



# 유저 샷 분석 2

## Introduction

QED앱과 콘솔에 내재화 시킬때 필요한 Q-Fit 사용자 Input 데이터의 평균값 신뢰성을 파악 및 드라이버, 아이언, 웨지등을 분류하기 위해 샷 분석을 하였습니다. 주어진 데이터는 2022년 6, 8월 샷 데이터를 통일된 컬럼들에 맞추어 병합하였고 대략 200만개의 데이터 행 (전체 유저의 전체 샷 개수)을 활용하였습니다. 전체 샷 데이터중에서 30~100개의 유저수는 6485명, 100~200개의 유저수는 3572개, 200~300개의 유저수는 1279개, 300개 이상의 유저수는 1279개 입니다. 샷 데이터 갯수별/전처리된 비거리별/특성별로 정규성 비율(%), 평균값, 최소값, 최대값, 표준편차값을 나타내었습니다. 해당 값들을 통해서 한 사용자 기준으로 드라이버/아이언/웨지 샷들중 드라이버 샷을 구분해보겠습니다.

## Analysis

한 유저 기준 300개이상 가진 샷데이터(원본 데이터)

	back spin	club speed	lanch angle	face angle	club path	total distance	side dist
총 데이터 수	573549	573549	573549	573549	573549	573549	573549
평균 데이터수(유저당)	451	451	451	451	451	451	451
정규성 수	64	23	29	170	147	9	194
비정규성 수	1209	1250	1244	1103	1126	1264	1079
정규성(%)	5.03%	1.81%	2.28%	13.35%	11.55%	0.71%	15.24%
평균값	2400.30	30.67	17.28	0.90	0.81	119.69	0.71
min값	-12763	0.0	-24.51	-89.94	-102.30	0.54	-108.65
max값	16738	72.80	78.70	91.94	129.10	339.74	109.36
표준편차값	497.25	2.43	2.35	2.53	1.69	14.33	6.97

전체적으로 특성들의 정규성 비율은 현저히 낮고 최소값과 최대값의 값들이 평균값과의 차이가 있습니다. 하지만, 예상보다는 표준편차값의 크기가 크지는 않았습니다.

한 유저 기준 300개이상 가진 샷데이터(비거리 상위 10%만)

	back spin	club speed	lanch angle	face angle	club path	total distance	side dist
총 데이터 수	57926	57926	57926	57926	57926	57926	57926
평균 데이터수(유저당)	45.5	45.5	45.5	45.5	45.5	45.5	45.5
정규성 수	941	510	927	1138	971	185	1074
비정규성 수	331	762	345	134	301	1087	198
정규성(%)	73.98%	40.09%	72.88%	89.47%	76.34%	14.54%	84.43%
평균값	2184.20	35.36	15.88	0.07	0.45	164.58	-2.74
min값	-7404	0.0	-18.4	-57.42	-37.62	5.61	-85.45
max값	8604	65.80	44.37	35.01	85.30	339.74	93.96
표준편차값	276.78	2.17	1.04	1.18	1.23	6.01	6.35

정규성 비율은 각각의 특성들이 높고 최소값, 최대값들이 평균값과의 차이가 현저히 줄었습니다. 특성들마다의 표준편차의 값들도 원본데이터보다 대략 두배이상 줄었습니다.

그외에도 데이터 갯수별/비거리 전처리별에 따른 다양한 상황적인 결과가 있지만 많은 관계로 위의 두개만 나타내었습니다.

## Result

한 유저 기준으로 원본 데이터와 전처리 데이터(비거리 상위 10%)의 정규성 비율(%), 평균값, 최소값, 최대값, 표준편차의 수치값을 비교 하였습니다. 원본데이터와 대비해서 전반적으로 전처리 데이터의 정규성 비율(%)은 특성별로 많이 증가하였고 원본데이터의 평균값, 최소값, 최대값에 대비해서 전반적으로 차이가 많이 있습니다. 표준편차값에서도 각각의 특성마다 대략 두배 이상 줄어듦을 볼수 있습니다. 이러한 결과는 드라이버, 웨지, 아이언 샷중에서 드라이버 샷라고 구별하기에 의미있는 결과라고 볼수 있습니다.

다음 단계는 비거리 전처리별로 상위 5%와 10%만의 데이터를 분류하여 300개 이상의 샷데이터를 가진 사용자 대략 1270명중에 각각 10명씩 랜덤으로 추출하여 각각의 특성들의 평균값과 실제 해당 샷데이터의 사용자들의 드라이버샷을 비교하여 테스트 할 예정입니다.

## Data Source

1. 2208\_150m\_driver.xlsx(2022년 8월 샷데이터, sheet1/sheet3)
2. 드라이버추출\_220610.xlsx(2022년 6월 샷데이터, 스윙옵션/스윙옵틱스)