

# **Analyzing regression models for predicting the outcomes of Major League Baseball games**

Jeremy Moon

*August 2, 2024*

*Institute for Computing in Research*

MLB · Fri, Aug 2, 7:05 PM



Boston Red Sox

Google

at



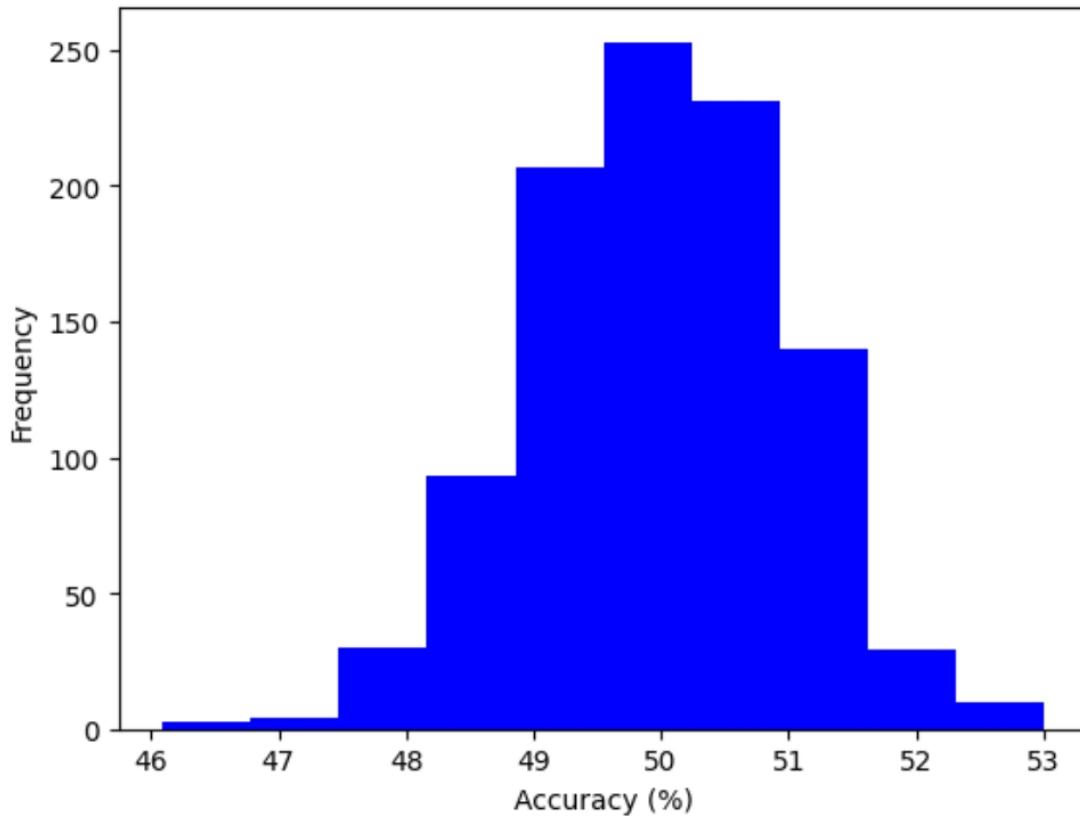
Texas Rangers

# Why?

# 1

Set a standard

# Random guessing



count	1000.000000
mean	49.993333
std	1.009721
min	46.090535
25%	49.341564
50%	50.000000
75%	50.699588
max	53.004115

# 2

Collect data

away_team	away_score	away_team_xwoba_diff	away_lineup_xwoba	away_team_xfp	away_starter_xfp	away_close_win_pct	away_run_diff	l
SF	5	-0.01395	0.324623035548816	4.3175	3.6649265262021	0.481528046421663	-50.85	/
CHC	3	0.0179	0.320554970785412	3.9005	3.5721621836341	0.577740192623914	151.35	\$
NYY	3	0.0218	0.332433249977573	3.812	3.4600865800866	0.485525467440361	123.2	-
PHI	4	-0.0244	0.299168151537865	4.396	4.7980995323981	0.414121487028104	-129.15	0
SD	3	-0.0263	0.275540411693033	4.2655	4.3873105909221	0.471744186046512	-160.35	l
COL	7	-0.01285	0.312190243126617	4.263	3.5588136977694	0.521931818181818	20.41	
ATL	0	-0.0131	0.309743940246347	4.576	4.563091734629	0.478513650151668	-113.95	l
MIA	2	-0.0131	0.330383062524465	4.627	4.5964105547476	0.467974308300395	-42.9	0
TOR	2	0.0112	0.332435802227687	4.242	4.9326034376918	0.460795638819467	1.8	l
PIT	3	-0.0081	0.324551829633789	4.202	3.7098924731183	0.496565284692298	-29.9	0
SEA	0	0.0023	0.326119810644743	4.425	3.972082482028	0.579172015404365	-8.45	0
KC	1	-0.00935	0.316878485682368	4.528	4.2619339111593	0.571895325203252	-50.2	0
LAA	2	-0.0099	0.32339481849641	4.426	4.5652940043342	0.546438648224363	-4	0
CLE	8	0.03065	0.334295037142721	3.5315	2.9073007393523	0.553896103896104	182	
DET	2	0.01005	0.300510010000007	4.6175	3.0000001100170	0.401500000000000	50.05	

# Runs scored

response variable ( $y$ -value)

away_team	away_score	away_team_xwoba_diff	away_lineup_xwoba	away_team_xfip	away_starter_xfip	away_close_win_pct	away_run_diff	l
SF	5	-0.01395	0.324623035548816	4.3175	3.6649265262021	0.481528046421663	-50.85	1
CHC	3	0.0179	0.320554970785412	3.9005	3.5721621836341	0.577740192623914	151.35	1
NYY	3	0.0218	0.332433249977573	3.812	3.4600865800866	0.485525467440361	123.2	1
PHI	4	-0.0244	0.299168151537865	4.396	4.7980995323981	0.414121487028104	-129.15	0
SD	3	-0.0263	0.275540411693033	4.2655	4.3873105909221	0.471744186046512	-160.35	1
COL	7	-0.01285	0.312190243126617	4.263	3.5588136977694	0.521931818181818	20.41	1
ATL	0	-0.0131	0.309743940246347	4.576	4.563091734629	0.478513650151668	-113.95	0
MIA	2	-0.0131	0.330383062524465	4.627	4.5964105547476	0.467974308300395	-42.9	0
TOR	2	0.0112	0.332435802227687	4.242	4.9326034376918	0.460795638819467	1.81	1
PIT	3	-0.0081	0.324551829633789	4.202	3.7098924731183	0.496565284692298	-29.91	0
SEA	0	0.0023	0.326119810644743	4.425	3.972082482028	0.579172015404365	-8.45	1
KC	1	-0.00935	0.316878485682368	4.528	4.2619339111593	0.571895325203252	-50.21	1
LAA	2	-0.0099	0.32339481849641	4.426	4.5652940043342	0.546438648224363	-40	0
CLE	8	0.03065	0.334295037142721	3.5315	2.9073007393523	0.553896103896104	182	1
DET	0	0.01005	0.30051001000000007	4.6175	3.0000001100170	0.40150000000000007	50.05	1

# xwOBA

expected weighted on base average

away_team	away_score	away_team_xwoba_diff	away_lineup_xwoba	away_team_xfp	away_starter_xfp	away_close_win_pct	away_run_diff	l
SF		-0.01395	0.324623035548816	4.3175	3.6649265262021	0.481528046421663	-50.85	/
CHC		0.0179	0.320554970785412	3.9005	3.5721621836341	0.577740192623914	151.35	\$
NYY		0.0218	0.332433249977573	3.812	3.4600865800866	0.485525467440361	123.2	-
PHI		-0.0244	0.299168151537865	4.396	4.7980995323981	0.414121487028104	-129.15	0
SD		-0.0263	0.275540411693033	4.2655	4.3873105909221	0.471744186046512	-160.35	l
COL		-0.01285	0.312190243126617	4.263	3.5588136977694	0.521931818181818	20.41	
ATL		-0.0131	0.309743940246347	4.576	4.563091734629	0.478513650151668	-113.95	!
MIA		-0.0131	0.330383062524465	4.627	4.5964105547476	0.467974308300395	-42.9	0
TOR		0.0112	0.332435802227687	4.242	4.9326034376918	0.460795638819467	1.81	
PIT		-0.0081	0.324551829633789	4.202	3.7098924731183	0.496565284692298	-29.9	!
SEA		0.0023	0.326119810644743	4.425	3.972082482028	0.579172015404365	-8.45	!
KC		-0.00935	0.316878485682368	4.528	4.2619339111593	0.571895325203252	-50.21	l
LAA		-0.0099	0.32339481849641	4.426	4.5652940043342	0.546438648224363	-4	0
CLE		0.03065	0.334295037142721	3.5315	2.9073007393523	0.553896103896104	182	
DET				4.6175	0.0000000000000000	0.4015000000000000	50.05	

# xFIP

expected fielding independent pitching

away_team	away_score	away_team_xwoba_diff	away_lineup_xwoba	away_team_xfip	away_starter_xfip	away_close_win_pct	away_run_diff	l
SF	5	-0.01395	0.324623035548816	4.3175	3.6649265262021	0.481528046421663	-50.85	1
CHC	3	0.0179	0.320554970785412	3.9005	3.5721621836341	0.577740192623914	151.35	1
NYY	3	0.0218	0.332433249977573	3.812	3.4600865800866	0.485525467440361	123.2	1
PHI	4	-0.0244	0.299168151537865	4.396	4.7980995323981	0.414121487028104	-129.15	0
SD	3	-0.0263	0.275540411693033	4.2655	4.3873105909221	0.471744186046512	-160.35	1
COL	7	-0.01285	0.312190243126617	4.263	3.5588136977694	0.521931818181818	20.41	1
ATL	0	-0.0131	0.309743940246347	4.576	4.563091734629	0.478513650151668	-113.95	0
MIA	2	-0.0131	0.330383062524463	4.627	4.5964105547476	0.467974308300395	-42.9	0
TOR	2	0.0112	0.332435802227687	4.242	4.9326034376918	0.460795638819467	1.81	1
PIT	3	-0.0081	0.324551829633789	4.202	3.7098924731183	0.496565284692298	-29.9	0
SEA	0	0.0023	0.326119810644743	4.425	3.972082482028	0.579172015404365	-8.45	1
KC	1	-0.00935	0.316878485682368	4.528	4.2619339111593	0.571895325203252	-50.21	1
LAA	2	-0.0099	0.32339481849643	4.426	4.5652940043342	0.546438648224363	-4	0
CLE	8	0.03065	0.334295037142721	3.5315	2.9073007393523	0.553896103896104	182	1
DET	2	0.01005	0.30051001000000003	4.01500	4.01500	0.501500	0.0	0

# Close win %

win percentage in 1-run games

away_team	away_score	away_team_xwoba_diff	away_lineup_xwoba	away_team_xfp	away_starter_xfp	away_close_win_pct	away_run_diff
SF	5	-0.01395	0.324623035548816	4.3175	3.664926526202	0.481528046421663	-50.85
CHC	3	0.0179	0.320554970785412	3.9005	3.572162183634	0.577740192623914	151.35
NYY	3	0.0218	0.332433249977573	3.812	3.460086580086	0.485525467440361	123.2
PHI	4	-0.0244	0.299168151537865	4.396	4.798099532398	0.414121487028104	-129.15
SD	3	-0.0263	0.275540411693033	4.2655	4.387310590922	0.471744186046512	-160.35
COL	7	-0.01285	0.312190243126617	4.263	3.558813697769	0.521931818181818	20.41
ATL	0	-0.0131	0.309743940246347	4.576	4.56309173462	0.478513650151668	-113.95
MIA	2	-0.0131	0.330383062524465	4.627	4.596410554747	0.467974308300395	-42.9
TOR	2	0.0112	0.332435802227687	4.242	4.932603437691	0.460795638819467	1.81
PIT	3	-0.0081	0.324551829633789	4.202	3.709892473118	0.496565284692298	-29.91
SEA	0	0.0023	0.326119810644743	4.425	3.97208248202	0.579172015404365	-8.45
KC	1	-0.00935	0.316878485682368	4.528	4.261933911159	0.571895325203252	-50.21
LAA	2	-0.0099	0.32339481849641	4.426	4.565294004334	0.546438648224363	-40
CLE	8	0.03065	0.334295037142721	3.5315	2.907300739352	0.553896103896104	182
DET	2	0.01005	0.30051001000000007	4.6175	3.0000001100000007	0.5000000000000001	50.05

# Run differential

runs scored – runs allowed

away_team	away_score	away_team_xwoba_diff	away_lineup_xwoba	away_team_xfp	away_starter_xfp	away_close_win_pct	away_run_diff
SF	5	-0.01395	0.324623035548816	4.3175	3.6649265262021	0.481528046421661	-50.85
CHC	3	0.0179	0.320554970785412	3.9005	3.5721621836341	0.577740192623914	151.35
NYY	3	0.0218	0.332433249977573	3.812	3.4600865800866	0.485525467440361	123.2
PHI	4	-0.0244	0.299168151537865	4.396	4.7980995323981	0.414121487028104	-129.15
SD	3	-0.0263	0.275540411693033	4.2655	4.3873105909221	0.471744186046512	-160.35
COL	7	-0.01285	0.312190243126617	4.263	3.5588136977694	0.521931818181818	20.4
ATL	0	-0.0131	0.309743940246347	4.576	4.563091734629	0.478513650151668	-113.95
MIA	2	-0.0131	0.330383062524465	4.627	4.5964105547476	0.467974308300395	-42.9
TOR	2	0.0112	0.332435802227687	4.242	4.9326034376918	0.460795638819461	1.8
PIT	3	-0.0081	0.324551829633789	4.202	3.7098924731183	0.496565284692298	-29.9
SEA	0	0.0023	0.326119810644743	4.425	3.972082482028	0.579172015404361	-8.45
KC	1	-0.00935	0.316878485682368	4.528	4.2619339111593	0.571895325203251	-50.2
LAA	2	-0.0099	0.32339481849641	4.426	4.5652940043342	0.546438648224363	-4
CLE	8	0.03065	0.334295037142721	3.5315	2.9073007393523	0.553896103896104	182
DET	2	0.01005	0.300510010000007	4.6175	3.0000001100170	0.401500010016001	100



ESPN

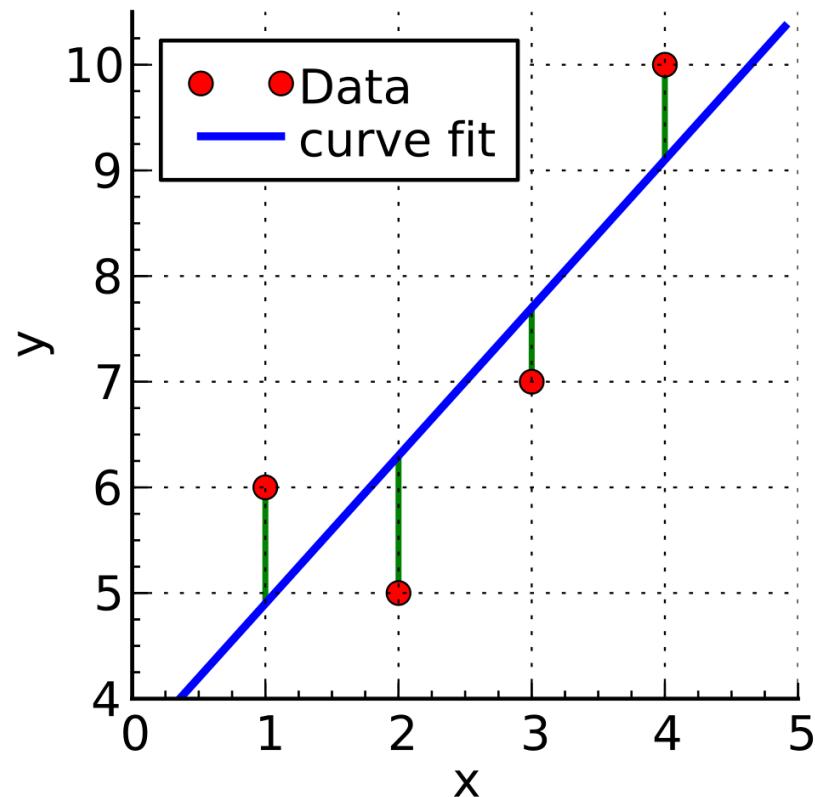


Arizona Sports

# 3

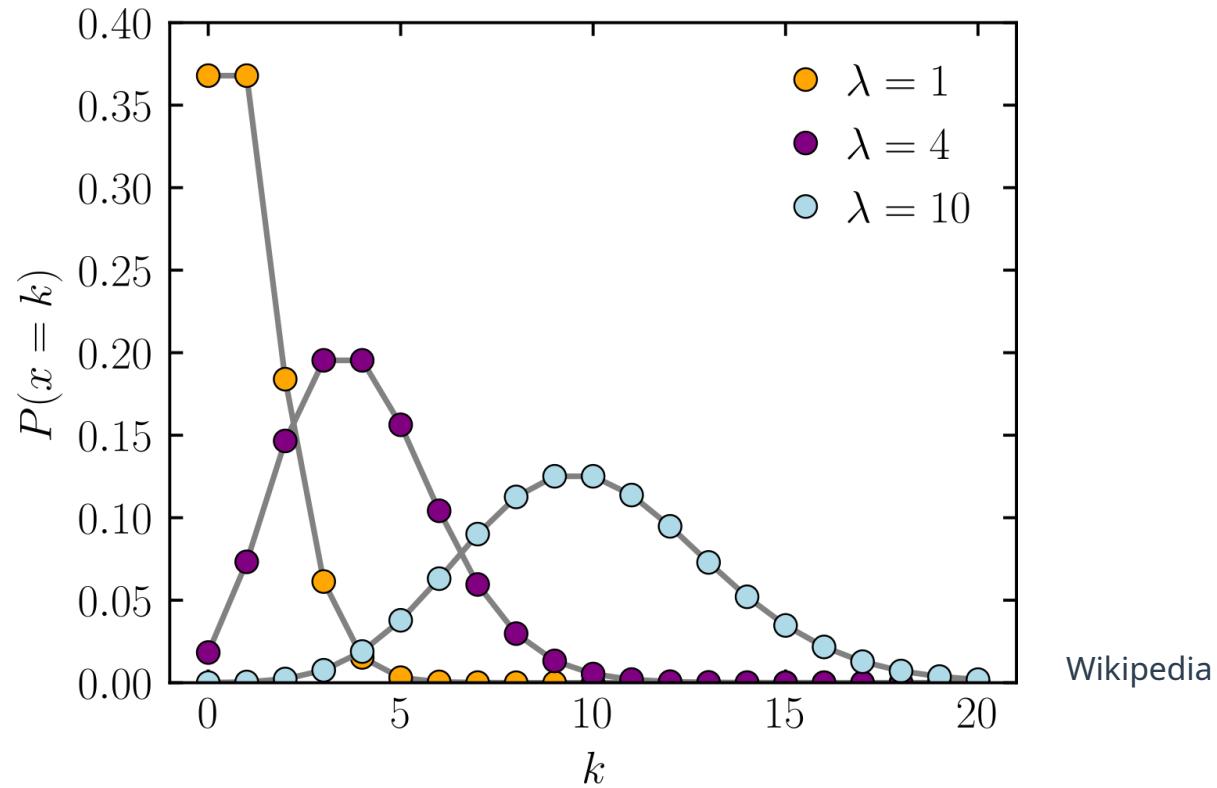
Select a model

# Linear regression

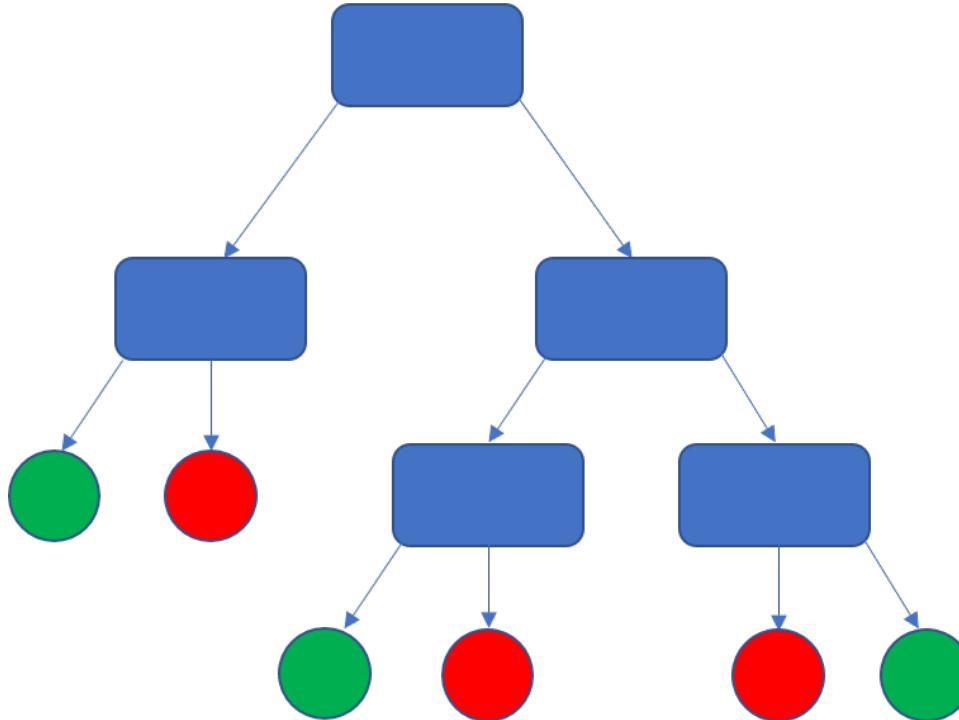


Wikimedia Commons

# Poisson regression

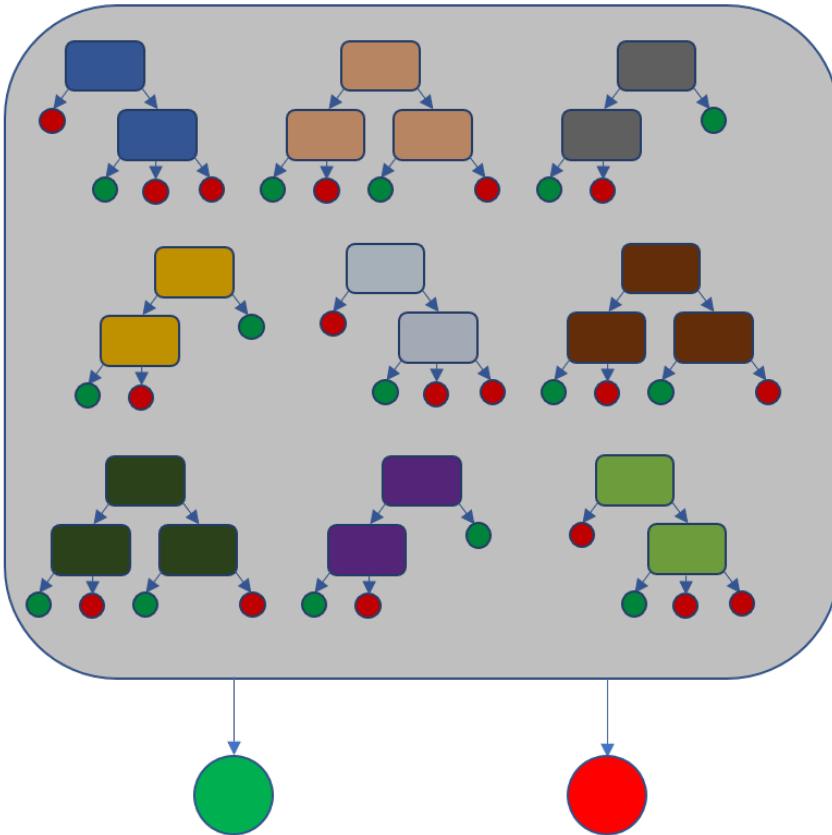


# Random forest regression



Wikimedia Commons

# Random forest regression



Wikimedia Commons

# 4

## Test

## LINEAR REGRESSION MODEL

---

Accuracy: 58.2716049382716%



## POISSON REGRESSION MODEL

---

Accuracy: 58.18930041152264%



## RANDOM FOREST REGRESSION MODEL

---

Accuracy: 58.0246913580247%



## LINEAR REGRESSION MODEL

Accuracy: 58.2716049382716%

## POISSON REGRESSION MODEL

Accuracy: 58.18930041152264%

## RANDOM FOREST REGRESSION MODEL

Accuracy: 58.0246913580247%

Wikimedia Commons



**57-59%**

Season records:

ATL: 151-11	TOR: 148-14	LAD: 138-24	SEA: 132-30	PHI: 131-31
HOU: 128-34	MIL: 123-39	MIN: 118-44	SF: 118-44	MIA: 108-54
NYY: 103-59	SD: 100-62	TB: 99-63	NYM: 96-66	BAL: 89-73
CHC: 81-81	BOS: 73-89	TEX: 70-92	DET: 62-100	STL: 61-101
CLE: 55-107	LAA: 54-108	CIN: 52-110	KC: 36-126	PIT: 31-131
AZ: 23-139	CWS: 21-141	WSH: 17-145	COL: 10-152	OAK: 2-160

Season records:

ATL: 151-11	TOR: 148-14	LAD: 138-24	SEA: 132-30	PHI: 131-31
HOU: 128-34	MIL: 123-39	MIN: 118-44	SF: 118-44	MIA: 108-54
NYY: 103-59	SD: 100-62	TB: 99-63	NYM: 96-66	BAL: 89-73
CHC: 81-81	BOS: 73-89	TEX: 70-92	DET: 62-100	STL: 61-101
CLE: 55-107	LAA: 54-108	CIN: 52-110	KC: 36-126	PIT: 31-131
AZ: 23-139	CWS: 21-141	WSH: 17-145	COL: 10-152	OAK: 2-160

Season records:

ATL: 151-11	TOR: 148-14	LAD: 138-24	SEA: 132-30	PHI: 131-31
HOU: 128-34	MIL: 123-39	MIN: 118-44	SF: 118-44	MIA: 108-54
NYY: 103-59	SD: 100-62	TB: 99-63	NYM: 96-66	BAL: 89-73
CHC: 81-81	BOS: 73-89	TEX: 70-92	DET: 62-100	STL: 61-101
CLE: 55-107	LAA: 54-108	CTN: 52-110	KC: 36-126	PTT: 31-131
AZ: 23-139	CWS: 21-141	WSH: 17-145	COL: 10-152	OAK: 2-160

Accuracy: 54.032921810699584%

---

Season records:

ATL: 101-61	PHI: 99-63	LAD: 97-65	TOR: 94-68	SF: 94-68
MIL: 90-72	HOU: 90-72	MIA: 89-73	SEA: 89-73	NYM: 87-75
BAL: 84-78	MIN: 84-78	NYY: 84-78	TB: 84-78	SD: 83-79
LAA: 83-79	CHC: 81-81	STL: 81-81	BOS: 81-81	TEX: 81-81
CIN: 79-83	CLE: 78-84	DET: 78-84	CWS: 77-85	KC: 74-88
PIT: 70-92	AZ: 69-93	WSH: 59-103	COL: 48-114	OAK: 42-120

Accuracy: 54.032921810699584%

Season records:

ATL: 101-61	PHI: 99-63	LAD: 97-65	TOR: 94-68	SF: 94-68
MIL: 90-72	HOU: 90-72	MIA: 89-73	SEA: 89-73	NYM: 87-75
BAL: 84-78	MIN: 84-78	NYY: 84-78	TB: 84-78	SD: 83-79
LAA: 83-79	CHC: 81-81	STL: 81-81	BOS: 81-81	TEX: 81-81
CIN: 79-83	CLE: 78-84	DET: 78-84	CWS: 77-85	KC: 74-88
PIT: 70-92	AZ: 69-93	WSH: 59-103	COL: 48-114	OAK: 42-120

# In conclusion...

# In conclusion...

MLB · Fri, Aug 2, 7:05 PM



4.40 : 4.59



Boston Red Sox

Texas Rangers