|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 데이터 | 날짜 | 날짜끝 | 데이터수 | 영화수 |
| Kobis\_final\_table | 2003-01-01 | 2018-08-02 | 53,253 | 4,663 |
| Sorting\_all\_movie | 2017-01-02 | 2018-06-29 | 1,350,000 | 561 |
| Bef\_sorting\_all\_movie | 2017-01-03 | 2018-06-27 | 106,218 | 474 |

각 데이터 구분

Kobis\_final\_table

Columns

movie index : 영화 구분에 좋겠지?

스크린수, 스크린점유율, 상영횟수

좌석수 좌석점유율

좌석판매율, 매출액, 매출액증감

관객수, 관객수 증감

누적 매출액, 누적 관객수, 순위

리뷰데이터 사용 – 개봉전 특징 개봉후 특징 나눠서 가져가기

리뷰의 변화에 따른 매출액 변화

간단한 평균 사용

매출데이터로 바꿔서

특정 단어 빈도에 따른 잠재성 평가

기본 데이터와 맞지 않는 문제 : 리뷰 데이터 사용과 사용전 비교

데이터 분석 방법 –

1. 목표 – 매출액 예측을 통해서 영화의 매출 파악?
   * 1. 매츨액 예측
     2. 매출이 높은 영화의 특징
     3. 수요 예측 – 데이터 안맞을거 같은데?
2. 가설
3. 방법
   1. 리뷰데이터 기간별 특징
4. 문제점
   1. 리뷰데이터를 사용하기에 데이터 기간이 맞지 않다.

[통상적으로 업계에서는 영화 흥행에 영향을 미치는 내외적 요소를 다음과 같이 손꼽고 있다](https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiOrteServHistIFrame.kci?sereArticleSearchBean.artiId=ART002253444&sereArticleSearchBean.orteFileId=KCI_FI002253444). 영화 흥행에 미치는 내적 요인(제작 단계)에는 감독의 역량, 출연 배우, 예상 관람등급, 원작 유무, 장르, 시나리오의 완성도, 총제작비 등이 있다. 외적 요인(배급 및 상영 단계)에는 배급사의 규모와 영향력, 개봉 시기, 스크린 수, 해외영화제 수상 여부, 전략적인 마케팅 등이 있다.

앞서 언급된 영화들은 흥행요소를 두루 갖췄으나 결과적으로는 손익분기점을 넘지 못했다. 두 영화의 공통점은 바로 시나리오의 완성도에 의문점을 남겼다는 점이다. [영화 한 편의 모든 가능성을 담고 있는 설계도](https://eiec.kdi.re.kr/publish/nara/column/view.jsp?idx=5897)가 미흡하다면, 제아무리 화려한 제작인력과 배우가 투입되더라도 흥행으로 이어지지 못할 수 있음을 보여준 단적인 예다.

[영화를 평가하거나 흥행도 예측이 필요한 시점은 영화 출시 전후가 아니라, 시나리오가 완성된 시점이라는 주장에 힘이 실리는 배경](http://www.dbpia.co.kr/Journal/ArticleDetail/NODE02466494)이다. 이는 영화제작 과정에서 발생하는 고정비가 발생하는 시점의 특이성 때문이다. 고정비의 상당 부분은 제작준비 및 제작 단계에서 발생한다. 고정비가 곧 제작비라고 봐도 무방하다. 반면, 소재 발굴 및 시나리오 개발 단계에서는 비용이 거의 들지 않는다. [투자, 캐스팅 등이 시나리오를 기반으로 이뤄지는 만큼, 영화제작 기간 중 가장 큰 노력과 시간을 들여야 하는 부분](https://edu.kocca.kr/edu/bbs/B0000020/view.do?nttId=41398&searchCnd=&searchWrd=&gubun=&delCode=&useAt=&replyAt=&menuNo=500011&sdate=&edate=&viewType=&type=&siteId=&option1=&option5=&pageIndex=15)이 바로 이 시나리오다. 이런 논리에 따르면 영화 제작 전, 즉 시나리오 투자 여부를 결정하는 시점에서 비용과 수익에 대한 정밀한 계산이 필요하다는 결론을 도출해 볼 수 있다.