

# WordCloudZuckerberg

April 17, 2020

```
[1]: # Roll no: 33140 (Sahil Naphade)
# Batch: L9
# Assignment 8(B): WordCloud on Mark Zuckerberg text file
import pandas as pd
from wordcloud import WordCloud
import nltk
from nltk import word_tokenize
import matplotlib.pyplot as plt
from nltk.corpus import stopwords
# nltk.download('stopwords')

[2]: data = open("G:/College/SI-VI DataSets/TextMining/MarkZuckerberg.txt", "r").read()
data = data.lower()

[3]: stop = stopwords.words('english')

[4]: def create_word_cloud(string):
    cloud = WordCloud(width=800, height=800, background_color =
→ "white", stopwords = stop).generate(string)
    plt.figure(figsize = (10, 10), facecolor = None)
    plt.axis('off')
    plt.imshow(cloud)

[5]: create_word_cloud(data)
```

estimated andrew rb various year  
world people us ranked  
phase since worth campuses powerful chairman saverin  
involved harvard legal march  
co dormitory friends mcollum named  
cofounders currently time may person  
introduced room among based  
forbes influential internet eduardo became  
college elliot  
roommates expanded rapidly born university involvement  
company american  
officer upon net one claimed chief brought  
wealthiest december list fellow  
zuckerberg due  
part hughes share students entrepreneur  
billion magazine february award  
dustin moskovitz mark group rg users  
disputes