Analyzing the Modern NFL Running Back

Mike Gallaspy and Kevin Kannappan

DSE 241 Final Project Proposal

January 31, 2020

Dataset: Kaggle's "Big Data Bowl"

NFL's "Next Gen Stats"

- Taken from Kaggle for a competition to predict yards gained
- "Running" or "rushing" plays from 2017- mid 2019
    - Hand-off was made to either a Running Back (RB), Wide Receiver (WR) or Tight End (TE)
- Contains rich information:
    - Player pre-snap X,Y coordinates
        - Offensive/Defensive Formations
    - Player acceleration and speed
    - Play direction and orientation
    - Play result

Rushing * Selected to Pro Bowl, + First-Team All-Pro Share & more ▾ Hic

					Games		Rushing						
Rk	Player	Tm	Age	Pos	G	GS	Att	Yds ▼	TD	1D	Lng	Y/A	Y/G
1	Derrick Henry*	TEN	25	RB	15	15	303	1540	16	73	74	5.1	102.7
2	Nick Chubb*	CLE	24	RB	16	16	298	1494	8	62	88	5.0	93.4
3	Christian McCaffrey*+	CAR	23	RB	16	16	287	1387	15	57	84	4.8	86.7
4	Ezekiel Elliott*	DAL	24	RB	16	16	301	1357	12	78	33	4.5	84.8
5	Chris Carson	SEA	25	RB	15	15	278	1230	7	75	59	4.4	82.0
6	Lamar Jackson*+	BAL	22	QB	15	15	176	1206	7	71	47	6.9	80.4
7	Leonard Fournette	JAX	24	RB	15	15	265	1152	3	55	81	4.3	76.8
8	Josh Jacobs	OAK	21	RB	13	13	242	1150	7	53	51	4.8	88.5
9	Joe Mixon	CIN	23	RB	16	15	278	1137	5	56	41	4.1	71.1
10	Dalvin Cook*	MIN	24	RB	14	14	250	1135	13	60	75	4.5	81.1
11	Marlon Mack	IND	23	RB	14	12	247	1091	8	67	63	4.4	77.9
12	Aaron Jones	GNB	25	RB	16	16	236	1084	16	55	56	4.6	67.8
13	Carlos Hyde	HOU	29	RB	16	14	245	1070	6	55	58	4.4	66.9
14	Mark Ingram*	BAL	30	RB	15	15	202	1018	10	53	53	5.0	67.9
15	Phillip Lindsay	DEN	25	RB	16	16	224	1011	7	38	40	4.5	63.2
16	Saquon Barkley	NYG	22	RB	13	13	217	1003	6	45	68	4.6	77.2
17	Sony Michel	NWE	24	RB	16	14	247	912	7	55	26	3.7	57.0
18	Adrian Peterson	WAS	34	RB	15	15	211	898	5	41	32	4.3	59.9
19	David Montgomery	CHI	22	RB	16	8	242	889	6	50	55	3.7	55.6
20	Todd Gurley	LAR	25	RB	15	15	223	857	12	51	25	3.8	57.1

What do we already know about running backs?

Predominately, they “run” or “rush” with the ball to gain yardage to score touchdowns.

Rushing Yards & Touchdowns seem to be primary drivers for Pro Bowl selection

Dataset allows for us to visualize the *how*

Proposed tasks both explore and present new information for analysts and the public alike

1: RB Speed

Who are consistently the fastest RBs?

2: RB Direction

How well do RBs rush to both sides of the center?

3: Defenses

How do defensive formations affect the RB?

4: All together

Can we see a rich combination of data for RB rushes?

5: Analysis

Can we visually propose an alternative hypothesis for Pro Bowl consideration?

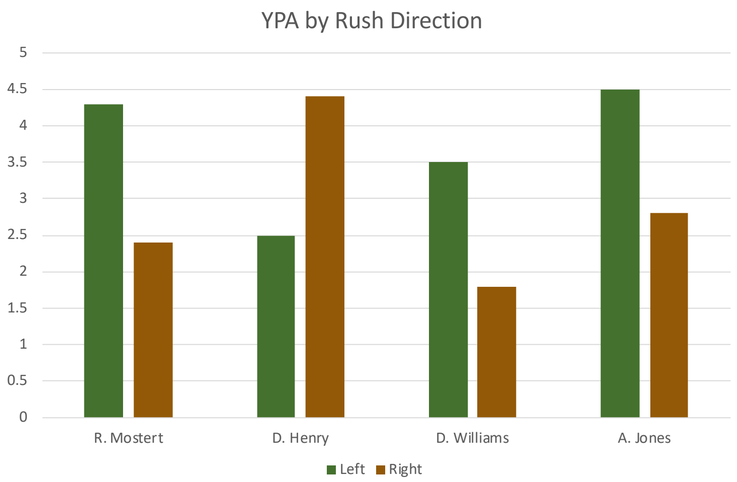
Pre-requisite steps

01	02
Reduce Dataset	Create rush vectors
Create rush-level dataset by dropping peripheral snap-level information	Leverage X, Y coordinates, speed, direction, etc. to develop rush vectors to plot

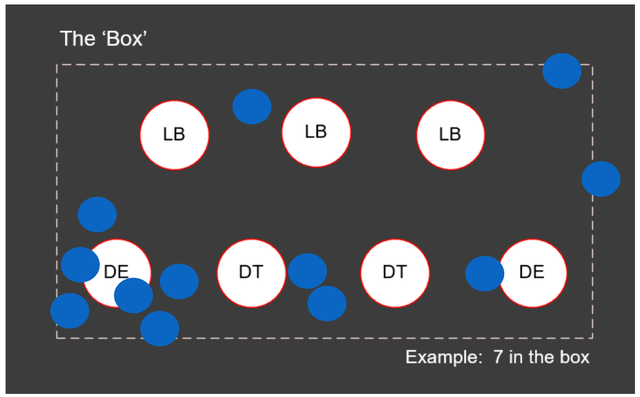
Programmatic Solution: Leveraging filtering, sorting and movement

Some theoretical plots & sketches…

Tasks 1 & 2:



Task 3:



May zoom or filter for specific set or specific players. Lines and dots may move.

Tasks 4 & 5: