



유저 샷 분석

유저 샷 분석

1. 서론 (Introduction)

(주) QED 콘솔에 내재화할 Q-Fit 사용자 입력 데이터의 평균값 신뢰성을 파악하기 위해 유저 샷 분석을 진행하였습니다. 본 분석에서는 2022년 6월과 8월의 샷 데이터를 통합하여 약 200만 개의 데이터 행을 활용하였으며, 데이터를 정제하여 드라이버 샷을 중심으로 분석하였습니다.

가설 설정:

1. **가설 1:** 데이터 전처리를 통해 이상치를 제거하면 정규성을 가지는 유저 비율이 증가할 것이다.
2. **가설 2:** 샷 데이터 수가 많을수록 정규성을 가지는 유저의 비율이 감소할 것이다.
3. **가설 3:** 특정 기준(비거리, Back Spin 등)으로 데이터를 전처리하면 일부 특성에서 정규성이 향상될 것이다.

2. 데이터 전처리

데이터는 아래와 같이 4가지 유형으로 분류하여 전처리되었습니다.

번호	데이터 유형	Total Distance 기준	Back Spin 기준
1	원본 데이터	전체	전체
2	드라이버만	상위 90%	하위 90%
3	드라이버만	상위 80%	하위 80%
4	드라이버만	상위 70%	하위 70%

데이터 샘플링 과정에서 유저당 샷 개수에 따른 정규성을 분석하였습니다. 각 샷 데이터 유형별로 랜덤하게 1,000명을 선택하여 **Back Spin, Club Speed, Launch Angle, Face Angle, Club Path, Side Distance** 특성에 대해 정규성을 분석하였습니다.

3. 분석 결과

가설 1 검증:

- 원본 데이터 대비 전처리된 데이터에서 정규성을 가지는 유저 비율이 전반적으로 증가함.
- 특히, Total Distance와 Back Spin의 이상치를 제거한 데이터일수록 정규성을 가진 유저 비율이 높아짐.

가설 2 검증:

- 샷 데이터 수가 많아질수록 정규성을 가지는 유저의 비율이 감소하는 경향을 보임.
- 특히, 샷 개수가 300개 이상인 그룹에서는 정규성이 급격히 낮아짐.

가설 3 검증:

- Total Distance와 Back Spin을 기준으로 전처리했을 때 Face Angle과 Club Path 특성에서 정규성 증가가 두드러짐.
- 그러나 Back Spin 특성의 경우 정규성을 가지는 유저 비율이 감소하는 경향을 보임.

4. 결론 및 향후 연구 방향

본 연구를 통해 이상치를 제거할수록 Q-Fit 데이터의 신뢰성이 향상될 가능성이 있음을 확인하였습니다. 그러나 데이터 손실이 발생할 위험도 있으므로, 적절한 전처리 기준을 설정해야 합니다.

향후 연구 방향:

1. 샷 개수에 따른 평균, 표준편차, 최대값, 최소값을 분석하여 드라이버, 웨지, 아이언 샷을 구분하는 추가 연구 진행.
2. 정규성 외에도 데이터 품질 지표(예: 변동성, 일관성 등)를 고려한 심층 분석.
3. 이상치 제거에 따른 데이터 손실을 최소화하는 최적의 전처리 방법 도출.

5. 데이터 출처

- 2208_150m_driver.xlsx (2022년 8월 샷 데이터, Sheet1/Sheet3)
- 드라이버추출_220610.xlsx (2022년 6월 샷 데이터, 스윙옵션/스윙옵틱스)