# project preprocessing

## Takshsheel Goswami

## 2024-04-23

## Including Plots

You can also embed plots, for example:

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
salary_data = read.csv("salary_Proj.csv")
nflverse_data = read.csv("nflverse.csv")

names(salary_data)[1] <- "player_display_name"
total <- merge(salary_data,nflverse_data,by="player_display_name")
```

```
#For each player, select all instances. Averaege all stats. Return avg.
total = subset(total, select = -c(player_name,player_id,X,Position,season_type,headshot_url,recent_team

library(dplyr)
total <- total %>%
  group_by(player_display_name,position,position_group) %>%
  summarise_all(mean)

#Combine position back
table(total$position_group)
```

```
##
##   DB   DL   LB   OL   QB   RB SPEC   TE   WR
##    2    2    2    3   53   55    9   51   87
```