

# A Comprehensive Analysis of Cohesive Fantasy Team Selections

Anmol Sapru

```
library(tidyverse)
library(nflreadr)
library(nflfastR)
library(ggcorrplot)
library(ggimage)
library(patchwork)
library(gt)
```

To start off, the data from the BBM II and BBM III were loaded in. These offer a pretty good understanding of trends in the BBM landscape, while both having more data than can be found in the original BBM data set. When loading in the data, the BBM Ill data set had the playoff team variable be mutated to match what was offered by the same variable in the BBM II data set.

The roster data were added in from the nflread package, taking the information from Week 1 in order to best capture what drafters would have been thinking about when making picks for their stacks. There were some issues in matching some names and positions for hybrid players between this roster data and the BBM data set, so these were quickly fixed in a mutate statement and then the two joined.

```
rosters_2021 <- load_rosters_weekly(2021) |>
  filter(week == 1) |>
  select(team, position, full_name) |>
  mutate(full_name = case_when(
    full_name == "A.J. Dillon" ~ "AJ Dillon",
    full_name == "Christopher Herndon" ~ "Chris Herndon",
    full_name == "DJ Chark" ~ "D.J. Chark",
    full_name == "D.J. Moore" ~ "DJ Moore",
    full_name == "Jeffery Wilson" ~ "Jeff Wilson",
```

```

full_name == "Josh Palmer" ~ "Joshua Palmer",
full_name == "K.J. Hamler" ~ "KJ Hamler",
full_name == "Will Fuller" ~ "William Fuller",
full_name == "Jamycal Hasty" ~ "JaMycal Hasty",
.default = full_name
)) |>
mutate(position = case_when(
  full_name == "Cordarrelle Patterson" ~ "WR",
  full_name == "Juwan Johnson" ~ "WR",
  position == "FB" ~ "RB",
  .default = position
))

rosters_2022 <- load_rosters_weekly(2022) |>
filter(week == 1) |>
select(team, position, full_name) |>
mutate(full_name = case_when(
  full_name == "A.J. Dillon" ~ "AJ Dillon",
  full_name == "Christopher Herndon" ~ "Chris Herndon",
  full_name == "D.J. Chark" ~ "DJ Chark",
  full_name == "D.J. Moore" ~ "DJ Moore",
  full_name == "D'Wayne Eskrid" ~ "Dee Eskrid",
  full_name == "Jeffery Wilson" ~ "Jeff Wilson",
  full_name == "Joe Fortson" ~ "Jody Fortson",
  full_name == "Josh Palmer" ~ "Joshua Palmer",
  full_name == "K.J. Hamler" ~ "KJ Hamler",
  full_name == "Mitchell Trubisky" ~ "Mitch Trubisky",
  full_name == "Robby Anderson" ~ "Robbie Anderson",
  full_name == "Scott Miller" ~ "Scotty Miller",
  full_name == "Will Fuller" ~ "William Fuller",
  full_name == "Jamycal Hasty" ~ "JaMycal Hasty",
  .default = full_name
)) |>
mutate(position = case_when(
  full_name == "Taysom Hill" ~ "TE",
  position == "FB" ~ "RB",
  .default = position
))

BBMII_regular_season <- left_join(BBMII_regular_season, rosters_2021,
join_by(player_name == full_name, position_name == position))

```

```
BBMIII_regular_season <- left_join(BBMIII_regular_season, rosters_2022,
                                     join_by(player_name == full_name, position_name == position))
```

Adding these teams to the BBM data set was helpful for determining which NFL teams were being stacked in fantasy teams. A list of tournament\_entry ids was collected to filter the data used later on.

Averaging the points scored by each player for their team during the regular season gives a good idea of how well each player does for the fantasy team they were drafted too. Summing up these points score by position gives a position group score for each team. The correlation point below shows a variance in how related the scores for positions were for both BBM seasons, varying between positions from basically none up to a decently strong positive correlation between some position groups scoring on their teams. This leads into the idea of stacking players on the same team, a strategy that can make or break scoring though since a bad game for one player on a team could likely mean that the same is happening for others on the team.

```
BBMII_team_position_scores <- BBMII_regular_season_stack_players |>
  group_by(player_name) |>
  summarize(pick_points = mean(pick_points)) |>
  left_join(rosters_2021, join_by(player_name == full_name)) |>
  group_by(team, position) |>
  mutate(team_position_score = sum(pick_points)) |>
  arrange(team, position) |>
  group_by(team, position) |>
  summarize(team_position_score = mean(team_position_score), .groups = "drop") |>
  pivot_wider(names_from = position, values_from = team_position_score) |>
  filter(!is.na(team)) |>
  select(QB, RB, WR, TE)

BBMIII_team_position_scores <- BBMIII_regular_season_stack_players |>
  group_by(player_name) |>
  summarize(pick_points = mean(pick_points)) |>
  left_join(rosters_2021, join_by(player_name == full_name)) |>
  group_by(team, position) |>
  mutate(team_position_score = sum(pick_points)) |>
  arrange(team, position) |>
  group_by(team, position) |>
  summarize(team_position_score = mean(team_position_score), .groups = "drop") |>
  pivot_wider(names_from = position, values_from = team_position_score) |>
  filter(!is.na(team)) |>
```

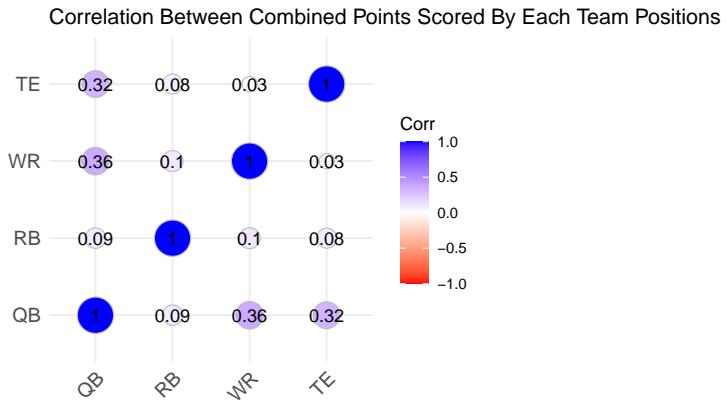
```

select(QB, RB, WR, TE)

combined_team_position_scores <-
  bind_rows(BBMII_team_position_scores, BBMIII_team_position_scores)

ggcorrplot(cor(combined_team_position_scores),
  method = "circle",
  title = "Correlation Between Combined Points Scored By Each Team Positions",
  colors = c("red", "white", "blue"),
  lab = TRUE)

```



To get a better idea of what stacks are most common, the list of stacked teams from earlier was utilized. With 18 rounds in BBM, it's not too surprising that nearly every team has at least one stack and many have multiple. The most common stacks seem to trend towards the better teams in the league and-possibly more generally with teams that have a top player at a position.

```

image_urls <- load_teams()

player_count_graph_data <- bind_rows(stacked_teams_BBMII, stacked_teams_BBMIII) |>
  filter(!is.na(team)) |>
  count(team) |>
  left_join(image_urls, join_by(team == team_abbr)) |>
  select(team, n, team_division, team_logo_espn)

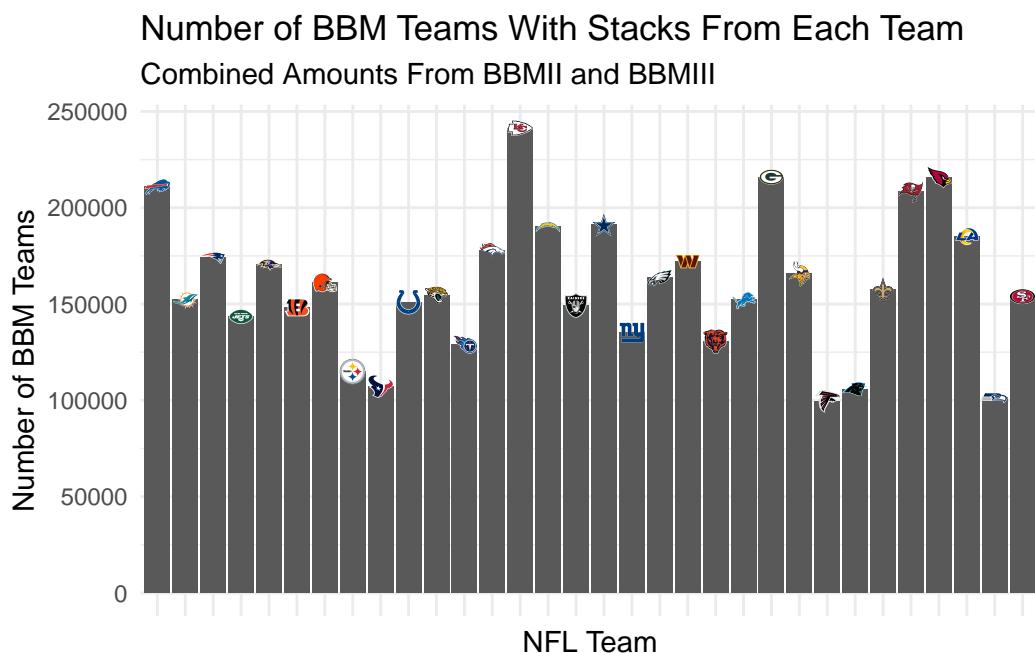
division_order <- c("AFC East", "AFC North", "AFC South", "AFC West",
  "NFC East", "NFC North", "NFC South", "NFC West")

```

```

player_count_graph_data$team <- factor(player_count_graph_data$team,
                                         levels = unique(player_count_graph_data$team)
                                         [order(match(player_count_graph_data$team_division, division_order))]
                                         )
ggplot(player_count_graph_data, aes(x = team, y = n)) +
  geom_col() +
  geom_image(aes(image = team_logo_espn)) +
  labs(
    title = "Number of BBM Teams With Stacks From Each Team",
    subtitle = "Combined Amounts From BBMII and BBMIII",
    x = "NFL Team",
    y = "Number of BBM Teams"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_blank(),
        axis.ticks.x = element_blank())

```



Another way to investigate the stacks is how the frequency of positions within in each team's stacks vary. These differences can play a role into the types of stacks that exist and the idea of general variances in stacking results by team.

```

BBMII_regular_season <- BBMII_regular_season |>
  mutate(year = 2021)

BBMIII_regular_season <- BBMIII_regular_season |>
  mutate(year = 2022)

BBMII_regular_season_counts <- BBMII_regular_season |>
  count(team, position_name) |>
  filter(!is.na(team)) |>
  left_join(image_urls, join_by(team == team_abbr)) |>
  select(team, position_name, n, team_division, team_logo_espn)

BBMIII_regular_season_counts <- BBMIII_regular_season |>
  count(team, position_name) |>
  filter(!is.na(team)) |>
  left_join(image_urls, join_by(team == team_abbr)) |>
  select(team, position_name, n, team_division, team_logo_espn)

BBMII_regular_season_counts$team <- factor(BBMII_regular_season_counts$team,
                                             levels = unique(BBMII_regular_season_counts$team
                                             [order(match(BBMII_regular_season_counts$team_division, division_order

BBMIII_regular_season_counts$team <- factor(BBMIII_regular_season_counts$team,
                                             levels = unique(BBMIII_regular_season_counts$team
                                             [order(match(BBMIII_regular_season_counts$team_division, division_order

BBMII_regular_season_counts$position_name <-
  factor(BBMII_regular_season_counts$position_name,
         levels = c("QB", "RB", "WR", "TE"))

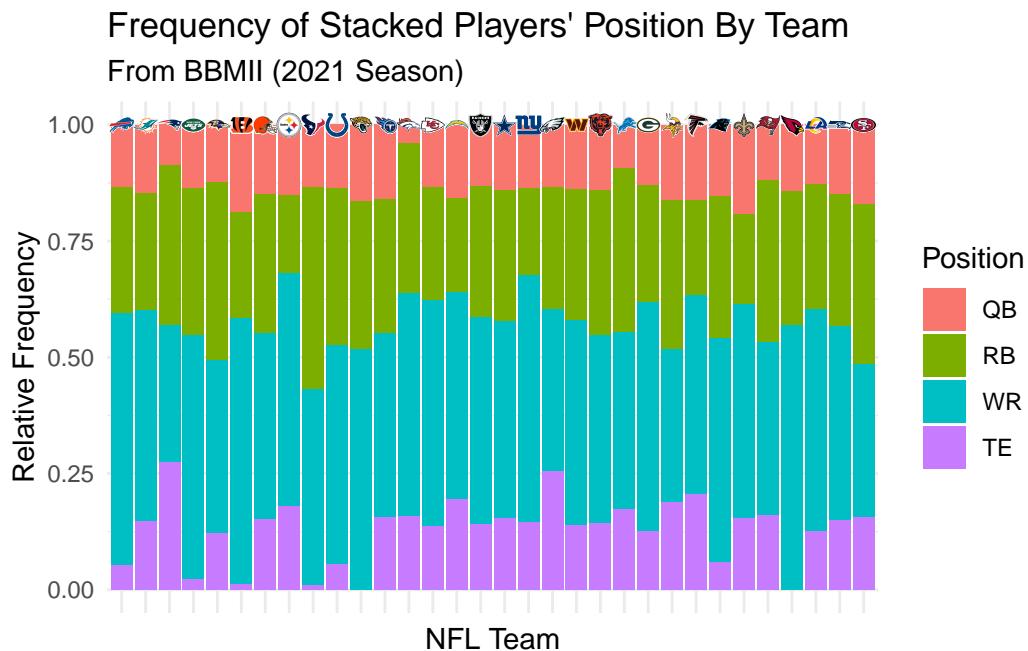
BBMII_regular_season_counts |>
  ggplot(aes(x = team, y = n, fill = position_name)) +
  geom_col(position = "fill") +
  geom_image(aes(image = team_logo_espn, y = 1)) +
  labs(
    title = "Frequency of Stacked Players' Position By Team",
    subtitle = "From BBMII (2021 Season)",
    x = "NFL Team",
    y = "Relative Frequency",
    fill = "Position"
  ) +

```

```

theme_minimal() +
theme(axis.text.x = element_blank(),
      axis.ticks.x = element_blank())

```



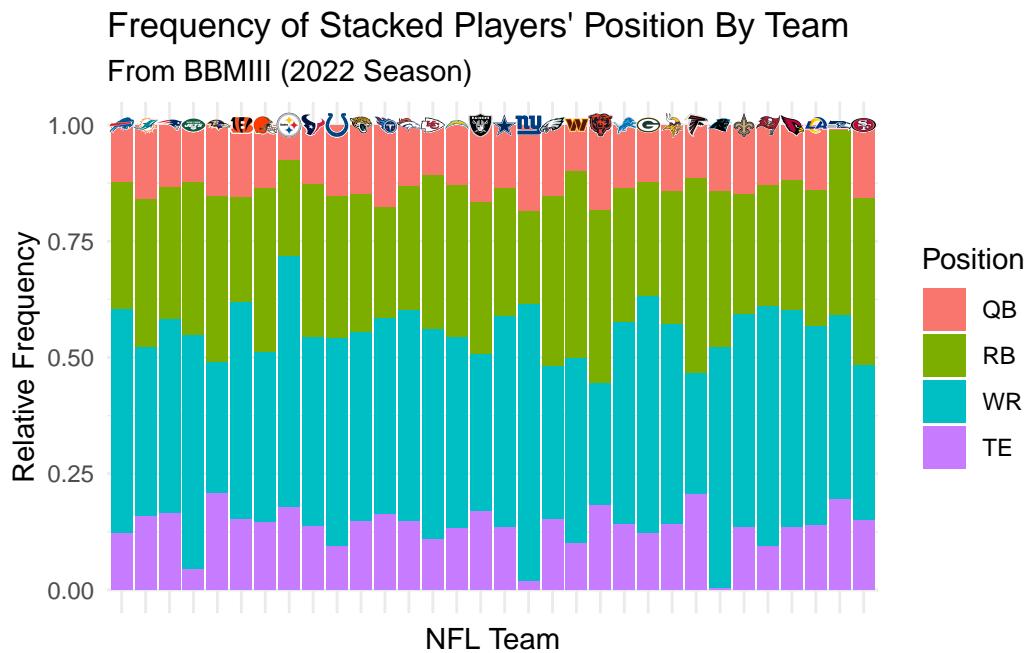
```

BBMIII_regular_season_counts$position_name <-
  factor(BBMIII_regular_season_counts$position_name,
         levels = c("QB", "RB", "WR", "TE"))

BBMIII_regular_season_counts |>
  ggplot(aes(x = team, y = n, fill = position_name)) +
  geom_col(position = "fill") +
  geom_image(aes(image = team_logo_espn, y = 1)) +
  labs(
    title = "Frequency of Stacked Players' Position By Team",
    subtitle = "From BBMIII (2022 Season)",
    x = "NFL Team",
    y = "Relative Frequency",
    fill = "Position"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_blank(),

```

```
axis.ticks.x = element_blank())
```



QB	RB	WR	TE	QB_RB	QB_WR
2661099	3052954	5539827	1782845	996384	2162604
QB_TE	RB_WR	RB_TE	WR_TE	QB_RB_WR	QB_RB_TE
880299	1918091	738735	1187686	684381	282555
QB_WR_TE	RB_WR_TE	QB_RB_WR_TE			
625800	341460	202719			

Positional.Stack	Count
QB and RB only	232167
QB and WR only	1055142
QB and TE only	174663
RB and WR only	1094969
RB and TE only	317439
WR and TE only	423145
QB, RB, and WR only	481662
QB, RB, and TE only	79836
QB, WR, and TE only	423081
RB, WR, and TE only	138741

The following is the main part of the analysis. All three of the following have the same structure, but I had the goal of seeing if there was a difference between what I thought of as high (first 4 rounds), mid-high (first 8 rounds), and medium level (first 12 rounds) stacks. The medium level includes both of these other levels, while mid-high includes the high level stacks. The graphs compare what is the overall ration of teams that were able to make the playoff, compared to the ratio of teams with stacks for certain characteristics (number of players stacked or the number NFL teams that those stacks came from).

```

BBMII_regular_season_stack_players_12_round <- BBMII_regular_season |>
  filter(team_pick_number <= 12) |>
  filter(paste0(tournament_entry_id, team) %in%
    paste0((BBMII_regular_season |>
      filter(team_pick_number <= 12) |>
      count(tournament_entry_id, team) |>
      filter(n>=2))$tournament_entry_id,
    (BBMII_regular_season |>
      filter(team_pick_number <= 12) |>
      count(tournament_entry_id, team) |>
      filter(n>=2))$team))

BBMIII_regular_season_stack_players_12_round <- BBMIII_regular_season |>
  filter(team_pick_number <= 12) |>
  filter(paste0(tournament_entry_id, team) %in%
    paste0((BBMIII_regular_season |>
      filter(team_pick_number <= 12) |>
      count(tournament_entry_id, team) |>
      filter(n>=2))$tournament_entry_id,
    (BBMIII_regular_season |>
      filter(team_pick_number <= 12) |>
      count(tournament_entry_id, team) |>
      filter(n>=2))$team))

stacked_teams_BBMII_12_round <- BBMII_regular_season_stack_players_12_round |>
  select(tournament_entry_id, team, playoff_team)

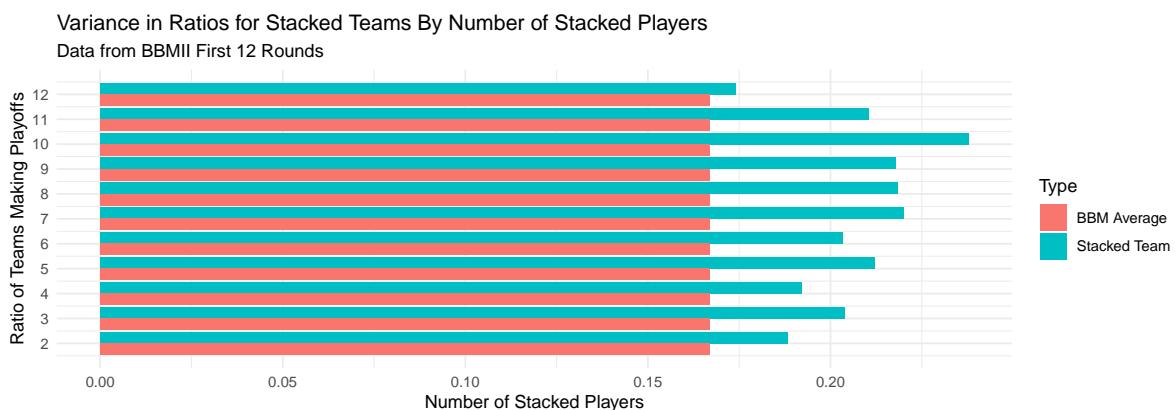
stacked_teams_BBMIII_12_round <- BBMIII_regular_season_stack_players_12_round |>
  select(tournament_entry_id, team, playoff_team)

stacked_teams_BBMII_12_round |>
```

```

count(tournament_entry_id) |>
right_join(BBMII_regular_season, join_by(tournament_entry_id)) |>
group_by(n) |>
count(playoff_team) |>
filter(!is.na(n)) |>
pivot_wider(names_from = playoff_team, values_from = nn) |>
mutate(ratio = `1` / `0`,
       average_ratio = (nrow(BBMII_regular_season) |>
                           filter(playoff_team == TRUE))) /
(nrow(BBMII_regular_season))) |>
select(n, ratio, average_ratio) |>
rename(`Stacked Team` = ratio,
      `BBM Average` = average_ratio) |>
pivot_longer(cols = c(`Stacked Team`, `BBM Average`),
             names_to = "type", values_to = "ratio") |>
ggplot(aes(x = n, y = ratio, fill = type)) +
geom_col(position = "dodge") +
coord_flip() +
scale_x_continuous(n.breaks = 11) +
labs(
  title = "Variance in Ratios for Stacked Teams By Number of Stacked Players",
  subtitle = "Data from BBMII First 12 Rounds",
  x = "Ratio of Teams Making Playoffs",
  y = "Number of Stacked Players",
  fill = "Type"
) +
theme_minimal()

```



```

stacked_teams_BBMIII_12_round |>
  count(tournament_entry_id) |>
  right_join(BBMIII_regular_season, join_by(tournament_entry_id)) |>
  group_by(n) |>
  count(playoff_team) |>
  filter(!is.na(n)) |>
  pivot_wider(names_from = playoff_team, values_from = nn) |>
  mutate(ratio = `1` / `0`,  

         average_ratio = (nrow(BBMIII_regular_season) |>  

                           filter(playoff_team == TRUE))) /  

                           (nrow(BBMIII_regular_season))) |>
  select(n, ratio, average_ratio) |>
  rename(`Stacked Team` = ratio,  

         `BBM Average` = average_ratio) |>
  pivot_longer(cols = c(`Stacked Team`, `BBM Average`),  

               names_to = "type", values_to = "ratio") |>
  ggplot(aes(x = n, y = ratio, fill = type)) +  

  geom_col(position = "dodge") +  

  coord_flip() +  

  scale_x_continuous(n.breaks = 11) +  

  labs(  

    title = "Variance in Ratios for Stacked Teams By Number of Stacked Players",  

    subtitle = "Data from BBMIII First 12 Rounds",  

    x = "Ratio of Teams Making Playoffs",  

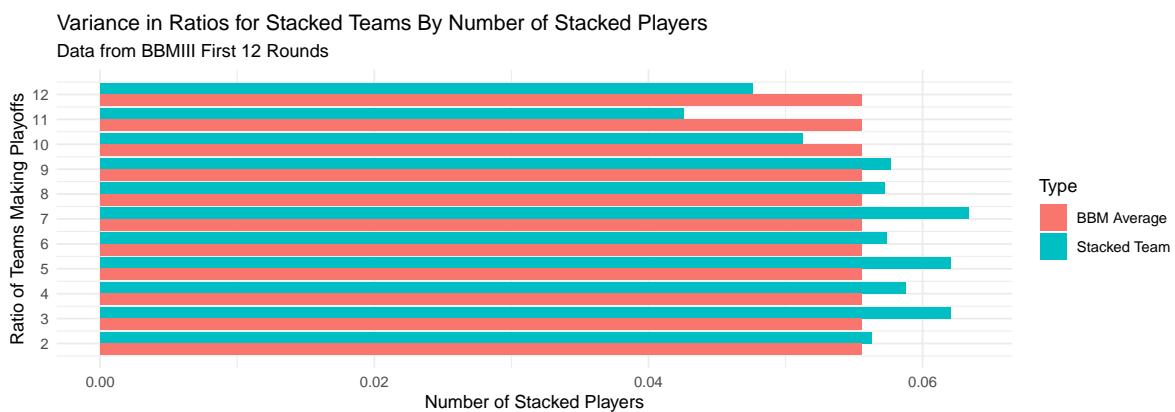
    y = "Number of Stacked Players",  

    fill = "Type"  

  ) +  

  theme_minimal()

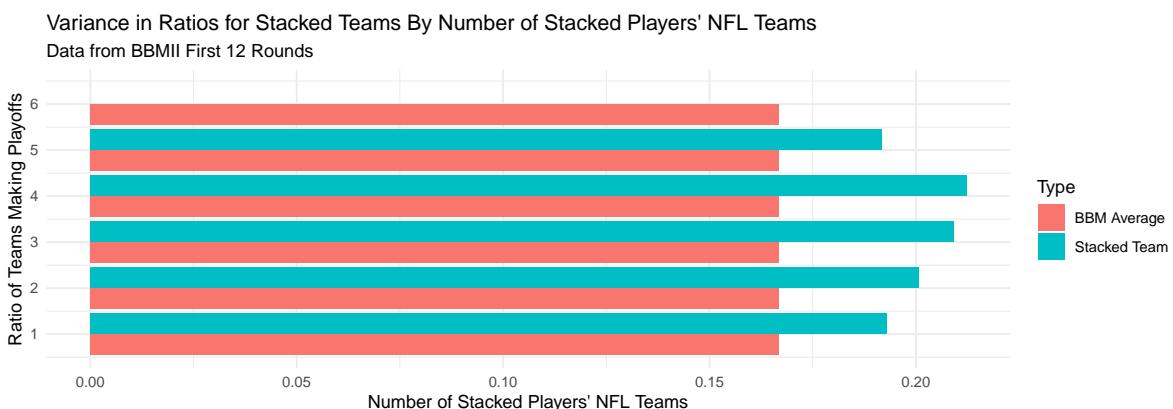
```



```

stacked_teams_BBMII_12_round |>
  count(tournament_entry_id, team) |>
  count(tournament_entry_id) |>
  right_join(BBMII_regular_season, join_by(tournament_entry_id)) |>
  group_by(n) |>
  count(playoff_team) |>
  filter(!is.na(n)) |>
  pivot_wider(names_from = playoff_team, values_from = nn) |>
  mutate(ratio = `1` / `0`,
         average_ratio = (nrow(BBMII_regular_season) |>
                           filter(playoff_team == TRUE)) /
                           (nrow(BBMII_regular_season))) |>
  select(n, ratio, average_ratio) |>
  rename(`Stacked Team` = ratio,
         `BBM Average` = average_ratio) |>
  pivot_longer(cols = c(`Stacked Team`, `BBM Average`),
               names_to = "type", values_to = "ratio") |>
  ggplot(aes(x = n, y = ratio, fill = type)) +
  geom_col(position = "dodge") +
  coord_flip() +
  scale_x_continuous(n.breaks = 11) +
  labs(
    title = "Variance in Ratios for Stacked Teams By Number of Stacked Players' NFL Teams",
    subtitle = "Data from BBMII First 12 Rounds",
    x = "Ratio of Teams Making Playoffs",
    y = "Number of Stacked Players' NFL Teams",
    fill = "Type"
  ) +
  theme_minimal()

```



```

stacked_teams_BBMIII_12_round |>
  count(tournament_entry_id, team) |>
  count(tournament_entry_id) |>
  right_join(BBMIII_regular_season, join_by(tournament_entry_id)) |>
  group_by(n) |>
  count(playoff_team) |>
  filter(!is.na(n)) |>
  pivot_wider(names_from = playoff_team, values_from = nn) |>
  mutate(ratio = `1` / `0`,
         average_ratio = (nrow(BBMIII_regular_season) |>
                           filter(playoff_team == TRUE)) /
                           (nrow(BBMIII_regular_season))) |>
  select(n, ratio, average_ratio) |>
  rename(`Stacked Team` = ratio,
         `BBM Average` = average_ratio) |>
  pivot_longer(cols = c(`Stacked Team`, `BBM Average`),
               names_to = "type", values_to = "ratio") |>
  ggplot(aes(x = n, y = ratio, fill = type)) +
  geom_col(position = "dodge") +
  coord_flip() +
  scale_x_continuous(n.breaks = 11) +
  labs(
    title = "Variance in Ratios for Stacked Teams By Number of Stacked Players' NFL Teams",
    subtitle = "Data from BBMIII First 12 Rounds",
    x = "Ratio of Teams Making Playoffs",
    y = "Number of Stacked Players' NFL Teams",
    fill = "Type"
  ) +
  theme_minimal()

```



```

BBMII_regular_season_stack_players_8_round <- BBMII_regular_season |>
  filter(team_pick_number <= 8) |>
  filter(paste0(tournament_entry_id, team) %in%
  paste0((BBMII_regular_season |>
    filter(team_pick_number <= 8) |>
    count(tournament_entry_id, team) |>
    filter(n >= 2))$tournament_entry_id,
(BBMII_regular_season |>
  filter(team_pick_number <= 8) |>
  count(tournament_entry_id, team) |>
  filter(n >= 2))$team))

BBMIII_regular_season_stack_players_8_round <- BBMIII_regular_season |>
  filter(team_pick_number <= 8) |>
  filter(paste0(tournament_entry_id, team) %in%
  paste0((BBMIII_regular_season |>
    filter(team_pick_number <= 8) |>
    count(tournament_entry_id, team) |>
    filter(n >= 2))$tournament_entry_id,
(BBMIII_regular_season |>
  filter(team_pick_number <= 8) |>
  count(tournament_entry_id, team) |>
  filter(n >= 2))$team))

stacked_teams_BBMII_8_round <- BBMII_regular_season_stack_players_8_round |>
  select(tournament_entry_id, team, playoff_team)

stacked_teams_BBMIII_8_round <- BBMIII_regular_season_stack_players_8_round |>
  select(tournament_entry_id, team, playoff_team)

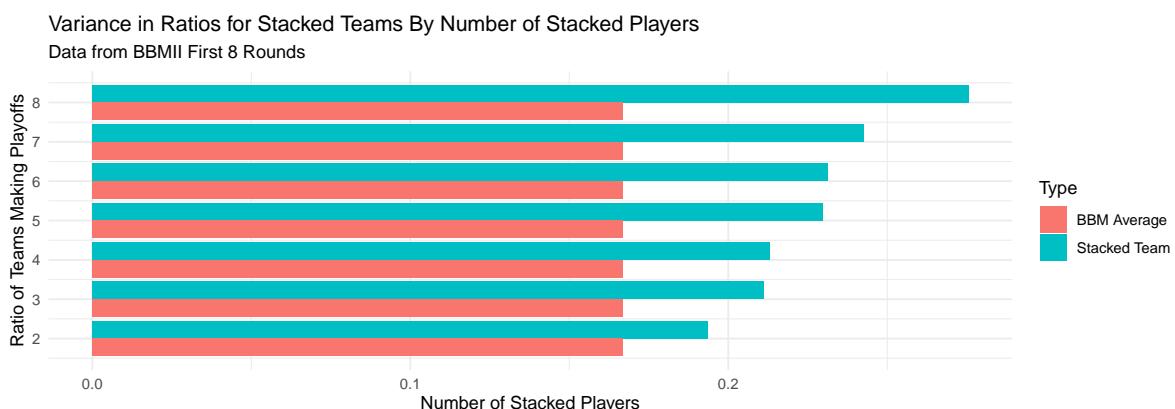
stacked_teams_BBMII_8_round |>
  count(tournament_entry_id) |>
  right_join(BBMII_regular_season, join_by(tournament_entry_id)) |>
  group_by(n) |>
  count(playoff_team) |>
  filter(!is.na(n)) |>
  pivot_wider(names_from = playoff_team, values_from = nn) |>
  mutate(ratio = `1` / `0`,
    average_ratio = (nrow(BBMII_regular_season |>
      filter(playoff_team == TRUE))) /
      (nrow(BBMII_regular_season))) |>

```

```

select(n, ratio, average_ratio) |>
  rename(`Stacked Team` = ratio,
         `BBM Average` = average_ratio) |>
  pivot_longer(cols = c(`Stacked Team`, `BBM Average`),
               names_to = "type", values_to = "ratio") |>
  ggplot(aes(x = n, y = ratio, fill = type)) +
  geom_col(position = "dodge") +
  coord_flip() +
  scale_x_continuous(n.breaks = 7) +
  labs(
    title = "Variance in Ratios for Stacked Teams By Number of Stacked Players",
    subtitle = "Data from BBMII First 8 Rounds",
    x = "Ratio of Teams Making Playoffs",
    y = "Number of Stacked Players",
    fill = "Type"
  ) +
  theme_minimal()

```



```

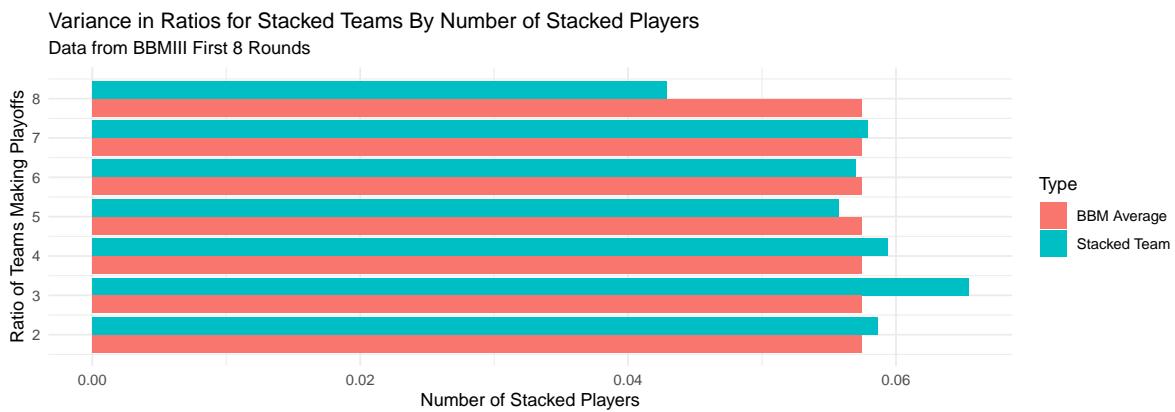
stacked_teams_BBMIII_8_round |>
  count(tournament_entry_id) |>
  right_join(BBMIII_regular_season, join_by(tournament_entry_id)) |>
  group_by(n) |>
  count(playoff_team) |>
  filter(!is.na(n)) |>
  pivot_wider(names_from = playoff_team, values_from = nn) |>
  mutate(ratio = `1` / `0`,
         average_ratio = (nrow(BBMII_regular_season |>
                                filter(playoff_team == TRUE))) /

```

```

        (nrow(BBMIII_regular_season))) |>
select(n, ratio, average_ratio) |>
rename(`Stacked Team` = ratio,
`BBM Average` = average_ratio) |>
pivot_longer(cols = c(`Stacked Team`, `BBM Average`),
names_to = "type", values_to = "ratio") |>
ggplot(aes(x = n, y = ratio, fill = type)) +
geom_col(position = "dodge") +
coord_flip() +
scale_x_continuous(n.breaks = 7) +
labs(
  title = "Variance in Ratios for Stacked Teams By Number of Stacked Players",
  subtitle = "Data from BBMIII First 8 Rounds",
  x = "Ratio of Teams Making Playoffs",
  y = "Number of Stacked Players",
  fill = "Type"
) +
theme_minimal()

```



```

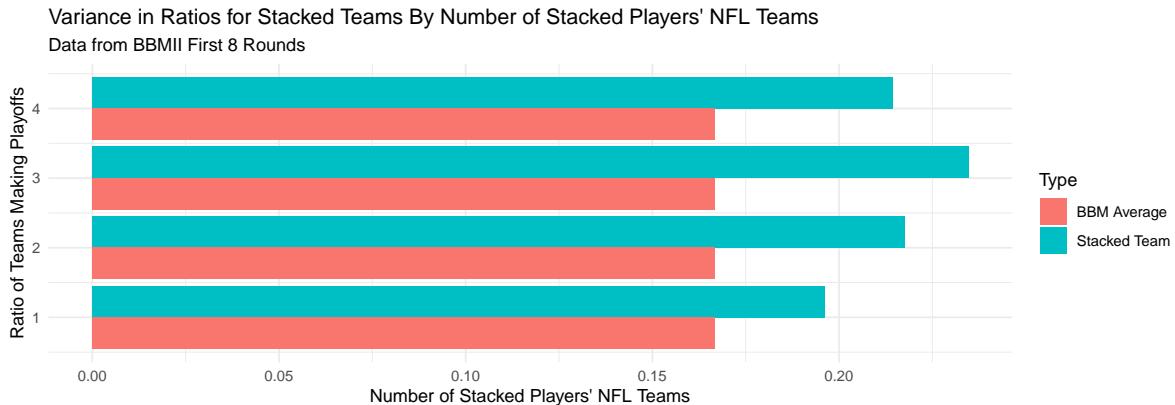
stacked_teams_BBMII_8_round |>
count(tournament_entry_id, team) |>
count(tournament_entry_id) |>
right_join(BBMII_regular_season, join_by(tournament_entry_id)) |>
group_by(n) |>
count(playoff_team) |>
filter(!is.na(n)) |>
pivot_wider(names_from = playoff_team, values_from = nn) |>
mutate(ratio = `1` / `0`,

```

```

average_ratio = (nrow(BBMII_regular_season |>
                      filter(playoff_team == TRUE))) / 
  (nrow(BBMII_regular_season)) |>
  select(n, ratio, average_ratio) |>
  rename(`Stacked Team` = ratio,
         `BBM Average` = average_ratio) |>
  pivot_longer(cols = c(`Stacked Team`, `BBM Average`),
               names_to = "type", values_to = "ratio") |>
  ggplot(aes(x = n, y = ratio, fill = type)) +
  geom_col(position = "dodge") +
  coord_flip() +
  scale_x_continuous(n.breaks = 7) +
  labs(
    title = "Variance in Ratios for Stacked Teams By Number of Stacked Players' NFL Teams",
    subtitle = "Data from BBMII First 8 Rounds",
    x = "Ratio of Teams Making Playoffs",
    y = "Number of Stacked Players' NFL Teams",
    fill = "Type"
  ) +
  theme_minimal()

```



```

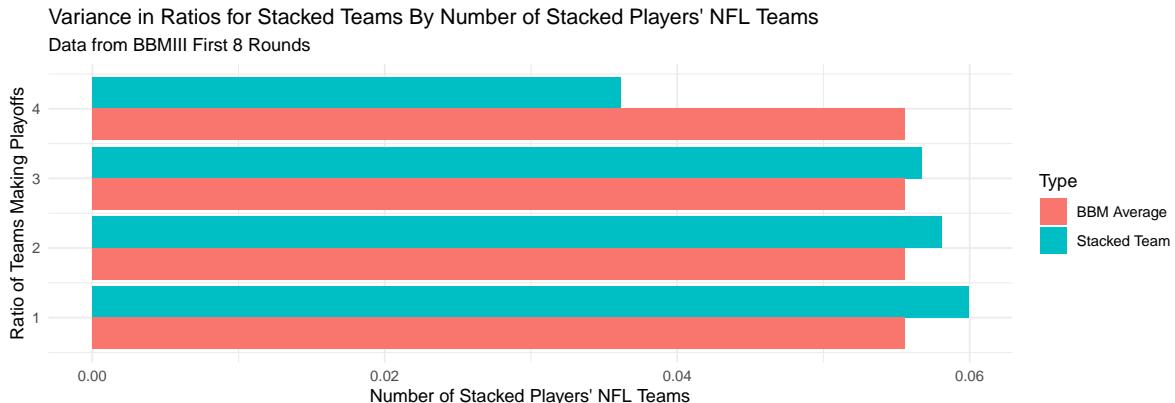
stacked_teams_BBMIII_8_round |>
  count(tournament_entry_id, team) |>
  count(tournament_entry_id) |>
  right_join(BBMIII_regular_season, join_by(tournament_entry_id)) |>
  group_by(n) |>
  count(playoff_team) |>
  filter(!is.na(n)) |>

```

```

pivot_wider(names_from = playoff_team, values_from = nn) |>
  mutate(ratio = `1` / `0`,
         average_ratio = (nrow(BBMIII_regular_season) |>
                           filter(playoff_team == TRUE)) /
                           (nrow(BBMIII_regular_season))) |>
  select(n, ratio, average_ratio) |>
  rename(`Stacked Team` = ratio,
         `BBM Average` = average_ratio) |>
  pivot_longer(cols = c(`Stacked Team`, `BBM Average`),
               names_to = "type", values_to = "ratio") |>
  ggplot(aes(x = n, y = ratio, fill = type)) +
  geom_col(position = "dodge") +
  coord_flip() +
  scale_x_continuous(n.breaks = 7) +
  labs(
    title = "Variance in Ratios for Stacked Teams By Number of Stacked Players' NFL Teams",
    subtitle = "Data from BBMIII First 8 Rounds",
    x = "Ratio of Teams Making Playoffs",
    y = "Number of Stacked Players' NFL Teams",
    fill = "Type"
  ) +
  theme_minimal()

```



```

BBMII_regular_season_stack_players_4_round <- BBMII_regular_season |>
  filter(team_pick_number <= 4) |>
  filter(paste0(tournament_entry_id, team) %in%
  paste0((BBMII_regular_season |>
    filter(team_pick_number <= 4) |>

```

```

    count(tournament_entry_id, team) |>
      filter(n>=2))$tournament_entry_id,
    (BBMII_regular_season |>
      filter(team_pick_number <= 4) |>
      count(tournament_entry_id, team) |>
      filter(n>=2))$team))

BBMIII_regular_season_stack_players_4_round <- BBMIII_regular_season |>
  filter(team_pick_number <= 4) |>
  filter(paste0(tournament_entry_id, team) %in%
  paste0((BBMIII_regular_season |>
    filter(team_pick_number <= 4) |>
    count(tournament_entry_id, team) |>
    filter(n>=2))$tournament_entry_id,
  (BBMIII_regular_season |>
    filter(team_pick_number <= 4) |>
    count(tournament_entry_id, team) |>
    filter(n>=2))$team))

stacked_teams_BBMII_4_round <- BBMII_regular_season_stack_players_4_round |>
  select(tournament_entry_id, team, playoff_team)

stacked_teams_BBMIII_4_round <- BBMIII_regular_season_stack_players_4_round |>
  select(tournament_entry_id, team, playoff_team)

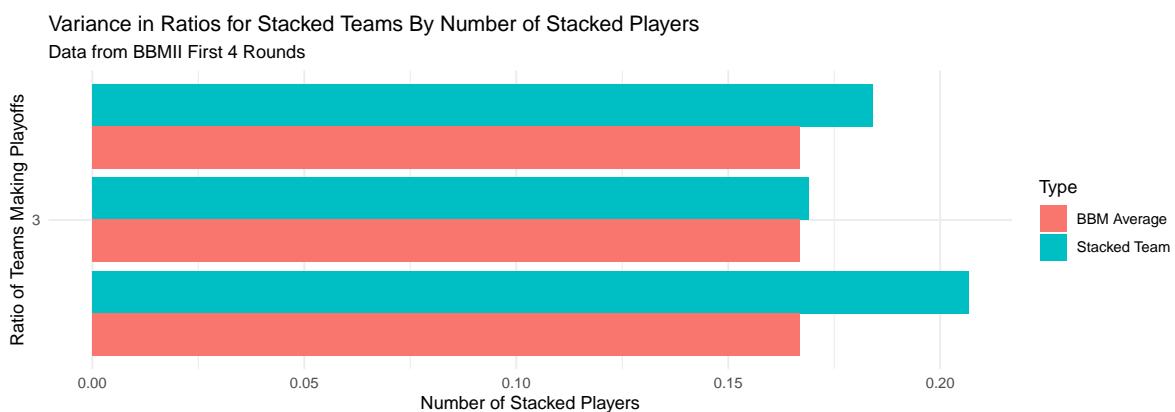
stacked_teams_BBMII_4_round |>
  count(tournament_entry_id) |>
  right_join(BBMII_regular_season, join_by(tournament_entry_id)) |>
  group_by(n) |>
  count(playoff_team) |>
  filter(!is.na(n)) |>
  pivot_wider(names_from = playoff_team, values_from = nn) |>
  mutate(ratio = `1` / `0`,
         average_ratio = (nrow(BBMII_regular_season) |>
                           filter(playoff_team == TRUE))) /
    (nrow(BBMII_regular_season))) |>
  select(n, ratio, average_ratio) |>
  rename(`Stacked Team` = ratio,
         `BBM Average` = average_ratio) |>
  pivot_longer(cols = c(`Stacked Team`, `BBM Average`),
               names_to = "type", values_to = "ratio") |>

```

```

ggplot(aes(x = n, y = ratio, fill = type)) +
  geom_col(position = "dodge") +
  coord_flip() +
  scale_x_continuous(n.breaks = 3) +
  labs(
    title = "Variance in Ratios for Stacked Teams By Number of Stacked Players",
    subtitle = "Data from BBMII First 4 Rounds",
    x = "Ratio of Teams Making Playoffs",
    y = "Number of Stacked Players",
    fill = "Type"
  ) +
  theme_minimal()

```



```

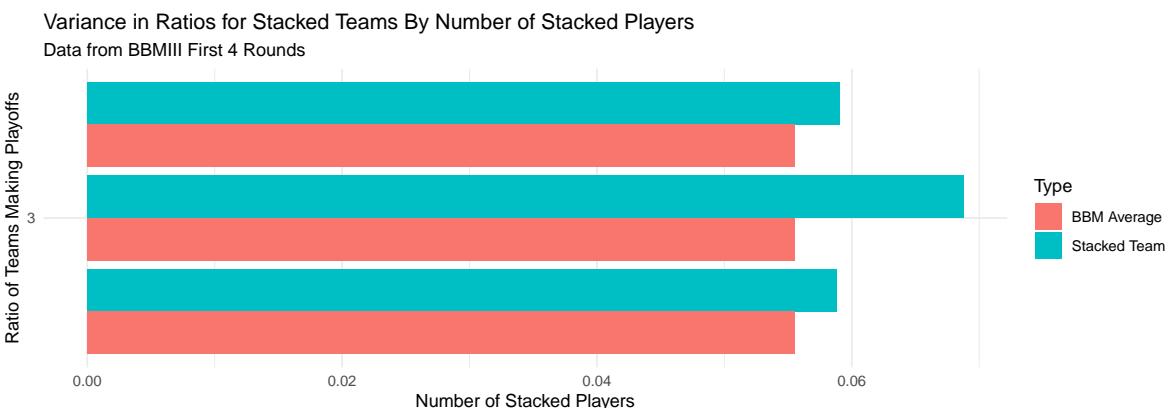
stacked_teams_BBMIII_4_round |>
  count(tournament_entry_id) |>
  right_join(BBMIII_regular_season, join_by(tournament_entry_id)) |>
  group_by(n) |>
  count(playoff_team) |>
  filter(!is.na(n)) |>
  pivot_wider(names_from = playoff_team, values_from = nn) |>
  mutate(ratio = `1` / `0`,
         average_ratio = (nrow(BBMIII_regular_season) |>
                           filter(playoff_team == TRUE)) /
                           (nrow(BBMIII_regular_season))) |>
  select(n, ratio, average_ratio) |>
  rename(`Stacked Team` = ratio,
         `BBM Average` = average_ratio) |>
  pivot_longer(cols = c(`Stacked Team`, `BBM Average`),

```

```

            names_to = "type", values_to = "ratio") |>
ggplot(aes(x = n, y = ratio, fill = type)) +
geom_col(position = "dodge") +
coord_flip() +
scale_x_continuous(n.breaks = 3) +
labs(
  title = "Variance in Ratios for Stacked Teams By Number of Stacked Players",
  subtitle = "Data from BBMIII First 4 Rounds",
  x = "Ratio of Teams Making Playoffs",
  y = "Number of Stacked Players",
  fill = "Type"
) +
theme_minimal()

```



```

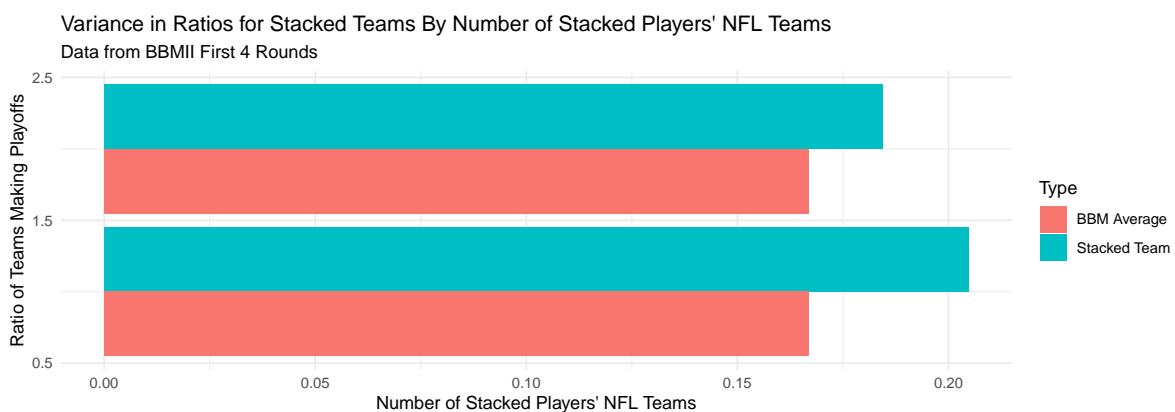
stacked_teams_BBMII_4_round |>
  count(tournament_entry_id, team) |>
  count(tournament_entry_id) |>
  right_join(BBMII_regular_season, join_by(tournament_entry_id)) |>
  group_by(n) |>
  count(playoff_team) |>
  filter(!is.na(n)) |>
  pivot_wider(names_from = playoff_team, values_from = nn) |>
  mutate(ratio = `1` / `0`,
        average_ratio = (nrow(BBMII_regular_season) |>
                           filter(playoff_team == TRUE))) /
        (nrow(BBMII_regular_season))) |>
  select(n, ratio, average_ratio) |>
  rename(`Stacked Team` = ratio,

```

```

`BBM Average` = average_ratio) |>
pivot_longer(cols = c(`Stacked Team`, `BBM Average`),
             names_to = "type", values_to = "ratio") |>
ggplot(aes(x = n, y = ratio, fill = type)) +
geom_col(position = "dodge") +
coord_flip() +
scale_x_continuous(n.breaks = 3) +
labs(
  title = "Variance in Ratios for Stacked Teams By Number of Stacked Players' NFL Teams",
  subtitle = "Data from BBMII First 4 Rounds",
  x = "Ratio of Teams Making Playoffs",
  y = "Number of Stacked Players' NFL Teams",
  fill = "Type"
) +
theme_minimal()

```



```

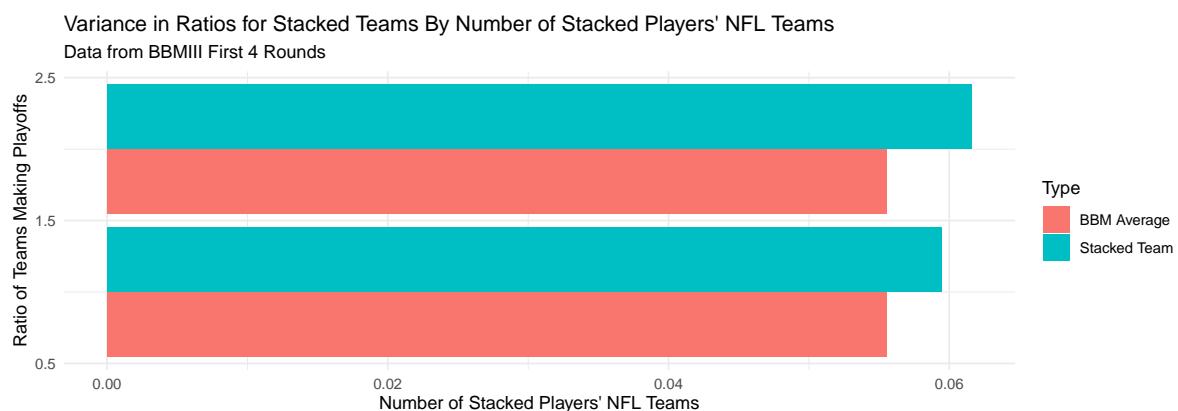
stacked_teams_BBMIII_4_round |>
  count(tournament_entry_id, team) |>
  count(tournament_entry_id) |>
  right_join(BBMIII_regular_season, join_by(tournament_entry_id)) |>
  group_by(n) |>
  count(playoff_team) |>
  filter(!is.na(n)) |>
  pivot_wider(names_from = playoff_team, values_from = nn) |>
  mutate(ratio = `1` / `0`,
        average_ratio = (nrow(BBMIII_regular_season) |>
                           filter(playoff_team == TRUE))) /
  (nrow(BBMIII_regular_season))) |>

```

```

select(n, ratio, average_ratio) |>
  rename(`Stacked Team` = ratio,
         `BBM Average` = average_ratio) |>
  pivot_longer(cols = c(`Stacked Team`, `BBM Average`),
               names_to = "type", values_to = "ratio") |>
  ggplot(aes(x = n, y = ratio, fill = type)) +
  geom_col(position = "dodge") +
  coord_flip() +
  scale_x_continuous(n.breaks = 3) +
  labs(
    title = "Variance in Ratios for Stacked Teams By Number of Stacked Players' NFL Teams",
    subtitle = "Data from BBMIII First 4 Rounds",
    x = "Ratio of Teams Making Playoffs",
    y = "Number of Stacked Players' NFL Teams",
    fill = "Type"
  ) +
  theme_minimal()

```



The results show that having stacked players greatly benefits the chances of your BBM team making it on to the next round. However, this varies by the level of stacking and spreading out your stacking too much along too many NFL teams' players leads to a dropoff for the medium level player stacks.

The process was a nice one to get a practice some more data science skills especially with trying out some new packages and getting that coding rust out with an interesting data set. Although I was hoping to get some more analysis done here about a wider range of questions about stacking players from different teams in a variety of ways, this was not able to be done due to time constraints. So to end off I wanted to write a little more about some of the ideas that I had envisioned to add (and hopefully will be able to capitalize on with more data next

summer!):

- Does stacking from the higher standing teams lead to better results, or is there little correlation?
  - If the former is the case, this could influence a strategy of heavily focusing on players from teams that are anticipated to be playoff contenders this year. This of course comes with an issue of better players, especially quarterbacks, being on better teams, but there is definitely no lack of talent to stack from the teams that may end up at the bottom of the standings.
- Are triples/quads from the same team efficient stacking decisions? Is there a round where this drops off and stacking with that final player from a team isn't as helpful to advancing?
  - Triples and quads of positions from the same team can be risky but might also be a good way of smoothing out the scoring that the group is able to put up when there is an off-day for some players.
- What double/triple stack is the most effective?
  - Similar to the previous idea, there are some stacks that could build upon each other, but get destroyed by an off game, while others might be more smoothing but with lower booming potential.

Thank you guys on the Underdog Team!