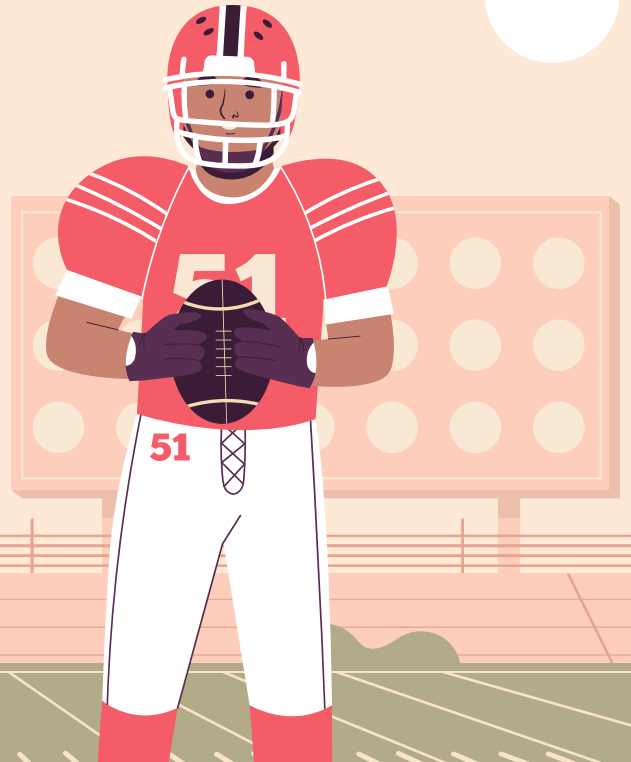
GROUND GAME METRICS: STATISTICAL EVALUATION OF RUNNING BACKS

Jordan Britton, Abigail Finsten, Andrew Ebert



INTRODUCTION

MOTIVATION

We chose to investigate running back numbers and how age plays a significant role in their performance.
The running back position has been an integral part of NFL offenses since the sport was invented. However we've seen a major devaluation of the position in the last decade:

- Franchise tags being used to extend their contracts instead of long-term deals (with elite running backs even having a hard time getting signed to major deals)
- 4 drafts in the last decade not featuring a running back taken in the first round
- Guys being released/traded after career-defining seasons
- "Running back by committee" taking over the league so no one is heavily featured/relied upon

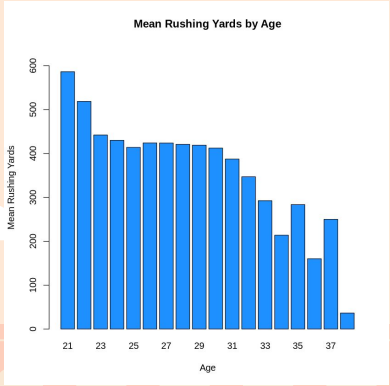
We've all heard about the saying that running backs "hitting a wall" at 30 years old. This has been circulating for a long time, but why is the devaluation happening to the position so far before that point? Does the fall off for running backs happen much earlier than we think, and have teams noticed this?

INTRODUCING THE DATA

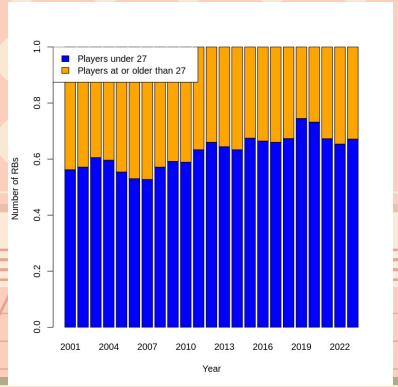
Our dataset consisted of all the full-season statistics for all running backs from 2001 to 2023. The sample is not a random sample, however it is a sample of the total rushing seasons throughout the history of the NFL, in order to track the importance of the run game through time. The data was presumably taken from the NFL official site, and aggregated into a CSV file, which we loaded into a publicly accessible folder in Drive so we could easily manipulate it.

	Player_Name	Age	Games_Played	Games_Started	Rushing_Attempts	Rushing_Yards	Rushing_Touchdowns	Rushing_First_Downs	Longest_Rush	Yards_per_Attempt	Yards_
	<chr>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<dbl>	
1	Stephen Davis	27	16	16	356	1432	5	75	32	4.0	
2	Corey Dillon	27	16	16	340	1315	10	69	96	3.9	
3	LaDainian Tomlinson	22	16	16	339	1236	10	68	54	3.6	
4	Curtis Martin	28	16	16	333	1513	10	78	47	4.5	
5	Priest Holmes	28	16	16	327	1555	8	81	41	4.8	
6	Eddie George	28	16	16	315	939	5	40	27	3.0	

HIGHLIGHTS FROM EDA



Data on the average rushing yards per age. The interesting part is how the average for rushing yards for 21-year-olds is considerably higher than that of any other age group, then it all seems to stabilize around 400 yards. Then we see a drop at around 30 that continues until the edge of the graph.

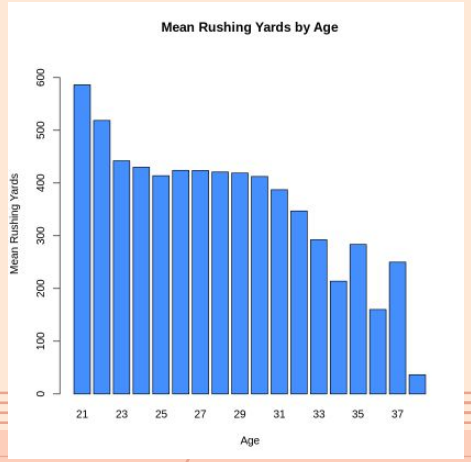


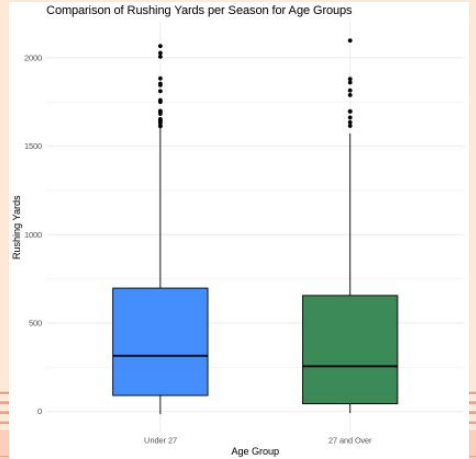
From the chart to the left, we can see that the proportion of 27+ year old running backs has been shrinking consistently over the last 22 years.

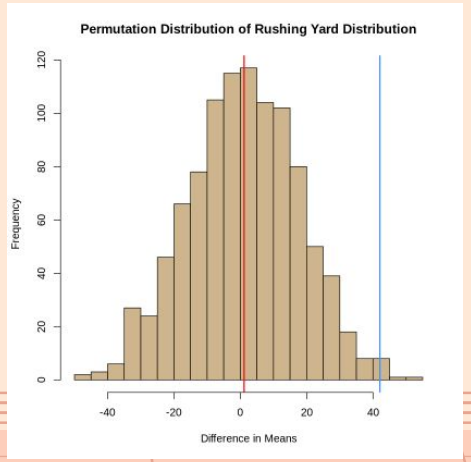
DIFFERENCE IN TWO MEANS

Do running backs under 27 have more yards in a season, on average, than running backs 27 and over?

- H0: Running backs under 27 rush for as many yards per season on average compared to running backs 27 or older: $\mu_{under27} = \mu_{27+}$
- Ha: Running backs under 27 rush for more yards per season on average compared to running backs 27 or older: $\mu_{under27} > \mu_{27+}$







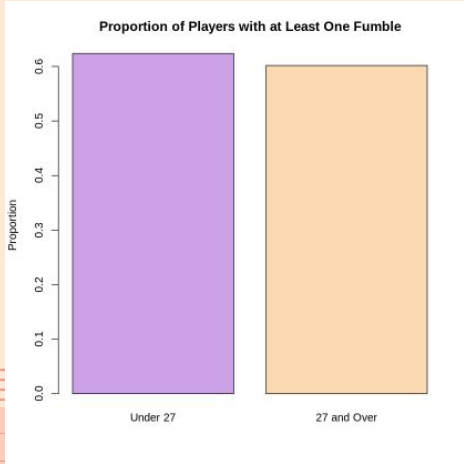
Permutation Test:
p-val = 0.009 < 0.05 :
Reject H0, accept Ha

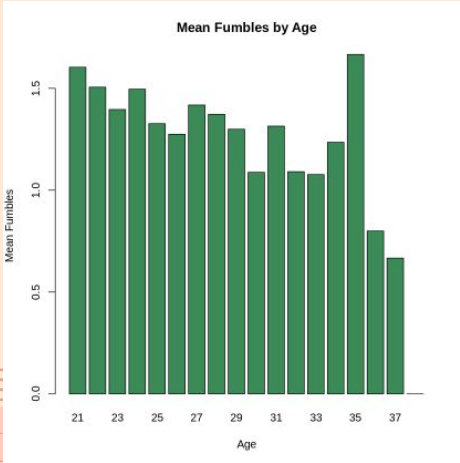
Parametric Test:
95% confidence that the interval 9.451 < μ < 74.693 contains the difference in means for the population
****Interval doesn't contain 0****
p-val = 0.012

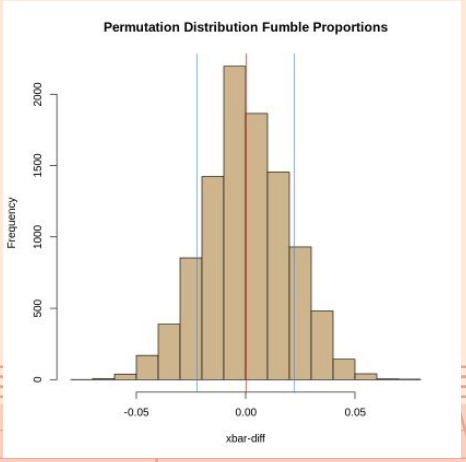
DIFFERENCE IN TWO PROPORTIONS

Do running backs under 27 have a greater proportion of having a fumble than running backs 27 or older?

- H0: the proportion of running backs under 27 who have at least one fumble is the same as the proportion of running backs 27 or older who have at least one fumble: $p_{fumble\_over\_27} = p_{fumble\_under\_27}$
- Ha: the proportion of running backs under 27 who have at least one fumble is greater than the proportion of running backs 27 or older who have at least one fumble: $p_{fumble\_over\_27} \neq p_{fumble\_under\_27}$







Permutation Test:
p-val = 0.142 > 0.05 :
Fail to reject H0, test is inconclusive

Parametric Test:
95% confidence that the interval -0.016 < p < 0.060 contains the difference in proportions for the population
****Interval contains 0****
p-val = 0.255

RESULTS

CONCLUSION

From the analysis done, we can see that there is a difference in performance for running backs as they get older, with a significant change happening around 27 years of age(in regards to rush yards).

FUTURE WORK

The next steps could be to study the performance of running backs around the same age. We saw from an analysis done in a previous section that running backs average somewhere around 3.5-5 yards per carry independent of age, which isn't really a large margin of performance. So how much better is an "elite" running back in comparison to one that is average? The difference might be low, considering how few NFL teams want to lean on "feature backs" today, so this would be an interesting next step in understanding the evolution of the position and the league as a whole.