```python
# -*- coding: utf-8 -*-

import os
from os import listdir
from os.path import isfile, join
import pathlib
import gzip

DEBUG = False


def humanbytes(B):
    """Return the given bytes as a human friendly KB, MB, GB, or TB string."""
    B = float(B)
    KB = float(1024)
    MB = float(KB ** 2)  # 1,048,576
    GB = float(KB ** 3)  # 1,073,741,824
    TB = float(KB ** 4)  # 1,099,511,627,776

    if B < KB:
        return '{0} {1}'.format(B, 'Bytes' if 0 == B > 1 else 'Byte')
    elif KB <= B < MB:
        return '{0:.2f} KB'.format(B / KB)
    elif MB <= B < GB:
        return '{0:.2f} MB'.format(B / MB)
    elif GB <= B < TB:
        return '{0:.2f} GB'.format(B / GB)
    elif TB <= B:
        return '{0:.2f} TB'.format(B / TB)


def getFileTypeInfo(x):
    lastSuffix = pathlib.Path(x).suffix
    lastSuffix2 = pathlib.Path(x.rstrip(lastSuffix)).suffix

```

```
36      lastSuffix = lastSuffix.lstrip('.')
37      lastSuffix2 = lastSuffix2.lstrip('.')
38
39      if DEBUG:
40          print('getFileTypeInfo -> inputValue: ' + x)
41          print('getFileTypeInfo -> lastSuffix: ' +
    lastSuffix)
42          print('getFileTypeInfo -> lastSuffix2:' +
    lastSuffix2)
43
44      retStr = ''
45
46      if eligibleforcompression(x):
47          retStr = '[*]'
48      else:
49          retStr = '---'
50
51      if len(lastSuffix) > 0:
52          retStr = retStr + "[" + lastSuffix.rjust(5, ' '
    ) + "]"
53      else:
54          retStr = retStr + '[   ]'
55
56      if len(lastSuffix2) > 0:
57          retStr = retStr + "[" + lastSuffix2.rjust(5,
    ' ') + "]"
58      else:
59          retStr = retStr + '-------'
60
61      return retStr
62
63
64  def eligibleforcompression(x):
65      eligExtensions = ['.csv', '.pkl']
66      eligible = False
67
68      for ext in eligExtensions:
69          if DEBUG:
```

```
 70             print("eligibleforcompress -> eligible
    extension:    " + ext)
 71             print("eligibleforcompress -> extension
    received: " + pathlib.Path(x).suffix)
 72         if ext == pathlib.Path(x).suffix:
 73             eligible = True
 74     if DEBUG:
 75         print(f'{x} is eligible for compression.
    Return with {str(eligible)}')
 76     return eligible
 77
 78
 79 def compressfile(x, abspath, removeOriginal=False):
 80     if DEBUG:
 81         print(f'compressfile has been called on: {x}')
 82
 83     if eligibleforcompression(x):
 84         if DEBUG:
 85             print(x + " is eligible for compression")
 86     else:
 87         if DEBUG:
 88             print(x + " is not eligible for
    compression")
 89         return 0
 90
 91
 92     file = open(join(abspath, x), "rb")
 93     if DEBUG:
 94         print(x + " file has been opened")
 95     data = file.read()
 96     bindata = bytearray(data)
 97     print("======> compressing file: " + x)
 98     compressName = join(abspath, x) + ".gz"
 99     with gzip.open(compressName, "wb") as f:
100         f.write(bindata)
101
102     if removeOriginal:
103         if DEBUG:
```

```
104             print(f'Compressed file. Removing original
    : {x}')
105         os.remove(join(abspath, x))
106 
107     return 1
108 
109 
110 def exploreDirectory(absPath, compress=False,
    removeOriginal=False):
111     onlyFiles = [f for f in listdir(absPath) if isfile
    (join(absPath, f))]
112     onlyDirs = [f for f in listdir(absPath) if not
    isfile(join(absPath, f))]
113     onlyFiles.sort()
114     onlyDirs.sort()
115     numCompressedFiles = 0
116 
117     if DEBUG:
118         print(f'exploreDirectory called with directory
    :       {str(absPath)}')
119         print(f'exploreDirectory called with compress
    :        {str(compress)}')
120         print(f'exploreDirectory called with
    removeOriginal:  {str(removeOriginal)}')
121 
122     if compress:
123         print("Scanning [" + absPath + "] for data
    files to compress")
124         if len(onlyFiles) > 0:
125             print(f"======> {len(onlyFiles)} files
    found in directory.")
126             for x in onlyFiles:
127                 numCompressedFiles += compressfile(x,
    absPath, removeOriginal=removeOriginal)
128             print(f"======> {numCompressedFiles} file(
    s) have been compressed.")
129         else:
130             print("======> No files found in directory
```

```
130 .")
131 
132         print('')
133         for x in onlyDirs:
134             exploreDirectory(join(absPath,x), compress
    =True, removeOriginal=removeOriginal)
135 
136 
137         if DEBUG and compress:
138             print("Explore directory completed
    compression.")
139 
140 
141     else:
142 
143         if len(onlyFiles) == 0:
144             print(f'[D] {absPath} [Empty directory]')
145         else:
146             print(f'[D] {absPath}')
147             for x in onlyFiles:
148                 print(f'{getFileTypeInfo(x)}--> {x} ({
    humanbytes(os.stat(join(absPath, x)).st_size)})')
149 
150         print('')
151         for x in onlyDirs:
152             exploreDirectory(join(absPath, x))
153 
```