

LAB 6 : Text Data Preparation and Exploration

วัดถูประสงค์

นศ. สามารถเขียนโปรแกรมด้วยภาษา Python เพื่อ Text Data Preparation and Exploration

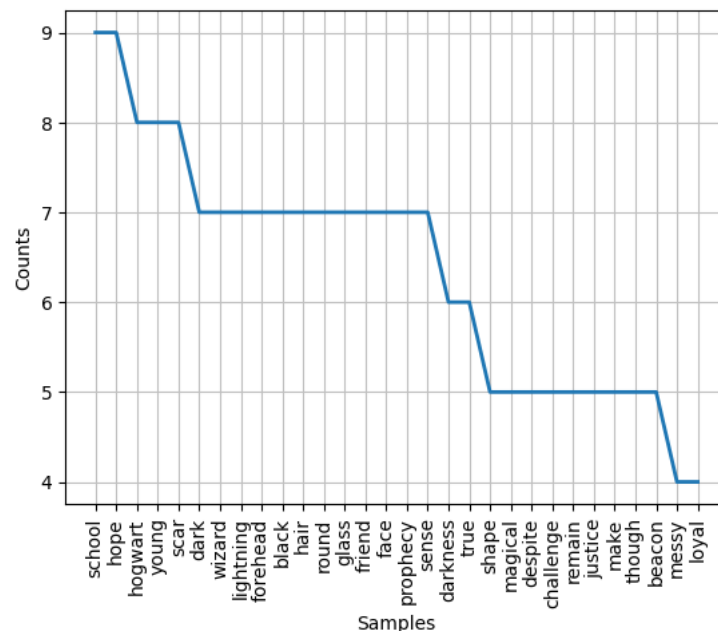
* ไม่อนุญาตให้ส่งงานหลังกำหนด และนศ.ต้องตรวจสอบสถานะการส่งงาน หลัง submit งานทุกครั้ง

ตอนที่ 1

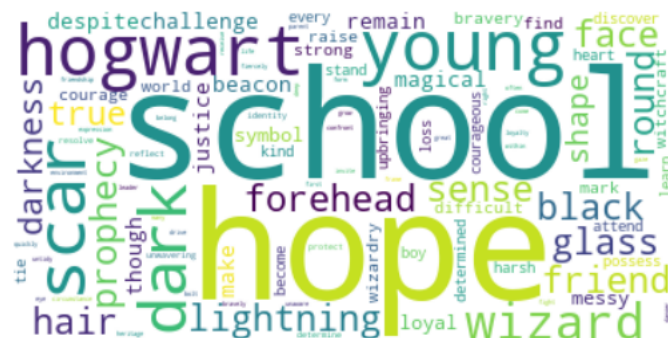
คำสั่ง กำหนดให้ใช้ไฟล์ข้อมูลชื่อ unknown1.csv สำหรับนักศึกษาที่มีรหัสลงท้ายด้วยเลขคี่ และไฟล์ข้อมูลชื่อ unknown2.csv สำหรับนักศึกษาที่มีรหัสลงท้ายด้วยเลขคู่ และให้ใช้ไฟล์ต้นแบบสำหรับการเขียนโปรแกรม ที่ชื่อว่า

Assignment6_TextDataExploration.ipynb

1. ให้ทำการอ่านข้อมูลจากไฟล์ และเตรียมข้อมูลให้พร้อมก่อนการประมวลผล
2. ให้ทำการสร้างกราฟแสดงความถี่ของคำ และให้ทำการ Capture ภาพของกราฟแสดงความถี่ของคำ มาแสดงในช่องว่างที่เว้นไว้



3. ให้ทำการสร้าง WordCloud แสดงความถี่ของคำ และให้ทำการ Capture ภาพของ WordCloudแสดงความถี่ของคำ มาแสดงในช่องว่างที่เว้นไว้



4. ให้สรุปว่าเนื้อหาของข้อมูลในไฟล์ข้อมูลที่ได้รับมอบหมาย กล่าวเกี่ยวกับประเด็นใด โรงเรียนเวทีย์มันต์ (Harry Potter)

ตอนที่ 2

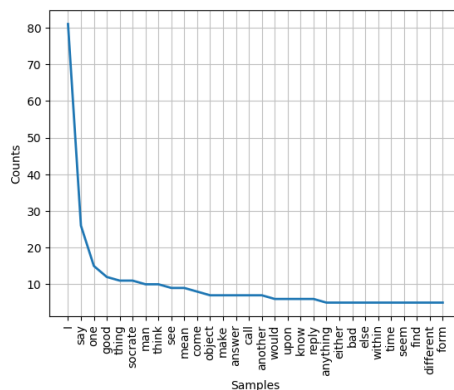
- ให้นักศึกษาเลือกประเด็นที่สนใจสร้าง Wordcloud และระบุประเด็นหรือหัวข้อที่สนใจ ปรัชญาในอดีต
- ทำการสร้างไฟล์ข้อความภาษาอังกฤษ จัดเก็บในไฟล์ .csv จำนวนไม่น้อยกว่า 10 ระเบียบ โดยในเนื้อหาให้พยายามตัดคำที่เป็นส่วนหนึ่งของประเด็นหรือหัวข้อที่สนใจ หรือให้ปรากฏน้อยที่สุด
- ให้ทำการ Capture ภาพของตัวอย่างข้อมูลที่สร้างขึ้น

```

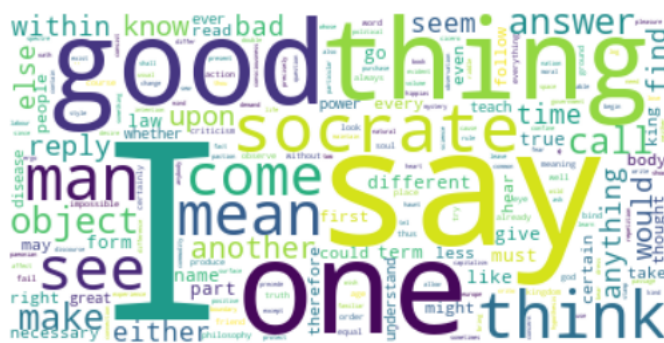
author                                sentence_spacy
302934      Marx      spectre haunt europe spectre communism
110380      Spinoza   experience teach I usual surrounding social li...
331701      Nietzsche intention within short time confront fellow ma...
322679      Keynes   capitalism faint heart
212550      Deleuze   year I refer europe certain way thinking corre...
...
17092       Plato    I guess socrate say I right hippocrate sophist...
206978      Deleuze   I representation relate representation free fa...
113759      Leibniz   I I could gain true understanding dogma find s...
219967      Merleau-Ponty I meet person say alain prepared admit eye pre...
17807       Plato    I meet person say alain prepared admit eye pre...
I confident
[140 rows x 2 columns]

```

- ให้ทำการอ่านข้อมูลจากไฟล์ที่สร้าง และเตรียมข้อมูลให้พร้อมก่อนการประมวลผล
- ให้ทำการสร้างกราฟแสดงความถี่ของคำ และให้ทำการ Capture ภาพของกราฟแสดงความถี่ของคำ มาแสดงในช่องว่างที่เว้นไว้



- ให้ทำการสร้าง WordCloud แสดงความถี่ของคำ และให้ทำการ Capture ภาพของ WordCloud แสดงความถี่ของคำ มาแสดงในช่องว่างที่เว้นไว้



- ให้สรุปว่าเนื้อหาของข้อมูลในไฟล์ข้อมูลที่ได้รับมอบหมาย กล่าวเกี่ยวกับประเด็นใด ปรัชญาในอดีต