## 네이버 영화에서 제목과 별점정보만 추출하기

## In [1]:

import requests from bs4 import BeautifulSoup

#### In [2]:

response = requests.get('https://movie.naver.com/movie/running/current.naver#') response

#### Out[2]:

<Response [200]>

#### In [3]:

html = response.text html

#### Out[3]:

'\r\n\r\n\r\n\r\n\r\n\r\n\r\n\r\n\t\r\n\t\r\n\t\r\n\t\r\n\t\r\n \t\r\n\r\n<!DOCTYPE html>\r\n<html lang="ko">\r\n<head>\r \n<meta charset="utf-8">\r\n<meta http-equiv="X-UA-Compati ble" content="IE=edge">\r\n<meta http-equiv="imagetoolbar" content="no">\r\n<title>현재 상영영화 : 네이버 영화</title>\r\n\r meta property="me2:image" content="https://ssl.pstatic.net/i mgmovie/today/naverme/naverme\_profile.jpg"/>\r\n<meta pr operty="me2:post\_tag" content="네이버영화 현재상영작"/>\r\n< meta property="me2:category1" content="네이버영화"/>\r\n<m eta property="me2:category2" content="현재상영작"/>\r\n\r\n\r\n\t\r\n\t\t<meta property="og:title" content="현재상영작 : 네이버 영화" />\r\n\t\r\n\t\r\n\t\r\n\r\n\r\n<meta property="og:t ype" content="article"/>\r\n\r\n<meta property="og:url" conte nt="https://movie.naver.com/movie/running/current.naver"/> \r\n\r\n\t\r\n\t\r\n\t\t<meta property="og:image" content ="https://ssl.pstatic.net/static/m/movie/icons/OG 270 270.pn g"/>\r\n\t\r\n\r\n\t\r\n\t<meta property="og:description" c ontent="상영 중 영화의 예매율/평점/좋아요 순 정보 제공."/>\r\n\r\n\r \n\r\n<meta property="og'article'thumbnail!!rl" content=

#### In [4]:

```
soup=BeautifulSoup(html, 'html.parser')
soup
```

## Out[4]:

```
<!DOCTYPE html>
<html lang="ko">
<head>
<meta charset="utf-8"/>
<meta content="IE=edge" http-equiv="X-UA-Compatible"/>
<meta content="no" http-equiv="imagetoolbar"/>
<title>현재 상영영화: 네이버 영화</title>
<meta content="https://ssl.pstatic.net/imgmovie/today/naver me/naverme_profile.jpg" property="me2:image">
<meta content="네이버영화 현재상영작" property="me2:post_tagg">
<meta content="네이버영화" property="me2:category1">
<meta content="현재상영작" property="me2:category2"/>
<meta content="현재상영작: 네이버 영화" property="og:title"/>
<meta content="article" property="og:type"/>
<meta content="https://movie.naver.com/movie/running/curre nt.naver" property="og:url"/>
<meta content="https://ssl.pstatic.net/static/m/movie/icons/O"</pre>
```

## In [5]:

```
title=soup.select('#content > div.article > div:nth-child(1) > div.lst_wrap > ul > li > dl
title
title[0].text

titleList=[]
for i in range(0,10):
   titleList.append(title[i].text)
print(titleList)
```

```
['씽2게더', '경관의 피', '스파이더맨: 노 웨이 홈', '킹스맨: 퍼스트 에이전트', ^'드라이브 마이 카', '라임크라임', '해탄적일천', '해피 뉴 이어', '해피 아워', '코다']
```

## In [6]:

```
starRate = soup.select('#content > div.article > div:nth-child(1) > div.lst wrap > ul > l
starRate
starRate[0].text
Out[6]:
'9.41'
In [7]:
starRateArray = []
for i in range(0,10):
 starRateArray.append(starRate[i].text)
print(starRateArray)
float_list = list(map(float, starRateArray)) #문자열배열을 숫자배열로 변환
print(float list)
['9.41', '7.71', '8.92', '8.09', '8.73', '8.62', '8.48', '7.12', '9.10', '9.23']
[9.41, 7.71, 8.92, 8.09, 8.73, 8.62, 8.48, 7.12, 9.1, 9.23]
In [8]:
for i in range(0,10):
 print("[결과, 예매순 상위 10개]")
 print("{}.{} / 별점:{}".format(i+1,titleList[i], starRateArray[i]))
[결과, 예매순 상위 10개]
1.씽2게더 / 별점: 9.41
[결과, 예매순 상위 10개]
2.경관의 피 / 별점: 7.71
[결과, 예매순 상위 10개]
3.스파이더맨: 노 웨이 홈 / 별점: 8.92
[결과, 예매순 상위 10개]
4.킹스맨: 퍼스트 에이전트 / 별점: 8.09
「결과. 예매순 상위 10개]
5.드라이브 마이 카 / 별점 : 8.73
「결과、예매순 상위 10개]
6.라임크라임 / 별점: 8.62
[결과, 예매순 상위 10개]
7.해탄적일천 / 별점: 8.48
[결과, 예매순 상위 10개]
8.해피 뉴 이어 / 별점: 7.12
[결과, 예매순 상위 10개]
9.해피 아워 / 별점: 9.10
[결과, 예매순 상위 10개]
10.코다 / 별점: 9.23
```

# 예매순 상위 10개의 별점 점수에 대해 그래프로 시각화하기

## In [9]:

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

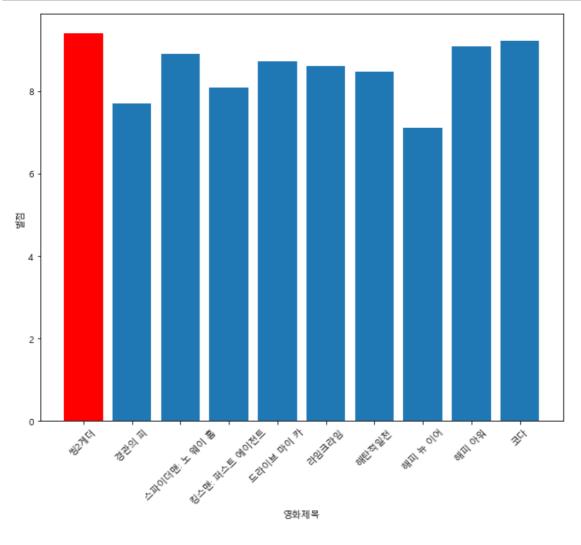
## In [39]:

```
#x축: titleList[]
plt.figure(figsize=(10, 8))
x_pos = np.arange(len(titleList))

#x축 위에 올릴 세로 bar
#plt.bar(x축, 세로로 쓸 내용)
bar_list = plt.bar(x_pos, float_list)

#한글 폰트 지정
plt.rc('font', family = 'Malgun Gothic')
#print(plt.rcParams['font.family'])

bar_list[len(float_list)-10].set_color('r') #bar_list의 첫번째 바를 색을 r로 지정
plt.xlabel('영화제목')
plt.ylabel('병접')
plt.xticks(x_pos, titleList, rotation =45)
plt.show()
```



In [14]:

import matplotlib matplotlib.matplotlib\_fname()

Out[14]:

'C:\\Users\\YJ\\anaconda3\\lib\\site-packages\\matplotlib\\mpl-d
ata\\matplotlibrc'

In []: