NFL Draft

December 16, 2019

1 Introduction

We wanted to see how the draft position of players in the NFL relates to their succesfulnes in the NFl. If you are drafted first are you really the best player drafted that year. The first thing that we had to do was collect stats about players in the NFL and draft information. Then we calculated the succesfulness of every player by year based on their stats. We then standardized the scores based on position, since the different positions have very different successes, the max kicker score is close to 1200 while the max QB score is 140. We then averaged all of the seasons of each player and added the offset for Pro-Bowl selections to create a standardized success for each player. The NFL is one of the largest entertainment industries in the United States earning over $8 Billion in revenue in the 2018 season. Teams are always trying to get the most out of draft picks and other resources so understanding the draft and the potential success of draft picks is extremely important to a teams success. Data Science is important to coaches and other executives of a team to maximize the impact that they have on the league. If you are unfamilier with the structure of the nfl draft and the positions of player here are some useful links to provide important information about the nfl.

How the NFL Draft Works
Official NFL Website
Information about the NFL
Game Information

[340]:

```
import pandas as pd
import sqlite3
import seaborn as sns
import sklearn
from sklearn import linear_model
import numpy as np
import matplotlib.pyplot as plt
import statsmodels.formula.api as sm
import warnings
warnings.filterwarnings('ignore')
```

2 Collecting Data

To get started we first had to get a list of every player drafted in the nfl, we chose to cut the lower bound for the year at 1970 since before then the data collected is inconsistent at best. To do this we scraped the draft information from the NFL official website. The NFL website seperates drafts by year so we were able to collect only the data that we wanted to work with in an organized manner. Once we had every player drafted we then had to collect the stats for each individual player, we were able to use a large database of stats collected from kaggle. The statistical data was stored in csv files which we then converted to sql tables for ease of use. The stats collected included individual statistics for recieving, passing, rushing, kicking, and defensive plays broken apart by year and player. Once we had this data we had every player drafted and their corresponding carrer statistics. There isn't code for player statistic because the information was already neatly packaged in csv files. We also needed to collect the number of pro-bowls each player went to since that is an important indicator of the success of a player.

3 Code for scraping draft information

Since we stored all important data in sql databases most of the code involved in data collection and cleaning will be shown in markdown due to the fact that the files are large the code only needs to be run once to structure the data in a usable manner.

from bs4 import BeautifulSoup as bs import pandas as pd import sqlite3 import requests from os import path import re import urllib

def scrape(url, year):

```
g = requests.get(url)
page = bs(g.content)
table = page.findAll('table')
draft = pd.read_html(str(table))
rounds = 1
returned = pd.DataFrame(columns = ['pick', 'team', 'player', 'position', 'college', 'year', 'r
for i in draft:
    result = i
    result.columns = ['pick', 'team', 'player', 'position', 'college']
    result['year'] = year
    result['round'] = rounds
    rounds += 1
    returned = returned.append(result)
return returned
```

if **name** == "**main**": drafts = pd.DataFrame(columns = ['pick', 'team', 'player', 'position', 'college', 'year', 'round']) for year in range(1970,2016): d = scrape('http://www.nfl.com/draft/history/fulldraft?season='+str(year)+'&round=round2#round2', year) drafts = drafts.append(d) drafts = drafts.reset_index() conn = sqlite3.connect('Player.sqlite') drafts.to_sql('DraftInfo', con=conn)

3.1 Pro-Bowl scraper

from bs4 import BeautifulSoup as bs import pandas as pd import sqlite3 import requests from os import path import re import urllib

def scrape(url):