# Trends in NFL Player Representation from SEC Schools (2011–2025)

Omar Sepulveda, Riccardo Gutierrez, Naomi Menard

2025-11-06

Github Link : https://github.com/rdgutierrez2005/Stat107Project.git

```
# This chunk loads the script to attach necessary packages like tidyverse and gridExtra.
# It sets up the environment for data manipulation and visualization.
# Outcome: Packages are installed and enabled

source("00_requirements.R")
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.1     v stringr   1.5.2
## v ggplot2   4.0.0     v tibble    3.3.0
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.1.0
## -- 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
##
## Attaching package: 'gridExtra'
##
##
## The following object is masked from 'package:dplyr':
##
##     combine
```

## Abstract

The NCAA SEC is one of the premier conferences in college football from recent years, but how much is each program represented in the draft? We collected our data from Pro-Football-Reference.com and combined each program in the SEC data sets into a singular Excel file and analyzed the data by distinguishing which players came from which college program, seeing if the player was drafted, then drafted within the years 2011-2025. From there we used bar graphs and histograms to further comprehend the trends of the collegiate programs alumni to further analyze when each collegiate program was at its peak during the given time period. Hopefully our analysis can help future students choose which program and legacy they want to lead. For example whether they want to compete on an elite program or compete against an elite program within the conference.

# Introduction

The purpose of this analysis is to highlight which schools in the NCAA SEC (South Eastern Conference) have gained representation in the draft and which programs have lost representation in the draft over the last 15 years. This will answer the question: "How has the number of NFL players from each SEC school changed over time (e.g., 2011–2015, 2016–2020, and 2021–2025), and which schools have shown the most growth or decline in player representation?" This analysis may benefit various groups, including NFL scouts/teams, sports analysts/journalists, or other researchers looking to see which college improved the most. The goal is to understand patterns and changes throughout the 2011 - 2015, 2016-2020, and 2021-2025 time periods.

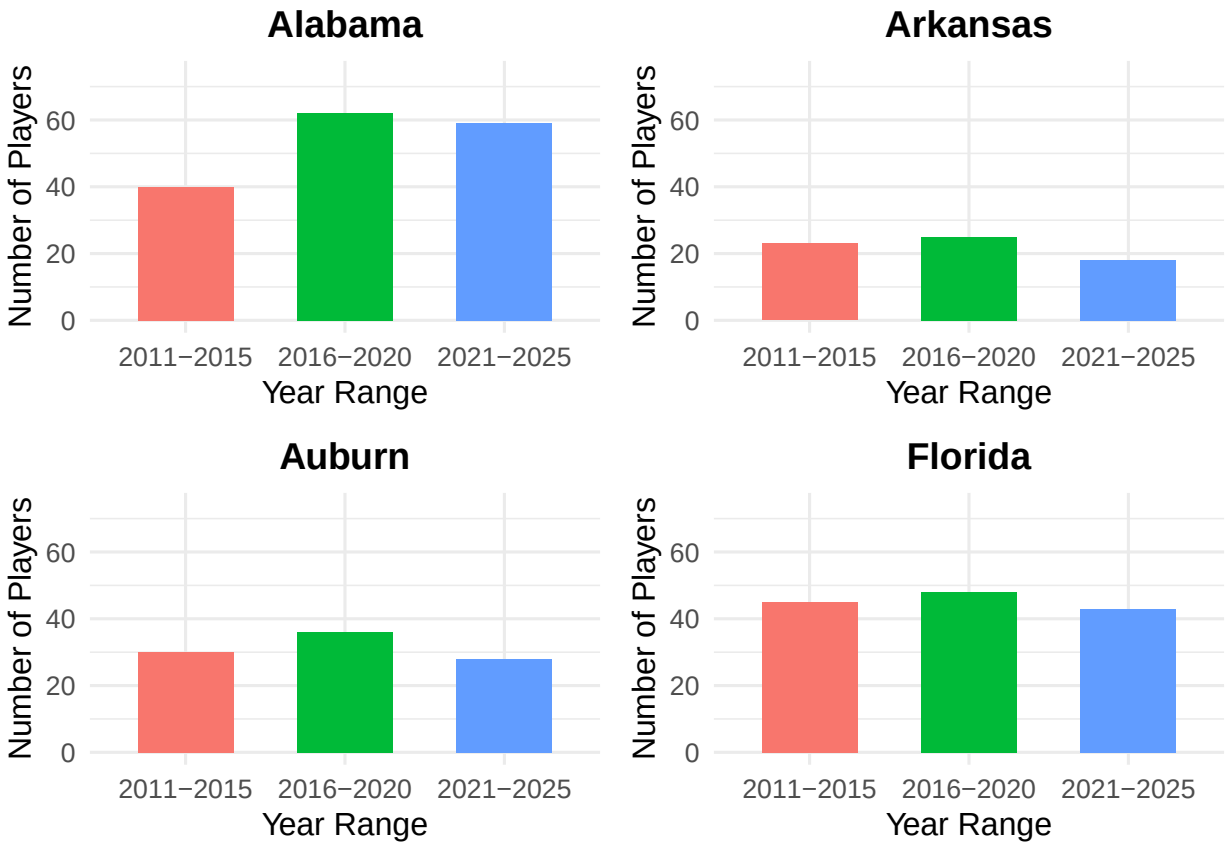
# Data

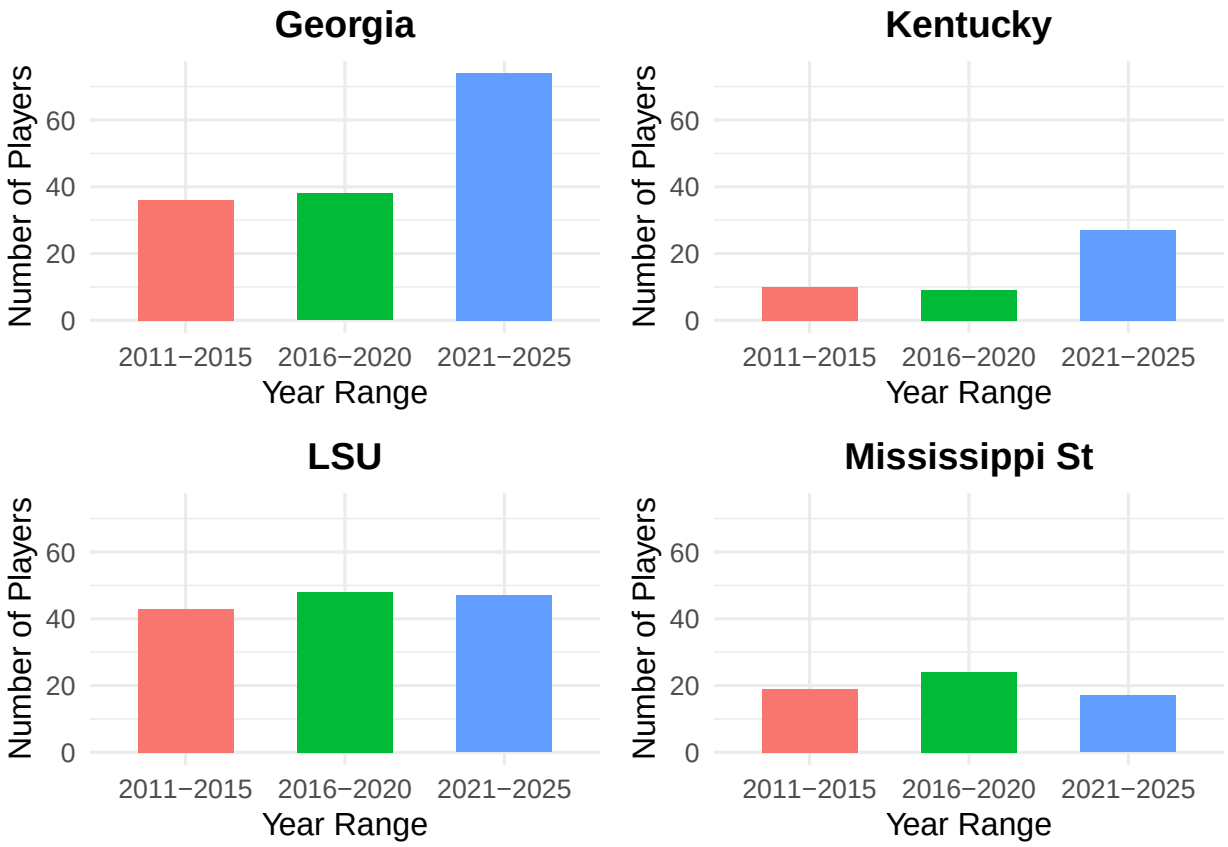
```
# This chunk sources the data cleaning script to process the raw dataset.
# It takes raw data, keeps/renames columns, filters to 2011+, and saves the cleaned data.
# Outcome: A cleaned dataset is made and saved as "cleaned_data.RDS" for later use.

source("01_funct_DataCleaning.R")
```

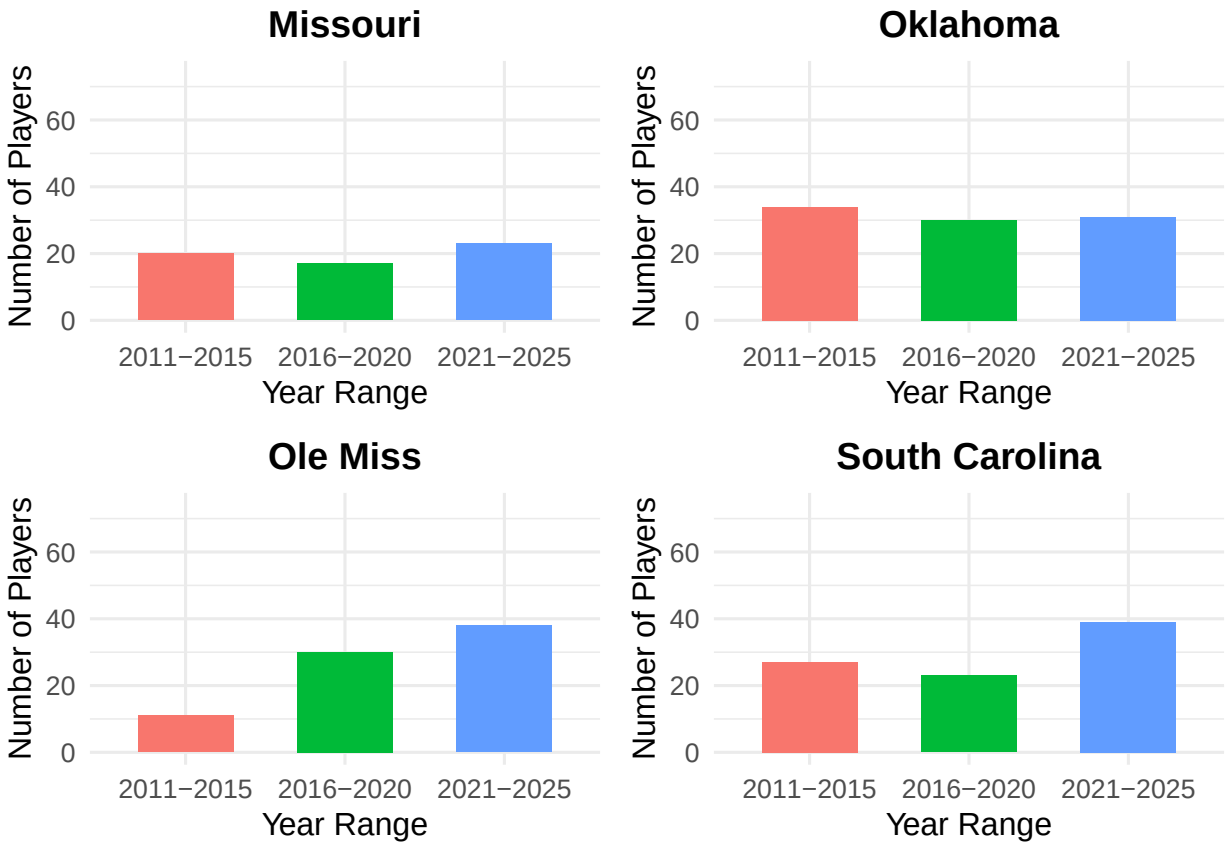
The data set used for this analysis, "College_Dataset.xlsx," contains information on NFL players and the colleges they attended. The data was obtained from Pro-Football-Reference, a reliable source that collects and publishes official NFL statistics, including player backgrounds and school affiliations. Relevant variables include each player's name, the college they attended, and the range of years they played in the NFL. During the cleaning process, unnecessary columns such as AP1, PB, St, wAV, Ht, and Wt were removed, keeping only the relevant variables. Additionally, since the data set included players from earlier eras, we extracted the starting year from each player's career range and filtered the data to include only those who began their NFL careers in 2011 or later. The data is entirely real and will not be generated through any randomized simulation.

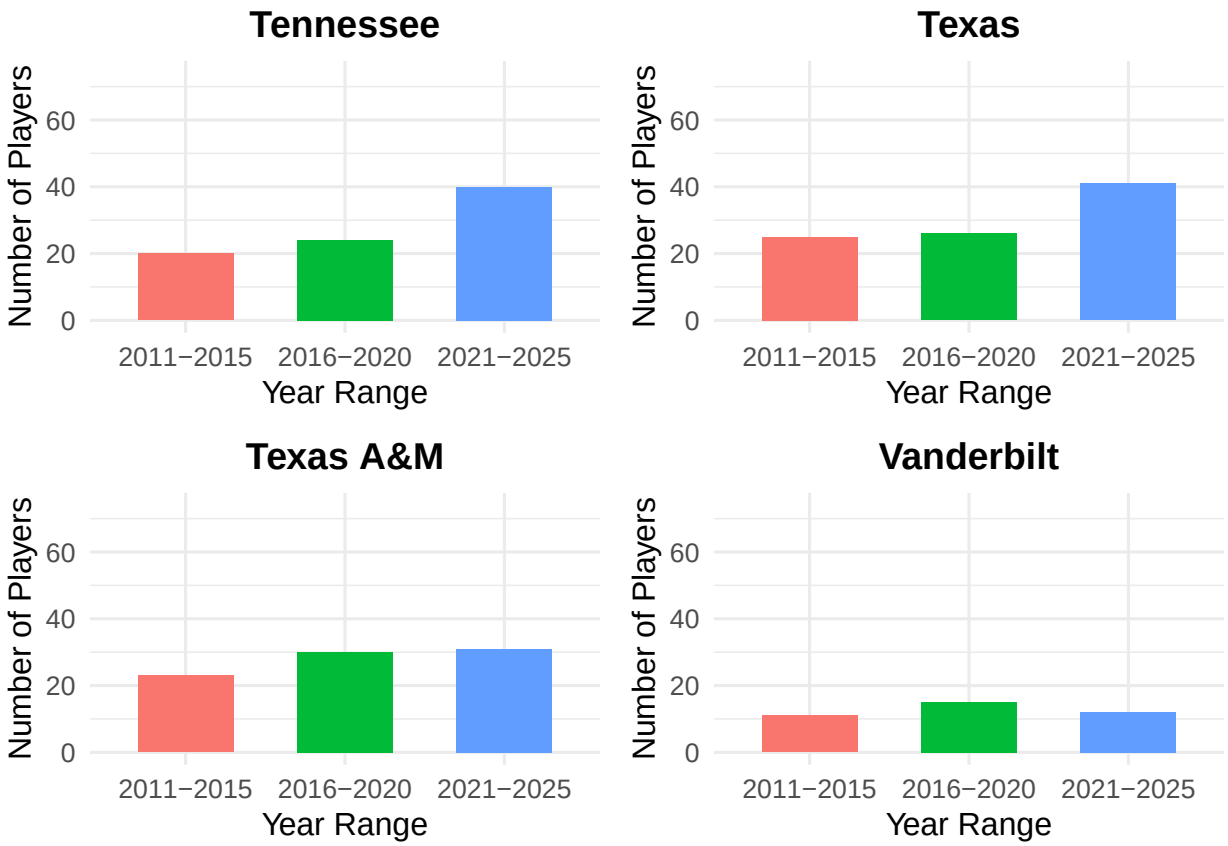
# Visualization & Analysis

```
# This chunk sources the plotting script to create bar plots.
# It creates bar charts for each SEC school, showing player counts across three year gaps.
# Outcome: Bar plots are created and displayed, visualizing trends in player representation.
source("02_funct_Plots.R")
```


Alabama
Number of Players
0
20
40
60
2011−2015
2016−2020
2021−2025
Year Range
Arkansas
Number of Players
0
20
40
60
2011−2015
2016−2020
2021−2025
Year Range
Auburn
Number of Players
0
20
40
60
2011−2015
2016−2020
2021−2025
Year Range
Florida
Number of Players
0
20
40
60
2011−2015
2016−2020
2021−2025
Year Range


Georgia
Kentucky
LSU
Mississippi St
Number of Players
0
20
40
60
2011−2015
2016−2020
2021−2025
Year Range


Missouri
Oklahoma
Ole Miss
South Carolina
Number of Players
0
20
40
60
2011−2015
2016−2020
2021−2025
Year Range




```
# This chunk sources the table script to summarize data numerically.
# It creates a tibble table with player counts by school and year range.
# Outcome: A table is generated and printed, providing exact counts for comparison.
source("03_numeric_Table.R")
```

```
## # A tibble: 16 x 4
##    School         `2011-2015` `2016-2020` `2021-2025`
##    <chr>                <int>       <int>       <int>
##  1 Alabama                 40          62          59
##  2 Arkansas                23          25          18
##  3 Auburn                  30          36          28
##  4 Florida                 45          48          43
##  5 Georgia                 36          38          74
##  6 Kentucky                10           9          27
##  7 LSU                     43          48          47
##  8 Mississippi St          19          24          17
##  9 Missouri                20          17          23
## 10 Oklahoma                34          30          31
## 11 Ole Miss                11          30          38
## 12 South Carolina          27          23          39
## 13 Tennessee               20          24          40
## 14 Texas                   25          26          41
## 15 Texas A&M               23          30          31
## 16 Vanderbilt              11          15          12
```

These preliminary visualizations support our future analysis by allowing us to compare the number of NFL

players produced by various SEC colleges across different time periods. By examining the bar plots, we can identify early patterns of growth or decline in player representation for each school. In later stages of the analysis, we will compare these bar plots in greater detail to draw conclusions about which colleges have shown the most progress or decline in producing NFL players over the years. Further visualizations we discussed implementing include histograms that display the trends for each collegiate program within the conference during the year 2011-2025 time period.

## Analysis

For the analysis, we plan to use descriptive statistics and visual models to track changes in the number of NFL players from each SEC school over the three time periods. We'll look at the overall trend for each school using a simple linear regression, which will help show whether the number of NFL players from that school is going up or down over time. We'll also use percentage change calculations to measure growth rates and rank the schools based on overall gains or declines. These methods should make it clear which schools have built a stronger NFL presence over time. We expect that powerhouse programs like Alabama and Georgia will show a steady growth, while schools like Vanderbilt and Auburn may show slight growth or even a decline. For further analysis discussing the histograms hopefully the year by year displays for each collegiate program in the conference will further contribute to our theory of steady growth for the elite programs while the lesser known programs won't display such trends. We expect to find that most if not all programs within the conference will display some form of growth within the 2011-2025 time period and then a variation of growth and decay during the covid year as not all programs held the same appeal during the pandemic. With big name schools such as Alabama, Georgia and Texas will display steady growth and the values will remain relatively high compared to lesser known schools such as Auburn and Vanderbilt.

## Contributions

Omar Sepulveda: Gathered data sets, completed data section, code used to clean data set, code for the visuals, completed analysis section, and wrote step by step process in README.md

Riccardo Gutierrez: Gathered data sets, completed abstract section, and contributed to visualization and analysis section.

Naomi Menard: Completed introduction, gathered data sets, contributed to data, visuals/code , and analysis section.