## 

## 

## 

## 

## 

## 

## 

## 1. Data Analysis and Insights

## 1.1 Consumer-Centric Analysis

* 관객수
* Budget 대비 Revenue: 수식: revenue/budget
* 평점
* Runtime

## 1.2 Popularity and Ratings Correlation

* 가장 인기 있는 영화와 사용자 평가 간의 상관관계 분석
* 영화 런닝타임이 길수록 사용자의 평점이 높아지는지 낮아지는지 런닝타임과 평점 간의 상관관계 분석
* 영화 장르별로 평균평점을 냈을 때, 가장 평점이 높은 장르는 무엇인지, 그 이유는 무엇일지 원인 분석
* 영화리뷰 작성 시간대별로 리뷰내용이나 평점에 어떤 영향을 미치는지 연관관계 분석

## 1.3 Genre and Keywords Analysis

* 특정 장르나 키워드가 많은 영화를 파악하고, 장르와 키워드 간의 연관성을 분석
* 특정 키워드가 영화 평점에 미치는 영향
* 영화 정보에 따른 유행성 있는 태그라인 생성 모델 제작

## 1.4 External Database and Ratings

* 외부 데이터베이스와 연결된 영화를 분석
* 평점과 외부 영화 정보 간의 관계를 분석
* 키워드 및 평점 기반으로 영화의 트렌드를 분석

## 1.5 Economic Impact Analysis

* 영화 산업의 경제적 변화를 분석
* 어떤 제작사가 가장 성공적인지를 분석
* 특정 시기에 개봉한 영화들의 성과를 분석

## 2. Predictive Modeling

## 2.1 Revenue Prediction Models

* 영화 수익 예측 모델: 제작국가, 제작 예산 규모, 장르 등의 요인을 입력하여 수익 예측하는 모델 만들기 + casting 정보도 활용 가능하면 좋을 듯

## 2.2 Success Prediction Models

* 영화 흥행 여부 예측 모델: 제작국가, 제작 예산 규모, 장르 등의 요인을 입력하여 흥행 여부 예측 모델 만들기 + casting 정보도 활용 가능하면 좋을 듯

## 2.3 Rating Prediction Models

* 영화 평점 예측 모델: 예산, 상영시간, 장르, 국가, 인기도 등의 요인을 입력하여 회귀모델 제작

## 2.4 Genre Classification Models

* 영화 개요를 활용하여 장르 분류 모델 제작

## 2.5 Tagline Generation Models

* 영화 정보에 따른 유행성 있는 태그라인 생성 모델 제작

## 2.6 Recommendation Systems

* 영화 추천 시스템

## 3. Detailed Analysis Topics

## 3.1 Popularity Factors

* 영화 인기도에 영향을 주는 요인을 분석 및 인사이트 도출

## 3.2 Series vs Non-Series Movies

* 시리즈 영화 대비 시리즈 영화가 아닌 영화의 흥행 차이 분석

## 3.3 Genre Profitability

* 장르별 수익성 분석: 장르별 수익성과 인기도를 분석하여 어떤 장르가 수익성이 높은지 파악 > 영화 투자사 입장에서의 인사이트

## 3.4 Release Timing and Success

* 개봉 시기와 흥행 성적의 상관관계 분석: 영화 개봉 시기와 흥행 성적의 상관관계 분석

## 3.5 Production Company Success Patterns

* 제작사별 성공 패턴 분석: 제작사별로 성공적인 영화의 패턴을 분석하여 어떤 제작사가 어떤 유형의 영화에서 성공적인지 파악 > 제작사별로 강점을 분석하여 제작사 선정 인사이트 도출

## 4. Additional Analyses

## 4.1 Revenue and Vote Average Analysis

* Revenue에 영향을 미친 요소 분석
* Vote\_average에 영향을 미친 요소 분석 및 시각화
* Revenue와 vote\_average 간의 상관관계 분석

## 4.2 User Ratings and Recommendations

* Ratings 파일의 user가 영화에 매긴 rating을 바탕으로 높은 점수를 매긴 영화와 유사한 데이터(production\_company, genre, belongs\_to\_collection, keyword, title 등)를 가진 영화를 추천

## 4.3 Runtime and Ratings by Genre

* 영화 장르(gener\_name)별 러닝타임에 사용자의 평점이 어떻게 달라졌는지 분석
* 태그라인에 쓰인 단어의 갯수가 평점에 얼마나 달라졌는지 분석
* 계절별 흥행한 영화 장르가 평점에 영향을 미쳤는지 분석

### **1. 소비자 관점에서의 데이터 분석**

* **관객수 및 수익성 분석**
  + 관객수
  + Budget 대비 Revenue
  + Revenue/Budget
* **평점과 런타임**
  + 평점
  + 런타임
  + 평점과 런타임의 상관관계 분석
* **국내/외 영화 구분**
  + 국내 영화
  + 해외 영화

### **2. 영화 콘텐츠 및 속성 분석**

* **장르 및 키워드**
  + 특정 장르 및 키워드 분석
  + 장르와 키워드 간의 연관성
* **제작사 및 제작 정보**
  + 성공적인 제작사 분석
  + 제작사별 평균 평점
* **개봉 시기와 성과**
  + 특정 시기 개봉 영화 성과 분석
  + 계절별 흥행 장르 분석

### **3. 영화 평점 및 리뷰 분석**

* **평점 관련 분석**
  + 평점에 영향을 미치는 요소
  + 평점과 외부 영화 정보의 관계
  + 리뷰 작성 시간대별 영향
* **트렌드 및 인기**
  + 영화 트렌드 분석
  + 인기도에 영향을 주는 요인

### **4. 데이터 모델링 및 예측**

* **수익 및 평점 예측 모델**
  + 수익 예측 모델 (제작국가, 예산, 장르 등)
  + 평점 예측 모델 (예산, 상영시간, 인기도 등)
* **흥행 예측 모델**
  + 흥행 여부 예측 모델
  + 시리즈 영화 vs 비시리즈 영화 흥행 비교
* **장르 분류 및 추천 시스템**
  + 영화 개요를 통한 장르 분류
  + 영화 추천 시스템 (태그라인 생성 포함)

### **5. 추가적인 분석 및 딥러닝 적용**

* **포스터 이미지를 활용한 딥러닝**
  + 이미지 분석을 통한 트렌드 예측
  + 포스터 이미지와 영화 성과의 관계
* **상관관계 분석**
  + Revenue에 영향을 미친 요소
  + Vote\_average에 영향을 미친 요소
  + Revenue와 Vote\_average 간의 상관관계

## **별도의 페이지로 분리된 주제**

아래의 주제들은 각각 별도로 다루어질 수 있습니다.

### **A. 영화 제작사 및 수익 분석**

* 제작사별 수익성
* 제작사의 성공 패턴

### **B. 시기별 유행 및 장르 분석**

* + 시대/월별 유행 장르
  + 계절별 흥행 영화 장르

### **C. 키워드 및 텍스트 분석**

* + 태그라인의 단어 갯수와 평점의 관계
  + 특정 키워드의 평점 영향