**상위권 유지 요인 분석 및 투자 전략**

**목차**

**프로젝트 배경 및 목적**

이 프로젝트는 전 세계적으로 가장 큰 인기를 누리는 잉글랜드 프리미어리그 경기들을 지켜보며 리그 상위권 클럽들이 어떻게 지속적으로 성공을 거두고 있는지에 궁금증에서 시작되어, 프리미어리그에서 강력한 입지를 구축한 상위권 클럽들의 핵심 성공 요인을 데이터 분석을 통해 도출하는 것을 목표로 합니다. 특히, 본 분석은 투자자의 관점에서 프리미어리그 구단 운영의 수익성과 지속 가능성을 분석하고, 최적의 투자 전략을 수립하는 데 중점을 둡니다.이러한 분석 결과를 바탕으로 투자자가 장기적으로 높은 수익을 얻을 수 있는 구단 투자 전략을 도출하는 것이 본 연구의 최종 목표입니다.

**분석 주요 문제**

본 프로젝트를 통해 해결하고자 하는 주요 문제는 다음과 같습니다.

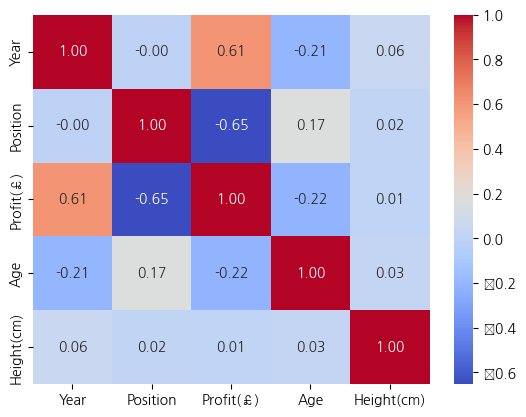
* 프리미어리그의 상위권 클럽들은 어떠한 요인들로 인해 지속적인 수익 창출이 가능한가?
* 투자자가 최대한의 수익을 얻기 위해 고려해야 할 핵심 요소는 무엇인가?
* 리그 성적과 구단의 재정적 성장 간의 관계는 어떻게 형성 되는가?

**데이터 소개 및 분석**

* 2014-2015 시즌부터 2023-2024 시즌까지 최근 10년간의 프리미어리그 데이터 활용
* 총 5가지 데이터로 구성: 리그 순위, 시즌별 팀명, 팀의 평균나이, 수익, 선수 연봉 및 키
* 병합한 데이터는 리그 순위, 팀의 수익, 평균 나이
* 선수 연봉와 키 데이터는 해당 시즌 모든 선수를 포함하지 않아 누락값 발생 이유로 병합하지 않았습니다.

**EDA 분석**

* 리그 순위와 팀 수익 간의 상관관계가 유의미하다 (p-value: 7.95e-09)
* Position(순위)과 Profit(수익)의 상관계수는 -0.65를 기록하여, 다른 변수에 비해 강한 상관관계를 가지고 있습니다.



*상관관계 히트맵*

**머신러닝 분석**

**선형 회귀 분석:**

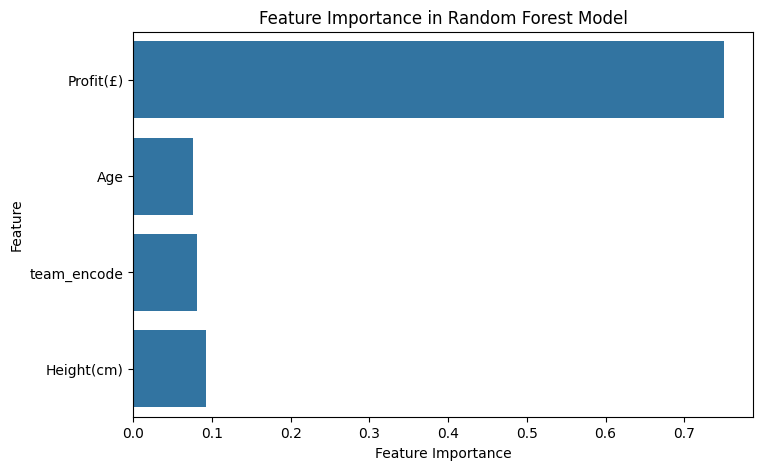
* R²: 0.573
* RMSE: 3.4366
* 결과: 예측 정확도 낮음

**랜덤 포레스트 회귀 분석:**

* R²: 0.8657
* RMSE: 1.9271
* 결과: 예측 오차 감소 및 정확도 크게 향상

**순위예측 모델**

상위권(7위이상) 도달을 위해서 최적의 수익은 £90 million(9천만 파운드)이상이며, 최소값은 £50 million(5천만 파운드)입니다.



피처 중요도 그래프에서 Profit(수익)이 핵심적인 피처임을 확인하였습니다. 상위권을 도달하기 위해서는 수익 외에 다른 피처들은 중요하지 않다는 결과가 나왔습니다.

**결론**

순위 예측 모델은 단순한 예측을 넘어 구단 운영 및 마케팅에 유용한 정보를 제공할 수 있습니다. 투자자의 관점에서 이를 활용해 구단 운영진과 투자 우선순위를 토론 및 결정하여 투자 위험을 최소화하고 안정적인 수익을 추구할 수 있습니다.

**마케팅 전략**

1. **아시아 시장 투자**

* 아시아 전 세계 팬덤이 가장 큰 축구 팬덤 보유
* 비시즌 기간의 월드투어를 통해 수익과 클럽 인지도 상승 효과
* 아시아 유소년 영입 혹은 아카데미 설립하여 장기적으로 잠재적 팬 확보

1. **팬 서비스 확장**

* 시즌 시작 전, 중반, 종료 시점에 예측 모델 활용하여 팬들과 순위 예측 토론 및 콘텐츠 제작
* 팬들과의 소통 채널을 다양화
* 경기 후 혹은 경기 외적으로 선수들의 비하인드 영상 제작으로 기존 팬 유지와 잠재적 팬 확보를 할 수 있습니다. 예를 들어, 한국에서 손흥민 선수의 경기 전, 후의 직캠 영상이 조회수가 높아 많은 사람들, 특히 한국팬들을

**개선사항**

추가 변수: 날씨, 경기 결과, 선수 이적료, 팬 특성 등 추가 변수 고려 시 모델 정확도 향상 기대.

데이터 보완**:** 전 시즌 선수 키, 연봉 데이터 확보 시 분석 신뢰성 증대