



유저샷분석_2

1. 연구 목적 (Research Purpose)

본 연구는 (주) QED 앱과 콘솔 내에 적용될 Q-Fit 사용자 입력 데이터의 신뢰성을 검토하고, 드라이버, 아이언, 웨지 샷을 구분하기 위한 분석을 수행하는 것을 목적으로 합니다. 이를 위해 2022년 6월과 8월의 샷 데이터를 통합하여 약 200만 개의 데이터 행을 활용하였습니다. 데이터 분석을 통해 각 샷 유형의 정규성 비율, 평균값, 최소값, 최대값, 표준편차를 측정하고 드라이버 샷을 식별할 수 있는 기준을 도출하고자 합니다.

2. 가설 설정 (Hypothesis)

- 가설 1:** 샷 데이터의 개수가 많을수록 정규성을 가지는 유저의 비율이 감소할 것이다.
- 가설 2:** 비거리를 기준으로 상위 데이터를 선택하여 전처리하면, 정규성을 가지는 유저 비율이 증가할 것이다.
- 가설 3:** 비거리 전처리를 수행하면 데이터의 평균값과 최소/최대값 간 차이가 줄어들어 샷 유형을 보다 정확하게 분류할 수 있을 것이다.
- 가설 4:** 드라이버 샷의 정규성 비율은 Face Angle 및 Club Path와 같은 특성에서 더욱 뚜렷하게 나타날 것이다.

3. 데이터 분석 (Analysis)

한 유저 기준 300개 이상 가진 샷 데이터 (원본 데이터)

샷 데이터 수가 많은 경우, 전체적인 특성들의 정규성 비율은 낮으며 최소값과 최대값이 평균값과 차이가 큼니다.

특성	총 데이터 수	평균 데이터 수 (유저당)	정규성 수	비정규성 수	정규성 비율(%)	평균값	최소값
Back Spin	573549	451	64	1209	5.03%	2400.30	-12763
Club Speed	573549	451	23	1250	1.81%	30.67	0.0
Launch Angle	573549	451	29	1244	2.28%	17.28	-24.51
Face Angle	573549	451	170	1103	13.35%	0.90	-89.94
Club Path	573549	451	147	1126	11.55%	0.81	-102.30
Total Distance	573549	451	9	1264	0.71%	119.69	0.54
Side Distance	573549	451	194	1079	15.24%	0.71	-108.65

한 유저 기준 300개 이상 가진 샷 데이터 (비거리 상위 10%만)

비거리 상위 10%만을 선택한 경우, 정규성 비율이 대폭 증가하였으며 최소값과 최대값이 평균값과의 차이가 감소하였습니다.

특성	총 데이터 수	평균 데이터 수 (유저당)	정규성 수	비정규성 수	정규성 비율(%)	평균값	최소값
Back Spin	57926	45.5	941	331	73.98%	2184.20	-7404
Club Speed	57926	45.5	510	762	40.09%	35.36	0.0
Launch Angle	57926	45.5	927	345	72.88%	15.88	-18.4
Face Angle	57926	45.5	1138	134	89.47%	0.07	-57.42
Club Path	57926	45.5	971	301	76.34%	0.45	-37.62
Total Distance	57926	45.5	185	1087	14.54%	164.58	5.61
Side Distance	57926	45.5	1074	198	84.43%	-2.74	-85.45

4. 결론 및 향후 연구 방향 (Conclusion & Future Research)

본 연구를 통해 비거리 기준으로 데이터를 전처리하면 정규성 비율이 증가하고, 평균값과 최소/최대값의 차이가 줄어드는 경향을 확인하였습니다. 이는 드라이버 샷을 보다 정확하게 구별하는 데 기여할 수 있는 결과입니다.

결론:

- 데이터 전처리를 통해 이상치를 제거하면 정규성 비율이 증가하는 것이 확인됨.

2. 샷 데이터 개수가 많을수록 정규성 비율이 감소하는 경향을 보임.
3. 특정 특성(예: Face Angle, Club Path)에서 정규성 비율이 상대적으로 높게 나타남.

향후 연구 방향:

1. 비거리 전처리 기준을 상위 5% 및 10%로 구분하여 추가 분석.
2. 300개 이상의 샷 데이터를 보유한 약 1270명의 사용자 중 랜덤 10명을 선정하여 평균값 비교 및 드라이버 샷 판별 테스트 수행.
3. 정규성 분석뿐만 아니라, 데이터 분포 및 신뢰도 지표 분석을 통한 샷 유형 분류 모델 개선.

5. 데이터 출처 (Data Source)

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