Descriptive Analysis
4/22/2017

Teams Winning the Tournament

Which 5 teams have won the most NCAA titles?

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(plyr)
```

```
## -------------------------------------------------------------------------

## You have loaded plyr after dplyr - this is likely to cause problems.
## If you need functions from both plyr and dplyr, please load plyr first, then dplyr:
## library(plyr); library(dplyr)

## -------------------------------------------------------------------------

##
## Attaching package: 'plyr'

## The following objects are masked from 'package:dplyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize
```

```r
# Get the champion for all years
tourney <- read.csv('./data/TourneyCompactResults.csv')
teams <- read.csv('./data/Teams.csv')
titleTeams <- merge(tourney, teams, by.x="Wteam", by.y="Team_Id")
titleTeams <- titleTeams %>% filter(Daynum == 154)
titleTeams <- titleTeams[, c('Team_Name', 'Season')] %>% arrange(desc(Season))

# Count the frequency of winning teams from 1985-2017
titleCount <- count(titleTeams, 'Team_Name') %>% arrange(desc(freq))
```

Team	Titles
Duke	5
Connecticut	4
North Carolina	4
Kentucky	3
Florida	2

Most Titles (1985–2017)
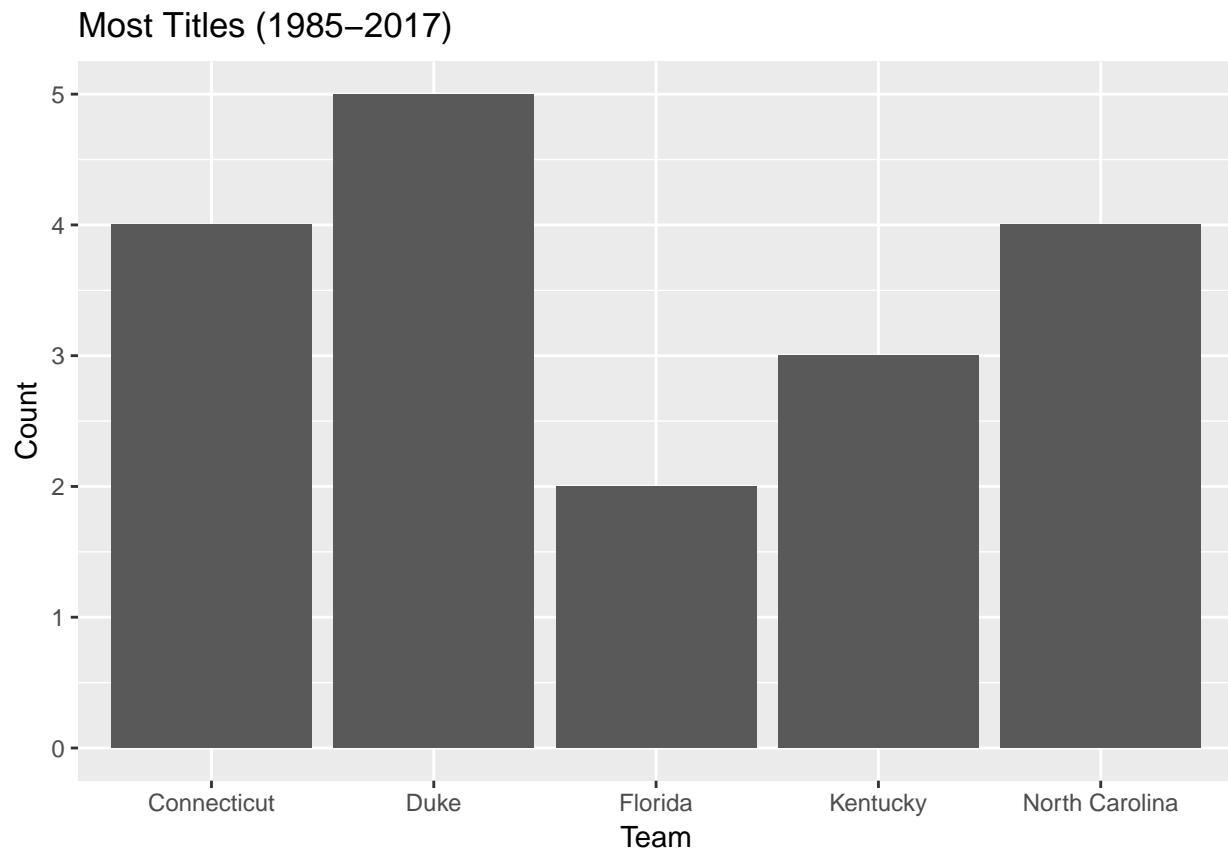
Teams in the Finals

Which teams have played in the championship game the most times?

```r
# We have the teams that have won the title, now we want teams that have lost in the finals.
finalsLosingTeams <- merge(tourney, teams, by.x="Lteam", by.y="Team_Id")
finalsLosingTeams <- finalsLosingTeams %>% filter(Daynum == 154)
finalsLosingTeams <- finalsLosingTeams[, c('Team_Name', 'Season')] %>% arrange(desc(Season))

# Combine all the winning teams with all the losing teams in the finals
finalsAppearances <- rbind(titleTeams, finalsLosingTeams)

# Count the frequency of winning teams from 1985-2017
finalsAppearancesCount <- count(finalsAppearances, 'Team_Name') %>% arrange(desc(freq))
```

Team	Finals Appearances
Duke	9
North Carolina	5
Kansas	5
Kentucky	5
Connecticut	4

## Most Finals Appearances (1985–2017)
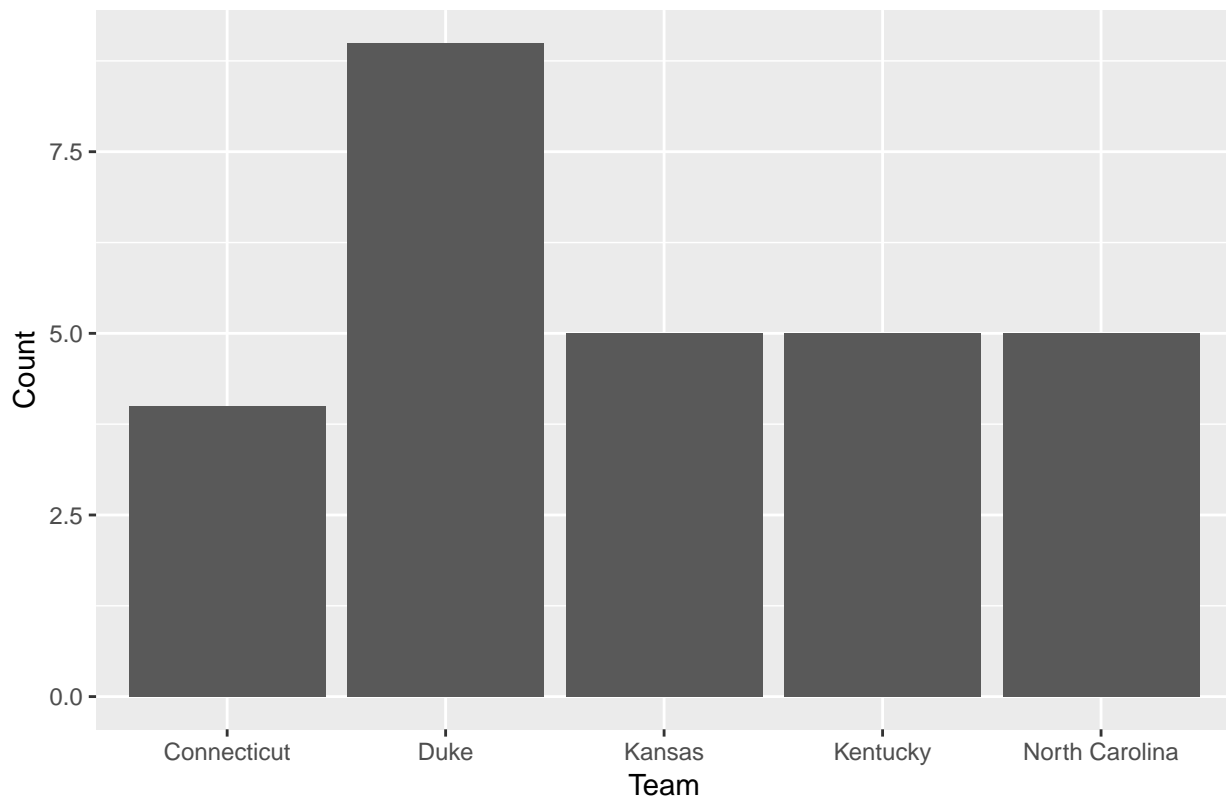
Teams in the Final Four

How about the 5 teams that have appeared in the Final Four the most times?

```
finalFourTeams <- tourney %>% filter(Daynum == 145 | Daynum == 146)
finalFourWinningTeams <- merge(finalFourTeams, teams, by.x="Wteam", by.y="Team_Id")
finalFourLosingTeams <- merge(finalFourTeams, teams, by.x="Lteam", by.y="Team_Id")
finalFourTeams <- rbind(finalFourWinningTeams, finalFourLosingTeams)
finalFourTeams <- finalFourTeams[, c('Team_Name', 'Season')] %>% arrange(desc(Season))

# Count the frequency
finalFourCount <- count(finalFourTeams, 'Team_Name') %>% arrange(desc(freq))
```

Team	Appearances
North Carolina	17
Kentucky	16
Kansas	14
Duke	14
Arizona	10

Most Final Four Appearances (1985–2017)
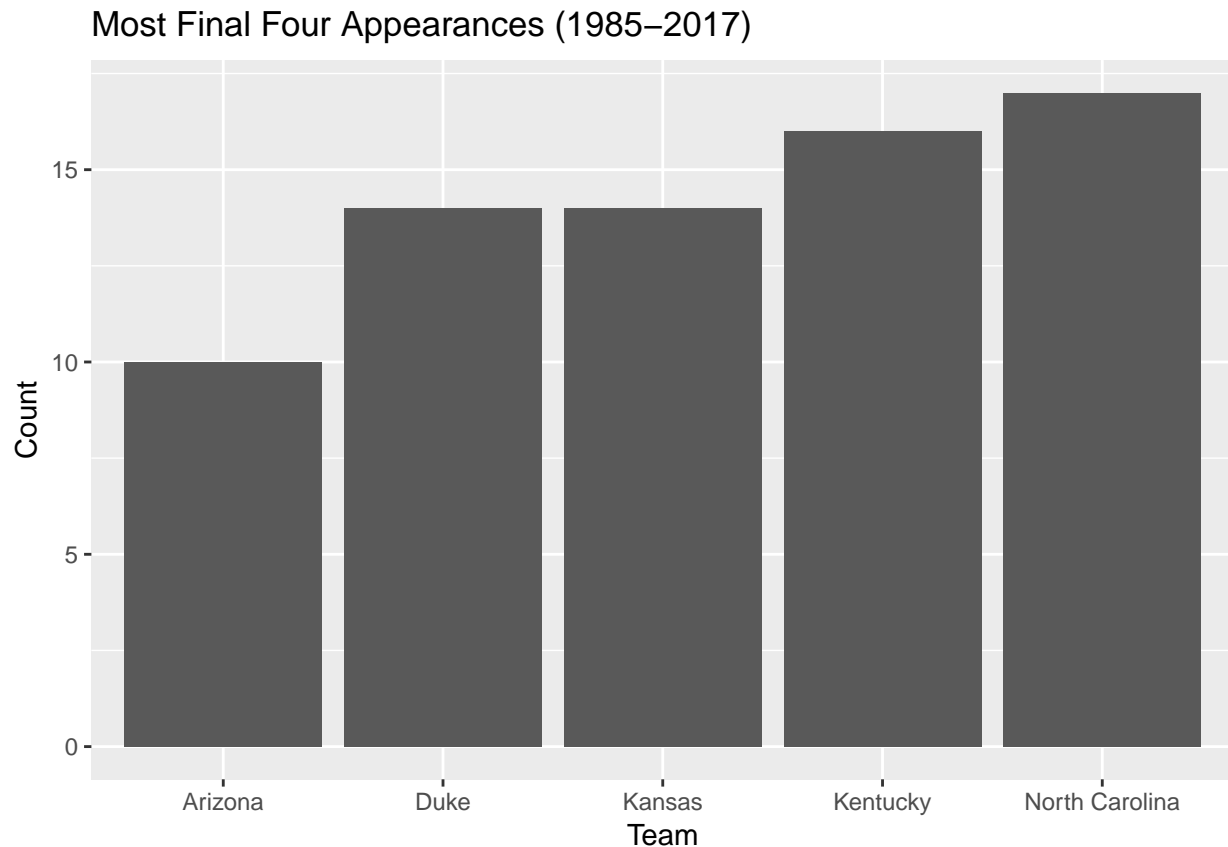
Significance of Statistics over Time

How have TOPG and PCT for the winning team from 2003-2017 changed over time?

```r
# All team season stats from 1985-2017
seasonStats <- read.csv('./data/FinalStats.csv')[,2:15]

# All championship teams from 1985-2017 with Team_Id column
titleTeams <- tourney %>% filter(Daynum == 154)
titleTeams <- titleTeams[, c('Wteam', 'Season')] %>% arrange(desc(Season))

# Function to return the season stats for each title team
getTitleTeamStats <- function(seasonStats, titleTeams) {
  datalist = list()
```

```r
  for (i in 1:nrow(titleTeams)) {
    datalist[[i]] <- subset(seasonStats, Season == titleTeams[i, 'Season'] & TeamID == titleTeams[i, 'W-
  }
  result <- bind_rows(datalist)
  result <- merge(result, teams, by.x="TeamID", by.y="Team_Id")
  result <- result[, !(names(result) %in% c('TeamID'))] %>% arrange(desc(Season))
  return(result)
}

# Title team season stats for years 2003-2017
titleTeamStats <- getTitleTeamStats(seasonStats, titleTeams)
```

Title Team Season Stats (2003–2017)
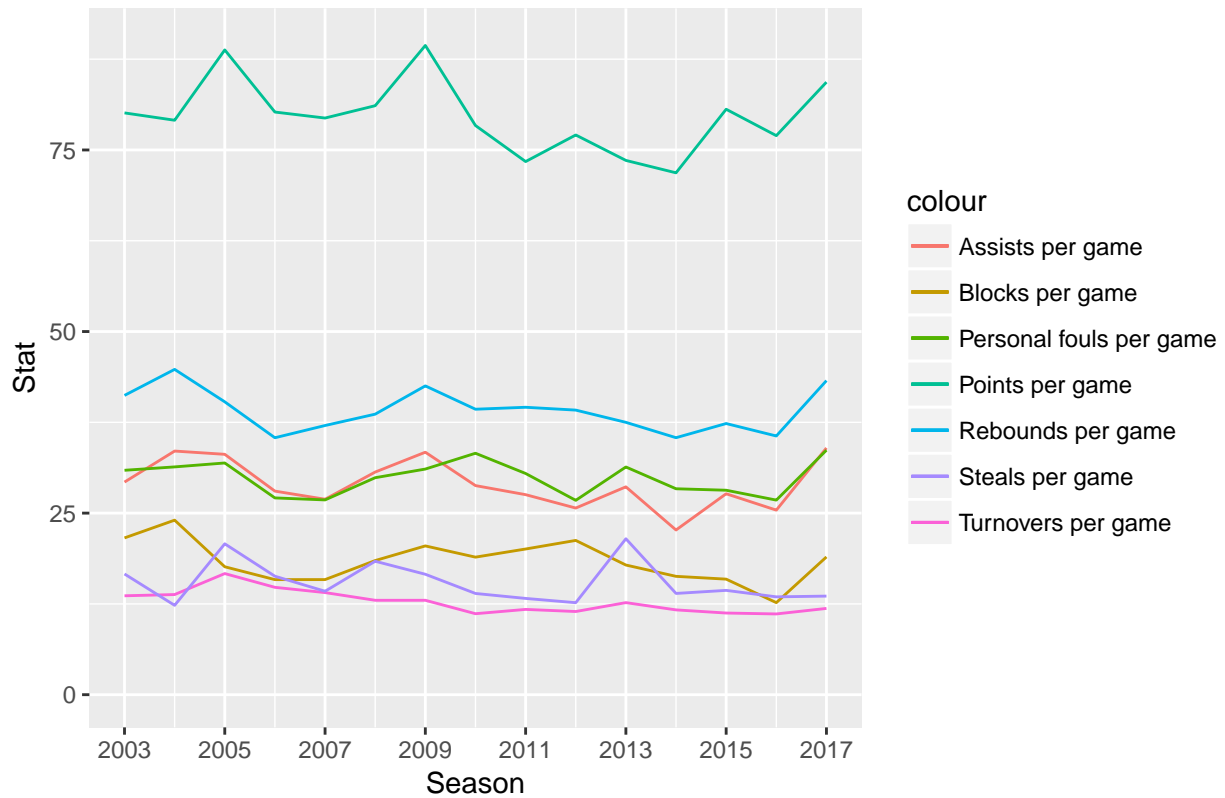
Title Team Season Percentage–based Stats (2003–2017)