

야구에서 'CrHome'과 'CrRuns'의 관계

2019003954 남유진

1. 상관관계, 인과관계 예측

두 변수는 강한 양의 상관관계를 띌 것이다.

CrRuns 독립변수는 CrHome 종속변수에 유의미한 영향을 줄 것이다.

2. 상관분석과 회귀분석 코드

```
1 PROC CORR DATA=SASHELP.BASEBALL PEARSON SPEARMAN PLOTS=MATRIX;
2   VAR CrHome CrRuns;
3
4
5 PROC REG DATA=SASHELP.BASEBALL;
6   MODEL CrHome=CrRuns;
7
8 RUN;
```

3. 상관분석과 회귀분석 출력 결과

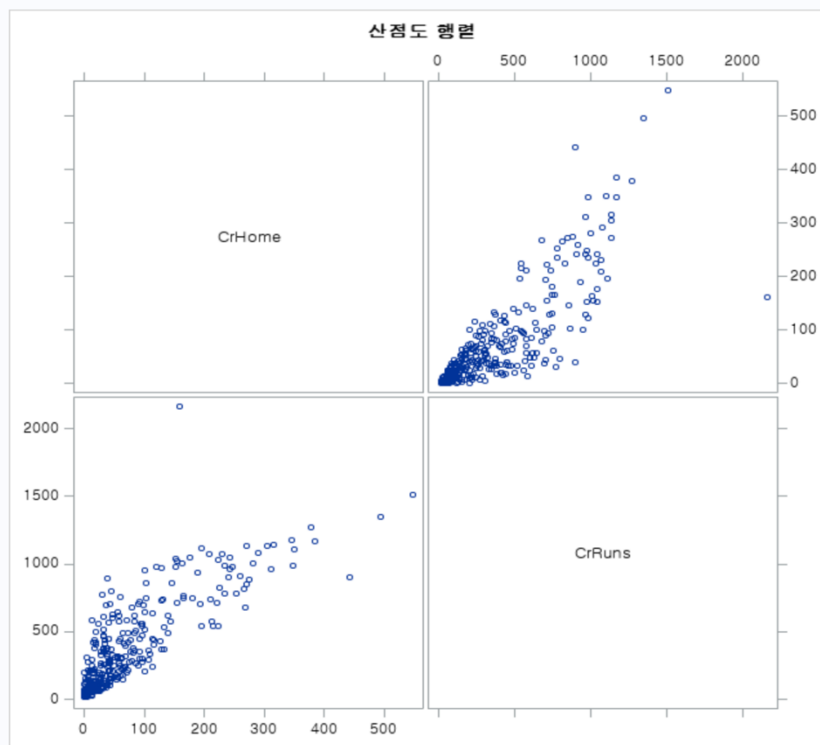
CORR 프로시저

2 개의 변수: CrHome CrRuns

단순 통계량							
변수	N	평균	표준편차	중위수	최솟값	최댓값	레이블
CrHome	322	74.09006	90.06513	40.00000	0	548.00000	Career Home Runs
CrRuns	322	374.28571	336.42504	266.00000	18.00000	2165	Career Runs

피어슨 상관 계수, N = 322 H0: Rho=0 가정하에서 Prob > r		
	CrHome	CrRuns
CrHome Career Home Runs	1.00000	0.82093 <.0001
CrRuns Career Runs	0.82093 <.0001	1.00000

스피어만 상관 계수, N = 322 H0: Rho=0 가정하에서 Prob > r		
	CrHome	CrRuns
CrHome Career Home Runs	1.00000	0.84076 <.0001
CrRuns Career Runs	0.84076 <.0001	1.00000



The REG Procedure
Model: MODEL1
Dependent Variable: CrHome Career Home Runs

Number of Observations Read	322
Number of Observations Used	322

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1754812	1754812	661.37	<.0001
Error	320	849053	2653.28966		
Corrected Total	321	2603864			

Root MSE	51.51009	R-Square	0.6739
Dependent Mean	74.09006	Adj R-Sq	0.6729
Coeff Var	69.52362		

Parameter Estimates					
Variable	Label	DF	Parameter Estimate	Standard Error	t Value Pr > t
Intercept	Intercept	1	-8.16785	4.29777	-1.90 0.0583
CrRuns	Career Runs	1	0.21977	0.00855	25.72 <.0001

4. 상관분석 결과내용

Pearson 상관계수=0.82093 / Spearman 상관계수=0.84076

상관관계가 모두 양수이며 1에 가까우므로 강한 양의 상관관계를 가진다.

5. 회귀분석 결과내용

유의확률이 0.0001보다 작으므로 귀무가설 기각. 독립변수로 종속변수를 유의미하게 설명할 수 있다. 결정계수가 0.6739이므로 현재 회귀선모델은 67%만큼 데이터를 잘 설명하고 있다.