Word cloud using python

Report submitted in partial fulfilment of the requirements to complete Term Work & Practical work of Project Based Learning (PBL).

As prescribed by

SAVITRIBAI PHULE PUNE UNIVERSITY

By

1 Darshan Ravindra Khairnar PRN NO: 72286401K

2 Swarna Ganesh Makune PRN NO: 72286488E

3 Parth Amol Pawar PRN NO: 72286612H

4 Shubham Bhanudas Godambe PRN NO: 72286255F

5 Purva Pravin Dhoakarikar PRN NO: 72286195J

6 Prathamesh Deepak Musale PRN NO: 72286533D

7 Suraj Arjun Sidankar PRN NO: 72286769H

Under the supervision of

MRS. H. A. NAVARE

**SINHGAD COLLEGE OF ENGINEERING**



First Year Engineering Department

SINHGAD COLLEGE OF ENGINEERING

44/1, Vadgaon (BK), Of Sinhgad Road Pune – 411041

Certificate

This is certify to following students,

Darshan Ravindra Khairnar Roll no. 106031

Swarna Ganesh Makune Roll no. 106032

Parth Amol Pawar Roll no. 106033

Shubham Bhanudas Godambe Roll no. 106034

Purva Pravin Dhokarikar Roll no. 106035

Prathamesh Deepak Musale Roll no. 106036

Suraj Arjun Sidankar Roll no. 106037

have completed all the Term work and Practical Work in the subject **Project Based Learning (PBL)** satisfactorily in the the department of First Year Engineering Sciences as prescribed by Savitribai Phule Pune University ,in the academic year 2021-2022.

Mrs. H. A. Navare Mr. Ganesh Pathak Prof. S. D. lokhande

Faculty-in-charge Head of Department principal

**INDEX**

* 1. ABSTRACT
* 2. MOTIVATION
* 3. INTRODUCTION
* 4. FLOWCHART
* 5. IMPLEMENTATION
* 6. RESULT
* 7. CONCLUSIONS
* 8. FUTURE SCOPE
* 9. REFERENCE

**ABSTRACT**

**Topic name:** Word Cloud using Python.

**Aim:** Breakdown of large and boring texts into illustrative way with less frequent words using Word Cloud.

**Motivation:** We chose the topic Word-Cloud using Python as it had vast day-to-day as well as industrial applications.

**Introduction:** The word cloud can be used to create an easy pictorial representation from a long information containing text.

This pictorial representation gives a summarized info of text and looks attractive to please the target audience.

**Platforms used:** The coding of the program is done in python language using ‘Pycharm’ and various libraries of python.

**Implementation**: WORDCLOUD is used in advertising pamphlets and in hoarding that can be seen on various shops.

It is also used in product advertising and marketing purpose.

Thus your audience can summarize their view of a topic.

**MOTIVATION**

From all the seven topics we received from all the members of group we had to decide one topic as our aim for the project.

We discussed and surveyed a lot about all the topics.

From those topics we chose the topic Word-Cloud using Python as it had vast day-to-day as well as industrial applications.

It too related to some of our topics like network marketing.

As it covered a lot of interest and ideas of all our members in the group we decided to select this topic.

After some further discussions we finalized this topic.

**INTRODUCTION**

A Word Cloud is a collection or cluster of words depicted in different sizes.

The bigger and bolder the word appears, the more often it is selected/voted for by an audience member.

Word Clouds are a powerful way to visualize what your audience really thinks about a topic.

Creating a simple, visual display of text data can be used to display the most prominent or frequent words in a body of text.

A ‘word cloud’ is a visual representation of word frequency.

The more commonly the term appears within the text being analyzed, the larger the word appears in the image generated.

Word clouds are increasingly being employed as a simple tool to identify the focus of written material.

They have been used in politics, business and education, for example, to visualize the content of political speeches.

The audience feels part of the presentation by contributing their answers.

The audience is interested to see the responses from their fellow audience members.

Word clouds can be good icebreakers and can provide an entry point for a topic of discussion.

Example of Word Cloud:



**FLOWCHART**

Selecting the text and amount of text for word cloud

Check for null values

Plotting the word cloud

Creating the word cloud

Adding text to a variable

**IMPLEMENTATION**

**LOGIC:** Word Clouds are visual displays of text data – simple text analysis.

Word Clouds display the most prominent or frequent words in a body of text (such as a State of the Union Address).

Typically, a Word Cloud will ignore the most common words in the language (“a”, “an”, “the” etc).

**PLATFORMS:** WORDCLOUD is used in advertising pamphlets and in

hoarding that can be seen on various shops.

It is also used in product advertising and marketing purpose.

**TOOLS and TECHNOLOGY:**

1. Computer language used – Python (Version 3.10)

2. Integrated Development Learning Environment (IDLE) used - a. Vs Code,

b. Pycharm

3. Module Required – a. Matplotlib ,

b. wordcloud,

c. Numpy.

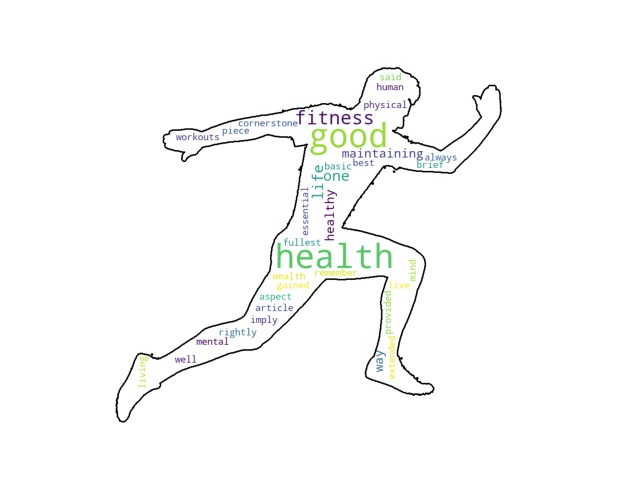
**RESULTS**

The word cloud converts the given input text in easy attractive pictorial form.

**Example 1:**

Input Text: Health is wealth, rightly said. Good health is an essential aspect of human life. Fitness is the cornerstone of good health. The best way of living a good life is by maintaining good health and fitness. One can live life to the fullest only by maintaining good health and fitness. Good health and fitness imply both mental and physical well-being. The basic way to good health and fitness is doing workouts. We should always remember that a healthy mind can only be gained by being healthy. In this article, we have provided one brief and one extended piece.

Output:



**Example 2:**

Input Text: Milk is a very crucial part of our daily diet. It has great nourishing value as it contains several nutrients like calcium, proteins, potassium, phosphorous and many vitamins. Therefore it becomes an essential part of every kid’s life. Milk is not only important for growing children but also for people of all age groups.

We get milk from mostly cows or buffaloes. There is also other milk giving animals like sheep, goats, and camels. Yogurt, cheese, and chocolates are also milk products.

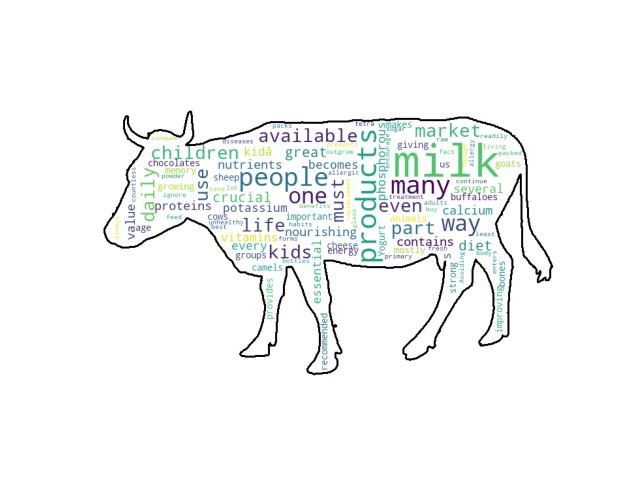
Milk makes our bones strong, provides us with energy and is even recommended for improving memory. Milk also prevents many diseases in this way. Though milk has countless benefits, people ignore them and continue living their life in an unhealthy way.

Some people are allergic to milk and milk products. Avoiding milk and milk products is the primary treatment, many children outgrow milk allergy.

There are some products such as powder milk available in the markets. People tend to use those products as a replacement for milk. In fact, not just adults use it readily but they also feed it to their kids. Kids even like to consume it raw and in that way, a lot of sugar enters their body. One must not encourage such habits.

Milk is available in many forms in the market, such as in tetra packs and bottles. But it is always best to buy fresh milk and not the packed one. We must have at least one glass of milk daily.

Output:



**Example 3:**

Input text: Listed pricing is Maximum Retail Price (inclusive of all taxes).

1TB = 1 trillion bytes; actual formatted capacity less.

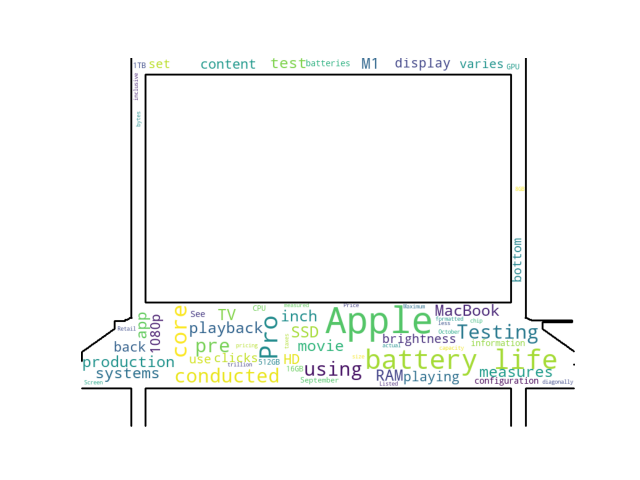
Screen size is measured diagonally.

Testing conducted by Apple in October 2020 using pre-production 13-inch MacBook Pro systems with Apple M1 chip, 8GB of RAM and 512GB SSD. The Apple TV app movie playback test measures battery life by playing back 1080p HD content with display brightness set to 8 clicks from the bottom. Battery life varies by use and configuration. See apple.com/in/batteries for more information.

Testing conducted by Apple in September 2021 using pre-production 14-inch MacBook Pro systems with Apple M1 Pro, 8-core CPU, 14-core GPU, 16GB of RAM and 512GB SSD. The Apple TV app movie playback test measures battery life by playing back 1080p HD content with display brightness set to 8 clicks from the bottom. Battery life varies by use and configuration. See apple.com/in/batteries for more information.

Testing conducted by Apple in September 2021 using pre-production 16-inch MacBook Pro systems with Apple M1 Pro, 10-core CPU, 16-core GPU, 16GB of RAM and 1TB SSD. The Apple TV app movie playback test measures battery life by playing back 1080p HD content with display brightness set to 8 clicks from the bottom. Battery life varies by use and configuration. See apple.com/in/batteries for more information.

Output:

**CONCLUSION**

* Understand what your audience/employees feel about a topic/situation
* Your audience can summarize their view of a topic
* Allows you to identify what is important to your audience
* You can measure audience understanding of a topic.

**FUTURE-SCOPE**

1. In future, some automatic color selection for different words according to their size and occurrence in the text can be developed.

2. In future, animations can be added to word cloud to improve the presentation of information to please the audience.

3. In future, we can add auto selection of the shape (png) of the word cloud using an A.I. that analysis the text for doing so.

4. It can be made convenient for even small scale industries to use it.

**REFERENCE**

1. Used data from ***Datacamp.com***

2. YouTube Channel for reference-: ***Python Enthusiast***

3. Photo Gallery used -: ***Pinterest.***

4. A copy of Jim Flanagan's Search Referral Zeitgeist was available at ***archive.org*** but has since been blocked. In the comments of a ***blog entry Archived*** 2006-04-26 at the ***Wayback Machine***, a user identified as Steve Minutillo attribute the idea to Jim Flanagan, stating that Flanagan's site had such displays in 2002.