



게임을 출시한다면?

팀 명	6조
팀 원	김홍일, 윤택한, 이남경, 권희은
프로젝트 기간	2022.07.07 - 2022.07.10

데이터셋 기본정보

1) 데이터 파싱 날짜

- ① 21년 6월 16일~17일 파싱 (2010년 ~ 2021년 6월)
- ② 22년 7월 9일 파싱 (2010년 ~ 2022년 7월)

2) 24개의 columns, 2,312,944개의 Row

(21년 6월 16일~17일 파싱)

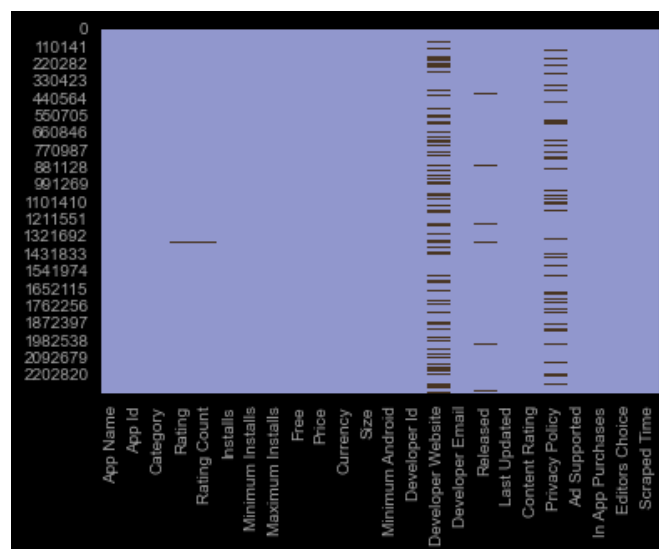
Column	Dtype	Column	Dtype
App Name	object	Minimum Android	object
App Id	object	Developer Id	object
Category	object	Developer Website	object

Column	Dtype	Column	Dtype
Rating	float64	Developer Email	object
Rating Count	float64	Released	object
Installs	object	Last Updated	object
Minimum Installs	float64	Content Rating	object
Maximum Installs	float64	Privacy Policy	object
Free	bool	Ad Supported	bool
Price	float64	In App Purchases	bool
Currency	object	Editors Choice	bool
Size	object	Scraped Time	object

3) 요약통계량

	Rating	Rating Count	Minimum Installs	Maximum Installs	Price
count	2,290,061	2,290,061	2,312,837	2,312,944	2,312,944
mean	2	2,865	183,445	320,202	0
std	2	212,163	15,131,440	23,554,950	3
min	0	0	0	0	0
25%	0	0	50	84	0
50%	3	6	500	695	0
75%	4	42	5,000	7,354	0
max	5	138,557,600	10,000,000,000	12,057,630,000	400

4) 결측치



개발자 웹 사이트 주소와 보안정책에 결측치가 있음을 알 수 있다.

분석 과정에서 개발자 개인정보 columns는 제거, 보안정책이 결측치라면 로그인 필요없는 앱으로 column추가

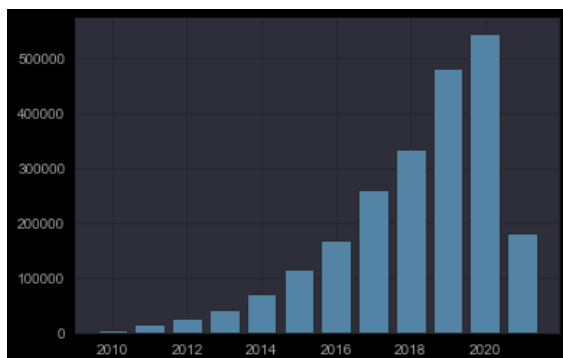
4) 전처리

1. 필요없는 변수 제거(Installs, Developer Id, Developer Website, Developer Email, Scraped Time)
2. 앱 이름에 결측치 있는 데이터 제거
3. Game변수 추가(대분류가 게임인 앱 True)
4. Currency가 NaN값이면 다른 변수에서도 결측치가 보여서 유의미한 데이터라고 볼 수 없다 → 데이터 제거
5. Released가 결측치이면 플레이스토어 확인 결과 삭제된 앱으로 분석에 필요하지 않다고 판단 → 데이터 제거
6. Released, Last updated 날짜 데이터로 변경
7. Released, Last updated의 연도만 가져온 Year변수 추가

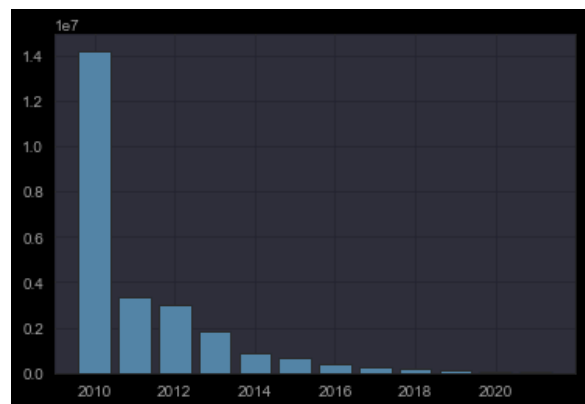
1. 게임업계 동향파악

1) 앱 출시량과 다운로드 횟수

앱 출시량



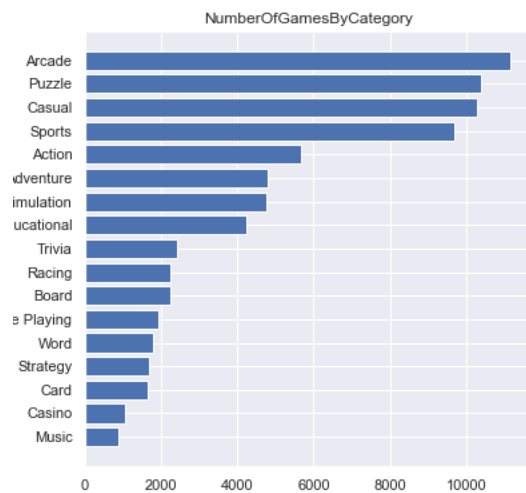
한 앱당 다운로드 횟수



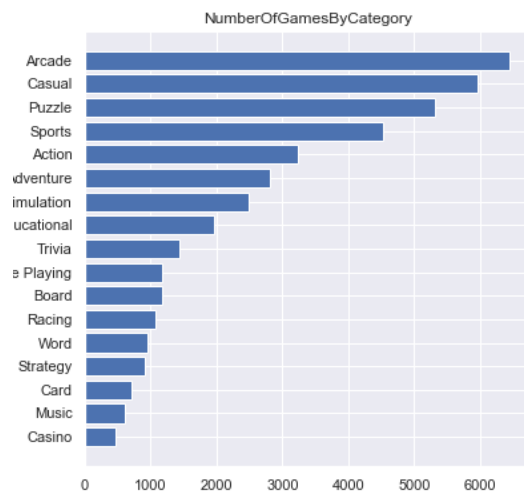
매해 앱 출시량이 증가하는 것에 비해 한 앱당 다운로드 받는 횟수는 감소하고 있다.

2) 카테고리별 출시량과 다운로드 횟수

- 출시량

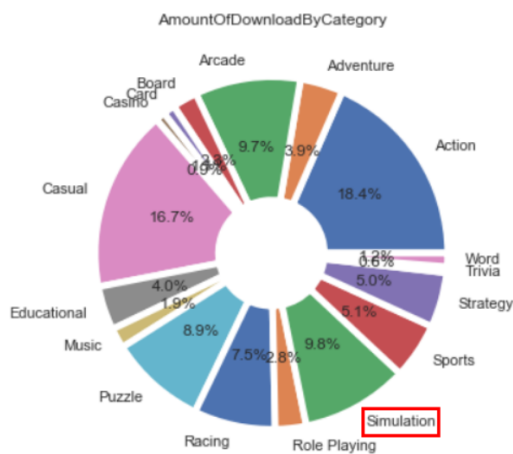


전체 기간

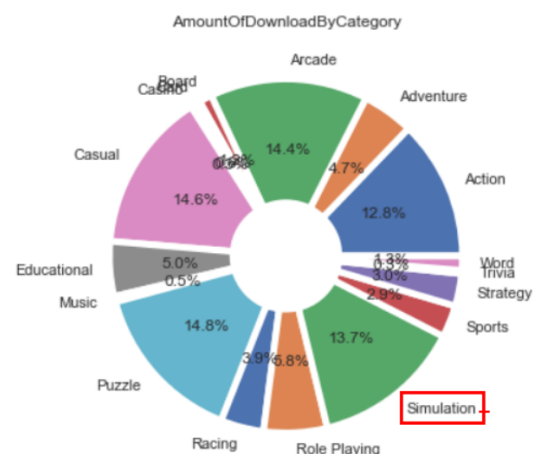


최근 3년

• 다운로드 횟수



전체 기간



최근 3년



전체 게임 및 최근 3년 게임 카테고리별 개수 분석 결과 모두 Arcade, Puzzle, Casual이 Top3 유지.

해당 게임들은 다른 게임에 비해 상대적으로 빠른 개발 속도와 결과를 볼 수 있다는 특징이 있다.



시뮬레이션 카테고리는 출시 수에 비해 높은 다운로드 수를 보여, 아직 시장이 포화되지 않은것에 비해 사람들의 관심이 높은 카테고리로 보인다.



2010년 출시한 게임중 인기 게임:
말하는 고양이 토크툼



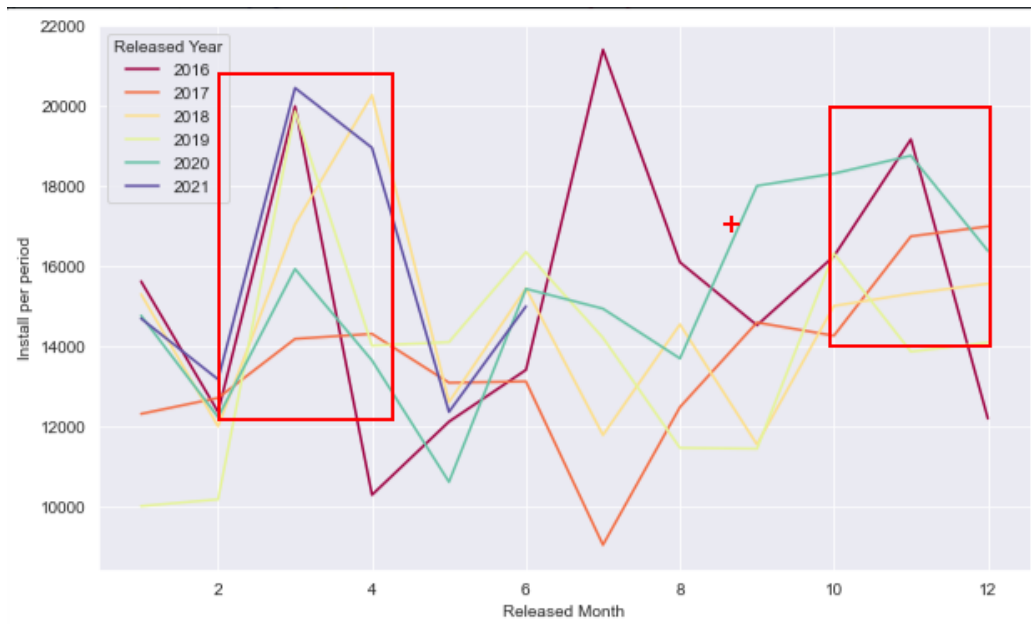
2021년 출시한 게임중 인기 게임:
High Heels

3) 게임 카테고리별 게임명 빈도 수

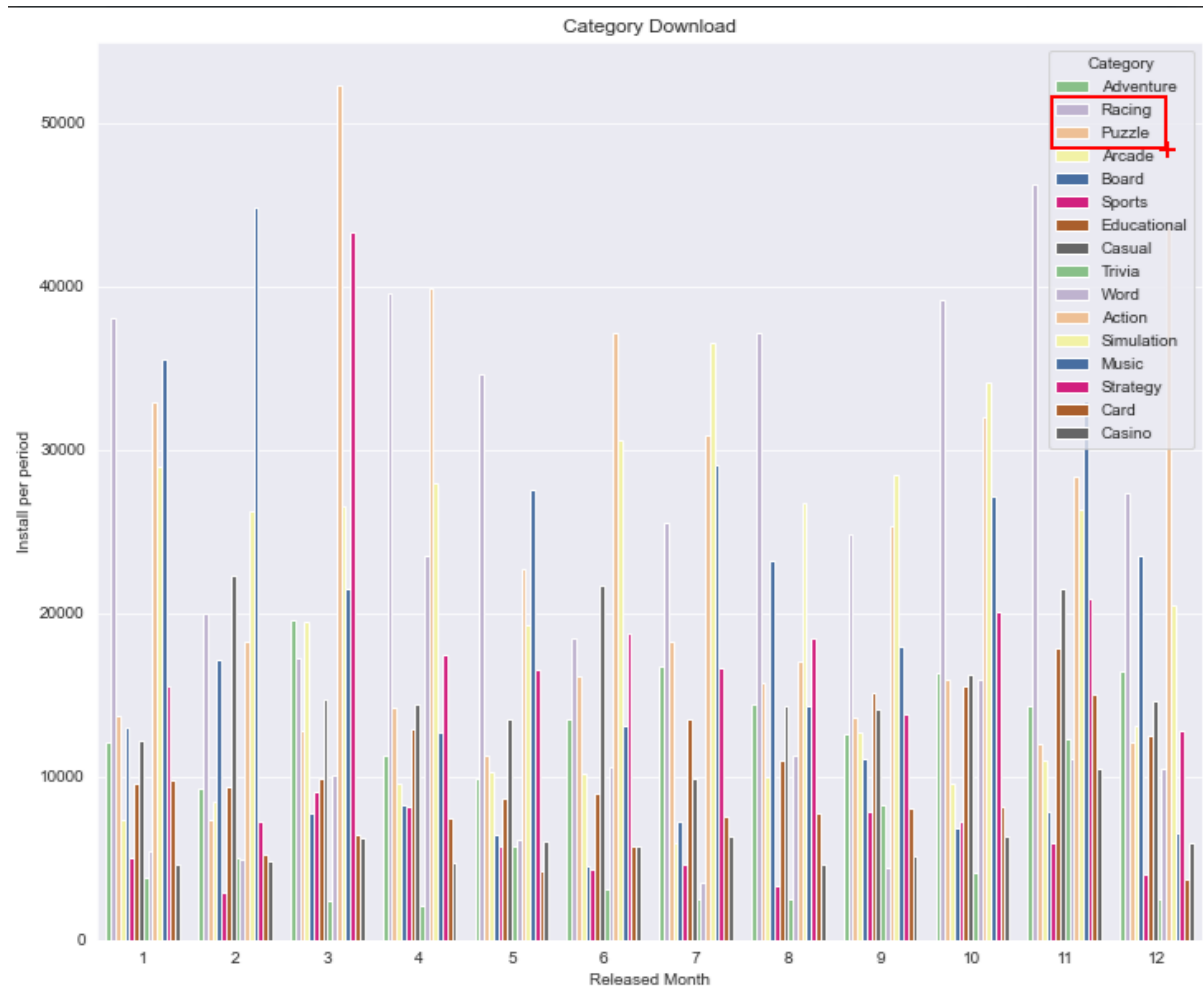


Top3 게임 카테고리인 Arcade, Puzzle, Casual의 경우
VS, Match 등 대결을 통해 승패를 결정하는 단어를 많이 사용하고 있다고 볼 수 있다.

4) 월별 다운로드



월별 다운로드 횟수



월별, 카테고리별 다운로드 횟수



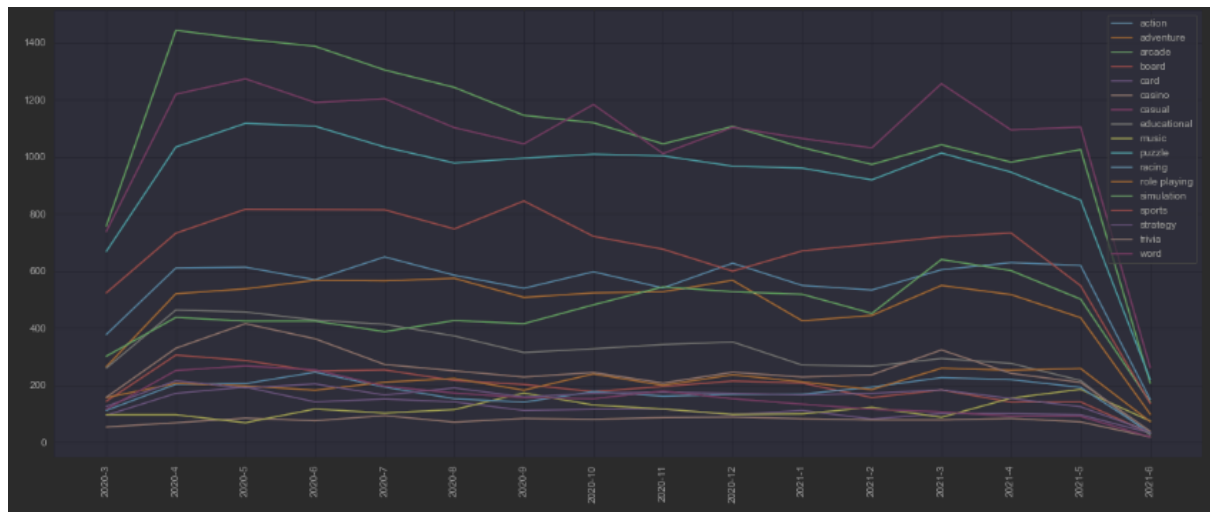
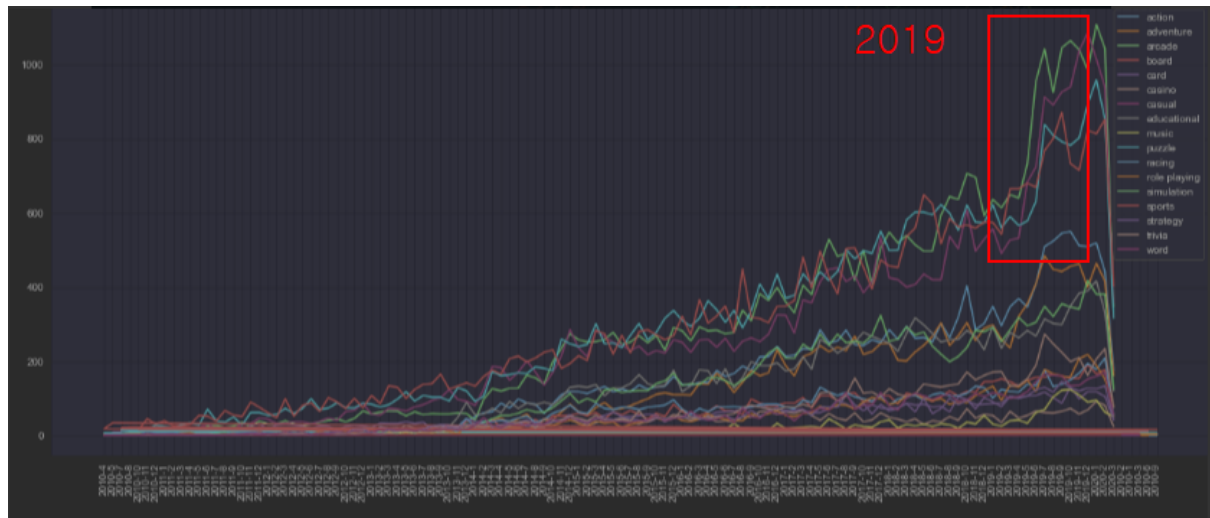
2-4월과 10-12월에 상대적으로 다운로드가 횡수가 많음을 알 수 있다.

다운로드 횡수가 많은 카테고리는 2-4월 퍼즐, 10-12월 레이싱으로 볼 수 있다.

2. 코로나 전 후

1) 게임 앱 출시량 비교

*2020년 3월 10일 팬데믹 선언을 기점으로 코로나 전 후로 구분



코로나 후 (2020년 3월~2021년 6월)



코로나 이전에 급격한 변화가 있었기때문에 상대적으로 코로나 전 후 게임앱 출시량의 급격한 변화는 없다고 보인다.

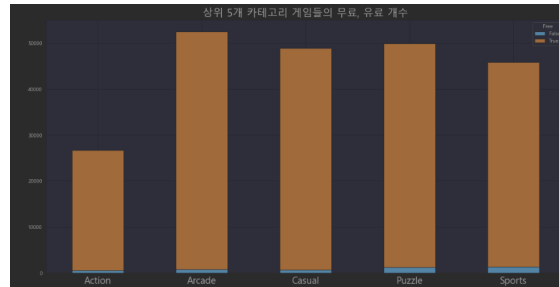
계속 성장하고 있음을 알 수 있다.

2) 상위 5개의 게임 카테고리 비교

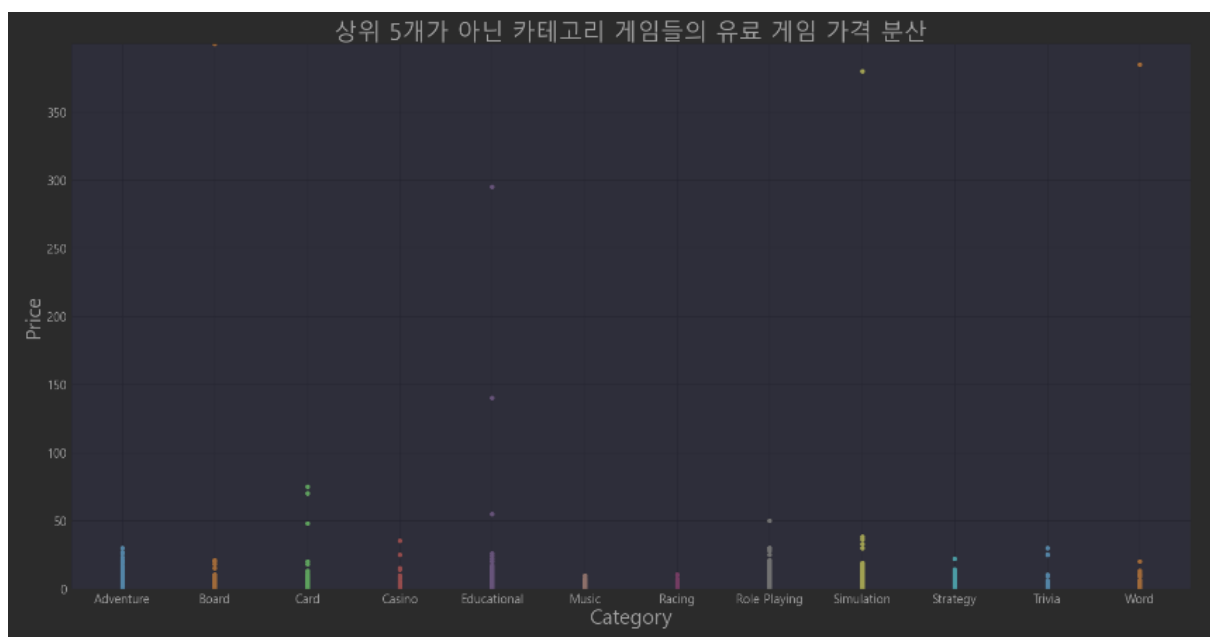
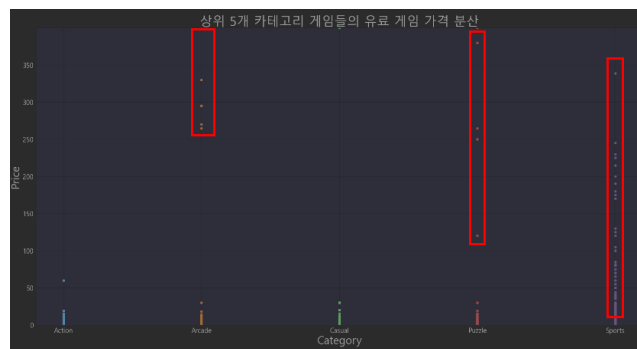
코로나 이후 상대적으로 출시량이 늘어난 게임 카테고리 상위 5개 선정

***Arcade, Casual, Puzzle, Sports, Action**

- 앱 가격

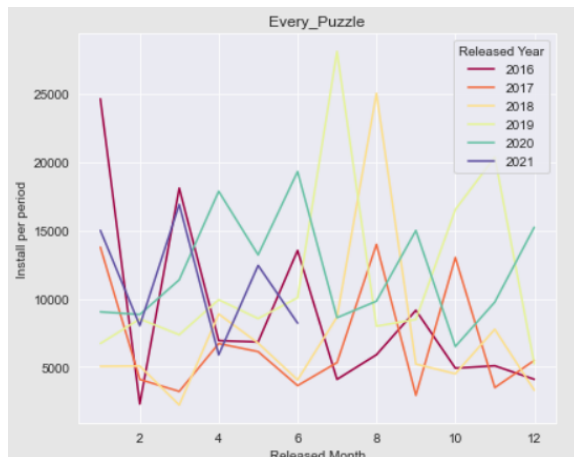
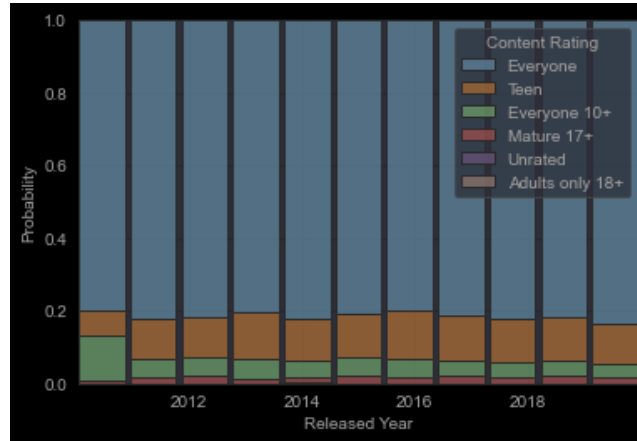


대부분의 앱이 설치시에 무료로 제공됨을 알 수 있다.

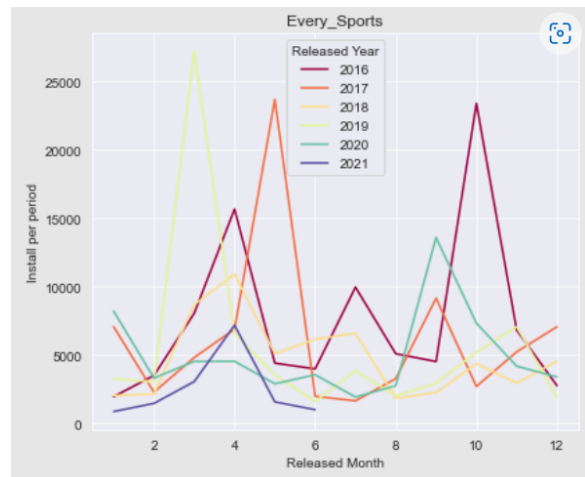




상위 카테고리의 경우 앱이, 그 외 카테고리 보다 가격이 비싼 앱의 비율이 높은것을 알 수 있다.



퍼즐 카테고리



스포츠 카테고리



모든 연령이 사용하도록 앱을 만들고 있음을 알 수 있다.

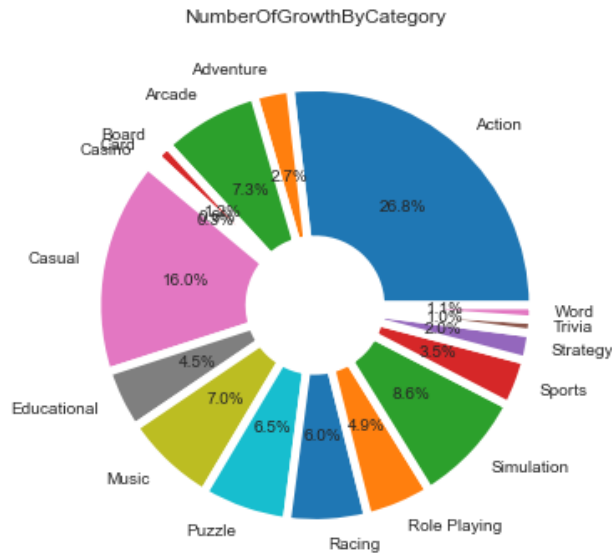
앱 사용 연령이 Every인 데이터 비교해봤을 때, 퍼즐은 주로 6-8월에 유행을 하고, sports 는 대략 2월과 10월 때 많이 한다고 볼 수 있다.

3. 지난 1년간 변화

21.06.13일 파싱 데이터에 있는 게임 앱 약 7만개를 2022.07.09(현재) 일자로 재 파싱
약 1년간의 변화 추적

- Installs Count Current: 현재 다운로드 량

다운로드 증가량 비교



2010~2021년과 비교했을때, 코로나 이후 다운로드와 출시 개수가 상위였던 **Action, Casual, Arcade**카테고리가 2022년 7월 기준 여전히 강세임을 알 수 있다.

결론



1. 매해 앱 출시량이 증가하는 것에 비해 한 앱당 다운로드 받는 횟수는 감소하고 있다.
2. 코로나로 인한 게임앱 출시량의 급격한 변화는 없다고 볼 수 있다.
3. 최근 3년 게임 카테고리 중 Action, Casual, Arcade가 높은 순위를 유지하고 있다.



게임을 출시한다면, 다음을 특징을 갖춰 출시하는것이 좋을것으로 보인다.

1. 무료 앱
2. 전체 사용자
3. 대결에 관한 단어 포함해서 이름 짓기
4. **Arcade, Casual, Puzzle, Sports, Action**의 카테고리라면 유료 앱 출시도 고려해볼만하다.
5. 인디게임이라면 2월~4월에 출시해 6월 구글 인디게임페스티벌을 목표로 하고, 메이저게임이라면 출시 개수가 많은 달을 피하는것도 좋은 방법일 것 같다.

제언

- 1) 다운로드 횟수가 누적된 데이터로 연도별로 앱 다운로드 횟수를 비교할 수 없었다.
- 2) 앱(게임) 시장의 추세가 아닌 사용자의 관심도를 보고자 한다면, 연도별 다운로드 횟수, 매출액 등의 변수들이 필요할 것으로 보인다.