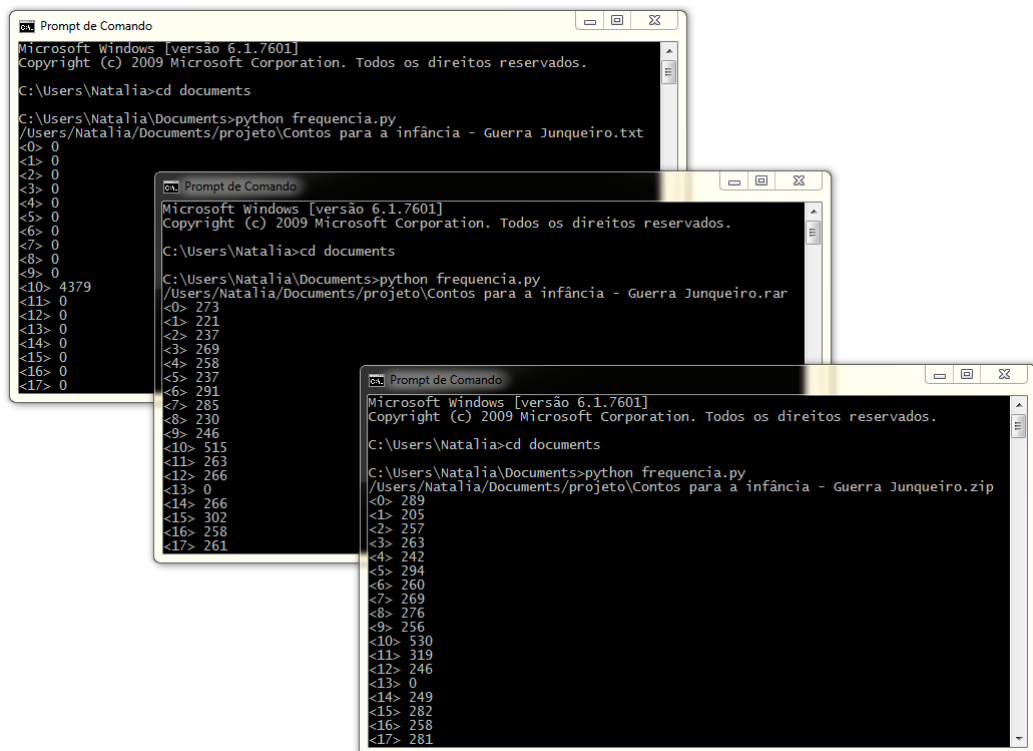
O arquivo utilizado para testar o formato *.txt* é o livro *Contos para Infância*, do autor Guerra Junqueiro e foi obtido através do Projeto Gutenberg.



Nome	Tipo	Tamanho
Contos para a infância - Guerra Junqueiro	Arquivo ZIP do WinRAR	71 KB
Contos para a infância - Guerra Junqueiro	Documento de Texto	182 KB
Contos para a infância - Guerra Junqueiro	Arquivo do WinRAR	67 KB

Figura 7: Arquivo *.txt* original e compactados

Os testes foram realizados para os três formatos de compactação do arquivo, conforme Figura 8: Testes para arquivo *.txt* original e compactado.



```
Microsoft Windows [Versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Contos para a infância - Guerra Junqueiro.txt
<0> 0
<1> 0
<2> 0
<3> 0
<4> 0
<5> 0
<6> 0
<7> 0
<8> 0
<9> 0
<10> 4379
<11> 0
<12> 0
<13> 0
<14> 0
<15> 0
<16> 0
<17> 0

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Contos para a infância - Guerra Junqueiro.rar
<0> 273
<1> 221
<2> 237
<3> 269
<4> 258
<5> 237
<6> 291
<7> 285
<8> 230
<9> 246
<10> 515
<11> 263
<12> 266
<13> 0
<14> 266
<15> 302
<16> 258
<17> 261

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Contos para a infância - Guerra Junqueiro.zip
<0> 289
<1> 205
<2> 257
<3> 263
<4> 242
<5> 294
<6> 260
<7> 269
<8> 276
<9> 256
<10> 530
<11> 319
<12> 246
<13> 0
<14> 249
<15> 282
<16> 258
<17> 281
```

Figura 8: Testes para arquivo *.txt* original e compactado

```

Prompt de Comando
Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto\Contos para a infância - Guerra Junqueiro.txt

<0> 0      27543      A 435      a 17360      <127> 0      <160> 0      <200> 0      <230> 0
<1> 0      282      B 95      b 1458      <128> 0      <161> 0      <201> 50      <231> 474
<2> 0      22      C 153      c 4304      <129> 0      <162> 0      <202> 0      <232> 5
<3> 0      1      D 199      d 6380      <130> 0      <163> 0      <203> 0      <233> 306
<4> 0      2      E 397      e 17229      <131> 0      <164> 0      <204> 0      <234> 47
<5> 0      % 1      F 126      f 1592      <132> 0      <165> 0      <205> 0      <235> 0
<6> 0      & 0      G 132      g 1724      <133> 0      <166> 0      <206> 0      <236> 0
<7> 0      ' 456      H 48      h 2856      <134> 0      <167> 0      <207> 0      <237> 11
<8> 0      ( 25      I 127      i 8094      <135> 0      <168> 0      <208> 0      <238> 0
<9> 0      ) 25      J 70      j 439      <136> 0      <169> 0      <209> 0      <239> 0
<10> 4379      * 132      K 6      k 127      <137> 0      <170> 5      <210> 0      <240> 0
<11> 0      + 0      L 79      l 4416      <138> 0      <171> 166      <211> 2      <241> 0
<12> 0      , 2777      M 203      m 6202      <139> 0      <172> 0      <212> 0      <242> 1
<13> 0      ' 1887      N 240      n 7218      <140> 0      <173> 0      <213> 0      <243> 96
<14> 0      . 1913      O 358      o 14285      <141> 0      <174> 0      <214> 0      <244> 35
<15> 0      / 35      P 249      p 3409      <142> 0      <175> 0      <215> 0      <245> 75
<16> 0      0 22      Q 112      q 1650      <143> 0      <176> 0      <216> 0      <246> 0
<17> 0      1 64      R 107      r 9152      <144> 0      <177> 0      <217> 0      <247> 0
<18> 0      2 14      S 183      s 9842      <145> 0      <178> 0      <218> 0      <248> 0
<19> 0      3 12      T 169      t 6733      <146> 0      <179> 0      <219> 0      <249> 0
<20> 0      4 15      U 94      u 6451      <147> 0      <180> 0      <220> 0      <250> 3
<21> 0      5 13      V 72      v 2190      <148> 0      <181> 0      <221> 0      <251> 0
<22> 0      6 12      W 14      w 260      <149> 0      <182> 0      <222> 0      <252> 0
<23> 0      7 8      X 2      x 274      <150> 0      <183> 0      <223> 0      <253> 0
<24> 0      8 11      Y 29      y 309      <151> 0      <184> 0      <224> 18
<25> 0      9 13      Z 1      z 632      <152> 0      <185> 0      <225> 373
<26> 0      : 191      [ 2      { 4      <153> 0      <186> 0      <226> 3
<27> 0      ; 219      \ 0      | 0      <154> 0      <187> 369      <227> 1069
<28> 0      < 0      ] 2      } 4      <155> 0      <188> 0      <228> 0
<29> 0      = 0      ^ 4      ~ 254      <156> 0      <189> 0      <229> 0
<30> 0      > 0      _ 202      <157> 0      <190> 0
<31> 0      ? 154      ` 0      <158> 0      <191> 0
      @ 2      <159> 0      <192> 0
      <193> 14
      <194> 2
      <195> 0
      <196> 0
      <197> 0
      <198> 0
      <199> 0

```

Figura 9: Resultado da contagem de Bytes do arquivo .txt

```

Prompt de Comando
Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto\Contos para a infância - Guerra Junqueiro.zip

<0> 289      289      A 249      a 275      <127> 263      <161> 281      <191> 252      <221> 295
<1> 205      249      B 257      b 259      <128> 219      <162> 284      <192> 240      <222> 319
<2> 257      255      C 275      c 313      <129> 275      <163> 290      <193> 287      <223> 272
<3> 263      298      D 266      d 273      <130> 259      <164> 272      <194> 295      <224> 296
<4> 242      247      E 254      e 313      <131> 277      <165> 263      <195> 281      <225> 258
<5> 294      274      F 270      f 304      <132> 276      <166> 323      <196> 279      <226> 287
<6> 260      295      G 299      g 304      <133> 307      <167> 294      <197> 279      <227> 260
<7> 269      272      H 277      h 307      <134> 286      <168> 279      <198> 290      <228> 282
<8> 276      290      I 260      i 296      <135> 312      <169> 295      <199> 247      <229> 308
<9> 256      250      J 277      j 279      <136> 271      <170> 270      <200> 279      <230> 318
<10> 530      263      K 255      k 264      <137> 258      <171> 295      <201> 271      <231> 305
<11> 319      331      L 278      l 282      <138> 284      <172> 317      <202> 250      <232> 243
<12> 246      282      M 296      m 260      <139> 276      <173> 286      <203> 317      <233> 277
<13> 0      281      N 287      n 316      <140> 241      <174> 293      <204> 308      <234> 303
<14> 249      284      O 292      o 267      <141> 286      <175> 294      <205> 288      <235> 294
<15> 282      271      P 255      p 284      <142> 288      <176> 280      <206> 305      <236> 290
<16> 258      286      Q 264      q 301      <143> 279      <177> 297      <207> 289      <237> 313
<17> 281      308      R 262      r 275      <144> 266      <178> 330      <208> 274      <238> 278
<18> 296      286      S 304      s 302      <145> 271      <179> 272      <209> 271      <239> 308
<19> 286      283      T 282      t 296      <146> 257      <180> 250      <210> 294      <240> 271
<20> 279      316      U 271      u 335      <147> 298      <181> 265      <211> 288      <241> 287
<21> 275      300      V 285      v 323      <148> 249      <182> 290      <212> 294      <242> 281
<22> 296      308      W 282      w 285      <149> 270      <183> 291      <213> 315      <243> 313
<23> 259      308      X 290      x 254      <150> 286      <184> 280      <214> 292      <244> 294
<24> 279      266      Y 331      y 314      <151> 288      <185> 294      <215> 290      <245> 283
<25> 261      276      Z 290      z 269      <152> 265      <186> 285      <216> 303      <246> 277
<26> 263      311      [ 260      { 270      <153> 283      <187> 321      <217> 299      <247> 278
<27> 301      279      \ 279      | 283      <154> 300      <188> 297      <218> 267      <248> 261
<28> 279      297      ^ 258      } 300      <155> 292      <189> 287      <219> 329      <249> 305
<29> 288      263      _ 267      ~ 254      <156> 267      <190> 283      <220> 273      <250> 265
<30> 257      302      ` 248      <157> 288      <158> 287      <251> 280
<31> 319      @ 240      <159> 269      <159> 269      <252> 281
      <160> 235      <253> 248
      <254> 279
      <255> 250

```

Figura 10: Resultado da contagem de bytes para o arquivo .txt compactado em formato .zip


```

Prompt de Comando
Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents
C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Contos para a infância - Guerra Junqueiro.rar
<0> 273      246      A 268      a 262      <127> 224      <161> 247      <191> 282      <221> 281
<1> 221      ! 262      B 295      b 266      <128> 243      <162> 283      <192> 262      <222> 248
<2> 237      " 296      C 284      c 274      <129> 265      <163> 265      <193> 255      <223> 286
<3> 269      # 267      D 262      d 295      <130> 237      <164> 237      <194> 242      <224> 252
<4> 258      $ 270      E 253      e 286      <131> 280      <165> 267      <195> 289      <225> 282
<5> 237      % 257      F 238      f 271      <132> 275      <166> 281      <196> 244      <226> 297
<6> 291      & 249      G 309      g 267      <133> 261      <167> 286      <197> 268      <227> 273
<7> 285      ' 253      H 282      h 278      <134> 280      <168> 260      <198> 287      <228> 263
<8> 230      ( 244      I 242      i 244      <135> 285      <169> 260      <199> 271      <229> 254
<9> 246      ) 253      J 257      j 274      <136> 274      <170> 271      <200> 244      <230> 275
<10> 515      * 279      K 258      k 281      <137> 264      <171> 257      <201> 280      <231> 240
<11> 263      + 230      L 283      l 272      <138> 301      <172> 243      <202> 282      <232> 251
<12> 266      , 254      M 247      m 262      <139> 300      <173> 311      <203> 252      <233> 272
<13> 0      - 262      N 265      n 329      <140> 257      <174> 237      <204> 265      <234> 276
<14> 266      . 279      O 255      o 264      <141> 289      <175> 284      <205> 277      <235> 241
<15> 302      / 301      P 241      p 237      <142> 335      <176> 257      <206> 285      <236> 287
<16> 258      0 248      Q 298      q 310      <143> 299      <177> 290      <207> 283      <237> 241
<17> 261      1 261      R 286      r 261      <144> 215      <178> 256      <208> 257      <238> 262
<18> 237      2 274      S 261      s 273      <145> 247      <179> 269      <209> 257      <239> 265
<19> 280      3 260      T 276      t 322      <146> 246      <180> 279      <210> 316      <240> 280
<20> 260      4 259      U 275      u 263      <147> 291      <181> 252      <211> 295      <241> 274
<21> 289      5 264      V 259      v 264      <148> 275      <182> 262      <212> 272      <242> 281
<22> 273      6 263      W 276      w 304      <149> 264      <183> 248      <213> 246      <243> 265
<23> 251      7 241      X 267      x 260      <150> 275      <184> 253      <214> 267      <244> 266
<24> 281      8 262      Y 237      y 277      <151> 298      <185> 264      <215> 269      <245> 270
<25> 291      9 278      Z 256      z 300      <152> 263      <186> 253      <216> 270      <246> 271
<26> 264      : 256      [ 270      { 253      <153> 282      <187> 272      <217> 257      <247> 278
<27> 255      ; 281      \ 272      | 280      <154> 256      <188> 252      <218> 264      <248> 287
<28> 254      < 268      ] 299      } 278      <155> 255      <189> 291      <219> 286      <249> 262
<29> 280      = 271      ^ 291      ~ 271      <156> 273      <190> 244      <220> 257      <250> 262
<30> 257      > 270      _ 292      <157> 301      <251> 257
<31> 282      ? 277      ` 237      <158> 261      <252> 247
      @ 241      <159> 275      <253> 271
      <160> 268      <254> 272
      <255> 268

```

Figura 11: Resultado da contagem de bytes para o arquivo .txt compactado em formato .rar

Na Figura 7: Arquivo .txt original e compactados, podemos perceber a diferença de tamanho para os arquivos compactados, enquanto a compactação em .zip reduziu o arquivo de 182KB para 71KB, a compactação do .rar reduziu o mesmo arquivo para 67KB. Para a análise da frequência de bytes, do arquivo .txt utilizamos, após identificação dos caracteres que apareceram com maior frequência, os caracteres de *espaço* e *'a'* para comparação entre os três tipos de arquivo.

Percebemos que, para o arquivo não compactado, o caractere de *espaço* apareceu 27543 vezes, para 289 vezes no formato .zip e 246 vezes no formato .rar; para o caractere *'a'*, a ocorrência identificada no arquivo original foi de 17360 vezes, para 275 vezes no .zip e 262 vezes no .rar.

	.txt	.zip	.rar	Frequência de Bytes: Arquivo original/Compactado	
				.zip	.rar
<i>espaço</i>	27543	289	246	95,30	111,96
<i>a</i>	17360	275	262	63,13	66,26

Sendo assim, é possível identificar que, após a compactação do arquivo, tanto pela frequência de bytes, quanto pelo tamanho dos arquivos, para um formato .txt, o formato de compactação .rar compacta melhor que o .zip.

✓ .pdf:

Para os testes com arquivo pdf foi selecionado o livro *A Descoberta do Mundo*, de autora Clarice Lispector. O arquivo foi disponibilizado pela equipe do *LeLivros*, para uso parcial em pesquisas e estudos acadêmicos.




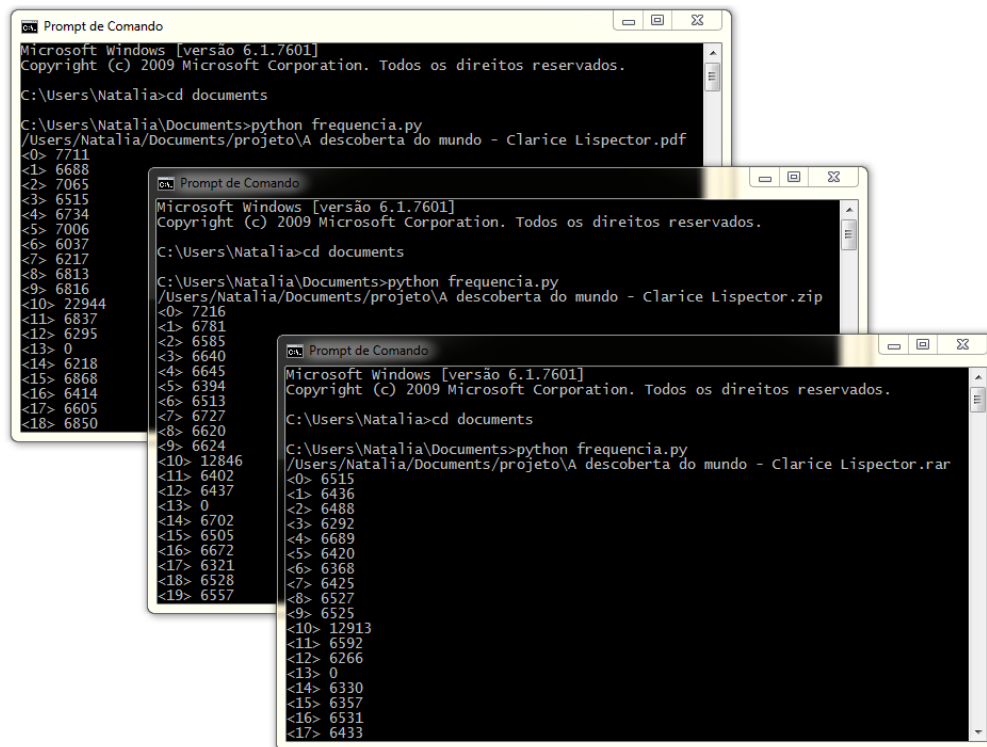
Biblioteca Documentos		
projeto		
Nome	Tipo	Tamanho
 A descoberta do mundo - Clarice Lispector	Arquivo ZIP do WinRAR	1.621 KB
 A descoberta do mundo - Clarice Lispector	Arquivo do WinRAR	1.614 KB
 A descoberta do mundo - Clarice Lispector	Adobe Acrobat Document	1.718 KB

Figura 12: Arquivo .pdf original e compactados

Os testes foram realizados para os três formatos do arquivo, conforme Figura 13: Compactação de arquivo .pdf.



```
Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/A descoberta do mundo - Clarice Lispector.pdf
<0> 7711
<1> 6688
<2> 7065
<3> 6515
<4> 6734
<5> 7006
<6> 6037
<7> 6217
<8> 6813
<9> 6816
<10> 22944
<11> 6837
<12> 6295
<13> 0
<14> 6218
<15> 6868
<16> 6414
<17> 6605
<18> 6850

Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/A descoberta do mundo - Clarice Lispector.zip
<0> 7216
<1> 6781
<2> 6585
<3> 6640
<4> 6645
<5> 6394
<6> 6513
<7> 6727
<8> 6620
<9> 6624
<10> 12846
<11> 6402
<12> 6437
<13> 0
<14> 6702
<15> 6505
<16> 6672
<17> 6321
<18> 6528
<19> 6557

Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/A descoberta do mundo - Clarice Lispector.rar
<0> 6515
<1> 6436
<2> 6488
<3> 6292
<4> 6689
<5> 6420
<6> 6368
<7> 6425
<8> 6527
<9> 6525
<10> 12913
<11> 6592
<12> 6266
<13> 0
<14> 6330
<15> 6357
<16> 6531
<17> 6433
```

Figura 13: Compactação de arquivo .pdf


```
Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/A descoberta do mundo - Clarice Lispector.pdf

<0> 7711      24460      A 6189      a 9061      <127> 6378      <161> 6517      <191> 6715      <221> 7101
<1> 6688      ! 6366      B 6953      b 6945      <128> 6645      <162> 5620      <192> 6682      <222> 6223
<2> 7065      " 6762      C 6512      c 6793      <129> 6866      <163> 5631      <193> 6609      <223> 6778
<3> 6515      # 6574      D 6884      d 7874      <130> 6717      <164> 5976      <194> 6450      <224> 7070
<4> 6734      $ 6475      E 6590      e 12206     <131> 5945      <165> 6152      <195> 6944      <225> 7220
<5> 7006      % 6705      F 7761      f 6266      <132> 6739      <166> 6226      <196> 6127      <226> 6824
<6> 6037      & 5858      G 6472      g 6799      <133> 6531      <167> 6914      <197> 6114      <227> 7129
<7> 6217      ' 6243      H 6469      h 6285      <134> 6076      <168> 6355      <198> 6134      <228> 6536
<8> 6813      ( 6333      I 5964      i 6793      <135> 6858      <169> 6293      <199> 6890      <229> 6939
<9> 6816      ) 6026      J 6359      j 7615      <136> 6067      <170> 6732      <200> 6647      <230> 6404
<10> 22944     * 6635      K 6308      k 6069      <137> 6083      <171> 6310      <201> 6380      <231> 6892
<11> 6837      + 6493      L 6726      l 6607      <138> 6398      <172> 5899      <202> 6306      <232> 5826
<12> 6295      , 6313      M 6201      m 7110      <139> 6105      <173> 5912      <203> 6786      <233> 6751
<13> 0          . 6400      N 6364      n 10451     <140> 6228      <174> 6637      <204> 5973      <234> 6287
<14> 6218      / 6819      O 7428      o 9403      <141> 6008      <175> 6812      <205> 6447      <235> 6332
<15> 6868      : 11758     P 7350      p 8236      <142> 6472      <176> 6273      <206> 6420      <236> 6222
<16> 6414      ; 0 16519   Q 5787      q 6793      <143> 6588      <177> 5978      <207> 6734      <237> 6682
<17> 6605      <1 8570     R 8866      r 8468      <144> 6336      <178> 6065      <208> 6360      <238> 6959
<18> 6850      2 8177     S 6728      s 8226      <145> 6569      <179> 5977      <209> 5558      <239> 6187
<19> 6332      3 7482     T 7001      t 10464     <146> 6222      <180> 6126      <210> 5902      <240> 7259
<20> 6304      4 7910     U 6903      u 6620      <147> 6070      <181> 6150      <211> 6966      <241> 7008
<21> 6709      5 8151     V 6040      v 6583      <148> 6202      <182> 6378      <212> 6193      <242> 6767
<22> 6473      6 7232     W 6685      w 6635      <149> 6454      <183> 6861      <213> 6198      <243> 7003
<23> 6475      7 7934     X 6187      x 7977      <150> 6184      <184> 7128      <214> 5490      <244> 6395
<24> 5823      8 7965     Y 6119      y 7218      <151> 6906      <185> 6831      <215> 7014      <245> 6188
<25> 6355      9 7370     Z 5853      z 6191      <152> 5621      <186> 6415      <216> 6018      <246> 6143
<26> 5844      : 5979     [ 6857      { 6101      <153> 6354      <187> 7065      <217> 5946      <247> 6845
<27> 6092      ; 6202     \ 7031      | 6894      <154> 6427      <188> 6738      <218> 6248      <248> 6891
<28> 6631      < 10758    ] 7407      } 6626      <155> 6543      <189> 6015      <219> 6867      <249> 6788
<29> 5935      = 6276     ^ 6713      ~ 7332      <156> 6839      <190> 6988      <220> 6851      <250> 6389
<30> 6551      > 9801     _ 7007      <157> 6054      <158> 6845      <251> 6920
<31> 6835      ? 7013     ` 6301      <159> 6722      <252> 7012
                                <160> 6253      <253> 6774
                                <254> 6301
                                <255> 5690
```

Figura 14: Resultado da contagem de bytes para o arquivo .pdf

```
Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/A descoberta do mundo - Clarice Lispector.rar

<0> 6515      6633      A 6602      a 6464      <127> 6645      <161> 6278      <191> 6849      <221> 6667
<1> 6436      ! 6435      B 6498      b 6382      <128> 6419      <162> 6376      <192> 6360      <222> 6608
<2> 6488      " 6308      C 6227      c 6335      <129> 6440      <163> 6372      <193> 6512      <223> 6839
<3> 6292      # 6469      D 6335      d 6323      <130> 6418      <164> 6311      <194> 6404      <224> 6606
<4> 6689      $ 6454      E 6380      e 6380      <131> 6491      <165> 6506      <195> 6412      <225> 6571
<5> 6420      % 6390      F 6453      f 6408      <132> 6362      <166> 6375      <196> 6274      <226> 6450
<6> 6368      & 6275      G 6301      g 6288      <133> 6437      <167> 6400      <197> 6517      <227> 6452
<7> 6425      ' 6389      H 6485      h 6254      <134> 6346      <168> 6305      <198> 6257      <228> 6399
<8> 6527      ( 6605      I 6316      i 6527      <135> 6389      <169> 6380      <199> 6709      <229> 6572
<9> 6525      ) 6263      J 6386      j 6352      <136> 6378      <170> 6277      <200> 6552      <230> 6423
<10> 12913    * 6393      K 6426      k 6198      <137> 6427      <171> 6416      <201> 6367      <231> 6721
<11> 6592      + 6499      L 6402      l 6584      <138> 6378      <172> 6373      <202> 6312      <232> 6448
<12> 6266      , 6375      M 6367      m 6500      <139> 6467      <173> 6406      <203> 6506      <233> 6581
<13> 0          . 6537      N 6531      n 6534      <140> 6358      <174> 6390      <204> 6337      <234> 6276
<14> 6330      / 6466      O 6489      o 6541      <141> 6265      <175> 6515      <205> 6314      <235> 6640
<15> 6357      : 6403      P 6404      p 6543      <142> 6371      <176> 6307      <206> 6411      <236> 6481
<16> 6531      ; 0 6456   Q 6487      q 6417      <143> 6526      <177> 6373      <207> 6627      <237> 6746
<17> 6433      <1 6332     R 6437      r 6393      <144> 6425      <178> 6361      <208> 6214      <238> 6472
<18> 6498      2 6374     S 6370      s 6459      <145> 6426      <179> 6368      <209> 6379      <239> 6840
<19> 6461      3 6342     T 6334      t 6255      <146> 6361      <180> 6383      <210> 6317      <240> 6538
<20> 6296      4 6408     U 6421      u 6469      <147> 6353      <181> 6409      <211> 6731      <241> 6335
<21> 6561      5 6268     V 6457      v 6519      <148> 6437      <182> 6511      <212> 6281      <242> 6454
<22> 6414      6 6388     W 6260      w 6476      <149> 6434      <183> 6677      <213> 6454      <243> 6546
<23> 6560      7 6393     X 6389      x 6361      <150> 6366      <184> 6532      <214> 6329      <244> 6609
<24> 6218      8 6355     Y 6388      y 6466      <151> 6486      <185> 6517      <215> 6579      <245> 6599
<25> 6335      9 6458     Z 6307      z 6458      <152> 6291      <186> 6339      <216> 6491      <246> 6596
<26> 6230      : 6451     [ 6416      { 6507      <153> 6450      <187> 6535      <217> 6485      <247> 6770
<27> 6566      ; 6557     \ 6362      | 6488      <154> 6340      <188> 6456      <218> 6520      <248> 6638
<28> 6374      < 6331     ] 6391      } 6782      <155> 6468      <189> 6483      <219> 6929      <249> 6523
<29> 6429      = 6401     ^ 6594      ~ 6709      <156> 6406      <190> 6586      <220> 6627      <250> 6633
<30> 6426      > 6431     _ 6596      <157> 6402      <158> 6354      <251> 6713
<31> 6565      ? 6728     ` 6332      <159> 6663      <159> 6663      <252> 6669
                                <160> 6405      <253> 6656
                                <254> 6828
                                <255> 6642
```

Figura 15: Resultado da contagem de bytes para o arquivo .pdf compactado em formato .rar

```

C:\Users\Natalia>cd documents
C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/A descoberta do mundo - Clarice Lispector.zip
<0> 7216      6822  A 6510  a 6070  <127> 6732  <161> 6611  <191> 6571  <221> 6620
<1> 6781      6671  B 6559  b 6347  <128> 6725  <162> 6226  <192> 6519  <222> 6382
<2> 6585      6429  C 6729  c 6265  <129> 6591  <163> 6357  <193> 6347  <223> 6912
<3> 6640      6432  D 6332  d 6336  <130> 6428  <164> 6394  <194> 6371  <224> 6450
<4> 6645      6514  E 6180  e 6463  <131> 6376  <165> 6178  <195> 6419  <225> 6362
<5> 6394      6223  F 6081  f 6320  <132> 6468  <166> 6170  <196> 6274  <226> 6389
<6> 6513      6385  G 6469  g 6436  <133> 6285  <167> 6560  <197> 6081  <227> 6797
<7> 6727      6555  H 6453  h 6291  <134> 6360  <168> 6551  <198> 6278  <228> 6408
<8> 6620      6393  I 6408  i 6382  <135> 6575  <169> 6482  <199> 6758  <229> 6639
<9> 6624      6306  J 6387  j 6399  <136> 6471  <170> 6500  <200> 6470  <230> 6369
<10> 12846    6518  K 6324  k 6186  <137> 6371  <171> 6361  <201> 6245  <231> 6985
<11> 6402      6211  L 6396  l 6489  <138> 6220  <172> 6190  <202> 6400  <232> 6456
<12> 6437      6192  M 6282  m 6606  <139> 6208  <173> 6451  <203> 6518  <233> 6516
<13> 0         6167  N 6712  n 6498  <140> 6175  <174> 6510  <204> 6331  <234> 6449
<14> 6702      6310  O 6522  o 6305  <141> 6307  <175> 6529  <205> 6454  <235> 6821
<15> 6505      6249  P 6600  p 6395  <142> 6694  <176> 6234  <206> 6581  <236> 6599
<16> 6672      6322  Q 6267  q 6384  <143> 6778  <177> 6335  <207> 6681  <237> 6748
<17> 6321      6224  R 6298  r 6517  <144> 6766  <178> 6297  <208> 6727  <238> 6774
<18> 6528      2 6362  S 6340  s 6435  <145> 6339  <179> 6331  <209> 6274  <239> 7150
<19> 6557      3 6214  T 6482  t 6641  <146> 6313  <180> 6409  <210> 6486  <240> 6606
<20> 6528      4 6478  U 6567  u 6620  <147> 6474  <181> 6407  <211> 6582  <241> 6436
<21> 6298      5 6149  V 6275  v 6848  <148> 6402  <182> 6669  <212> 6323  <242> 6350
<22> 6108      6 6432  W 6374  w 6474  <149> 6381  <183> 6660  <213> 6358  <243> 6695
<23> 6218      7 6278  X 6181  x 6511  <150> 6394  <184> 6309  <214> 6476  <244> 6492
<24> 6295      8 6604  Y 6355  y 6575  <151> 6545  <185> 6472  <215> 6722  <245> 6504
<25> 6182      9 6643  Z 6287  z 6560  <152> 6426  <186> 6530  <216> 6443  <246> 6678
<26> 6345      6628  [ 6367  { 6679  <153> 6396  <187> 6617  <217> 6544  <247> 7125
<27> 6394      6613  \ 6195  | 6713  <154> 6266  <188> 6319  <218> 6474  <248> 6569
<28> 6510      6652  ] 6482  } 6646  <155> 6411  <189> 6468  <219> 6951  <249> 6693
<29> 6714      6499  ^ 6358  ~ 6694  <156> 6431  <190> 6556  <220> 6435  <250> 6526
<30> 6660      6710  _ 6387  <157> 6690  <251> 6984
<31> 6580      6833  ` 6323  <158> 6710  <252> 6518
      6922      <159> 6648  <253> 6834
      <160> 6676  <254> 6623
      <255> 7036

```

Figura 16: Resultado da contagem de bytes para o arquivo .pdf compactado em formato .zip

A Figura 12: Arquivo .pdf original e compactados, nos permite identificar a redução de tamanho para os arquivos .pdf, compactados em .zip e .rar. Ao comparar os valores que em .pdf foram de 1718 KB, em .zip teve uma redução para 1621KB e 1614KB em .rar.

Para analisar a frequência de Bytes, os caracteres escolhidos foram o numeral '0' e a letra 'e', onde o '0' aparecia 16519 vezes no arquivo .pdf e passou a aparecer 6322 vezes no .zip e 6456 no .rar; e a letra 'e' aparecia 12206 vezes no arquivo .pdf e reduziu para 6463 aparecimentos no .zip e 6380 no .rar.

	.pdf	.zip	.rar	Frequência de Bytes: Arquivo original/Compactado	
				.zip	.rar
0	16519	6322	6456	2,61	2,56
e	12206	6463	6380	1,89	1,91

Após análise do tamanho dos arquivos e da frequência dos bytes '0' e 'e' pudemos perceber que, apesar do arquivo original ter sido compacto, o formato de arquivo .pdf não teve uma redução significativa em nenhum dos dois formatos.

✓ .jpg:

A imagem abaixo utilizada para os testes em *.jpg* foi obtida através do *Google Imagens* e está disponível para download no *Pinterest*.

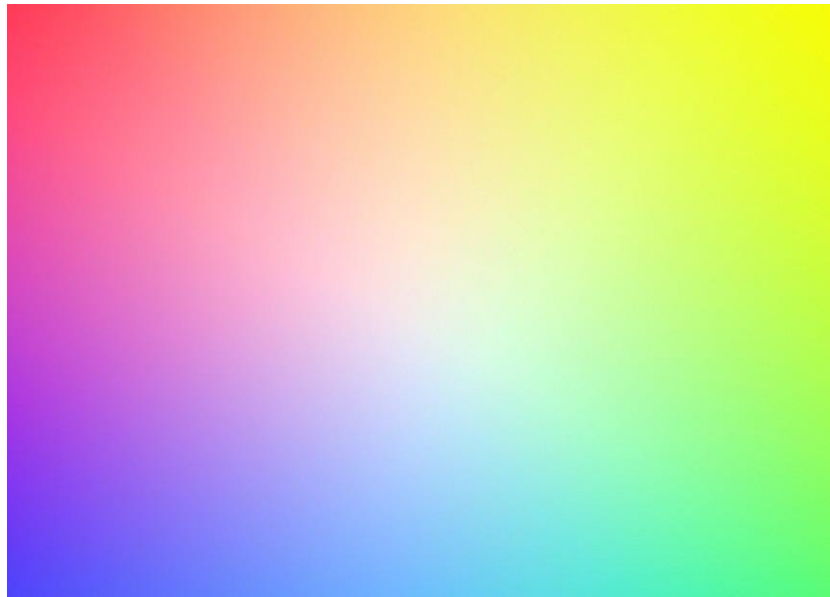


Figura 17: Imagem utilizada para o teste de contagem de frequência de Byte no formato *.jpg*




Biblioteca Documentos		
projeto		
Nome	Tipo	Tamanho
 degradê	Arquivo ZIP do WinRAR	39 KB
 degradê	Arquivo do WinRAR	39 KB
 degradê	Imagem JPEG	40 KB

Figura 18: Arquivo *.jpg* original e compactados

Os testes foram realizados para os três formatos do arquivo, conforme Figura 19: Compactação de arquivo *.jpg*.

```

Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/degradê.jpg
<0> 216
<1> 122
<2> 103
<3> 109
<4> 90
<5> 183
<6> 278
<7> 59
<8> 91
<9> 98
<10> 333
<11> 146
<12> 257
<13> 0
<14> 77
<15> 74
<16> 54
<17> 202
<18> 145
<19> 341
<20> 107

Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/degradê.zip
<0> 191
<1> 158
<2> 130
<3> 159
<4> 141
<5> 117
<6> 140
<7> 148
<8> 141
<9> 158
<10> 338
<11> 136
<12> 177
<13> 0
<14> 160
<15> 165
<16> 134
<17> 141
<18> 159
<19> 124
<20> 141

Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/degradê.rar
<0> 168
<1> 145
<2> 138
<3> 141
<4> 142
<5> 123
<6> 163
<7> 178
<8> 135
<9> 150
<10> 299
<11> 150
<12> 169
<13> 0
<14> 154
<15> 158
<16> 143
<17> 126
<18> 159
<19> 167
<20> 146

```

Figura 19: Compactação de arquivo .jpg

```

Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/degradê.jpg
<0> 216 40 A 107 a 168 <127> 123 <161> 141 <191> 142 <221> 84
<1> 122 ! 85 B 81 b 288 <128> 83 <162> 275 <192> 133 <222> 101
<2> 103 " 168 C 134 c 243 <129> 127 <163> 183 <193> 295 <223> 119
<3> 109 # 93 D 196 d 121 <130> 163 <164> 132 <194> 88 <224> 125
<4> 90 $ 59 E 241 e 131 <131> 217 <165> 186 <195> 157 <225> 128
<5> 183 % 104 F 205 f 192 <132> 82 <166> 356 <196> 313 <226> 147
<6> 278 & 324 G 70 g 113 <133> 106 <167> 264 <197> 137 <227> 117
<7> 59 ' 178 H 95 h 228 <134> 231 <168> 290 <198> 337 <228> 50
<8> 91 ( 145 I 131 i 239 <135> 67 <169> 349 <199> 75 <229> 128
<9> 98 ) 179 J 208 j 325 <136> 132 <170> 388 <200> 90 <230> 23
<10> 333 * 336 K 103 k 296 <137> 328 <171> 206 <201> 97 <231> 62
<11> 146 + 110 L 275 l 272 <138> 257 <172> 146 <202> 160 <232> 147
<12> 257 , 204 M 383 m 312 <139> 163 <173> 297 <203> 80 <233> 178
<13> 0 - 156 N 302 n 121 <140> 224 <174> 128 <204> 111 <234> 133
<14> 77 . 55 O 88 o 192 <141> 303 <175> 121 <205> 114 <235> 100
<15> 74 / 86 P 128 p 120 <142> 53 <176> 227 <206> 132 <236> 88
<16> 54 0 242 Q 276 q 123 <143> 65 <177> 273 <207> 64 <237> 145
<17> 202 1 318 R 178 r 93 <144> 48 <178> 46 <208> 130 <238> 26
<18> 145 2 196 S 345 s 73 <145> 115 <179> 115 <209> 210 <239> 61
<19> 341 3 183 T 453 t 203 <146> 58 <180> 239 <210> 164 <240> 137
<20> 107 4 182 U 315 u 123 <147> 165 <181> 318 <211> 251 <241> 160
<21> 276 5 297 V 169 v 124 <148> 157 <182> 201 <212> 176 <242> 63
<22> 281 6 419 W 169 w 54 <149> 224 <183> 193 <213> 289 <243> 28
<23> 69 7 121 X 218 x 124 <150> 100 <184> 89 <214> 316 <244> 118
<24> 319 8 144 Y 115 y 59 <151> 90 <185> 97 <215> 87 <245> 112
<25> 194 9 68 Z 258 z 99 <152> 180 <186> 83 <216> 279 <246> 105
<26> 265 : 210 [ 223 { 52 <153> 215 <187> 40 <217> 83 <247> 28
<27> 236 ; 133 \ 140 } 88 <154> 208 <188> 136 <218> 322 <248> 179
<28> 60 < 61 ] 40 ~ 111 <155> 316 <189> 113 <219> 140 <249> 32
<29> 94 = 67 ^ 113 <156> 149 <190> 70 <220> 65 <250> 163
<30> 53 > 61 _ 78 <157> 293 <251> 51
<31> 73 ? 104 ` 312 <158> 56 <252> 85
@ 52 <159> 94 <253> 121
<160> 107 <254> 108
<255> 109

```

Figura 20: Resultado da contagem de bytes para o arquivo .jpg


```

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto\degradê.zip

<0> 191 129 A 148 [ 175 <127> 169 <161> 152 <191> 176 <221> 195
<1> 158 ! 141 B 153 \ 155 <128> 134 <162> 130 <192> 138 <222> 167
<2> 130 " 152 C 146 ] 156 <129> 141 <163> 162 <193> 156 <223> 178
<3> 159 # 133 D 179 ^ 152 <130> 143 <164> 161 <194> 157 <224> 158
<4> 141 $ 144 E 150 _ 162 <131> 177 <165> 137 <195> 148 <225> 180
<5> 117 % 146 F 135 ` 151 <132> 164 <166> 159 <196> 143 <226> 147
<6> 140 & 157 G 167 a 167 <133> 140 <167> 161 <197> 154 <227> 144
<7> 148 ' 179 H 146 b 145 <134> 137 <168> 128 <198> 162 <228> 159
<8> 141 ( 141 I 152 c 137 <135> 153 <169> 156 <199> 143 <229> 173
<9> 158 ) 180 J 150 d 152 <136> 158 <170> 158 <200> 134 <230> 154
<10> 338 * 145 K 149 e 148 <137> 171 <171> 162 <201> 135 <231> 173
<11> 136 + 129 L 148 f 131 <138> 126 <172> 152 <202> 161 <232> 147
<12> 177 , 154 M 185 g 130 <139> 166 <173> 144 <203> 144 <233> 153
<13> 0 - 163 N 159 h 145 <140> 113 <174> 177 <204> 146 <234> 176
<14> 160 . 172 O 135 i 150 <141> 147 <175> 155 <205> 140 <235> 185
<15> 165 / 149 P 135 j 158 <142> 156 <176> 143 <206> 165 <236> 156
<16> 134 0 151 Q 153 k 160 <143> 164 <177> 155 <207> 157 <237> 167
<17> 141 1 129 R 136 l 170 <144> 163 <178> 158 <208> 170 <238> 147
<18> 159 2 135 S 144 m 163 <145> 131 <179> 137 <209> 166 <239> 167
<19> 124 3 154 T 147 n 158 <146> 156 <180> 163 <210> 123 <240> 177
<20> 141 4 155 U 166 o 156 <147> 163 <181> 140 <211> 164 <241> 139
<21> 170 5 148 V 147 p 157 <148> 174 <182> 147 <212> 144 <242> 140
<22> 143 6 162 w 159 q 143 <149> 163 <183> 167 <213> 134 <243> 143
<23> 188 7 172 X 131 r 144 <150> 178 <184> 158 <214> 148 <244> 143
<24> 138 8 155 Y 150 s 178 <151> 160 <185> 173 <215> 157 <245> 136
<25> 145 9 141 Z 130 t 145 <152> 135 <186> 171 <216> 169 <246> 154
<26> 151 : 150 [ 175 u 165 <153> 136 <187> 161 <217> 161 <247> 155
<27> 154 ; 172 \ 155 v 149 <154> 145 <188> 141 <218> 168 <248> 163
<28> 147 < 154 ] 156 w 166 <155> 185 <189> 175 <219> 145 <249> 166
<29> 147 = 164 ^ 152 x 156 <156> 145 <190> 157 <220> 142 <250> 184
<30> 153 > 157 _ 162 y 184 <157> 149 <251> 162
<31> 165 ? 167 ` 151 z 168 <158> 142 <252> 168
@ 132 _ 187 <159> 184 <253> 164
_ 181 <160> 136 <254> 185
j 150 <255> 167
~ 163
_ 163

```

Figura 21: Resultado da contagem de bytes para o arquivo .jpg compactado em formato .zip

```

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto\degradê.rar

<0> 168 140 A 149 a 166 <127> 159 <161> 138 <191> 172 <221> 156
<1> 145 ! 158 B 180 b 141 <128> 127 <162> 135 <192> 150 <222> 165
<2> 138 " 157 C 156 c 154 <129> 133 <163> 141 <193> 152 <223> 159
<3> 141 # 165 D 148 d 159 <130> 135 <164> 147 <194> 146 <224> 150
<4> 142 $ 163 E 151 e 144 <131> 136 <165> 140 <195> 148 <225> 163
<5> 123 % 152 F 152 f 140 <132> 135 <166> 144 <196> 157 <226> 171
<6> 163 & 141 G 145 g 151 <133> 165 <167> 163 <197> 144 <227> 154
<7> 178 ' 148 H 145 h 125 <134> 166 <168> 146 <198> 155 <228> 158
<8> 135 ( 131 I 128 i 138 <135> 164 <169> 136 <199> 155 <229> 159
<9> 150 ) 170 J 131 j 161 <136> 142 <170> 156 <200> 143 <230> 156
<10> 299 * 136 K 169 k 162 <137> 142 <171> 156 <201> 161 <231> 163
<11> 150 + 161 L 151 l 125 <138> 130 <172> 165 <202> 148 <232> 132
<12> 169 , 164 M 132 m 138 <139> 159 <173> 166 <203> 156 <233> 171
<13> 0 - 157 N 149 n 150 <140> 145 <174> 157 <204> 143 <234> 149
<14> 154 . 146 O 155 o 161 <141> 135 <175> 182 <205> 175 <235> 173
<15> 158 / 166 P 131 p 134 <142> 150 <176> 163 <206> 172 <236> 164
<16> 143 0 164 Q 147 q 163 <143> 162 <177> 144 <207> 143 <237> 164
<17> 126 1 138 R 142 r 124 <144> 134 <178> 156 <208> 153 <238> 164
<18> 159 2 156 S 140 s 172 <145> 166 <179> 157 <209> 167 <239> 148
<19> 167 3 128 T 143 t 133 <146> 149 <180> 161 <210> 151 <240> 142
<20> 146 4 126 U 150 u 137 <147> 122 <181> 161 <211> 145 <241> 140
<21> 138 5 140 V 132 v 156 <148> 143 <182> 170 <212> 148 <242> 155
<22> 153 6 150 W 185 w 169 <149> 147 <183> 151 <213> 164 <243> 150
<23> 163 7 196 X 162 x 147 <150> 157 <184> 142 <214> 161 <244> 152
<24> 141 8 161 Y 139 y 169 <151> 162 <185> 173 <215> 142 <245> 178
<25> 156 9 129 Z 174 z 140 <152> 142 <186> 160 <216> 166 <246> 176
<26> 139 : 150 [ 168 { 157 <153> 140 <187> 164 <217> 137 <247> 157
<27> 157 ; 153 \ 169 | 141 <154> 145 <188> 171 <218> 171 <248> 164
<28> 130 < 165 ] 164 } 158 <155> 177 <189> 163 <219> 180 <249> 134
<29> 152 = 154 ^ 176 ~ 164 <156> 145 <190> 159 <220> 172 <250> 171
<30> 141 > 168 _ 153 <157> 143 <251> 159
<31> 149 ? 165 ` 155 <158> 131 <252> 163
@ 136 <159> 163 <253> 179
_ 181 <160> 139 <254> 146
j 150 <255> 133
~ 163
_ 163

```

Figura 22: Resultado da contagem de bytes para o arquivo .jpg compactado em formato .rar

A Figura 18: Arquivo *.jpg* original e compactados nos mostra a redução de tamanho entre os formatos *.jpeg*, *.zip* e *.rar*, e, através dela, podemos concluir que a compactação identificada foi mínima, considerando que um arquivo que antes tinha 40KB passou a ter 39KB em ambos os formatos de compactação.

Na Figura 20: Resultado da contagem de bytes para o arquivo *.jpg*, Figura 21: Resultado da contagem de bytes para o arquivo *.jpg* compactado em formato *.zip* e Figura 22: Resultado da contagem de bytes para o arquivo *.jpg* compactado em formato *.rar*, foram destacados os caracteres encontrados em maior número no formato *.jpg*: '6' e 'm'. Para o formato *.jpg*, eles apareceram 419 vezes e 312 vezes, respectivamente. Para o formato *.zip*, os caracteres passaram a aparecer 162 e 163 vezes, respectivamente. E, para o formato *.rar*, a aparição dos caracteres ocorreu em 150 e 138 vezes, respectivamente.

	<i>.jpg</i>	<i>.zip</i>	<i>.rar</i>	Frequência de Bytes: Arquivo original/Compactado	
				<i>.zip</i>	<i>.rar</i>
6	419	162	150	2,58	2,79
m	312	163	138	1,91	2,26

Com base nos valores identificados na Figura 18: Arquivo *.jpg* original e compactados e na Figura 20: Resultado da contagem de bytes para o arquivo *.jpg*, Figura 21: Resultado da contagem de bytes para o arquivo *.jpg* compactado em formato *.zip* e Figura 22: Resultado da contagem de bytes para o arquivo *.jpg* compactado em formato *.rar*, podemos concluir que, como o *.jpg* é um formato de fácil compactação por ser gerado em blocos, a imagem utilizada no teste já possuía um alto nível de compactação, considerando que a diferença entre os tamanhos do arquivo original e dos compactados foi de apenas 1KB e a redução das frequências é inferior a 3 vezes menor.

✓ *.png*:

A imagem abaixo foi utilizada para os testes, foi encontrada através do *Google Imagens* e está disponível no *Pinterest*.



Biblioteca Documentos

projeto




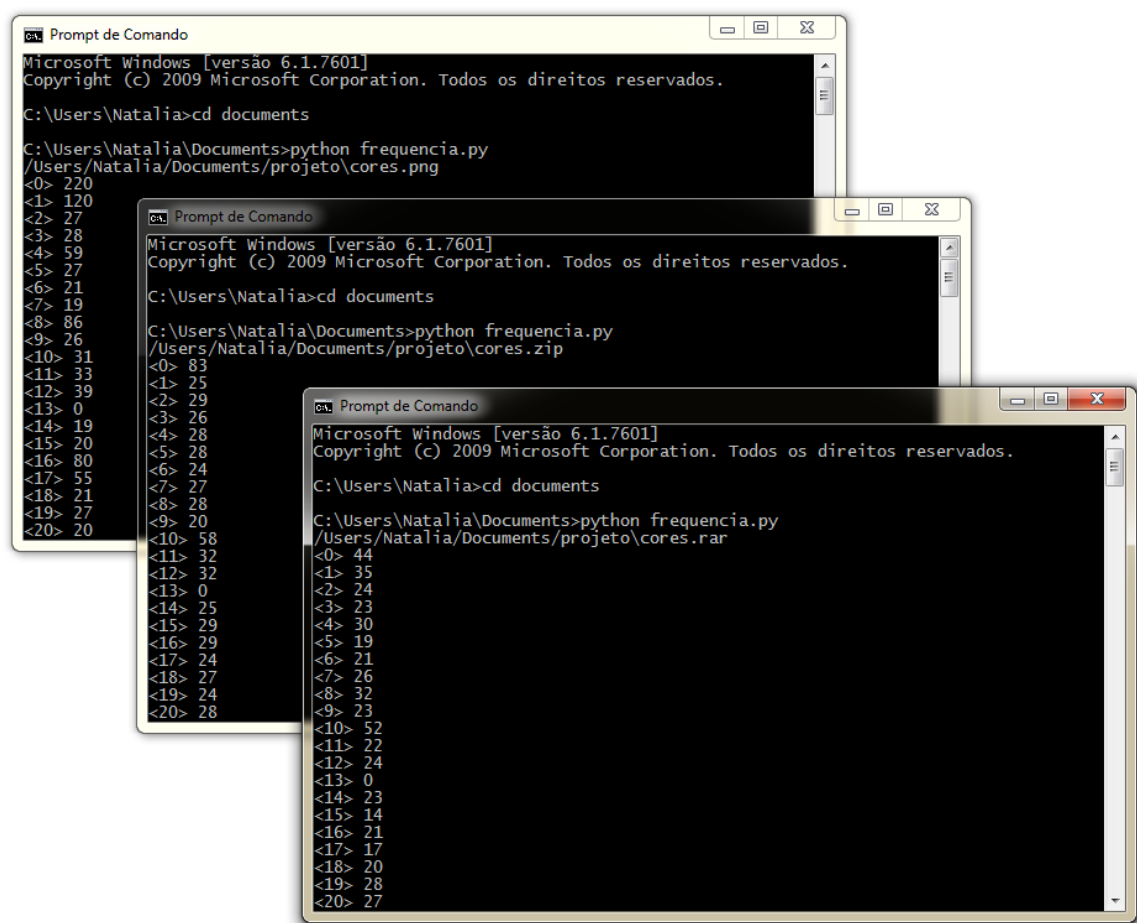
Nome	Tipo	Tamanho
 cores	Arquivo ZIP do WinRAR	7 KB
 cores	Arquivo do WinRAR	7 KB
 cores	Imagem PNG	7 KB

Figura 23: Arquivo .png original e compactados

Os testes foram realizados para os três formatos do arquivo, conforme Figura 24: Compactação de arquivo .png.



```
Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto\cores.png
<0> 220
<1> 120
<2> 27
<3> 28
<4> 59
<5> 27
<6> 21
<7> 19
<8> 86
<9> 26
<10> 31
<11> 33
<12> 39
<13> 0
<14> 19
<15> 20
<16> 80
<17> 55
<18> 21
<19> 27
<20> 20

Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto\cores.zip
<0> 83
<1> 25
<2> 29
<3> 26
<4> 28
<5> 28
<6> 24
<7> 27
<8> 28
<9> 20
<10> 58
<11> 32
<12> 32
<13> 0
<14> 25
<15> 29
<16> 29
<17> 24
<18> 27
<19> 24
<20> 28

Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto\cores.rar
<0> 44
<1> 35
<2> 24
<3> 23
<4> 30
<5> 19
<6> 21
<7> 26
<8> 32
<9> 23
<10> 52
<11> 22
<12> 24
<13> 0
<14> 23
<15> 14
<16> 21
<17> 17
<18> 20
<19> 28
<20> 27
```

Figura 24: Compactação de arquivo .png


```

C:\Users\Natalia>cd documents
C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto\cores.png
<0> 220 57 A 40 a 28 <127> 17 <161> 20 <191> 19 <221> 15
<1> 120 ! 53 B 48 b 35 <128> 30 <162> 17 <192> 19 <222> 28
<2> 27 " 54 C 28 c 32 <129> 14 <163> 12 <193> 24 <223> 20
<3> 28 # 30 D 28 d 18 <130> 37 <164> 24 <194> 28 <224> 32
<4> 59 $ 48 E 24 e 57 <131> 13 <165> 45 <195> 35 <225> 24
<5> 27 % 25 F 15 f 25 <132> 56 <166> 14 <196> 38 <226> 16
<6> 21 & 22 G 32 g 40 <133> 25 <167> 14 <197> 21 <227> 21
<7> 19 ' 12 H 23 h 11 <134> 24 <168> 22 <198> 49 <228> 30
<8> 86 ( 79 I 37 i 20 <135> 27 <169> 19 <199> 16 <229> 21
<9> 26 ) 42 J 39 j 18 <136> 57 <170> 23 <200> 34 <230> 26
<10> 31 * 29 K 31 k 29 <137> 22 <171> 23 <201> 24 <231> 22
<11> 33 + 11 L 22 l 32 <138> 26 <172> 22 <202> 18 <232> 18
<12> 39 , 107 M 19 m 65 <139> 24 <173> 17 <203> 63 <233> 25
<13> 0 - 24 N 25 n 7 <140> 49 <174> 13 <204> 37 <234> 16
<14> 19 . 14 O 20 o 26 <141> 19 <175> 14 <205> 24 <235> 23
<15> 20 / 29 P 29 p 21 <142> 18 <176> 28 <206> 21 <236> 18
<16> 80 0 27 Q 12 q 18 <143> 32 <177> 32 <207> 23 <237> 15
<17> 55 1 54 R 45 r 18 <144> 28 <178> 73 <208> 21 <238> 14
<18> 21 2 19 S 10 s 33 <145> 29 <179> 37 <209> 18 <239> 16
<19> 27 3 27 T 24 t 12 <146> 40 <180> 10 <210> 25 <240> 37
<20> 20 4 16 U 20 u 16 <147> 17 <181> 25 <211> 15 <241> 13
<21> 21 5 29 V 22 v 6 <148> 41 <182> 69 <212> 27 <242> 25
<22> 38 6 35 W 16 w 15 <149> 18 <183> 20 <213> 15 <243> 20
<23> 10 7 12 X 29 x 35 <150> 64 <184> 18 <214> 32 <244> 28
<24> 42 8 37 Y 65 y 26 <151> 21 <185> 27 <215> 10 <245> 18
<25> 22 9 20 Z 25 z 29 <152> 11 <186> 10 <216> 19 <246> 22
<26> 19 : 27 [ 19 { 25 <153> 33 <187> 6 <217> 18 <247> 13
<27> 21 ; 9 \ 19 | 29 <154> 14 <188> 20 <218> 23 <248> 20
<28> 33 < 32 ] 5 } 11 <155> 15 <189> 20 <219> 90 <249> 31
<29> 16 = 27 ^ 13 ~ 15 <156> 27 <190> 20 <220> 8 <250> 12
<30> 19 > 20 _ 10 <157> 22 <251> 23
<31> 23 ? 23 ` 10 <158> 27 <252> 14
@ 25 <159> 20 <253> 15
<160> 16 <254> 21
<255> 44

```

Figura 25: Resultado da contagem de bytes para o arquivo .png

```

C:\Users\Natalia>cd documents
C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto\cores.zip
<0> 83 25 A 22 a 24 <127> 35 <161> 35 <191> 26 <221> 38
<1> 25 ! 29 B 26 b 21 <128> 29 <162> 21 <192> 33 <222> 23
<2> 29 " 25 C 23 c 35 <129> 27 <163> 27 <193> 32 <223> 30
<3> 26 # 29 D 21 d 29 <130> 26 <164> 27 <194> 25 <224> 34
<4> 28 $ 23 E 26 e 33 <131> 24 <165> 28 <195> 19 <225> 31
<5> 28 % 18 F 17 f 33 <132> 31 <166> 30 <196> 25 <226> 32
<6> 24 & 29 G 20 g 20 <133> 36 <167> 22 <197> 29 <227> 36
<7> 27 ' 23 H 25 h 20 <134> 29 <168> 23 <198> 23 <228> 34
<8> 28 ( 24 I 32 i 24 <135> 28 <169> 28 <199> 16 <229> 31
<9> 20 ) 19 J 32 j 32 <136> 29 <170> 29 <200> 23 <230> 30
<10> 58 * 26 K 34 k 32 <137> 15 <171> 28 <201> 34 <231> 32
<11> 32 + 27 L 35 l 20 <138> 23 <172> 27 <202> 19 <232> 14
<12> 32 , 24 M 29 m 24 <139> 22 <173> 25 <203> 35 <233> 29
<13> 0 - 22 N 20 n 27 <140> 33 <174> 20 <204> 22 <234> 24
<14> 25 . 28 O 32 o 20 <141> 28 <175> 22 <205> 21 <235> 29
<15> 29 / 21 P 26 p 15 <142> 29 <176> 26 <206> 28 <236> 35
<16> 29 0 23 Q 23 q 33 <143> 27 <177> 19 <207> 29 <237> 23
<17> 24 1 28 R 19 r 26 <144> 26 <178> 33 <208> 17 <238> 24
<18> 27 2 32 S 27 s 33 <145> 29 <179> 14 <209> 23 <239> 21
<19> 24 3 30 T 28 t 38 <146> 27 <180> 31 <210> 26 <240> 22
<20> 28 4 24 U 31 u 29 <147> 25 <181> 26 <211> 32 <241> 29
<21> 14 5 25 V 36 v 22 <148> 25 <182> 29 <212> 21 <242> 28
<22> 29 6 20 W 21 w 24 <149> 22 <183> 25 <213> 27 <243> 39
<23> 33 7 23 X 29 x 24 <150> 27 <184> 25 <214> 25 <244> 22
<24> 28 8 25 Y 31 y 26 <151> 24 <185> 20 <215> 26 <245> 21
<25> 23 9 32 Z 26 z 30 <152> 27 <186> 25 <216> 27 <246> 25
<26> 33 : 25 [ 36 { 17 <153> 23 <187> 24 <217> 28 <247> 31
<27> 22 ; 26 \ 20 | 23 <154> 28 <188> 32 <218> 34 <248> 21
<28> 33 < 15 ] 38 ^ 22 <155> 20 <189> 24 <219> 19 <249> 29
<29> 16 = 26 ^ 31 ~ 21 <156> 26 <190> 30 <220> 19 <250> 33
<30> 16 > 36 _ 26 <157> 21 <251> 36
<31> 29 ? 29 ` 24 <158> 22 <252> 18
@ 14 <159> 23 <253> 21
<160> 25 <254> 18
<255> 26

```

Figura 26: Resultado da contagem de bytes para o arquivo .png compactado em formato .zip

```

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto\cores.rar
<0> 44 30 - - a 22 <127> 22 <161> 35 <191> 31 <221> 26
<1> 35 ! 26 A 23 a 22 <128> 30 <162> 28 <192> 15 <222> 37
<2> 24 " 33 B 20 b 26 <129> 33 <163> 26 <193> 26 <223> 31
<3> 23 # 29 C 23 c 31 <130> 16 <164> 28 <194> 25 <224> 32
<4> 30 $ 25 D 16 d 13 <131> 25 <165> 19 <195> 29 <225> 18
<5> 19 % 23 E 43 e 27 <132> 28 <166> 22 <196> 26 <226> 18
<6> 21 & 23 F 20 f 26 <133> 26 <167> 22 <197> 20 <227> 21
<7> 26 ' 28 G 25 g 25 <134> 31 <168> 22 <198> 25 <228> 27
<8> 32 ( 27 H 30 h 39 <135> 23 <169> 19 <199> 23 <229> 29
<9> 23 ) 28 I 15 i 29 <136> 20 <170> 22 <200> 34 <230> 25
<10> 52 * 29 J 28 j 24 <137> 30 <171> 38 <201> 23 <231> 27
<11> 22 + 24 K 29 k 23 <138> 31 <172> 29 <202> 26 <232> 34
<12> 24 , 26 L 24 l 32 <139> 20 <173> 18 <203> 36 <233> 10
<13> 0 - 34 M 32 m 18 <140> 26 <174> 32 <204> 17 <234> 32
<14> 23 . 39 N 29 n 35 <141> 23 <175> 24 <205> 24 <235> 35
<15> 14 / 33 O 27 o 29 <142> 26 <176> 24 <206> 33 <236> 30
<16> 21 0 24 P 27 p 20 <143> 33 <177> 20 <207> 29 <237> 24
<17> 17 1 21 Q 30 q 21 <144> 25 <178> 27 <208> 22 <238> 22
<18> 20 2 17 R 27 r 27 <145> 28 <179> 25 <209> 31 <239> 38
<19> 28 3 22 S 33 s 30 <146> 28 <180> 19 <210> 18 <240> 24
<20> 27 4 27 T 27 t 24 <147> 19 <181> 33 <211> 25 <241> 34
<21> 24 5 32 U 24 u 31 <148> 20 <182> 33 <212> 28 <242> 26
<22> 24 6 20 V 29 v 29 <149> 16 <183> 24 <213> 34 <243> 33
<23> 27 7 24 W 27 w 29 <150> 23 <184> 30 <214> 24 <244> 22
<24> 24 8 12 X 36 x 17 <151> 21 <185> 24 <215> 35 <245> 26
<25> 30 9 23 Y 18 y 20 <152> 26 <186> 22 <216> 23 <246> 22
<26> 34 : 27 Z 29 z 26 <153> 23 <187> 28 <217> 20 <247> 29
<27> 26 ; 28 [ 21 { 23 <154> 24 <188> 21 <218> 23 <248> 21
<28> 32 < 31 \ 26 | 17 <155> 25 <189> 32 <219> 13 <249> 23
<29> 27 > 36 ] 24 } 39 <156> 22 <190> 21 <220> 33 <250> 35
<30> 24 ~ 31 ^ 29 ~ 25 <157> 26 <251> 25
<31> 29 ? 31 _ 21 <158> 22 <252> 32
@ 25 <159> 26 <253> 30
. 18 <160> 22 <254> 26
<255> 30

```

Figura 27: Resultado da contagem de bytes para o arquivo .png compactado em formato .rar

Na Figura 23: Arquivo .png original e compactados é possível perceber que quando o arquivo .png foi compactado não houve redução de tamanho e o arquivo que possuía 7KB se manteve com os 7KB no formato .zip e no formato .rar, isto não necessariamente se deve ao formato possuir baixo poder de compressão e sim aos pixels da imagem, que, por ser muito colorida, minimiza os bytes que podem ser suprimidos.

Na Figura 25: Resultado da contagem de bytes para o arquivo .png, Figura 26: Resultado da contagem de bytes para o arquivo .png compactado em formato .zip e Figura 27: Resultado da contagem de bytes para o arquivo .png compactado em formato .rar escolhemos os caracteres ‘,’ e ‘m’, que foram identificados com maior frequência no formato .png, para realizarmos a análise. No formato .png a ‘,’ apareceu 107 vezes, para 24 no .zip e 26 no .rar; já o caractere ‘m’, a frequência encontrada foi de 65 aparições no arquivo .png para 24 no .zip e 18 no .rar.

	.png	.zip	.rar	Frequência de Bytes: Arquivo original/Compactado	
				.zip	.rar
,	107	24	26	4,46	4,11
m	65	24	18	2,7	3,61

De acordo com o que foi identificado para o arquivo png, podemos concluir que, apesar da frequência de Bytes ter diminuído em média 4 vezes nos caracteres analisados para ambos os formatos, o tamanho dos arquivos se manteve constante, sendo assim, ainda que a compressão não tenha sido capaz de reduzir o tamanho da imagem a frequência de Bytes de alterou.

✓ .mp3:

O arquivo .mp3 utilizado para a análise da contagem de frequência foi a música *Apenas mais uma de amor*, do Lulu Santos.




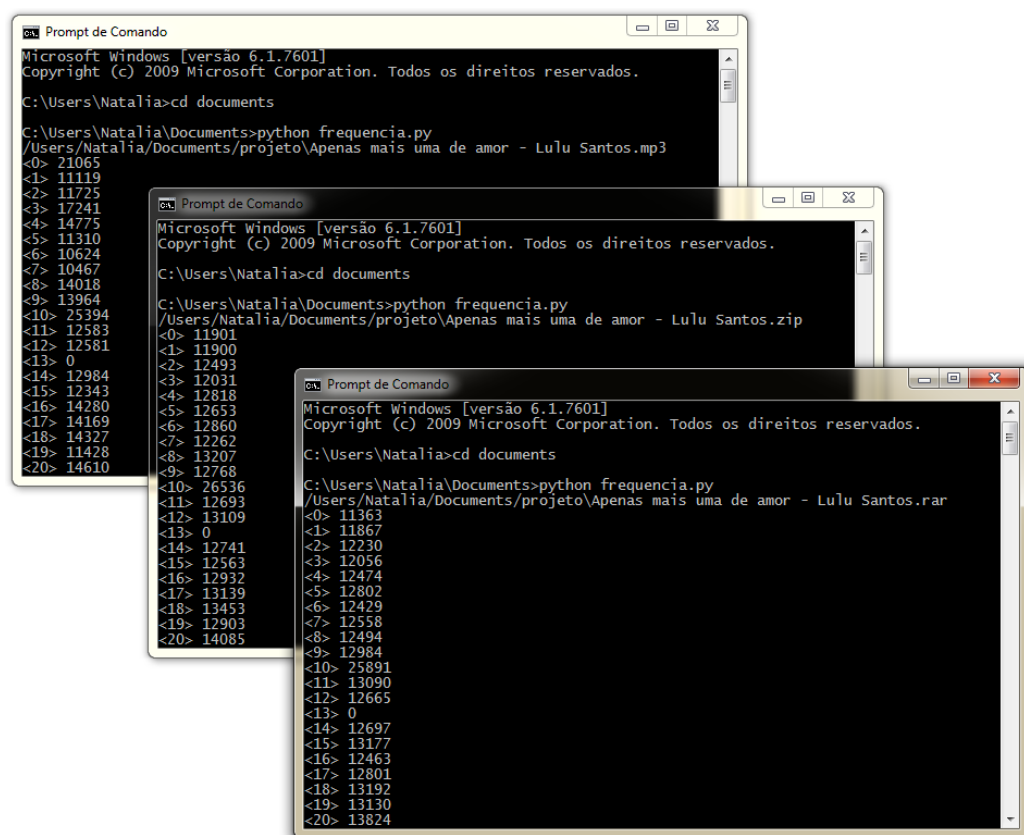
Biblioteca Documentos		
projeto		
Nome	Tipo	Tamanho
 Apenas mais uma de amor - Lulu Santos	Arquivo ZIP do WinRAR	3.384 KB
 Apenas mais uma de amor - Lulu Santos	Arquivo do WinRAR	3.378 KB
 Apenas mais uma de amor - Lulu Santos	Som no Formato MP3	3.406 KB

Figura 28: Arquivo .mp3 original e compactados

Os testes foram realizados para os três formatos do arquivo, conforme Figura 29: Compactação de arquivo .mp3.



```

Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents
C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Apenas mais uma de amor - Lulu Santos.mp3
<0> 21065
<1> 11119
<2> 11725
<3> 17241
<4> 14775
<5> 11310
<6> 10624
<7> 10467
<8> 14018
<9> 13964
<10> 25394
<11> 12583
<12> 12581
<13> 0
<14> 12984
<15> 12343
<16> 14280
<17> 14169
<18> 14327
<19> 11428
<20> 14610

Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents
C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Apenas mais uma de amor - Lulu Santos.zip
<0> 11901
<1> 11900
<2> 12493
<3> 12031
<4> 12818
<5> 12653
<6> 12860
<7> 12262
<8> 13207
<9> 12768
<10> 26536
<11> 12693
<12> 13109
<13> 0
<14> 12741
<15> 12563
<16> 12932
<17> 13139
<18> 13453
<19> 12903
<20> 14085

Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents
C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Apenas mais uma de amor - Lulu Santos.rar
<0> 11363
<1> 11867
<2> 12230
<3> 12056
<4> 12474
<5> 12802
<6> 12429
<7> 12558
<8> 12494
<9> 12984
<10> 25891
<11> 13090
<12> 12665
<13> 0
<14> 12697
<15> 13177
<16> 12463
<17> 12801
<18> 13192
<19> 13130
<20> 13824
  
```

Figura 29: Compactação de arquivo .mp3

```
Prompt de Comando
Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Apenas mais uma de amor - Lulu Santos.mp3

<0> 21065 | 14456 | A 13611 | a 14045 | <127> 15303 | <161> 18386 | <191> 14718 | <221> 13295
<1> 11119 | 13244 | B 16295 | b 13088 | <128> 16238 | <162> 15149 | <192> 9238 | <222> 11631
<2> 11725 | 16162 | C 13603 | c 12802 | <129> 10695 | <163> 13764 | <193> 9988 | <223> 13931
<3> 17241 | 12922 | D 15869 | d 12385 | <130> 11710 | <164> 15722 | <194> 10899 | <224> 8567
<4> 14775 | 20744 | E 16677 | e 14537 | <131> 10444 | <165> 15057 | <195> 10658 | <225> 10267
<5> 11310 | 15378 | F 14060 | f 12352 | <132> 12122 | <166> 14436 | <196> 12346 | <226> 11647
<6> 10624 | 14094 | G 15146 | g 12518 | <133> 11858 | <167> 14551 | <197> 12355 | <227> 12035
<7> 10467 | 11558 | H 14871 | h 14152 | <134> 11612 | <168> 15390 | <198> 12026 | <228> 11968
<8> 14018 | 16189 | I 21105 | i 16956 | <135> 11859 | <169> 16337 | <199> 12093 | <229> 12763
<9> 13964 | 14798 | J 16709 | j 15357 | <136> 13858 | <170> 15885 | <200> 12467 | <230> 12254
<10> 25394 | 14752 | K 14519 | k 15243 | <137> 15774 | <171> 15522 | <201> 12363 | <231> 12633
<11> 12583 | 14047 | L 14184 | l 12604 | <138> 14747 | <172> 14046 | <202> 13656 | <232> 15697
<12> 12581 | 14935 | M 16435 | m 21076 | <139> 13131 | <173> 16912 | <203> 13623 | <233> 14322
<13> 0 | 13715 | N 12732 | n 14375 | <140> 12546 | <174> 13202 | <204> 12600 | <234> 15006
<14> 12984 | 12928 | O 11840 | o 12615 | <141> 14205 | <175> 13167 | <205> 12347 | <235> 15500
<15> 12343 | 10619 | P 14524 | p 10684 | <142> 13368 | <176> 10408 | <206> 12147 | <236> 12679
<16> 14280 | 12673 | Q 17110 | q 13760 | <143> 12498 | <177> 13339 | <207> 12164 | <237> 14115
<17> 14169 | 11901 | R 15839 | r 13271 | <144> 12522 | <178> 13311 | <208> 11369 | <238> 12670
<18> 14327 | 12974 | S 14620 | s 12707 | <145> 14564 | <179> 12608 | <209> 13992 | <239> 14835
<19> 11428 | 11910 | T 14967 | t 14767 | <146> 26577 | <180> 15460 | <210> 14067 | <240> 8126
<20> 14610 | 14160 | U 16802 | u 13902 | <147> 12564 | <181> 14896 | <211> 14551 | <241> 10610
<21> 13650 | 12919 | V 14420 | v 12229 | <148> 15135 | <182> 14564 | <212> 14403 | <242> 10897
<22> 11970 | 12662 | W 12997 | w 12005 | <149> 15710 | <183> 14143 | <213> 14496 | <243> 11708
<23> 10811 | 11803 | X 13202 | x 9509 | <150> 13287 | <184> 10685 | <214> 13499 | <244> 13563
<24> 12737 | 12251 | Y 15437 | y 11491 | <151> 11836 | <185> 13753 | <215> 13539 | <245> 14424
<25> 11405 | 12659 | Z 16132 | z 13300 | <152> 12120 | <186> 13572 | <216> 12055 | <246> 11929
<26> 12362 | 12605 | [ 15644 | { 12181 | <153> 13599 | <187> 13164 | <217> 13067 | <247> 13462
<27> 11757 | 12312 | \ 12026 | | 10208 | <154> 14301 | <188> 10588 | <218> 14648 | <248> 9085
<28> 14557 | 12233 | ] 15977 | } 13037 | <155> 13484 | <189> 15004 | <219> 16462 | <249> 11851
<29> 12092 | 13284 | ^ 11206 | ~ 10760 | <156> 11976 | <190> 11856 | <220> 12099 | <250> 14285
<30> 11425 | 11199 | _ 12607 | | 13076 | <157> 13076 | <251> 22291
<31> 10940 | 12146 | ` 19219 | | 11507 | <252> 10983
@ 13528 | | | | <158> 12055 | <253> 16064
| | | | <159> 12827 | <254> 14928
| | | | <160> 12827 | <255> 41305
```

Figura 30: Resultado da contagem de bytes para o arquivo .mp3

```
Prompt de Comando
Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Apenas mais uma de amor - Lulu Santos.zip

<0> 11901 | 12910 | A 12869 | a 12796 | <127> 14423 | <161> 13793 | <191> 14120 | <221> 13613
<1> 11900 | 13248 | B 13702 | b 13502 | <128> 11940 | <162> 13789 | <192> 12233 | <222> 13281
<2> 12493 | 13411 | C 13509 | c 13584 | <129> 12852 | <163> 13642 | <193> 13202 | <223> 14415
<3> 12031 | 13444 | D 13351 | d 13583 | <130> 13056 | <164> 13916 | <194> 12584 | <224> 12732
<4> 12818 | 13699 | E 13801 | e 13874 | <131> 12747 | <165> 14573 | <195> 12892 | <225> 12905
<5> 12653 | 13858 | F 13391 | f 13071 | <132> 12982 | <166> 14044 | <196> 12717 | <226> 13149
<6> 12860 | 13212 | G 13630 | g 13388 | <133> 13373 | <167> 13812 | <197> 13906 | <227> 13685
<7> 12262 | 13255 | H 13383 | h 13530 | <134> 13331 | <168> 13682 | <198> 13303 | <228> 13517
<8> 13207 | 13428 | I 13942 | i 13871 | <135> 12899 | <169> 14477 | <199> 13306 | <229> 13965
<9> 12768 | 14265 | J 14195 | j 14025 | <136> 12699 | <170> 15007 | <200> 13422 | <230> 13291
<10> 26536 | 14289 | K 14227 | k 14085 | <137> 13482 | <171> 14232 | <201> 13786 | <231> 13789
<11> 12693 | 13611 | L 13502 | l 13526 | <138> 14191 | <172> 13908 | <202> 13939 | <232> 13168
<12> 13109 | 13619 | M 13663 | m 14185 | <139> 13588 | <173> 14415 | <203> 13939 | <233> 13896
<13> 0 | 13642 | N 13654 | n 13127 | <140> 13165 | <174> 13942 | <204> 12889 | <234> 14124
<14> 12741 | 13717 | O 13757 | o 13532 | <141> 13260 | <175> 13977 | <205> 13632 | <235> 14031
<15> 12563 | 13703 | P 13417 | p 12470 | <142> 13374 | <176> 12596 | <206> 13458 | <236> 12992
<16> 12932 | 12769 | Q 14050 | q 13190 | <143> 13536 | <177> 13880 | <207> 13811 | <237> 13785
<17> 13139 | 113101 | R 14442 | r 13957 | <144> 13145 | <178> 13870 | <208> 13146 | <238> 13482
<18> 13453 | 123648 | S 14256 | s 13475 | <145> 13747 | <179> 13677 | <209> 13785 | <239> 13985
<19> 12903 | 13279 | T 14341 | t 13543 | <146> 14277 | <180> 13627 | <210> 13921 | <240> 12950
<20> 14085 | 413557 | U 14837 | u 13624 | <147> 13525 | <181> 14244 | <211> 13912 | <241> 13506
<21> 13552 | 513839 | V 14173 | v 13332 | <148> 13841 | <182> 14437 | <212> 13812 | <242> 13670
<22> 13195 | 613176 | W 13918 | w 13467 | <149> 14700 | <183> 13604 | <213> 14621 | <243> 13835
<23> 13016 | 713091 | X 13531 | x 12674 | <150> 14391 | <184> 12898 | <214> 13907 | <244> 13503
<24> 13311 | 813088 | Y 13992 | y 13260 | <151> 13831 | <185> 13786 | <215> 14064 | <245> 14178
<25> 13265 | 913580 | Z 13956 | z 13562 | <152> 12916 | <186> 14005 | <216> 12803 | <246> 13400
<26> 13680 | :13217 | [ 14082 | { 13558 | <153> 13623 | <187> 13760 | <217> 13790 | <247> 14262
<27> 12863 | :13214 | \ 13578 | | 13335 | <154> 13855 | <188> 13060 | <218> 13860 | <248> 13500
<28> 13067 | <12992 | ] 13983 | } 14130 | <155> 13182 | <189> 13593 | <219> 14373 | <249> 14120
<29> 12743 | =13162 | ^ 13504 | ~ 13930 | <156> 13622 | <190> 14150 | <220> 12983 | <250> 14487
<30> 12828 | >13442 | _ 13957 | | 13793 | <157> 13793 | <251> 14092
<31> 13284 | ?14197 | ` 12557 | | 13534 | <158> 13534 | <252> 13930
@ 12630 | | | | <159> 13999 | <253> 14450
| | | | <160> 12771 | <254> 14432
| | | | | <255> 14780
```

Figura 31: Resultado da contagem de bytes para o arquivo .mp3 compactado em formato .zip


```

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Apenas mais uma de amor - Lulu Santos.rar
<0> 11363 | 12493 | A 12777 | a 13050 | <127> 14175 | <161> 13687 | <191> 14127 | <221> 13271
<1> 11867 | 12851 | B 13057 | b 13549 | <128> 11919 | <162> 13812 | <192> 11975 | <222> 13556
<2> 12230 | 13363 | C 13180 | c 13574 | <129> 12747 | <163> 13882 | <193> 12999 | <223> 14199
<3> 12056 | 13142 | D 13561 | d 13112 | <130> 13118 | <164> 13843 | <194> 13439 | <224> 12610
<4> 12474 | 13806 | E 13684 | e 13682 | <131> 12907 | <165> 14276 | <195> 13262 | <225> 13388
<5> 12802 | 13633 | F 13545 | f 13253 | <132> 12945 | <166> 14137 | <196> 13185 | <226> 13552
<6> 12429 | 13433 | G 13622 | g 13695 | <133> 13698 | <167> 14290 | <197> 13596 | <227> 13716
<7> 12558 | 13618 | H 13333 | h 13616 | <134> 13168 | <168> 13722 | <198> 13635 | <228> 13299
<8> 12494 | 13451 | I 14198 | i 13923 | <135> 13389 | <169> 14481 | <199> 13631 | <229> 13887
<9> 12984 | 14068 | J 14126 | j 13859 | <136> 12771 | <170> 14389 | <200> 12798 | <230> 13620
<10> 25891 | 14083 | K 13736 | k 13666 | <137> 13452 | <171> 14041 | <201> 13630 | <231> 13812
<11> 13090 | 13742 | L 13448 | l 13105 | <138> 13759 | <172> 13731 | <202> 13864 | <232> 13388
<12> 12665 | 13471 | M 13770 | m 13724 | <139> 13813 | <173> 13948 | <203> 14013 | <233> 14058
<13> 0 | 13723 | N 13832 | n 12907 | <140> 13287 | <174> 13507 | <204> 13436 | <234> 13938
<14> 12697 | 13582 | O 14094 | o 13586 | <141> 13636 | <175> 13994 | <205> 13729 | <235> 13781
<15> 13177 | 13791 | P 13264 | p 12960 | <142> 13465 | <176> 12696 | <206> 13455 | <236> 13394
<16> 12463 | 12785 | Q 13873 | q 13528 | <143> 13650 | <177> 13881 | <207> 14102 | <237> 13577
<17> 12801 | 13188 | R 14020 | r 13600 | <144> 12513 | <178> 13600 | <208> 12989 | <238> 13226
<18> 13192 | 13468 | S 14093 | s 13416 | <145> 13580 | <179> 13754 | <209> 13837 | <239> 13843
<19> 13130 | 13267 | T 14238 | t 13455 | <146> 13976 | <180> 13439 | <210> 13769 | <240> 13121
<20> 13824 | 13894 | U 14500 | u 13774 | <147> 13640 | <181> 13948 | <211> 13920 | <241> 13707
<21> 13473 | 13394 | V 14058 | v 13183 | <148> 13913 | <182> 13614 | <212> 14022 | <242> 13708
<22> 13286 | 13162 | W 13707 | w 13253 | <149> 14286 | <183> 13581 | <213> 14089 | <243> 13829
<23> 13415 | 12981 | X 13403 | x 13181 | <150> 13851 | <184> 13204 | <214> 13895 | <244> 13846
<24> 13238 | 13100 | Y 13941 | y 13888 | <151> 13815 | <185> 13395 | <215> 13793 | <245> 14374
<25> 13251 | 13359 | Z 14122 | z 13681 | <152> 13063 | <186> 13682 | <216> 12833 | <246> 13764
<26> 13569 | 13481 | [ 13601 | { 13639 | <153> 13420 | <187> 13406 | <217> 13488 | <247> 14101
<27> 12905 | 13001 | \ 13366 | | 13643 | <154> 13961 | <188> 13521 | <218> 13623 | <248> 13753
<28> 13312 | 13500 | ] 13769 | ^ 14367 | <155> 13607 | <189> 13735 | <219> 13879 | <249> 13958
<29> 12900 | 13584 | _ 13706 | ~ 14283 | <156> 13495 | <190> 14171 | <220> 13309 | <250> 14499
<30> 13309 | 14247 | 13944 | <157> 13565 | <191> 14171 | <221> 13271 | <251> 14249
<31> 14079 | 14155 | 12390 | <158> 13451 | <192> 13812 | <222> 13556 | <252> 14104
| <159> 14528 | <193> 12999 | <223> 14199 | <253> 14154
| <160> 12764 | <194> 13439 | <224> 12610 | <254> 14231
| <195> 13262 | <196> 13185 | <225> 13388 | <255> 14086
| <197> 13596 | <198> 13635 | <226> 13552 |
| <199> 13631 | <200> 12798 | <227> 13716 |
| <201> 13630 | <202> 13864 | <228> 13299 |
| <203> 14013 | <204> 13436 | <229> 13887 |
| <205> 13729 | <206> 13455 | <230> 13620 |
| <207> 14102 | <208> 12989 | <231> 13812 |
| <209> 13837 | <210> 13769 | <232> 13388 |
| <211> 13920 | <212> 14022 | <233> 14058 |
| <213> 14089 | <214> 13895 | <234> 13938 |
| <215> 13793 | <216> 12833 | <235> 13781 |
| <217> 13488 | <218> 13623 | <236> 13394 |
| <219> 13879 | <220> 13309 | <237> 13577 |
| <221> 13271 | <222> 13556 | <238> 13226 |
| <223> 14199 | <224> 12610 | <239> 13843 |
| <225> 13388 | <226> 13552 | <240> 13121 |
| <227> 13716 | <228> 13299 | <241> 13707 |
| <229> 13887 | <230> 13620 | <242> 13708 |
| <231> 13812 | <232> 13388 | <243> 13829 |
| <233> 14058 | <234> 13938 | <244> 13846 |
| <235> 13781 | <236> 13394 | <245> 14374 |
| <237> 13577 | <238> 13226 | <246> 13764 |
| <239> 13843 | <240> 13121 | <247> 14101 |
| <241> 13707 | <242> 13708 | <248> 13753 |
| <243> 13829 | <244> 13846 | <249> 13958 |
| <245> 14374 | <246> 13764 | <250> 14499 |
| <247> 14101 | <248> 13753 | <251> 14249 |
| <249> 13958 | <250> 14499 | <252> 14104 |
| <251> 14249 | <252> 14104 | <253> 14154 |
| <253> 14154 | <254> 14231 | <255> 14086 |
| <255> 14086 |

```

Figura 32: Resultado da contagem de bytes para o arquivo .mp3 compactado em formato .rar

Para a compressão de um arquivo .mp3, identificamos através da Figura 28: Arquivo .mp3 original e compactados, que apesar do formato já ser compactado, houve compressão e diminuição do tamanho dos arquivos, que originalmente tinha 3406KB e passaram a ter 3384KB no formato .zip e 3378KB no formato .rar. Ainda assim, o valor reduzido é quase insignificante para o tamanho total do arquivo.

Nos testes de frequência de bytes, identificados na Figura 30: Resultado da contagem de bytes para o arquivo .mp3, Figura 31: Resultado da contagem de bytes para o arquivo .mp3 compactado em formato .zip e Figura 32: Resultado da contagem de bytes para o arquivo .mp3 compactado em formato .rar, foram escolhidos os caracteres 'Q' e 'm', que tiveram uma redução significativa se comparada à redução de tamanho total dos arquivos. Para o caractere 'Q', que aparecia 17110 vezes no arquivo original, passou a aparecer 14050 vezes no arquivo .zip e 13873 vezes no .rar. Para o caractere 'm', a incidência de 21076 vezes no arquivo principal caiu para 14185 vezes no .zip e 13724 vezes no .rar.

	.png	.zip	.rar	Frequência de Bytes: Arquivo original/Compactado	
				.zip	.rar
Q	17110	14050	13873	1,22	1,23
m	21076	14185	13724	1,48	1,53

Na tabela acima é possível perceber que, apesar do arquivo ter sido comprimido, a diferença de tamanho é extremamente insignificante, considerando que os caracteres analisados não atingiram redução de nem 50% da frequência.

✓ .mp4:




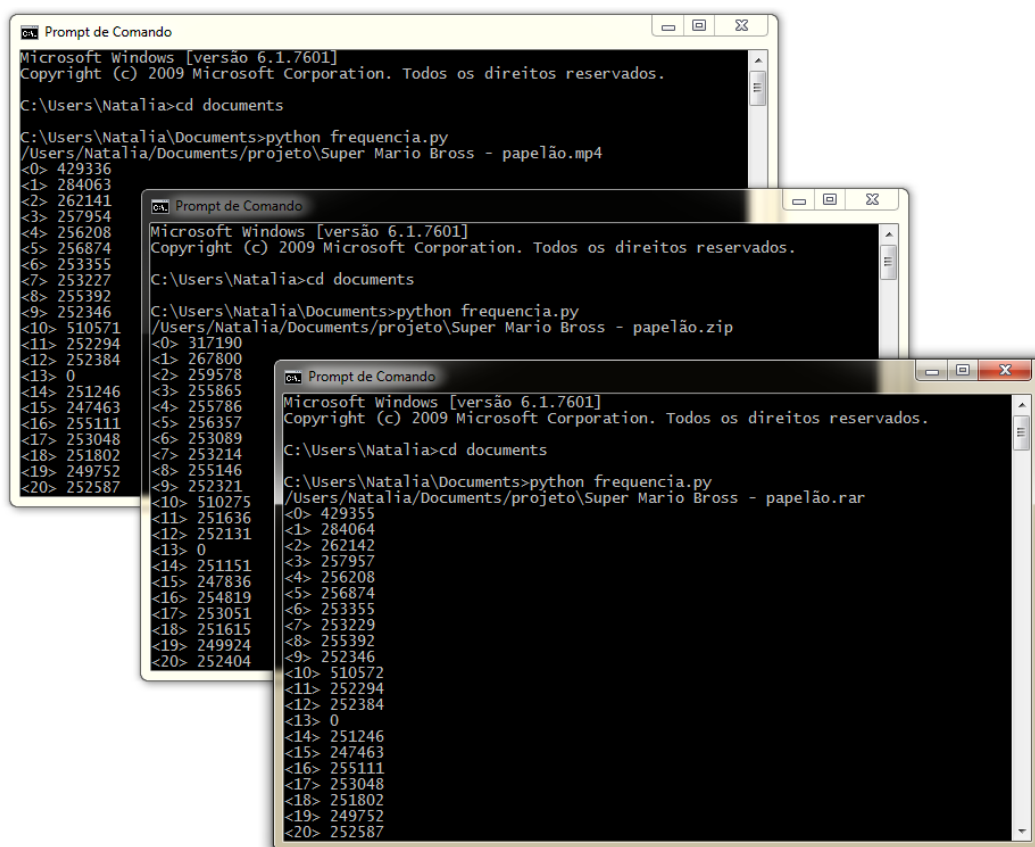
Biblioteca Documentos		
projeto		
Nome	Tipo	Tamanho
 Super Mario Bross - papelão	Arquivo ZIP do WinRAR	62.732 KB
 Super Mario Bross - papelão	Arquivo do WinRAR	62.846 KB
 Super Mario Bross - papelão	Vídeo MP4	62.845 KB

Figura 33: Arquivo .mp4 original e compactados

Os testes foram realizados para os três formatos do arquivo, conforme Figura 34: Compactação de arquivo .mp4.



```

Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Super Mario Bross - papelão.mp4
<0> 429336
<1> 284063
<2> 262141
<3> 257954
<4> 256208
<5> 256874
<6> 253355
<7> 253227
<8> 255392
<9> 252346
<10> 510571
<11> 252294
<12> 252384
<13> 0
<14> 251246
<15> 247463
<16> 255111
<17> 253048
<18> 251802
<19> 249752
<20> 252587
  
```

Figura 34: Compactação de arquivo .mp4

```
Prompt de Comando
Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Super Mario Bros - papelão.mp4
<0> 429336 257262 A 256150 a 251062 <127> 249840 <161> 248740 <191> 250109 <221> 251709
<1> 284063 262215 B 251444 b 250920 <128> 265925 <162> 251271 <192> 256796 <222> 248680
<2> 262141 250591 C 247589 c 248703 <129> 252358 <163> 248386 <193> 249385 <223> 250877
<3> 257954 250067 D 250376 d 249793 <130> 250896 <164> 249852 <194> 249518 <224> 250892
<4> 256208 250721 E 250380 e 250301 <131> 248741 <165> 250363 <195> 247498 <225> 247670
<5> 256874 250402 F 249239 f 249498 <132> 250953 <166> 250474 <196> 247740 <226> 248330
<6> 253355 250201 G 247609 g 250129 <133> 249983 <167> 248951 <197> 248702 <227> 248918
<7> 253227 249063 H 250995 h 249954 <134> 250070 <168> 251602 <198> 250750 <228> 248836
<8> 255392 253088 I 248747 i 249816 <135> 247521 <169> 250018 <199> 248396 <229> 250324
<9> 252346 250121 J 251791 j 251344 <136> 253393 <170> 253070 <200> 249240 <230> 247100
<10> 510571 252490 K 250641 k 250869 <137> 249376 <171> 252291 <201> 249452 <231> 249595
<11> 252294 251266 L 250396 l 249456 <138> 250833 <172> 250186 <202> 250159 <232> 249094
<12> 252384 251848 M 248035 m 250255 <139> 248932 <173> 250705 <203> 249748 <233> 248609
<13> 0 250484 N 250978 n 249420 <140> 250620 <174> 251682 <204> 247625 <234> 249143
<14> 251246 250484 O 248392 o 248865 <141> 249743 <175> 249297 <205> 249813 <235> 249280
<15> 247463 250246 P 253207 p 251881 <142> 249651 <176> 252591 <206> 249633 <236> 249358
<16> 255111 251069 Q 249856 q 249544 <143> 247756 <177> 249465 <207> 249209 <237> 248951
<17> 253048 249094 R 250838 r 249842 <144> 251936 <178> 251933 <208> 252149 <238> 250709
<18> 251802 249851 S 251215 s 250908 <145> 249753 <179> 249252 <209> 248935 <239> 249315
<19> 249752 249563 T 252095 t 250252 <146> 250489 <180> 249180 <210> 249018 <240> 249356
<20> 252587 250047 U 252083 u 250411 <147> 247572 <181> 250635 <211> 248384 <241> 247926
<21> 253535 250349 V 251355 v 249838 <148> 258896 <182> 251075 <212> 252577 <242> 248562
<22> 253048 248431 W 250459 w 250349 <149> 253992 <183> 249092 <213> 252210 <243> 248808
<23> 251065 248832 X 251509 x 249264 <150> 251882 <184> 250386 <214> 249410 <244> 248337
<24> 250799 249720 Y 250900 y 249206 <151> 249218 <185> 249388 <215> 250283 <245> 249573
<25> 250210 250259 Z 251295 z 250369 <152> 249299 <186> 250593 <216> 248690 <246> 248312
<26> 257068 249858 [ 250478 { 250369 <153> 248337 <187> 249649 <217> 249585 <247> 248874
<27> 248870 249598 \ 249761 | 248773 <154> 251760 <188> 249218 <218> 248853 <248> 248668
<28> 249226 248639 ^ 250649 } 248572 <155> 251242 <189> 249251 <219> 249886 <249> 248433
<29> 249018 248884 _ 250478 ~ 248454 <156> 249877 <190> 250216 <220> 249366 <250> 247826
<30> 248034 247146 253653 <157> 249495 <251> 248627
<31> 248905 248459 253653 <158> 251361 <252> 247553
260389 <159> 250383 <253> 247948
255111 <160> 254699 <254> 247830
<255> 255615
```

Figura 35: Resultado da contagem de bytes para o arquivo .mp4

```
Prompt de Comando
Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Super Mario Bros - papelão.zip
<0> 317190 257439 A 255957 a 250795 <127> 252320 <161> 248876 <190> 250403 <221> 251829
<1> 267800 261598 B 251475 b 250775 <128> 267159 <162> 251418 <191> 250583 <222> 249061
<2> 259578 250466 C 247818 c 248776 <129> 252330 <163> 248416 <192> 256645 <223> 251168
<3> 255865 248725 D 250380 d 249689 <130> 250829 <164> 249899 <193> 249475 <224> 251087
<4> 255786 250613 E 250327 e 250363 <131> 249001 <165> 250402 <194> 249721 <225> 247833
<5> 256357 250375 F 249307 f 249333 <132> 250946 <166> 250524 <195> 247694 <226> 248801
<6> 253089 250342 G 247813 g 249955 <133> 249900 <167> 249111 <196> 247973 <227> 249131
<7> 253214 249101 H 250944 h 249649 <134> 250345 <168> 251663 <197> 249019 <228> 249112
<8> 255146 251355 I 248846 i 249686 <135> 247706 <169> 250242 <198> 250841 <229> 250403
<9> 252321 250139 J 251737 j 251121 <136> 253125 <170> 253038 <199> 248553 <230> 247609
<10> 510275 252511 K 250600 k 250877 <137> 249452 <171> 252252 <200> 249431 <231> 249870
<11> 251636 251234 L 250433 l 249436 <138> 250876 <172> 250357 <201> 249585 <232> 249202
<12> 252131 251815 M 248178 m 250451 <139> 249102 <173> 251009 <202> 250271 <233> 248852
<13> 0 250628 N 250876 n 249423 <140> 250544 <174> 251845 <203> 249987 <234> 249232
<14> 251151 250464 O 248504 o 249033 <141> 249801 <175> 249542 <204> 247992 <235> 249476
<15> 247836 250359 P 252830 p 251090 <142> 249771 <176> 252544 <205> 249965 <236> 249635
<16> 254819 251102 Q 249822 q 249498 <143> 247922 <177> 249511 <206> 249966 <237> 249244
<17> 253051 249262 R 250813 r 250037 <144> 252006 <178> 252111 <207> 249378 <238> 250862
<18> 251615 249926 S 251122 s 250949 <145> 249753 <179> 249651 <208> 252348 <239> 249700
<19> 249924 249663 T 251784 t 250265 <146> 250518 <180> 249478 <209> 248983 <240> 249796
<20> 252404 250020 U 251962 u 250396 <147> 247651 <181> 250705 <210> 249145 <241> 248333
<21> 253232 250421 V 251406 v 249856 <148> 258562 <182> 251211 <211> 248751 <242> 248888
<22> 250535 248630 W 250477 w 250372 <149> 254137 <183> 249241 <212> 252423 <243> 249236
<23> 250398 248905 X 251250 x 249252 <150> 251833 <184> 249761 <213> 252391 <244> 248675
<24> 250854 249781 Y 250709 y 249311 <151> 249549 <185> 249644 <214> 249544 <245> 249965
<25> 250197 250299 Z 251081 z 250350 <152> 249657 <186> 250703 <215> 250623 <246> 248662
<26> 256525 250083 [ 250500 { 249066 <153> 248436 <187> 249709 <216> 248992 <247> 249332
<27> 248812 249666 \ 249577 | 248699 <154> 251965 <188> 249418 <217> 249894 <248> 249067
<28> 249200 248761 ^ 250529 } 248630 <155> 251252 <189> 249381 <218> 249086 <249> 248792
<29> 248968 249203 _ 250288 ~ 249151 <156> 250109 <219> 250249 <250> 248028
<30> 248319 247355 248869 <157> 249695 <220> 249557 <251> 248858
<31> 249087 248866 253439 <158> 251486 <252> 247982
259993 <159> 250529 <253> 248440
255111 <160> 254429 <254> 248419
<255> 255655
```

Figura 36: Resultado da contagem de bytes para o arquivo .mp4 compactado em formato .zip


```

Prompt de Comando
Microsoft Windows [versão 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Natalia>cd documents

C:\Users\Natalia\Documents>python frequencia.py
/Users/Natalia/Documents/projeto/Super Mario Bros - papelão.rar
<0> 429355 257767 A 256150 a 251065 <127> 249840 <161> 248740 <221> 251709
<1> 284064 ! 262216 B 251445 b 250920 <128> 265925 <162> 251271 <191> 250109 <222> 248680
<2> 262142 250591 C 247589 c 248703 <129> 252358 <163> 251271 <192> 256797 <223> 250877
<3> 257957 # 250067 D 250376 d 249793 <130> 250896 <164> 249852 <193> 249385 <224> 250892
<4> 256208 $ 250721 E 250380 e 250303 <131> 248742 <165> 250363 <194> 249518 <225> 247670
<5> 256874 % 250403 F 249239 f 249498 <132> 250953 <166> 250474 <195> 247498 <226> 248330
<6> 253355 & 250201 G 247609 g 250130 <133> 249983 <167> 248953 <196> 247741 <227> 248919
<7> 253229 ' 249063 H 250995 h 249954 <134> 250070 <168> 251602 <197> 248702 <228> 248836
<8> 255392 ( 253088 I 248747 i 249817 <135> 247521 <169> 250018 <198> 250751 <229> 250324
<9> 252346 ) 250121 J 251792 j 251344 <136> 253393 <170> 253070 <200> 249240 <230> 247100
<10> 510572 * 252490 K 250642 k 250869 <137> 249376 <171> 252291 <201> 249452 <231> 249595
<11> 252294 + 251266 L 250396 l 249457 <138> 250833 <172> 250186 <202> 250159 <232> 249094
<12> 252384 , 251848 M 248036 m 250256 <139> 248932 <173> 250705 <203> 249748 <233> 248609
<13> 0 - 250485 N 250978 n 249420 <140> 250620 <174> 251682 <204> 247626 <234> 249143
<14> 251246 . 250485 O 248392 o 248868 <141> 249743 <175> 249297 <205> 249813 <235> 249280
<15> 247463 / 250246 P 253207 p 251886 <142> 249651 <176> 252592 <206> 249633 <236> 249358
<16> 255111 0 251070 Q 249856 q 249544 <143> 247756 <177> 249465 <207> 249210 <237> 248951
<17> 253048 1 249094 R 250839 r 249846 <144> 251937 <178> 251933 <208> 252149 <238> 250709
<18> 251802 2 249851 S 251216 s 250911 <145> 249753 <179> 249252 <209> 248935 <239> 249315
<19> 249752 3 249564 T 252095 t 250253 <146> 250490 <180> 249180 <210> 249018 <240> 249356
<20> 252587 4 250048 U 252083 u 250412 <147> 247572 <181> 250635 <211> 248385 <241> 247926
<21> 253535 5 250349 V 251355 v 249838 <148> 258896 <182> 251075 <212> 252577 <242> 248562
<22> 250597 6 248431 W 250459 w 250349 <149> 253992 <183> 249092 <213> 252212 <243> 248810
<23> 251066 7 248832 X 251509 x 249264 <150> 251882 <184> 250386 <214> 249410 <244> 248337
<24> 250800 8 249720 Y 250900 y 249206 <151> 249218 <185> 249388 <215> 250283 <245> 249574
<25> 250210 9 250259 Z 251295 z 250369 <152> 249299 <186> 250594 <216> 248691 <246> 248312
<26> 257069 : 249858 [ 250478 { 248774 <153> 248337 <187> 249649 <217> 249585 <247> 248874
<27> 248870 ; 249598 \ 249761 } 248572 <154> 251760 <188> 249218 <218> 248853 <248> 248668
<28> 249226 < 248639 ] 250649 ^ 248454 <155> 251242 <189> 249251 <219> 249886 <249> 248433
<29> 249019 > 248885 A 250478 _ 248642 <156> 249877 <190> 250216 <220> 249366 <250> 247827
<30> 248034 ~ 247146 ` 248819 <157> 249495 <251> 248627
<31> 248905 ? 248459 253654 <158> 251361 <252> 247553
@ 260390 <159> 250383 <253> 247949
<160> 254699 <254> 247830
<255> 255615

```

Figura 37: Resultado da contagem de bytes para o arquivo .mp4 compactado em formato .rar

Com o auxílio da Figura 33: Arquivo .mp4 original e compactados é possível perceber que, apesar da compressão do arquivo .mp4 para o formato .zip ter gerado uma compressão mínima, a compressão do mesmo para o formato .rar, além de não ter reduzido seu tamanho, o tornou 1KB maior do que no formato original.

Na análise da frequência de bytes, foram utilizados os caracteres '!' e '@', que, como podemos ver na Figura 35: Resultado da contagem de bytes para o arquivo .mp4, Figura 36: Resultado da contagem de bytes para o arquivo .mp4 compactado em formato .zip e Figura 37: Resultado da contagem de bytes para o arquivo .mp4 compactado em formato .rar, possuíam 262215 e 260389 aparições no arquivo principal, respectivamente; estes valores foram reduzidos para 261598 e 259993 aparições para os caracteres no arquivo .zip e sofreram um aumento de um byte para cada caractere, que passaram a ter frequência de 262216 e 260390 vezes, respectivamente no arquivo .rar.

	.png	.zip	.rar	Frequência de Bytes: Arquivo original/Compactado	
				.zip	.rar
!	262215	261598	262216	1	0,99
@	260389	259993	260390	1	0,99

Com o auxílio da tabela acima, podemos verificar que, pelo fato de o formato de arquivo .mp4 ser naturalmente comprimido, as diferenças geradas na frequência dos bytes em cada arquivo se torna insignificante tanto para a diminuição, quanto para o aumento de frequência de cada caractere.

Conclusão

Após análise da frequência dos bytes nos formatos de arquivo *.txt*, *.pdf*, *.jpg*, *.png*, *.mp3*, *.mp4* e suas compressões para os formatos *.zip* e *.rar*, pudemos perceber que nem todo formato de arquivo é suficientemente compactável, e que nem sempre a frequência dos bytes se torna menor em um arquivo após sua compactação. Identificamos também que, para a maioria dos formatos, nos testes realizados, a compactação para *.rar* foi superior, tornando-a melhor no quesito compressão, porém, para o único formato (*.mp4*) onde o *.rar* não teve maior índice de compressão, ele também tornou o arquivo maior que o formato original.

Também foi identificada uma semelhança para as compactações de imagens, onde em ambos os formatos (*.jpg* e *.png*) as compactações tiveram o mesmo tamanho (em *.jpg* a redução de 1KB e em *.png* sem reduções).

Considerando que só foram realizados testes com um arquivo de cada formato, para que a pesquisa tenha maior garantia é recomendado realizar testes com mais de um arquivo do mesmo formato, e também, como futuros trabalhos, seria interessante o estudo sobre outros formatos menos utilizados, tanto para arquivos quanto para formatos de compactação, ou até mesmo a comparação entre arquivos de mesmo tipo, seja ele texto, imagem, áudio ou vídeo.