install.packages(“tinytex”) tinytex::install_tinytex() — title: “Untitled” output: pdf_document date: “2025-01-09” —

Load necessary libraries

library(dplyr) getwd() setwd(“~/Desktop/538”) game_stats <- read.csv(“GameStats.csv”) View(game_stats)

Remove games played on a neutral field

game_stats <- game_stats %>% filter(X != “N”)

Create a unique game identifier

game_stats <- game_stats %>% mutate(GameID = rep(1:(n()/2), each = 2))

Split the dataset into home and away teams

home_stats <- game_stats %>% filter(X == “@”) away_stats <- game_stats %>% filter(X != “@”)

Merge the home and away datasets

cleaned_stats <- merge(home_stats, away_stats, by = “GameID”, suffixes = c(“_home”, “_away”))

Calculate HomeWins based on the sum of PassTD, RushTD, and KickPts

cleaned_stats <- cleaned_stats %>% mutate(HomeWins = ifelse((PassTD_home + RushTD_home + KickPts_home) > (PassTD_away + RushTD_away + KickPts_away), 1, 0))

Create the cleaned dataset with required columns

cleaned_stats <- cleaned_stats %>% select(Date = Date_home, Home = School_home, Away = School_away, HomeWins, HPassCmp = PassCmp_home, APassCmp = PassCmp_away, HPassAtt = PassAtt_home, APassAtt = PassAtt_away, HPassPct = PassPct_home, APassPct = PassPct_away, HPassYds = PassYds_home, APassYds = PassYds_away, HPassTD = PassTD_home, APassTD = PassTD_away, HRushAtt = RushAtt_home, ARushAtt = RushAtt_away, HRushYds = RushYds_home, ARushYds = RushYds_away, HRushAvg = RushAvg_home, ARushAvg = RushAvg_away, HRushTD = RushTD_home, ARushTD = RushTD_away, HXPM = XPM_home, AXPM = XPM_away, HXPA = XPA_home, AXPA = XPA_away, HXPPercent = XPPercent_home, AXPPercent = XPPercent_away, HFGM = FGM_home, AFGM = FGM_away, HFGA = FGA_home, AFGA = FGA_away, HFGPercent = FGPercent_home, AFGPercent = FGPercent_away, HKickPts = KickPts_home, AKickPts = KickPts_away, HFum = Fum_home, AFum = Fum_away, HInt = Int_home, AInt = Int_away, HTotalTO = TotalTO_home, ATotalTO = TotalTO_away)

Write the cleaned dataset to a CSV file

output_path <- "~/Desktop/538/CleanedGameStats.csv" write.csv(cleaned_stats, output_path, row.names = FALSE)

head(cleaned_stats)