



Google Play EDA

DSL EDA 프로젝트 2조

박지호 안세정 유채원 이서현 한예림

목차

0 | 목차

1 | 데이터 소개

- 1.1 App data 소개
- 1.2 Review data 소개

2 | 데이터 전처리

- 2.1 App data 전처리
- 2.2 Review data 전처리

3 | 기초 분석

- 3.1 데이터 살펴보기
- 3.2 Sentiment별 Word Cloud

4 | 카테고리별 분석

- 4.1 카테고리별 키워드
- 4.2 인기 카테고리 인사이트
제시

5 | 카테고리 내 분석

- 5.1 'GAME' 카테고리

6 | 결론

- 6.1 추천 카테고리
- 6.2 카테고리별 전략 제시
- 6.3 한계점

7 | 질의응답

데이터 변수 소개

googleplaystore.csv

- App : 앱 이름
- Category : 앱의 카테고리
- Rating : 평점
- Reviews : 리뷰 수
- Size : 크기
- Installs : 설치 수
- Type : 유료/무료 여부
- Price : 가격 (무료인 경우 0)
- Content Rating: 이용 나이
- Genres : 앱의 장르 (카테고리와 거의 동일)
- Last Updated : 최근 업데이트 일 (19년도 기준)
- Current Ver : 현재 버전
- Android Ver : 안드로이드 버전

googleplaystore_user_reviews.csv

- App : 앱 이름
- Translated_Review : 리뷰(eng)
- Sentiment : 감정 (Positive/Negative/Neutral(Preprocessed))
- Sentiment_Polarity : Sentiment polarity score
- Sentiment_Subjectivity : Sentiment subjectivity score
 - Polarity : $(p-n)/(p+n)$
 - Subjectivity : $(p+n)/N$
 - p : 긍정 단어의 수
 - n : 부정 단어의 수
 - N : 단어의 수



App Data

| | App | Category | Rating | Reviews | Size | Installs | Type | Price | Content Rating | Genres | Last Updated | Current Ver | Android Ver |
|---|---|----------------|--------|---------|------|-------------|------|-------|----------------|---------------------------|------------------|--------------------|--------------|
| 0 | Photo Editor & Candy Camera & Grid & ScrapBook | ART_AND_DESIGN | 4.1 | 159 | 19M | 10,000+ | Free | 0 | Everyone | Art & Design | January 7, 2018 | 1.0.0 | 4.0.3 and up |
| 1 | Coloring book moana | ART_AND_DESIGN | 3.9 | 967 | 14M | 500,000+ | Free | 0 | Everyone | Art & Design;Pretend Play | January 15, 2018 | 2.0.0 | 4.0.3 and up |
| 2 | U Launcher Lite – FREE Live Cool Themes, Hide ... | ART_AND_DESIGN | 4.7 | 87510 | 8.7M | 5,000,000+ | Free | 0 | Everyone | Art & Design | August 1, 2018 | 1.2.4 | 4.0.3 and up |
| 3 | Sketch - Draw & Paint | ART_AND_DESIGN | 4.5 | 215644 | 25M | 50,000,000+ | Free | 0 | Teen | Art & Design | June 8, 2018 | Varies with device | 4.2 and up |
| 4 | Pixel Draw - Number Art Coloring Book | ART_AND_DESIGN | 4.3 | 967 | 2.8M | 100,000+ | Free | 0 | Everyone | Art & Design;Creativity | June 20, 2018 | 1.1 | 4.4 and up |

Review Data

| | App | Translated_Review | Sentiment | Sentiment_Polarity | Sentiment_Subjectivity |
|-------|---------------------------|---|-----------|--------------------|------------------------|
| 13442 | Beautiful Widgets Pro | S7 I love bw used years. SUDDENLY keeps ghosti... | Positive | 0.033333 | 0.600000 |
| 6064 | Allrecipes Dinner Spinner | Allrecipes great place find recipes. However, ... | Positive | 0.612500 | 0.700000 |
| 16945 | Bowmasters | I got I really like it! Too much blood too! An... | Negative | -0.020833 | 0.466667 |
| 51320 | GO Notifier | Lollipop update fix!!!! After lollipop update ... | Neutral | 0.000000 | 0.550000 |
| 55011 | Google Analytics | | NaN | NaN | NaN |

데이터 전처리

1. App Data: type 변환 및 결측치 보간

- Size 열
 - MB(메가바이트)로 단위 통일
 - string → float 변환
- Installs 열
 - string → int 변환
- Price 열
 - string → float 변환

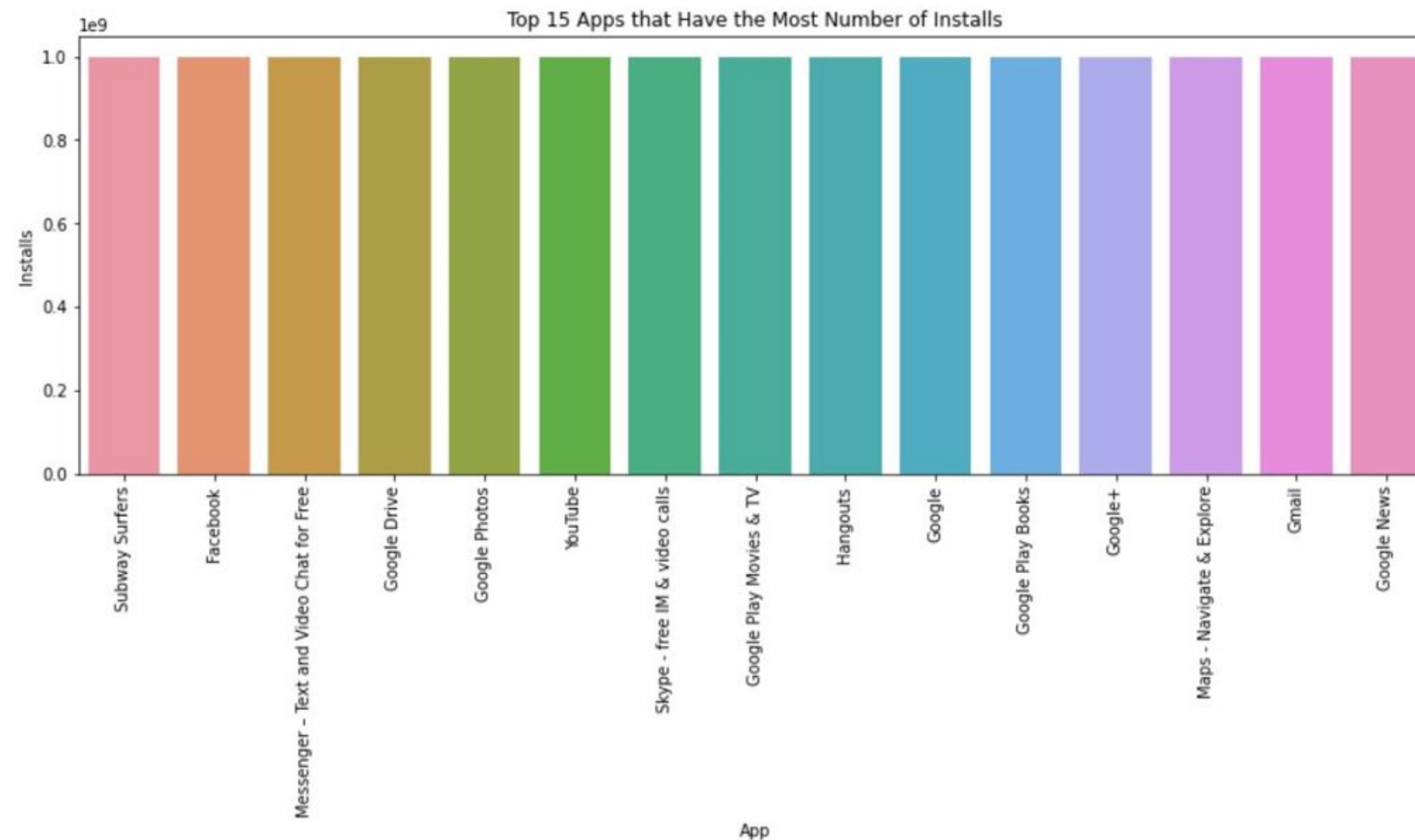
데이터 전처리

2. Review Data: NaN값 제거



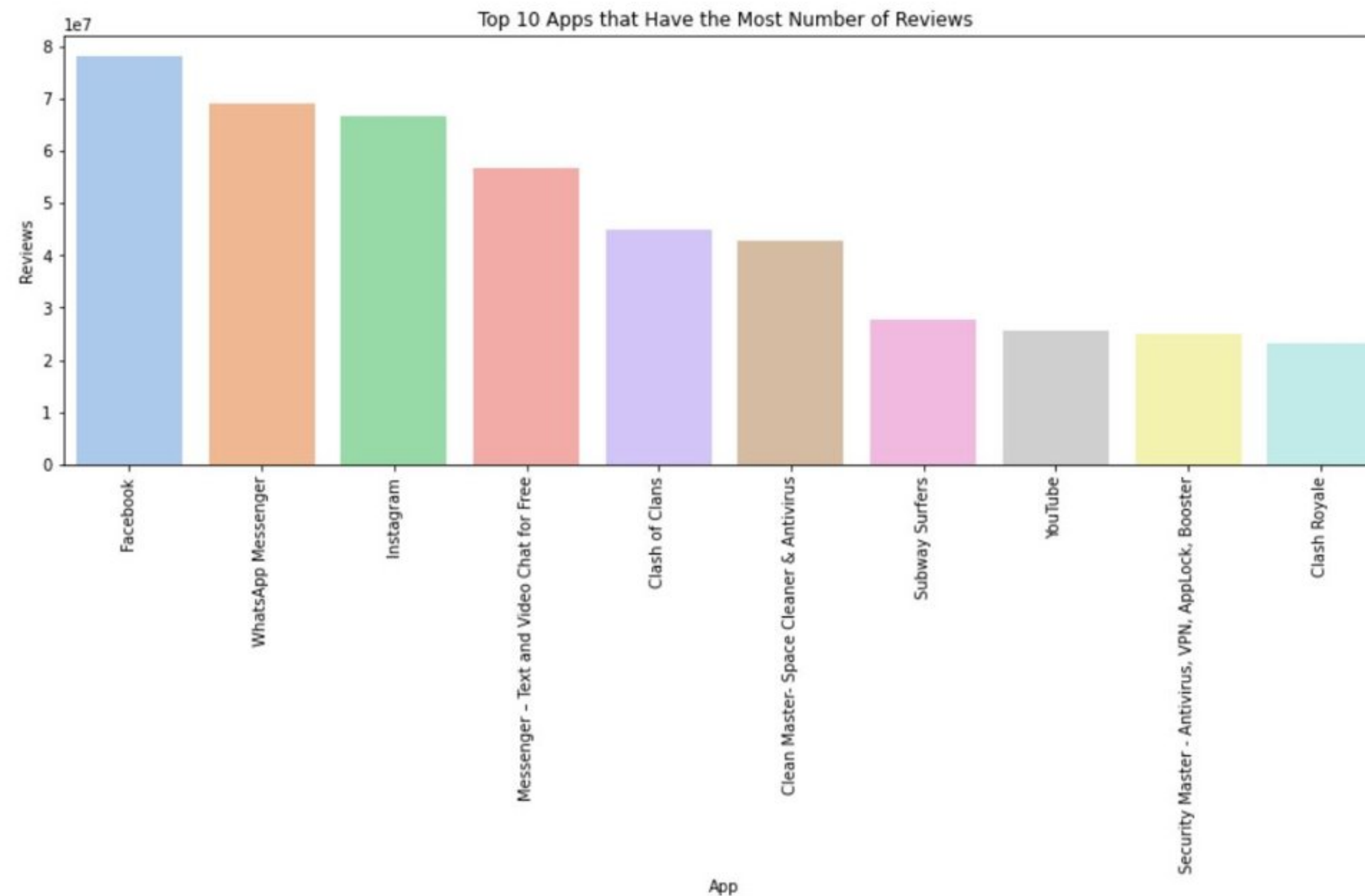
데이터 살펴보기 (*App Data*)

1. 다운로드 수 최다 앱



데이터 살펴보기 (*App Data*)

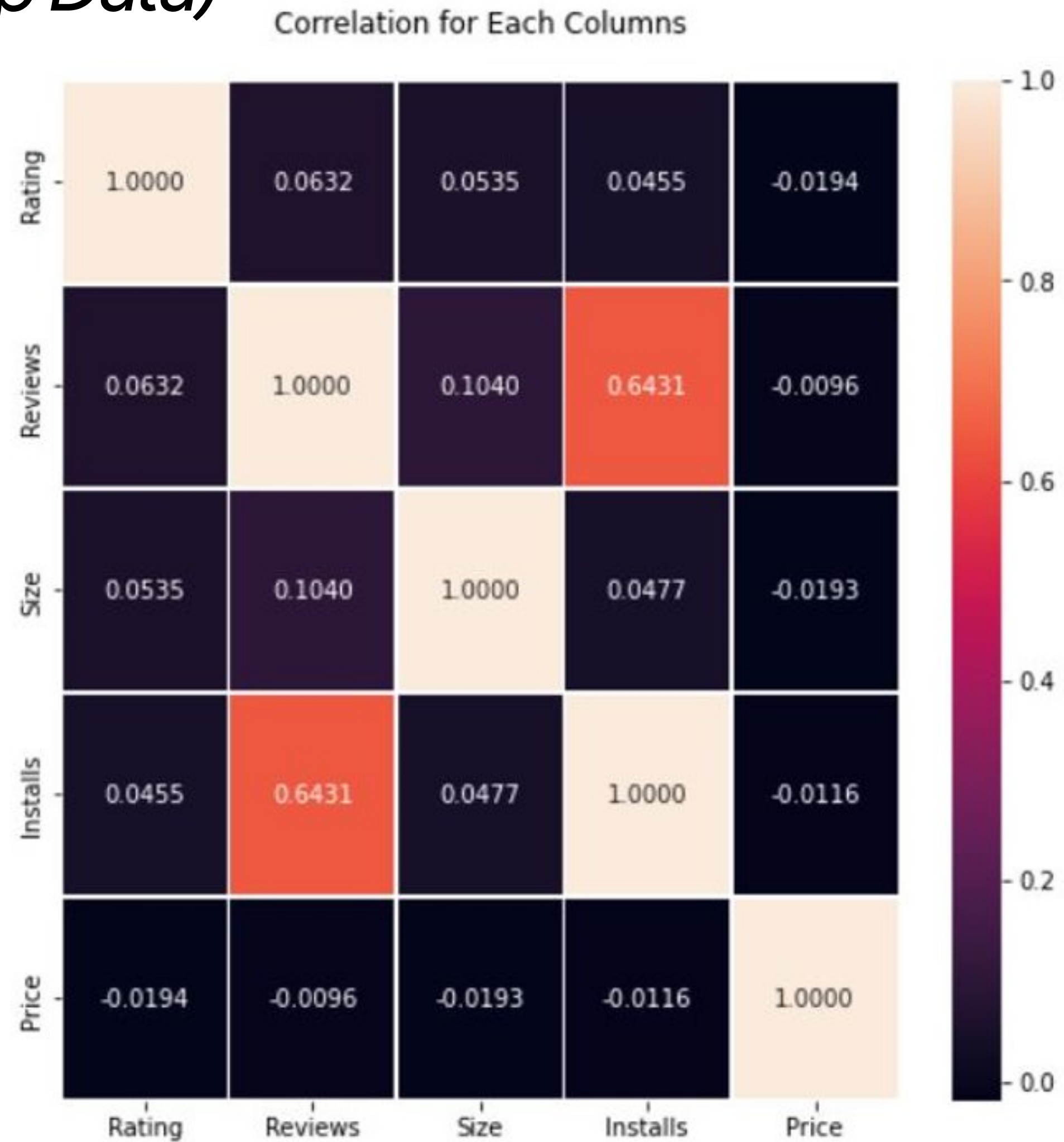
2. 리뷰 수 최다 앱





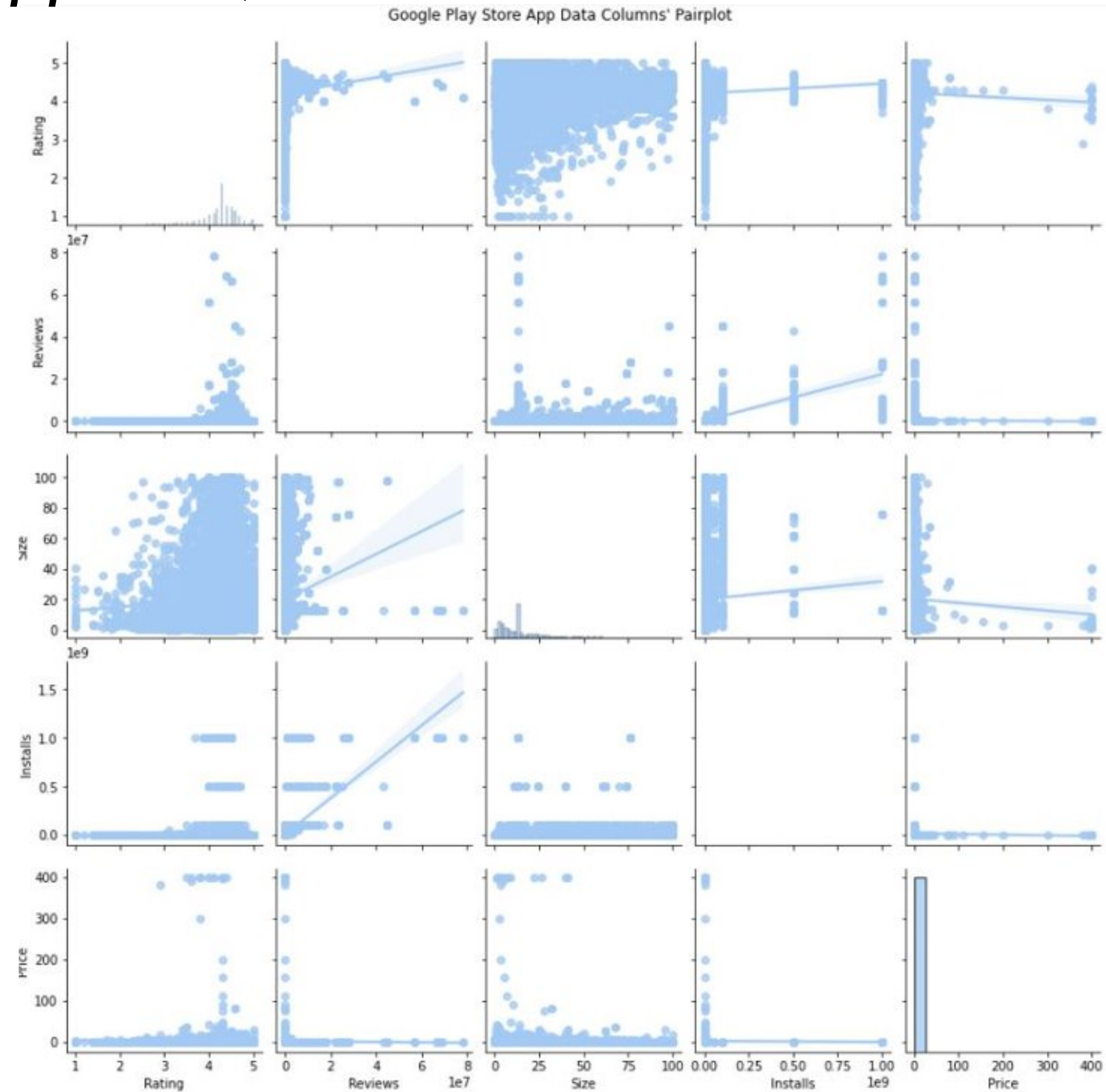
데이터 살펴보기 (*App Data*)

3. 상관관계 분석



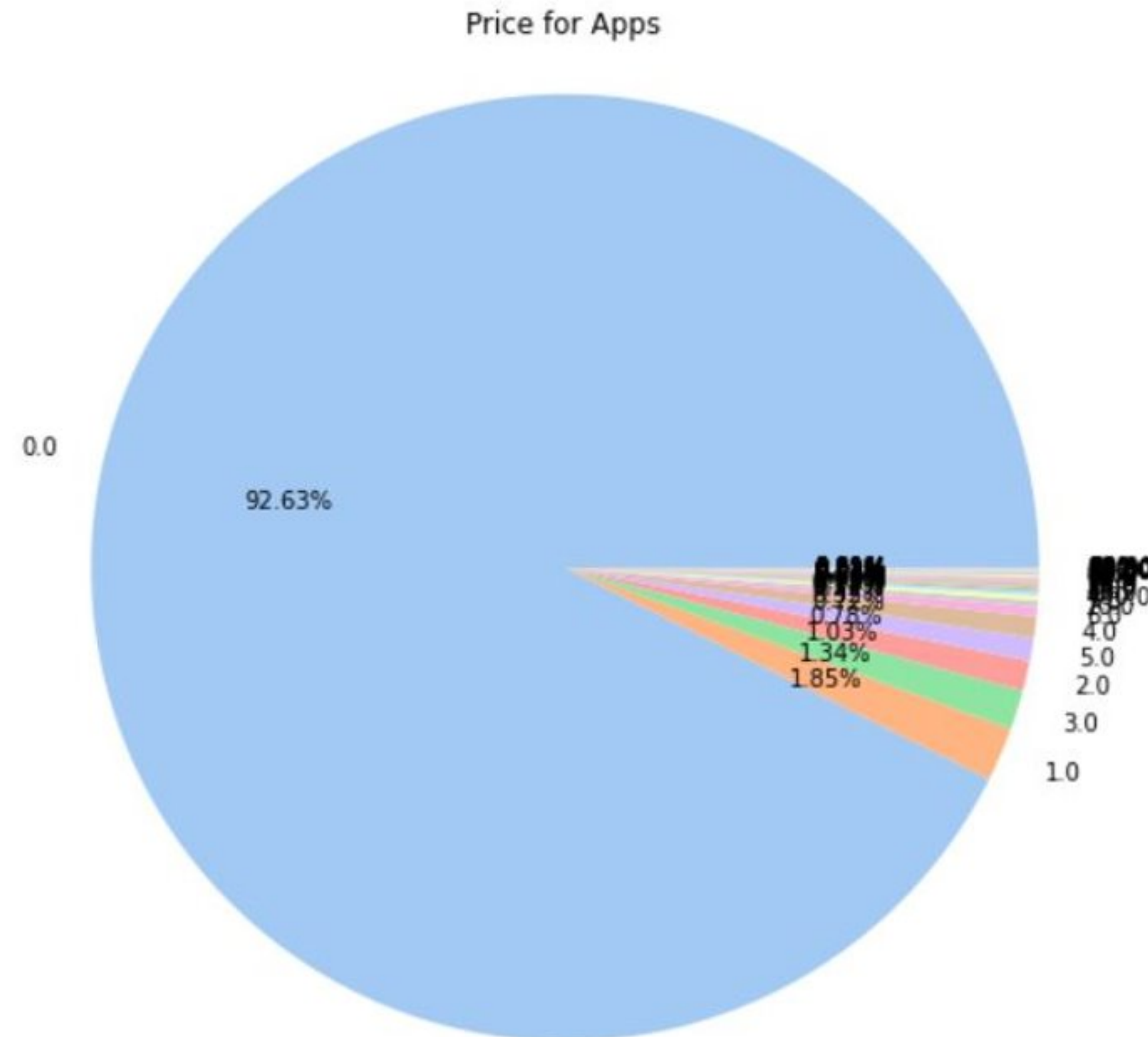
데이터 살펴보기 (*App Data*)

3. 상관관계 분석



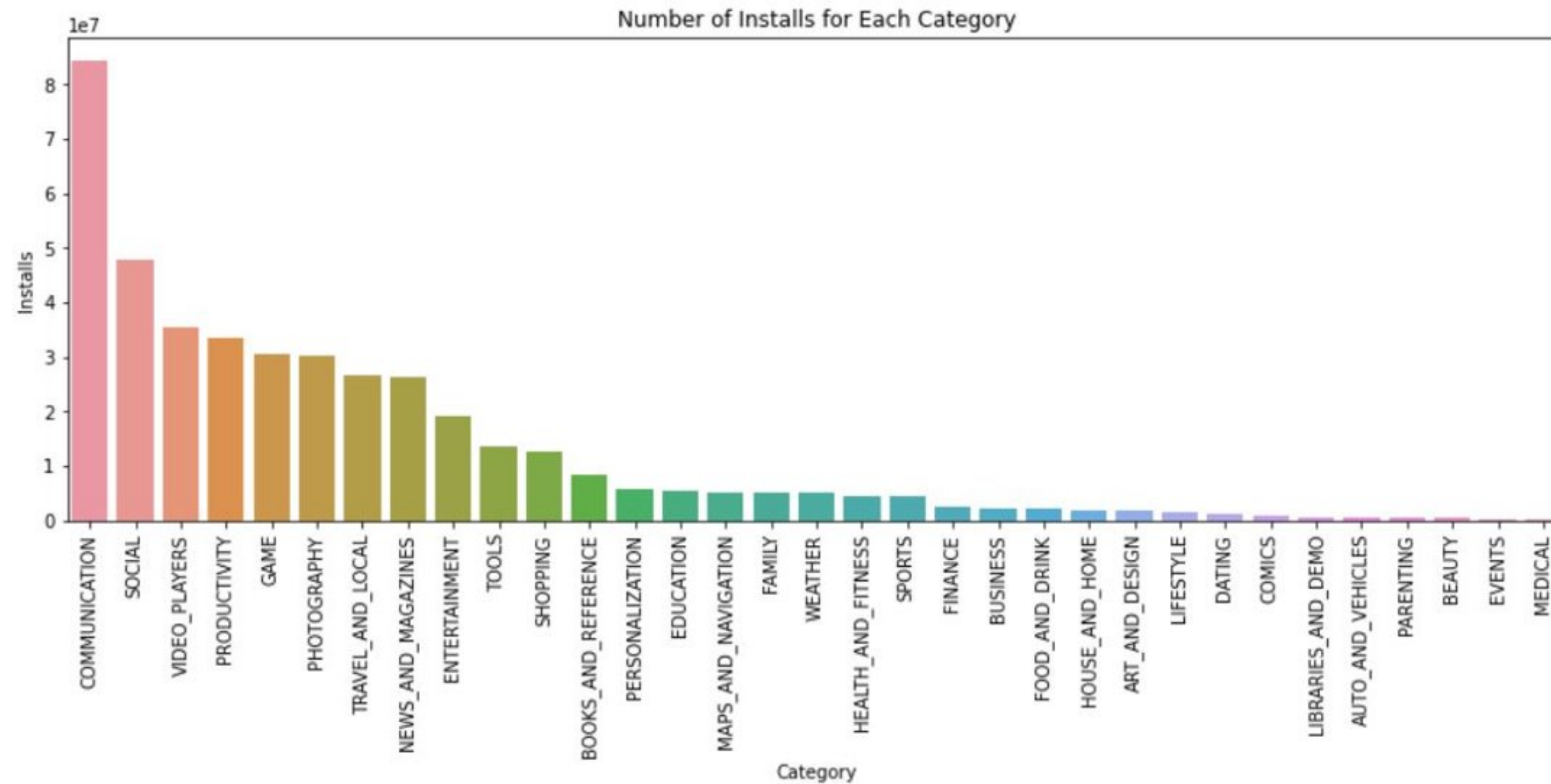
데이터 살펴보기 (*App Data*)

4. 다운로드 수 상위 랭킹 앱 가격



데이터 살펴보기 (*App Data*)

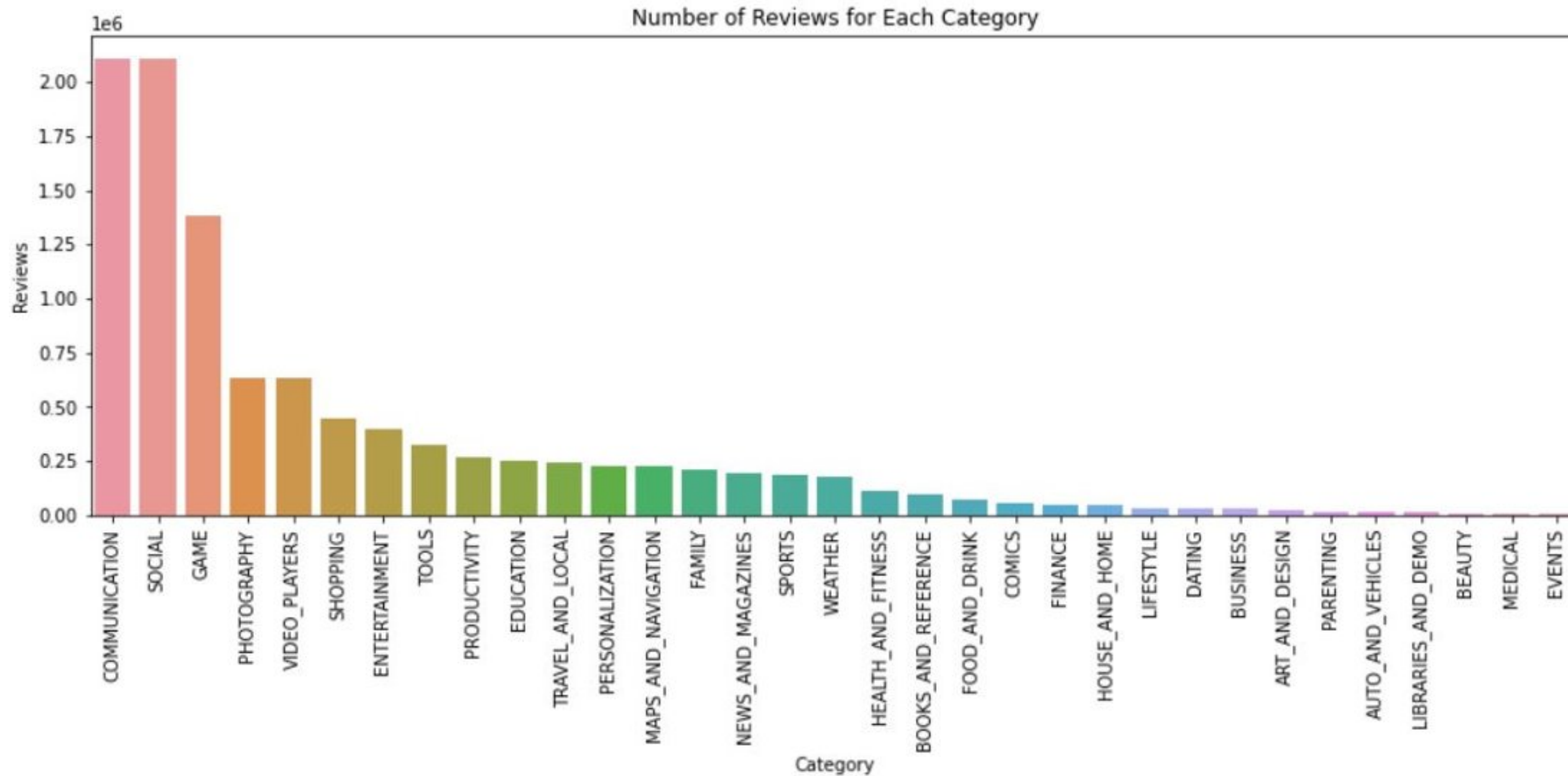
5. 카테고리별 다운로드 수 순위

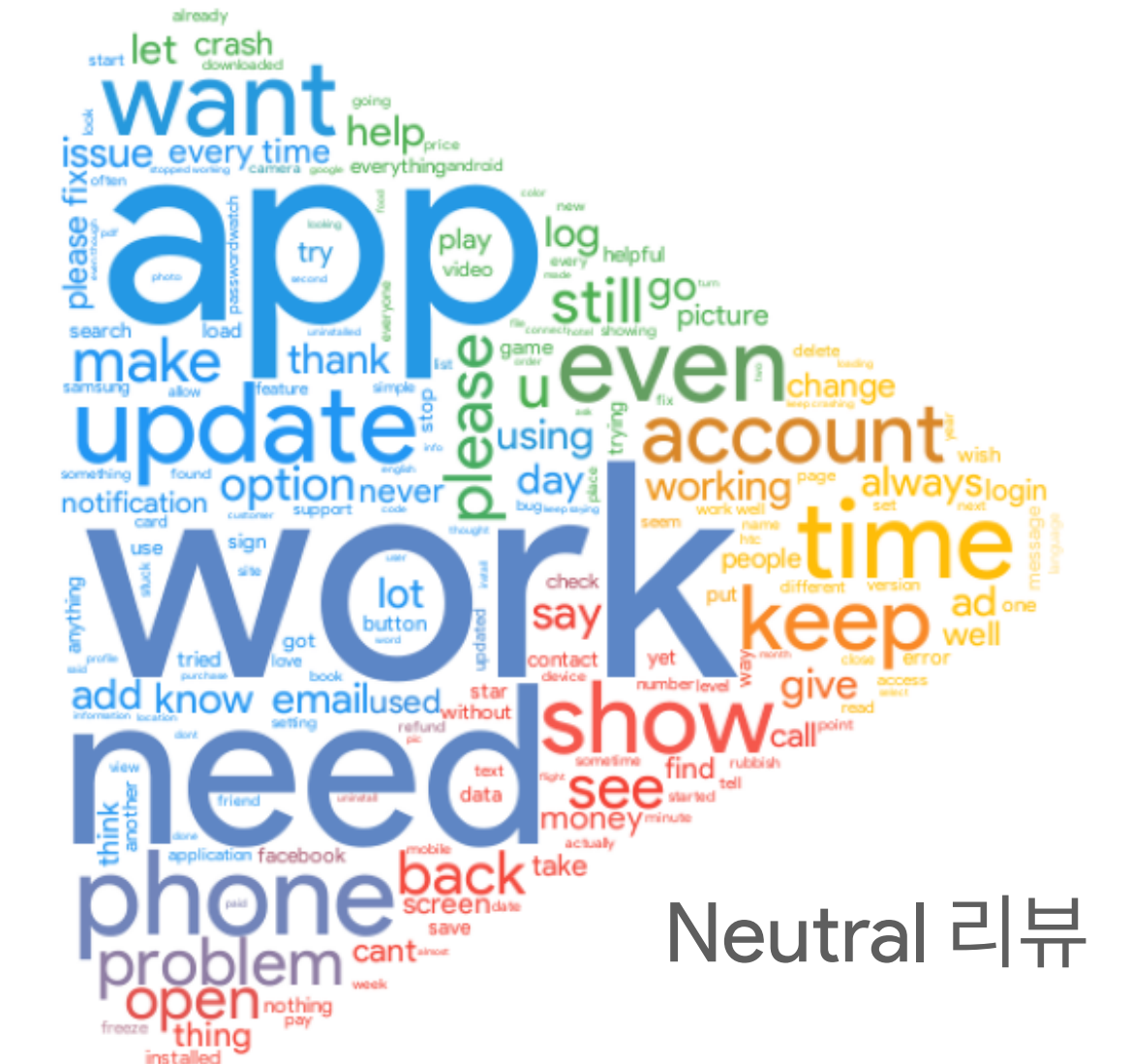




데이터 살펴보기 (*App Data*)

6. 카테고리별 리뷰 수 순위





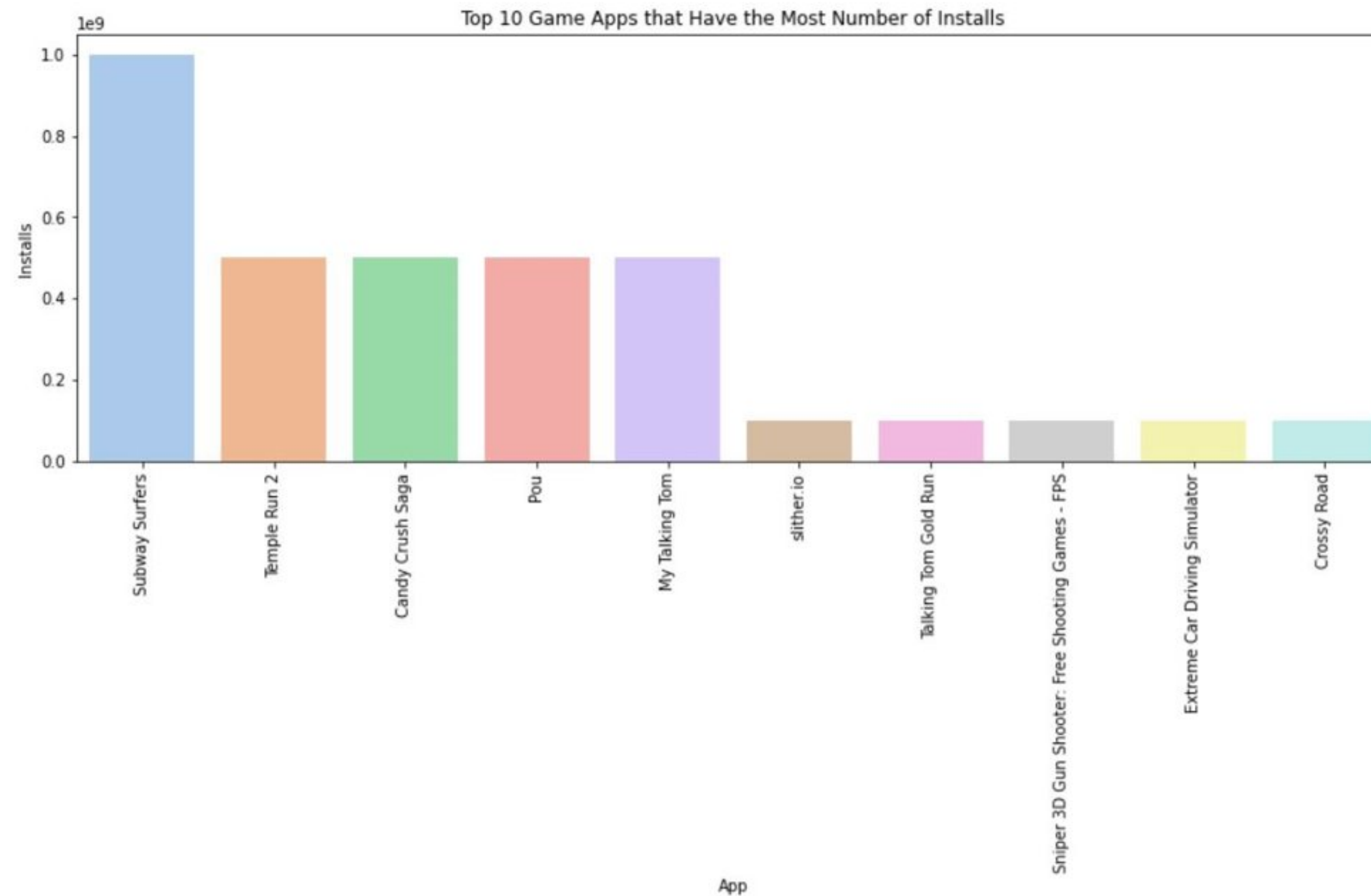
Analysis by Category

살펴볼 카테고리

- ‘ART AND DESIGN’
- ‘DATING’
- ‘COMMUNICATION’
- ‘GAME’
- ‘FAMILY’
- ‘SOCIAL’
- ‘WEATHER’
- ‘MAPS AND NAVIGATION’

카테고리 내 분석

: 'GAME' 카테고리



카테고리 내 분석

: 'GAME' 카테고리

1. 안드로이드 어플 크기의 평균은 12MB

게임 어플 크기의 평균은 40MB
(상위권 게임 어플은 대부분 60MB 이상)

→ 사이즈가 크다 = 콘텐츠가 많다

⇒ 자체 콘텐츠 중요

2. 그럼에도 자체 콘텐츠가 너무 무거우면 안됨

Installs 상위 5개 게임 어플 모두 'Hyper Casual' 종류의 게임

→ 다운 받자마자 별도의 튜토리얼 없이 쉽게 즐길 수 있는 게임 선호

결론

1. 추천 카테고리

- 'ENTERTAINMENT'
- 'PHOTOGRAPHY'
- 'SHOPPING'
- 'GAME'
- 'EDUCATION'

질의응답

