

```

1 !sudo apt-get install -y fonts-nanum
2 !sudo fc-cache -fv
3 !rm ~/.cache/matplotlib -rf

```

```

Reading package lists... Done
Building dependency tree
Reading state information... Done
The following package was automatically installed and is no longer required:
  libnvidia-common-460
Use 'sudo apt autoremove' to remove it.
The following NEW packages will be installed:
  fonts-nanum
0 upgraded, 1 newly installed, 0 to remove and 20 not upgraded.
Need to get 9,604 kB of archives.
After this operation, 29.5 MB of additional disk space will be used.
Get:1 http://archive.ubuntu.com/ubuntu bionic/universe amd64 fonts-nanum all 20170925-1 [9,604 B]
Fetched 9,604 kB in 1s (8,265 kB/s)
debconf: unable to initialize frontend: Dialog
debconf: (No usable dialog-like program is installed, so the dialog based frontend cannot be used)
debconf: falling back to frontend: Readline
debconf: unable to initialize frontend: Readline
debconf: (This frontend requires a controlling tty.)
debconf: falling back to frontend: Teletype
dpkg-preconfigure: unable to re-open stdin:
Selecting previously unselected package fonts-nanum.
(Reading database ... 157604 files and directories currently installed.)
Preparing to unpack .../fonts-nanum_20170925-1_all.deb ...
Unpacking fonts-nanum (20170925-1) ...
Setting up fonts-nanum (20170925-1) ...
Processing triggers for fontconfig (2.12.6-0ubuntu2) ...
/usr/share/fonts: caching, new cache contents: 0 fonts, 1 dirs
/usr/share/fonts/truetype: caching, new cache contents: 0 fonts, 3 dirs
/usr/share/fonts/truetype/humor-sans: caching, new cache contents: 1 fonts, 0 dirs
/usr/share/fonts/truetype/liberation: caching, new cache contents: 16 fonts, 0 dirs
/usr/share/fonts/truetype/nanum: caching, new cache contents: 10 fonts, 0 dirs
/usr/local/share/fonts: caching, new cache contents: 0 fonts, 0 dirs
/root/.local/share/fonts: skipping, no such directory
/root/.fonts: skipping, no such directory
/var/cache/fontconfig: cleaning cache directory
/root/.cache/fontconfig: not cleaning non-existent cache directory
/root/.fontconfig: not cleaning non-existent cache directory
fc-cache: succeeded

```

▼ 1. 문자열 처리

```

1 text = '챗봇(chatbot) 혹은 채터봇(chatterbot)은 음성이나 문자를 통한 인간과의 대화를 통해서 특정한

1 text_upper = text.upper()

```

```
2 text_upper
```

'챗봇(CHATBOT) 혹은 채터봇(CHATTERBOT)은 음성이나 문자를 통한 인간과의 대화를 통해서 특정한 작업을 수행하도록 제작된 컴퓨터 프로그램이다. 토크봇(TALKBOT), 채터박스(CHATTERBOX) 혹은 그냥 봇(BOT)라고도 한다.'

```
1 text_lower = text.lower()
```

```
2 text_lower
```

'챗봇(chatbot) 혹은 채터봇(chatterbot)은 음성이나 문자를 통한 인간과의 대화를 통해서 특정한 작업을 수행하도록 제작된 컴퓨터 프로그램이다. 토크봇(talkbot), 채터박스(chatterbox) 혹은 그냥 봇(bot)라고도 한다.'

```
1 text.find('컴퓨터')
```

```
75
```

```
1 pos = text.find('컴퓨터')
```

```
2 text[pos : pos+len('컴퓨터 프로그램')]
```

```
'컴퓨터 프로그램'
```

```
1 '인공지능' in text
```

```
False
```

```
1 len(text)
```

```
137
```

```
1 text.count('bot')
```

```
4
```

▼ 2. WordCloud 생성

```
1 from google.colab import files
```

```
2 files.upload()
```



\xea\xb7\xb8\xeb\x9f\xac\xeb\x82\x98 \xec\x9d\xb4\xea\xb2\x83\xec\x9d\x80
\xed\x99\x95\xec\x8b\xa4\xed\x95\x98\xea\xb2\x8c \xec\xa0\x95\xed\x95\xb4\xec\xa7\x84
\xea\xb2\x83\xec\x9d\xb4 \xec\x95\x84\xeb\x8b\x88\xeb\xaf\x80\xeb\xa1\x9c
\xea\xb3\x84\xec\x86\x8d \xeb\xb3\x80\xed\x95\x98\xea\xb3\xa0 \xec\x9e\x88\xeb\x8b\xa4.
\xeb\x98\x90\xed\x95\x9c, \xec\xb1\x97\xeb\xb4\x87\xec\x9d\x84
\xeb\xa7\x8c\xeb\x93\xa4\xeb\x95\x8c \xed\x9d\x94\xed\x9e\x88 \xed\x95\x98\xeb\x8a\x94
\xec\x8b\xa4\xed\x8c\xa8\xec\x9a\x94\xec\x9d\xb8\xec\x9d\xb4
\xec\x9e\x88\xec\x9c\xbc\xeb\x8b\x88 \xec\xa3\xbc\xec\x9d\x98\xed\x95\xb4\xec\x95\xbc
\xed\x95\x9c\xeb\x8b\xa4.[9]\r\n\r\n\xec\xb6\x9c\xeb\xa0\xa5
\xec\x83\x9d\xec\x84\xb1\xec\x97\x90 \xec\x82\xac\xec\x9a\xa9\xeb\x90\x98\xeb\x8a\x94
\xeb\x8d\xb0\xec\x9d\xb4\xed\x84\xb0\xeb\xb2\xa0\xec\x9d\xb4\xec\x8a\xa4\xea\xb0\x80
\xea\xb3\xa0\xec\xa0\x95\xeb\x90\x98\xea\xb3\xa0
\xec\xa0\x9c\xed\x95\x9c\xeb\x90\x98\xec\x96\xb4 \xec\x9e\x88\xea\xb8\xb0
\xeb\x95\x8c\xeb\xac\xb8\xec\x97\x90 \xed\x95\x99\xec\x8a\xb5\xeb\x90\x98\xec\xa7\x80
\xec\x95\x8a\xea\xb1\xb0\xeb\x82\x98 \xec\xa0\x80\xec\x9e\xa5\xeb\x90\x98\xec\xa7\x80
\xec\x95\x8a\xec\x9d\x80 \xec\xa7\x88\xeb\xac\xb8\xec\x9d\x80
\xec\xb1\x97\xeb\xb4\x87\xec\x9d\xb4 \xec\xb2\x98\xeb\xa6\xac\xeb\xa5\xbc
\xec\x8b\xa4\xed\x8c\xa8\xed\x95\xa0 \xec\x88\x98 \xec\x9e\x88\xeb\x8b\xa4.
[10]\r\n\xec\xb1\x97\xeb\xb4\x87\xec\x9d\x80 \xec\x96\xb8\xec\x96\xb4
\xed\x94\x84\xeb\xa1\x9c\xec\x84\xb8\xec\x8a\xa4\xec\x97\x90 \xed\x81\xac\xea\xb2\x8c
\xec\x98\x81\xed\x96\xa5\xec\x9d\x84 \xeb\xb0\x9b\xec\x9c\xbc\xeb\xa9\xb0
\xec\x95\xa1\xec\x84\xbc\xed\x8a\xb8 \xeb\xb0\x8f \xec\x8b\xa4\xec\x88\x98\xec\x99\x80
\xea\xb0\x99\xec\x9d\x80
\xeb\xb6\x88\xea\xb7\x9c\xec\xb9\x99\xec\x84\xb1\xec\x9c\xbc\xeb\xa1\x9c
\xec\x9d\xb8\xed\x95\xb4 \xec\xa0\x9c\xed\x95\x9c\xeb\x90\xa0 \xec\x88\x98
\xec\x9e\x88\xeb\x8b\xa4.[11]\r\n\xec\xb1\x97\xeb\xb4\x87\xec\x9d\x80
\xeb\x8f\x99\xec\x8b\x9c\xec\x97\x90 \xec\x97\xac\xeb\x9f\xac
\xec\xa7\x88\xeb\xac\xb8\xec\x9d\x84 \xec\xb2\x98\xeb\xa6\xac\xed\x95\xa0 \xec\x88\x98
\xec\x97\x86\xec\x9c\xbc\xeb\xaf\x80\xeb\xa1\x9c \xec\x82\xac\xeb\x9e\x8c\xea\xb3\xbc
\xeb\x8c\x80\xed\x99\x94\xed\x95\xa0 \xeb\x95\x8c \xeb\xb3\xb4\xeb\x8b\xa4
\xed\x9a\xa8\xec\x9c\xa8\xec\x84\xb1\xec\x9d\xb4 \xeb\x96\xa8\xec\x96\xb4\xec\xa7\x88
\xec\x88\x98 \xec\x9e\x88\xeb\x8b\xa4.\r\n\xec\xb1\x97\xeb\xb4\x87\xec\x9d\x80
\xed\x95\x99\xec\x8a\xb5\xea\xb2\xbd\xed\x97\x98\xec\x9d\x84 \xec\x9c\x84\xed\x95\xb4
\xeb\xa7\x8e\xec\x9d\x80 \xec\x96\x91\xec\x9d\x98 \xeb\x8c\x80\xed\x99\x94
\xeb\x8d\xb0\xec\x9d\xb4\xed\x84\xb0\xea\xb0\x80
\xed\x95\x84\xec\x9a\x94\xed\x95\x98\xeb\x8b\xa4.\r\n\xec\xb1\x97\xeb\xb4\x87\xec\x9d\x80
\xec\x82\xac\xec\x9a\xa9\xec\x9e\x90\xec\x99\x80 \xec\x9d\xb4\xec\xa0\x84\xec\x9d\x98
\xeb\x8c\x80\xed\x99\x94\xeb\xa5\xbc \xec\xa3\xbc\xea\xb3\xa0
\xeb\xb0\x9b\xec\x95\x84\xec\x95\xbc \xed\x95\x98\xeb\x8a\x94
\xeb\xb9\x84\xec\x84\xa0\xed\x98\x95 \xeb\x8c\x80\xed\x99\x94\xeb\xa5\xbc
\xec\xb2\x98\xeb\xa6\xac\xed\x95\x98\xeb\x8a\x94\xeb\x8d\xb0
\xec\x96\xb4\xeb\xa0\xa4\xec\x9b\x80\xec\x9d\xb4 \xec\x9e\x88\xeb\x8b\xa4.
[12]\r\n\xec\x9d\xbc\xeb\xb0\x98\xec\xa0\x81\xec\x9c\xbc\xeb\xa1\x9c \xea\xb8\xb0\xec\xa1\xb4
\xec\x84\x9c\xeb\xb9\x84\xec\x8a\xa4\xec\x9d\x98 \xea\xb8\xb0\xec\x88\xa0
\xec\xa3\xbc\xeb\x8f\x84\xec\xa0\x81 \xeb\xb3\x80\xed\x99\x94\xec\x97\x90\xec\x84\x9c
\xeb\xb0\x9c\xec\x83\x9d\xed\x95\x98\xeb\x8a\x94 \xea\xb2\x83\xec\xb2\x98\xeb\x9f\xbc
\xec\x9d\xbc\xeb\xb6\x80 \xec\x86\x8c\xeb\xb9\x84\xec\x9e\x90\xeb\x8a\x94
\xec\xb1\x97\xeb\xb4\x87\xec\x97\x90 \xeb\xb6\x88\xed\x8e\xb8\xed\x95\xa8\xec\x9d\x84
\xeb\x8a\x90\xeb\x81\xbc\xea\xb3\xa0 \xec\x9e\x90\xec\x8b\xa0\xec\x9d\x98
\xec\x9a\x94\xec\xb2\xad\xec\x9d\xb4 \xea\xb8\xb0\xea\xb3\x84\xec\x97\x90
\xec\x9d\x98\xed\x95\xb4 \xec\xb2\x98\xeb\xa6\xac\xeb\x90\x98\xea\xb3\xa0
\xec\x9e\x88\xec\x9d\x8c\xec\x9d\x84 \xeb\xb6\x84\xeb\xaa\x85\xed\x9e\x88 \xec\x95\x8c
\xec\x88\x98 \xec\x9e\x88\xeb\x8b\xa4.[11]'




```
1 !pip install koNLpy
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple
Collecting koNLpy
  Downloading konlpy-0.6.0-py2.py3-none-any.whl (19.4 MB)
    |████████████████████████████████████████| 19.4 MB 1.4 MB/s
Collecting JPype1>=0.7.0
  Downloading JPype1-1.4.0-cp37-cp37m-manylinux_2_5_x86_64.manylinux1_x86_64.whl (453 kB)
    |████████████████████████████████████████| 453 kB 56.9 MB/s
Requirement already satisfied: numpy>=1.6 in /usr/local/lib/python3.7/dist-packages (from koNLpy)
Requirement already satisfied: lxml>=4.1.0 in /usr/local/lib/python3.7/dist-packages (from koNLpy)
Requirement already satisfied: typing-extensions in /usr/local/lib/python3.7/dist-packages (from koNLpy)
Installing collected packages: JPype1, koNLpy
Successfully installed JPype1-1.4.0 koNLpy-0.6.0
```

```
1 from wordcloud import WordCloud
2 from wordcloud import STOPWORDS
3 import matplotlib.pyplot as plt
4
5 from PIL import Image
6 import numpy as np
```

```
1 f = open('./챗봇.txt', 'r', encoding='UTF8')
2 text = f.read()
3 text
```