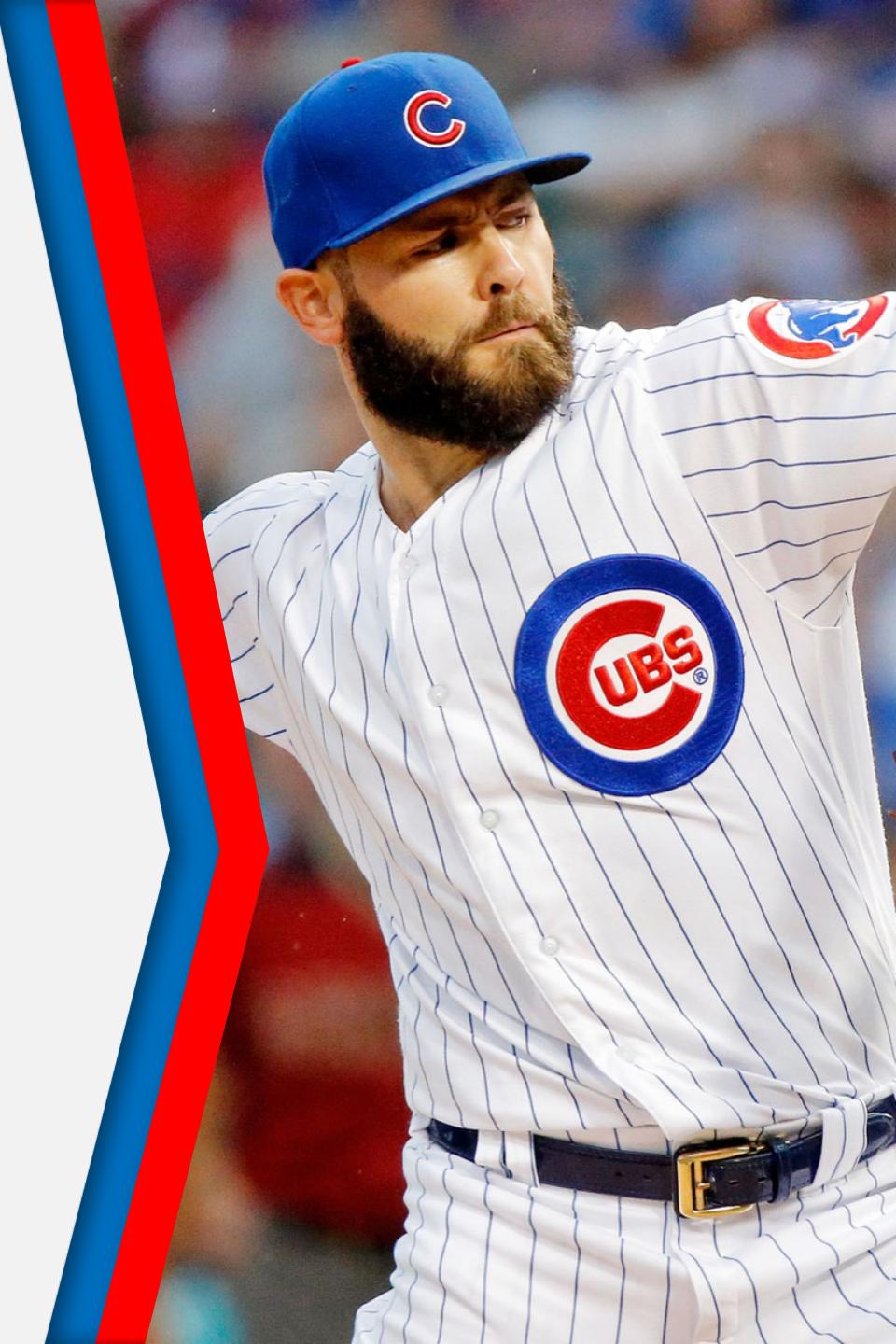


Predicting Starting Pitcher Salary

Chris Gillespie



Objective

Recommend Annual Salary for Free Agent Starting Pitcher to Chicago Cubs



Ballpark Strategy

Data

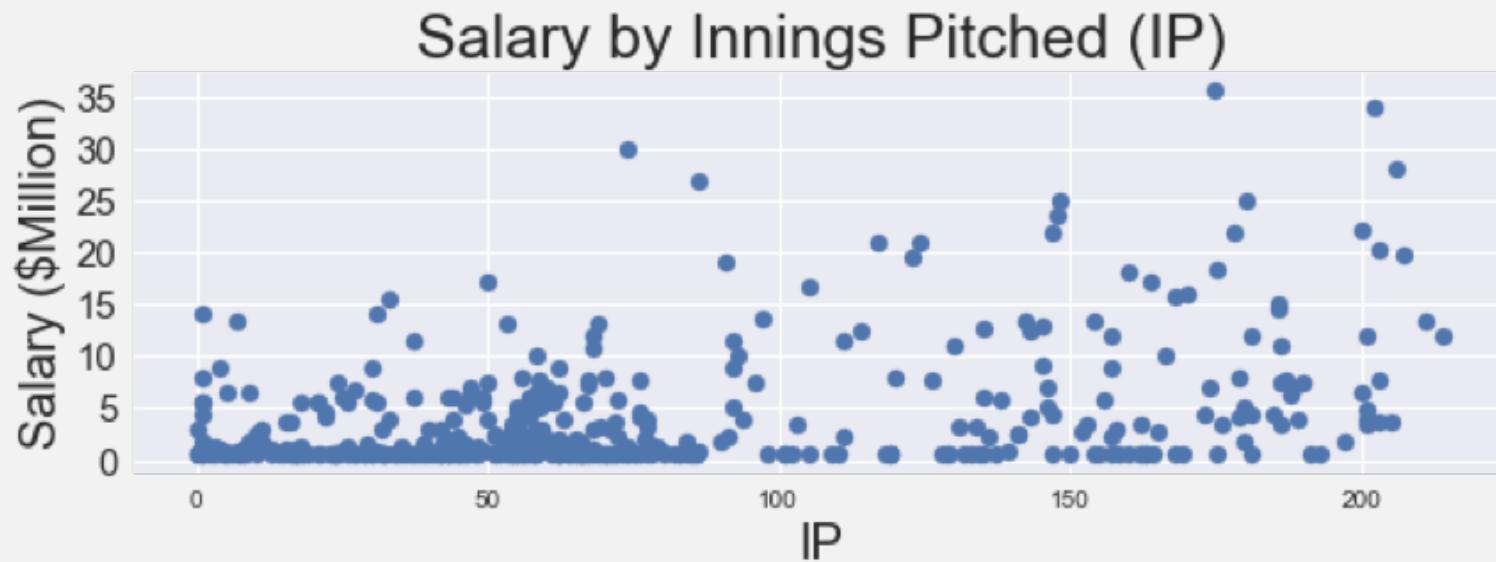


2017

Focusing on Starting Pitchers

Data

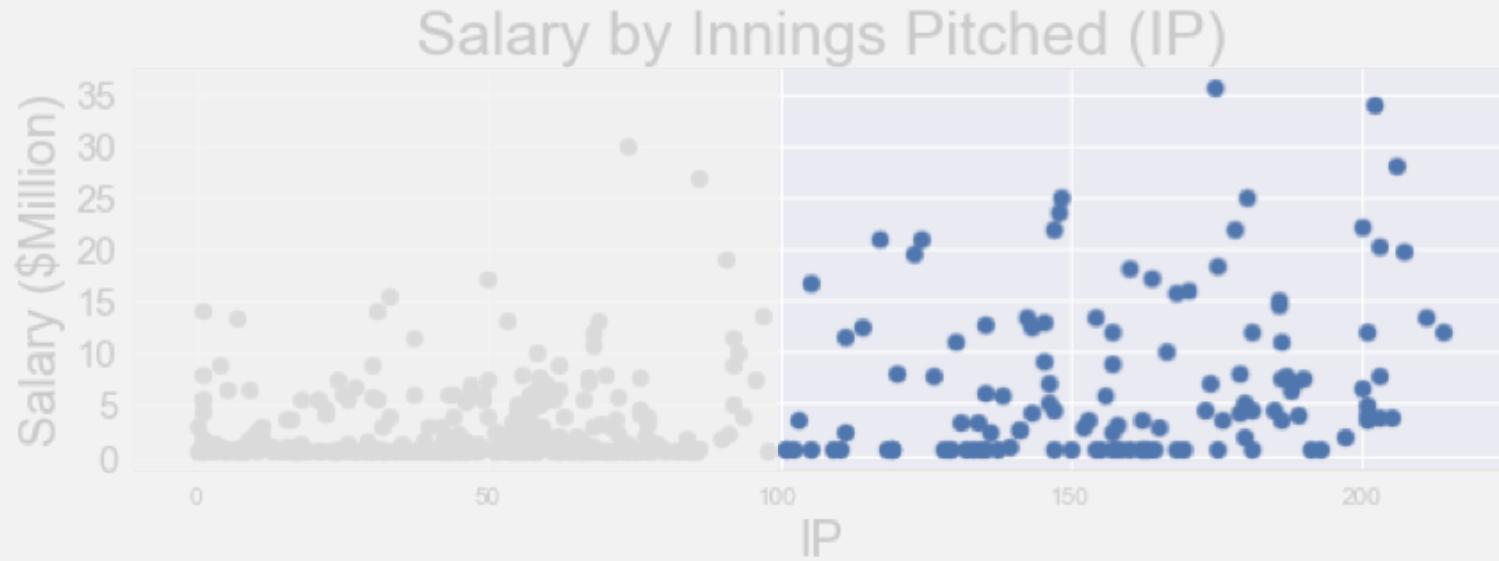
Clean



Focusing on Starting Pitchers

Data

Clean



Focusing on Starting Pitchers

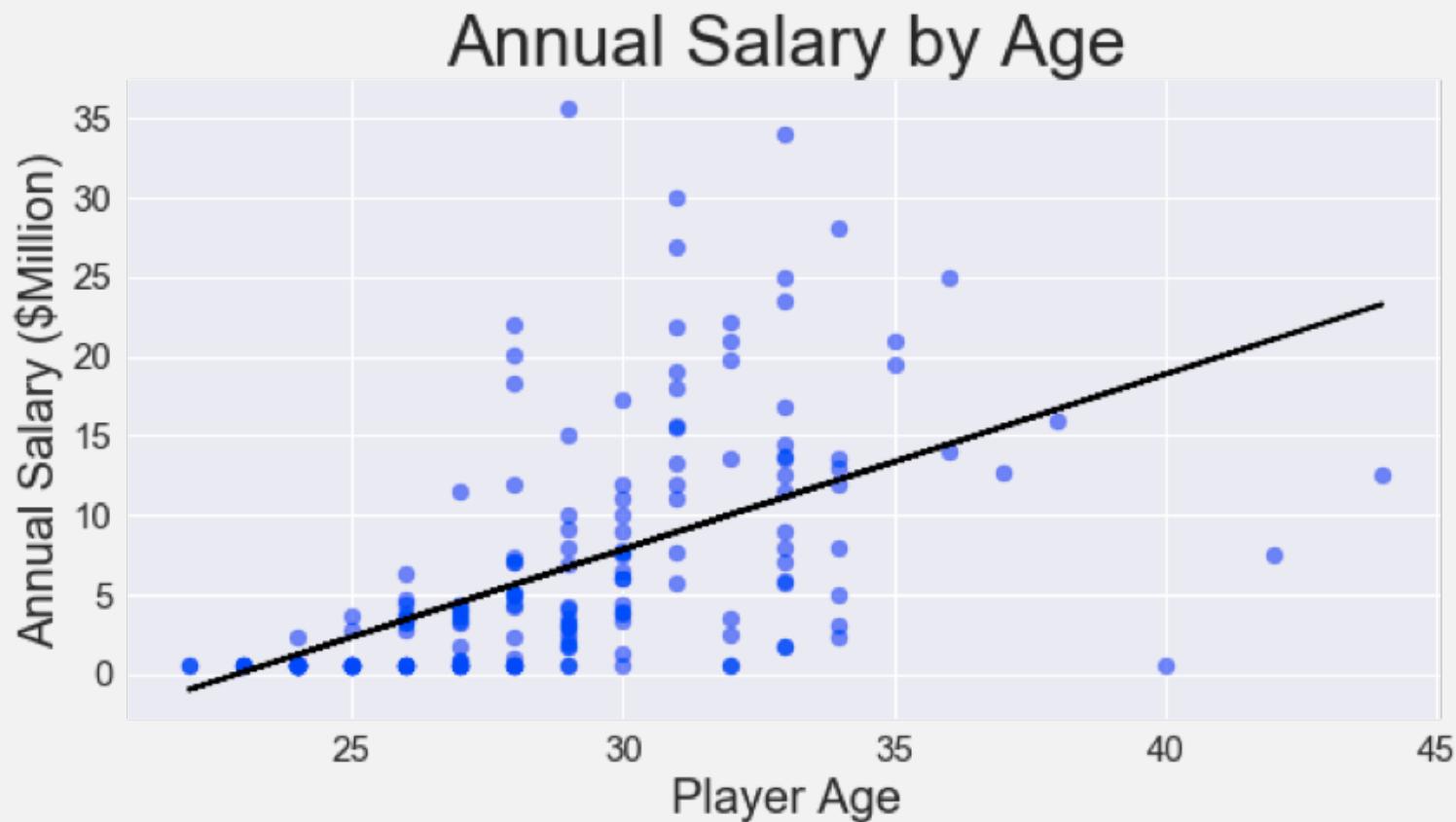
Data

Clean

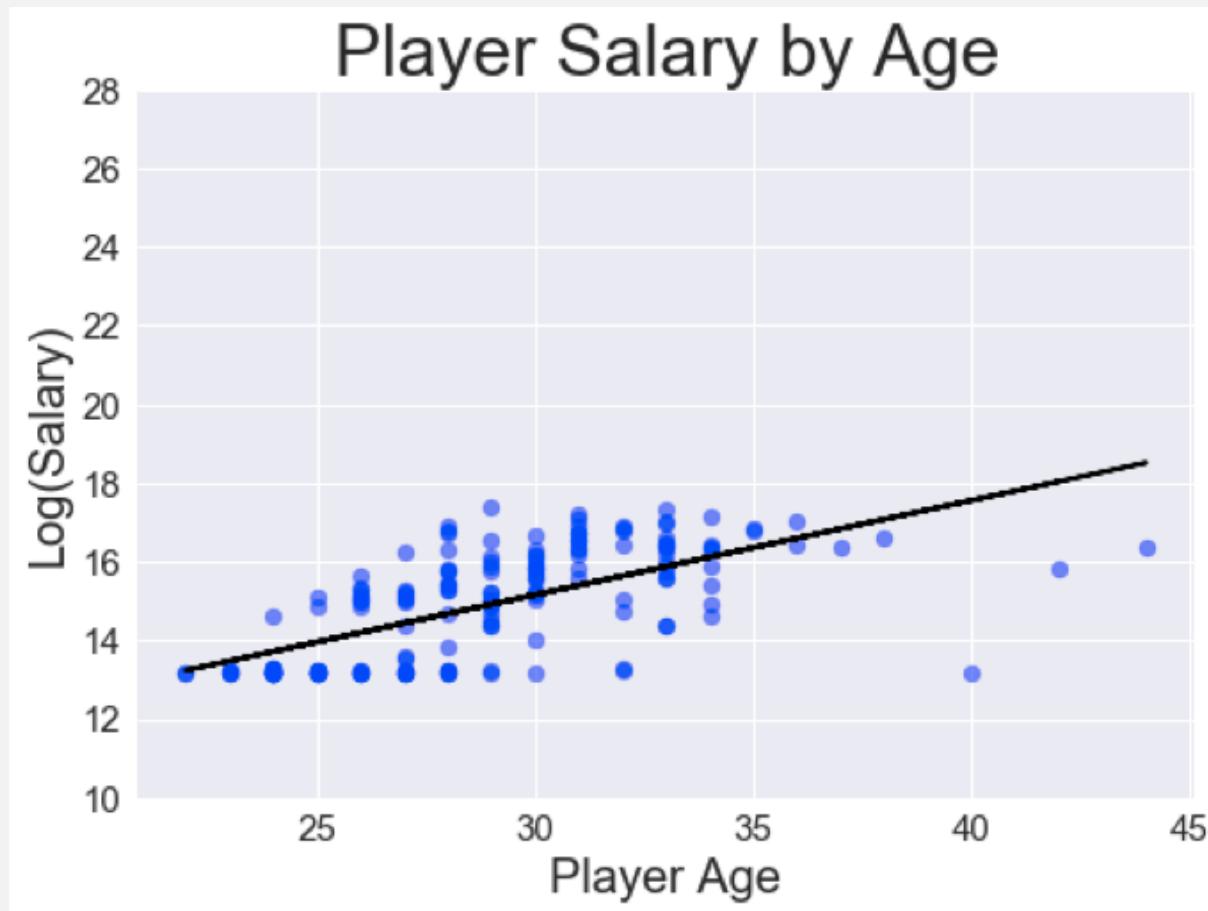
Position Players

728	Jon Jay*	32	CHC	1.0	1	0	0	0.00	4.64	0.23	-0.32	100.0	4.09	0	0.0	.00	0.0	0.0	1	.547	.500	\$8,000,000	Free Agency
729	Pierce Johnson	26	CHC	1.0	1	0	2	18.00	4.64	0.23	-0.32	100.0	4.09	-2	-0.1	.99	0.0	-0.1	-1	.362	.99		
730	Michael Martinez	34	CLE	1.0	1	0	0	0.00	4.72	0.30	-0.32	108.0	4.42	0	0.1	.02	0.0	0.0	1	.550	.500		Waivers
731	Miguel Montero	33	CHC	1.0	1	0	0	0.00	4.77	0.16	-0.33	100.0	4.28	0	0.0	.03	0.0	0.0	1	.549	.500	\$14,000,000	Traded
732	Mitch Moreland	31	BOS	1.0	1	0	0	0.00	4.97	0.59	-0.34	108.0	4.36	0	0.0	.00	0.0	0.0	1	.550	.500	\$5,500,000	Free Agency
733	J. A. Nichols	28	TEX	1.0	1	0	4	36.00	5.68	-0.08	-0.40	114.0	6.11	-3	-0.3	.00	0.1	-0.1	-3	.249	.498	4532.000	Amateur Draft
734	Hernan Perez	26	MIL	1.0	1	0	0	0.00	4.88	0.13	-0.34	104.0	4.59	1	0.1	.00	0.0	0.0	1	.553	.500	\$545,700	Waivers
735	Andrew Romine	31	DET	1.0	2	0	0	0.00	4.53	-0.45	-0.32	106.7	4.98	1	0.1	1.25	0.0	0.1	1	.528	.500	\$1,300,000	Traded
736	Carlos Ruiz	38	SEA	1.0	1	0	1	9.00	5.28	0.11	-0.37	100.0	4.81	0	0.0	.00	0.0	0.0	0	.454	.500	\$4,500,000	Traded

Annual Salary by Age



What if I Log Transform Salary?



Annual Salary Based on Age & IP

Data

Clean

Model

Player Salary by Age & IP

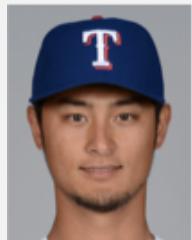


Log(Salary) Based on Age & IP



Predicted 2017 Salary: **\$5.64 Million**

Polynomial Features



Predicted 2017 Salary: **\$7.92 Million**



Actual



Year	Projected Salary	Actual Salary
2013	\$1.0M	\$1.0M
2014	\$1.5M	\$1.5M
2015	\$2.0M	\$2.0M
2016	\$2.5M	\$2.5M
2017	\$3.0M	\$11.0M

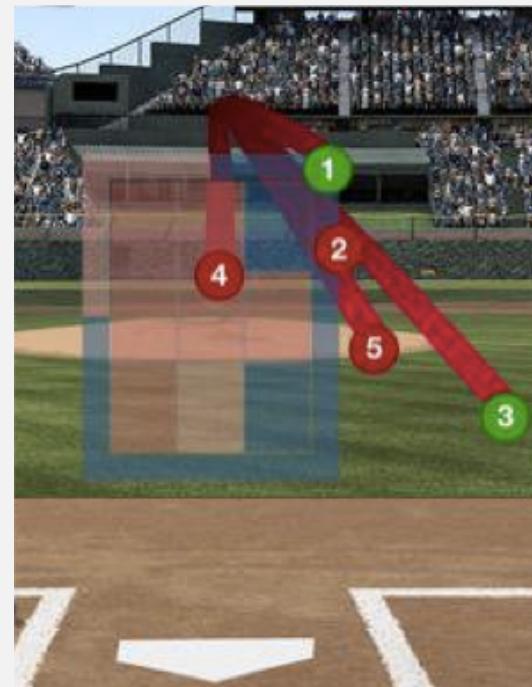
Actual 2017 Salary: \$11 Million

How Did I Do?



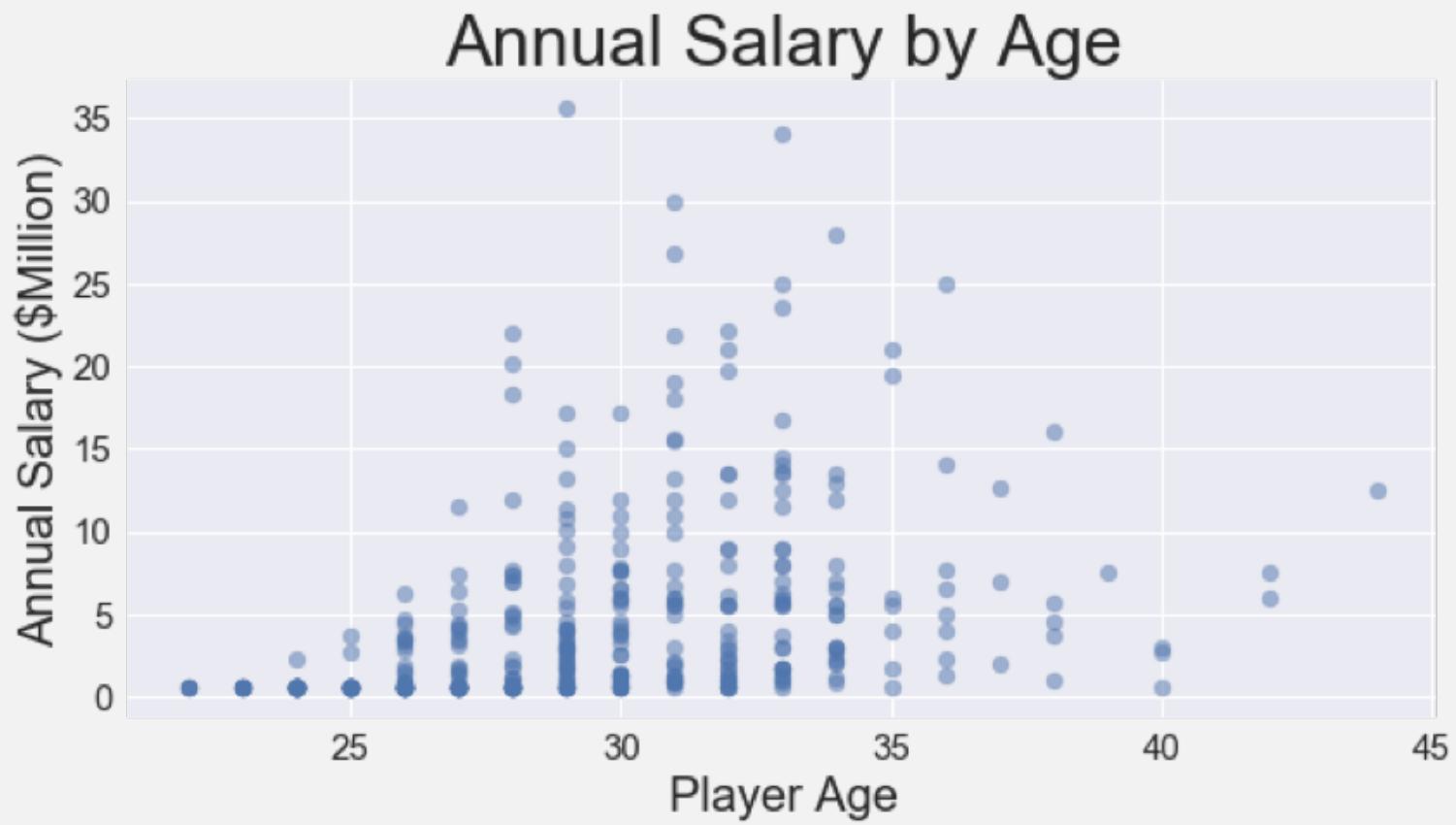
Further Investigation

- 1) 5-year average statistics
- 2) Consider weighting years
- 3) Explore different feature transformations
- 4) Apply different model
- 5) Explore a model for relief pitchers

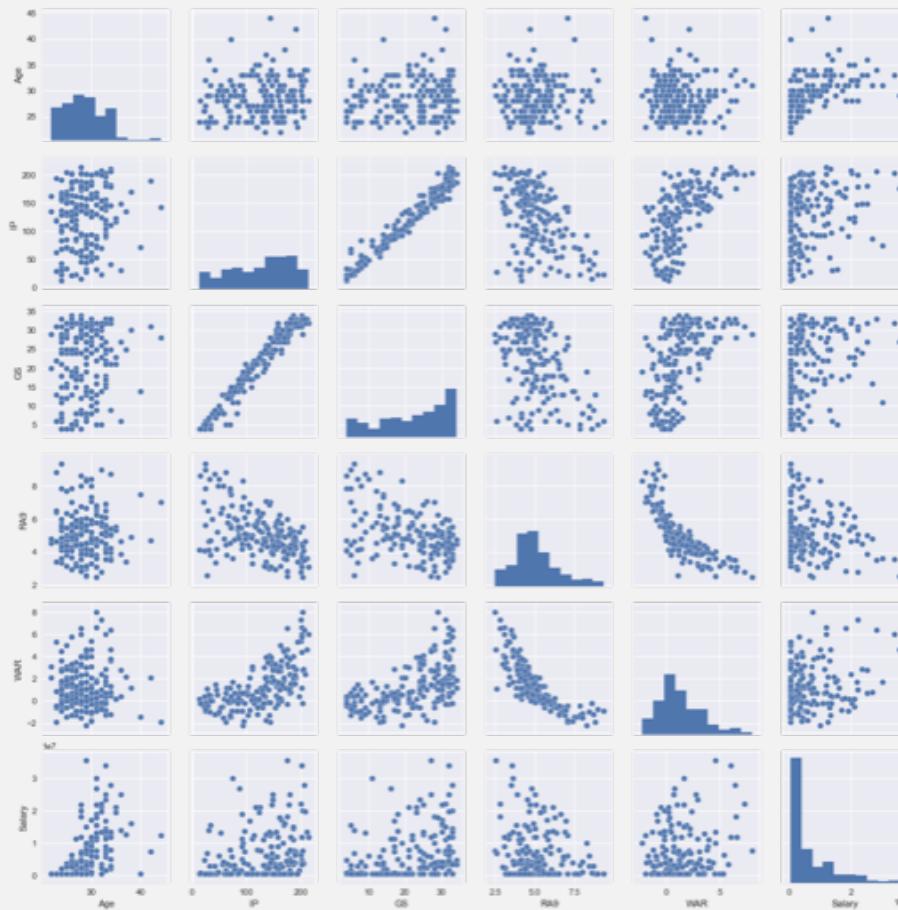


Appendix

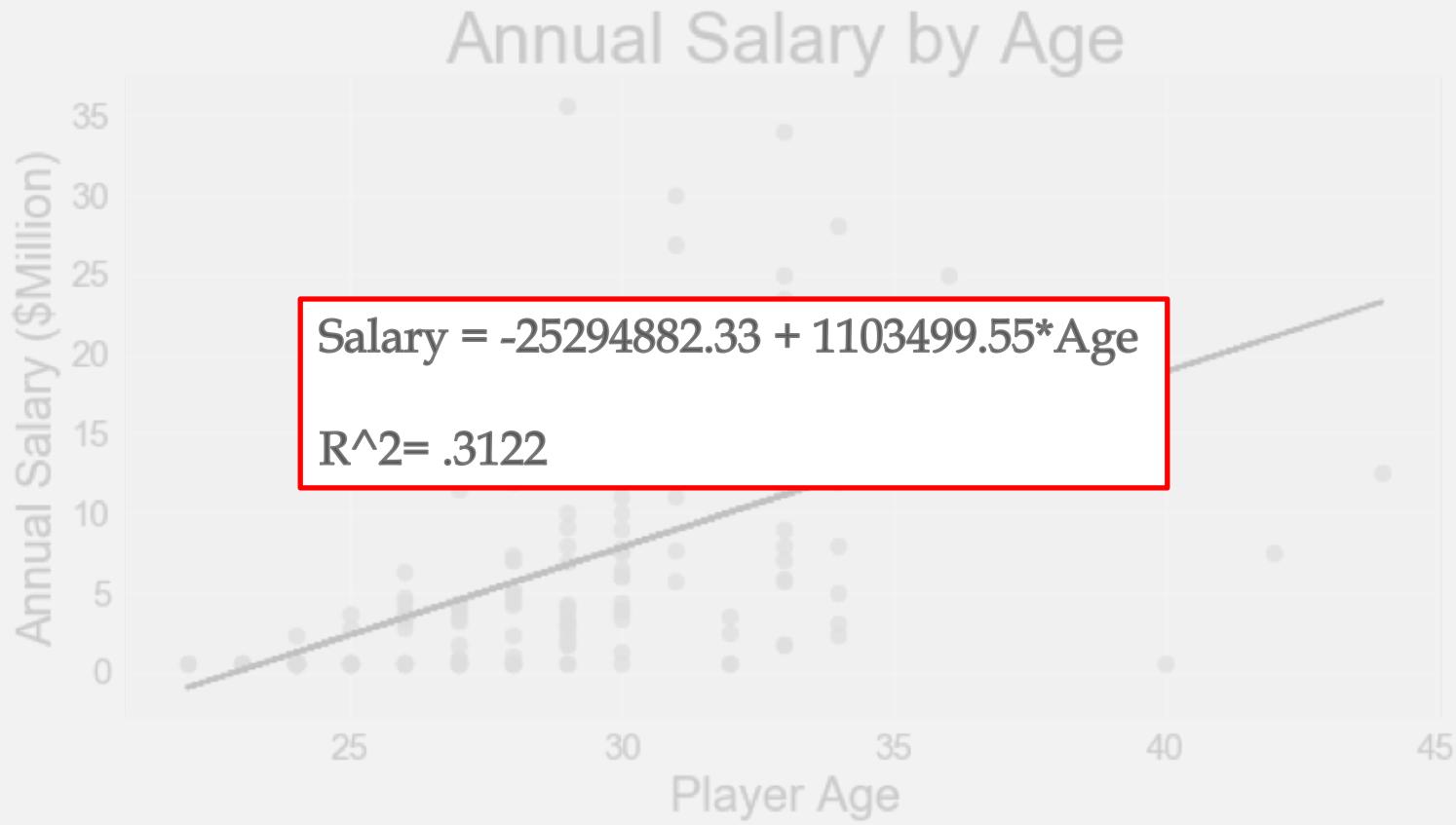
Annual Salary by Age



Feature Pair Plot



Annual Salary by Age

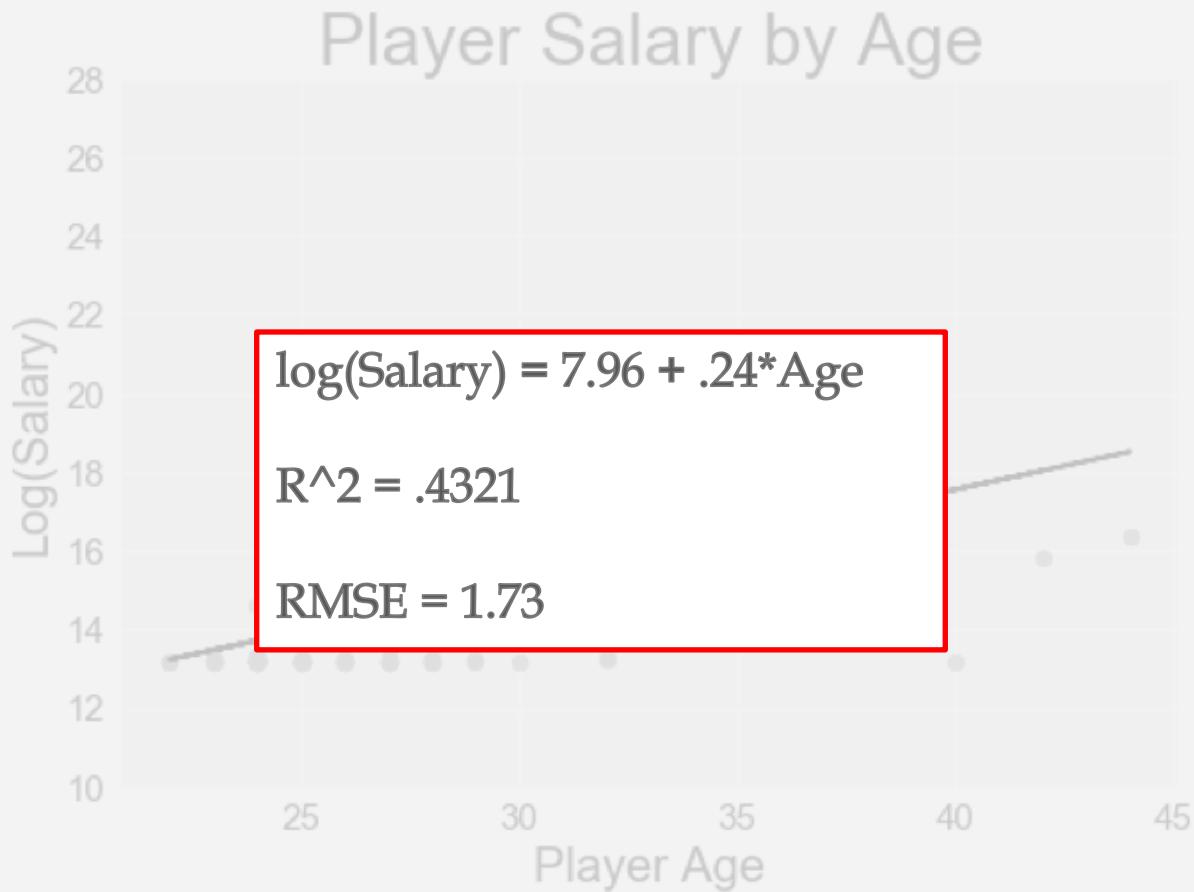


Annual Salary by Age

Data

Clean

Model



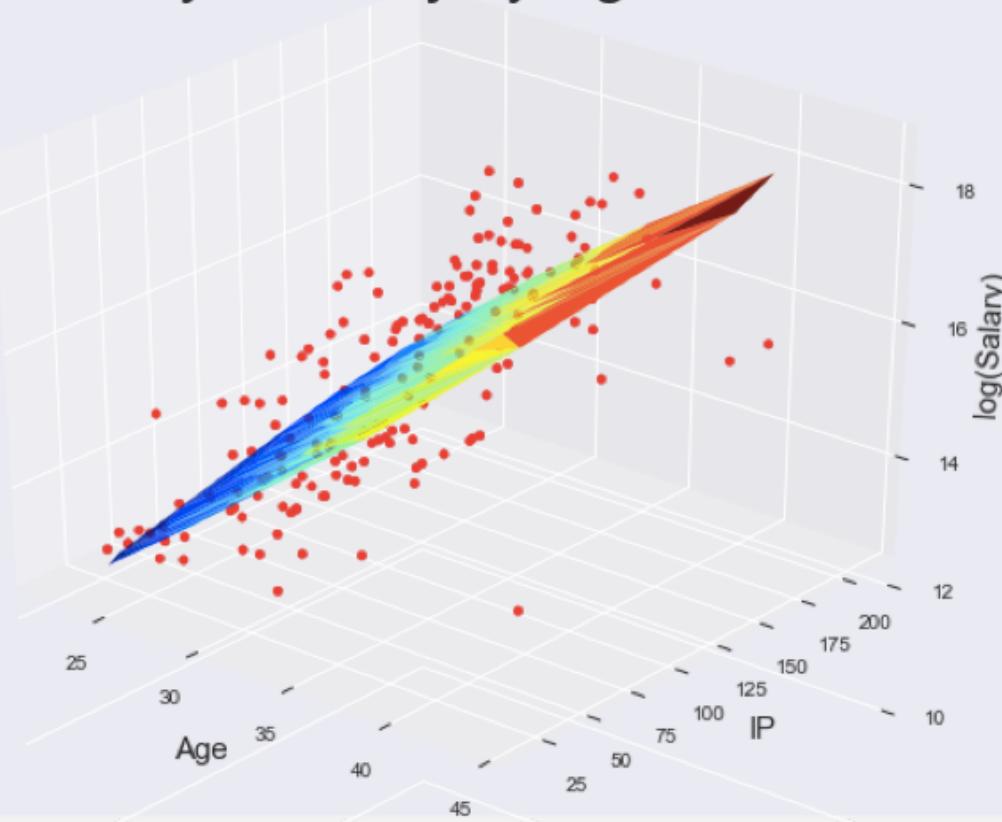
Log(Salary) Based on Age & IP

Data

Clean

Model

Player Salary by Age & IP



Annual Salary Based on Age & IP



Player Salary by Age & IP

$$\log(\text{Salary}) = 7.46 + .227 * \text{Age} + .007 * \text{IP}$$

$$\text{RMSE} = 1.62$$



Polynomial Features Residuals

