

EDA

daily_sales(매출) 테이블 사용
쿼리 입력 -> 데이터 추출 ->
파이썬으로 시각화 진행

결측치 확인 및 기초통계량 확인

```
1 daily_sales.isnull().sum()
```

✓ 0.0s

```
date          0
WID           0
item_id       0
buy_count     0
revenue       0
dtype: int64
```

- 매출 테이블에는 결측치가 없음을 확인

	buy_count	revenue
count	97133.000	97133.000
mean	2.108	12750.505
std	4.327	56464.874
min	1.000	1200.000
25%	1.000	1200.000
50%	1.000	2400.000
75%	1.000	6000.000
max	180.000	2882400.000

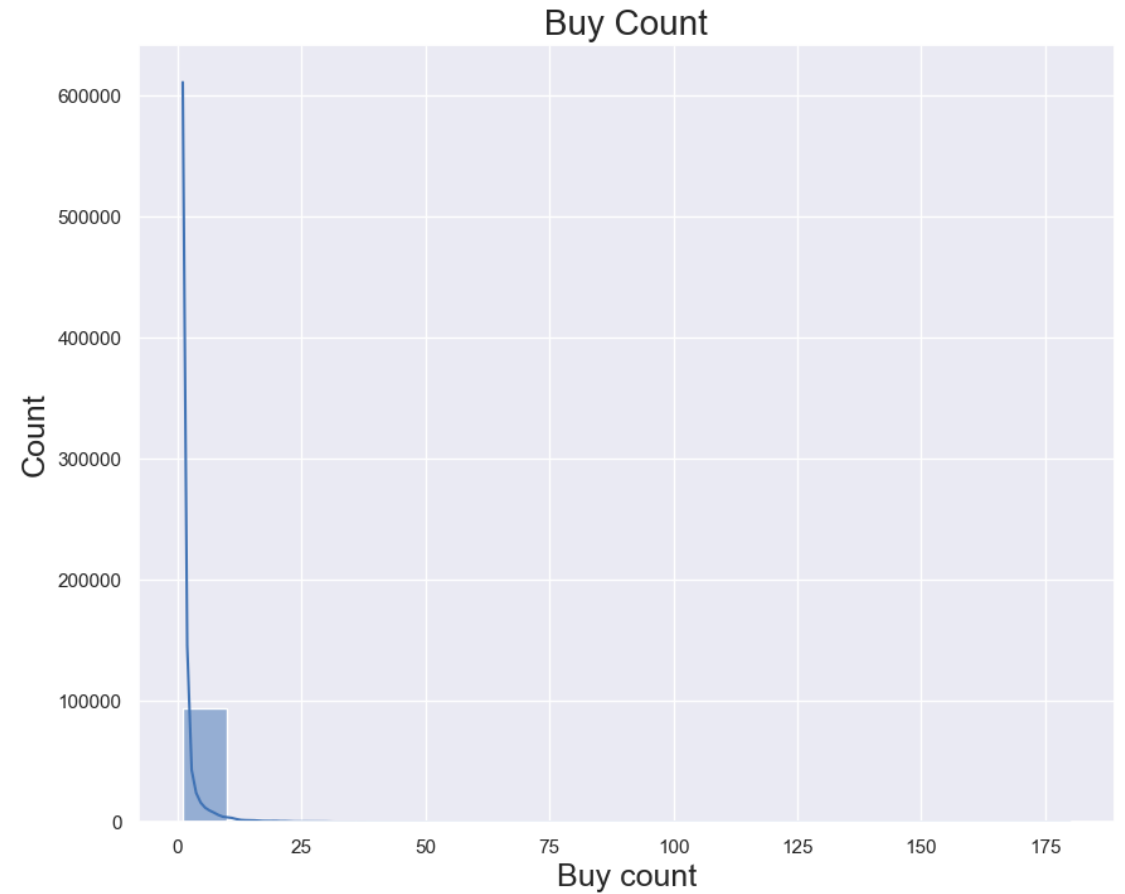
- 구매 횟수와 수익의 기초통계량 확인

25%, 50%, 75% 의 값이 모두 1이므로 대부분의 값이 1에

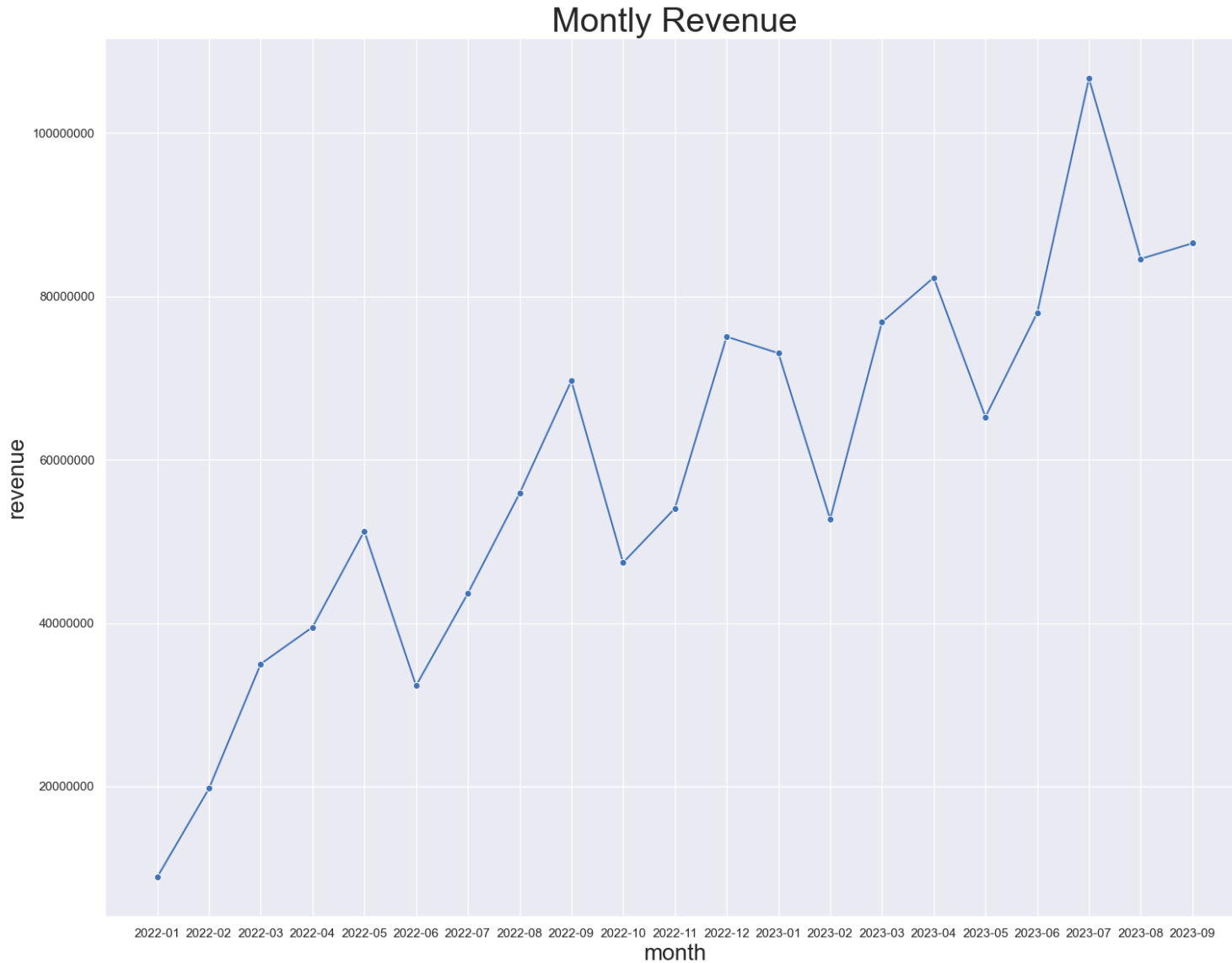
몰려있음을 확인 가능

구매횟수 시각화

- 게임 유저들은 대체로 구매횟수가 1회에 그친다는 것을 확인 할 수 있다.



월별 매출 추이



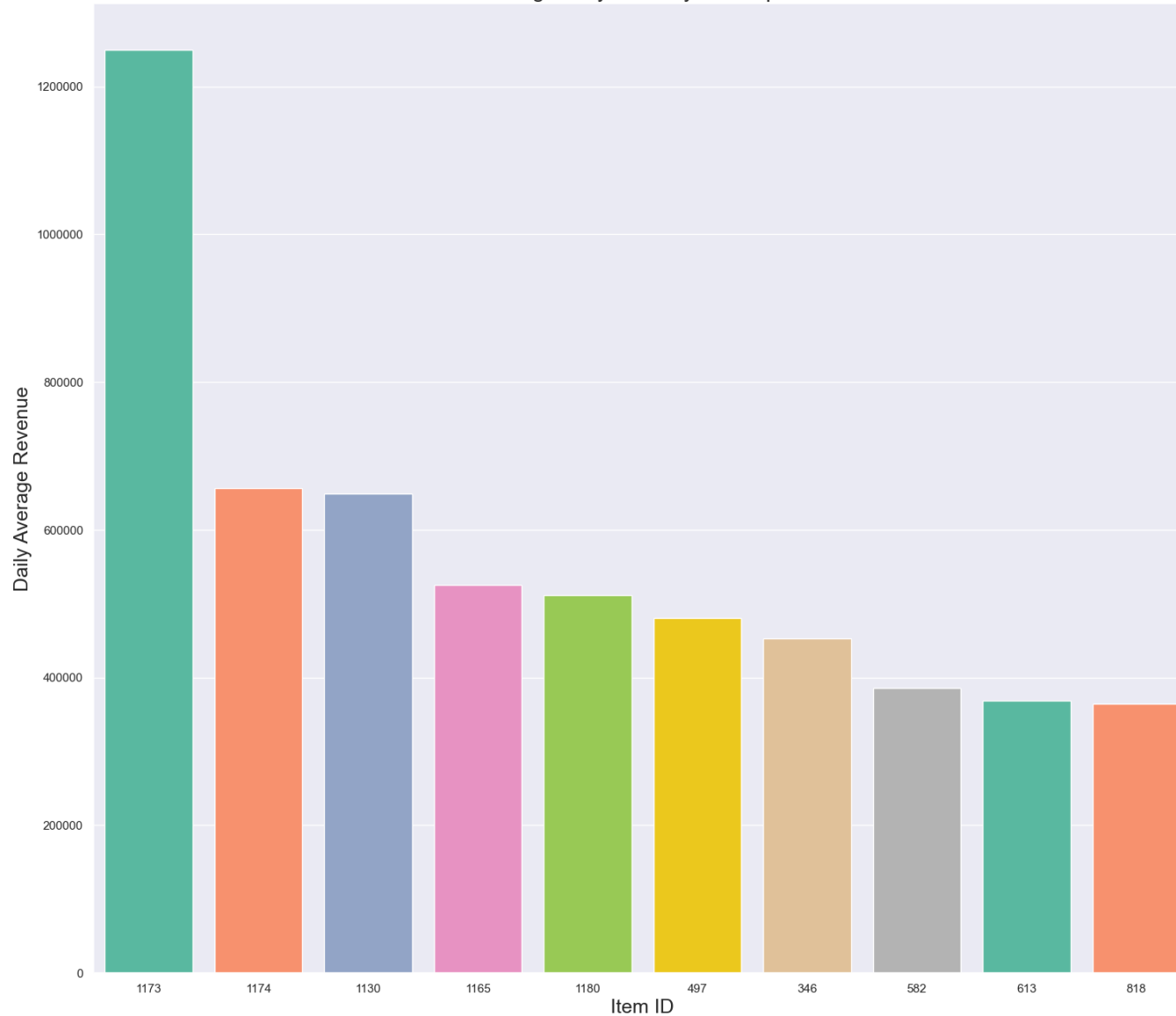
```
SELECT FORMAT_DATE('%Y-%m', date) month, SUM(revenue) month_revenue
FROM project.daily_sales
GROUP BY 1
```

매출은 꾸준히 증가하는 추세이고

이는 기업에게 긍정적인 상황이다.

아이템별 하루 평균 매출 top 10

Average daily sales by item top 10



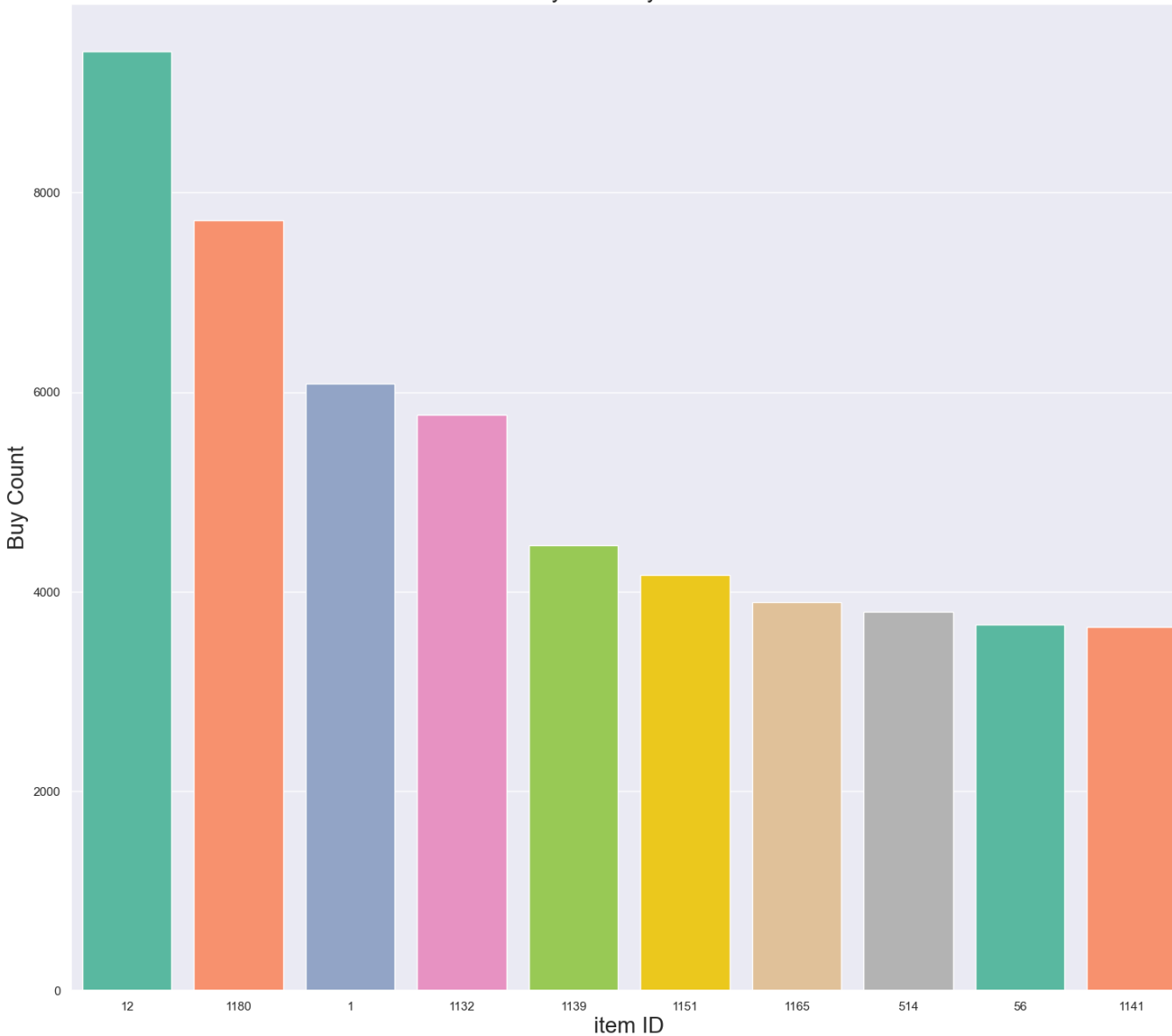
```
SELECT item_id, ROUND(avg(sum_revenue),2) daily_avg_revenue
FROM (SELECT date, item_id, sum(revenue) sum_revenue FROM project.daily_sales
GROUP BY 1,2)
GROUP BY 1
ORDER BY 2 DESC
LIMIT 10
```

Item ID

1173, 1174, 1130, 1165, 1180, 497, 346, 582, 613
,818이 하루 평균 높은 매출을 기록했다.

아이템별 구매 횟수 높은 순 top 10

Buy Count by item ID



```
SELECT DISTINCT(item_id), sum(buy_count) sum_buy_count FROM project.daily_sales  
GROUP BY 1  
ORDER BY 2 DESC  
LIMIT 10
```

Item ID

12, 1180, 1, 1132, 1139, 1151, 1165, 514, 56, 1141

아이템이 높은 구매 횟수를 기록 했다.

1165, 1180가 공통으로 나타남

Item ID 1165와 1180은 높은 매출과 높은 구매 횟수를 기록하는 것으로 나타납니다.

어떠한 요인들이 두 제품에게 영향을 미쳤는지 파악하기 위해

두 제품의 특성을 확인, 두 제품이 어떤 카테고리에 속해 있는지 확인,

특정 광고 및 할인 행사가 두 제품에게 긍정적으로 작동 했는지 확인하고

이를 기반한 기업의 의사결정이 가능합니다.