Konlpy 및 WordCloud 설치

**1.Konlpy 설치**

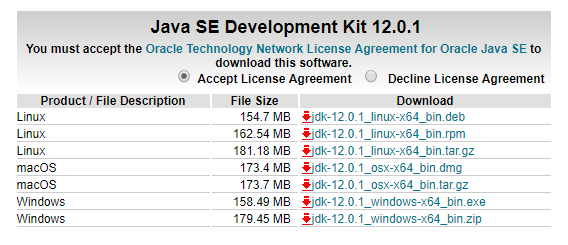
Terminal 접속후 pip install konlpy 입력하여 설치

Konlpy – 한국어 처리에 필요

이후 JAVA JDK를 설치 해줘야한다.

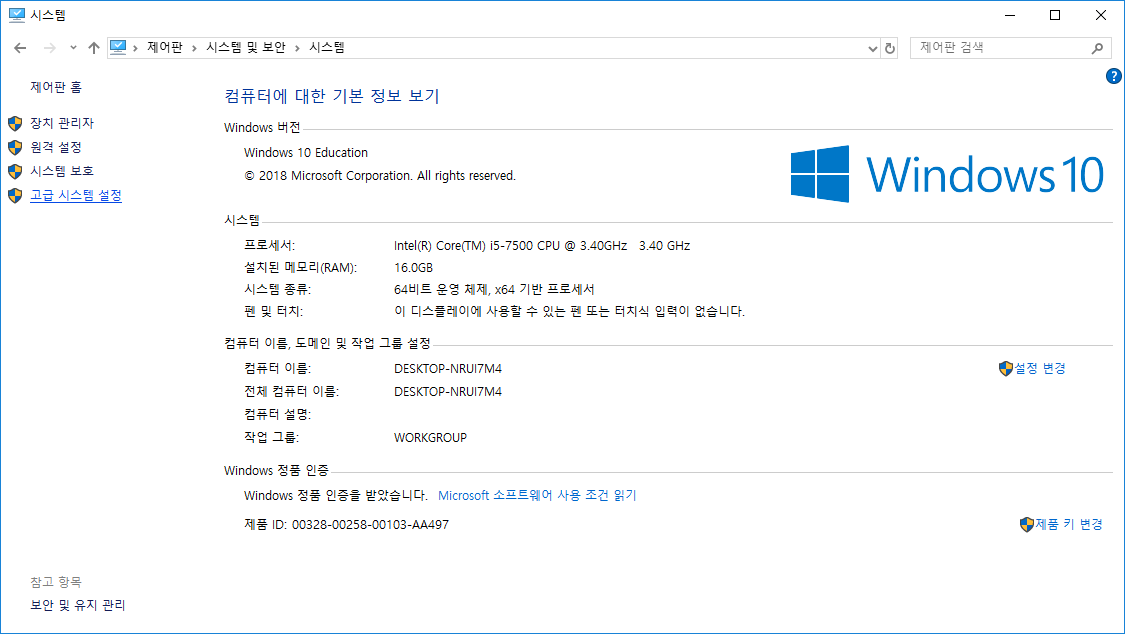
<https://www.oracle.com/technetwork/java/javase/downloads/jdk12-downloads-5295953.html>

접속하여 자신에게 맞는 운영체제를 골라서 실행한다.

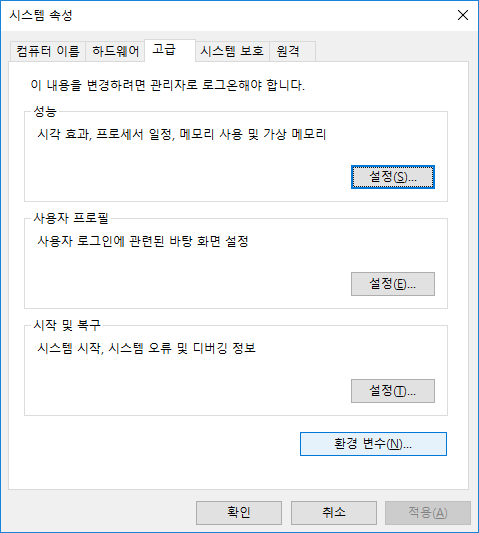


Jdk 설치

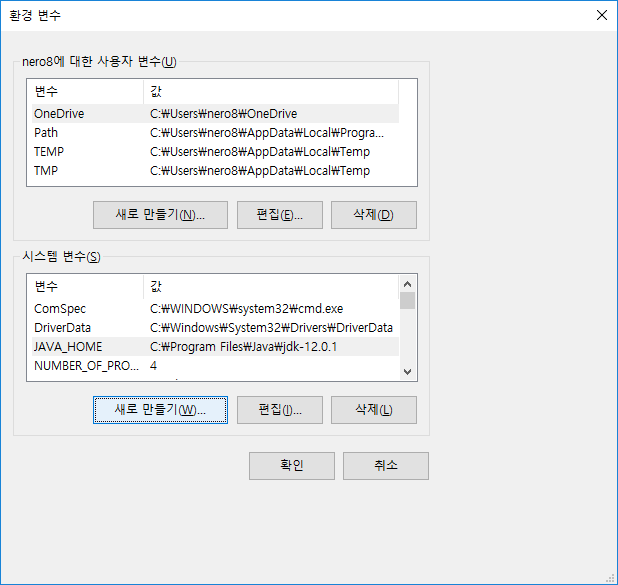
이후 제어판 -> 시스템 및 보안 -> 시스템을 들어가서 JAVA\_HOME을 설정해주어야 한다.



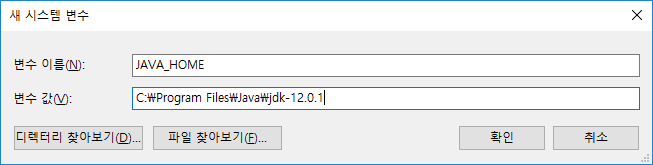
시스템-> 고급 시스템 설정 클릭



고급 탭-> 환경 변수 클릭



새로 만들기 클릭

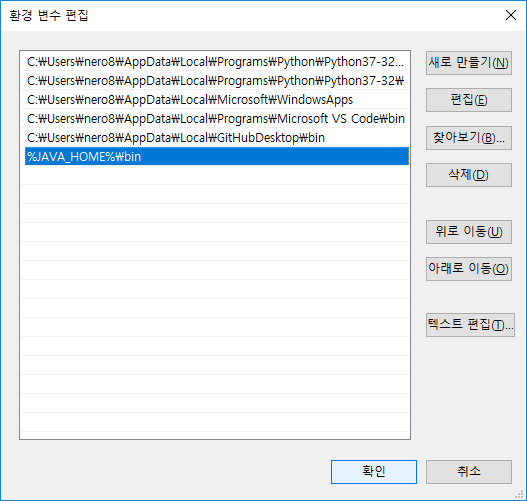


JAVA\_HOME 설정

JAVA\_HOME 설정시 변수 값은 자신의 jdk 가 설치된 경로를 입력해주어야 한다.

보통 C:\Program Files\Java\jdk-12.0.1(설치 버전에 따라 다름) 경로에 설치 된다.

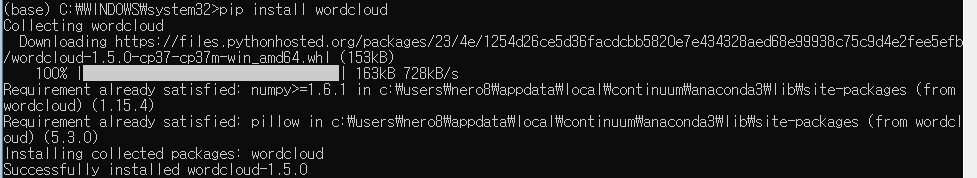
확인을 누르고 사용자 변수에 있는 Path 클릭 후 편집을 눌러서 ‘%JAVA\_HOME%\bin’을 입력해준다.



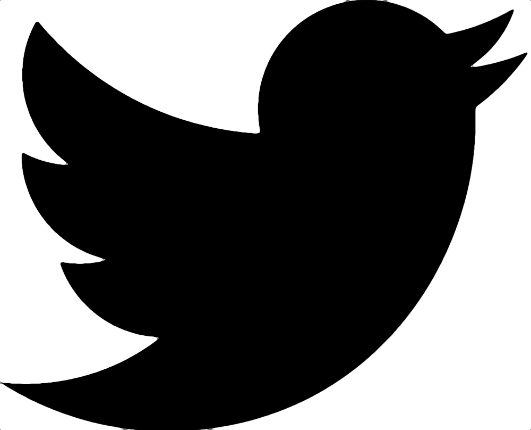
**2. WordCloud 설치**

Anaconda prompt에서 Pip install wordcloud 입력.

오류 발생 시 pip install pillow 입력



**3. WordCloud 실행 해보기**



WordCloud를 사용할 때에는 검은색으로 채워진 사진을 사용한다.

기존 파란색 이미지 사용시 그림에 제대로 적용이 되지 않음.

from wordcloud import WordCloud, STOPWORDS

import numpy as np

from PIL import Image

text = open('./09. alice.txt').read()

alice\_mask = np.array(Image.open('./twitter\_mask.png'))

stopwords = set(STOPWORDS)

stopwords.add("said")

import matplotlib.pyplot as plt

import platform

path = "c:/Windows/Fonts/malgun.ttf"

from matplotlib import font\_manager, rc

if platform.system() == 'Darwin':

rc('font', family='AppleGothic')

elif platform.system() == 'Windows':

font\_name = font\_manager.FontProperties(fname=path).get\_name()

rc('font', family=font\_name)

else:

print('Unknown system... sorry~~~~')

%matplotlib inline

plt.figure(figsize=(8,8))

plt.imshow(alice\_mask, cmap=plt.cm.gray, interpolation='bilinear')

plt.axis('off')

plt.show()

wc = WordCloud(background\_color='white', max\_words=2000, mask=alice\_mask,

stopwords = stopwords)

wc = wc.generate(text)

wc.words\_

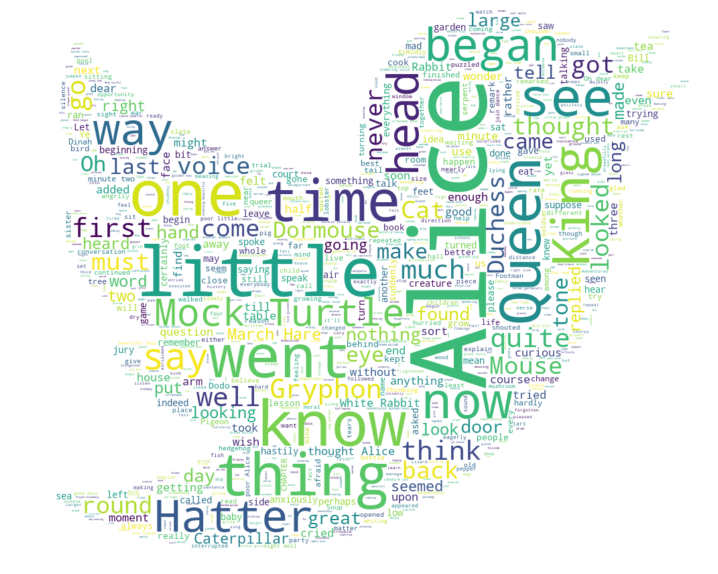
plt.figure(figsize=(12,12))

plt.imshow(wc, interpolation='bilinear')

plt.axis('off')

plt.show()

현재는 alice.text 에 저장된 단어들중 사용된 비중이 높은 것들은 크기가 크게 나타남을 보여준다.



적용된 결과

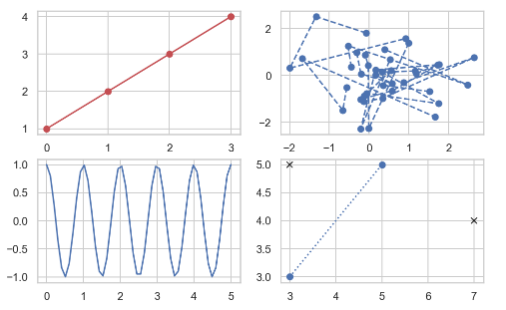
Wordcloud 이외의 시각화 라이브러리

1. Matplotlib

터미널에서 ‘Pip3 install matplotlib’ 입력하여 설치.

기본적으로 가장 많이 사용되는 시각화 라이브러리이다.

간단한 그래프나 차트 등으로 보여준다.



1. Folium

지도 데이터에 위치정보를 시각화하는 라이브러리이다.

