

The slide features a light gray background with decorative elements. In the top-left corner, there is a black triangle containing a smaller cyan triangle. In the bottom-right corner, there is a large cyan triangle. The main title is centered in a bold, black, sans-serif font.

2019 야구공작소 테이터 스터디 - 1주차 -

2019.09.xx

일정 안내

- 9월 3주차부터 시작
- 월요일 / 목요일 동일 회차
 - 같은 내용 ▶ 하루만!
 - 재수강(?) 자유
- 일정 연장 가능
 - 대학교 시험 기간...
 - 포스트시즌(?)

일	월	화	수	목	금	토
9월						
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					
10월						
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

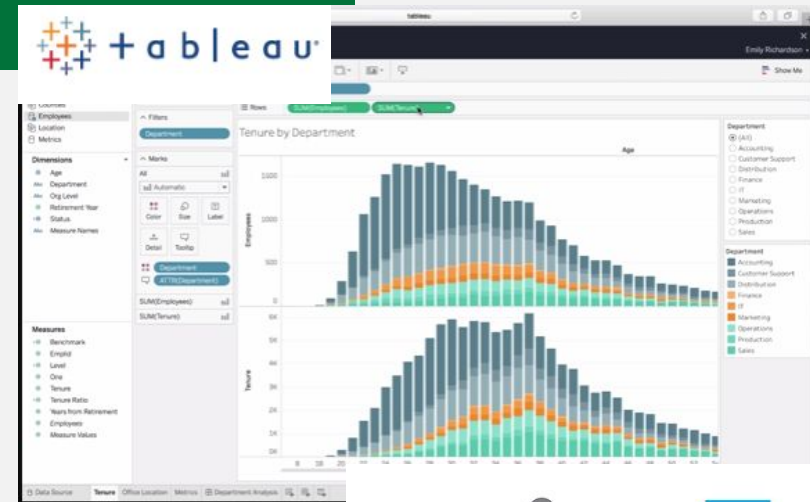
추석

과제!

- 매주 간단한 자습용 문제 내드림.
- 토론, 논의, 공유 자유롭게 해봅시다.
- 희망사항 : 마지막엔 프로젝트

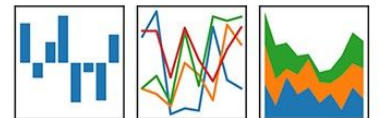
학습 목표

- 야구 데이터 뽐내기!
- 엑셀로 뽐내기
- 태블로로 뽐내기
- 코딩(프로그래밍)으로 뽐내기
- 데이터 수집하기



pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



야구 데이터 어디서 볼까

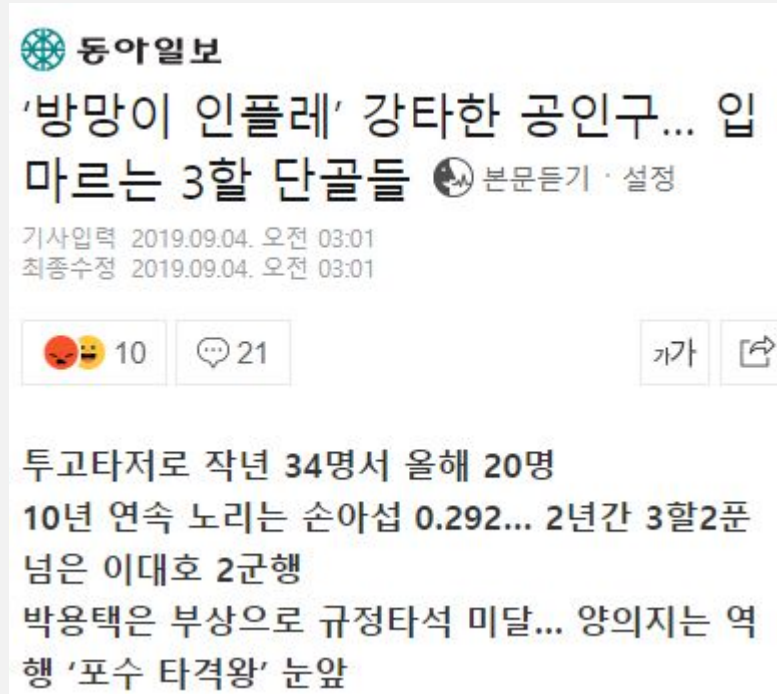
- KBO
 - 스탯티즈 (<http://www.statiz.co.kr>)
 - KBO 공식 홈페이지 (<http://www.koreabaseball.com>)
 - 레전드닷컴 (<http://www.legend2i.com>)
- 메이저리그
 - MLB.com (<https://mlb.com>)
 - 베이스볼 레퍼런스 (<https://bbref.com>)
 - 팬그래프 (<https://fangraphs.com>)

최고의 데이터 분석 도구 1순위 : 엑셀

- 엑셀이 없으면 [구글 스프레드시트](#)! (공짜)
- (비추천) [리브레 오피스\(Libre Office\)](#) (공짜2)
- (절대 비추천) 한컴 오피스

예제 - 공인구 변경의 여파를 알아보자(1)

- 3할 타자 지금은 몇 명?
- 스탯티즈 접속
 - 기록실 - 시즌기록실
- '타율'로 정렬
 - 규정타석 3할 이상은?



예제 - 공인구 변경의 여파를 알아보기(2)

- 작년 대비 홈런은 얼마나 줄어들었을까?
- 비교방법 - 예시: 100타석 중 2번 홈런
 - (1) 홈런/타석
 - ex) “홈런이 나온 타석은 전체의 2%”
 - (2) 타석/홈런
 - ex) “50타석에 1번 꼴로 홈런이 나왔다”
- 스탯티즈에서 확인 가능
- (2)는 직접 계산해보자

2019 연도 ▾ 18 ▾ ~19 ▾ 팀:전체 ▾ 포지션 ▾ 정규 ▾ 규정 ▾ 상황 ▾

[자동 : 전체]

기본 확장 가치 클러치 타석 타구1 타구2 파워

리그 기록

순	이름	팀	정렬 WAR*	G	타석	타수	득점	안타	2타	3타	홈런
1	리그	18	216.55	18225	56903	50547	7994	14445	2639	187	1756
2	리그	19	187.42	15350	49230	43476	5895	11663	2053	207	901

예제 - 공인구 변경의 여파를 알아보자(3)

- 프로야구 역사에서 올해는 투고타저? 타고투저?
- 1982년부터 2019년까지 다음 기록을 비교
 - 타율
 - 출루율
 - OPS
 - 타석/홈런
- 그래프를 그려본다면?

The screenshot shows a web-based search interface for baseball statistics. At the top, there is a row of filters: '2019' (year), '연도' (year), '82' (year), '~19' (year), '팀:전체' (team: all), '포지션' (position), '정규' (regular), '규정' (regulation), '상황' (situation), and '옵션' (option). Below this, the text '상세옵션' (detailed options) is visible. The main search area contains several dropdown menus and checkboxes: '나이: 17 ~ 51' (age), '검색' (search), '타석: 전체' (at-bats: all), '환산: 없음' (conversion: none), '표기: 리그' (notation: league), '출신고: 전체' (outfield: all), '출신대: 전체' (birth year: all), '전체' (all), '용병' (foreign player), '중간라인: 보기' (middle line: view), '제거' (remove), '조건1: 타율' (condition 1: batting average), '이상' (above), '조건2: 타점' (condition 2: at-bats), and '이상' (above). A blue '검색' (search) button is located at the bottom left.

엑셀 팁

- 스탯티즈에서 복사-붙여넣기 하면 시즌이 2자리 수만...
- if함수를 써서 값을 바꿉시다

리그 기록			
순	이름	팀	WAR*
1	리그	15	220.59
2	리그	16	219.73
3	리그	18	216.55
4	리그	17	215.86
5	리그	14	215.73
6	리그	13	214.78
7	리그	9	194.75
8	리그	0	194.38
9	리그	99	193.96
10	리그	1	193.83



순	이름	C3 2015 x	시즌	팀	WAR*
1	리그	=if(D3>19, 1900+D3, 2000+D3)			15 220.59
2	리그		2016	16	219.73
3	리그		2018	18	216.55
4	리그		2017	17	215.86
5	리그		2014	14	215.73
6	리그		2013	13	214.78
7	리그		2009	9	194.75
8	리그		2000	0	194.38
9	리그		1999	99	193.96
10	리그		2001	1	193.83



예제 - ERA와 승리의 상관 관계는?

- ‘야구는 투수 놀음’
 - 정말 그럴까?

- 연도별 팀 ERA와 승률 비교

- 대체로 낮을 수록 승률이 높긴 한데…?

시즌기록실

종합 타격 투구 **승비**

2019 연도 82 ~19 팀:전체 정규 규정

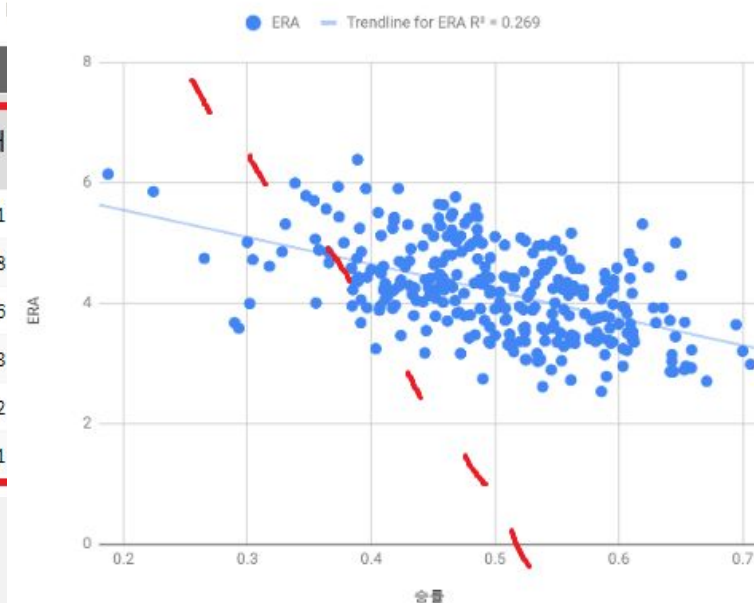
2013년 이전은 세부/상황별 기록 지원하지 않습니다.

기본 확장 방어율 가치 WP 타석

팀 기록

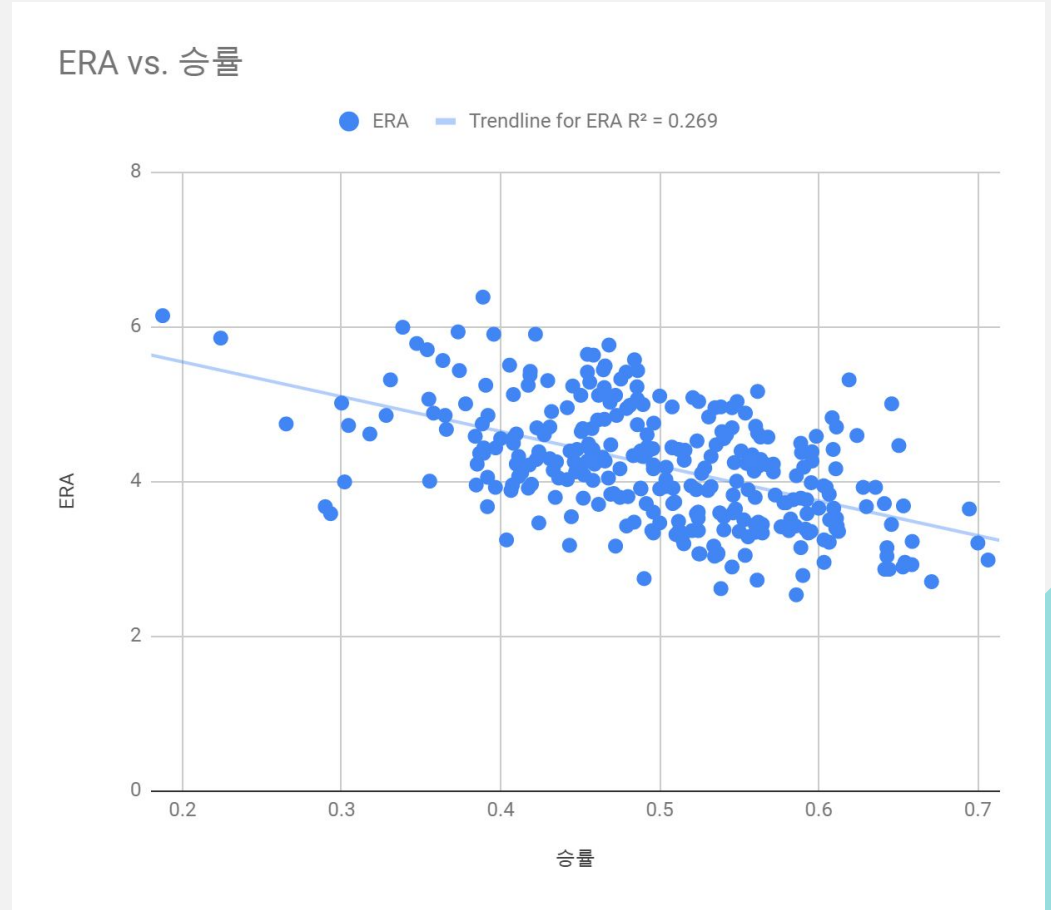
순	이름	팀	비율 ERA	출장	완투	완봉	선발	승	패
1	OB	84	2.53	229	14	2	100	58	41
2	OB	86	2.61	238	29	12	108	56	48
3	삼성	82	2.70	152	23	8	79	53	26
4	MBC	83	2.72	190	31	6	100	55	43
5	롯데	86	2.74	234	33	6	108	50	52
6	MBC	86	2.78	267	11	3	108	59	41

ERA vs. 승률



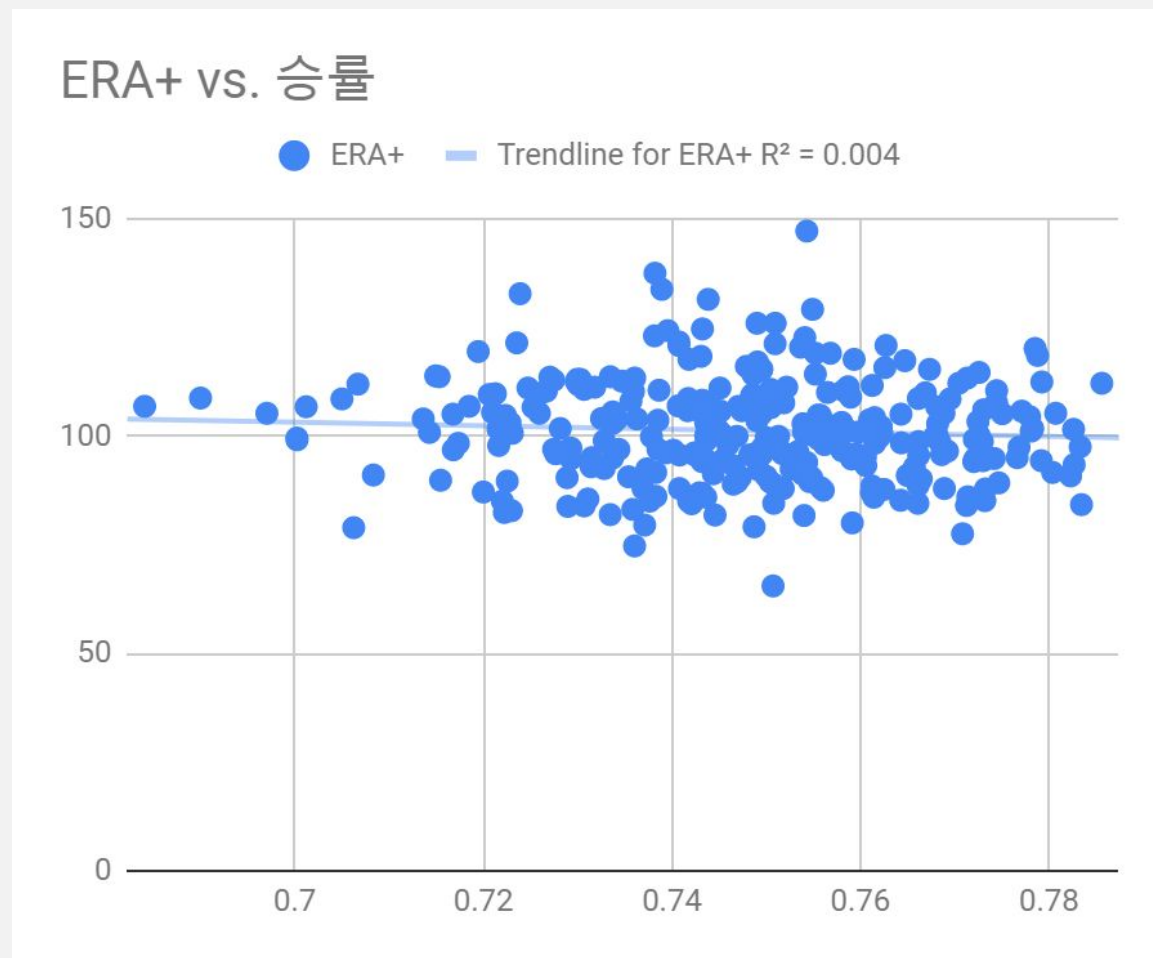
상관 계수 R, 결정 계수 R^2

- 1에 가까울 수록 ‘상관도 높음’
 - $-1 \leq R \leq +1$
 - $0 \leq R^2 \leq 1$
- R = 추세선(Trend Line)의 기울기
 - 정비례(\nearrow) = 양의 상관관계
 - 반비례(\searrow) = 음의 상관관계
- 보통 $R^2 \geq 0.5$ 일 때 ‘상관도가 높음’
 - ($R \geq 0.7$ 혹은 $R \leq -0.7$)
 - ‘A로 B의 50%를 설명할 수 있다’

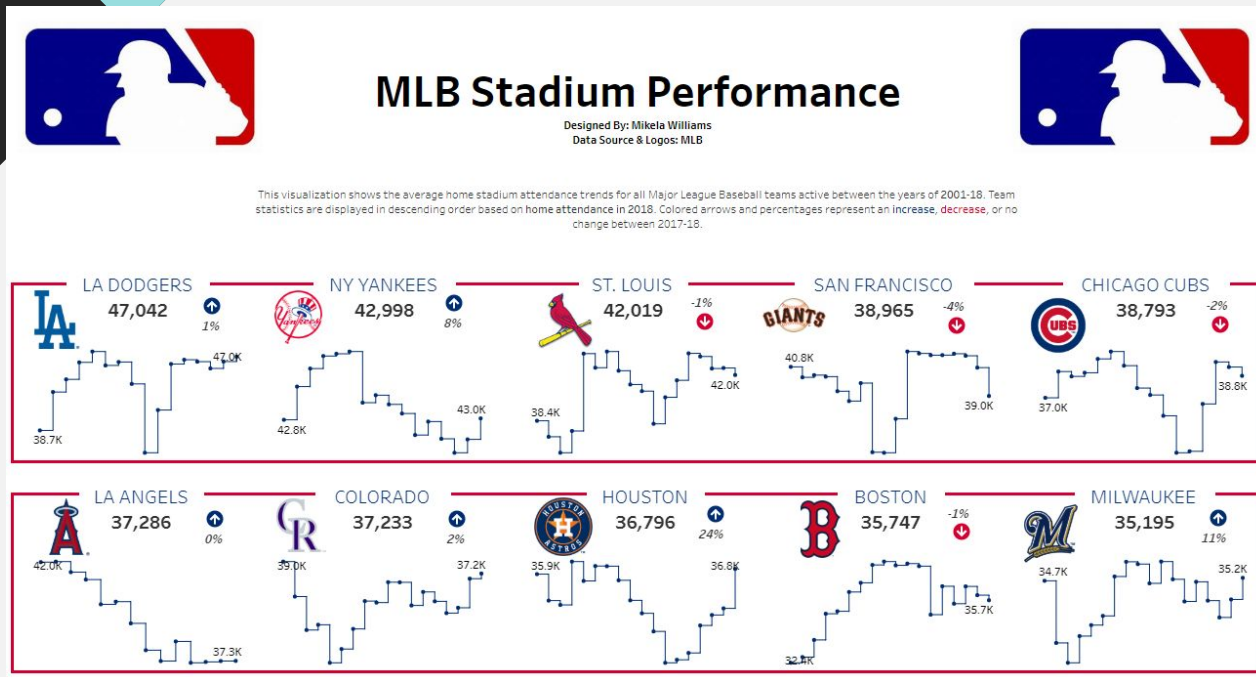


예제 - 조정 방어율로 하면?

- 류현진이 정점을 찍던 그 ERA+...
 - 100 : 리그 평균
 - >100 : 평균 이상
 - <100 : 평균 이하
- 투고타저/타고투저 변화 반영
- 결론은?



Tableau



<https://public.tableau.com/ko-kr/gallery/mlb-stadium-performance-0>

애슬릿 미디어에서도 씁니다...
SK도... KT도...

How the **NBA** has transitioned from 2pt shooting to 3pt shooting. Is old-school basketball being replaced by long distance shots?



<https://public.tableau.com/ko-kr/gallery/rise-nba-3-pointer>

Tableau 실습

- Fangraphs 접속
- 1960년~2019년 팀 ERA-승률 비교
- Tableau!
 - 다운로드 : <https://public.tableau.com/ko-kr/s/>

כל