

22, 23season KBO DATA

김도현 박상혁 장희수 홍인화



TABLE OF CONTENTS



01

KBO 관련 DATA 크롤링

02

DATA 시각화 및 상관관계 비교

03

INSIGHT

22, 23 KBO DATA



01

KBO 데이터 크롤링

22, 23 KBO DATA

타자의 연봉과 성적이 비례할까?

- 연봉과 타율 비교
- 연봉과 안타수 비교

타자의 안타와 홈런의 상관관계는?

- 타자의 1,2,3루타율과 홈런율의 비교

투수 성적과 팀 성적의 상관관계는?

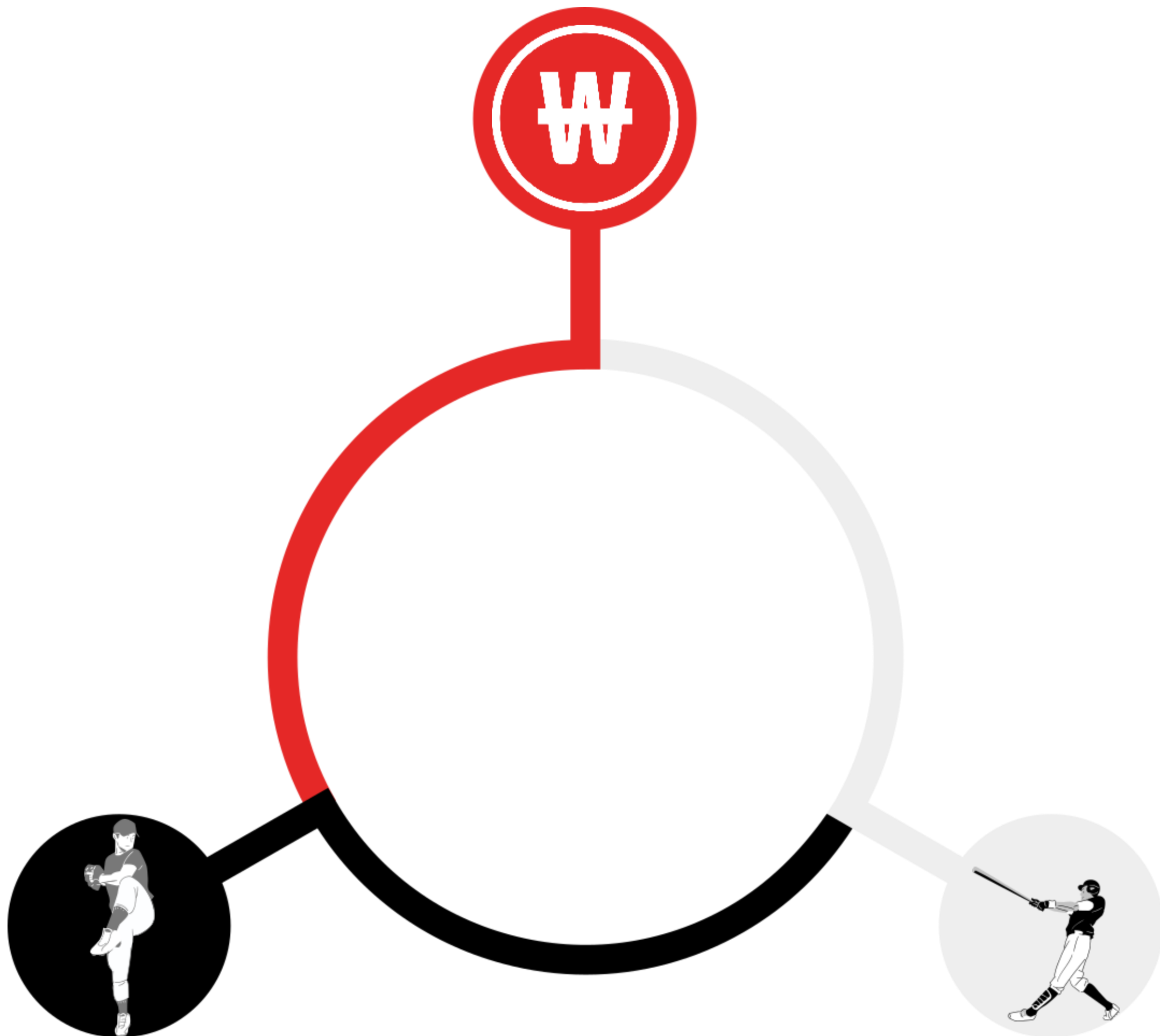
- 승률과 평균자책점과의 상관관계
- 승률과 평균피안타와의 상관관계

타자평균연봉과 투수평균연봉 비교

- T test


인구가 많은 지역이 야구부도 많을까?

- 전국인구 수와 학생 야구부수 지역별 비교



1. KBO, 야구관련 데이터 크롤링

- KBO





일정/결과순위기록선수NEWSKBO퓨처스리그KBO 마켓

기록

TOP5
주간/월간 랭킹
기록실
투수 vs 타자
역대기록실
주간/일간 예상 달성기록
기록 정경 현황

선수페이지 (투수)



선수명: 김광현
생년월일: 1988년 07월 22일
신장/체중: 188cm/88kg
입단 계약금: 50000만원
지명순위: 07 SK 1차

등번호: No.29
포지션: 투수(좌투좌타)
경력: 덕성초(안산리틀)-안산중양중-안산공고-SK-세인트루이스
연봉: 100000만원
입단년도: 07SK

타자

투수

기록용어

기본기록통산기록일자별기록경기별기록상황별기록등록일수

2022 성적

팀명	ERA	G	CG	SHO	W	L	SV	HLD	WPCT	TBF	NP	IP	H	2B	3B	HR
SSG	2.13	28	0	0	13	3	0	0	0.813	694	2634	173 1/3	141	16	2	10


SAC	SF	BB	IBB	SO	WP	BK	R	ER	BSV	WHIP	AVG	QS
6	2	45	2	153	1	0	48	41	0	1.07	0.222	19

최근 10경기

일자	상대	결과	ERA	TBF	IP	H	HR	BB	HBP	SO	R	ER	AVG
10.05	두산	패	6.00	24	6	6	1	1	0	8	4	4	0.261

KBO 퓨처스리그 기록 보기 >

- 한국 소프트볼 야구 협회



협회소개미디어센터일정결과정보기록실국가대표로그인회원가입

야구정보

팀 선수정보를 비롯한 야구정보를 제공합니다.

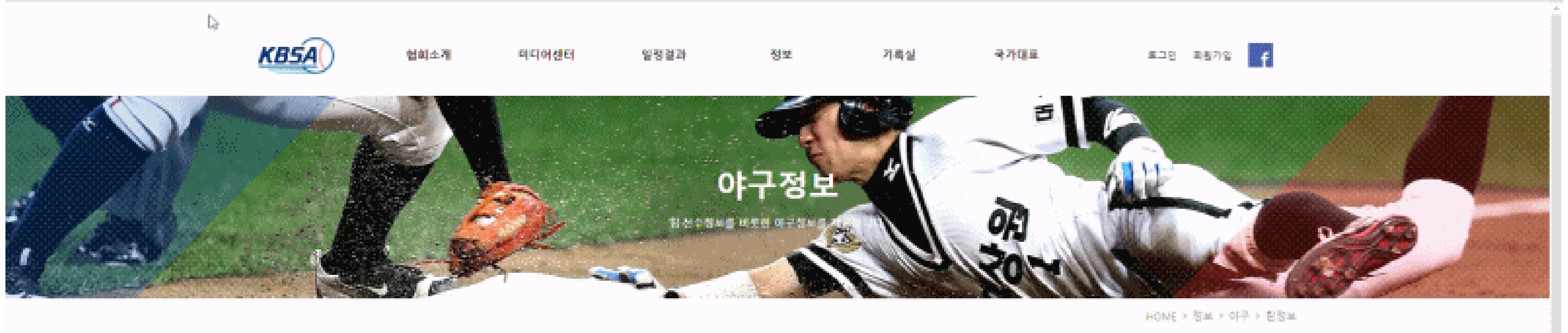
HOME > 정보 > 야구 > 팀정보

팀정보

대학부19세 이하부16세 이하부13세 이하부일반부생활체육스포츠클럽부

2023지역전체검색

년도	팀명	팀구분	지역	선수수	감독	
66	2023	강원 영랑초	13세 이하부	강원	16	권용석
65	2023	강원 알산초	13세 이하부	강원	13	정인규
64	2023	경기 SDF투칸스(U13)	13세 이하부	경기	9	
63	2023	경기 백마초	13세 이하부	경기	15	
62	2023	경기 삼일초	13세 이하부	경기	10	
61	2023	경기 수원신곡초	13세 이하부	경기	15	



팀정보

대학부

19세 이하부

16세 이하부

13세 이하부

일반부

생활체육

스포츠클럽부

2022 ▾

지역전체 ▾

검색

	연도	팀명	팀구분	지역	선수수	감독
89	2022	TK8C(U19)	19세 이하부	경남	18	이승훈
88	2022	장흥고	19세 이하부	강원	18	최재호
87	2022	김천고	19세 이하부	강원	25	김철수
86	2022	계성고	19세 이하부	부산	36	최현규
85	2022	철거고	19세 이하부	서울	54	신현남
84	2022	영기상업고	19세 이하부	서울	53	최학현
83	2022	24기원고	19세 이하부	서울	41	최종호

1
kbo.sort_values('원화',ascending=False).head(20)

Pythor

	선수명	등번호	생년월일	포지션	신장/체중	경력	입단 계약금	연봉	지명순위	입단년도	Team	원화
494	구자욱	No.5	1993년 02월 12일	외야수(우투좌타)	189cm/75kg	본리초-경북중-대구고-삼성-상무	13000만원	200000만원	12 삼성 2라운드 12순위	12삼성	삼성 라이온즈	200000.0
804	채은성	No.22	1990년 02월 06일	외야수(우투우타)	186cm/92kg	순천북초-순천이수중-효천고-LG		180000만원	09 LG 육성선수	14LG	한화 이글스	180000.0
74	추신수	No.17	1982년 07월 13일	외야수(좌투좌타)	180cm/97kg	부산수영초-부산중-부산고-SK		170000만원	07 해외진출선수 특별지명	00시애틀	SSG 랜더스	170000.0
595	박세웅	No.21	1995년 11월 30일	투수(우투우타)	182cm/85kg	대구경운초-경운중-경북고-KT	20000만원	150000만원	14 KT 1차	14KT	롯데 자이언츠	150000.0

타자 선수들의 연봉 확인하기

1
2022년 기준 타자 연봉 TOP10
2
taza_sal.sort_values('연봉(원화)_2022',ascending = False).

]

	Key	연봉(원화)_2022
43	SSG_추신수	270000.0
39	SSG_한유섬	240000.0
4	KIA_나성범	200000.0
2	NC_박건우	190000.0
26	NC_손아섭	150000.0
47	두산_김재환	150000.0
20	LG_김현수	150000.0
17	두산_허경민	120000.0
37	SSG_최정	120000.0
23	NC_양의지	100000.0

투수 선수들의 연봉 확인하기

1
2022년 기준 투수 연봉 TOP10
2
tusu_sal.sort_values('연봉(원화)_2022',ascending = False).

]

	Key	연봉(원화)_2022
1	SSG_김광현	810000.0
7	NC_루친스키	187147.0
6	SSG_폰트	128663.0
8	삼성_류캐년	128663.0
4	LG_켈리	105270.0
5	키움_요키시	105270.0
15	KIA_양현종	100000.0
21	KT_데스파이네	81876.0
3	삼성_수아레즈	81876.0
2	LG_플렛코	58483.0

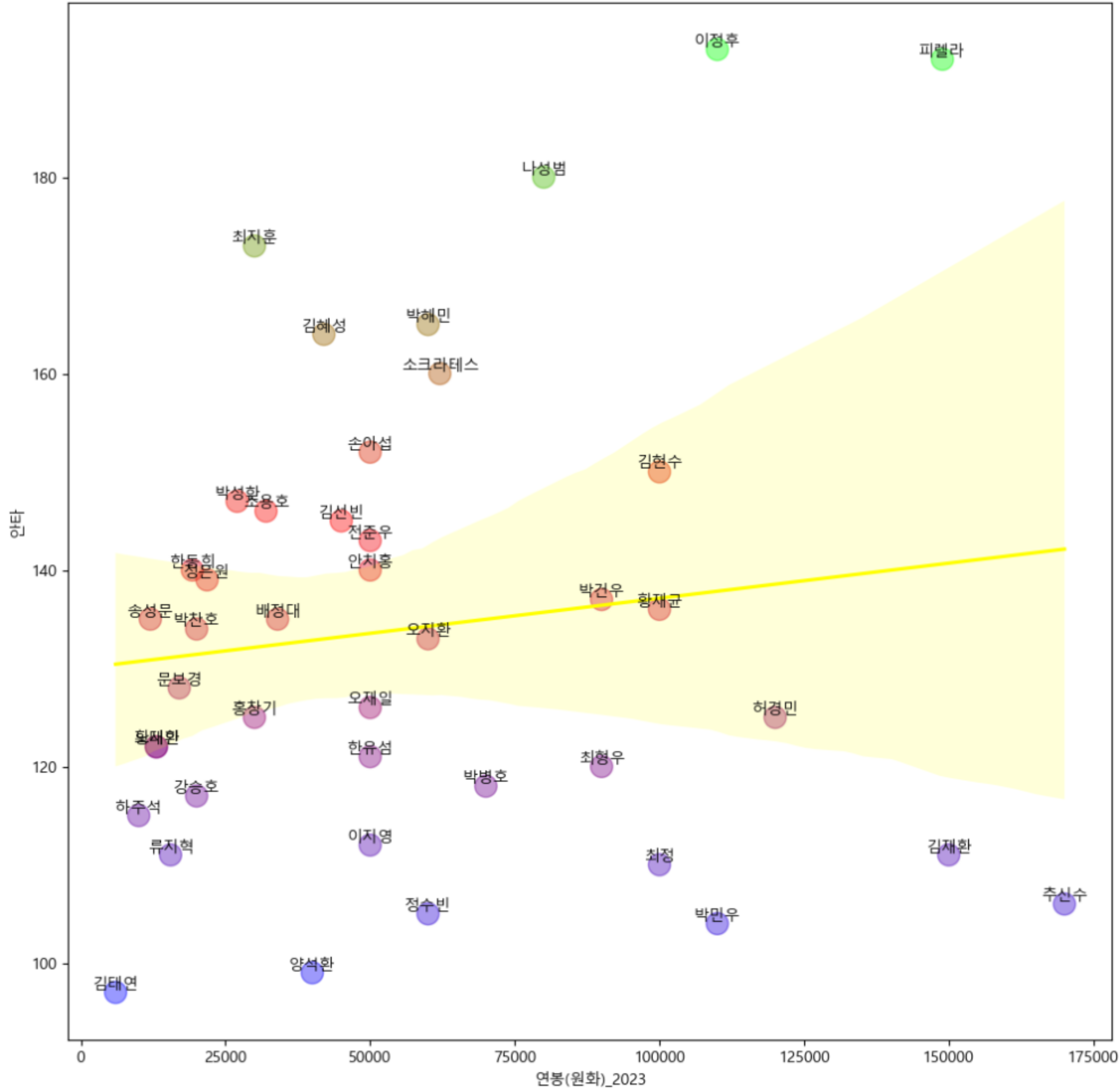
02

DATA 시각화 및 상관관계 비교

22, 23 KBO DATA



선수별 안타와 연봉(2023)의 산점도



* 노란색 선 : 오차가 최소가 되는점

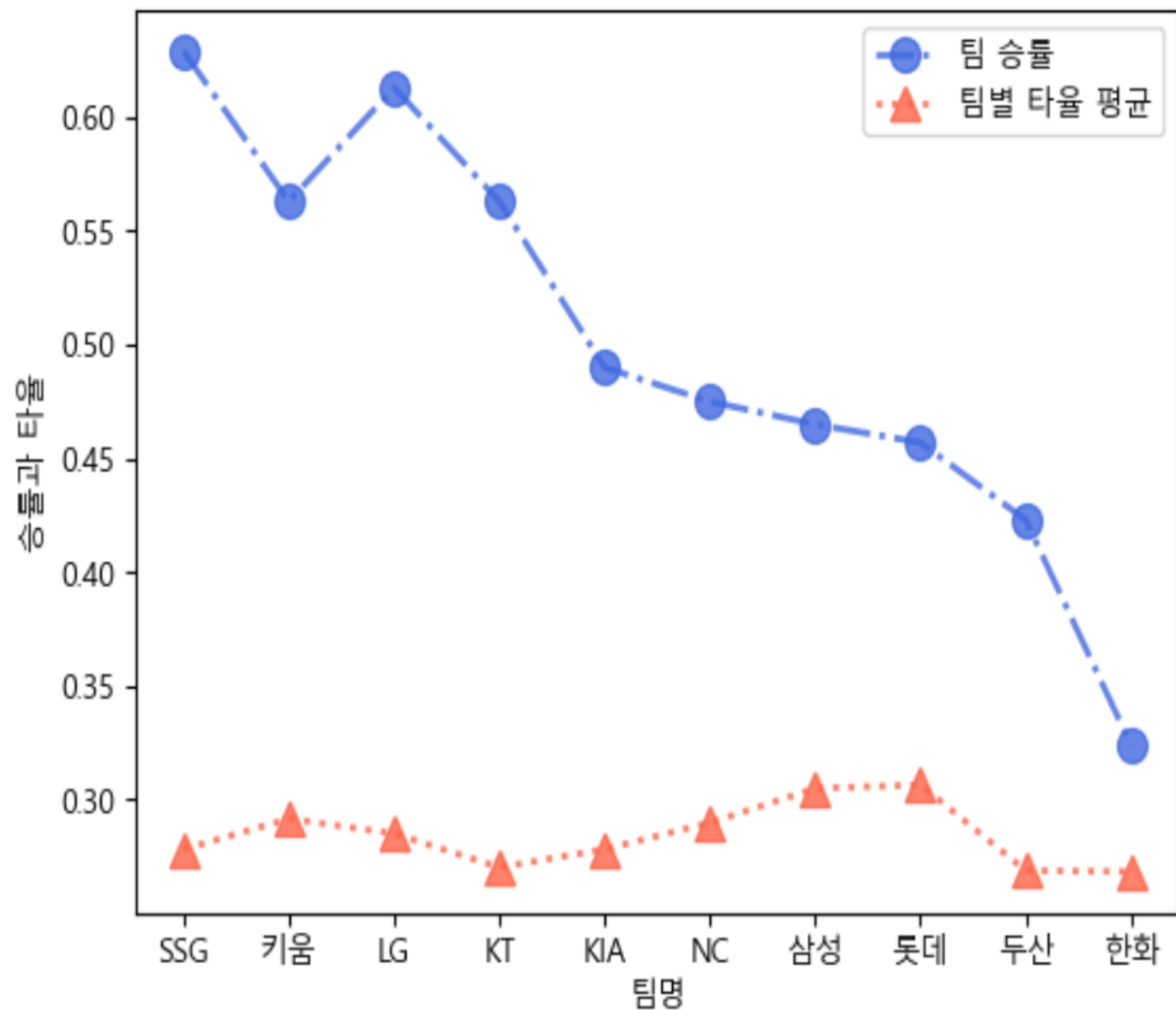
선수별 안타와 연봉의 산점도

PearsonRResult(statistic=0.1254189105354499
pvalue=0.4287045519539125)
-> p값 0.05 초과 -> 귀무가설 기각 불가

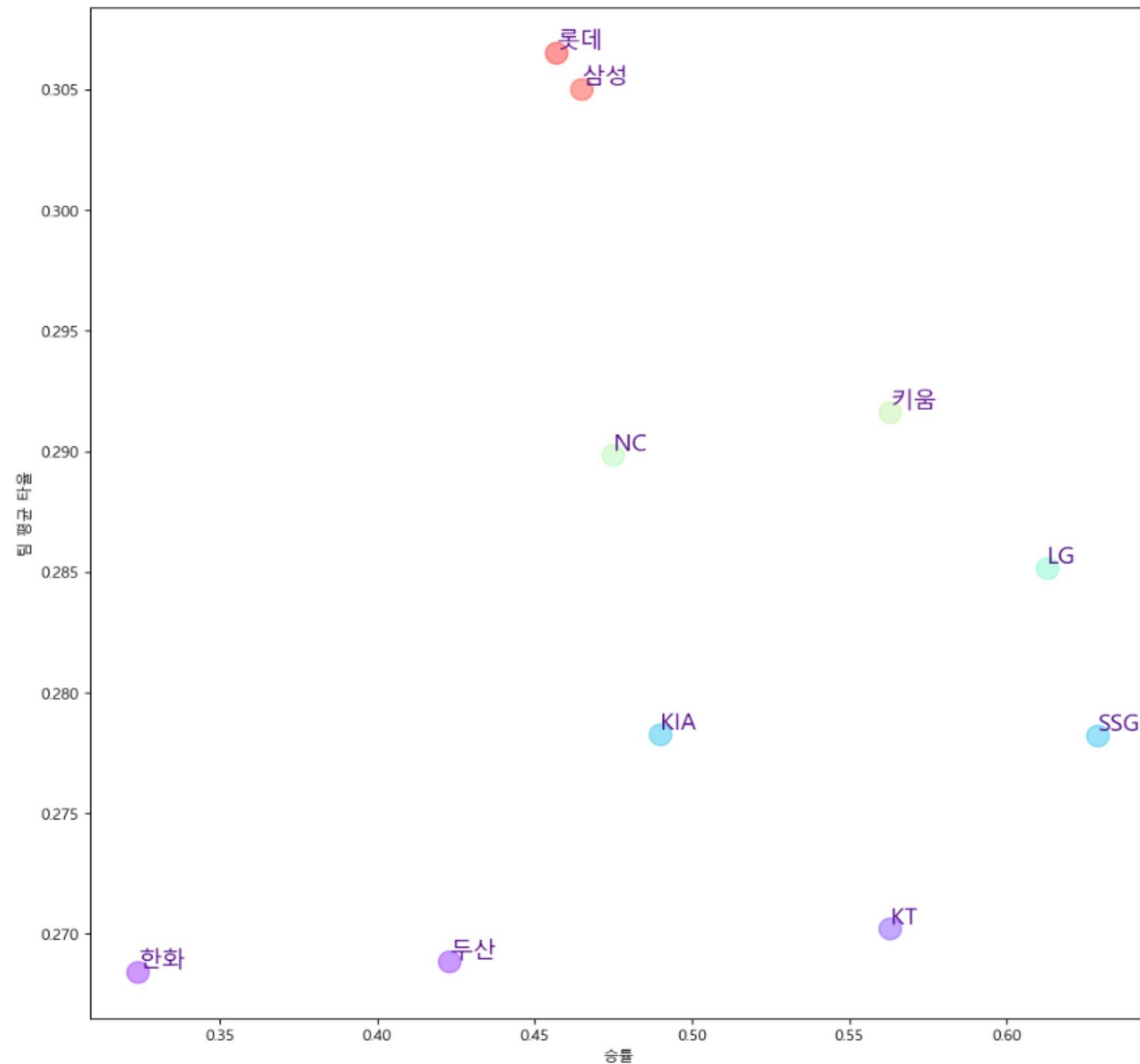


팀승률과 타자의 성적

- 팀 승률과 팀별 타자 타율의 평균



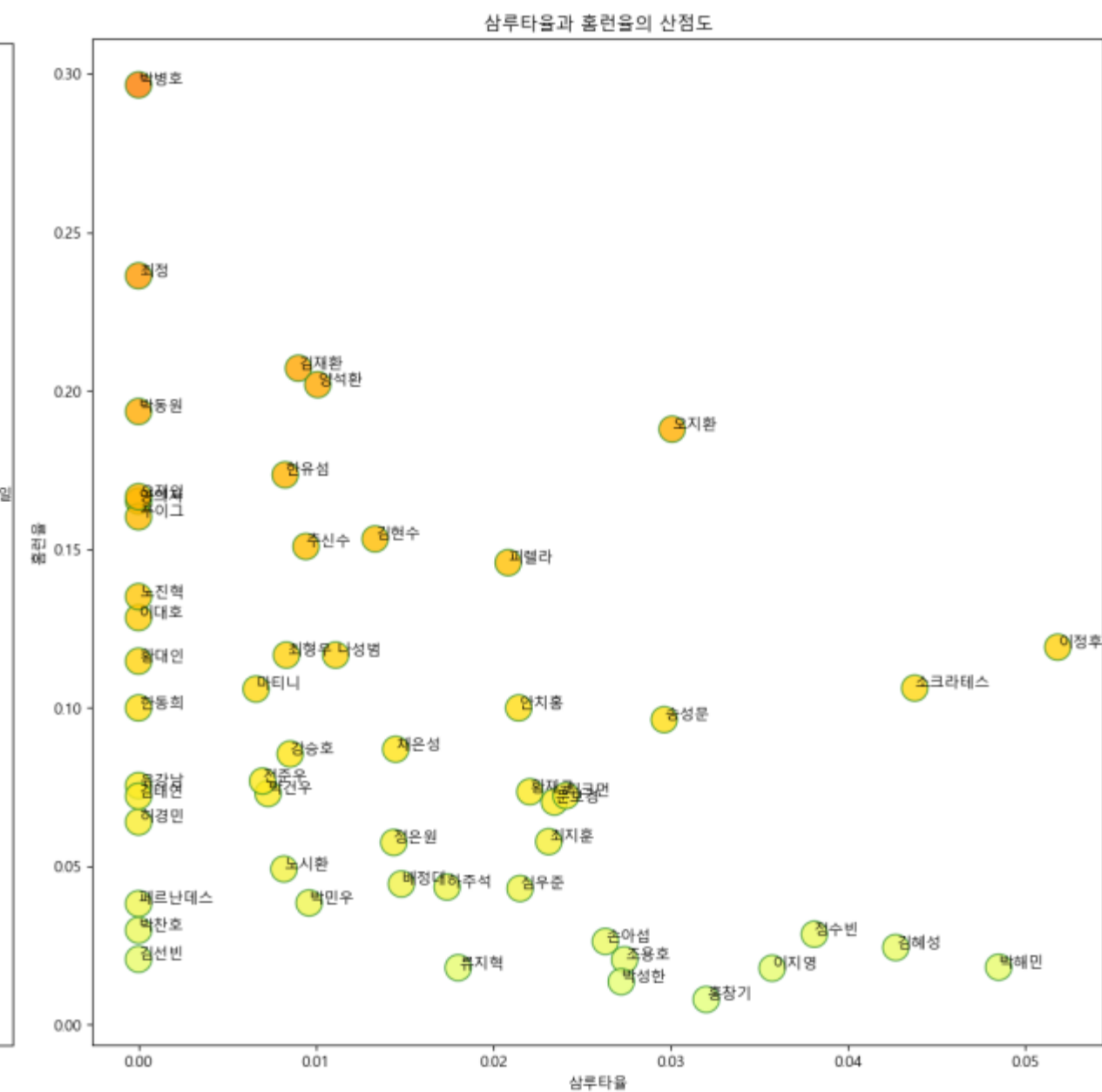
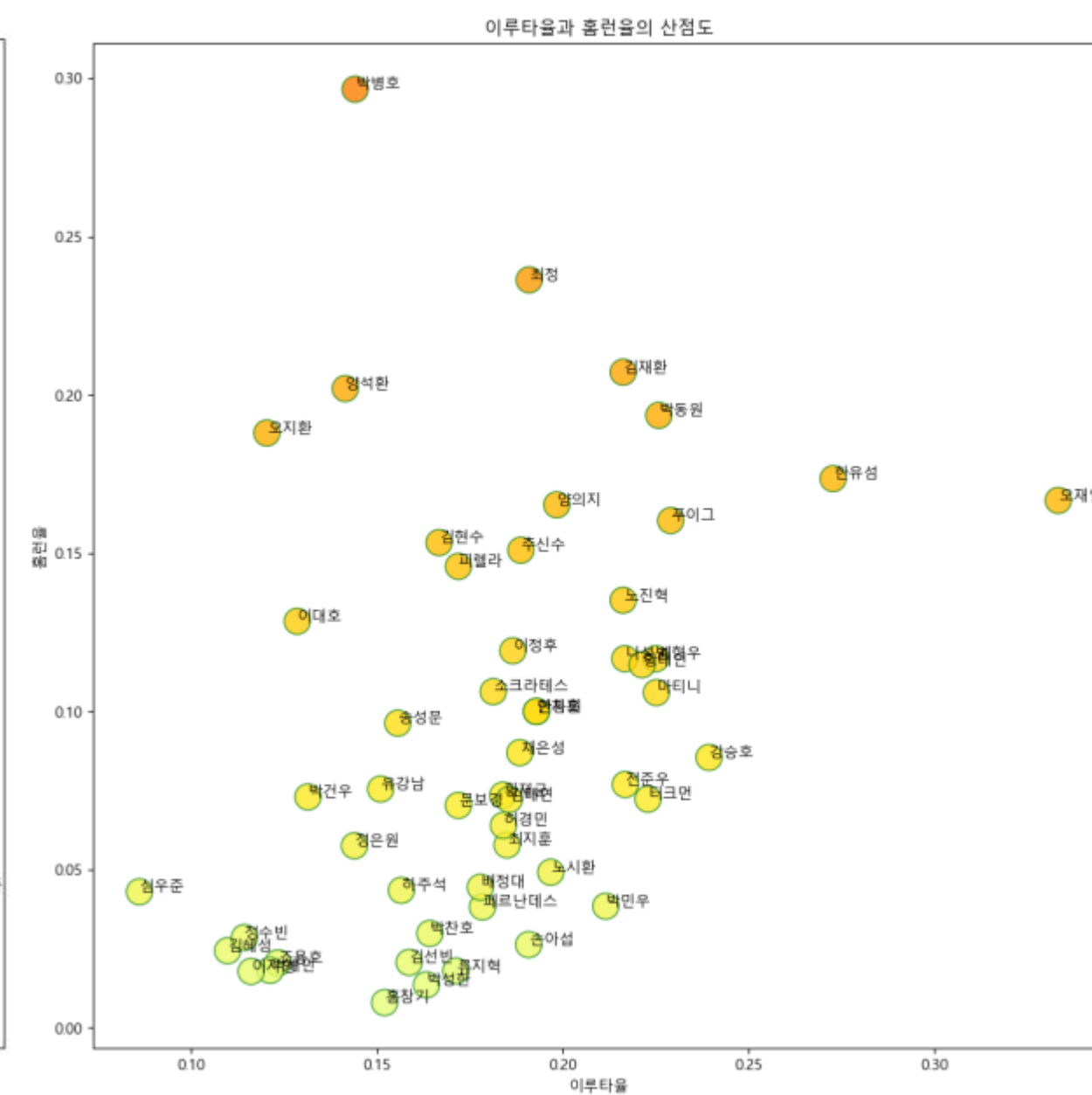
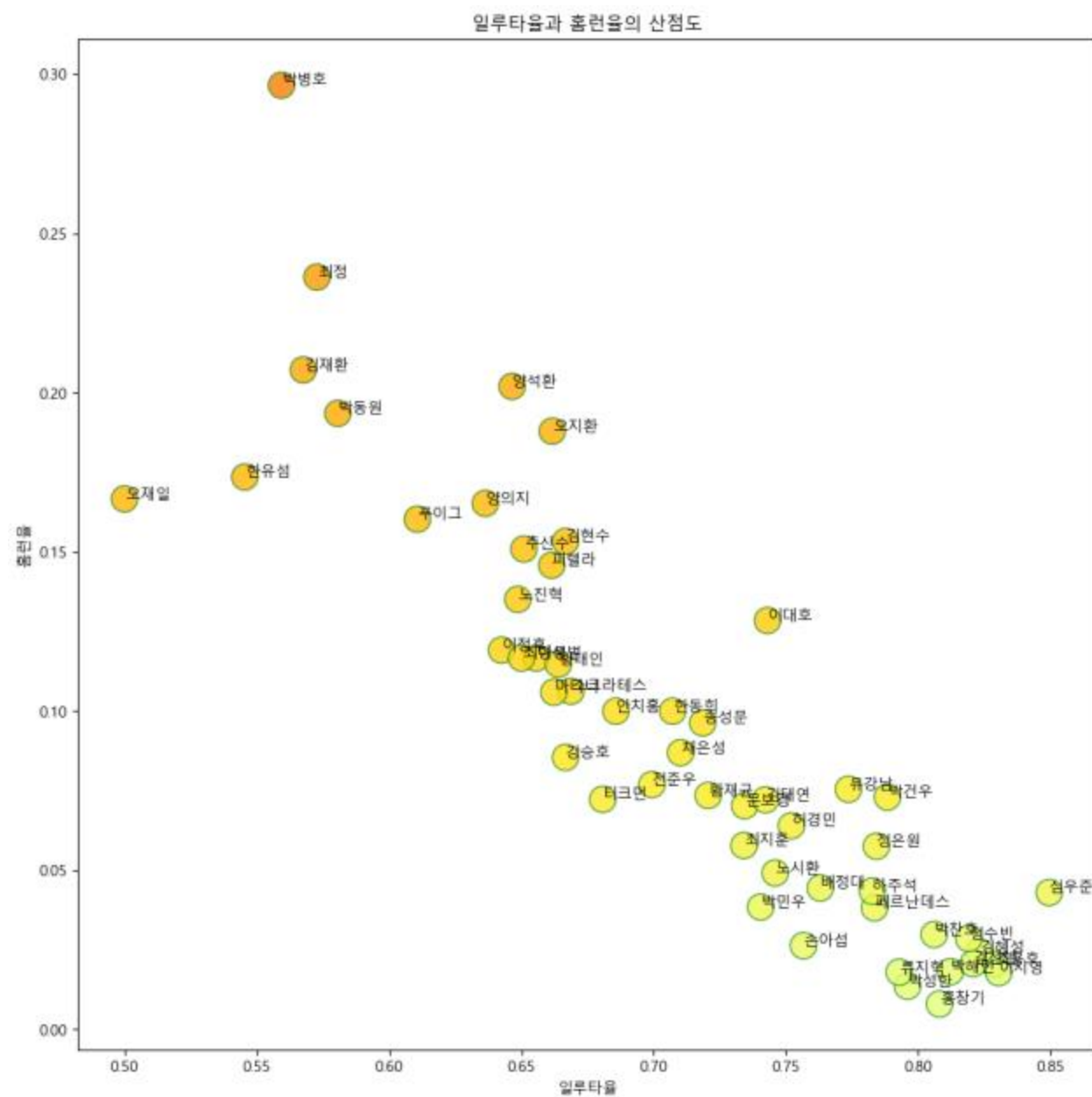
- 팀 승률과 팀평균 타율의 산점도



PearsonRResult(statistic=0.09464163633867723, pvalue=0.7948158391290246) -> p값 0.05 초과 -> 귀무가설 기각 불가



타자의 홈런과 안타 비교



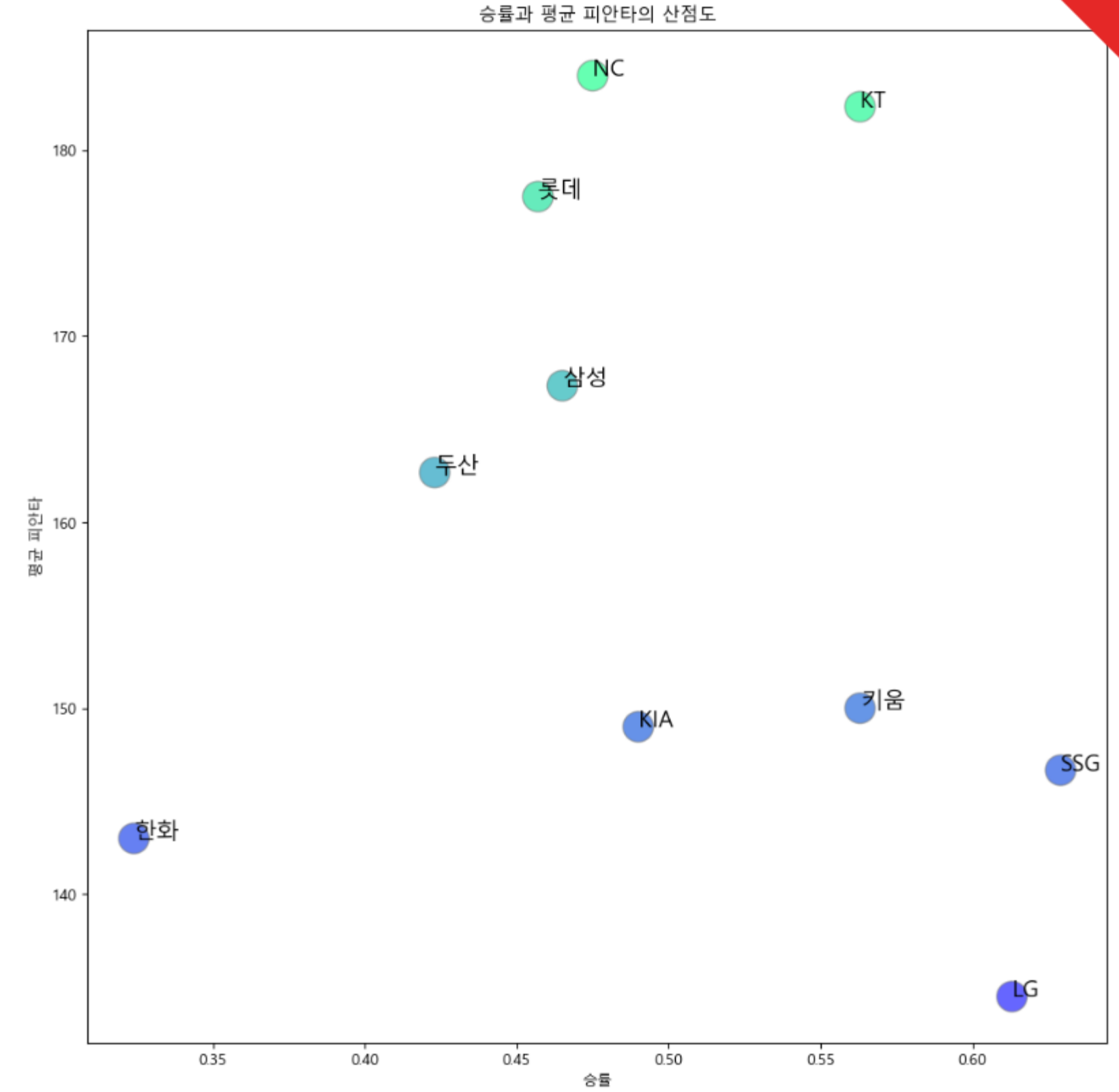
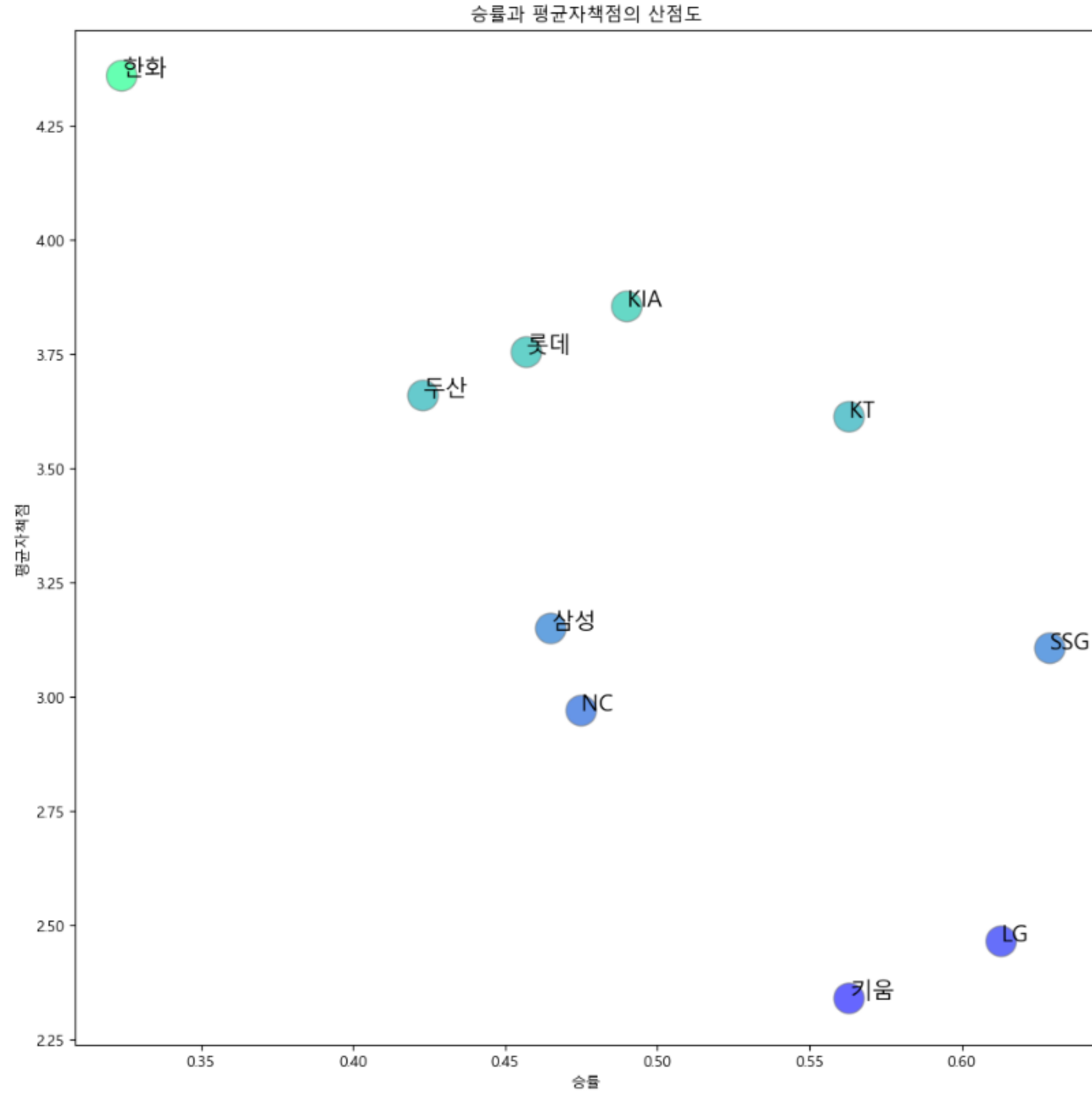
PearsonRResult(statistic=-0.8890524919338495, pvalue=6.161290642972242e-19)

PearsonRResult(statistic=0.34536418735058866, pvalue=0.011315618701714693)

PearsonRResult(statistic=-0.38969991151132444, pvalue=0.003921677353487378)



투수성적과 팀 승률 상관관계



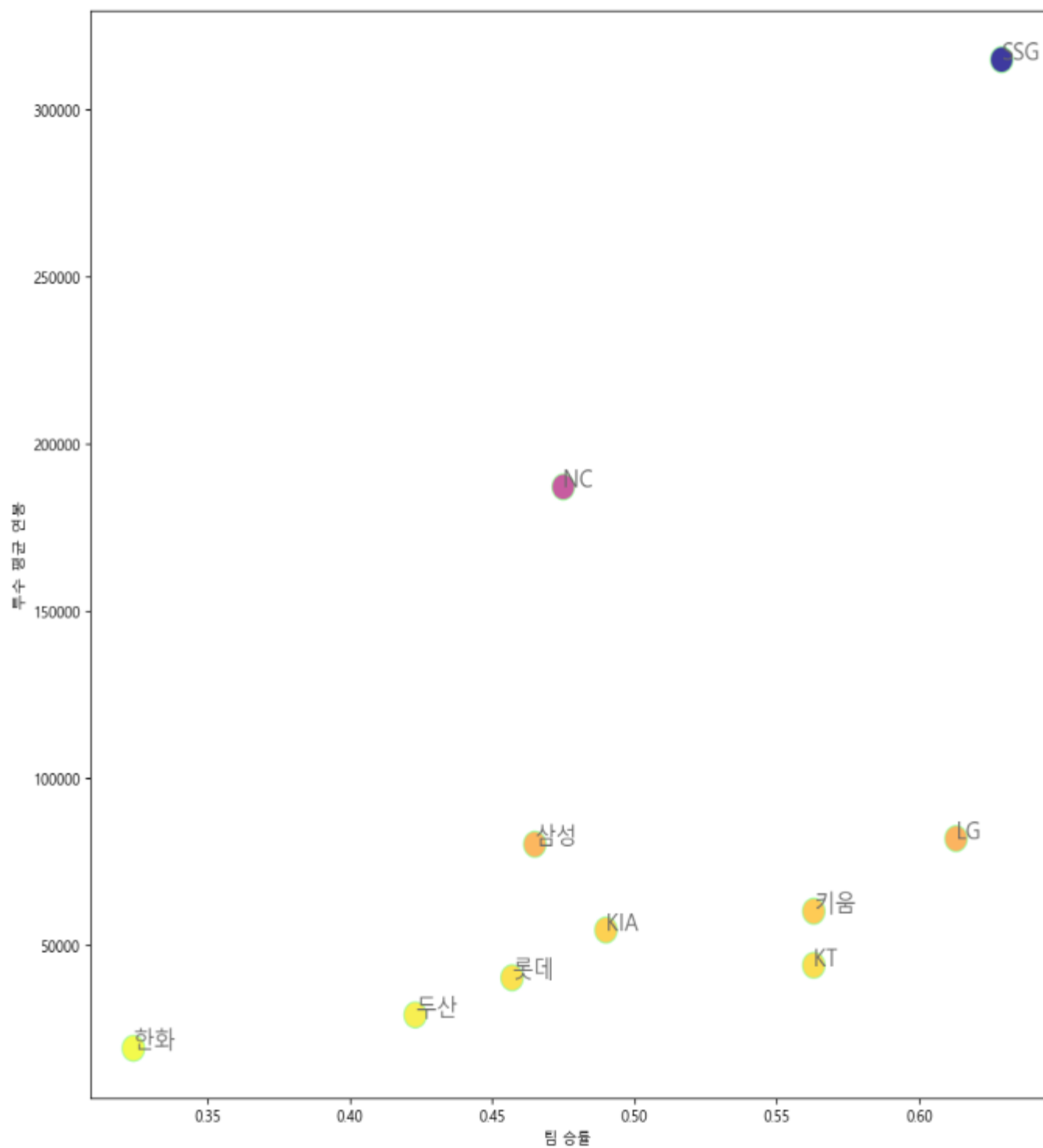
PearsonRResult(statistic=-0.7235079637962233, pvalue=0.01802397861071952)

PearsonRResult(statistic=-0.17221821883153499, pvalue=0.634248614402186)

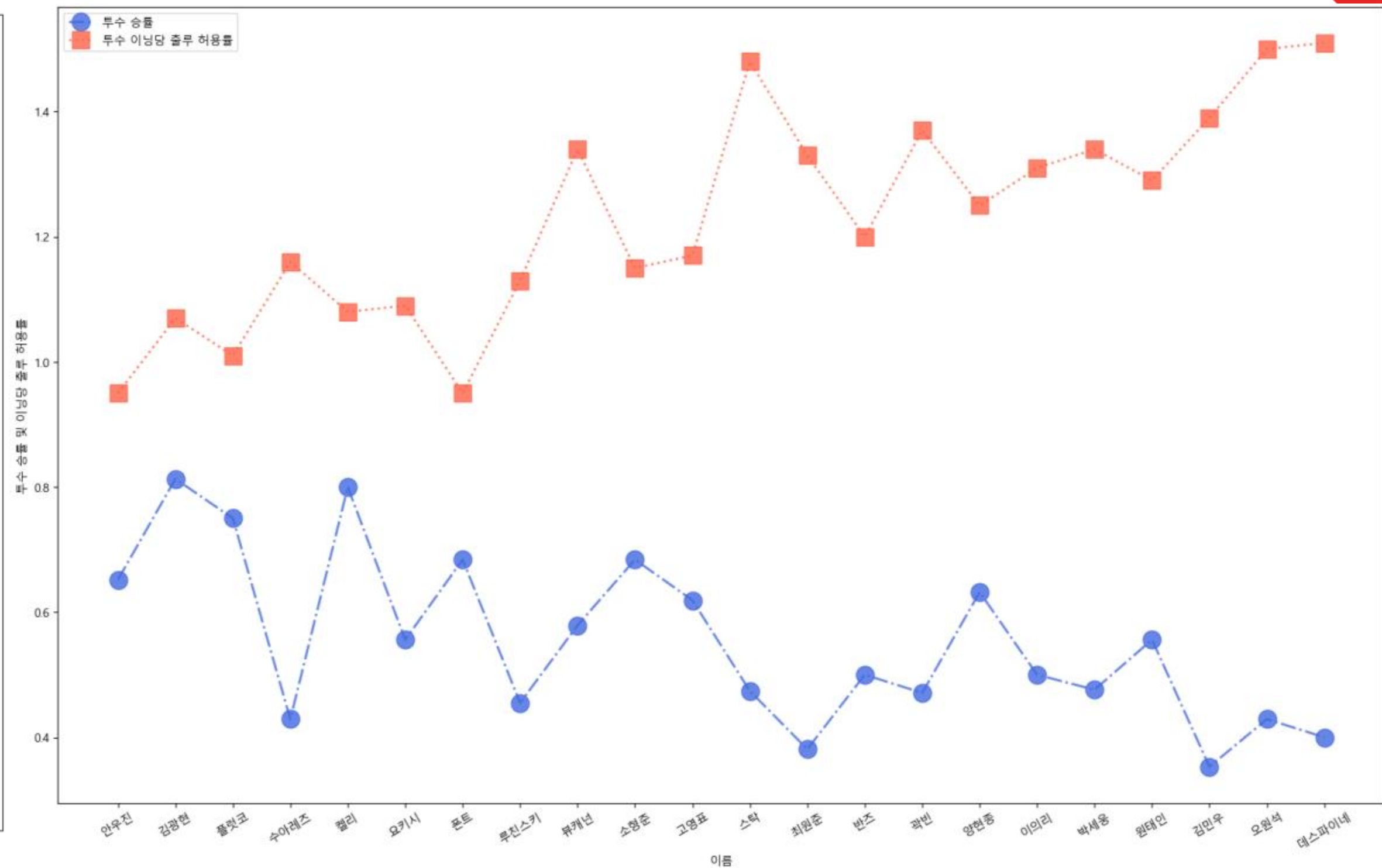


투수의 성적과 연봉의 상관관계 비교

- 팀 승률과 투수 평균 연봉의 산점도



-투수 승률과 이닝당 출루 허용률



타자평균연봉과 투수평균연봉 비교 (T TEST)

```
# 타자 평균 연봉과 투수 평균 연봉의 등분산 검정
from scipy.stats import bartlett
bartlett(taza_salary, tusu_salary)
# p값이 매우 작으므로 귀무가설을 기각한다. 즉, 등분산이 아니다.
```

```
BartlettResult(statistic=33.51151630715792, pvalue=7.084328265365405e-09)
```

```
# 타자 평균 연봉과 투수 평균 연봉의 t-test
# 귀무가설 : 타자 평균 연봉과 투수 평균 연봉은 같다.
# 대립가설 : 타자 평균 연봉과 투수 평균 연봉은 같지 않다.
import scipy.stats

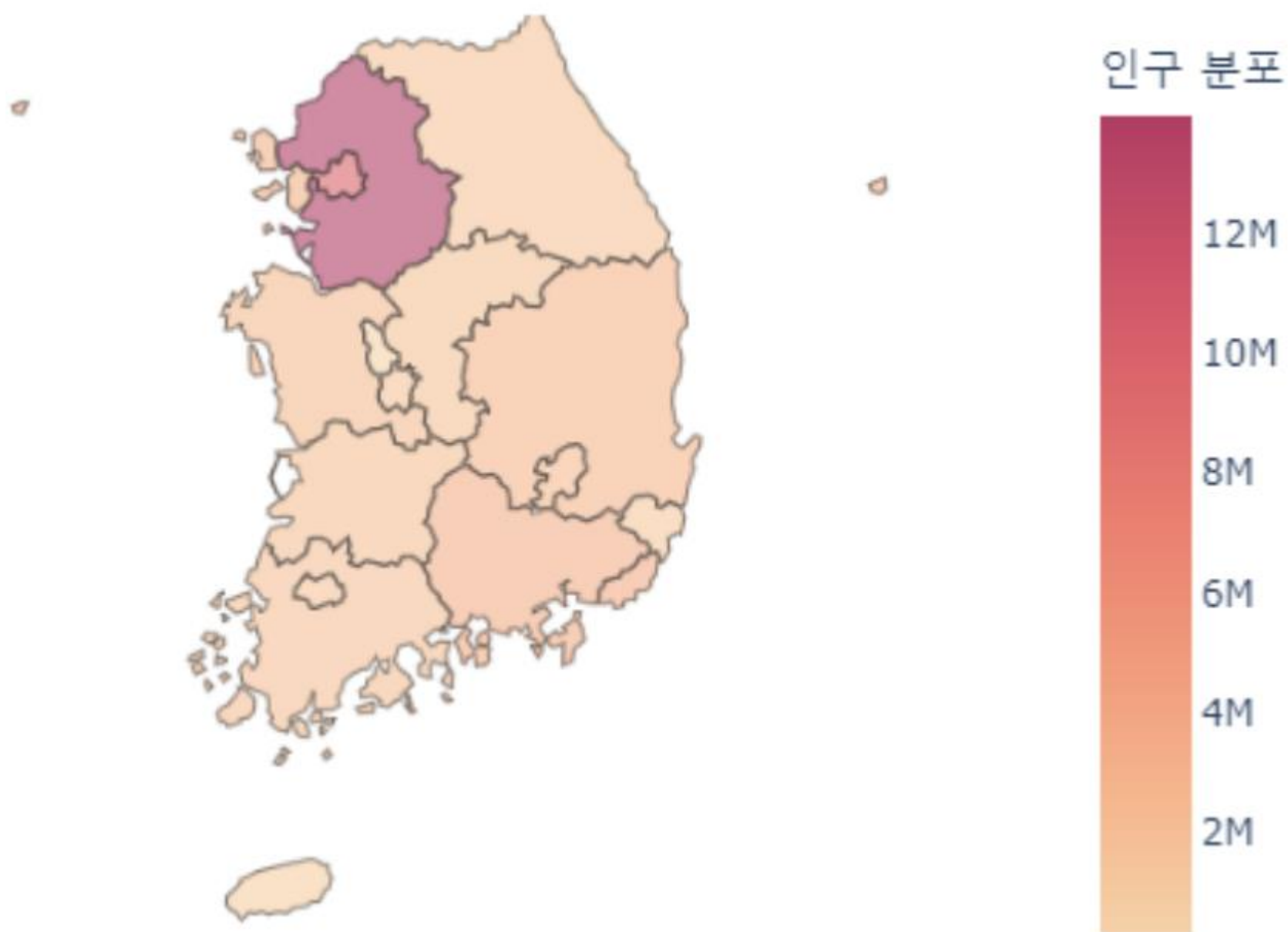
scipy.stats.ttest_ind(taza_salary, tusu_salary, equal_var=False)
# p값이 0.05보다 크므로 귀무가설을 기각할 수 없다. 즉 차이가 없다라고 할 수 있다.
```

```
Ttest_indResult(statistic=-0.8183109786758093, pvalue=0.4214297255981392)
```

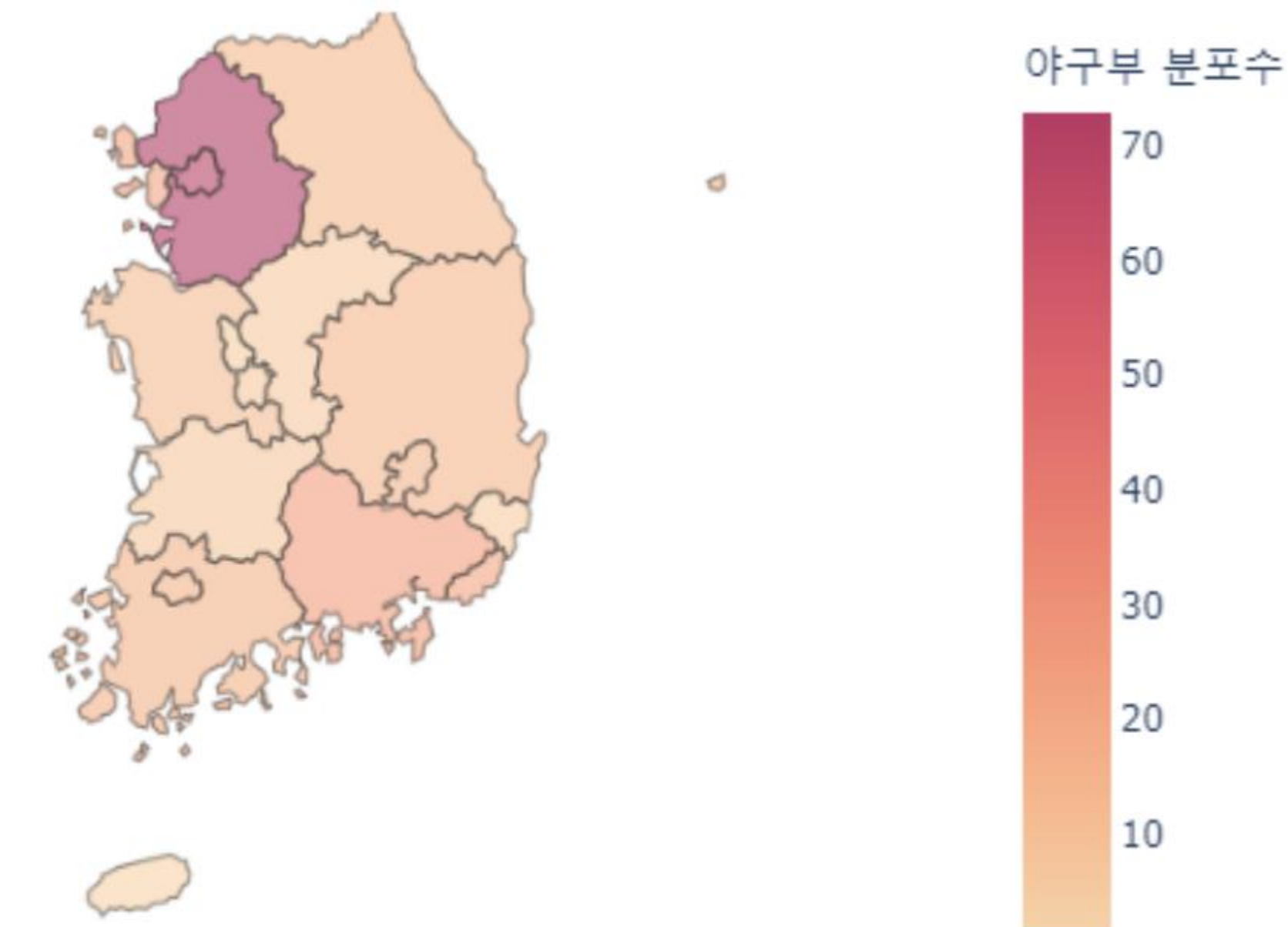


인구분포 / 전국 야구부 분포

인구분포



전국 야구부 분포





03

INSIGHT

22,23 KBO DATA

Limit Point



DATA의 부족



고려요인 다수



시간부족





THANK YOU

22, 23 KBO DATA