

# March Madness

```
library(readr)
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

```
files <- list.files("March Madness Regular Season Data", pattern = "\\.csv$", full.names = TRUE)

regular_season_df <- files |>
  lapply(function(f) read_csv(f, col_types = cols(.default = "c")))) |>
  bind_rows()
```

```
getwd()
```

[1] "/home/guest/sta345-final"

```
list.files()
```

```
[1] "March Madness Regular Season Data"  
[2] "MarchMadness_files"  
[3] "MarchMadness.qmd"  
[4] "MarchMadness.rmarkdown"  
[5] "MarchMadness.tex"  
[6] "sta345-final.Rproj"  
[7] "tourney_results_mod_through_2025.csv"
```

```
list.files("March Madness Regular Season Data")
```

```
[1] "ncaa_tr_regular_2000_2001.csv" "ncaa_tr_regular_2002_2003.csv"  
[3] "ncaa_tr_regular_2004_2005.csv" "ncaa_tr_regular_2006_2007.csv"  
[5] "ncaa_tr_regular_2008_2009.csv" "ncaa_tr_regular_2010_2011.csv"  
[7] "ncaa_tr_regular_2010_2025.csv" "ncaa_tr_regular_2012_2013.csv"  
[9] "ncaa_tr_regular_2014_2015.csv" "ncaa_tr_regular_2016_2017.csv"  
[11] "ncaa_tr_regular_2018_2019.csv" "ncaa_tr_regular_2020_2021.csv"  
[13] "ncaa_tr_regular_2022_2023.csv" "ncaa_tr_regular_2024_2025.csv"
```

```
tourney_result <- read_csv("tourney_results_mod_through_2025.csv")
```

Rows: 2585 Columns: 15

```
-- Column specification -----  
Delimiter: ","  
chr (5): wloc, wteam_school, lteam_school, wteam_region, lteam_region  
dbl (10): season, daynum, wteam, wscore, lteam, lscore, numot, wteam_seed, l...  
  
i Use `spec()` to retrieve the full column specification for this data.  
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```