**SEO tool to analyze live web pages**

**Introduction:**

SEO stands for search engine optimization—it’s the process of making your website and all its webpages easier for your customers to find on search engine results pages (SERPs). Your business will rank for a keyword and successful SEO aims to make your site rank as high as possible (at least the first page of the SERPs) for said keyword. This process can be very complicated and involved, but there are basic things you can do right now to make your site rank higher for your specific keyword. This applies if you’re a startup, a small business, or even a Fortune 500 company. Different search engines will have their own mechanisms that calculate the score of a keyword on the page based on its density and other factors and thus work out its ranking in the calculated order.

**Working:**

The user provides the keyword of interest and understands its density patterns across the various components of the HTML page. The analysis is done in a batch mode where number of webpages and the results are saved as reports and these key words can also be seen how trending they are using “Google Trends” to keep going on the right track, since it is necessary to have the right density of the right word to reach the customers. This is implemented in python and the result is stored in an excel sheet along with the chart to have a pictorial representation of the ratio of the words.

**Source Code:**

from xlrd import open\_workbook

import urllib.request

import xlwt

import xlsxwriter

import requests

from bs4 import BeautifulSoup

from xlutils.copy import copy

import sqlite3

## sample.xls is the input file with url and the words to search for

rb = open\_workbook('sample.xls')

book = xlsxwriter.Workbook('chart\_pie1.xlsx')

##format specifications for the output file

headings = ['Words', 'Count','Frequency']

bold = book.add\_format({'bold': 1})

format = book.add\_format()

format.set\_pattern(1)

format.set\_bg\_color('white')

##database for storage of word count data

conn=sqlite3.connect('mydb.db')

for s in rb.sheets():

##to drop existing tables with the same name

dqr="drop table "+s.name+";"

conn.execute(dqr)

##creation of tables

qr="create table "+s.name+"(word text, count int);"

conn.execute(qr)

print ('Sheet:',s.name,s.nrows,s.ncols)

url=s.cell(0,0).value

r = requests.get(url)

soup = BeautifulSoup(r.text,'html.parser')

type(soup)

text=soup.get\_text()

s1 = book.add\_worksheet()

print(s1.name)

s1.write\_row('A1',[url],bold)

s1.write\_row('A2', headings, bold)

#print(text)

sum1=0

count1={}

for row in range(s.nrows):

if (row>1):

print(s.cell(row,0).value)

value = (s.cell(row,0).value)

print(text.count(value))

count1[row+1]=text.count(value)

sum1=sum1+count1[row+1]

na='A'+str(row+1)

nb='B'+str(row+1)

v=[count1[row+1]]

##writing data into excel file

s1.write\_row(na,[value])

s1.write\_row(nb,v)

##inserin data into the database

qr1="insert into "+s.name+"(word, count) values (\""+str(value)+"\","+str(count1[row+1])+")"

#print(qr1)

conn.execute(qr1)

conn.commit()

##calculation of word count frequency

#print(sum1)

for val in count1:

#print(count1[val])

fq=count1[val]/sum1\*100

na='C'+str(val)

s1.write\_row(na,[fq])

##adding piechart for the data

chart = book.add\_chart({'type': 'pie'})

print("adding chart "+s1.name)

cat='='+s1.name+'!$A$3:$A$6'

print(cat)

va='='+s1.name+'!$C$3:$C$6'

chart.add\_series({'name': 'Pie Word frequency',

'categories': cat,

'values': va,

'points': [

{'fill': {'color': '#5ABA10'}},

{'fill': {'color': '#FE110E'}},

{'fill': {'color': '#CA5C05'}},

{'fill': {'color': 'yellow'}}

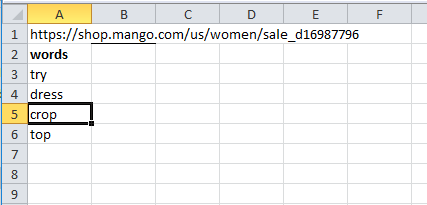
]})

s1.insert\_chart('D3', chart , {'x\_offset': 25, 'y\_offset': 10})

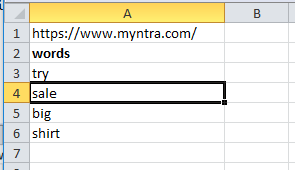
book.close()

Sample Input:

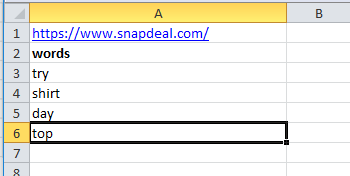
Sheet-1



Sheet-2

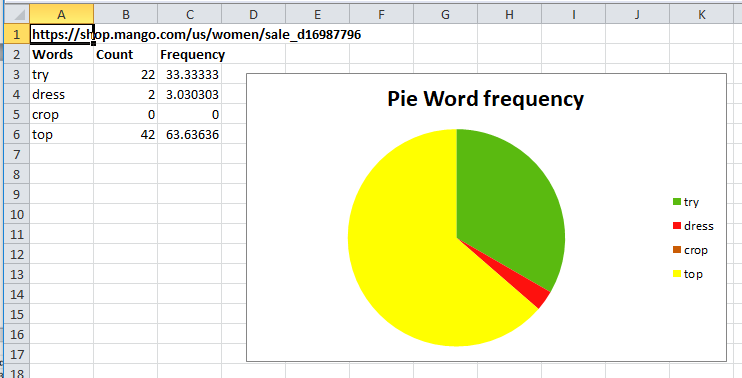


Sheet-3

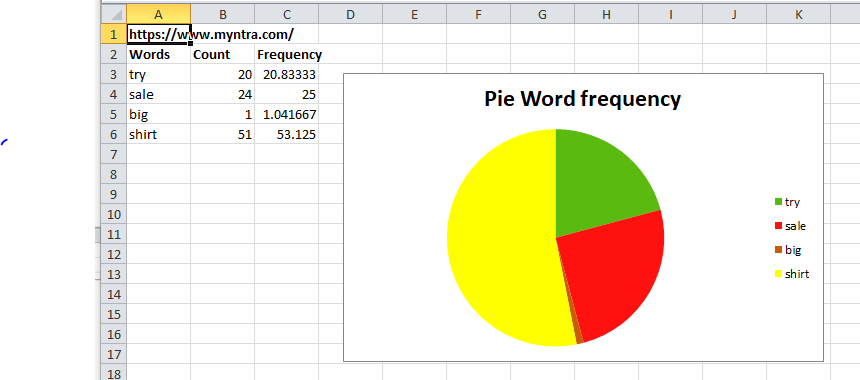


Output:

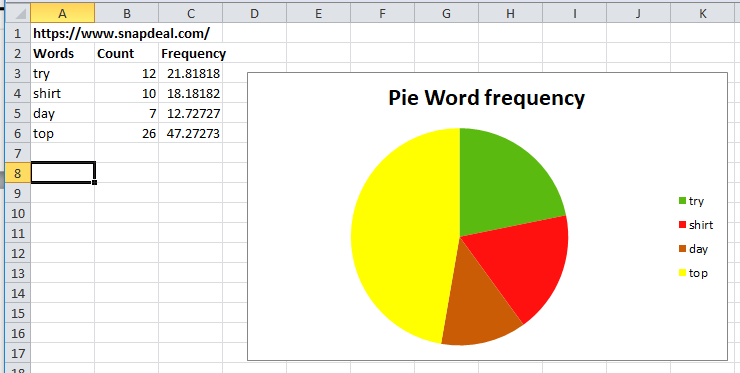
Sheet-1



Sheet-2



Sheet-3



Result on screen:

