

Exploratory Analysis Report

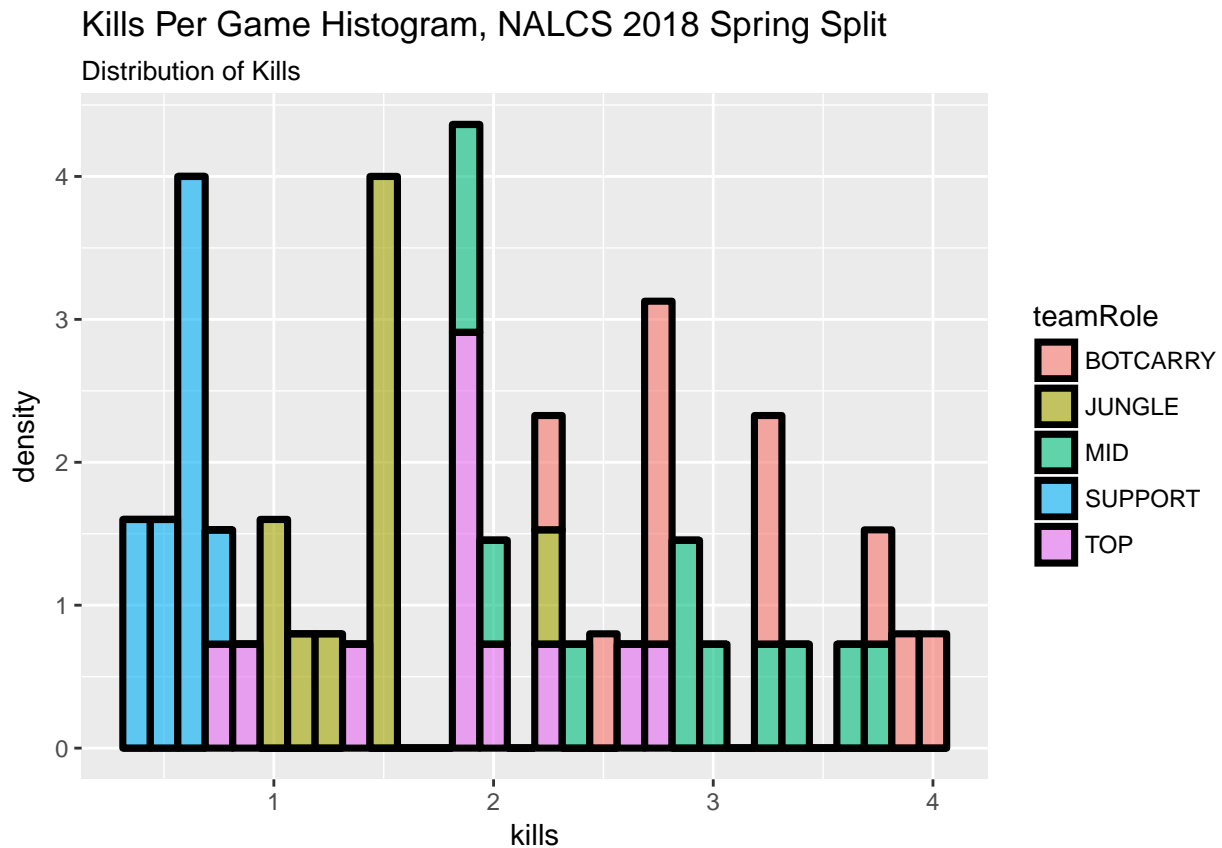
by Ryan Transfiguracion

Exploring Player and Team Role Performances

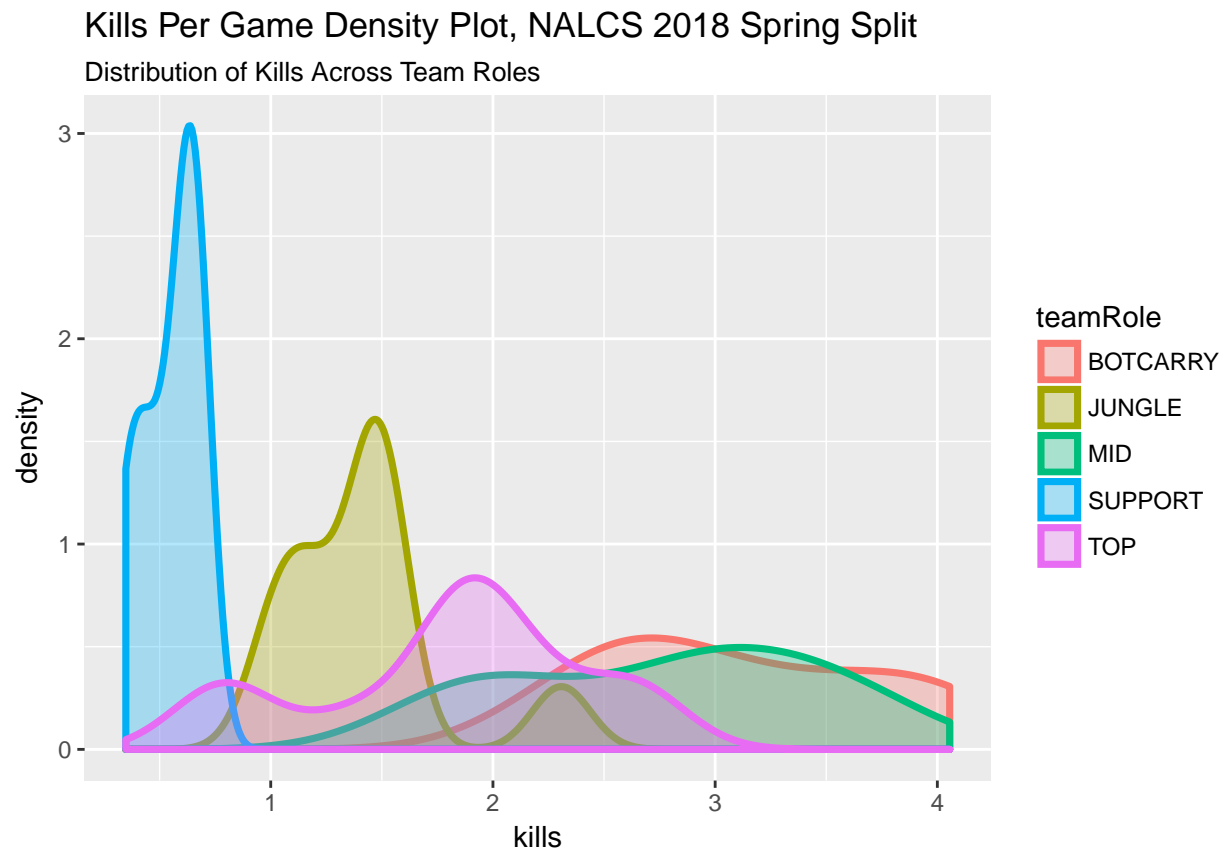
As we skimmed through the NALCS player averages data frame (first, by taking the match-by-match player stats dataset here, grouping by player name and team role, and then taking a mean summary of most of the numerical columns) and sorted the observations by individual columns, it was clear that certain team roles were generally grouped at the top, bottom, or middle of the lists of individual statistics.

First, let's look at some plots for *kills per game* in the NALCS:

```
nalcs_plot_player_avgs +  
  geom_histogram(mapping = aes(x = kills, y = ..density.., fill = teamRole),  
                 color = "black", size = 1.25, alpha = .6, binwidth = .125) +  
  labs(  
    title = "Kills Per Game Histogram, NALCS 2018 Spring Split",  
    subtitle = "Distribution of Kills")
```



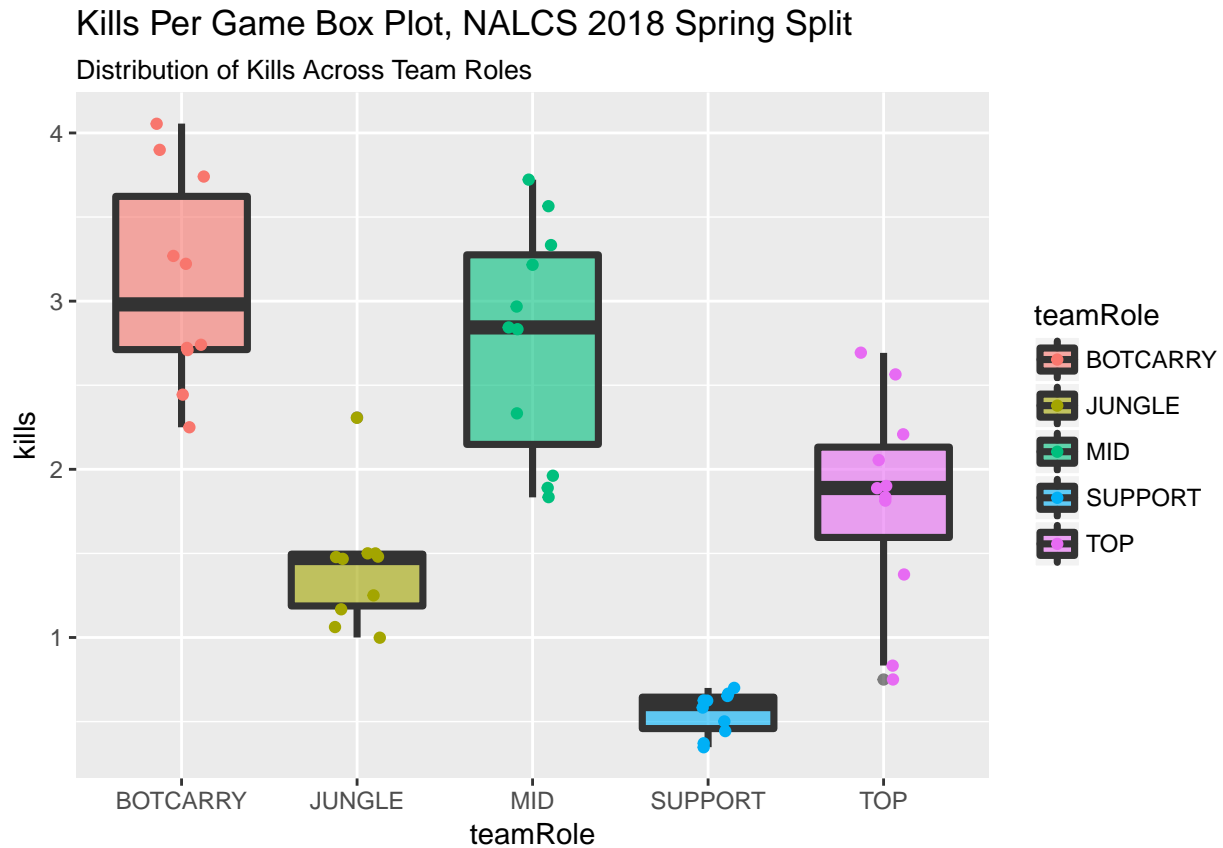
```
nalcs_plot_player_avgs +
  geom_density(mapping = aes(x = kills, color = teamRole, fill = teamRole),
              alpha = .3, size = 1.25) +
  labs(
    title = "Kills Per Game Density Plot, NALCS 2018 Spring Split",
    subtitle = "Distribution of Kills Across Team Roles")
```



```

nalcs_plot_player_avgs +
  geom_boxplot(mapping = aes(x = teamRole, y = kills, fill = teamRole),
               size = 1.25, alpha = .6) +
  geom_jitter(width = 0.15, mapping = aes(x = teamRole, y = kills, color = teamRole)) +
  labs(
    title = "Kills Per Game Box Plot, NALCS 2018 Spring Split",
    subtitle = "Distribution of Kills Across Team Roles")

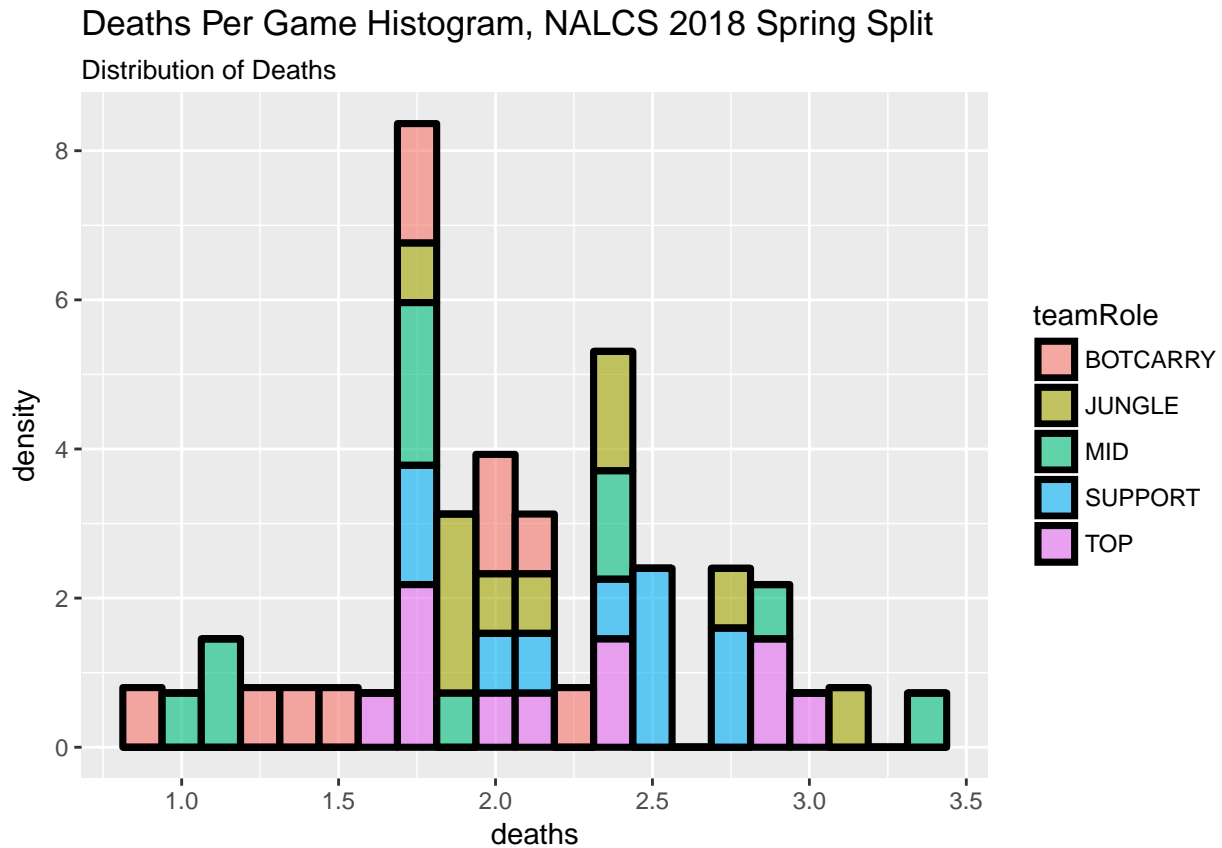
```



It's pretty clear in all three plots that Supports achieve the least number of kills per game, followed by Junglers and Top Laners, while Mid Laners and Bottom Lane Carries are mixed together at the top.

Now, let's look at *deaths per game*:

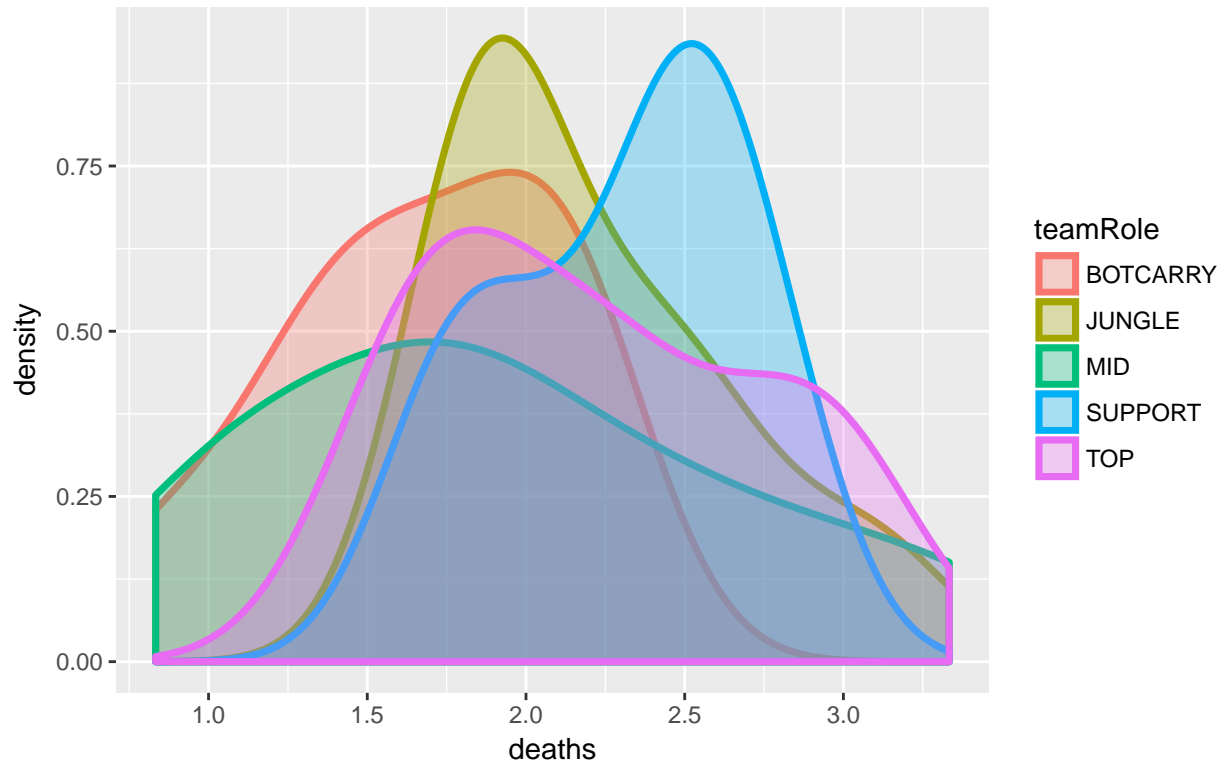
```
nalcs_plot_player_avgs +  
  geom_histogram(mapping = aes(x = deaths, y = ..density.., fill = teamRole),  
                 color = "black", size = 1.25, alpha = .6, binwidth = .125) +  
  labs(  
    title = "Deaths Per Game Histogram, NALCS 2018 Spring Split",  
    subtitle = "Distribution of Deaths")
```



```
nalcs_plot_player_avgs +  
  geom_density(mapping = aes(x = deaths, color = teamRole, fill = teamRole),  
               alpha = .3, size = 1.25) +  
  labs(  
    title = "Deaths Per Game Density Plot, NALCS 2018 Spring Split",  
    subtitle = "Distribution of Deaths Across Team Roles")
```

Deaths Per Game Density Plot, NALCS 2018 Spring Split

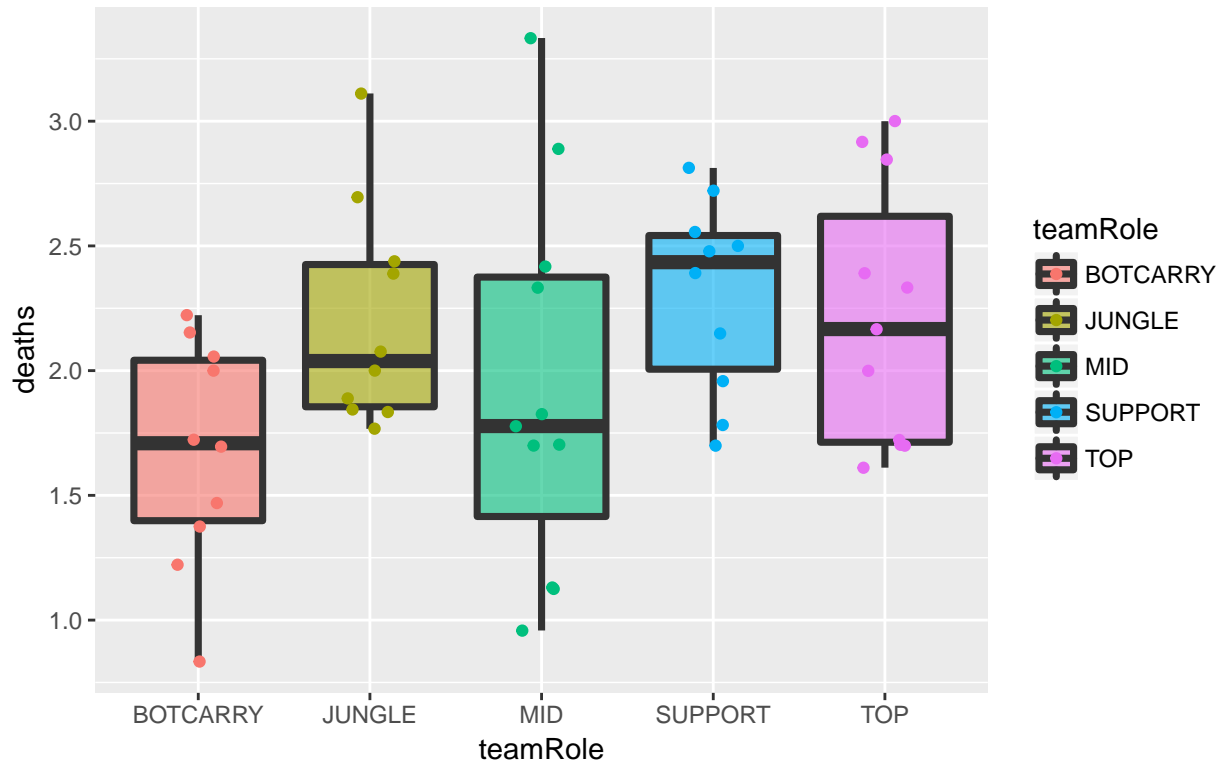
Distribution of Deaths Across Team Roles



```
nalcs_plot_player_avgs +  
  geom_boxplot(mapping = aes(x = teamRole, y = deaths, fill = teamRole),  
               size = 1.25, alpha = .6) +  
  geom_jitter(width = 0.15, mapping = aes(x = teamRole, y = deaths, color = teamRole)) +  
  labs(  
    title = "Deaths Per Game Box Plot, NALCS 2018 Spring Split",  
    subtitle = "Distribution of Assists Across Team Roles")
```

Deaths Per Game Box Plot, NALCS 2018 Spring Split

Distribution of Assists Across Team Roles



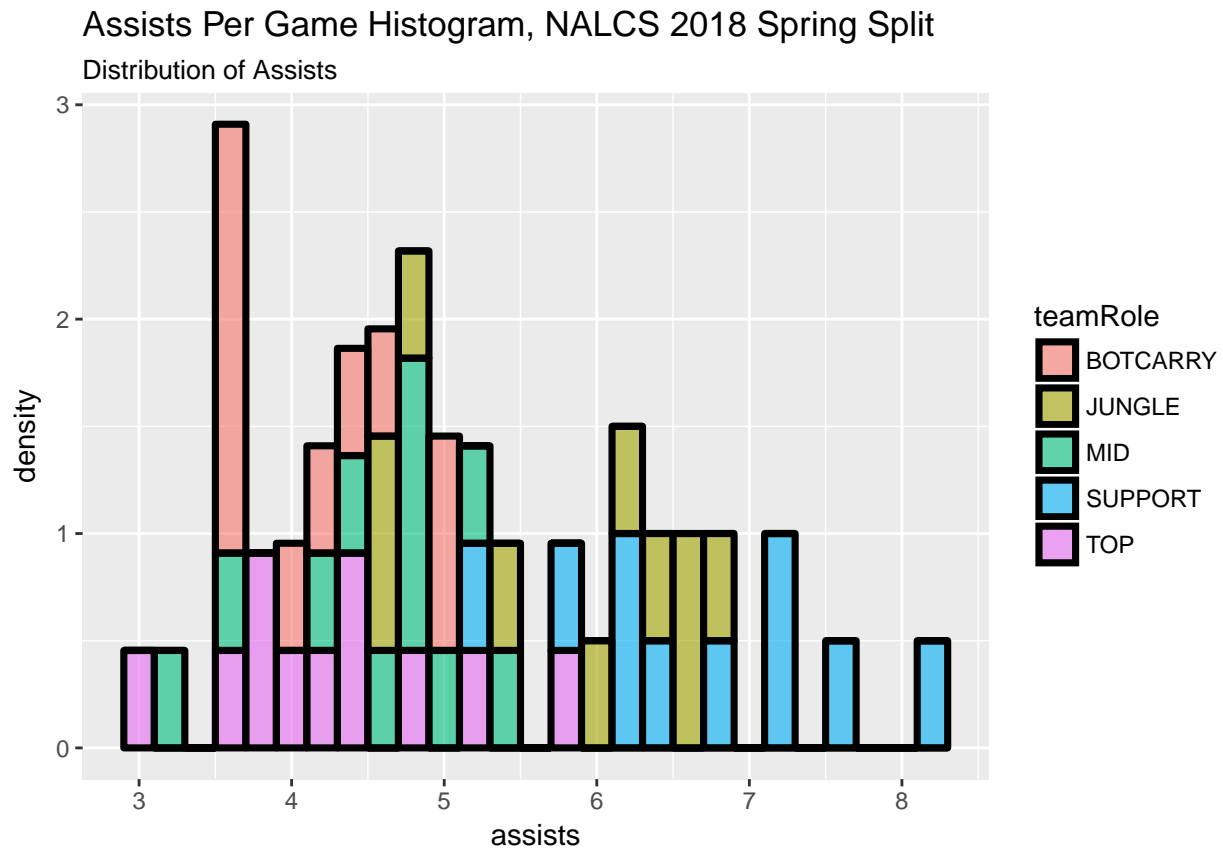
While team roles are clumped closer together for deaths per game compared to the kills per game statistic, we see that Top Laners and Supports are typically die the most often.

In a typical League of Legends game, the Top Laner is the most isolated role during the laning phase, and, thus, is most vulnerable to a gank (an amalgam of “gang flank”, or slang for “group kill”) by the enemy Jungler assisting with the enemy Top Laner.

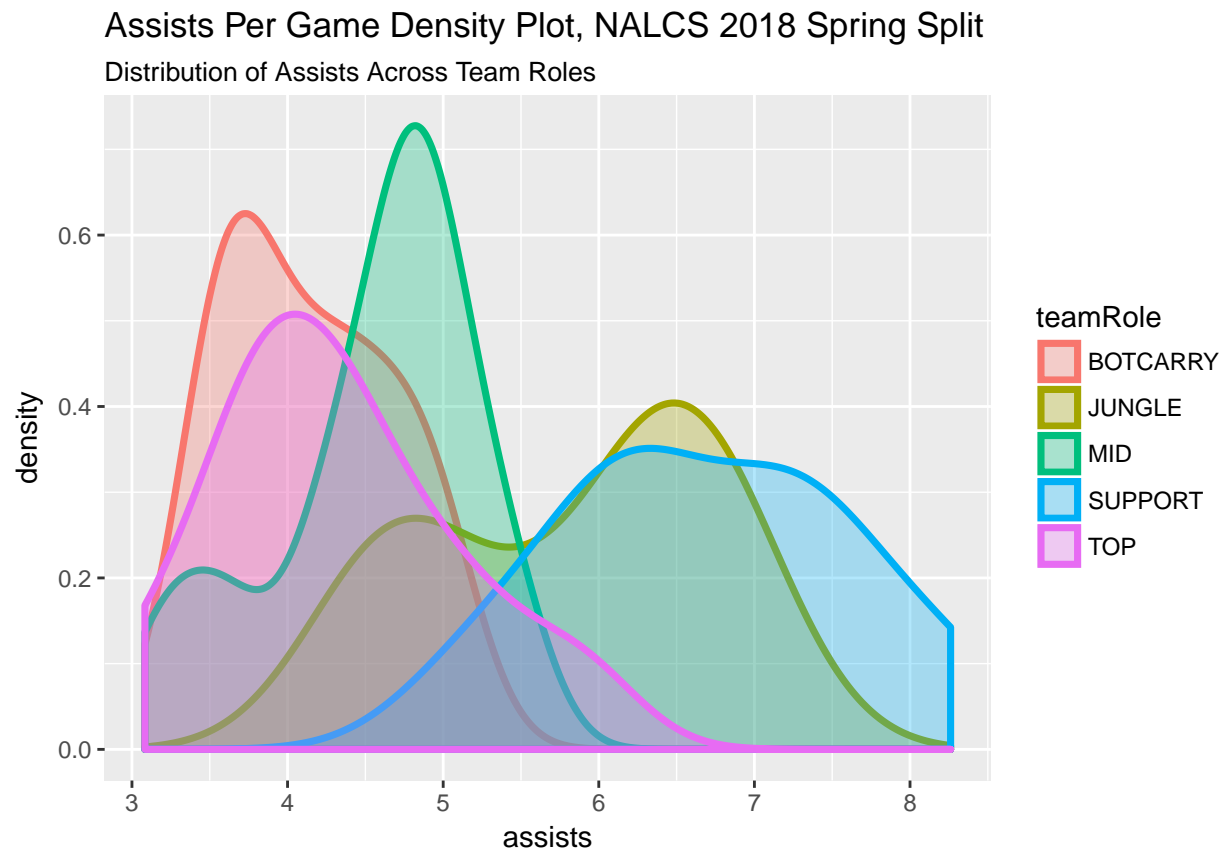
Supports are typically the weakest roles when combining both damage output and defense, and their protective abilities (e.g., healing, providing shields) are usually prioritized for their teammates. Thus, they are usually prioritized by the enemy team in group fights because of their abilities to help their teammates survive.

Now, let's look at *assists per game*. To clarify, an assist is usually achieved by either dealing damage or hindrance abilities (i.e. debuffs and crowd control [CC]) to a killed enemy champion without dealing the killing blow themselves, or providing protection or augmenting abilities (i.e., buffs) to a teammate who deals the killing blow to an enemy champion:

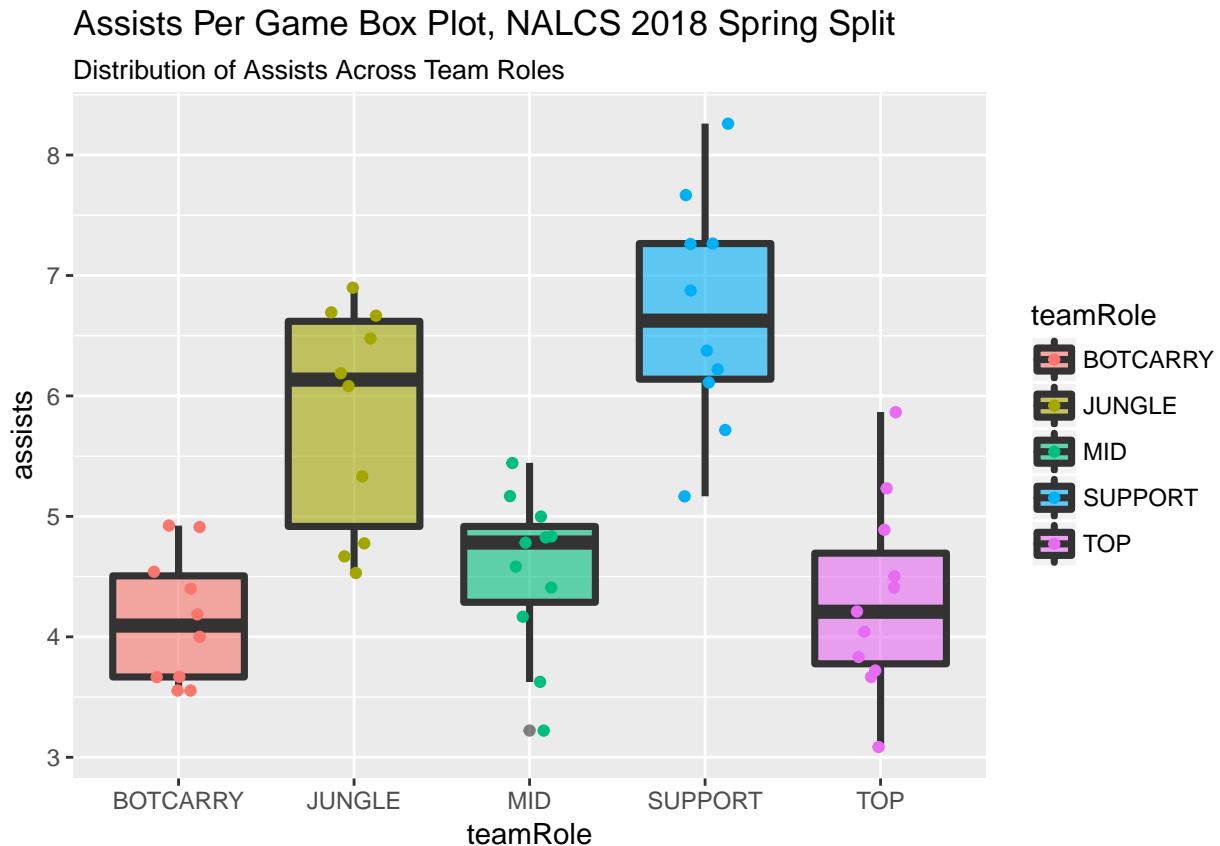
```
nalcs_plot_player_avgs +  
  geom_histogram(mapping = aes(x = assists, y = ..density.., fill = teamRole),  
                 color = "black", size = 1.25, alpha = .6, binwidth = .2) +  
  labs(  
    title = "Assists Per Game Histogram, NALCS 2018 Spring Split",  
    subtitle = "Distribution of Assists")
```



```
nalcs_plot_player_avgs +
  geom_density(mapping = aes(x = assists, color = teamRole, fill = teamRole),
               alpha = .3, size = 1.25) +
  labs(
    title = "Assists Per Game Density Plot, NALCS 2018 Spring Split",
    subtitle = "Distribution of Assists Across Team Roles")
```




```
nalcs_plot_player_avgs +
  geom_boxplot(mapping = aes(x = teamRole, y = assists, fill = teamRole),
               size = 1.25, alpha = .6) +
  geom_jitter(width = 0.15, mapping = aes(x = teamRole, y = assists, color = teamRole)) +
  labs(
    title = "Assists Per Game Box Plot, NALCS 2018 Spring Split",
    subtitle = "Distribution of Assists Across Team Roles")
```



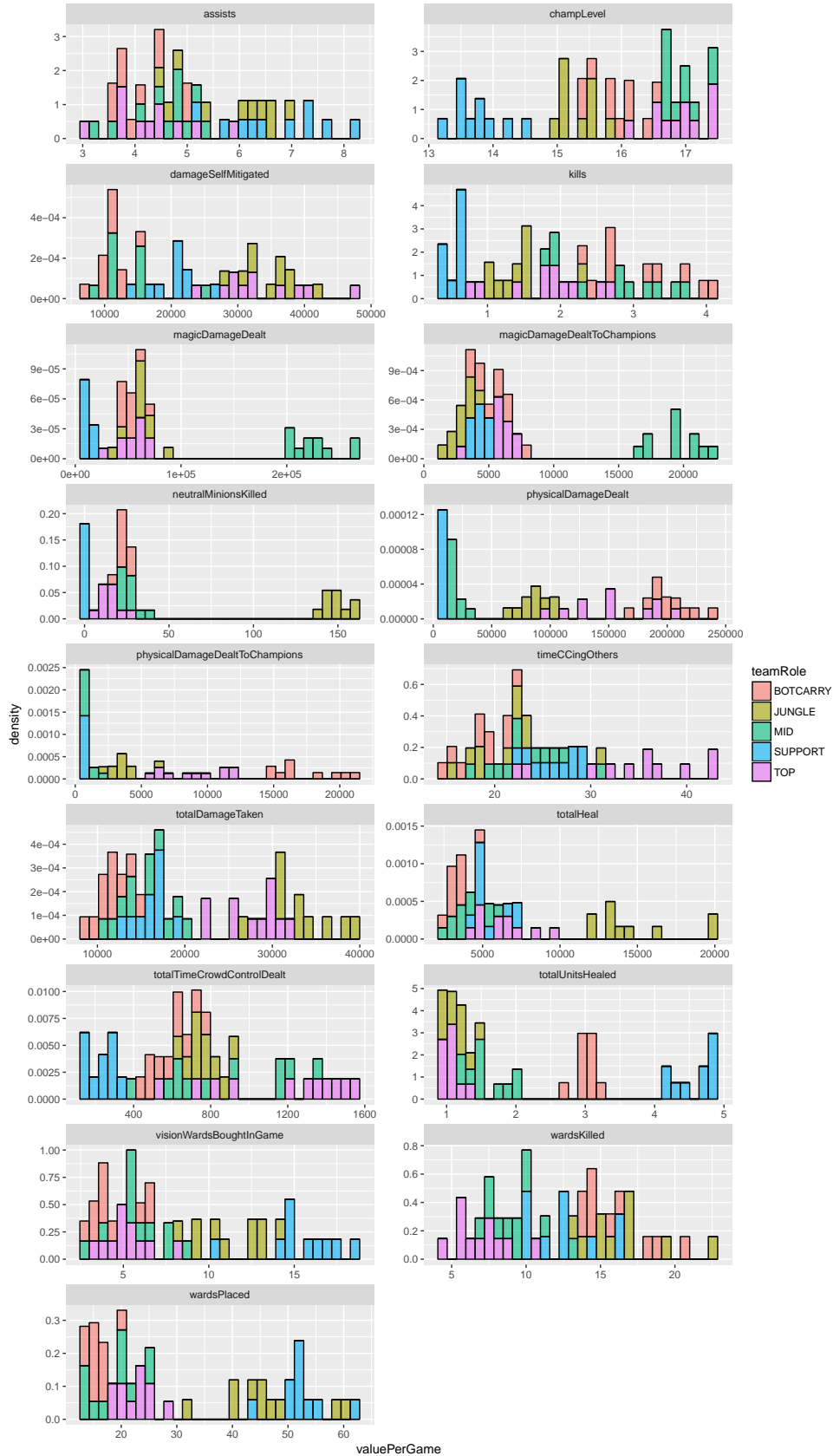
We see nearly the opposite of *kills per game*. Supports followed by Junglers average the most assists in a game. Supports typically, though lacking in damage output, provide the most protection, CC, buffs, and/or debuffs, while Junglers provide ganks for their teammates since they roam through most of the map.

Below are collections of histograms, density plots, and box plots faceted by 17 different variables, in alphabetical order, showing player averages per game through the entire 2018 Spring Split (regular season, tiebreakers, and playoffs):

```
# Histograms
nalcs_plot_player_avgs_gathered +
  geom_histogram(mapping = aes(x = valuePerGame, y = ..density.., fill = teamRole),
                 color = "black", alpha = .6) +
  facet_wrap(~varName, scales = "free", ncol = 2) +
  labs(
    title = "Player Averages Per Game Histograms, NALCS 2018 Spring Split",
    subtitle = "Distribution of Values")
```

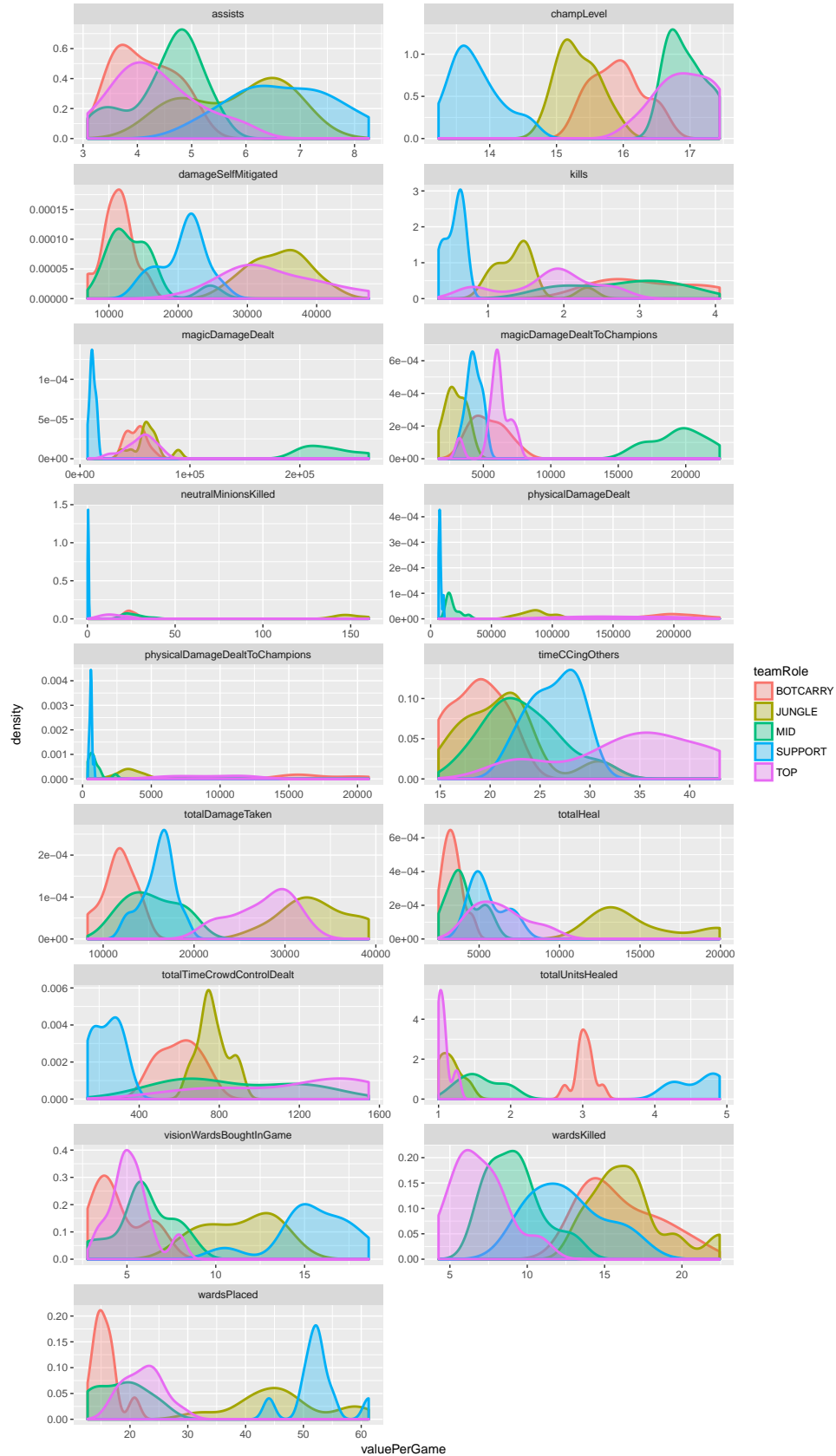
Player Averages Per Game Histograms, NALCS 2018 Spring Split

Distribution of Values



```
nalcs_plot_player_avgs_gathered +  
  geom_density(mapping = aes(x = valuePerGame, color = teamRole, fill = teamRole),  
    alpha = .3, size = 1) +  
  facet_wrap(~varName, scales = "free", ncol = 2) +  
  labs(  
    title = "Player Averages Per Game Density Plots, NALCS 2018 Spring Split",  
    subtitle = "Distribution of Values")
```

Player Averages Per Game Density Plots, NALCS 2018 Spring Split
Distribution of Values

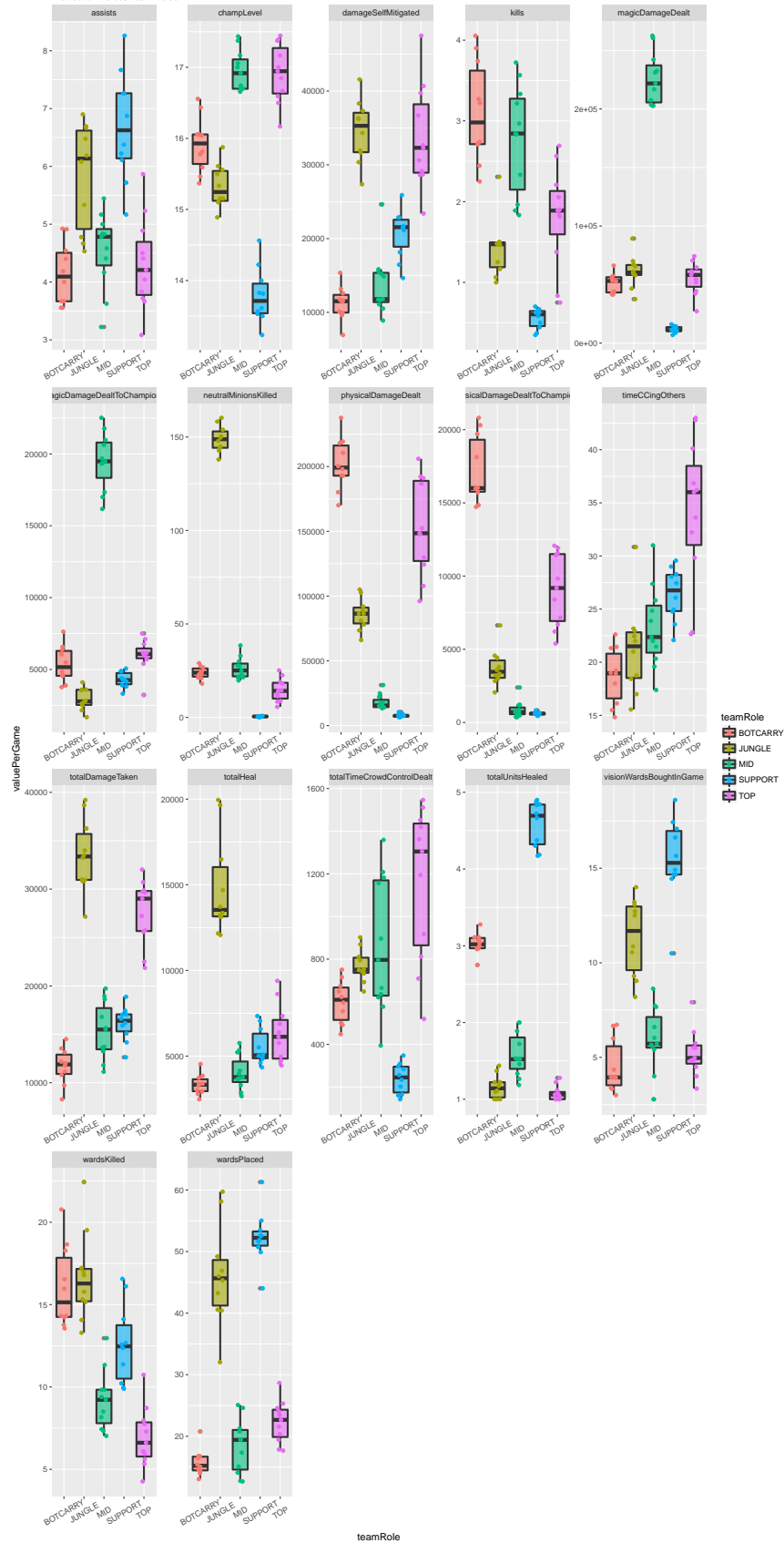


```

# Box Plots
nalcs_plot_player_avgs_gathered +
  geom_boxplot(mapping = aes(x = teamRole, y = valuePerGame, fill = teamRole),
    size = 1.00, alpha = .6) +
  geom_jitter(width = 0.15, mapping = aes(x = teamRole, y = valuePerGame, color = teamRole)) +
  facet_wrap(~ varName, scales = "free", ncol = 5) +
  theme(axis.text.x = element_text(angle=30, vjust=0.6)) +
  labs(
    title = "Player Averages per Game Box Plots, NALCS 2018 Spring Split",
    subtitle = "Distribution Across Team Roles")

```

Player Averages per Game Box Plots, NALCS 2018 Spring Split
Distribution Across Team Roles



As we can see through these plots, team roles can be clearly defined by certain characteristics and statistics:

Top Laner: least assists, least healing others, most true damage dealt, most CC time dealt

Jungler: most heal, most total damage taken and self-mitigated, highest vision score, most wards killed, most neutral minions killed

Mid Laner: most magic damage dealt to champions, one of highest finishing champ levels

Bottom Lane Carry: most total damage dealt to champions, most physical damage dealt, least heal, least total damage taken

Support: lowest kills, lowest damage dealt, highest assists, most unites healed, most vision wards bought, lowest finishing champ level

We will use this dataset and lean on such distinguishable statistics to eventually make a classification model.