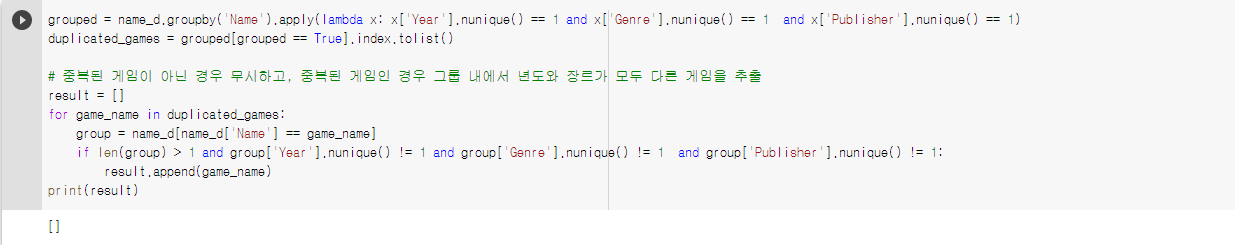
주제: 다음 분기 만들 게임을 설정

방법

1, 전처리

텍스트, 스크린샷, 다른, 여러개이(가) 표시된 사진

자동 생성된 설명 - 문제 : 결측치 / 해결방법 :

이름이 겹치는 것들을 모두 뽑음

그 중에서 게임 이름이 같을 경우, 년도와 장르가 다른 게임을 추출 – 없음

즉 이름이 같으면 년도와 장르가 같다

동일한 이름에 대해서 결측치가 채워지도록 넣음

텍스트이(가) 표시된 사진

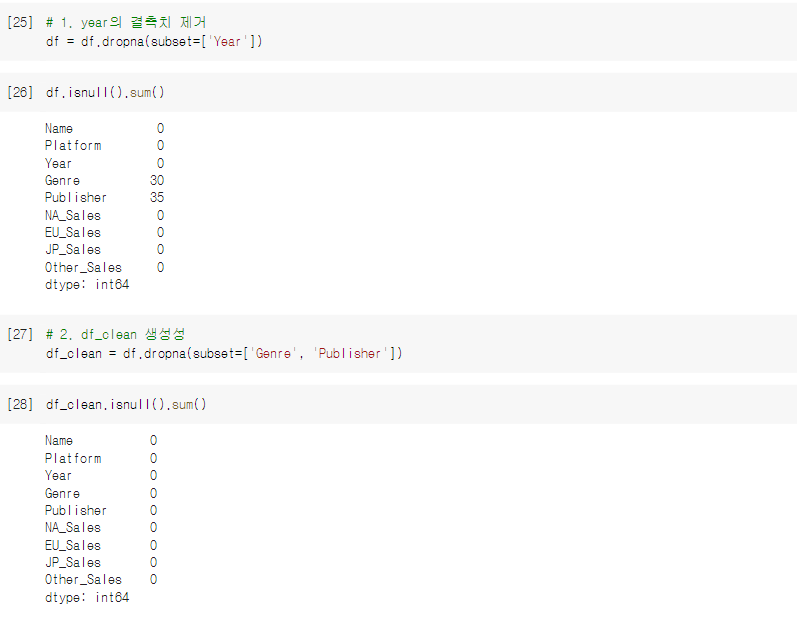
자동 생성된 설명

그래도 결측치가 존재

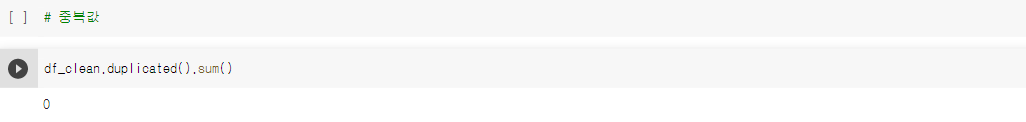
텍스트, 실내, 스크린샷이(가) 표시된 사진

자동 생성된 설명

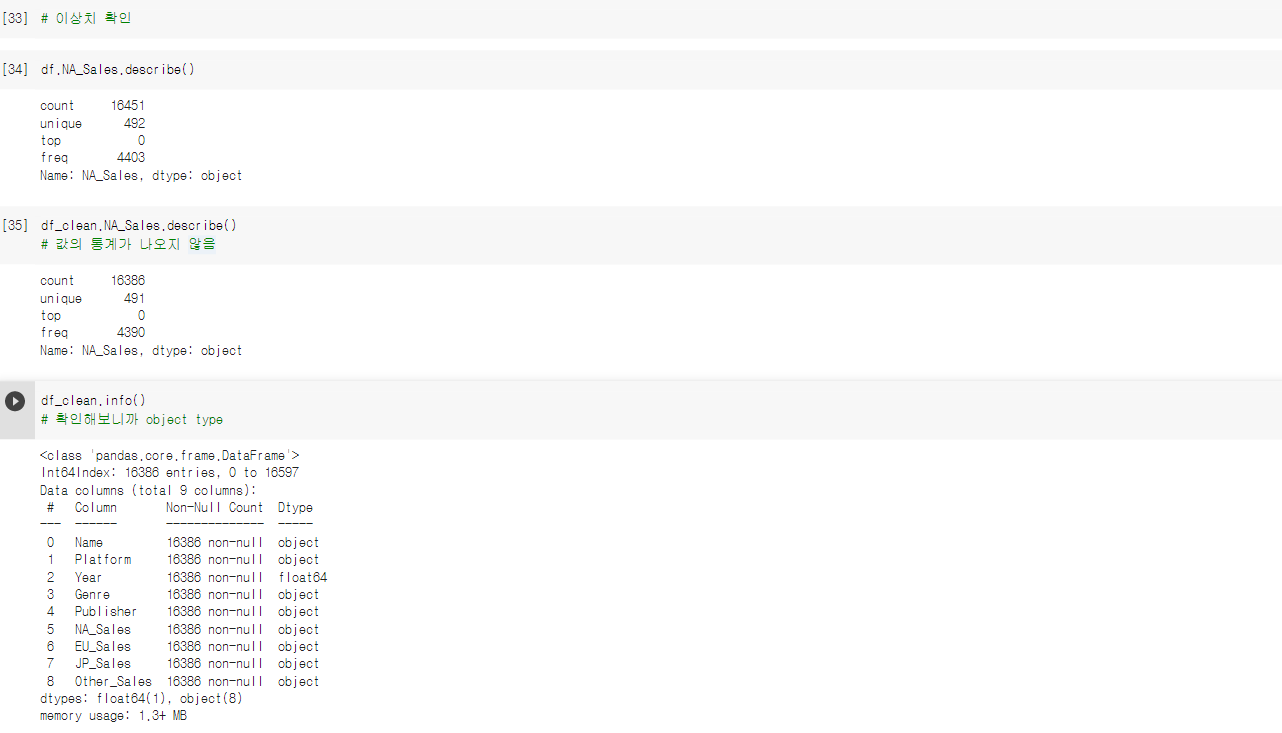
확인해보니 모두 결측인 경우에는 채우지 못함 – 년도, 장르, 배급사가 결측이면 채우지 못함.



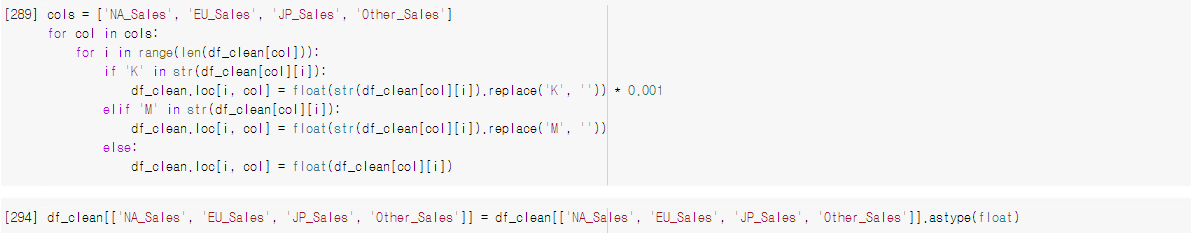
따라서 나머지 결측치는 모두 제거해주기

중복값

없음

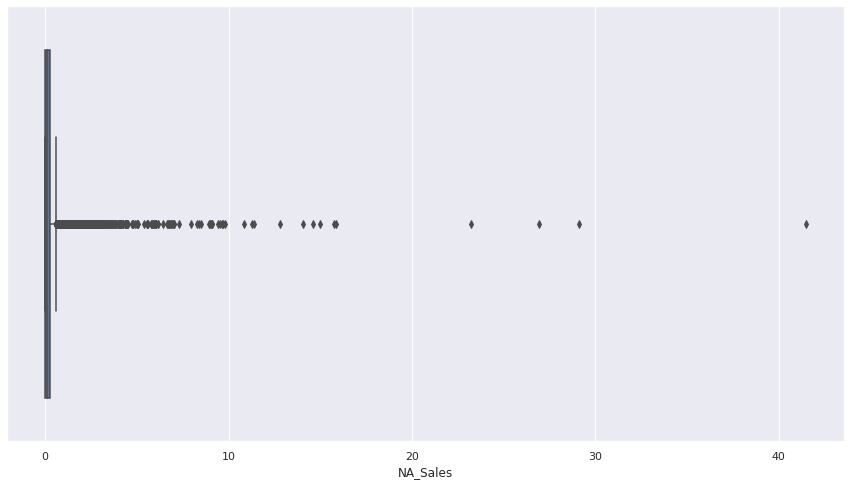
이상치

이상치를 확인해보기 위해 통계값을 확인했지만, object형태라 확인할 수 없음.

확인해보니텍스트이(가) 표시된 사진

자동 생성된 설명 K와 M이라는 문자가 있음

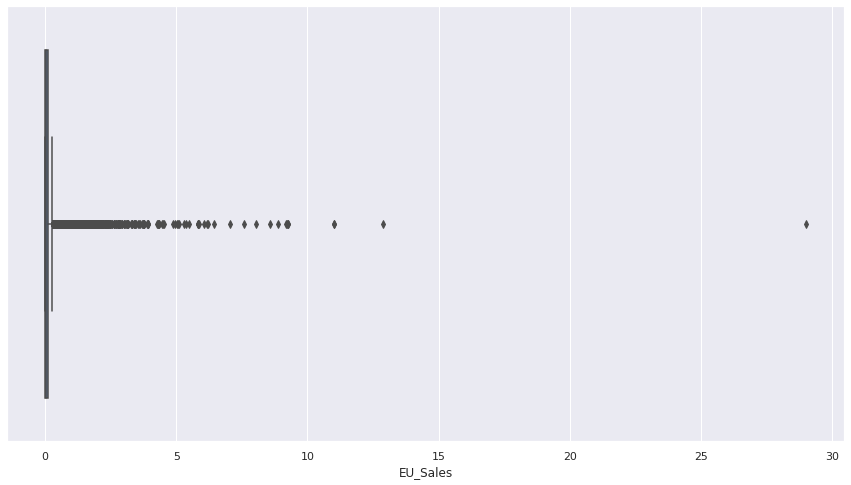
따라서 for문으로 M은 제거해주고, K는 제거해준 후 0.001을 곱해줌, float로 전환

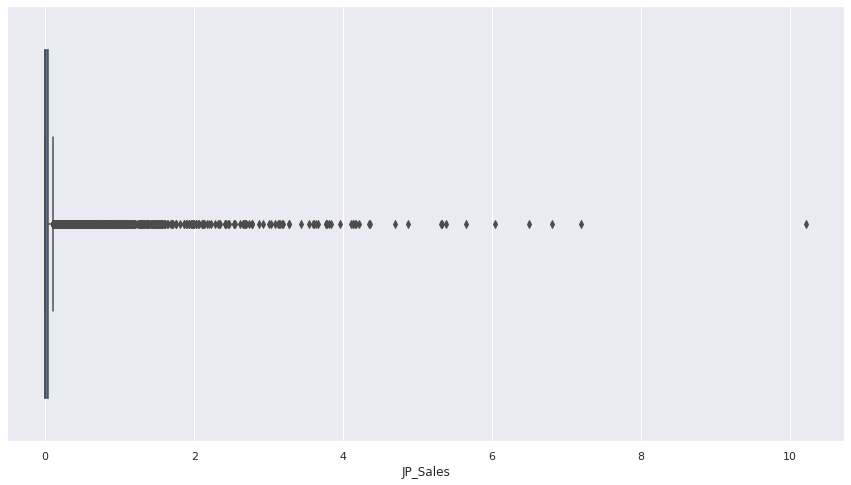
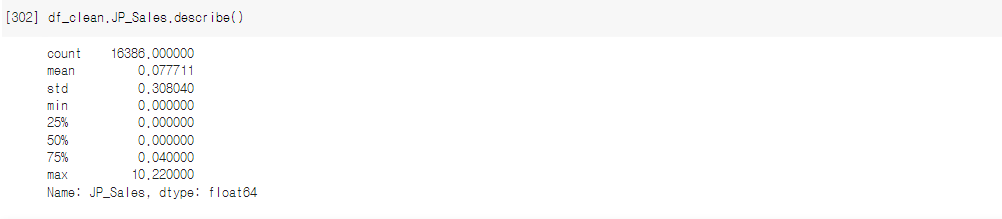


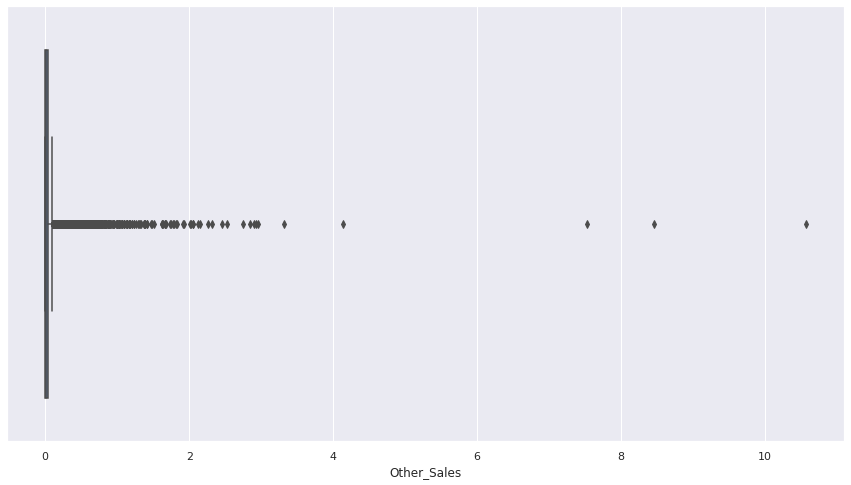
텍스트이(가) 표시된 사진

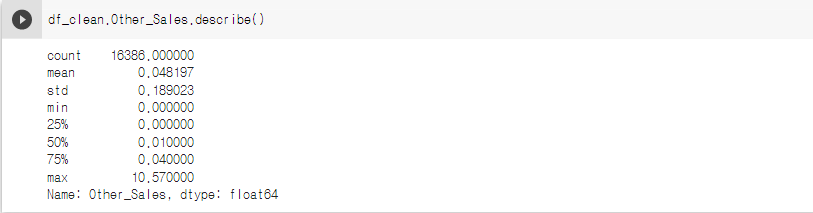
자동 생성된 설명

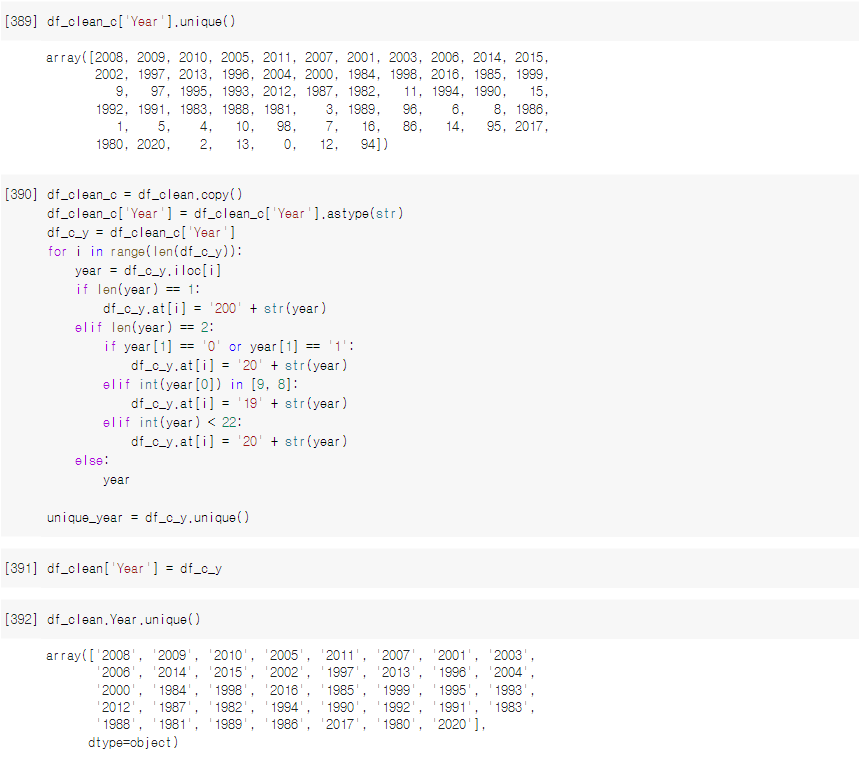
텍스트이(가) 표시된 사진

자동 생성된 설명





max값이 평균값에 비해서 크다 – 하지만 게임의 특성상 일부 게임이 흥행함. 따라서 이상치라고 판단하지 않고 분석을 진행함.

기타. 연도의 오류값

2. 분석

3. 결론