

5장. 단일변수 자료의 분석

학번	32233421	이름	이은지
----	----------	----	-----

제공된 user_behavior_dataset.csv 는 스마트폰 모델별 사용자 이용 패턴을 정리한 데이터셋이다. 다음과 같이 스마트폰 모델 정보는 models 에, 일별 앱 사용시간 (minute/day) 정보는 usage_time 에 저장 한 뒤 문제를 해결 하시오

```
import pandas as pd
df = pd.read_csv('d:/data/user_behavior_dataset.csv')

models = df['Device_Model']           # 스마트폰 모델 정보 (범주형 자료)
usage_time = df['App_Usage_Time']      # 앱사용 시간 (연속형 자료)
```

	A	B	C	D	E
1	User ID	Device_Model	Operating_System	App_Usage_Time	Screen On Time (
2	1	Google Pixel 5	Android		393
3	2	OnePlus 9	Android		268
4	3	Xiaomi Mi 11	Android		154
5	4	Google Pixel 5	Android		239
6	5	iPhone 12	iOS		187

(각 문제별로 코드 & 실행결과화면을 캡처를 제출. 실행 결과가 길 경우에는 결과의 앞쪽 일부만 자라서 캡쳐해도 됨)

※ 단일변수 범주형 데이터 탐색. (models)

데이터사이언스

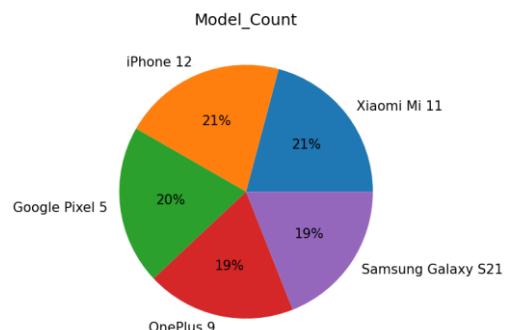
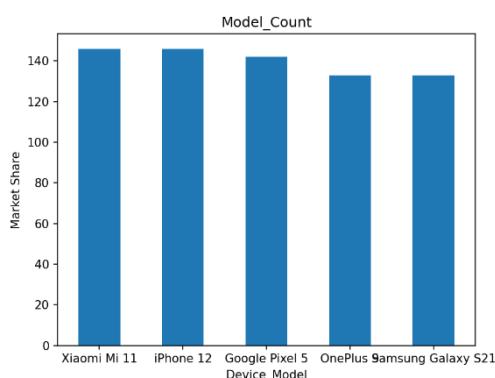
1. 모델별 도수분포표를 계산하여 md에 저장후 md의 내용을 보이시오.

```
>>> md= models.value_counts()  
>>> md  
Device_Model  
Xiaomi Mi 11      146  
iPhone 12         146  
Google Pixel 5    142  
OnePlus 9          133  
Samsung Galaxy S21 133  
Name: count, dtype: int64
```

2. md 부터 모델별 점유율을 계산하여 보이시오

```
>>> md= models.value_counts()/models.size  
>>> md  
Device_Model  
Xiaomi Mi 11      0.208571  
iPhone 12         0.208571  
Google Pixel 5    0.202857  
OnePlus 9          0.190000  
Samsung Galaxy S21 0.190000  
Name: count, dtype: float64
```

3. md를 이용하여 막대그래프와 원 그래프를 작성하여 보이시오



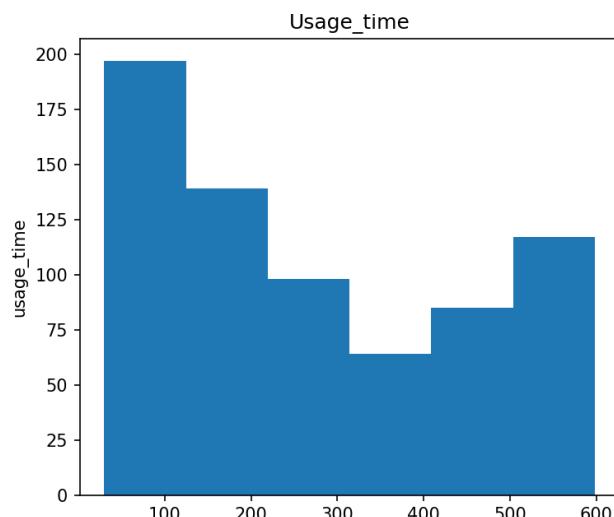
※ 단일변수 연속형 데이터 탐색. (usage_time)

데이터사이언스

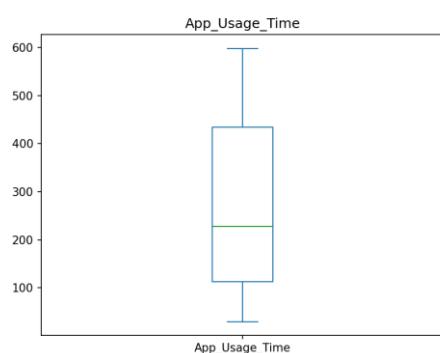
4. usage_time 에 대해 평균, 중앙값, 표준 편차, 사분위수를 보이시오

```
>>> usage_time.mean()
np.float64(271.12857142857143)
>>> usage_time.median()
np.float64(227.5)
>>> usage_time.describe()
count    700.000000
mean     271.128571
std      177.199484
min      30.000000
25%     113.250000
50%     227.500000
75%     434.250000
max     598.000000
Name: App_Usage_Time, dtype: float64
```

5. usage_time 에 대해 히스토그램을 작성하시오 (막대수:6)

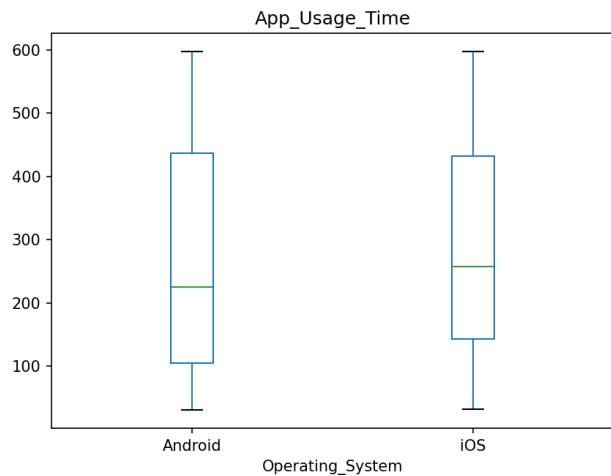


6. usage_time 에 대해 상자그림을 작성하시오



데이터사이언스

7. user_behavior_dataset.csv 의 Operating_System 컬럼에는 모델별 운영체제 정 보가 거장되어 있다. 이를 이용하여 usage_time 에 대해 상자그림을 작성하되 운영 체제별로 나누어 작성하시오 (코드5-10참조)



8. 5번과 6번의 그래프를 한화면에 출력하여 보이시오 (코드5-13 참조)

