**탐색적 데이터 분석 과정**

**필요 라이브러리 import**

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seabon as sns

%matplotlib inline

**데이터 읽어오기 및 전체적으로 살펴보기**

#titanic = sns.load\_dataset('titanic')

titanic = pd.read\_csv('data/titanic.csv')

**Head나 tail 확인**

titanic.head())

titanic.info()

**데이터 가공**

**이상치 및 결측치 확인**

예) 결측치를 찾아 평균값으로 채워라

예) 저장시 컬럼 리네임 및 열과 축을 변경하여 제출

**중복 데이터 확인**

print(df.duplicated())

print("\n")

중복 데이터 제거

drop\_duplicates()

데이터값 변경

Replace()

스케일링

https://m.blog.naver.com/PostView.nhn?blogId=demian7607&logNo=222009975984&proxyReferer=https:%2F%2Fwww.google.com%2F