# NFL Rookie Salary Guide



## **Motivation**

- The Superbowl is consistently the most watched primetime event on cable television and playing in the Superbowl comes not only with a title but also with a hefty bonus
- Spotrac.com is a sports database that gathers varying salary's in sports stating with when a player enter any professional league
- I am interested in studying varying factors that maximize a rookie's initial salary when they enter the NFL including the implementation of new salary guidelines and the team a rookie signs with

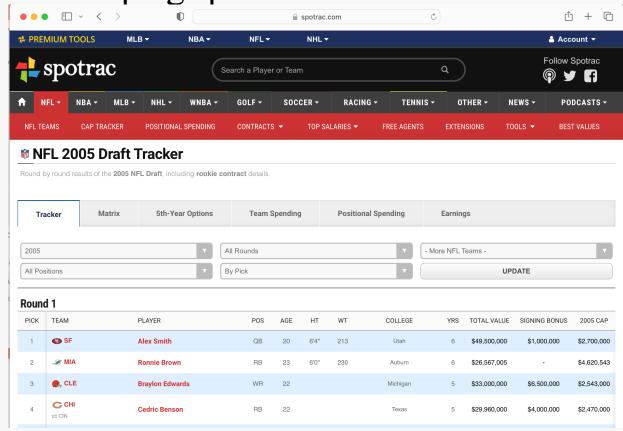


## **Research Questions**

- Does the length of a contract and the team a rookie signs to have an effect on total contract value?
- How did the 2011 CBA effect average contract factors effect contract values when a player enters the NFL? What other draft and contracts factors effect total contract value?



# Webscraping Spotrac

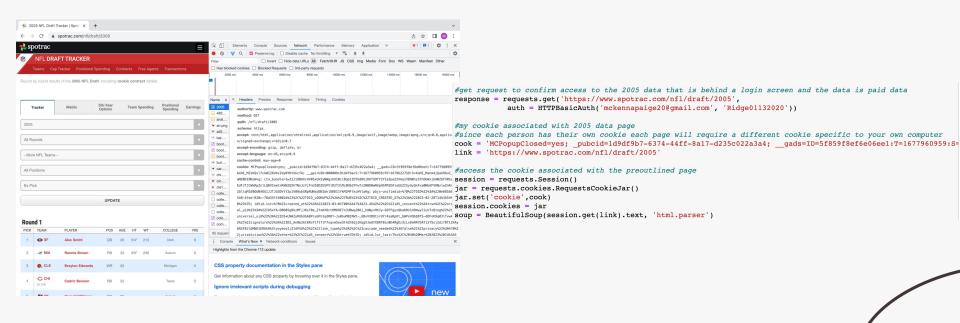




## Challenge in Web-Scraping: using a webdrvier to establish connection

```
#set webdriver
driver = webdriver.Chrome("/Users/mckennamoore/Downloads/chromedrive mac64(01)/chromedriver")
#website username
username = "Mckennapaige20@gmail.com"
#website password
password = "Ridge01132020"
#website login page
url = "https://www.spotrac.com/signin/"
#get request to establish connection with timebreak
driver.get(url)
driver.set page load timeout(70)
#tells website that for the key associated with the value'email'
driver.find element("id", "email").send keys(username)
#tells website that for the key assoicated with the value 'password'
driver.find element("id", "password").send keys(password)
#page load wait time
driver.implicitly wait(20)
#define if there is a popup on the login page before clicking the login button and if so, close the pop up
if EC.presence of element located(driver.find element(By.XPATH, '//*[@id="PopupSignupForm 0"]/div[2]/div[1]')):
    print("Pop up Occured")
   driver.find element(By.XPATH, '//*[@id="PopupSignupForm 0"]/div[2]/div[1]').click()
else:
    print("No pop up")
#click login button associated with XPATH that has the assigned login function
driver.find element(By.XPATH,'//*[@id="contactForm"]/div[2]/input').click()
driver.set page load timeout(70)
#once logged in confirm that I have authoirzation for the data I am accessing because it is paid data
response = requests.get('https://www.spotrac.com/signin',
            auth = HTTPBasicAuth('mckennapaige20@gmail.com', 'Ridge01132020'))
#perform final get request
driver.get("https://www.spotrac.com")
```

# Challenge in Web-Scraping: using cookies to access data





# Key Variables Web-scraping Spotrac.com

- Player name: key for merging to dataframes together
- Position: does position have an effect on salary
- Age: key for merging to dataframes together
- Team: Associated with different yearly salary caps
- College: Associated with division
- Years: Contract length
- Salary Variables



```
#find all values in dictionary with key class and value player which is the name of the value
allplayers= soup.find all('td', {'class':'player'})
#turn accessed values into a dataframe
                                                                                  Sample webscrape for one
allplayers table= pd.DataFrame(allplayers)
#rename column titles
                                                                                  year
allplayers table.columns='Pick Number', 'Player Name', 'Empty'
#ensure that the names of players, which is the targeted variable for this webscrape,
    # is a string variable joined first by a comma then by a space
allplayers table['Player Name'] = allplayers table['Player Name'].str.join(', ')
#find all values in dictionary with key class and value assiated with center
year= soup.find all('td', {'class':'center'})
 #turn these values into a matrix
year matrix=np.matrix(year)
#set the dimensions of the matric based on how many columns there are and how many rows of information
    # (number of players) are in a specified year
years=np.reshape(year matrix,(254,10))
#turn this matrix into a dataframe
year df=pd.DataFrame(years)
#drop undesired columns
year df.columns='Team', 'Position', 'Age', 'Height', 'Weight', 'College', 'Contract Length', 'Total Contract Value'
contract = year df.drop(['Height', 'Team', 'Weight'],1)
contract['Position'] = contract['Position'].str.join(', ')
#turn all values in column 'position' into a string variable
contract['Age'] = contract['Age'].str.join(', ')
#turn all values in column 'college' into a string variable
contract['College'] = contract['College'].str.join(', ')
#turn all values in column 'contract length' into string variable
contract['Contract Length']= contract['Contract Length'].str.join(', ')
#turn all values in 'total contract value' into string
contract['Total Contract Value'] = contract['Total Contract Value'].str.join(', ')
#turn all values in 'signing bonus' into string
contract['Signing Bonus']= contract['Signing Bonus'].str.join(', ')
#turn all values in 'rookie year cap' into string
contract['Rookie Year Cap']= contract['Rookie Year Cap'].str.join(', ')
#define 'player contracts' as the data frame with both data frames
player contracts =[allplayers table,contract]
#merge both data frames horizontally
player_contracts = pd.concat(player_contracts, axis=1)
#drop unwanted columns
player contracts05 = player contracts.drop(['Empty','Pick Number'],1)
```

# Sample cleaning of player names

	Pick Number	Player Name	Empty		Pick Number	Player Name	Empty
0		[Alex Smith]		0		Alex Smith	
1		[Ronnie Brown]		1		Ronnie Brown	
2		[Braylon Edwards]		2		Braylon Edwards	
3		[Cedric Benson]		3		Cedric Benson	
4		[Carnell Williams]		4		Carnell Williams	
249		[Madison Hedgecock]		249		Madison Hedgecock	
250		[David Bergeron]		250		David Bergeron	
251		[J.R. Russell]		251		J.R. Russell	
252		[Doug Nienhuis]		252		Doug Nienhuis	
253		[Andy Stokes]		253		Andy Stokes	

# Challenge in Web-Scraping: Cleaning Salary Information

Total Contract Value	Signing Bonus	Rookie Year Cap
\$49,500,000	\$1,000,000	\$2,700,000
\$26,567,005	-	\$4,620,543
\$33,000,000	\$6,500,000	\$2,543,000
\$29,960,000	\$4,000,000	\$2,470,000
\$31,000,000	\$2,900,000	\$2,327,000
\$952,900	\$27,900	\$239,300
\$1,388,600	\$3,600	\$230,900
\$558,250	\$18,250	\$239,125
\$942,500	\$17,500	\$230,000
\$1,402,500	\$17,500	\$229,375

Total Contract Value	Signing Bonus	Rookie Year Cap
49500000	1000000	2700000
26567005	-	4620543
33000000	6500000	2543000
29960000	4000000	2470000
31000000	2900000	2327000
952900	27900	239300
1388600	3600	230900
558250	18250	239125
942500	17500	230000
1402500	17500	229375



# Challenge in Web-Scraping: Cleaning Salary Information

- allcontracts['Total Contract Value'] = allcontracts['Total Contract Value'].str.replace( ',', ").str.replace('\$', ")
- allcontracts['Signing Bonus'] = allcontracts['Signing Bonus'].str.replace(
- ',', '').str.replace('\$', '')
- allcontracts['Rookie Year Cap'] = allcontracts['Rookie Year Cap'].str.replace(
- ',', '').str.replace('\$', '')



# Output of full webscrape for one year

Utah

North Carolina

Stanford

Louisville

Oregon State

William Penn

Player Name

Madison Hedgecock

David Bergeron

Doug Nienhuis

**Andy Stokes** 

J.R. Russell

0

249

250

251

252

253

Alex Smith

Position Age

QB

FB

LB

WR

Т

TE

23

23

23

23

20

1	Ronnie Brown	RB	23	Auburn	6	\$26,567,005	-	\$4,620,543
2	Braylon Edwards	WR	22	Michigan	5	\$33,000,000	\$6,500,000	\$2,543,000
3	Cedric Benson	RB	22	Texas	5	\$29,960,000	\$4,000,000	\$2,470,000
4	Carnell Williams	RB	23	Auburn	5	\$31,000,000	\$2,900,000	\$2,327,000

6

3

4

2

3

4

Contract Length Total Contract Value Signing Bonus

\$49,500,000

\$952,900

\$1,388,600

\$558,250

\$942,500

\$1,402,500

Rookie Year Cap

\$2,700,000

\$239,300

\$230,900

\$239,125

\$230,000

\$229,375

\$1,000,000

\$27,900

\$3,600

\$18,250

\$17,500

\$17,500

## Merging Dataframes

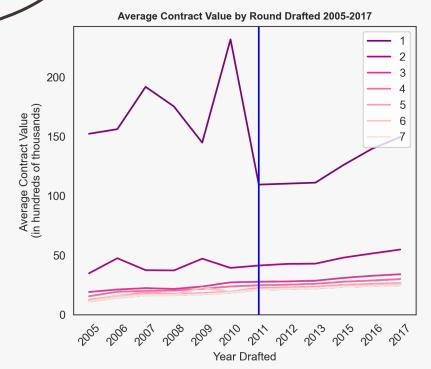
- Merging the yearly information dataframes:
- webscrape = pd.concat([player\_contracts17, player\_contracts16, player\_contracts15, player\_contracts14, player\_contracts13, player\_contracts12, player\_contracts11, player\_contracts10, player\_contracts09, player\_contracts08, player\_contracts07, player\_contracts06, player\_contracts05], axis=0, ignore\_index=True)
- Import external information and turning it into a dataframe
- Draft = pd.read\_csv('AllDrafts.csv', sep=',', header=0, dtype=str)
- Merge both dataframes on shared column titles
- allcontracts = pd.merge(Draft, webscrape, on=['Player Name','Position','College','Age', 'Contract Length'])
- allcontracts.to\_csv('allcontracts.csv')

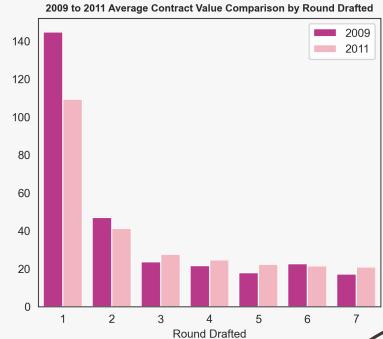


	Round Drafted	Pick Number	Year Drafted	CBA Binary	Team	Team Cap	Player Name	Position	Age	College	Division	Conference	Contract Length	Total Contract Value	Signing Bonus	Rookie Year Cap
,	<b>0</b> 1	1	2005	0	SFO	47574965	Alex Smith	QB	20	Utah	0	Mountain West Conference- West	6	\$49,500,000	\$1,000,000	\$2,700,000
	<b>1</b> 1	2	2005	0	MIA	35210067	Ronnie Brown	RB	23	Auburn	0	South Eastern Conference- West	6	\$26,567,005	-	\$4,620,543
:	<b>2</b> 1	3	2005	0	CLE	28072434	Braylon Edwards	WR	22	Michigan	0	Big 10 Conference- East	5	\$33,000,000	\$6,500,000	\$2,543,000
;	<b>3</b> 1	4	2005	0	СНІ	33891661	Cedric Benson	RB	22	Texas	0	Big 12 Conference	5	\$29,960,000	\$4,000,000	\$2,470,000
	<b>4</b> 1	5	2005	0	TAM	38305975	Carnell Williams	RB	23	Auburn	0	South Eastern Conference- West	5	\$31,000,000	\$2,900,000	\$2,327,000
	<b>.</b>															
24	9 7	252	2005	0	PHI	39907864	Madison Hedgecock	FB	23	North Carolina	0	Mountain West Conference- West	3	\$952,900	\$27,900	\$239,300
25	0 7	253	2005	0	TAM	38305975	David Bergeron	LB	23	Stanford	0	Atlantic Coast Conference- Atlantic	4	\$1,388,600	\$3,600	\$230,900
25	<b>1</b> 7	254	2005	0	SEA	43734928	J.R. Russell	WR	23	Louisville	0	Mountain West Conference- West	2	\$558,250	\$18,250	\$239,125
25	<b>2</b> 7	255	2005	0	NWE	41834687	Doug Nienhuis	т	23	Oregon State	o	Western Mountain West Conference- West	3	\$942,500	\$17,500	\$230,000
25	3 NaN	NaN	NaN	NaN	NaN	NaN	Andy Stokes	TE	23	William Penn	NaN	NaN	4	\$1,402,500	\$17,500	\$229,375

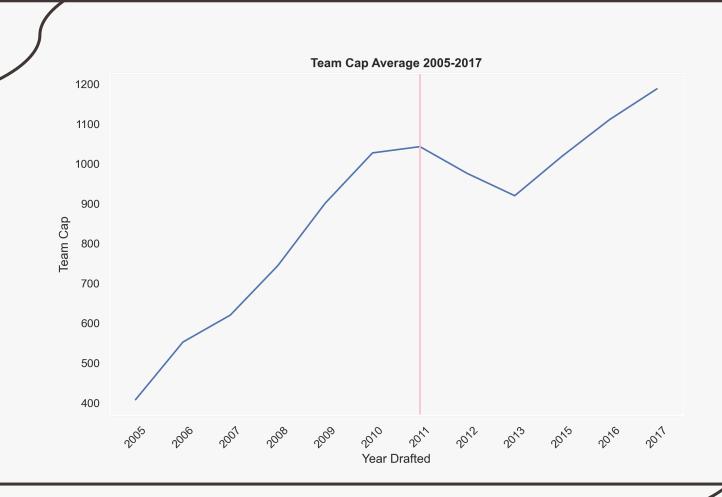
	Round Drafted	Pick Number	Year Drafted	CBA Binary	Team	Team Cap	Player Name	Position	Age	College	Division	Conference	Contract Length	Total Contract Value	Signing Bonus	Rookie Year Cap
o	1	1	2017	1	CLE	77946523	Myles Garrett	DE	21	Texas A&M	0	Big 12 Conference	4	30412255	20258004	5529501
1	1	2	2017	1	СНІ	110629152	Mitchell Trubisky	QB	22	North Carolina	0	Atlantic Coast Conference- Atlantic	4	29032424	19254490	5278622
2	1	3	2017	1	SFO	78589073	Solomon Thomas	DE	21	Stanford	0	Mountain West Conference- West	4	28154366	18615903	5118975
3	1	4	2017	1	JAX	151852626	Leonard Fournette	RB	22	Louisiana State	0	South Eastern Conference- West	4	27150842	17886067	4936516
4	1	5	2017	1	TEN	149109415	Corey Davis	WR	22	Western Michigan	0	Western Big 10 Conference- East	4	25394688	16608864	4617216
248	7	251	2005	0	STL	39773651	Ryan Fitzpatrick	QB	22	Harvard	0	Atlantic Coast Conference- Atlantic	3	953250	28250	239416
249	7	252	2005	0	PHI	39907864	Madison Hedgecock	FB	23	North Carolina	0	Mountain West Conference- West	3	952900	27900	239300
250	7	253	2005	0	TAM	38305975	David Bergeron	LB	23	Stanford	0	Atlantic Coast Conference- Atlantic	4	1388600	3600	230900
251	7	254	2005	o	SEA	43734928	J.R. Russell	WR	23	Louisville	0	Mountain West Conference- West	2	558250	18250	239125
252	7	255	2005	0	NWE	41834687	Doug Nienhuis	т	23	Oregon State	0	Western Mountain West Conference- West	3	942500	17500	230000

# Average Contract Values by Round Drafted











## Models

#### **Base Model:**

Log\_TotalValue =  $\beta_1$ (log\_TeamCap)<sub>i</sub> +  $\beta_2$ (CBABinary)i +  $\mu$ 

#### **Robust Standard Errors:**

Log\_TotalValue =  $\beta_1$ (log\_TeamCap)i +  $\beta_2$ (CBABinary x RoundDrafted)<sub>i</sub> +  $\Phi Z_i$  +  $\mu$  where Zi represents model controls Role, YearDrafted, Age, Length, and Division

#### **Feasble Generalized Least Squares:**

Log\_TotalValue =  $\beta_1$ (log\_TeamCap)i +  $\beta_2$ (CBABinary x RoundDrafted)<sub>i</sub> +  $\Phi$ Z<sub>i</sub> +  $\mu$ , weights = model.residuals

where Zi represents model controls Role, YearDrafted, Age, Length, and Division



	OLS	Controls with SE	FGLS
log of Team Salary Cap	0.280***	-0.117***	-0.085***
	(0.040)	(0.024)	(0.002)
Presence of 2011 CBA	0.257***	0.247***	0.247***
	(0.029)	(0.041)	(800.0)
Drafted Round 2		-1.383***	-1.379***
		(0.027)	(0.006)
Drafted Round 3		-1.895***	-1.843***
		(0.026)	(0.004)
Drafted Round 4		-2.002***	-1.960***
		(0.026)	(0.006)
Drafted Round 5		-2.196***	-2.112***
		(0.027)	(0.006)
Drafted Round 6		-2.201***	-2.173***
		(0.026)	(0.005)
Drafted Round 7		-2.263***	-2.191***
		(0.026)	(0.005)

		(0.034)	(0.007)
Post-2011 Round 5		0.651*** (0.034)	0.574*** (0.007)
Post-2011 Round 6		0.619*** (0.034)	0.603*** (0.007)
Post-2011 Round 7		0.660*** (0.033)	0.592*** (0.007)
Constant	9.578*** (0.728)	18.252*** (0.420)	17.603*** (0.042)
Position FE	No	Yes	Yes
Year Drafted	No	Yes	Yes
Division	No	Yes	Yes
Age	No	Yes	Yes
Observations	2,861	2,861	2,861
$\mathbb{R}^2$	0.102	0.888	1.000
Adjusted R <sup>2</sup>	0.101	0.886	1.000
Residual Std. Error	0.609 (df = 2858)	0.217 (df = 2811)	0.223 (df = 2811)

Controls with SE

0.476\*\*\*

(0.035)

0.566\*\*\*

(0.034)

0.558\*\*\*

161.455\*\*\* (df = 2; 2858) 455.393\*\*\* (df = 49; 2811) 1,050,594.000\*\*\* (df = 49; 2811)

**FGLS** 

0.526\*\*\*

(0.007) 0.529\*\*\*

(0.006)

0.542\*\*\*

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Source: Spotrac.com

OLS

Post-2011 Round 2

Post-2011 Round 3

Post-2011 Round 4

F Statistic

Note:

## **Discussions**

- My first attempt at controlling for heteroskedasticity was using robust standard errors on my OLS model.
- The log of team salary cap states that for every 1% change in team salary cap, the total value of a rookie's contract will decrease by 0.117%.
- The Presence of 2011 CBA variable indicates that if a candidate has the variable, or they are drafted after the implementation for the 2011 CBA, their rookie contract will by 0.247% higher than those without the variable.
- The interaction between round drafted and the CBA decreases the negative effect that not being drafted in the first round has on your salary. Pre-2011 round 2 salary decreases by 1.383%. The additional marginal affect of being drafted after 2011 in round 2 is 0.907% increase relative to the round 1 reference.



## **Discussions**

- My preferred model is my FGLS model because it controls for heteroskedasticity stemming from the residuals in my third OLS model.
- The log of team salary cap states that for every 1% change in team salary cap, the total value of a rookie's contract will decrease by 0.085%.
- The Presence of 2011 CBA variable indicates that if a candidate has the variable, or they are drafted after the implementation for the 2011 CBA, their rookie contract will by 0.247% higher than those without the variable.
- The interaction between round drafted and the CBA decreases the negative effect that not being drafted in the first round has on your salary. Pre-2011 round 2 salary decreases by 1.379%. The additional marginal affect of being drafted after 2011 in round 2 is 0.853% increase relative to the round 1 reference.



## Conclusion and Future Study

- The presence of the 2011 CBA has a great effect on decreasing the impact of not being drafted in the first round
- Increase in team salary cap has a negative effect on individual salary
- Assuming this is because the CBA decreased individual compensation
- Greatest effect of CBA is presented in candidates drafted in later rounds

