Vortex of Accuracy Version Notes

**Vortex of Accuracy**

* **0.x**
  + Version 0.0.1
    - Creation of spreadsheet, inclusion of AP/Coaches/CFP polls
  + Version 0.1
    - Inclusion of team stats (avg off/def yds/game, pts/game)
  + Version 0.2
    - Win/loss record
  + Version 0.3
    - Use of heavily flawed, semi-incomplete CBS Sports ranking to calculate opponent strength
* **1.x**
  + Version 1.0
    - Switch to ESPN FPI to determine average opponent strength
  + Version 1.1
    - Variable created to account for FCS teams
    - Version 1.1.1
      * Opponent FPI + inclusion of FCS opponent coefficient added to model
    - Version 1.1.2
      * Average FPI of losses added as variable
  + Version 1.2
    - Inclusion of FPI Team Efficiency Indices
  + Version 1.3
    - Addition of MOV stat
    - Version 1.3.1
      * Average Margin of Loss added, only measures loss margins
  + Version 1.4
    - Inclusion of 3rd down %, turnovers generated stats
  + Version 1.5
    - Inclusion of Top 25 win/loss record
    - Version 1.5.1
      * Also adding NonTop 25 W/L record, to mitigate any undeserved reward teams who play 0 Top 25 teams get from #1 ranking on Top 25 loss variable
  + Version 1.6
    - Creation of yard differential stat, derived from existing yards per game stats
    - Version 1.6.1
      * Efficiency stat of points per yard added
  + Version 1.7
    - Creation of conference champ stat, rewards teams winning their conference
    - Version 1.7.1
      * Tweaking of stat, punishes teams more for finishing 2nd compared to winning conference
    - Version 1.7.2
      * Added test rankings to serve as evaluation ground for potential new variables and tiebreakers in case of emergency
  + Version 1.8
    - Addition of SP+ data
    - Version 1.8.1
      * Inclusion of SP+ special teams data
    - Version 1.8.2
      * Addition of conference rank, stolen from Bill Connelly’s SP+ conference averages
    - Version 1.8.3
      * Creation of Mean of FPI and SP+ rankings
    - Version 1.8.4
      * Inclusion of Median
    - Version 2.8.5
      * Mean and Median of All Polls/Rankings
* 2**.x**
  + Version 2.0
    - First creation of R script
  + Version 2.1
    - Use of dense\_rank function so there are no gaps in the ranking columns (i.e. going from 12-13 in dense\_rank vs going from 12,12,12,12-16 in regular ranking function)
    - Version 2.1.1
      * Use of inverse for variables ranked in descending order, dense\_rank only works in ascending order
    - Version 2.1.2
      * Fixing data so negative variables (yd diff, MOV) do not result in teams falsely ranking higher than they should
    - Version 2.1.3
      * Labelling UConn as Independent for upcoming season due to conference realignment
    - Version 2.1.4
      * Added Adjusted MOV and Adjusted MOL variables to attempt to account for “higher quality” wins and losses
    - Version 2.1.5
      * Updated Adjusted MOL variable so as to be able to rank it in ascending order
  + Version 2.2
    - Adopted use of gt package to create final Top 25 table instead of grid.table
* 3**.x**
  + Version 3.0
    - scraping table straight from ESPN’s website
    - including offensive and defensive stats
      * including special teams stats for first time
        + FG accuracy, XP accuracy, kick and punt return yardage and TDs
    - included standard FPI rank and efficiency metrics but also added resume stats such as SOR, SOS, Game control, Avg WinProb
    - tested including 2019 data from 2020 opt outs (ODU, NMSU, UConn), decided against using this as final version, 2020 rankings included in 2021 VoP will utilize VoA rankings NOT including 2019 data from 2020 opt outs
    - Version 3.0.1
      * added offensive and defensive passing and rushing specific stats (YPA (pass and rush), YPC, sacks, INTs, etc)
  + Version 3.1
    - created IF statement which, depending on if week <= 4, scrapes ESPN data not just from current season, but previous season as well
    - Version 3.1.1
      * includes data from 2019 season for UConn, NMSU, and ODU who all opted out of 2020 season due to COVID, incorporates it into 2020 table so that the 2020 stats table has all 130 teams in it just like the current season table
      * NOTE: due to errors when testing new format, week 1 stats are slightly incorrect for some teams (most likely just teams which played on Friday of week 2). PNGs of tables containing team, VoA\_Output, and VoA\_Ranking are accurate, but stats are only accurate for teams which did not play before Friday of Week 2. VoA\_Output and VoA\_Ranking are manually updated in the spreadsheet based on the PNG tables
    - Version 3.1.2
      * Added code creating unintelligible chart for each conference, including independents
      * Changed model to begin phase out of previous season data after Week 3 (so in Week 4 running of VoA) instead of after Week 2
    - Version 3.1.3
      * Added Yards per play stat for both offense and defense
        + total yds/ (rush attempts + pass attempts)
        + added for 2021 week 5 VoA

will not be retroactively included in weeks 1-4 data, too much work and I do not feel like doing it, might not even be possible

* + - * also added a variable for taking the mean of FPI efficiency ranks
        + \*not\* efficiency values, but future updates will likely include this, meaning efficiency values will have to be dense\_rank() just like normal stats
        + mean of just offensive and defensive efficiencies, and separate mean of offensive, defensive, special teams efficiencies
      * FPI variables now weighted twice for current season
        + FPI mean of offensive, defensive, and special teams efficiencies double weighted, but not mean of just FPI offensive and defensive efficiencies
      * edited AAC unintelligible chart to better fit jpeg so title isn’t cut off
      * Updated gt tables so that VoA\_Output column has color scale matching values
        + removed beige background for full table
    - Version 3.1.4
      * added “points per play” stat for both offense and defense
        + fixed error where defensive YPP had not actually been added yet

2021 Weeks 0-5 VoA run without defensive YPP and def Pts/play as a result

* + - * added Adjusted MOV stat to attempt to account for opponent strength in MOV
        + MOV divided by ESPN’s FPI’s SOS
    - Version 3.1.5
      * added Adjusted MOV stat for ESPN’s FPI’s SOR
        + same as SOS-Adjusted MOV, but with SOR
    - Version 3.1.6
      * ESPN updated their team names to no longer include the nickname in the last fucking week of the season, so I did too, Unintelligible Charts from Week 13-on will reflect school names, not full school name plus nickname
      * added histograms reflecting team VoA Outputs for individual weeks, will be present for 2021 week 14 and week 15 (bowl week/CFP)
      * created scatterplot of VoA Output over VoA Ranking
        + not exporting for now, it’s not that meaningful
      * Moved stat page URLs to top of script, will make for easier updating
    - Version 3.1.7
      * added a Pythagorean win stat, using formula from Football Outsiders
        + only included for final 2021 week (Week 15), will not be included in whole season, will be included for whole of future seasons
      * fixed error in points per play stat where points per game was being used for calculation instead of total points
        + this probably didn’t affect the rankings too much, but may have negatively impacted teams which played an extra game (conference championship, maybe some bowl games)
    - Version 3.1.8
      * added adjusted net yards per passing attempt stat, using ESPN’s Bill Connelly’s formula
        + net yards per attempt, plus 20 yards per TD and -45 per INT
    - Version 3.1.9
      * ESPN updated their team names back to how it had been for all but the last fucking week of the season, so I had to too, meaning all team names include school name plus nickname
    - Version 3.1.10
      * updated script to export final data frame as csv, not excel file and to read in csv’s instead of xlsx files for previous week data
* 4**.x**
  + Version 4.0
    - switching to use cfbfastR as data source
    - reading in both regular and advanced stats, as well as including SP+ and FPI data as available
    - including recruiting and talent metrics as well, which are sourced from 247 Sports via cfbfastR
    - James Madison’s previous year data taken from **[stats.ncaa.org](http://stats.ncaa.org)**
      * advanced stats derived via taking the mean of each Group of 5 team for each stat, then assigning it to JMU
      * exception is 2021 SP+ Overall, Offense, and Defense ratings, which were downloaded from ESPN’s Bill Connelly’s SP+ google doc as a csv
    - adds basic linear model from {stats} package in R to create FPI/SP+ style metric
      * attempts to predict FPI/SP+ mean
    - VoA\_Output still included as mean rank of weighted stats, then included as predictor in linear model
    - VoA\_Ranking no longer determined from VoA\_Output alone but from VoA\_Rating, the name currently given to the linear model predicted output
    - 3 years worth of previous data is included and filtered out over the course of first 5 weeks
      * PY3 included for weeks 0 (preseason VoA) and 1
      * PY2 included week 0 through week 4
      * PY1 included through week 5
    - Unintelligible charts now represent VoA\_Rating and VoA Ranking from week to week
      * VoA\_Output no longer included in the charts
    - creates scatter plot of VoA\_Output and VoA\_Rating to examine contrast of prior method of deriving VoA\_Ranking and current method of using VoA\_Rating
    - final data frame now exported as csv instead of xlsx file
      * exported to VoAyear folder within Data folder instead of to Outputs folder
      * tables, charts, and graphs still exported to Outputs folder
    - **Version 4.0.1**
      * SP+ data unlikely to be available (at least not with any reliable regularity) via CFBData API and cfbfastR due to paywall issues
        + as such, current season SP+ data is now commented out and not used in weeks 1- end of season
        + FPI/SP+ mean no longer used as “predicted” variable in lm()
        + FPI alone serves as “predicted” variable used to create FPI/SP+ style metric
      * fixed error where Previous Year FPI/SP+ mean was used as “predicted” variable in model for Week 1 and beyond
        + switched to just FPI due to SP+ availability issues
      * fixed bug that could result in error where Advanced Stats might not be available for all teams even though Regular stats would be since all teams have played games
        + should be able to work regardless of which teams have missing data
        + currently only set up for week 1 in hope that any data issues will be resolved by the time the next week’s VoA is run
      * added small bit of code to remove season, conference, and current year recruiting points columns from csv containing all 3 previous years of JMU data
        + those columns added via current year dataframe
      * removed code that takes entire VoA\_Variables data frame and converts all NAs to 0s
        + is.na() = 0 is now done after pulling in each individual data frame (stats, advanced stats, FPI, SP+) before merging
      * fixed bug where ESPN’s new team names weren’t corrected to match what is standard in {cfbfastR}
      * updated rowmeans() function for Week 1 to properly reflect which columns are ranks and thus which to take the mean of when calculating VoA Output
      * fixed bug where JMU recruiting doesn’t show up in {cfbfastR} recruiting function (despite other FCS teams being there?) for current season, current season JMU recruiting is the median of the whole FBS recruiting