

G.O.A.T.S

An Elementary Adventure in Statistical Analysis



By:
Justin Wallander

Goals:

- Evaluate the greatest statistical performances of All-Time across multiple sports
- Determine whether or not the MVP for a given season generally has one of the best performances for that season/ in general

Challenges:

- Figuring out what data to run hypothesis tests on
- Keeping things simple/ focusing on the simple things

Data:

<https://www.pro-football-reference.com/years/2019/fantasy.html>

https://www.basketball-reference.com/leagues/NBA_2019_totals.html

pro-football-reference.com/years/2019/fantasy.html

Mental Modes: T... Venturky - Wind... Freight Market Se... FreightWaves Apps Script | Go... The Hitchhiker's... Imported DATA SCIENCE -... ATXIT Student Fo... Google Calendar -... g12

2019 NFL Season **Player Stats** Defensive Stats Weeks Leaders Playoffs Schedule Coaches Awards Draft Other Back to top

Fantasy Rankings

Selected to Pro Bowl, + First-Team All-Pro Share & more Glossary Toggle Per-Game Stats

Rk	Player	Tm	FantPos	Age	G	GS	Games				Passing				Rushing				Receiving				Fumbles				Scoring				Fantasy			
							Cmp	Att	Yds	TD	Int	Att	Yds	TD	Y/A	TD	Tgt	Rec	Yds	Y/R	TD	Fmb	FL	TD	2PM	2PP	FantPt	PPR	DKPt	FDPt	VBD	PosRank	OvrRank	
1	Christian McCaffrey**	CAR	RB	23	16	0	2	0	0	0	287	1387	4.83	15	142	116	1005	8.66	4	1	0	19	1	355	471.2	477.2	413.2	215	1	1				
2	Lamar Jackson**	BAL	QB	22	15	15	265	401	3127	36	6	176	1206	6.85	7	0	0	0	0	9	2	7	416	415.7	429.7	421.7	152	1	2					
3	Derrick Henry*	TEX	RB	25	15	15	0	0	0	0	303	1540	5.08	16	24	186	206	11.44	2	5	3	18	277	294.6	303.6	285.6	136	2	3					
4	Aaron Jones	GB	RB	25	16	16	0	0	0	0	236	1084	4.59	16	68	49	474	9.67	3	3	2	19	266	314.8	322.8	290.3	125	3	4					
5	Ezekiel Elliott*	DAL	RB	24	16	16	0	0	0	0	301	1357	4.51	12	71	54	420	7.78	2	3	2	14	258	317.7	319.7	284.7	117	4	5					
6	Dalvin Cook*	MIN	RB	24	14	14	0	0	0	0	250	1135	4.54	13	63	53	519	9.79	0	4	2	13	239	292.4	300.4	265.9	99	5	6					
7	Michael Thomas**	NO	WR	26	16	15	0	0	0	1	9	600	0	185	149	1725	11.58	9	1	0	9	228	274.6	277.6	300.1	97	1	7						
8	Toniia Kelce*	KAN	TE	30	16	0	0	0	0	1	4	400	1	136	97	1229	12.67	5	1	1	6	157	254.3	256.3	205.8	80	1	8						
9	Nick Chubb*	CLE	RB	24	16	16	0	0	0	0	298	1494	5.01	8	49	36	278	7.72	0	3	3	8	219	255.2	264.2	237.2	79	6	9					
10	Austin Ekeler	LAC	RB	24	16	8	0	0	0	0	132	557	4.22	3	108	92	993	10.79	8	3	2	11	217	309.0	317.0	263.0	76	7	10					
11	Mark Ingram*	BAL	RB	30	15	0	0	0	0	0	202	1018	5.04	10	29	26	247	9.50	5	2	2	15	217	242.5	250.5	229.5	76	8	11					
12	Dak Prescott	DAL	QB	26	16	16	388	596	4902	30	11	52	277	5.33	3	0	0	0	0	6	2	3	1	338	337.8	356.8	348.8	74	2	12				
13	Mark Andrews*	BAL	TE	23	15	4	0	0	0	0	0	0	0	0	0	98	64	852	13.11	10	2	1	10	143	207.2	212.2	175.2	66	2	13				
14	Russell Wilson*	SEA	QB	31	16	16	341	516	4110	31	5	75	342	4.56	3	0	0	0	0	8	2	3	1	329	328.6	341.6	333.6	65	3	14				
15	Chris Godwin*	TAM	WR	23	16	14	0	0	0	1	8	800	0	121	86	1333	15.50	9	0	0	9	190	276.1	279.1	233.1	61	2	15						
16	George Kittle**	SFO	TE	26	14	14	0	0	0	0	5	22	440	0	107	85	1053	12.39	5	1	0	5	138	222.5	225.5	180.0	61	3	16					
17	Deshon Watson*	NO	QB	24	15	15	333	485	3852	26	12	82	413	5.04	7	1	1	6	6	0	0	1	320	281.0	282.0	323.5	58	4	17					
18	Chris Carson	SEA	RB	25	15	15	0	0	0	0	278	1230	4.42	7	47	37	266	7.19	2	7	4	9	196	232.6	242.6	214.1	55	9	18					
19	Darren Waller	OKA	TE	27	16	16	0	0	0	2	5	250	0	117	90	1145	12.72	3	1	1	3	131	221.0	225.0	176.0	54	4	19						
20	Kenny Golladay*	DET	WR	26	16	16	0	0	0	0	0	0	0	0	116	65	1190	18.31	11	1	1	11	183	248.0	252.0	215.5	54	3	20					
21	Garrison Barkley	NYG	RB	22	13	13	0	0	0	0	217	1003	4.62	6	73	52	438	8.42	2	1	0	8	192	244.1	250.1	218.1	52	10	21					
22	Zach Ertz*	PHI	TE	29	15	15	0	0	0	0	0	0	0	0	0	135	88	916	10.41	6	1	1	6	128	215.6	219.6	171.6	51	5	22				
23	Joe Mixon	CIN	RB	23	15	15	0	0	0	0	0	278	1137	4.09	5	45	35	287	8.20	3	0	0	8	150	228.4	231.4	207.9	50	11	23				
24	Todd Gurley	LAR	RB	25	15	15	0	0	0	0	223	857	3.84	12	49	31	207	6.68	2	3	2	14	188	219.4	222.4	203.9	48	12	24					
25	Jarvis Cook*	NOB	TE	32	14	7	0	0	0	0	0	0	0	0	0	65	43	705	16.40	9	0	0	9	125	167.5	170.5	146.0	48	6	25				
26	Cooper Kueo	LAR	WR	26	16	14	0	1	0	0	2	4	200	1	134	94	1161	12.35	10	3	0	10	177	270.5	273.5	223.5	47	4	26					
27	Julio Jones*	ATL	WR	30	15	15	0	0	0	0	2	3	150	0	157	99	1294	14.08	6	1	0	6	175	274.1	277.1	224.6	46	5	27					
28	DeAndre Parker	SEA	WR	26	16	14	0	0	0	0	0	0	0	0	0	128	72	1202	16.69	9	0	0	9	174	246.2	249.2	210.2	45	6	28				
29	Leonard Fournette	JAX	RB	24	15	15	0	0	0	0	0	265	1152	4.35	3	100	76	322	6.87	0	1	1	3	183	259.4	266.4	221.4	43	13	29				

<

Tech:



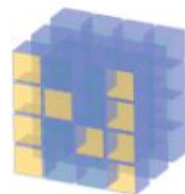
BeautifulSoup



Seaborn



pandas



NumPy

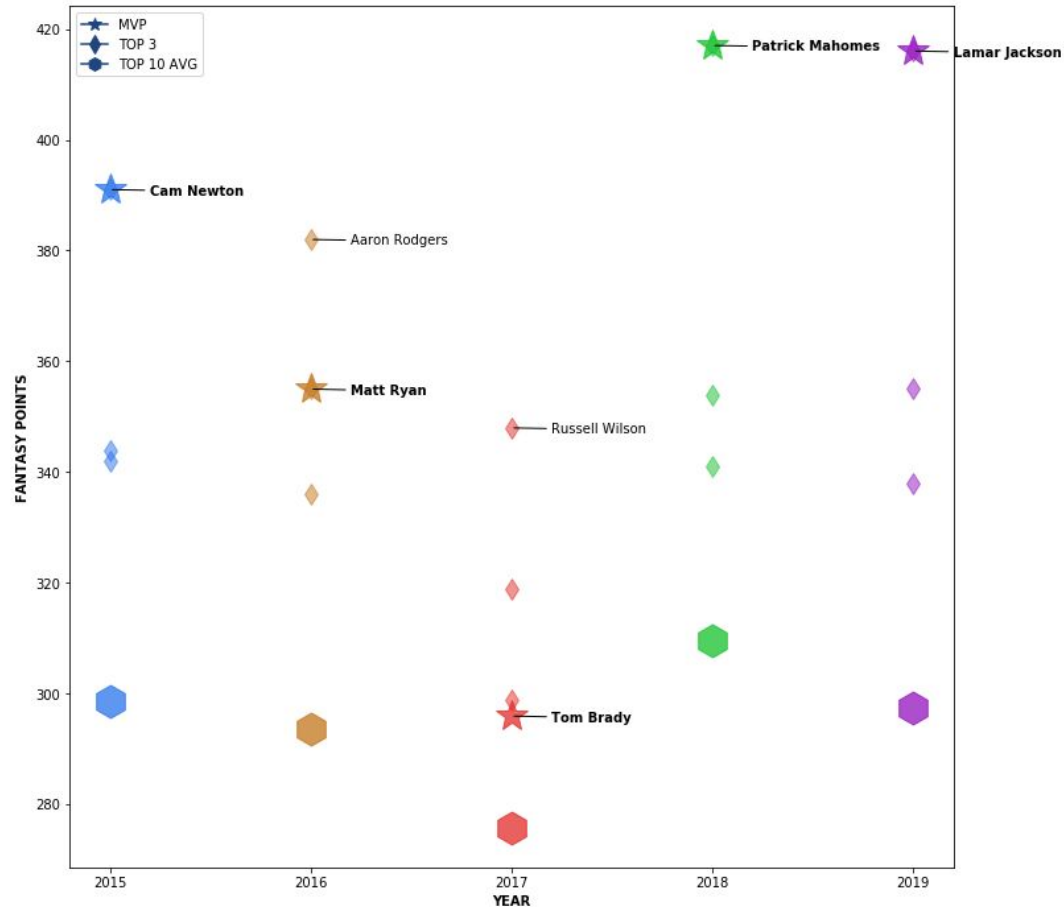
matplotlib

Experiment Design:

- Performance metric : Total Fantasy Point (basic formula) per season for each player who played 50% or greater games dating back to 1970 for NFL (as far back as database went) and 1977 for NBA (first year after NBA/ABA merger)
- Hypothesis test 1:
 - Null Hypothesis = MVP Player fantasy point production is in the top 10 percent of performances for that year 99% of the time.
 - Alternate Hypothesis = MVP Player fantasy point production is in the top 10 percent of performances for that year less than 99 % of the time.

Hypothesis Test 1:

NFL TOP Fantasy Point Production by Year

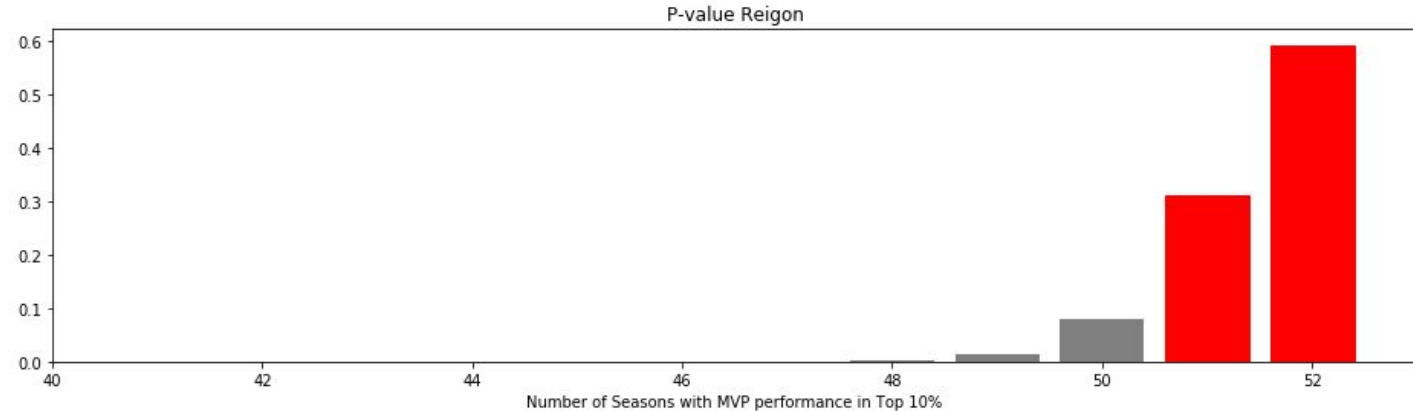
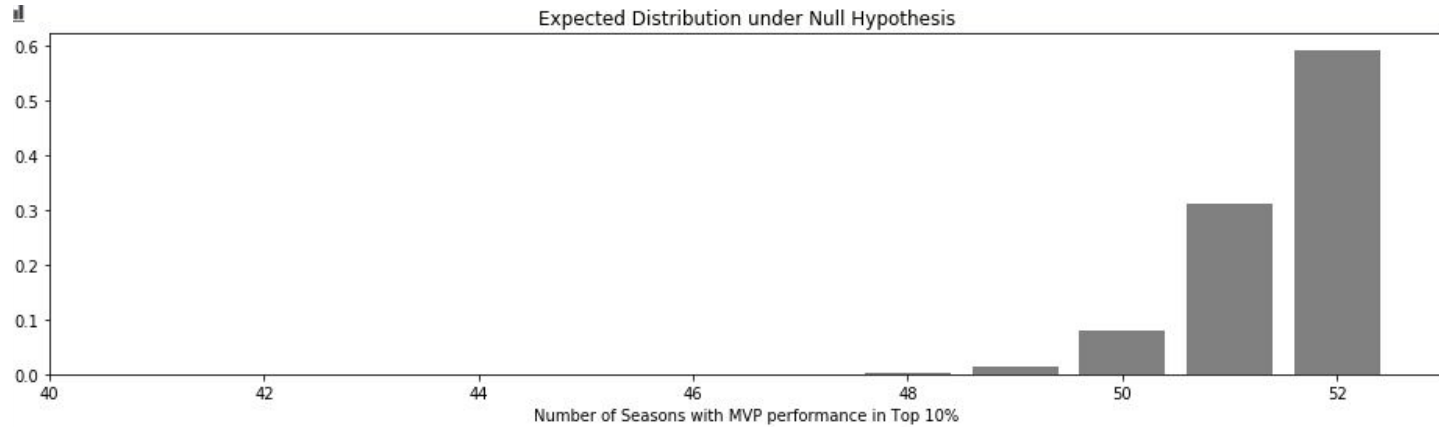


Hypothesis Test 1: Binomial Testing

P-value = 0.9

Alpha = 0.05

Result: I do not reject the null hypothesis

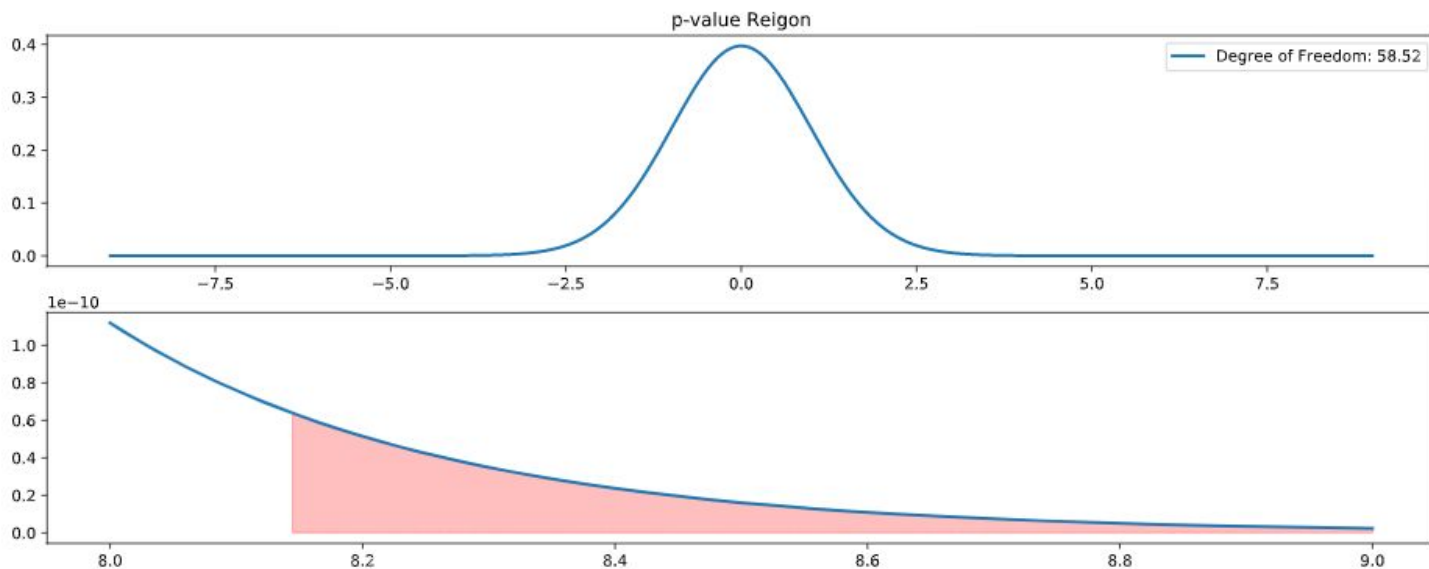


Experiment Design:

- Performance metric : Total Fantasy Point (basic formula) per season for each player who played 50% or greater games dating back to 1970 for NFL (as far back as database went) and 1977 for NBA (first year after NBA/ABA merger)
- Hypothesis test 2:
 - Null Hypothesis = There is no difference between mean mvp performance and the mean performance of the top 10 percent.
 - Alternate Hypothesis = There is a difference, mean MVP performances are greater than the mean Top 10 percent.

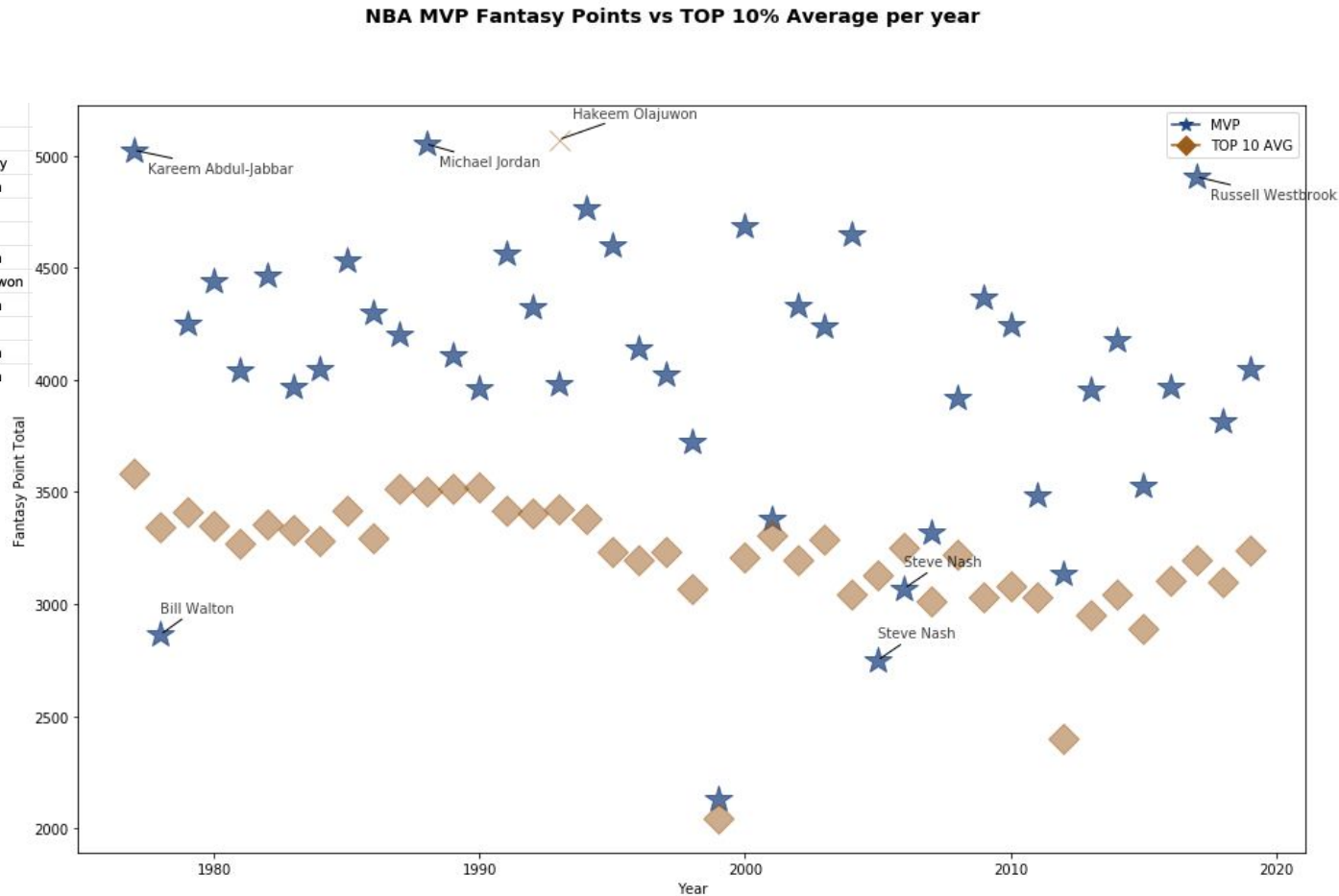
Hypothesis Test 2: Welsh's T-Test

- Test Stat = -8.14 Degrees of Freedom = 58.52 Alpha = 0.05
- The P-value for H_0 , different average performance between MVP and Top 10 percent: 0.000000000033
- The P-value for average MVP performance being greater than average top 10 percent performance: 0.000000000017
- I reject the Null Hypothesis and can conclude that the Average MVP performance is greater than the average performance in the Top 10%



Hypothesis Test 2:

Top 10 Performances:			
Player:	Year:	FantPt:	MVP:
Hakeem Olajuwon	1993	5072.1	False
Hakeem Olajuwon	1990	5058.8	False
Michael Jordan	1988	5052.3	True
Kareem Abdul-Jabbar	1977	5024.5	True
Michael Jordan	1989	4997.4	False
David Robinson	1994	4939.5	False
Michael Jordan	1987	4933.5	False
Russell Westbrook	2017	4905.8	True
Michael Jordan	1990	4805.5	False
Hakeem Olajuwon	1989	4793.5	False



Next Steps:

- Add MLB and College Sports
- Use different Performance Metrics/ Points of Comparison
- Rework web scraper to build a more robust database of stats
- ULTIMATE STEP: Use that database to build a machine learning model to help with fantasy drafts

Thanks!

Email : justin.wallander@gmail.com

Github: <https://github.com/justin-wallander>

LinkedIn: <https://www.linkedin.com/in/justin-wallander/>

