# Private School Estimates

## 2025-06-02

Data Cleaning and Organization

```r
library(readxl)
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
```

```r
library(tidyr)

# Load Data
spending <- read_excel("NCAA Public Data.xlsx", sheet = "Sheet1")
```

```
## New names:
## * '' -> '...9'
## * '' -> '...10'
```

```r
private_perf <- read_excel("NCAA Public Data.xlsx", sheet = "Sheet3")
```

```
## New names:
## * 'School' -> 'School...1'
## * '2017-18' -> '2017-18...2'
## * '2018-19' -> '2018-19...3'
## * '2019-20' -> '2019-20...4'
## * '2020-21' -> '2020-21...5'
## * '2021-22' -> '2021-22...6'
## * '2022-23' -> '2022-23...7'
## * '2023-24' -> '2023-24...8'
## * '' -> '...9'
## * 'School' -> 'School...10'
## * '2017-18' -> '2017-18...11'
## * '2018-19' -> '2018-19...12'
## * '2019-20' -> '2019-20...13'
```

```
## * '2020-21' -> '2020-21...14'
## * '2021-22' -> '2021-22...15'
## * '2022-23' -> '2022-23...16'
## * '2023-24' -> '2023-24...17'
```

```
win_loss <- read.csv("win_loss_percentages.csv")
srs <- read.csv("srs_values.csv")

spending
```

```
## # A tibble: 113 x 11
##     School         '2017-2018' '2018-2019' '2019-2020' '2020-2021' '2021-2022'
##     <chr>          <chr>       <chr>       <chr>       <chr>       <chr>
##  1 Air Force       2765376.0   2673055.0   2936227.0   N/A         N/A
##  2 Akron           2922459.0   3054013.0   3087252.0   2169057.0   2607362.0
##  3 Alabama         1.0276122E7 1.5966875E7 9429629.0   8387815.0   1.0900723E7
##  4 App State       1701046.0   2008818.0   1978719.0   1644277.0   2030131.0
##  5 Arizona         1.0265281E7 1.0771475E7 1.0795208E7 1.3586274E7 1.2527047E7
##  6 Arizona State   8097708.0   8894051.0   8733671.0   6993562.0   9501282.0
##  7 Arkansas        9195260.0   1.0440326E7 9883017.0   8201217.0   1.180217E7
##  8 Arkansas State  2375329.0   2336742.0   2166758.0   N/A         2195011.0
##  9 Army            N/A         N/A         N/A         848054.0    N/A
## 10 Auburn          9446381.0   1.0536654E7 1.177581E7  9917358.0   1.3813576E7
## # i 103 more rows
## # i 5 more variables: '2022-2023' <chr>, '2023-2024' <chr>, ...9 <lgl>,
## #   ...10 <lgl>, 'Potential Variables' <chr>
```

```
private_perf
```

```
## # A tibble: 3 x 17
##   School...1 '2017-18...2' '2018-19...3' '2019-20...4' '2020-21...5'
##   <chr>              <dbl>         <dbl>         <dbl>         <dbl>
## 1 Duke                24.4          26.9          22.6          13.0
## 2 Gonzaga             18.4          27.8          22.7          27.2
## 3 Villanova           26.6          14.3          16.8          17.5
## # i 12 more variables: '2021-22...6' <dbl>, '2022-23...7' <dbl>,
## #   '2023-24...8' <dbl>, ...9 <lgl>, School...10 <chr>, '2017-18...11' <dbl>,
## #   '2018-19...12' <dbl>, '2019-20...13' <dbl>, '2020-21...14' <dbl>,
## #   '2021-22...15' <dbl>, '2022-23...16' <dbl>, '2023-24...17' <dbl>
```

```
win_loss
```

```
##                   Name X2017.2018 X2018.2019 X2019.2020 X2020.2021 X2021.2022
## 1            Air_Force      0.387      0.438      0.375      0.200      0.379
## 2                Akron      0.438      0.515      0.774      0.652      0.706
## 3              Alabama      0.556      0.529      0.516      0.788      0.576
## 4            App_State      0.455      0.344      0.545      0.586      0.559
## 5              Arizona      0.771      0.531      0.656      0.654      0.892
## 6        Arizona_State      0.625      0.676      0.645      0.440      0.452
## 7             Arkansas      0.657      0.529      0.625      0.781      0.757
## 8       Arkansas_State      0.344      0.406      0.500      0.458      0.621
## 9                 Army      0.433      0.406      0.500      0.545      0.484
```

```
## 10            Auburn  0.765  0.750  0.806  0.481  0.824
## 11         Ball_State  0.594  0.485  0.581  0.435  0.452
## 12        Boise_State  0.719  0.394  0.625  0.679  0.771
## 13      Bowling_Green  0.500  0.647  0.677  0.538  0.419
## 14            Buffalo  0.750  0.889  0.625  0.640  0.633
## 15         California  0.250  0.258  0.438  0.310  0.375
## 16   Central_Michigan  0.583  0.657  0.438  0.304  0.233
## 17          Charlotte  0.207  0.276  0.552  0.360  0.548
## 18         Cincinnati  0.861  0.800  0.667  0.522  0.545
## 19            Clemson  0.714  0.588  0.516  0.667  0.515
## 20            Coastal  0.438  0.500  0.485  0.692  0.576
## 21           Colorado  0.531  0.639  0.656  0.719  0.636
## 22      Colorado_State  0.344  0.375  0.625  0.714  0.806
## 23        Connecticut  0.438  0.485  0.613  0.652  0.697
## 24       East_Carolina  0.333  0.323  0.355  0.421  0.500
## 25    Eastern_Michigan  0.629  0.469  0.500  0.333  0.323
## 26                FAU  0.387  0.515  0.531  0.565  0.559
## 27                FIU  0.438  0.588  0.594  0.346  0.469
## 28            Florida  0.618  0.556  0.613  0.600  0.588
## 29      Florida_State  0.657  0.784  0.839  0.720  0.548
## 30       Fresno_State  0.656  0.719  0.367  0.500  0.639
## 31            Georgia  0.545  0.344  0.500  0.538  0.188
## 32   Georgia_Southern  0.636  0.636  0.606  0.500  0.448
## 33      Georgia_State  0.686  0.706  0.594  0.727  0.621
## 34       Georgia_Tech  0.406  0.438  0.548  0.654  0.375
## 35             Hawaii  0.567  0.581  0.567  0.524  0.607
## 36            Houston  0.771  0.892  0.742  0.875  0.842
## 37           Illinois  0.438  0.364  0.677  0.774  0.697
## 38            Indiana  0.516  0.543  0.625  0.444  0.600
## 39               Iowa  0.424  0.657  0.645  0.710  0.722
## 40         Iowa_State  0.419  0.657  0.375  0.083  0.629
## 41      James_Madison  0.313  0.424  0.300  0.650  0.517
## 42             Kansas  0.795  0.722  0.903  0.700  0.850
## 43       Kansas_State  0.676  0.735  0.344  0.310  0.452
## 44         Kent_State  0.500  0.667  0.625  0.652  0.676
## 45           Kentucky  0.703  0.811  0.806  0.360  0.765
## 46          Louisiana  0.794  0.594  0.424  0.654  0.516
## 47    Louisiana_Monroe  0.500  0.543  0.310  0.269  0.419
## 48     Louisiana_Tech  0.515  0.606  0.733  0.750  0.706
## 49         Louisville  0.611  0.588  0.774  0.650  0.406
## 50           Marshall  0.694  0.622  0.531  0.682  0.364
## 51           Maryland  0.594  0.676  0.774  0.548  0.469
## 52      Massachusetts  0.394  0.344  0.452  0.533  0.469
## 53            Memphis  0.618  0.611  0.677  0.714  0.667
## 54         Miami_Ohio  0.471  0.469  0.406  0.522  0.438
## 55           Michigan  0.805  0.811  0.613  0.821  0.559
## 56      Michigan_State  0.857  0.821  0.710  0.536  0.639
## 57   Middle_Tennessee  0.758  0.344  0.258  0.217  0.703
## 58          Minnesota  0.469  0.611  0.484  0.483  0.433
## 59   Mississippi_State  0.676  0.676  0.645  0.545  0.529
## 60           Missouri  0.606  0.469  0.484  0.615  0.364
## 61            NC_State  0.636  0.667  0.625  0.560  0.344
## 62               Navy  0.625  0.387  0.467  0.833  0.656
## 63           Nebraska  0.667  0.528  0.219  0.259  0.313
```

```
## 64               Nevada  0.784  0.853  0.613  0.615  0.419
## 65           New_Mexico  0.559  0.438  0.576  0.273  0.406
## 66     New_Mexico_State  0.824  0.857  0.806  0.600  0.794
## 67       North_Carolina  0.703  0.806  0.424  0.621  0.744
## 68          North_Texas  0.526  0.636  0.645  0.643  0.781
## 69     Northern_Illinois  0.406  0.500  0.581  0.158  0.300
## 70                 Ohio  0.452  0.452  0.531  0.680  0.714
## 71           Ohio_State  0.735  0.571  0.677  0.677  0.625
## 72             Oklahoma  0.563  0.588  0.613  0.593  0.543
## 73       Oklahoma_State  0.583  0.375  0.563  0.700  0.500
## 74         Old_Dominion  0.781  0.743  0.406  0.652  0.406
## 75             Ole_Miss  0.375  0.606  0.469  0.571  0.406
## 76               Oregon  0.639  0.658  0.774  0.750  0.571
## 77         Oregon_State  0.500  0.581  0.581  0.606  0.097
## 78           Penn_State  0.667  0.438  0.677  0.440  0.452
## 79           Pittsburgh  0.250  0.424  0.485  0.455  0.344
## 80               Purdue  0.811  0.722  0.516  0.643  0.784
## 81              Rutgers  0.441  0.452  0.645  0.571  0.563
## 82       San_Diego_State  0.667  0.618  0.938  0.821  0.719
## 83        San_Jose_State  0.133  0.129  0.226  0.238  0.258
## 84        South_Alabama  0.438  0.500  0.645  0.607  0.636
## 85       South_Carolina  0.515  0.500  0.581  0.286  0.581
## 86         South_Florida  0.313  0.632  0.452  0.409  0.258
## 87 Southern_Mississippi  0.471  0.606  0.290  0.320  0.212
## 88               Temple  0.515  0.697  0.452  0.313  0.586
## 89            Tennessee  0.743  0.838  0.548  0.667  0.771
## 90                Texas  0.559  0.568  0.613  0.704  0.647
## 91             Texas_A&M  0.629  0.438  0.533  0.444  0.675
## 92          Texas_State  0.455  0.706  0.656  0.720  0.724
## 93           Texas_Tech  0.730  0.816  0.581  0.621  0.730
## 94               Toledo  0.676  0.758  0.531  0.700  0.765
## 95                 Troy  0.485  0.400  0.290  0.393  0.625
## 96                  UAB  0.606  0.571  0.594  0.759  0.771
## 97                  UCF  0.594  0.727  0.533  0.478  0.600
## 98                 UCLA  0.636  0.515  0.613  0.688  0.771
## 99                 UNLV  0.606  0.548  0.531  0.444  0.563
## 100                 UTEP  0.355  0.276  0.531  0.500  0.588
## 101                 UTSA  0.571  0.531  0.406  0.577  0.313
## 102                 Utah  0.657  0.548  0.516  0.480  0.355
## 103           Utah_State  0.500  0.800  0.765  0.690  0.529
## 104             Virginia  0.912  0.921  0.767  0.720  0.600
## 105         Virginia_Tech  0.636  0.743  0.500  0.682  0.639
## 106           Washington  0.618  0.750  0.469  0.192  0.531
## 107     Washington_State  0.387  0.344  0.500  0.519  0.595
## 108         West_Virginia  0.703  0.417  0.677  0.655  0.485
## 109     Western_Kentucky  0.711  0.588  0.667  0.724  0.594
## 110     Western_Michigan  0.531  0.250  0.406  0.238  0.258
## 111            Wisconsin  0.455  0.676  0.677  0.581  0.758
## 112              Wyoming  0.606  0.250  0.273  0.560  0.735
##     X2022.2023 X2023.2024   average
## 1         0.438      0.290 0.3581429
## 2         0.667      0.686 0.6340000
## 3         0.838      0.676 0.6398571
## 4         0.500      0.794 0.5404286
```

4

```
## 5           0.800        0.750 0.7220000
## 6           0.639        0.438 0.5592857
## 7           0.611        0.485 0.6350000
## 8           0.394        0.541 0.4662857
## 9           0.515        0.313 0.4565714
## 10          0.618        0.771 0.7164286
## 11          0.625        0.484 0.5222857
## 12          0.706        0.667 0.6515714
## 13          0.355        0.588 0.5320000
## 14          0.469        0.129 0.5907143
## 15          0.094        0.406 0.3044286
## 16          0.323        0.563 0.4430000
## 17          0.611        0.613 0.4524286
## 18          0.639        0.595 0.6612857
## 19          0.676        0.667 0.6204286
## 20          0.355        0.290 0.4765714
## 21          0.514        0.703 0.6282857
## 22          0.455        0.694 0.5732857
## 23          0.795        0.925 0.6578571
## 24          0.485        0.455 0.4102857
## 25          0.258        0.419 0.4187143
## 26          0.897        0.735 0.5984286
## 27          0.438        0.333 0.4580000
## 28          0.485        0.667 0.5895714
## 29          0.281        0.515 0.6205714
## 30          0.355        0.364 0.5142857
## 31          0.500        0.541 0.4508571
## 32          0.515        0.273 0.5162857
## 33          0.323        0.452 0.5870000
## 34          0.455        0.438 0.4734286
## 35          0.667        0.588 0.5858571
## 36          0.892        0.865 0.8398571
## 37          0.606        0.763 0.6170000
## 38          0.657        0.576 0.5658571
## 39          0.576        0.559 0.6132857
## 40          0.576        0.784 0.5032857
## 41          0.667        0.889 0.5371429
## 42          0.778        0.676 0.7748571
## 43          0.722        0.559 0.5425714
## 44          0.800        0.500 0.6314286
## 45          0.647        0.697 0.6841429
## 46          0.765        0.576 0.6175714
## 47          0.344        0.367 0.3931429
## 48          0.455        0.688 0.6361429
## 49          0.125        0.250 0.4862857
## 50          0.750        0.394 0.5767143
## 51          0.629        0.485 0.5964286
## 52          0.484        0.645 0.4744286
## 53          0.743        0.688 0.6740000
## 54          0.375        0.469 0.4500000
## 55          0.529        0.250 0.6268571
## 56          0.618        0.571 0.6788571
## 57          0.576        0.424 0.4685714
## 58          0.290        0.559 0.4755714
```

```
## 59         0.618      0.600 0.6127143
## 60         0.714      0.250 0.5002857
## 61         0.676      0.634 0.5917143
## 62         0.581      0.419 0.5668571
## 63         0.500      0.676 0.4517143
## 64         0.667      0.765 0.6737143
## 65         0.647      0.722 0.5172857
## 66         0.375      0.406 0.6660000
## 67         0.606      0.784 0.6697143
## 68         0.816      0.559 0.6580000
## 69         0.406      0.355 0.3865714
## 70         0.576      0.606 0.5730000
## 71         0.457      0.611 0.6218571
## 72         0.469      0.625 0.5705714
## 73         0.556      0.375 0.5217143
## 74         0.613      0.219 0.5457143
## 75         0.364      0.625 0.4880000
## 76         0.583      0.667 0.6631429
## 77         0.344      0.406 0.4450000
## 78         0.622      0.485 0.5401429
## 79         0.667      0.667 0.4702857
## 80         0.829      0.872 0.7395714
## 81         0.559      0.469 0.5285714
## 82         0.821      0.703 0.7552857
## 83         0.600      0.281 0.2664286
## 84         0.543      0.500 0.5527143
## 85         0.344      0.765 0.5102857
## 86         0.438      0.758 0.4657143
## 87         0.758      0.500 0.4510000
## 88         0.500      0.444 0.5010000
## 89         0.694      0.750 0.7158571
## 90         0.763      0.618 0.6388571
## 91         0.714      0.583 0.5737143
## 92         0.457      0.486 0.6005714
## 93         0.500      0.676 0.6648571
## 94         0.771      0.625 0.6894286
## 95         0.606      0.625 0.4891429
## 96         0.744      0.657 0.6717143
## 97         0.559      0.515 0.5722857
## 98         0.838      0.485 0.6494286
## 99         0.594      0.618 0.5577143
## 100        0.438      0.529 0.4595714
## 101        0.313      0.344 0.4364286
## 102        0.531      0.595 0.5260000
## 103        0.743      0.800 0.6895714
## 104        0.758      0.676 0.7648571
## 105        0.559      0.559 0.6168571
## 106        0.500      0.531 0.5130000
## 107        0.500      0.714 0.5084286
## 108        0.559      0.281 0.5395714
## 109        0.515      0.647 0.6351429
## 110        0.258      0.375 0.3308571
## 111        0.571      0.611 0.6184286
## 112        0.290      0.469 0.4547143
```

```
srs
```

```
##                     Name X2017.2018 X2018.2019 X2019.2020 X2020.2021 X2021.2022
## 1             Air_Force      -4.31      -4.28      -0.37     -12.98      -5.67
## 2                 Akron      -6.82       4.86       7.15       1.85       0.91
## 3               Alabama      12.34       9.45      11.12      19.58      14.62
## 4             App_State      -4.83      -3.73      -0.81      -5.84      -1.45
## 5               Arizona      15.67       8.32      19.49      15.02      22.75
## 6         Arizona_State      14.37      10.28      11.46       7.48       7.64
## 7              Arkansas      14.76      11.75      13.46      17.80      16.27
## 8        Arkansas_State      -9.65      -7.10      -2.92      -7.97      -3.27
## 9                  Army      -8.62      -7.57      -9.10       0.06     -10.53
## 10               Auburn      15.97      20.84      15.04      10.02      19.20
## 11           Ball_State      -1.22       3.39       4.12      -1.96      -7.99
## 12          Boise_State      11.21       3.61       8.59       9.75      11.93
## 13        Bowling_Green      -6.38       4.24       0.47      -0.17      -8.82
## 14              Buffalo       7.99      15.56       1.80       6.50       1.90
## 15           California      -3.40      -3.16       3.20       4.89       3.96
## 16     Central_Michigan      -1.39       2.79      -3.40     -12.34     -12.56
## 17            Charlotte     -10.13      -8.34       0.46      -5.68      -1.62
## 18           Cincinnati      20.25      14.53      12.33       3.99       7.44
## 19              Clemson      17.16      13.85      10.26      12.49       9.59
## 20              Coastal      -3.94      -0.50      -0.83      -2.74       0.66
## 21             Colorado       6.06       9.68      14.91      18.88       9.82
## 22       Colorado_State      -3.86      -0.11       6.70       8.06      11.51
## 23          Connecticut       1.08       6.81      11.15      15.77      16.40
## 24         East_Carolina      -8.72      -5.49      -2.27       1.00       0.49
## 25      Eastern_Michigan      -0.23       0.40      -3.12     -14.23     -10.86
## 26                  FAU      -5.35      -1.42      -1.66      -2.05       2.63
## 27                  FIU      -4.30      -4.55       0.90     -11.20      -6.47
## 28              Florida      16.66      15.42      14.94      13.70      11.93
## 29        Florida_State      15.78      17.99      17.28      17.12       6.39
## 30         Fresno_State       8.17       8.99       2.08      -3.57       8.36
## 31              Georgia      10.24       5.31       8.28       6.62      -0.63
## 32     Georgia_Southern      -0.54       3.75       1.69     -10.22      -7.04
## 33        Georgia_State       3.22       2.44       3.09      -1.70      -1.11
## 34         Georgia_Tech       5.25       6.93      10.49      12.79       2.24
## 35               Hawaii      -2.33      -1.30      -1.27      -2.98      -0.52
## 36              Houston      15.99      18.91      16.46      21.66      22.55
## 37             Illinois       8.50       8.95      15.27      23.25      16.69
## 38              Indiana       9.89      13.82      15.07      12.79      12.71
## 39                 Iowa       8.13      14.27      16.07      21.40      18.90
## 40           Iowa_State       6.88      18.07       9.10      -0.72      12.38
## 41        James_Madison      -5.47      -8.52      -9.23      -0.11      -4.13
## 42               Kansas      21.39      18.35      24.96      15.29      22.28
## 43         Kansas_State      12.85      15.39       8.70       0.76      10.61
## 44           Kent_State      -3.73       1.26       3.68       1.74      -0.13
## 45             Kentucky      17.24      21.43      15.78      11.77      20.90
## 46            Louisiana       7.58      -2.07      -4.60      -4.84      -2.71
## 47      Louisiana_Monroe      -6.28       0.81      -6.50     -10.59      -8.70
## 48       Louisiana_Tech       1.33       0.95       7.06       6.63       5.23
## 49           Louisville      14.30      17.28      17.94      10.35       4.23
## 50             Marshall       4.21      -0.64       1.70       5.70      -4.57
## 51             Maryland      13.78      16.01      17.82      13.94       9.34
```

7

```
## 52           Massachusetts  -1.65   -3.02    1.12    4.77   -0.67
## 53                 Memphis   1.81   10.70   12.16   13.81   14.67
## 54              Miami_Ohio  -3.56    0.77   -4.17   -2.67   -6.35
## 55                Michigan  19.03   21.82   17.94   22.00   14.26
## 56          Michigan_State  22.41   24.93   21.92    9.80   13.18
## 57        Middle_Tennessee  10.29   -6.09   -8.41  -10.78    4.50
## 58               Minnesota   7.00   12.52   16.39   11.04    5.93
## 59       Mississippi_State  11.71   15.96   12.21    9.24   11.83
## 60                Missouri  13.81    8.60    8.51   12.37    2.94
## 61                NC_State  12.91   14.94   12.05    9.60    3.77
## 62                    Navy  -6.91  -10.09   -7.36    1.02   -2.94
## 63                Nebraska  10.99   14.85    2.41    6.08    3.00
## 64                  Nevada  14.44   16.00    8.91    5.40    3.71
## 65              New_Mexico   5.81   -0.55    3.47  -12.38    1.25
## 66        New_Mexico_State   9.06   10.05    4.11   -1.21    7.11
## 67          North_Carolina  20.07   23.94    9.45   14.79   15.50
## 68             North_Texas   1.16   -0.25    7.32    7.58    8.98
## 69        Northern_Illinois  -6.71    2.58   -1.33  -17.90  -10.50
## 70                    Ohio  -2.52   -1.84    0.74    5.48    1.50
## 71              Ohio_State  17.67   13.89   19.37   18.25   14.17
## 72                Oklahoma  13.81   15.30   13.48   13.52   14.12
## 73           Oklahoma_State  12.45    8.09   12.01   13.65   13.28
## 74            Old_Dominion   9.27    3.74    0.02   -1.10   -1.32
## 75                Ole_Miss   6.85   12.32    7.37   11.18    6.72
## 76                  Oregon  10.70   13.95   18.22   16.15   10.05
## 77            Oregon_State   7.31    7.84    9.70   11.48   -1.44
## 78              Penn_State  15.85   12.55   16.12   13.53    8.72
## 79              Pittsburgh  -1.87    7.88    6.80    7.80    0.45
## 80                  Purdue  23.41   21.40   16.30   15.55   19.15
## 81                 Rutgers   5.50    8.89   15.72   14.14    8.78
## 82          San_Diego_State  12.23    5.29   18.38   14.36   13.13
## 83           San_Jose_State  -8.74  -16.37   -7.55  -17.50   -7.01
## 84           South_Alabama  -5.44   -4.46    0.08   -6.03    0.23
## 85          South_Carolina   9.45    8.34   10.27    3.39    7.14
## 86            South_Florida  -7.25    5.96    5.15   -0.02   -3.05
## 87    Southern_Mississippi  -4.26    2.58   -6.65   -6.88  -13.58
## 88                  Temple   7.25    8.17    6.59    2.57    4.61
## 89               Tennessee  18.20   21.55   11.18   16.87   20.46
## 90                   Texas  13.96   16.06   10.47   15.51   17.01
## 91                Texas_A&M  15.30    7.21    4.93    2.20   12.93
## 92              Texas_State  -5.84    1.54    5.42   -3.85    0.14
## 93               Texas_Tech  19.38   22.79   16.83   16.59   20.08
## 94                  Toledo   2.65    8.21    1.20    6.59    6.26
## 95                    Troy  -1.50   -6.12   -9.01  -12.06   -3.56
## 96                     UAB   4.90    0.36   -0.94    5.80   10.90
## 97                     UCF   4.76   13.37    5.13    5.94    6.23
## 98                    UCLA  12.07    7.17    9.83   16.30   19.30
## 99                    UNLV   7.20    1.58    6.79   -1.48    7.20
## 100                   UTEP  -5.31   -8.23   -1.00    1.07    0.27
## 101                   UTSA  -1.24    0.69   -3.18   -1.22  -10.46
## 102                   Utah  11.39    6.13    8.47   11.86    5.45
## 103             Utah_State   2.31   11.98   13.66   10.99   10.41
## 104               Virginia  22.20   25.46   12.16   15.27    8.88
## 105           Virginia_Tech  14.35   19.28    8.22   12.59   14.48
```

```
## 106        Washington        7.62    12.01    12.41     1.88     5.03
## 107   Washington_State        0.67    -1.67     5.13     8.12    12.57
## 108      West_Virginia       19.89     6.94    18.51    15.36    10.23
## 109   Western_Kentucky       11.08     3.14     4.20     5.19     3.30
## 110   Western_Michigan       -2.22    -6.26    -5.42   -11.51   -13.51
## 111          Wisconsin       10.38    17.90    15.95    17.22    13.13
## 112            Wyoming        4.26    -9.75    -3.05    -1.23     9.36
##    X2022.2023 X2023.2024    average
## 1        2.00      -4.77 -4.3400000
## 2        4.19       2.77  2.1300000
## 3       23.19      20.69 15.8557143
## 4       -1.50       6.04 -1.7314286
## 5       19.08      24.54 17.8385714
## 6       11.29       4.58  9.5857143
## 7       15.99       6.12 13.7357143
## 8       -8.07       0.28 -5.5285714
## 9       -7.13     -14.67 -8.2228571
## 10      14.35      22.46 16.8400000
## 11      -1.20      -6.68 -1.6485714
## 12      12.78      13.31 10.1685714
## 13      -8.46      -4.93 -3.4357143
## 14      -2.97     -15.12  2.2371429
## 15      -3.36       4.94  1.0100000
## 16     -11.97      -8.07 -6.7057143
## 17       3.90       4.09 -2.4742857
## 18      11.77      15.06 12.1957143
## 19      10.53      16.20 12.8685714
## 20      -8.93     -10.47 -3.8214286
## 21      11.17      15.95 12.3528571
## 22       5.64      13.40  5.9057143
## 23      22.95      26.70 14.4085714
## 24      -0.29      -0.62 -2.2714286
## 25     -11.05     -12.76 -7.4071429
## 26      13.92      13.26  2.7614286
## 27      -2.71      -8.01 -5.1914286
## 28      10.90      15.17 14.1028571
## 29      -0.38       9.52 11.9571429
## 30       2.60      -3.29  3.3342857
## 31       1.80       8.63  5.7500000
## 32      -3.25      -8.44 -3.4357143
## 33      -6.79      -4.59 -0.7771429
## 34       2.08       4.60  6.3400000
## 35       2.36      -0.07 -0.8728571
## 36      22.20      25.69 20.4942857
## 37      14.81      19.56 15.2900000
## 38      14.94       7.74 12.4228571
## 39      14.08      13.00 15.1214286
## 40      15.36      22.30 11.9100000
## 41       4.96       8.19 -2.0442857
## 42      19.20      16.35 19.6885714
## 43      15.79      11.06 10.7371429
## 44       7.53      -0.76  1.3700000
## 45      14.76      16.79 16.9528571
## 46       4.27      -0.60 -0.4242857
```

```
## 47         -8.64     -10.05 -7.1357143
## 48          0.10       4.83  3.7328571
## 49         -5.39       0.45  8.4514286
## 50          7.13      -5.42  1.1585714
## 51         14.58      10.60 13.7242857
## 52         -1.34       7.59  0.9714286
## 53         14.50       9.00 10.9500000
## 54         -7.46      -6.45 -4.2700000
## 55         12.43       5.28 16.1085714
## 56         14.25      17.72 17.7442857
## 57          2.95      -7.45 -2.1414286
## 58         -0.30       9.51  8.8700000
## 59         11.43      14.19 12.3671429
## 60         11.41       3.27  8.7014286
## 61         12.50      12.81 11.2257143
## 62         -2.79     -11.70 -5.8242857
## 63          7.12      14.36  8.4014286
## 64         10.01      12.85 10.1885714
## 65         10.82      15.04  3.3514286
## 66         -1.56      -8.62  2.7057143
## 67         12.94      21.36 16.8642857
## 68         10.71       8.41  6.2728571
## 69         -6.55      -9.16 -7.0814286
## 70          1.07       0.48  0.7014286
## 71         12.99      13.64 15.7114286
## 72         11.38      14.66 13.7528571
## 73         12.62       6.25 11.1928571
## 74         -0.41      -8.35  0.2642857
## 75          5.73       8.30  8.3528571
## 76         13.83      12.22 13.5885714
## 77          0.08       3.28  5.4642857
## 78         13.04       9.22 12.7185714
## 79         10.59      14.89  6.6485714
## 80         18.23      24.93 19.8528571
## 81         14.58       7.41 10.7171429
## 82         15.81      14.68 13.4114286
## 83          5.17      -3.82 -7.9742857
## 84          3.58      -5.92 -2.5657143
## 85         -2.28      10.98  6.7557143
## 86          1.95       7.37  1.4442857
## 87          3.31      -5.93 -4.4871429
## 88          4.68      -1.26  4.6585714
## 89         20.84      21.81 18.7014286
## 90         20.56      15.91 15.6400000
## 91         13.76      13.35  9.9542857
## 92         -2.32      -3.71 -1.2314286
## 93         12.58      14.96 17.6014286
## 94          5.76       1.78  4.6357143
## 95          1.39       1.24 -4.2314286
## 96         10.96       5.14  5.3028571
## 97         11.09      11.92  8.3485714
## 98         22.11       7.50 13.4685714
## 99          7.20       8.38  5.2671429
## 100        -0.74      -1.49 -2.2042857
```

```
## 101      -7.74       -5.26 -4.0585714
## 102      11.34       15.00  9.9485714
## 103      14.23       11.56 10.7342857
## 104      13.28       10.45 15.3857143
## 105       9.36       11.91 12.8842857
## 106       6.78       11.85  8.2257143
## 107      10.53       14.02  7.0528571
## 108      15.95        3.73 12.9442857
## 109      -0.81        1.19  3.8985714
## 110     -10.55       -9.30 -8.3957143
## 111      10.98       16.01 14.5100000
## 112       2.02        0.54  0.3071429
```

```r
# Reshape spending data to long format
spending_long <- spending %>%
  select(School, `2017-2018`:`2023-2024`) %>%
  pivot_longer(cols = -School, names_to = "Year", values_to = "Spending")




# Reshape win/loss data
colnames(win_loss) <- gsub("^X", "", colnames(win_loss))
colnames(win_loss) <- gsub("\\.", "-", colnames(win_loss))
win_long <- win_loss %>%
  pivot_longer(cols = starts_with("20"), names_to = "Year", values_to = "Win_Percentage") %>%
  rename(School = Name)
win_long$School <- gsub("_", " ", win_long$School)

# Reshape SRS data
colnames(srs) <- gsub("^X", "", colnames(srs))
colnames(srs) <- gsub("\\.", "-", colnames(srs))
srs_long <- srs %>%
  pivot_longer(cols = starts_with("20"), names_to = "Year", values_to = "SRS") %>%
  rename(School = Name)
srs_long$School <- gsub("_", " ", srs_long$School)

# Clean the school names
spending_long$School <- recode(spending_long$School,
  "Coastal California" = "Coastal",
  "Conecticut" = "Connecticut",
  "Lousiana Tech" = "Louisiana Tech",
  "Lousiana, Lafayette" = "Louisiana",
  "Lousiana, Monroe" = "Louisiana Monroe",
  "Miami (Ohio)" = "Miami Ohio",
  "Old Domionion" = "Old Dominion"
)

spending_long <- spending_long %>% filter(School != "LSU")

spending_long$School <- as.character(spending_long$School)
spending_long$Year <- as.character(spending_long$Year)

win_long$School <- as.character(win_long$School)
win_long$Year <- as.character(win_long$Year)
```

```r
srs_long$School <- as.character(srs_long$School)
srs_long$Year <- as.character(srs_long$Year)


clean_join_keys <- function(df) {
  df$School <- df$School %>%
    gsub("_", " ", .) %>%
    gsub("\\s+", " ", .) %>%
    trimws()

  df$Year <- trimws(df$Year)
  return(df)
}

spending_long <- clean_join_keys(spending_long)
win_long <- clean_join_keys(win_long)
srs_long <- clean_join_keys(srs_long)

# Find shared School-Year combinations across all three datasets
valid_keys <- inner_join(win_long, srs_long, by = c("School", "Year")) %>%
  inner_join(spending_long, by = c("School", "Year")) %>%
  select(School, Year)

# Filter each dataset to only those rows
spending_long <- inner_join(spending_long, valid_keys, by = c("School", "Year"))
win_long <- inner_join(win_long, valid_keys, by = c("School", "Year"))
srs_long <- inner_join(srs_long, valid_keys, by = c("School", "Year"))


colnames(spending)
```

```
##  [1] "School"              "2017-2018"          "2018-2019"
##  [4] "2019-2020"           "2020-2021"          "2021-2022"
##  [7] "2022-2023"           "2023-2024"          "...9"
## [10] "...10"               "Potential Variables"
```

```r
# Merge all three
public_data <- spending_long %>%
  inner_join(win_long, by = c("School", "Year")) %>%
  inner_join(srs_long, by = c("School", "Year")) %>%
  drop_na()

public_data
```

```
## # A tibble: 784 x 7
##    School    Year      Spending  average.x Win_Percentage average.y   SRS
##    <chr>     <chr>     <chr>         <dbl>          <dbl>     <dbl> <dbl>
##  1 Air Force 2017-2018 2765376.0     0.358          0.387     -4.34 -4.31
##  2 Air Force 2018-2019 2673055.0     0.358          0.438     -4.34 -4.28
##  3 Air Force 2019-2020 2936227.0     0.358          0.375     -4.34 -0.37
##  4 Air Force 2020-2021 N/A           0.358          0.2       -4.34 -13.0
##  5 Air Force 2021-2022 N/A           0.358          0.379     -4.34 -5.67
```

```
##  6 Air Force 2022-2023 2747369.0      0.358          0.438        -4.34   2
##  7 Air Force 2023-2024 N/A            0.358          0.29         -4.34  -4.77
##  8 Akron     2017-2018 2922459.0      0.634          0.438         2.13  -6.82
##  9 Akron     2018-2019 3054013.0      0.634          0.515         2.13   4.86
## 10 Akron     2019-2020 3087252.0      0.634          0.774         2.13   7.15
## # i 774 more rows
```

```
nrow(spending_long)
```

```
## [1] 784
```

```
nrow(win_long)
```

```
## [1] 784
```

```
nrow(srs_long)
```

```
## [1] 784
```

```
public_data$Spending <- gsub("[$,]", "", public_data$Spending)
public_data$Spending <- as.numeric(public_data$Spending)
```

```
## Warning: NAs introduced by coercion
```

Regression Model

```
model <- lm(Spending ~ Win_Percentage + SRS, data = public_data)
summary(model)
```

```
##
## Call:
## lm(formula = Spending ~ Win_Percentage + SRS, data = public_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8670633 -2033694  -392416  1325132 19396484
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     12195346     521755   23.37   <2e-16 ***
## Win_Percentage -15523660    1042489  -14.89   <2e-16 ***
## SRS               570167      18860   30.23   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3112000 on 714 degrees of freedom
##   (67 observations deleted due to missingness)
## Multiple R-squared:  0.5854, Adjusted R-squared:  0.5843
## F-statistic: 504.1 on 2 and 714 DF,  p-value: < 2.2e-16
```

```r
library(ggplot2)

# Add predictions and residuals
public_data <- public_data %>%
  mutate(
    Spending = as.numeric(Spending),
    Predicted_Spending = predict(model, newdata = public_data),
    Residual = Spending - Predicted_Spending
  )

# Average residual per school
residual_summary <- public_data %>%
  group_by(School) %>%
  summarize(Average_Residual = mean(Residual, na.rm = TRUE)) %>%
  arrange(desc(abs(Average_Residual)))

# Plot top 20 absolute residuals
top_residuals <- residual_summary %>%
  slice_max(order_by = abs(Average_Residual), n = 20)

ggplot(top_residuals, aes(x = reorder(School, Average_Residual), y = Average_Residual, fill = Average_R
  geom_col(show.legend = FALSE) +
  coord_flip() +
  labs(title = "Top 20 Schools by Spending Residual",
       x = "School",
       y = "Residual (Actual - Predicted Spending)") +
  scale_fill_manual(values = c("red", "green"))
```

## Top 20 Schools by Spending Residual
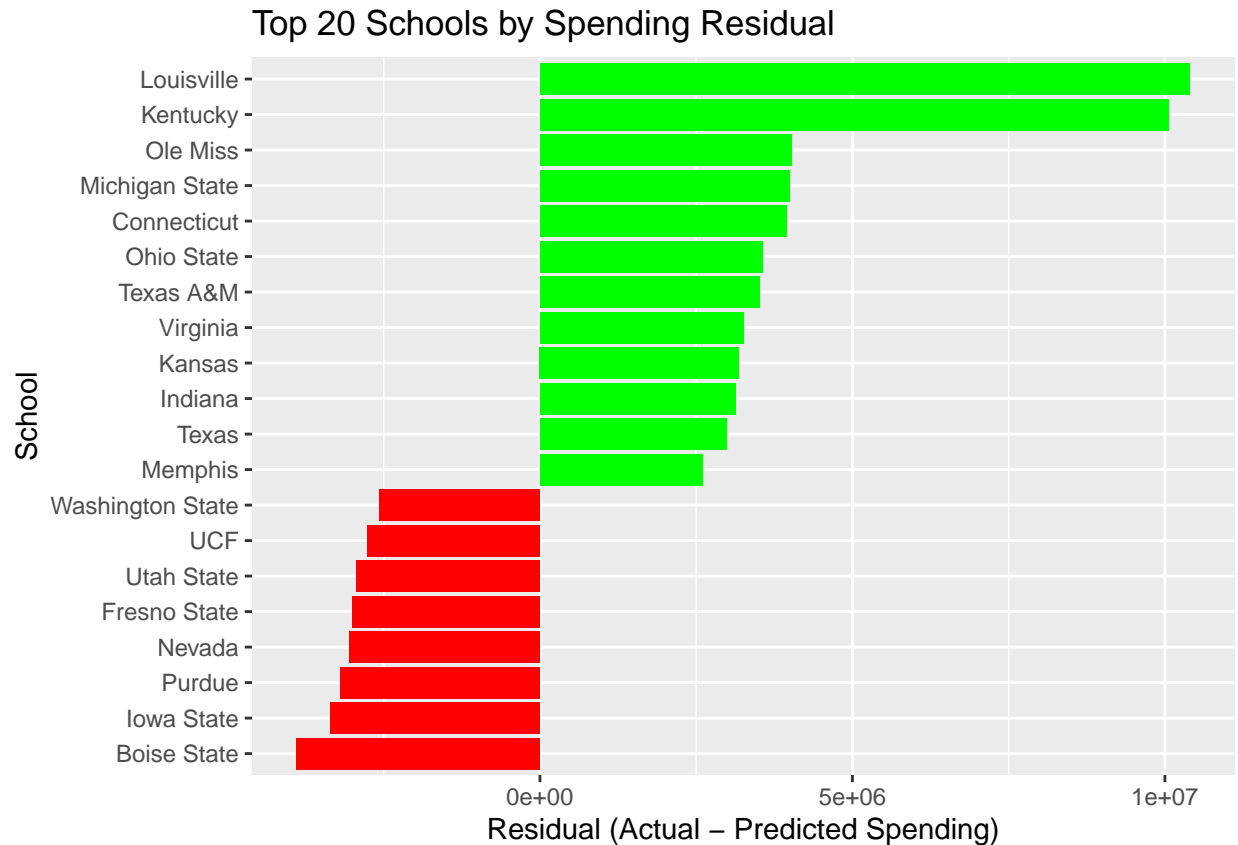


Main Takeaways: 58.5% of variation in spending is explained by Win% and SRS Big F-stat and low p-value means model is statistically significant Typical Error in prediction is about 3.1 million, median residual is -392416 Schools with negative residuals overperform their expectations Louisville and Kentucky overspend by 10 million

Private School Data

```r
# manually inputting private school data
private_data <- data.frame(
  School = rep(c("Duke", "Gonzaga", "Villanova"), each = 7),
  Year = rep(c("2017-2018", "2018-2019", "2019-2020", "2020-2021", "2021-2022", "2022-2023", "2023-2024
  Win_Percentage = c(
    0.85, 0.91, 0.78, 0.65, 0.73, 0.88, 0.84,    # Duke
    0.89, 0.94, 0.90, 0.95, 0.92, 0.91, 0.89,    # Gonzaga
    0.88, 0.76, 0.80, 0.82, 0.78, 0.84, 0.79     # Villanova
  ),
  SRS = c(
    20.1, 22.5, 18.3, 14.7, 17.2, 21.4, 20.7,    # Duke
    23.0, 24.8, 21.6, 25.1, 23.9, 22.4, 23.3,    # Gonzaga
    21.2, 17.5, 18.8, 20.1, 19.6, 21.0, 18.9     # Villanova
  )
)


# Prediction data set
private_data$Predicted_Spending <- predict(model, newdata = private_data)
```

Visualize Private Schools

```r
library(ggplot2)

# preds = data set with standard error
preds <- predict(model, newdata = private_data, se.fit = TRUE)

private_data$Predicted_Spending <- preds$fit
private_data$SE <- preds$se.fit
private_data$Lower_CI <- preds$fit - 1.96 * preds$se.fit
private_data$Upper_CI <- preds$fit + 1.96 * preds$se.fit

# All 3 no error bars
ggplot(private_data, aes(x = Year, y = Predicted_Spending, group = School, color = School)) +
  geom_line(size = 1.2) +
  geom_point(size = 2) +
  labs(
    title = "Predicted MBB Spending for Private Schools",
    x = "Year",
    y = "Predicted Spending ($)"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

# Predicted MBB Spending for Private Schools
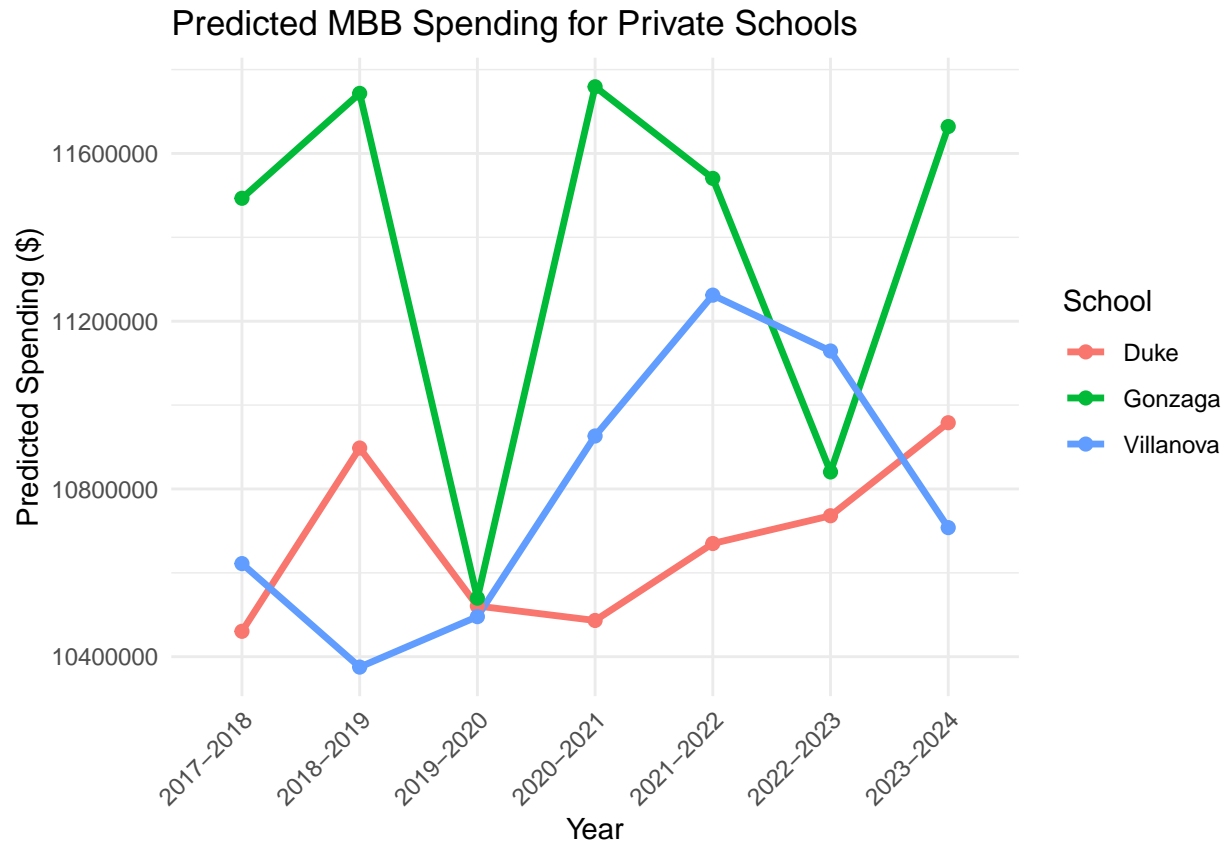


```r
# Duke plot
ggplot(private_data %>% filter(School == "Duke"),
       aes(x = Year, y = Predicted_Spending)) +
  geom_line(color = "navy", size = 1.2, group = 1) +
  geom_point(color = "navy", size = 2) +
  geom_errorbar(aes(ymin = Lower_CI, ymax = Upper_CI), width = 0.2, color = "navy") +
  labs(
    title = "Predicted MBB Spending for Duke",
    x = "Year",
    y = "Predicted Spending ($)"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

## Predicted MBB Spending for Duke
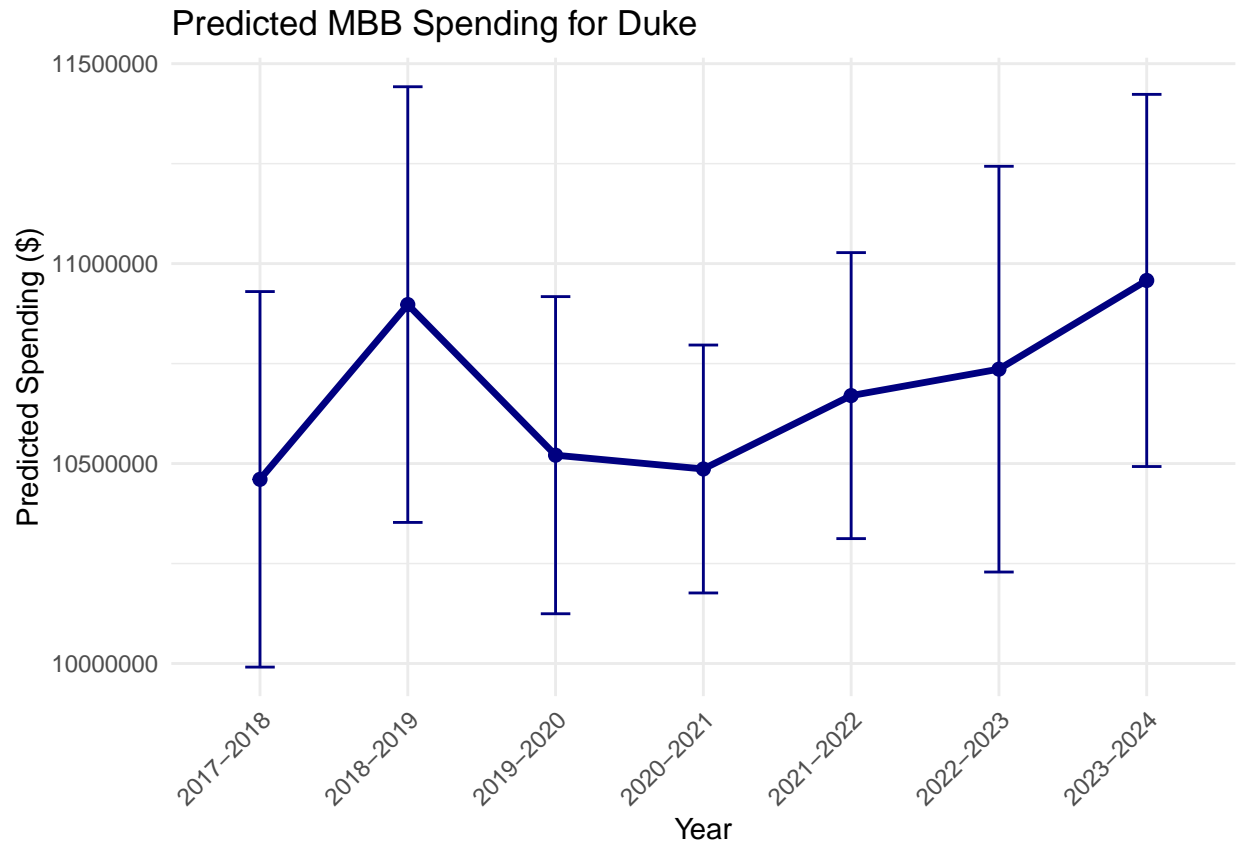


```
# Gonzaga Plot
ggplot(private_data %>% filter(School == "Gonzaga"),
       aes(x = Year, y = Predicted_Spending)) +
  geom_line(color = "darkgreen", size = 1.2, group = 1) +
  geom_point(color = "darkgreen", size = 2) +
  geom_errorbar(aes(ymin = Lower_CI, ymax = Upper_CI), width = 0.2, color = "darkgreen") +
  labs(
    title = "Predicted MBB Spending for Gonzaga",
    x = "Year",
    y = "Predicted Spending ($)"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

## Predicted MBB Spending for Gonzaga
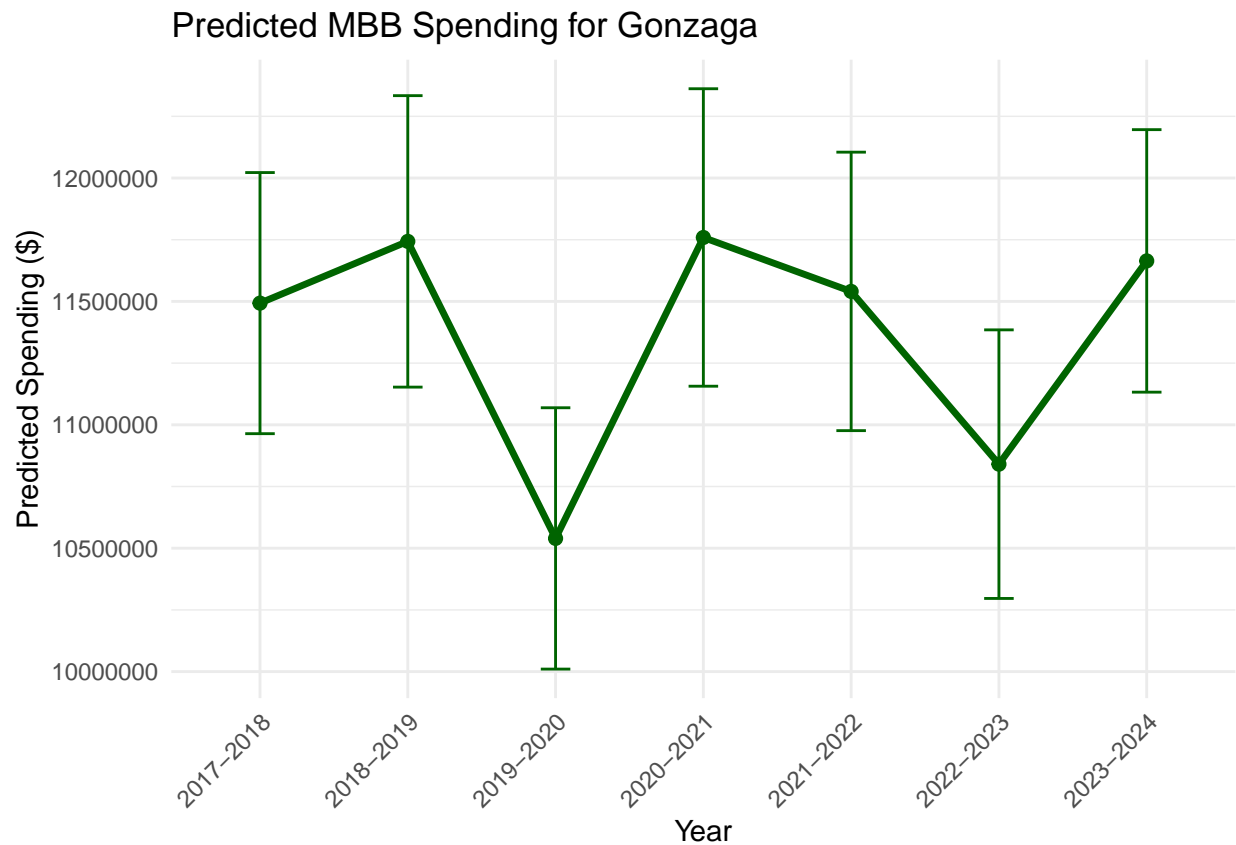


```r
# Villanova Plot
ggplot(private_data %>% filter(School == "Villanova"),
       aes(x = Year, y = Predicted_Spending)) +
  geom_line(color = "red", size = 1.2, group = 1) +
  geom_point(color = "red", size = 2) +
  geom_errorbar(aes(ymin = Lower_CI, ymax = Upper_CI), width = 0.2, color = "red") +
  labs(
    title = "Predicted MBB Spending for Villanova",
    x = "Year",
    y = "Predicted Spending ($)"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

Predicted MBB Spending for Villanova