

# ggplot2

Trevor Isaacson

2022-09-13

Instructions: Upload graph made with ggplot2 to GitHub. This can be any plot you want but must be done with ggplot2. You can use additional packages for data processing and graphing (e.g. you can use packages that build on ggplot2). Make sure that both the source code and final plot are available. If possible, share the data or the code used to obtain the data from online sources.

```
# read data
ncaaf_standings = read.csv("2021_NCAAF_standings.csv")

# rename and remove certain columns
ncaaf_standings = subset(ncaaf_standings, select = -c(Polls, X.9, X.10, X.11))
colnames(ncaaf_standings) = c("Rank", "School", "Conference", "Overall_Wins",
                              "Overall_Loses", "Overall_WinPerc", "Conf_Wins",
                              "Conf_Loses", "Conf_WinPerc", "Off_PPG",
                              "Def_PPG", "SRS", "SOS")

# remove first row b/c it isn't a school
ncaaf_standings = ncaaf_standings[-c(1), ]

# change variables types from chr to double
ncaaf_standings = ncaaf_standings %>%
  mutate(Overall_Wins = as.numeric(Overall_Wins)) %>%
  mutate(Overall_Loses = as.numeric(Overall_Loses)) %>%
  mutate(Overall_WinPerc = as.numeric(Overall_WinPerc)) %>%
  mutate(Conf_Wins = as.numeric(Conf_Wins)) %>%
  mutate(Conf_Loses = as.numeric(Conf_Loses)) %>%
  mutate(Conf_WinPerc = as.numeric(Conf_WinPerc)) %>%
  mutate(Off_PPG = as.numeric(Off_PPG)) %>%
  mutate(Def_PPG = as.numeric(Def_PPG)) %>%
  mutate(SRS = as.numeric(SRS)) %>%
  mutate(SOS = as.numeric(SOS))

# remove conference divisions
ncaaf_standings = ncaaf_standings %>%
  mutate(Conference = sub(".(East).*", "", Conference)) %>%
  mutate(Conference = sub(".(West).*", "", Conference)) %>%
  mutate(Conference = sub(".(Atlantic).*", "", Conference)) %>%
  mutate(Conference = sub(".(South).*", "", Conference)) %>%
  mutate(Conference = sub(".(North).*", "", Conference)) %>%
  mutate(Conference = sub(".(Mountain).*", "", Conference)) %>%
  mutate(Conference = sub(".(Coastal).*", "", Conference))

head(ncaaf_standings)
```

##	Rank	School	Conference	Overall_Wins	Overall_Loses
## 2	1	Wake Forest	ACC	11	3
## 3	2	Clemson	ACC	10	3
## 4	3	North Carolina State	ACC	9	3
## 5	4	Louisville	ACC	6	7
## 6	5	Florida State	ACC	5	7
## 7	6	Boston College	ACC	6	6

##	Overall_WinPerc	Conf_Wins	Conf_Loses	Conf_WinPerc	Off_PPG	Def_PPG	SRS	SOS
## 2	0.786	7	2	0.778	41.0	28.9	11.28	0.99
## 3	0.769	6	2	0.750	26.3	14.8	11.43	2.66
## 4	0.750	6	2	0.750	33.1	19.7	10.22	0.39
## 5	0.462	4	4	0.500	31.6	27.3	4.03	2.88
## 6	0.417	4	4	0.500	27.6	26.5	0.64	2.22
## 7	0.500	2	6	0.250	24.7	22.2	-1.00	-1.50

```

offense = ggplot(data = ncaaf_standings,
                  aes(x = Off_PPG, y = Overall_WinPerc, color = Conference)) +
  theme_minimal() +
  geom_point() +
  scale_color_viridis(discrete = TRUE) +
  xlab("Points Scored per Game") + ylab("Win Percentage") +
  xlim(10, 50) + ylim(0, 1) +
  plot_annotation(
    title = "Average Number of Points Scored and Overall Win Percentage",
    subtitle = "Comparing the average number of points scored for each FBS football team and
               their overall win percentage for the 2021 season",
    caption = "Data from Sports Reference"
  )

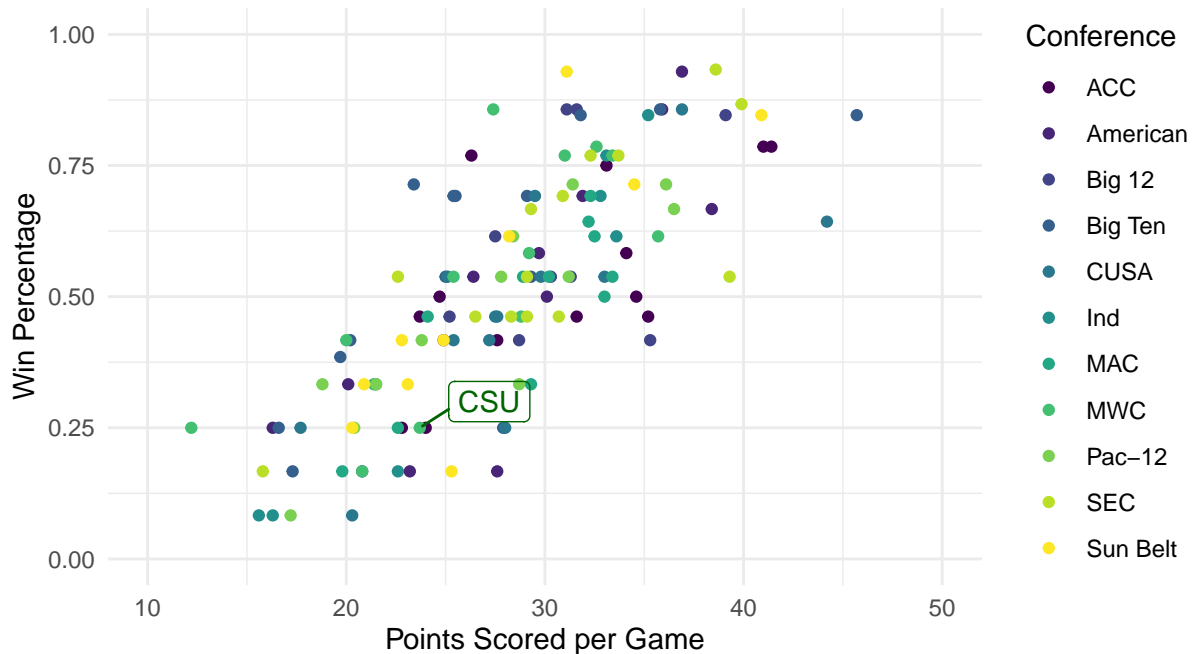
offense = offense +
  ggrepel::geom_label_repel(data = . %>% filter(School == "Colorado State"),
                            aes(label = "CSU"),
                            color = "darkgreen", show.legend = FALSE, fill = alpha(c("white"), 0),
                            nudge_y = 0.05, nudge_x = 3.5)

offense

```

## Average Number of Points Scored and Overall Win Percentage

Comparing the average number of points scored for each FBS football team and their overall win percentage for the 2021 season



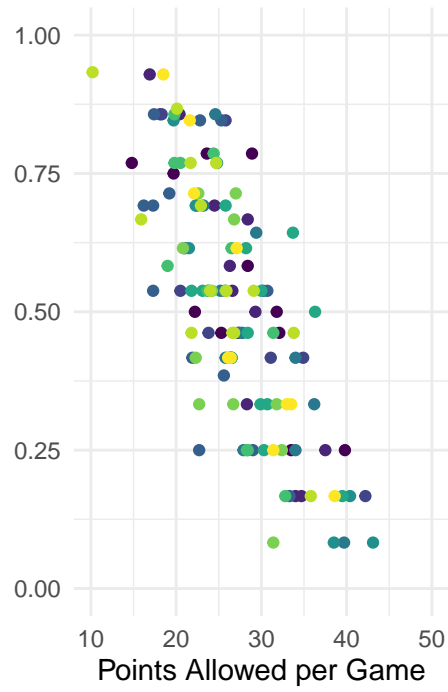
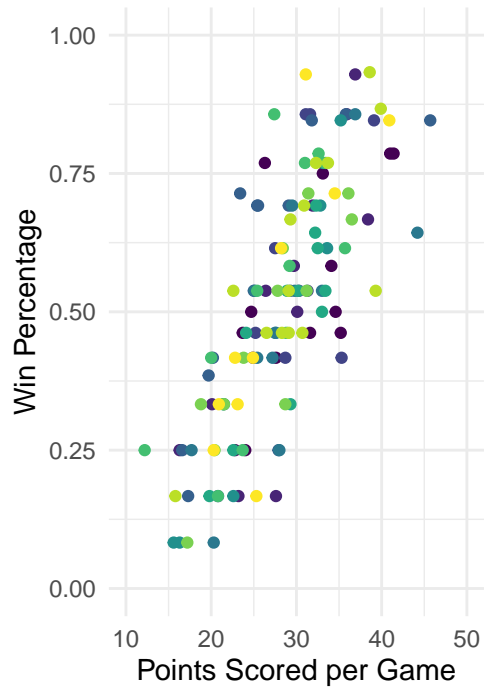
Data from Sports Reference

```
ncaaf_offense = ggplot(data = ncaaf_standings,
  aes(x = Off_PPG, y = Overall_WinPerc, color = Conference)) +
  theme_minimal() +
  geom_point() +
  xlab("Points Scored per Game") + ylab("Win Percentage") +
  theme(legend.position = "none") +
  scale_color_viridis(discrete = TRUE) +
  xlim(10, 50) + ylim(0, 1)

ncaaf_defense = ggplot(data = ncaaf_standings) +
  theme_minimal() +
  geom_point(aes(x = Def_PPG, y = Overall_WinPerc, color = Conference)) +
  scale_color_viridis(discrete = TRUE) +
  xlab("Points Allowed per Game") + ylab("") +
  xlim(10, 50) + ylim(0, 1)

ncaaf = ncaaf_offense + ncaaf_defense
ncaaf = ncaaf +
  plot_annotation(
    title = "",
    subtitle = "",
    caption = "Data from Sports Reference"
  )

ncaaf
```



#### Conference

- ACC
- American
- Big 12
- Big Ten
- CUSA
- Ind
- MAC
- MWC
- Pac-12
- SEC
- Sun Belt

Data from Sports Reference