Huffman-Coding

Aim

To implement Huffman coding to compress the data using Python.

Software Required

1. Anaconda - Python 3.7

Algorithm:

Step1:

Get the input String or assign the string to generate huffman code.

Step2:

Create a class and function to build the huffman code tree nodes.

Step3:

Find the individual characters in the string.

Step4:

Calculate frequency of occurrence and implement the huffman code function into the frequency.

Step5:

Print the characters and its huffmancode.

Program:

```
# Developed by:Kaushika.A
# reg no: 212221230048
```

```
# Get the input String
string = input()
```

```
# Create tree nodes
class node_tree(object):
    def __init__(self,left=None,right=None):
        self.left = left
        self.right=right
    def children(self):
        return(self.left,self.right)

def huffman_code_tree(node,left=True,binString=''):
    if type(node) is str:
        return {node:binString}
    (l,r) = node.children()
    d=dict()
    d.update(huffman_code_tree(l,True,binString+'0'))
    d.update(huffman_code_tree(r,False,binString+'1'))
    return d
```

```
# Main function to implement huffman coding
freq = {}
for c in string:
    if c in freq:
        freq[c] += 1
    else:
        freq[c] = 1
freq = sorted(freq.items(), key=lambda x: x[1], reverse=True)
nodes=freq
```

```
# Calculate frequency of occurrence
while len(nodes) > 1:
    (key1,c1)=nodes[-1]
    (key2,c2)=nodes[-2]
    nodes=nodes[:-2]
    node = node_tree(key1,key2)
    nodes.append((node,c1+c2))
    nodes = sorted(nodes,key=lambda x: x[1],reverse = True)
```

```
# Print the characters and its huffmancode
huffcode=huffman_code_tree(nodes[0][0])

print(' Char | Huffman code ')
print('----------------------')
for (char,frequency)in freq:
    print(' %-4r |%12s'%(char,huffcode[char]))
```

Output:

Input String

```
212221230048 kaushika
```

Huffman Coding

Char	Huffman code
'2'	10
'1'	001
'0'	000
'k'	011
'a'	010
'3'	11101
'4'	11100
'8'	11111
' '	11110
'u'	11001
's'	11000
'h'	11011
'i'	11010

Result

Thus the huffman coding was implemented to compress the data using python programming.