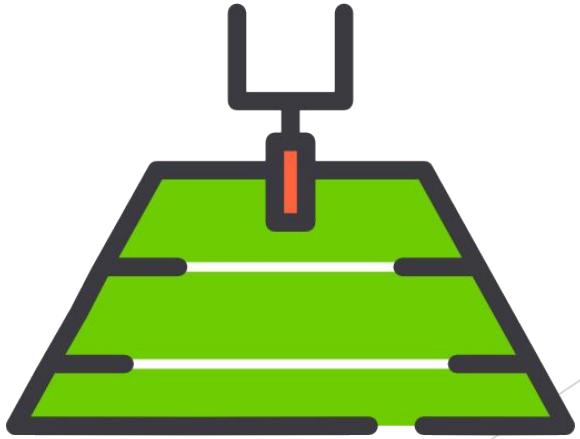
Group Project 2: NFL

Kellen Brown, Mason Griffith, Reed Stout, Sahil Khatri, Daniel Shin

Dataset

Pro Football Focus is a company dedicated to providing play by play data as well as analytics for fans and professionals alike
This dataset includes statistics from the first 6 weeks of the 2021 NFL Season

This set specifically contains data ranging from game situation to play information to outcomes
This allows for insight into which type of plays are most successful, the ability to compare teams, and a way to see how each team matches up.



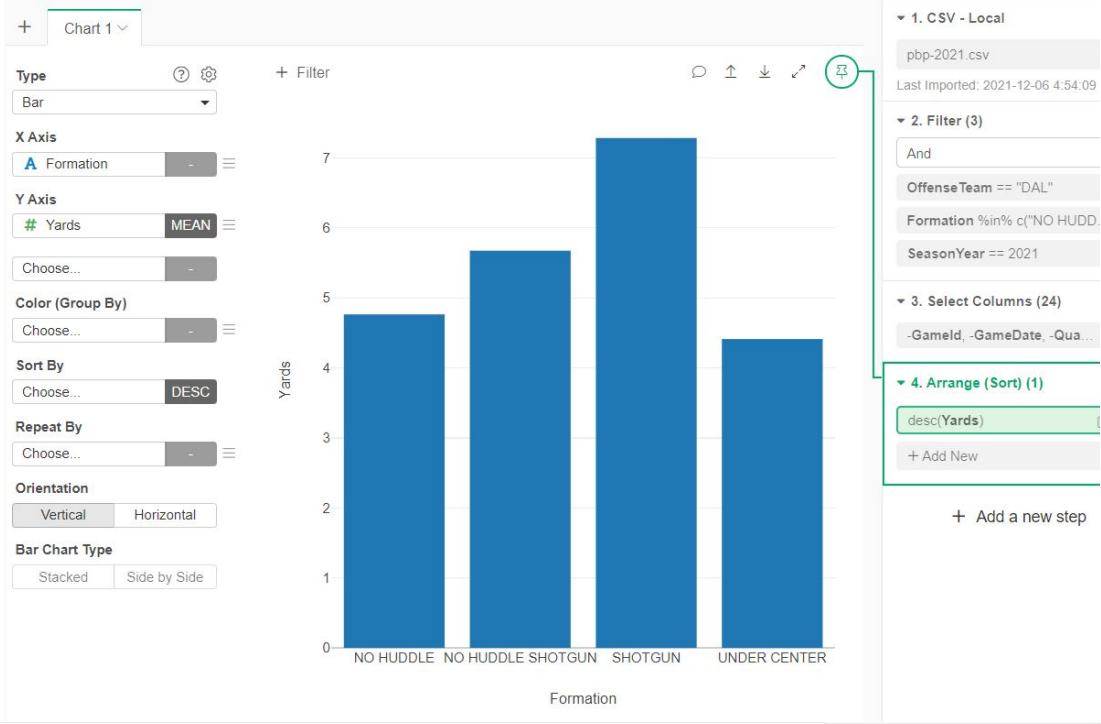
Question 1

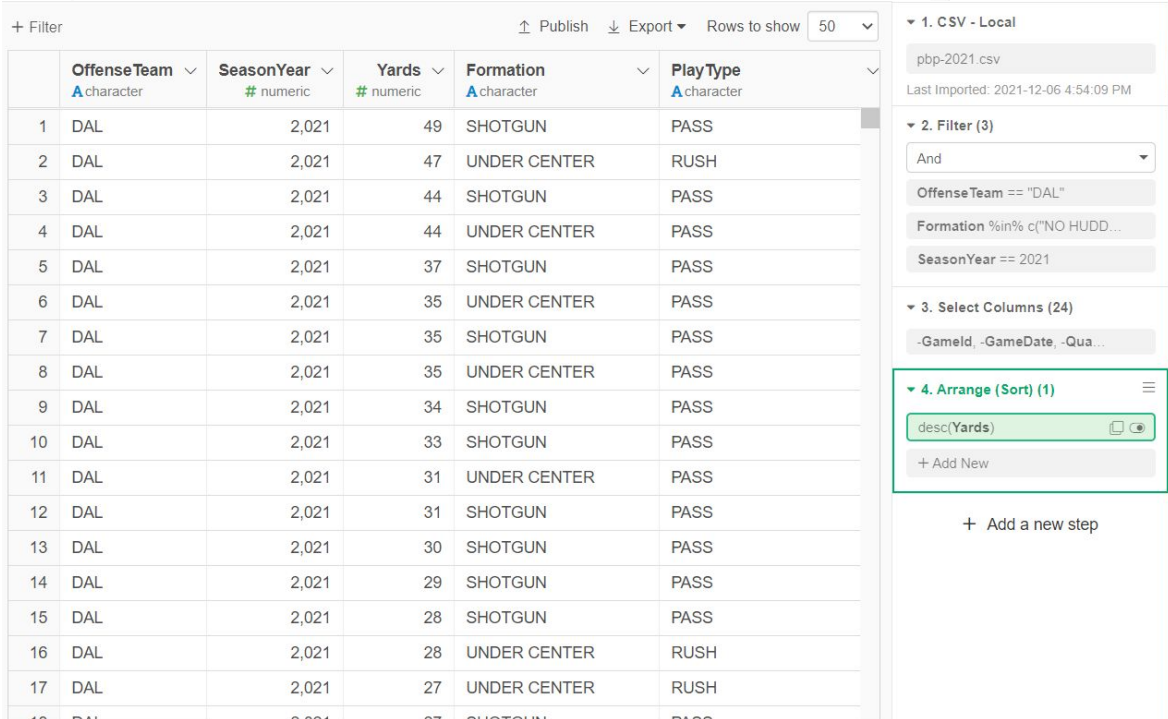
- What is the average yards per play in each formation for Dallas this season?
  - Formations split into Shotgun, Under Center, No Huddle Shotgun, No huddle, and Under Center
  - Useful for teams playing Dallas and for The Cowboys themselves
  - Lets teams playing them know which formations to prioritize in film
  - Lets Cowboys coaches know where they are strong and where they could improve in practice
  - No huddle stats lets teams know where they are strong in a time crunch

Dallas Cowboys Formations

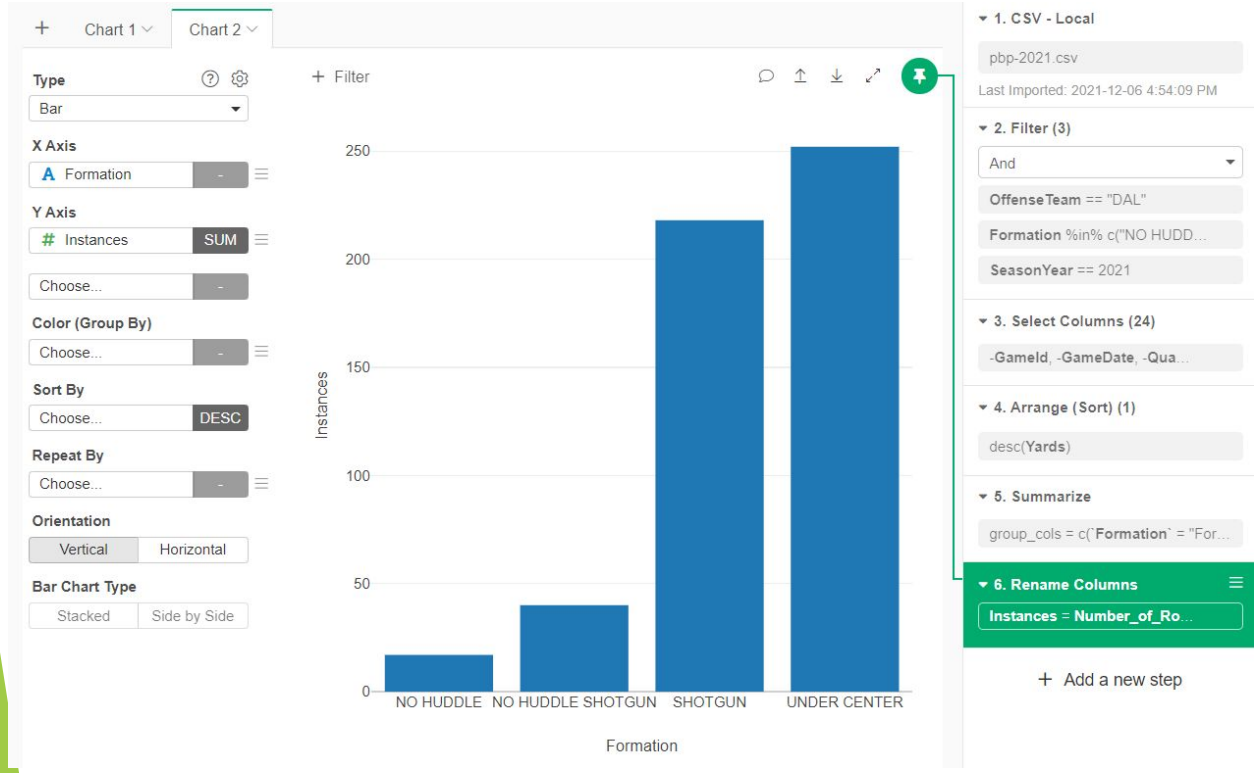
	Shotgun	Under Center	Expected Change
Huddle	+1.57	-1.3	+.135
No Huddle	-.04	-.95	-.495
Expected Change	+.765	-1.125	

*Numbers compared to average overall yards



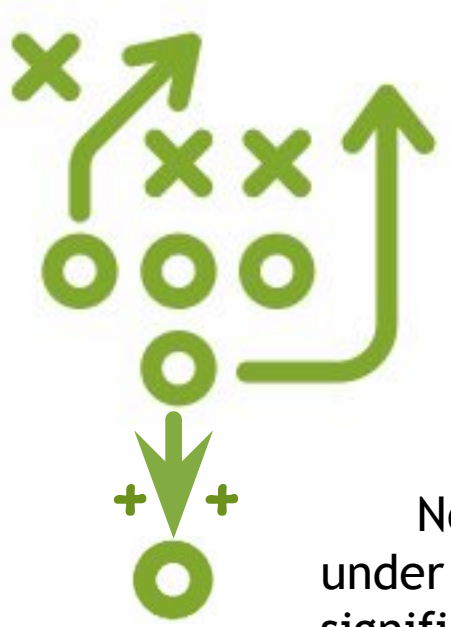


Additional Info



- We also thought that it would be interesting to see statistics on how often each formation appeared in game.
- The results are about what we expected, with No Huddle formations being significantly less frequent. Interestingly Shotgun plays were called less than under center ones despite having higher average yards.

Interpretation



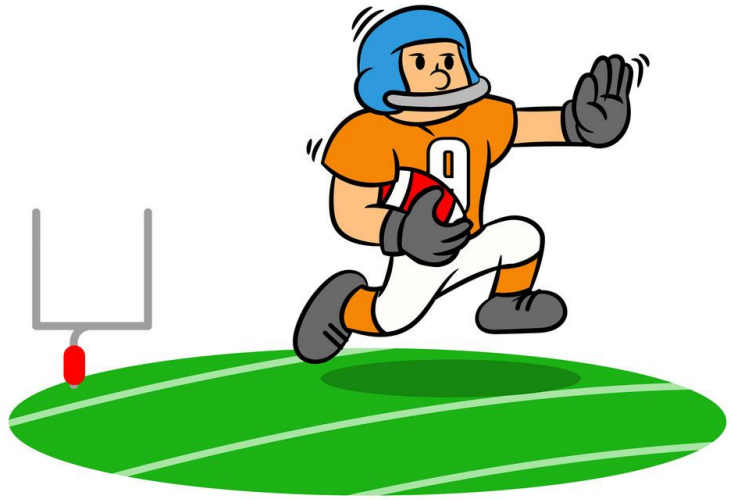
The results of our data manipulation showed us that the shotgun formation proved to result in the highest average yards and Under Center was the least.
This makes sense as most plays from the former are passes, however what is more interesting is the statistics from the no huddle plays.

Not only are under center plays more successful under these circumstances but no huddle shotguns are significantly less so.
In situations where you would call no huddle plays, time is most likely a concern.

Conventional wisdom says to pass here, which either means that the Cowboys have passed more from under center in no huddle or their runs are more successful. Either option provides insight into the Cowboys' strategy this season.

Question 2

- What factors have the largest impact on how successful a rush attempt will be for the NFL as a whole?
  - Factors to include are quarter, down, yards to go, yard line, rush direction, and yard line direction (side of the field)
  - This would allow teams to learn what to do in order to be more successful at rushing, whether that is which direction to go or what time of game to do it more
  - Conversely, it tells teams what strategies to avoid as well



Rushing Yard Factors

Summary of Fit — Showing 1 row

Null Deviance	DF for Null Model	Log Likelihood	AIC	BI
151,524.866695996	4,545	-14,371.9388220801	28,769.8776441603	28,853.3636832176

Parameter Estimates — Showing 12 rows

Term	Coefficient	Std Error	t Ratio	
(Intercept)	4.404809655569	0.5676340653004	7.759945931431	1.0426
Quarter	-0.07687294394489	0.07368829775495	-1.043217801021	0.2
Down	0.1027849859883	0.1481850370769	0.693625942375	0.4
YardLine	-0.03105544138749	0.006637352934968	-4.678889565128	0.000002
ToGo	0.09725398908326	0.02962516138948	3.282817190585	0.001
RushDirectionRIGHT GUARD	0.422034502971	0.2846105335955	1.482849203223	0.1
RushDirectionLEFT GUARD	0.3436356105033	0.2892544744915	1.188004476361	0.
RushDirectionRIGHT TACKLE	0.2933138251896	0.2953306680183	0.9931708994456	0.3
RushDirectionLEFT TACKLE	0.4987657100161	0.3005852774129	1.659315167758	0.09
RushDirectionLEFT END	1.205184013805	0.303945176907	3.965136167216	0.00007
RushDirectionRIGHT END	0.876531209029	0.3054197081707	2.869923536628	0.004
YardLineDirectionOPP	0.737430153404	0.3366182750456	2.190701480197	0.02

1. CSV - Local

pbp-2021.csv

Last Imported: 2021-12-07 1:09:00 PM

2. Filter (2)

And

SeasonYear == 2021

PlayType %in% c("RUSH")

3. Select Columns (19)

GameDate, Quarter, Offens...

4. Build GLM

Yards ~ Quarter + Down + YardLi...

+ Add a new step

Interpretation

The GLM we built is used to predict rushing yards on any given play for the NFL given several explanatory variables (quarter, down, yards to go, yardline, yardline direction, and rushing direction). The base run in our model is starting on your own side of the field and rushing up the middle for an average of 4.4 yards.

Statistically, the significant factors with a p value less than .05 are YardLine, yards to go, rushing at the left end, rushing at the right end, and yardline direction. Even so, the only factors that have large coefficients are rushing toward the left end, right end, and being on the opponent's half of the field.

Overall, rushing at the left defensive end appears to be the most effective strategy to increase yards per rush, and this could be compounded by doing it in the opponent's half. One reason we thought of for why rushing left could be more successful is because most running backs are right-handed, and this could make it easier for them to stiff-arm on the left (like Derrick Henry above).

Summary Conclusions

- **Dallas Cowboys Formations:** Plays called from the shotgun result in more yards per play than under center while going no huddle results in less yardage for shotgun plays but more for under center plays.
- **Rushing Yard Factors:** The three most important factors to increasing rushing yards per play are rushing at the left defensive end, right defensive end, and being in the opponent's half. The most effective is rushing at the left end, which increases average yards a carry by 1.21. This is very good to know as sometimes games come down to a single yard.