BASEBALL STATISTICS: Examples of the Mean, Median, and Mode

by *Bonnie H. Litwiller* and *David R. Duncan*
Professors of Mathematics
University of Northern Iowa
Cedar Falls, Iowa

When students refer to "the average" of a set of data, they usually are describing either the mean (sum of the scores divided by the number of scores), the median (the middle score), or the mode (the most frequent score). Teachers are always searching for examples of data for which the mean, median, and mode present dissimilar meanings of average. For example, suppose an 80-year-old grandfather and 78-year-old grandmother invite their three grandchildren for lunch. They are the twins (Shawn and Heather, age 3) and Gary (age 6). The mean age of those present at lunch is 34, the median age is 6, and the mode is 3. Which of these measures of central tendency gives the best representation of "average"?

We attempted to find a real world situation in which the mean, median, and mode would be greatly different. We located the names of all major league baseball players whose careers were contained wholly in the years 1920 through 1974. For each of these players, we located the total number of home runs that he hit in his entire *regular season* career. Complete information concerning these and other major league baseball statistics can be found in the publication *The Baseball Encyclopedia* published in 1976 by the Macmillan Publishing Company.

We used 1920 as the starting date because of the common agreement that this was the first year that the "live ball" was used. Prior to that time, the ball that was used was very difficult to hit a great distance. This change was apparently made to make baseball more exciting. Consequently, we did not include a player whose career began before 1920, since he did not have the opportunity to hit the live ball for his entire career. We made one exception - we included Babe Ruth who was able to hit *any* type of ball.

We stopped with those players whose careers ended prior to or in 1974 because their career records were then complete in our reference book.

Also we did not include any player who was primarily a pitcher, since he would not have the same opportunity to bat as would players in other positions.

Table I shows the frequency distribution for all of the 3049 players that we described previously.

TABLE I

Total Number of Home Runs	Number of Players	Cumulative Frequency
0	1138	1138
1	290	1428
2	172	1600
3	95	1695
4	77	1772
5	56	1828
6	71	1899
7	43	1942

Total Number of Home Runs	Number of Players	Cumulative Frequency	Total Number of Home Runs	Number of Players	Cumulative Frequency
8	50	1992	84	3	2799
9	47	2039	85	2	2801
10	22	2061	86	3	2804
11	22	2083	87	4	2808
12	30	2113	88	1	2809
13	32	2145	89	5	2814
14	32	2177	90	3	2817
15	26	2203	91	3	2820
16	28	2231	92	1	2821
17	19	2250	93	7	2828
18	16	2266	94	4	2832
19	26	2292	95	4	2836
20	17	2309	96	2	2838
21	10	2319	97	1	2839
22	27	2346	98	4	2843
23	11	2357	99	1	2844
24	19	2376	100	2	2846
25	10	2386	101	4	2850
26	13	2399	103	5	2855
27	12	2411	104	5	2860
28	19	2430	105	4	2864
29	19	2449	106	6	2870
30	9	2458	107	1	2871
31	17	2475	108	3	2874
32	19	2494	109	1	2875
33	6	2500	110	2	2877
34	6	2506	111	1	2878
35	8	2514	112	4	2882
36	9	2523	113	1	2883
37	8	2531	114	1	2884
38	13	2544	115	3	2887
39	11	2555	116	2	2889
40	11	2566	117	3	2892
41	9	2575	119	3	2895
42	6	2581	120	1	2896
43	6	2587	121	1	2897
44	12	2599	122	1	2898
45	8	2607	124	1	2899
46	10	2617	125	3	2902
47	10	2627	126	4	2906
48	7	2634	127	2	2908
49	5	2639	128	1	2909
50	8	2647	129	2	2911
51	4	2651	130	3	2914
52	6	2657	131	3	2917
53	4	2661	132	2	2919
54	3	2664	133	2	2921
55	4	2668	134	2	2923
56	5	2673	135	3	2926
57	7	2680	136	1	2927
58	10	2690	137	1	2928
59	3	2693	138	1	2929
60	4	2697	139	1	2930
61	12	2709	141	1	2931
62	3	2712	142	1	2932
63	3	2715	147	1	2933
64	4	2719	148	2	2935
65	7	2726	149	1	2936
66	6	2732	151	1	2937
67	4	2736	152	1	2938
68	4	2740	153	1	2939
69	7	2747	154	1	2940
70	1	2748	155	2	2942
71	1	2749	156	1	2943
72	3	2752	159	2	2945
73	7	2759	162	1	2946
74	1	2760	163	2	2948
75	1	2761	164	5	2953
76	2	2763	166	1	2954
77	8	2771	167	2	2956
78	3	2774	169	1	2957
79	7	2781	170	3	2960
80	5	2786	172	2	2962
81	3	2789	173	1	2963
82	3	2792	178	2	2965
83	4	2796	179	3	2968

Total Number of Home Runs	Number of Players	Cumulative Frequency
180	1	2969
181	1	2970
183	2	2972
184	1	2973
186	1	2974
189	1	2975
190	2	2977
192	2	2979
194	1	2980
199	1	2981
200	1	2982
202	3	2985
205	1	2986
206	2	2988
210	1	2989
211	2	2991
213	1	2992
219	2	2994
223	1	2995
226	1	2996
228	2	2998
235	1	2999
236	1	3000
237	1	3001
238	1	3002
239	1	3003
240	1	3004
242	2	3006
244	1	3007
247	1	3008
248	1	3009
253	2	3011
256	1	3012
264	1	3013
266	1	3014
275	1	3015
277	1	3016
279	1	3017
282	1	3018
286	1	3019
288	3	3022
300	1	3023
307	1	3024
318	1	3025
331	1	3026
336	1	3027
342	1	3028
358	1	3029
359	1	3030
361	1	3031
369	1	3032
370	1	3033
374	1	3034
377	1	3035
379	1	3036
382	1	3037
399	1	3038
407	1	3039
475	1	3040
493	1	3041
511	1	3042
512	2	3044
521	1	3045
534	1	3046
536	1	3047
660	1	3048
714	1	3049

Mean = 23.91
Median = 2
Mode = 0
Standard Deviation = 57.98

Observe that the mean, median, and mode differ greatly. The standard deviation is more than twice as large as the mean. Clearly, these data would not describe the normal curve.

In the normal curve, 68 percent (approximately two-thirds) of the data points should be within one standard deviation of the mean. In this baseball example, the data points less than the mean all lie within 41 percent of a standard deviation of the mean, while 2789 of the 3049 data points are within one standard deviation of the mean. This means that 91.5 percent of all the data points are within one standard deviation of the mean.

To better visualize these data, let the height of each player be proportional to the total number of home runs that he hit in his career. Also let the mean number of home runs represent a player whose "height" is 6 feet.

Suppose that all 3049 baseball players line up according to "height" (shortest to tallest) and begin running out of a dugout at the rate of one per second. If the first player runs out of the dugout at 1:00 p.m., the last player will appear at 1:50:48 p.m. (48 seconds after 1:50 p.m.)

We assume that for any given number of home runs, the players in that category run out in alphabetical order. The first player in line is Jimmy Adair (Chicago Cubs, 1931) who hit no homeruns in his brief 18-game career as a shortstop. He did however bat .276 in 76 times at bat. Adair's height is 0 feet.

For several minutes the players are of height 0. The last such player in this category is Frank Zupo (Baltimore Orioles, 1957-61). Zupo was a

little used catcher, as he appeared in 16 games and had 18 times at bat in his career. He appears at 1:18:57.

At 1:18:58 (58 seconds after 1:18), the first player whose height is greater than 0 appears. He is Joe Abreu (Cincinnati Reds, 1942). He played in 9 games and hit one home run. His height is .25 feet or 3.0 inches.

Luiz Alcaraz (Dodgers and Royals, 1967-70) who hit 4 home runs appears at 1:28:15. His height is one foot and he follows the "median player" by only 2 minutes and 51 seconds. The median player, Johnny Neun (Tigers and Red Sox, 1925-31) hit 2 home runs and is 6.0 inches tall. Observe that this small height represents the median in a group of baseball players whose "mean height" is 6 feet. Consequently, over half of the baseball players are 6 inches or less in "height" - a short group of baseball players!

The first player whose height is greater than or equal to -

1. 2 feet is Ruben Amaro, who hit 8 home runs and appears at 1:32:22.
2. 3 feet is Babe Barna, who hit 12 home runs and appears at 1:34:43.
3. 6 feet (the mean) is Pete Castiglione, who hit 24 home runs and appears at 1:39:17. With only 11 minutes and 31 seconds remaining of the 50 minutes and 48 seconds that it takes for the group to run from the dugout, the first player taller than the mean height appears.
4. 10 feet is Ed Coleman, who hit 40 home runs and appears at 1:42:35.
5. 20 feet is Bobby Avila, who hit 80 home runs and appears at 1:46:21.
6. 50 feet is Don Mincher, who hit 200 home runs and appears at 1:49:41.

There are 12 players whose heights exceed 100 feet. They are:

1. Al Kaline, who hit 399 home runs. He is 100 feet 2 inches in height and appears at 1:50:37.
2. Duke Snider, who hit 407 home runs. He is 102 feet 2 inches in height and appears at 1:50:38.
3. Stan Musial, who hit 475 home runs. He is 119 feet 2 inches in height and appears at 1:50:39.
4. Lou Gehrig, who hit 493 home runs. He is 123 feet 9 inches in height and appears at 1:50:40.
5. Mel Ott, who hit 511 home runs. He is 128 feet 3 inches in height and appears at 1:50:41.
6. Ernie Banks, who hit 512 home runs. He is 128 feet 6 inches in height and appears at 1:50:42.
7. Eddie Mathews, who hit 512 home runs. He is 128 feet 6 inches in height and appears at 1:50:44.
8. Ted Williams, who hit 521 home runs. He is 130 feet 9 inches in height and appears at 1:50:44.
9. Jimmie Foxx, who hit 534 home runs. He is 134 feet in height and appears at 1:50:45.
10. Mickey Mantle, who hit 536 home runs. He is 134 feet 6 inches in height and appears at 1:50:46.
11. Willie Mays, who hit 660 home runs. He is 165 feet 7 inches in height and appears at 1:50:47.
12. Babe Ruth, who hit 714 home runs. He is 179 feet 2 inches in height and appears at 1:50:48.

(Hank Aaron would be even taller, but he retired after our cut-off date.)

It is interesting to observe the heights of players who appear at

specific times. Table II displays this information.

TABLE II

Time	Player	Number of Home Runs	Height
1:00	Jimmy Adair	0	0
1:05	Bobby Floyd	0	0
1:10	Red Lutz	0	0
1:15	Norm Schlueter	0	0
1:20	Ox Eckhardt	1	3.0 in.
1:25	Ty La Forest	2	6.0 in.
1:30	Joe Lafata	5	1 ft. 3 in.
1:35	Lloyd Merriman	12	3 ft.
1:40	Horace Clarke	27	6 ft. 9 in.
1:45	Billy Johnson	61	15 ft. 4 in.
1:46	Sammy West	75	18 ft. 10 in.
1:47	Jim Hegan	92	23 ft. 1 in.
1:48	Paul Waner	112	28 ft. 1 in.
1:49	Ed Bailey	155	38 ft. 11 in.
1:50:00	Gus Zernial	237	59 ft. 6 in.
1:50:10	Joe Gordon	253	63 ft. 7 in.
1:50:20	Bob Johnson	288	72 ft. 3 in.
1:50:30	Joe DiMaggio	361	90 ft. 7 in.
1:50:35	Orlando Cepeda	379	95 ft. 1 in.
1:50:40	Lou Gehrig	493	123 ft. 9 in.
1:50:45	Jimmie Foxx	534	134 ft.
1:50:46	Mickey Mantle	536	134 ft. 6 in.
1:50:47	Willie Mays	660	165 ft. 7 in.
1:50:48	Babe Ruth	714	179 ft. 2 in.

Suggestions for the reader and his/her class:

1. Calculate a similar set of data for basketball players. For example, calculate the "heights" of basketball players using their total numbers of points.
2. Calculate the heights of players on your high school or favorite college baseball team.
3. Calculate the heights of baseball players using other statistics such as total hits or stolen bases.
4. Look for settings other than sports which give rise to "badly skewed" data. (The median is very close to one end of the range.) Published salaries of public institutions may be interesting to investigate for this purpose.