

1. OBTAINING THE DATA

The data for this dataset was obtained from the following sources.

- 1 [Pro Football Reference](#)
- 2 [Airyards](#)

1. Pro Football Reference was used to get the data on NCAA wide receivers . CSVs for each year from 2010-2017 were downloaded for the NCAA wide receivers, then joined together later. Only the final year in college was kept for each player. Combine data from 2010-2017 was also obtained from Pro Football Reference using the same method.

2. Airyards was used to collect the NFL data. Year-by-year data for all NFL wide receivers from 2011-2018 was downloaded, then joined together. Only the rookie year of each player was kept.

2. CLEANING THE DATA

After obtaining the various data, we had the following files.

1. College Data
 - “cfb_wr_2010.csv” - wide receiver data from the 2010 NCAA season
 - “cfb_wr_2011.csv.” - wide receiver data from the 2011 NCAA season
 - ⋮
 - “cfb_wr_2017.csv.” - wide receiver data from the 2017 NCAA season
2. Combine Data
 - “combine_wr.csv” - combine results for every wide receiver from 2010-2018
3. NFL Data
 - “airyards_2011.csv” - WR data from 2011 NFL season
 - “airyards_2012.csv” - WR data from 2012 NFL season
 - ⋮
 - “airyards_2018.csv” - WR data from 2018 NFL season

The code for merging those files into the final dataset (“finalDataset.csv”) can be found in “data-Cleaner.R”.

3. FINAL DATASET

See the table below for information on the variables included in the final dataset.

	Variable	Data Type	Description
1	Player	String	First and last name of a given player
2	Conf	String	The conference that the wide receiver’s college team played in (categorical)
3	Rec	Numeric	Number of receptions made by the player in their final year of college
4	Yds	Numeric	total number of receiving yards for the player in their final year of college
5	YdsCatch	Numeric	The average yards per catch for the player in their final year of college
6	TD	Numeric	The number of touchdowns caught by the player in their final year of college
7	NflYds	Numeric	Total receiving yards for the player in their NFL rookie season
8	NflTD	Numeric	Touchdowns caught by the player in their NFL rookie season
9	Height	Numeric	Height of the player in inches
10	Wt	Numeric	Weight of the player in pounds
11	fourtyYD	Numeric	The forty-yard dash time of a given player
12	Vertical	Numeric	The standing vertical leap for a given player
13	WrScore	Numeric	Statistic defined by $\text{NflYds} + 60(\text{NflTD})$
14	LogScore	Numeric	$\text{LogScore} = \log(\text{WrScore})$