

Rockies 2017 - Week #1

Jim Reed (jdreed@q.com)

April 3, 2017

Introduction

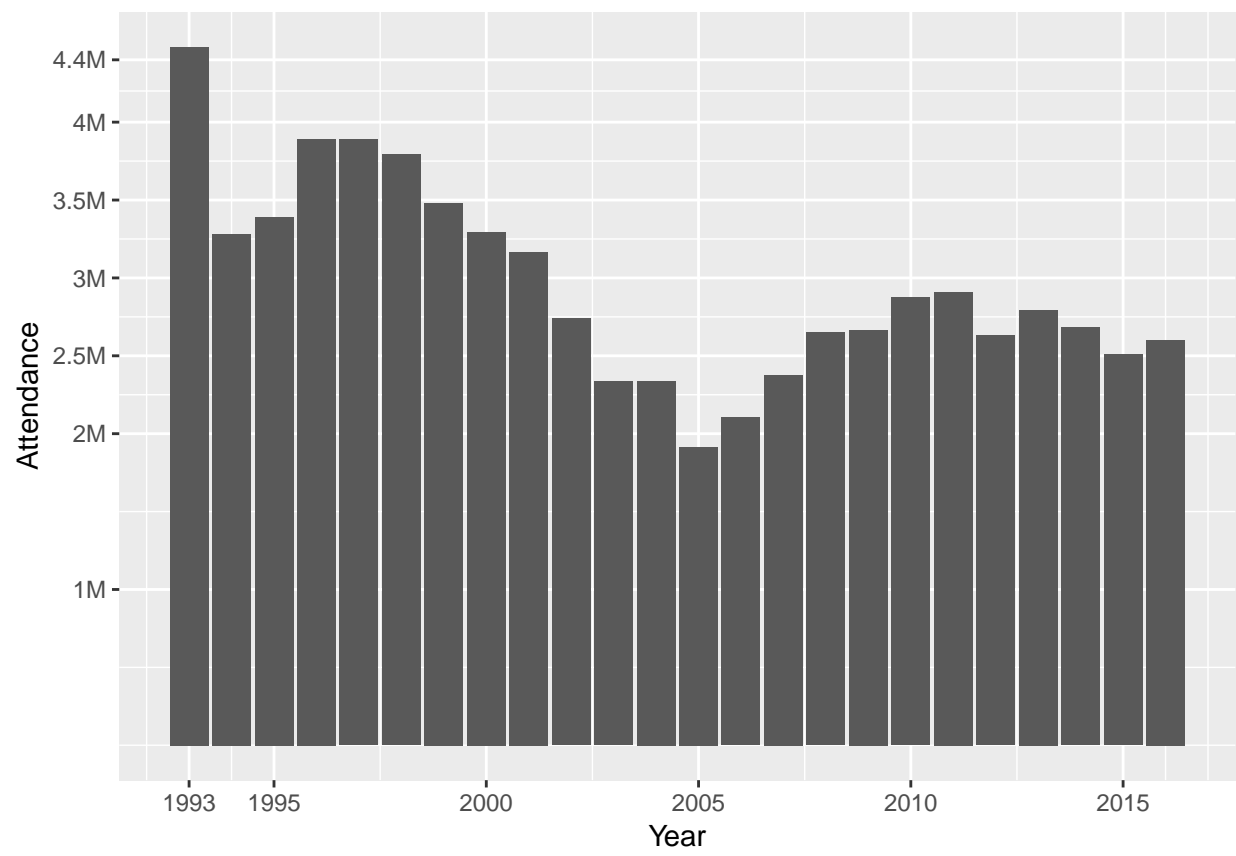
My purpose in writing this and future articles is twofold. The first is to try to impart some interesting information about the Colorado Rockies to you, the reader. My wife points out that each article should be told in the fashion of a story. So, using some statistics and data available on the internet, I will attempt to make each article a small story with beginning, middle and end.

Secondly, I am working to acquire more skills in data analytics, data science some call it. My tool of choice is R, a statistical framework and language. R is a powerful tool but has much complexity to master. One of the document types supported by R is the R Markdown, or, more specifically, R Notebooks. Think of an R Notebook as a lab notebook containing data and formulas as well as long form descriptions of what is observed. In this case, the data and code (programs) are hidden to make the overall results less geeky and perhaps more interesting to you, the reader. Of course, there will be add graphs and tables to illustrate a pertinent point.

A Short History of the Colorado Rockies 1993-2016

Year	G	W	L	Ties	W-L%	pythW-L%	Finish	GB	R	RA	Attendance
2016	162	75	87	0	0.463	0.492	3rd of 5	16	845	860	2602524
2015	162	68	94	0	0.420	0.438	5th of 5	24	737	844	2506789
2014	162	66	96	0	0.407	0.463	4th of 5	28	755	818	2680329
2013	162	74	88	0	0.457	0.466	5th of 5	18	706	760	2793828
2012	162	64	98	0	0.395	0.427	5th of 5	30	758	890	2630458
2011	162	73	89	0	0.451	0.476	4th of 5	21	735	774	2909777
2010	162	83	79	0	0.512	0.533	3rd of 5	9	770	717	2875245
2009	162	92	70	0	0.568	0.553	2nd of 5	3	804	715	2665080
2008	162	74	88	0	0.457	0.456	3rd of 5	10	747	822	2650218
2007	163	90	73	0	0.552	0.558	2nd of 5	0.5	860	758	2376250
2006	162	76	86	0	0.469	0.501	4th of 5	12	813	812	2104362
2005	162	67	95	0	0.414	0.431	5th of 5	15	740	862	1914389
2004	162	68	94	0	0.420	0.453	4th of 5	25	833	923	2338069
2003	162	74	88	0	0.457	0.480	4th of 5	26.5	853	892	2334085
2002	162	73	89	0	0.451	0.435	4th of 5	25	778	898	2737838
2001	162	73	89	0	0.451	0.509	5th of 5	19	923	906	3166821
2000	162	82	80	0	0.506	0.535	4th of 5	15	968	897	3295129
1999	162	72	90	0	0.444	0.442	5th of 5	28	906	1028	3481065
1998	162	77	85	0	0.475	0.484	4th of 5	21	826	855	3792683
1997	162	83	79	0	0.512	0.507	3rd of 4	7	923	908	3888453
1996	162	83	79	0	0.512	0.499	3rd of 4	8	961	964	3891014
1995	144	77	67	0	0.535	0.501	2nd of 4	1	785	783	3390037
1994	117	53	64	0	0.453	0.451	3rd of 4	6.5	573	638	3281511
1993	162	67	95	0	0.414	0.390	6th of 7	37	758	967	4483350

Bar Chart of Rockies Attendance 1993 - 2016



2017 Hope Springs Eternal

Whenever I hear the phrase *Hope Springs Eternal*, I think of Springtime (oddly) and the hope of a new baseball season. Will the Rockies have another memorable year like 1997 when they won the National League pennant?

The phrase comes from Alexander Pope's *An Essay on Man*

Hope springs eternal in the human breast; Man never is, but always to be blessed: The soul, uneasy and confined from home, Rests and expatiates in a life to come. – Alexander Pope, *An Essay on Man*

Well, here we are in week one of the 2017 MLB season. Will our Rockies rise to our hopes and expectations or make us wait for another time?

As some of you already know, I have caught the Sabermetrics fever over the winter, reading *Moneyball* by Michael Lewis.

Sabermetrics is defined by Wikipedia as:

Sabermetrics is the empirical analysis of baseball, especially baseball statistics that measure in-game activity. Sabermetricians collect and summarize the relevant data from this in-game activity to answer specific questions. The term is derived from the acronym SABR, which stands for the Society for American Baseball Research, founded in 1971. The term sabermetrics was coined by Bill James, who is one of its pioneers and is often considered its most prominent advocate and public face.

Another book, *Analyzing Baseball Data with R* by Max Marchi and Jim Albert gave me the technical framework to explore data sets available on the internet.

Following are some statistical tidbits I have manufactured on my own (based on historical data online) or gleaned from the internet. I will reference specific data from the internet with hyperlinks but generally, the data are either from the historical Lahman Database or Baseball Reference.

Rockies 2016, A Retrospective

The Rockies had an interesting season in 2016. Despite finishing number 2 of all MLB teams in OPS with .794 and Runs per Game, 5.22 (reference: 2016 Team Standard Batting). Despite all that success on offense, they still finished #20 with a win-loss record of 75-87 (.463). This, of course was way off the mark in making the playoffs. The old saw goes, “You gotta win at least ninety games to get into the playoffs.” The **Rockies** were fifteen games shy of that goal.

Pythagorean Win-Loss Theorem

The Pythagorean Win-Loss Theorem of Baseball is a creation of Bill James which relates the number of runs a team has scored (R) and surrendered (RA) to its actual winning percentage, based on the idea that runs scored compared to runs allowed is a better indicator of a team's (future) performance than a team's actual winning percentage. Feb 20, 2015

$$predictW - L = \frac{R^2}{R^2 + RA^2}$$

I won't go into the details (perhaps later in the season, if anyone is interested), except to make two points:

1. Pythagoras had nothing to do with this theorem. The formula looks somewhat similar to the calculation of the distance between two points in space or solving for the length of the hypotenuse of a right triangle. The usage of “Pythagorean” in this instance is all the idea of Bill James, credited as the *Father of Sabermetrics*.

2. The Rockies had a *predicted* Win-Loss ratio of 80-82. Simply put, the number of runs scored by and against the Rockies would indicate that they should have won at least eighty games in 2016.

Rockies 2017

The Rockies 40-man roster is young. In fact it's the fourth youngest team in Major League Baseball (out of 30). The oldest average is 29.9 (Toronto). The 2017 Rockies have an average of 27.6.

Rockies Roster - Current Career Statistics for non-Pitchers

Last	First	Age	Yrs	G	AB	R	H	2B	3B	HR	RBI	BB	SO	SB	AVG	SLG
Adames	Cristhian	25	2	33	68	5	14	1	1	0	3	3	16	0	0.206	0.250
Amarista	Alexi	28	6	541	1442	140	328	55	16	18	139	94	231	29	0.227	0.325
Arenado	Nolan	26	3	401	1534	204	431	106	10	70	243	82	240	6	0.281	0.500
Blackmon	Charlie	30	5	462	1664	234	480	84	14	45	169	91	282	84	0.288	0.437
Desmond	Ian	31	7	927	3480	424	917	185	21	110	432	232	890	122	0.264	0.424
Garneau	Dustin	29	1	22	70	6	11	3	0	2	8	6	14	0	0.157	0.286
Gonzalez	Carlos	31	8	914	3371	570	978	191	34	176	554	291	827	112	0.290	0.524
LeMahieu	DJ	28	5	526	1751	212	498	71	17	15	157	116	325	52	0.284	0.370
Murphy	Tom	26	1	11	35	5	9	1	0	3	9	4	10	0	0.257	0.543
Parra	Gerardo	30	9	988	3326	429	920	182	37	56	311	232	616	69	0.277	0.404
Reynolds	Mark	33	10	1258	4178	612	959	192	13	237	661	550	1519	59	0.230	0.452

Offensive Reason for Concern

In the 2016, the Rockies ended the year with the second best team OPS of .794. OPS is defined as:

On-base plus slugging (OPS) is a sabermetric baseball statistic calculated as the sum of a player's on-base percentage and slugging average. The ability of a player both to get on base and to hit for power, two important offensive skills, are represented.

This year's OPS average of non-pitchers on the 40-man roster is 0.655. This would put the Rockies in the below average range as defined by Bill James.

Category	Classification	OPS Range
A	Great	.9000 and Higher
B	Very Good	.8333 to .8999
C	Above Average	.7667 to .8333
D	Average	.7000 to .7666
E	Below Average	.6334 to .6999
F	Poor	.5667 to .6333
G	Very Poor	.5666 and Lower

Reason for concern, true, but the youth of the team and, whole season ahead, perhaps the Rockies will live up to their 2016 benchmark for batting efficiency again this year.

Topics for Additional Sabermetric Tidbits

The Rockies have consistently performed well at the plate. The most-often quoted reason for this is, "Because Coors Field is situated at the highest altitude in MLB, it is considered a hitter's park. Of course, the Rockies

turn in superior batting statistics.” Though, this is likely, the teams we face get to play at Coors Field, too. Is the reason for the Rockies’ struggle to get over a win-loss of .500, the skill of the Rockies pitching staff. Do the Rockies have a hard time attracting superior pitchers because of our “hitter’s park” issue?

Anyway, I would like to explore pitching metrics in future articles.

Let me know what you would like to see in future articles. Send me email at jdreed@q.com.

Yours truly,

Jim Reed