

Extra Feature Scraping

The purpose of this notebook is to scrape more features, including:

- Team Salaries
- Player Salaries
- Playoff Histories

```
In [ ]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import matplotlib.ticker as mtick
import sqlite3
import seaborn as sns
from matplotlib.offsetbox import OffsetImage, AnnotationBbox
from selenium import webdriver
from selenium.webdriver.common.keys import Keys
from bs4 import BeautifulSoup
from selenium.webdriver.common.by import By
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
import time
import requests
import shutil
import datetime
from scipy.stats import norm
from sklearn.preprocessing import OneHotEncoder, OrdinalEncoder
import os
import winsound
```

```
In [ ]: os.chdir('C:\\\\Users\\\\Travis\\\\OneDrive\\\\Data Science\\\\Personal_Projects\\\\Sports\\\\NBA')
```

```
In [ ]: options = np.arange(2000,2022,1)
```

Team Salaries

```
In [ ]: urlz = []
for i in options:
    url = 'https://hoopshype.com/salaries/' + str(i) + '-' + str(i+1) + '/'
    urlz.append(url)
urlz
```

```
Out[ ]: ['https://hoopshype.com/salaries/2000-2001/',
 'https://hoopshype.com/salaries/2001-2002/',
 'https://hoopshype.com/salaries/2002-2003/',
 'https://hoopshype.com/salaries/2003-2004/',
 'https://hoopshype.com/salaries/2004-2005/',
 'https://hoopshype.com/salaries/2005-2006/',
 'https://hoopshype.com/salaries/2006-2007/',
 'https://hoopshype.com/salaries/2007-2008/',
 'https://hoopshype.com/salaries/2008-2009/',
 'https://hoopshype.com/salaries/2009-2010/',
 'https://hoopshype.com/salaries/2010-2011/',
 'https://hoopshype.com/salaries/2011-2012/',
 'https://hoopshype.com/salaries/2012-2013/',
 'https://hoopshype.com/salaries/2013-2014/',
 'https://hoopshype.com/salaries/2014-2015/',
 'https://hoopshype.com/salaries/2015-2016/',
 'https://hoopshype.com/salaries/2016-2017/',
 'https://hoopshype.com/salaries/2017-2018/',
 'https://hoopshype.com/salaries/2018-2019/',
 'https://hoopshype.com/salaries/2019-2020/',
 'https://hoopshype.com/salaries/2020-2021/',
 'https://hoopshype.com/salaries/2021-2022/']
```

```
In [ ]: # if team_contracts not a folder, make it
if not os.path.exists('data/team/team_contracts'):
    os.makedirs('data/team/team_contracts')
```

```
In [ ]: for u in urlz:
    df = pd.read_html(u)[0]
    df.to_csv('data/team/team_contracts/' + str(u[-10:-1]) + '.csv')
```

Player Salaries

```
In [ ]: # if player_contracts not a folder, make it
if not os.path.exists('data/player/contracts'):
    os.makedirs('data/player/contracts')
```

```
In [ ]: player_urls = []
for i in options:
    url = 'https://hoopshype.com/salaries/players/' + str(i) + '-' + str(i+1) + '/'
    player_urls.append(url)
```

```
In [ ]: for u in player_urls:
    df = pd.read_html(u)[0]
    df.to_csv('data/player/contracts/' + str(u[-10:-1]) + '.csv')
```

Playoff History

```
In [ ]: playoff_history = 'https://www.basketball-reference.com/playoffs/series.html'
df = pd.read_html(playoff_history)[0]
df.to_csv('data/team/playoff_history.csv')
df
```

Out[]:

	Unnamed: 0_level_0	Unnamed: 1_level_0	Unnamed: 2_level_0	Unnamed: 3_level_0	Unnamed: 4_level_0	Winner	Unnamed: 7_level_0	
	Yr	Lg	Series	Unnamed: 3_level_1	Unnamed: 4_level_1	Team	W	Unnamed: 7_level_1
0	2022	NBA	Eastern Conf First Round	Apr 17 - Apr 27, 2022	NaN	Milwaukee Bucks (3)	4	NaN Chic
1	2022	NBA	Western Conf First Round	Apr 16 - Apr 29, 2022	NaN	Memphis Grizzlies (2)	4	NaN Tim
2	2022	NBA	Eastern Conf First Round	Apr 17 - Apr 25, 2022	NaN	Boston Celtics (2)	4	NaN
3	2022	NBA	Western Conf First Round	Apr 16 - Apr 27, 2022	NaN	Golden State Warriors (3)	4	NaN Nu
4	2022	NBA	Western Conf First Round	Apr 16 - Apr 28, 2022	NaN	Dallas Mavericks (4)	4	NaN Uta
...
998	1947	BAA	Quarterfinals	Apr 2 - Apr 9, 1947	NaN	New York Knicks (3)	2	NaN C
999	1947	BAA	Quarterfinals	Apr 2 - Apr 6, 1947	NaN	Philadelphia Warriors (2)	2	NaN Bo
1000	1947	BAA	Semifinals	Apr 2 - Apr 13, 1947	NaN	Chicago Stags (1)	4	NaN Wa
1001	1947	BAA	Semifinals	Apr 12 - Apr 14, 1947	NaN	Philadelphia Warriors (2)	2	NaN Ca
1002	1947	BAA	Finals	Apr 16 - Apr 22, 1947	NaN	Philadelphia Warriors (2)	4	NaN

1003 rows × 13 columns

```
In [ ]: winners = df[['Winner', 'Team']]
yr = df[['Unnamed: 0_level_0', 'Yr']]
round = df[['Unnamed: 2_level_0', 'Series']]
losers = df[['Loser', 'Team']]
favorite = df[['Unnamed: 11_level_0', 'Favorite']]
underdog = df[['Unnamed: 12_level_0', 'Underdog']]

new_df = pd.concat([yr, round, winners, losers, favorite, underdog], axis=1)
new_df
```

Out[]:

	Unnamed: 0_level_0	Unnamed: 2_level_0	Winner	Loser	Unnamed: 11_level_0	Unnamed: 12_level_0
	Yr	Series	Team	Team	Favorite	Underdog
0	2022	Eastern Conf First Round	Milwaukee Bucks (3)	Chicago Bulls (6)	MIL (-1100)	CHI (+700)
1	2022	Western Conf First Round	Memphis Grizzlies (2)	Minnesota Timberwolves (7)	MEM (-320)	MIN (+260)
2	2022	Eastern Conf First Round	Boston Celtics (2)	Brooklyn Nets (7)	BOS (-130)	BRK (+110)
3	2022	Western Conf First Round	Golden State Warriors (3)	Denver Nuggets (6)	GSW (-275)	DEN (+225)
4	2022	Western Conf First Round	Dallas Mavericks (4)	Utah Jazz (5)	UTA (-300)	DAL (+245)
...
998	1947	Quarterfinals	New York Knicks (3)	Cleveland Rebels (3)	NaN	NaN
999	1947	Quarterfinals	Philadelphia Warriors (2)	St. Louis Bombers (2)	NaN	NaN
1000	1947	Semifinals	Chicago Stags (1)	Washington Capitols (1)	NaN	NaN
1001	1947	Semifinals	Philadelphia Warriors (2)	New York Knicks (3)	NaN	NaN
1002	1947	Finals	Philadelphia Warriors (2)	Chicago Stags (1)	NaN	NaN

1003 rows × 6 columns

In []: `new_df = new_df.dropna()`
`new_df`

Out[]:

	Unnamed: 0_level_0	Unnamed: 2_level_0	Winner	Loser	Unnamed: 11_level_0	Unnamed: 12_level_0
	Yr	Series	Team	Team	Favorite	Underdog
0	2022	Eastern Conf First Round	Milwaukee Bucks (3)	Chicago Bulls (6)	MIL (-1100)	CHI (+700)
1	2022	Western Conf First Round	Memphis Grizzlies (2)	Minnesota Timberwolves (7)	MEM (-320)	MIN (+260)
2	2022	Eastern Conf First Round	Boston Celtics (2)	Brooklyn Nets (7)	BOS (-130)	BRK (+110)
3	2022	Western Conf First Round	Golden State Warriors (3)	Denver Nuggets (6)	GSW (-275)	DEN (+225)
4	2022	Western Conf First Round	Dallas Mavericks (4)	Utah Jazz (5)	UTA (-300)	DAL (+245)
...
893	Yr	Series	Team	Team	Favorite	Underdog
911	Yr	Series	Team	Team	Favorite	Underdog
933	Yr	Series	Team	Team	Favorite	Underdog
951	Yr	Series	Team	Team	Favorite	Underdog
981	Yr	Series	Team	Team	Favorite	Underdog

539 rows × 6 columns

In []: `new_df.to_csv('data/team/playoffs_history_cleaned.csv')`

In []: `finals = new_df[new_df['Unnamed: 2_level_0'] == 'Finals']`

Out[]:	Unnamed: 0_level_0	Unnamed: 2_level_0	Winner	Loser	Unnamed: 11_level_0	Unnamed: 12_level_0
	Yr	Series	Team	Team	Favorite	Underdog
30	2021	Finals	Milwaukee Bucks (3)	Phoenix Suns (2)	PHO (-180)	MIL (+160)
46	2020	Finals	Los Angeles Lakers (1)	Miami Heat (5)	LAL (-350)	MIA (+290)
62	2019	Finals	Toronto Raptors (2)	Golden State Warriors (1)	GSW (-270)	TOR (+230)
78	2018	Finals	Golden State Warriors (2)	Cleveland Cavaliers (4)	GSW (-1075)	CLE (+688)
94	2017	Finals	Golden State Warriors (1)	Cleveland Cavaliers (2)	GSW (-300)	CLE (+250)
110	2016	Finals	Cleveland Cavaliers (1)	Golden State Warriors (1)	GSW (-220)	CLE (+180)
126	2015	Finals	Golden State Warriors (1)	Cleveland Cavaliers (2)	GSW (-220)	CLE (+190)
142	2014	Finals	San Antonio Spurs (1)	Miami Heat (2)	SAS (-155)	MIA (+135)
158	2013	Finals	Miami Heat (1)	San Antonio Spurs (2)	MIA (-220)	SAS (+180)
174	2012	Finals	Miami Heat (2)	Oklahoma City Thunder (2)	OKC (-175)	MIA (+155)
190	2011	Finals	Dallas Mavericks (3)	Miami Heat (2)	MIA (-175)	DAL (+155)
206	2010	Finals	Los Angeles Lakers (1)	Boston Celtics (4)	LAL (-190)	BOS (+165)
222	2009	Finals	Los Angeles Lakers (1)	Orlando Magic (3)	LAL (-280)	ORL (+240)
238	2008	Finals	Boston Celtics (1)	Los Angeles Lakers (1)	LAL (-180)	BOS (+160)
254	2007	Finals	San Antonio Spurs (3)	Cleveland Cavaliers (2)	SAS (-450)	CLE (+360)
270	2006	Finals	Miami Heat (2)	Dallas Mavericks (4)	DAL (-150)	MIA (+130)
286	2005	Finals	San Antonio Spurs (2)	Detroit Pistons (2)	SAS (-250)	DET (+200)
302	2004	Finals	Detroit Pistons (3)	Los Angeles Lakers (2)	LAL (-700)	DET (+500)
318	2003	Finals	San Antonio Spurs (1)	New Jersey Nets (2)	SAS (-330)	NJN (+260)
334	2002	Finals	Los Angeles Lakers (3)	New Jersey Nets (1)	LAL (-750)	NJN (+550)

	Unnamed: 0_level_0	Unnamed: 2_level_0	Winner	Loser	Unnamed: 11_level_0	Unnamed: 12_level_0
	Yr	Series	Team	Team	Favorite	Underdog
350	2001	Finals	Los Angeles Lakers (2)	Philadelphia 76ers (1)	LAL (-2000)	PHI (+1600)
366	2000	Finals	Los Angeles Lakers (1)	Indiana Pacers (1)	LAL (-800)	IND (+550)
382	1999	Finals	San Antonio Spurs (1)	New York Knicks (8)	SAS (-900)	NYK (+600)
398	1998	Finals	Chicago Bulls (1)	Utah Jazz (1)	CHI (-125)	UTA (+105)
414	1997	Finals	Chicago Bulls (1)	Utah Jazz (1)	CHI (-600)	UTA (+400)
430	1996	Finals	Chicago Bulls (1)	Seattle SuperSonics (1)	CHI (-950)	SEA (+650)
446	1995	Finals	Houston Rockets (6)	Orlando Magic (1)	ORL (-150)	HOU (+130)
462	1994	Finals	Houston Rockets (2)	New York Knicks (2)	HOU (-220)	NYK (+180)
478	1993	Finals	Chicago Bulls (2)	Phoenix Suns (1)	CHI (-240)	PHO (+190)
494	1992	Finals	Chicago Bulls (1)	Portland Trail Blazers (1)	CHI (-250)	POR (+200)
510	1991	Finals	Chicago Bulls (1)	Los Angeles Lakers (3)	CHI (-200)	LAL (+170)
526	1990	Finals	Detroit Pistons (1)	Portland Trail Blazers (3)	DET (-400)	POR (+300)
542	1989	Finals	Detroit Pistons (1)	Los Angeles Lakers (1)	LAL (-120)	DET (+100)
558	1988	Finals	Los Angeles Lakers (1)	Detroit Pistons (2)	LAL (-260)	DET (+210)

```
In [ ]: finals.to_csv('data/team/playoffs_history_finals.csv')
```

```
In [ ]: fin_df = pd.read_csv('data/team/playoffs_history_finals.csv', skiprows = [0])
```

```
In [ ]: fin_df = fin_df.rename(columns={'Team': 'Winning_team', 'Team.1' : 'Losing_team'})
```

Out[]:	Unnamed: 0	Yr	Series	Winning_team	Losing_team	Favorite	Underdog	
	0	30	2021	Finals	Milwaukee Bucks (3)	Phoenix Suns (2)	PHO (-180)	MIL (+160)
	1	46	2020	Finals	Los Angeles Lakers (1)	Miami Heat (5)	LAL (-350)	MIA (+290)
	2	62	2019	Finals	Toronto Raptors (2)	Golden State Warriors (1)	GSW (-270)	TOR (+230)
	3	78	2018	Finals	Golden State Warriors (2)	Cleveland Cavaliers (4)	GSW (-1075)	CLE (+688)
	4	94	2017	Finals	Golden State Warriors (1)	Cleveland Cavaliers (2)	GSW (-300)	CLE (+250)
	5	110	2016	Finals	Cleveland Cavaliers (1)	Golden State Warriors (1)	GSW (-220)	CLE (+180)
	6	126	2015	Finals	Golden State Warriors (1)	Cleveland Cavaliers (2)	GSW (-220)	CLE (+190)
	7	142	2014	Finals	San Antonio Spurs (1)	Miami Heat (2)	SAS (-155)	MIA (+135)
	8	158	2013	Finals	Miami Heat (1)	San Antonio Spurs (2)	MIA (-220)	SAS (+180)
	9	174	2012	Finals	Miami Heat (2)	Oklahoma City Thunder (2)	OKC (-175)	MIA (+155)
	10	190	2011	Finals	Dallas Mavericks (3)	Miami Heat (2)	MIA (-175)	DAL (+155)
	11	206	2010	Finals	Los Angeles Lakers (1)	Boston Celtics (4)	LAL (-190)	BOS (+165)
	12	222	2009	Finals	Los Angeles Lakers (1)	Orlando Magic (3)	LAL (-280)	ORL (+240)
	13	238	2008	Finals	Boston Celtics (1)	Los Angeles Lakers (1)	LAL (-180)	BOS (+160)
	14	254	2007	Finals	San Antonio Spurs (3)	Cleveland Cavaliers (2)	SAS (-450)	CLE (+360)
	15	270	2006	Finals	Miami Heat (2)	Dallas Mavericks (4)	DAL (-150)	MIA (+130)
	16	286	2005	Finals	San Antonio Spurs (2)	Detroit Pistons (2)	SAS (-250)	DET (+200)
	17	302	2004	Finals	Detroit Pistons (3)	Los Angeles Lakers (2)	LAL (-700)	DET (+500)
	18	318	2003	Finals	San Antonio Spurs (1)	New Jersey Nets (2)	SAS (-330)	NJN (+260)
	19	334	2002	Finals	Los Angeles Lakers (3)	New Jersey Nets (1)	LAL (-750)	NJN (+550)
	20	350	2001	Finals	Los Angeles Lakers (2)	Philadelphia 76ers (1)	LAL (-2000)	PHI (+1600)

	Unnamed: 0	Yr	Series	Winning_team	Losing_team	Favorite	Underdog
21	366	2000	Finals	Los Angeles Lakers (1)	Indiana Pacers (1)	LAL (-800)	IND (+550)
22	382	1999	Finals	San Antonio Spurs (1)	New York Knicks (8)	SAS (-900)	NYK (+600)
23	398	1998	Finals	Chicago Bulls (1)	Utah Jazz (1)	CHI (-125)	UTA (+105)
24	414	1997	Finals	Chicago Bulls (1)	Utah Jazz (1)	CHI (-600)	UTA (+400)
25	430	1996	Finals	Chicago Bulls (1)	Seattle SuperSonics (1)	CHI (-950)	SEA (+650)
26	446	1995	Finals	Houston Rockets (6)	Orlando Magic (1)	ORL (-150)	HOU (+130)
27	462	1994	Finals	Houston Rockets (2)	New York Knicks (2)	HOU (-220)	NYK (+180)
28	478	1993	Finals	Chicago Bulls (2)	Phoenix Suns (1)	CHI (-240)	PHO (+190)
29	494	1992	Finals	Chicago Bulls (1)	Portland Trail Blazers (1)	CHI (-250)	POR (+200)
30	510	1991	Finals	Chicago Bulls (1)	Los Angeles Lakers (3)	CHI (-200)	LAL (+170)
31	526	1990	Finals	Detroit Pistons (1)	Portland Trail Blazers (3)	DET (-400)	POR (+300)
32	542	1989	Finals	Detroit Pistons (1)	Los Angeles Lakers (1)	LAL (-120)	DET (+100)
33	558	1988	Finals	Los Angeles Lakers (1)	Detroit Pistons (2)	LAL (-260)	DET (+210)

```
In [ ]: fin_df = fin_df.drop(columns = ['Unnamed: 0'])
```

```
In [ ]: # split Favorite column by ' '
fin_df['Favorite'] = fin_df['Favorite'].str.split(' ').str[0]
fin_df['Underdog'] = fin_df['Underdog'].str.split(' ').str[0]
```

```
In [ ]: fin_df['Winning_team'] = fin_df['Winning_team'].str.split('(').str[0]
fin_df['Winning_team'] = fin_df['Winning_team'].str[:-1]
fin_df['Losing_team'] = fin_df['Losing_team'].str.split('(').str[0]
fin_df['Losing_team'] = fin_df['Losing_team'].str[:-1]
fin_df
```

Out[]:

	Yr	Series	Winning_team	Losing_team	Favorite	Underdog
0	2021	Finals	Milwaukee Bucks	Phoenix Suns	PHO	MIL
1	2020	Finals	Los Angeles Lakers	Miami Heat	LAL	MIA
2	2019	Finals	Toronto Raptors	Golden State Warriors	GSW	TOR
3	2018	Finals	Golden State Warriors	Cleveland Cavaliers	GSW	CLE
4	2017	Finals	Golden State Warriors	Cleveland Cavaliers	GSW	CLE
5	2016	Finals	Cleveland Cavaliers	Golden State Warriors	GSW	CLE
6	2015	Finals	Golden State Warriors	Cleveland Cavaliers	GSW	CLE
7	2014	Finals	San Antonio Spurs	Miami Heat	SAS	MIA
8	2013	Finals	Miami Heat	San Antonio Spurs	MIA	SAS
9	2012	Finals	Miami Heat	Oklahoma City Thunder	OKC	MIA
10	2011	Finals	Dallas Mavericks	Miami Heat	MIA	DAL
11	2010	Finals	Los Angeles Lakers	Boston Celtics	LAL	BOS
12	2009	Finals	Los Angeles Lakers	Orlando Magic	LAL	ORL
13	2008	Finals	Boston Celtics	Los Angeles Lakers	LAL	BOS
14	2007	Finals	San Antonio Spurs	Cleveland Cavaliers	SAS	CLE
15	2006	Finals	Miami Heat	Dallas Mavericks	DAL	MIA
16	2005	Finals	San Antonio Spurs	Detroit Pistons	SAS	DET
17	2004	Finals	Detroit Pistons	Los Angeles Lakers	LAL	DET
18	2003	Finals	San Antonio Spurs	New Jersey Nets	SAS	NJN
19	2002	Finals	Los Angeles Lakers	New Jersey Nets	LAL	NJN
20	2001	Finals	Los Angeles Lakers	Philadelphia 76ers	LAL	PHI
21	2000	Finals	Los Angeles Lakers	Indiana Pacers	LAL	IND
22	1999	Finals	San Antonio Spurs	New York Knicks	SAS	NYK
23	1998	Finals	Chicago Bulls	Utah Jazz	CHI	UTA
24	1997	Finals	Chicago Bulls	Utah Jazz	CHI	UTA
25	1996	Finals	Chicago Bulls	Seattle SuperSonics	CHI	SEA
26	1995	Finals	Houston Rockets	Orlando Magic	ORL	HOU
27	1994	Finals	Houston Rockets	New York Knicks	HOU	NYK
28	1993	Finals	Chicago Bulls	Phoenix Suns	CHI	PHO
29	1992	Finals	Chicago Bulls	Portland Trail Blazers	CHI	POR
30	1991	Finals	Chicago Bulls	Los Angeles Lakers	CHI	LAL
31	1990	Finals	Detroit Pistons	Portland Trail Blazers	DET	POR
32	1989	Finals	Detroit Pistons	Los Angeles Lakers	LAL	DET

Yr	Series	Winning_team	Losing_team	Favorite	Underdog
33	1988 Finals	Los Angeles Lakers	Detroit Pistons	LAL	DET

```
In [ ]: def teamname_to_abb(teamname):
    if teamname == 'New York Knicks':
        return 'NYK'
    elif teamname == 'Los Angeles Lakers':
        return 'LAL'
    elif teamname == 'Boston Celtics':
        return 'BOS'
    elif teamname == 'Philadelphia 76ers':
        return 'PHI'
    elif teamname == 'Toronto Raptors':
        return 'TOR'
    elif teamname == 'Chicago Bulls':
        return 'CHI'
    elif teamname == 'Oklahoma City Thunder':
        return 'OKC'
    elif teamname == 'Charlotte Hornets':
        return 'CHO'
    elif teamname == 'Detroit Pistons':
        return 'DET'
    elif teamname == 'Indiana Pacers':
        return 'IND'
    elif teamname == 'Milwaukee Bucks':
        return 'MIL'
    elif teamname == 'Washington Wizards':
        return 'WAS'
    elif teamname == 'Orlando Magic':
        return 'ORL'
    elif teamname == 'Miami Heat':
        return 'MIA'
    elif teamname == 'New Orleans Pelicans':
        return 'NOP'
    elif teamname == 'Atlanta Hawks':
        return 'ATL'
    elif teamname == 'Charlotte Bobcats':
        return 'CHA'
    elif teamname == 'Cleveland Cavaliers':
        return 'CLE'
    elif teamname == 'Dallas Mavericks':
        return 'DAL'
    elif teamname == 'Denver Nuggets':
        return 'DEN'
    elif teamname == 'Golden State Warriors':
        return 'GSW'
    elif teamname == 'Houston Rockets':
        return 'HOU'
    elif teamname == 'Memphis Grizzlies':
        return 'MEM'
    elif teamname == 'Minnesota Timberwolves':
        return 'MIN'
    elif teamname == 'New Jersey Nets':
        return 'NJN'
    elif teamname == 'Oklahoma City Thunder':
        return 'OKC'
    elif teamname == 'Portland Trail Blazers':
        return 'POR'
```

```
In [ ]: fin_df['Winner_abv'] = fin_df.apply(lambda row: teamname_to_abb(row['Winning_team'])
fin_df['Loser_abv'] = fin_df.apply(lambda row: teamname_to_abb(row['Losing_team'])),
fin_df
```

Out[]:	Yr	Series	Winning_team	Losing_team	Favorite	Underdog	Winner_abv	Loser_abv
0	2021	Finals	Milwaukee Bucks	Phoenix Suns	PHO	MIL	MIL	None
1	2020	Finals	Los Angeles Lakers	Miami Heat	LAL	MIA	LAL	MIA
2	2019	Finals	Toronto Raptors	Golden State Warriors	GSW	TOR	TOR	GSW
3	2018	Finals	Golden State Warriors	Cleveland Cavaliers	GSW	CLE	GSW	CLE
4	2017	Finals	Golden State Warriors	Cleveland Cavaliers	GSW	CLE	GSW	CLE
5	2016	Finals	Cleveland Cavaliers	Golden State Warriors	GSW	CLE	CLE	GSW
6	2015	Finals	Golden State Warriors	Cleveland Cavaliers	GSW	CLE	GSW	CLE
7	2014	Finals	San Antonio Spurs	Miami Heat	SAS	MIA	None	MIA
8	2013	Finals	Miami Heat	San Antonio Spurs	MIA	SAS	MIA	None
9	2012	Finals	Miami Heat	Oklahoma City Thunder	OKC	MIA	MIA	OKC
10	2011	Finals	Dallas Mavericks	Miami Heat	MIA	DAL	DAL	MIA
11	2010	Finals	Los Angeles Lakers	Boston Celtics	LAL	BOS	LAL	BOS
12	2009	Finals	Los Angeles Lakers	Orlando Magic	LAL	ORL	LAL	ORL
13	2008	Finals	Boston Celtics	Los Angeles Lakers	LAL	BOS	BOS	LAL
14	2007	Finals	San Antonio Spurs	Cleveland Cavaliers	SAS	CLE	None	CLE
15	2006	Finals	Miami Heat	Dallas Mavericks	DAL	MIA	MIA	DAL
16	2005	Finals	San Antonio Spurs	Detroit Pistons	SAS	DET	None	DET
17	2004	Finals	Detroit Pistons	Los Angeles Lakers	LAL	DET	DET	LAL
18	2003	Finals	San Antonio Spurs	New Jersey Nets	SAS	NJN	None	NJN
19	2002	Finals	Los Angeles Lakers	New Jersey Nets	LAL	NJN	LAL	NJN
20	2001	Finals	Los Angeles Lakers	Philadelphia 76ers	LAL	PHI	LAL	PHI

	Yr	Series	Winning_team	Losing_team	Favorite	Underdog	Winner_abv	Loser_abv
21	2000	Finals	Los Angeles Lakers	Indiana Pacers	LAL	IND	LAL	IND
22	1999	Finals	San Antonio Spurs	New York Knicks	SAS	NYK	None	NYK
23	1998	Finals	Chicago Bulls	Utah Jazz	CHI	UTA	CHI	None
24	1997	Finals	Chicago Bulls	Utah Jazz	CHI	UTA	CHI	None
25	1996	Finals	Chicago Bulls	Seattle SuperSonics	CHI	SEA	CHI	None
26	1995	Finals	Houston Rockets	Orlando Magic	ORL	HOU	HOU	ORL
27	1994	Finals	Houston Rockets	New York Knicks	HOU	NYK	HOU	NYK
28	1993	Finals	Chicago Bulls	Phoenix Suns	CHI	PHO	CHI	None
29	1992	Finals	Chicago Bulls	Portland Trail Blazers	CHI	POR	CHI	POR
30	1991	Finals	Chicago Bulls	Los Angeles Lakers	CHI	LAL	CHI	LAL
31	1990	Finals	Detroit Pistons	Portland Trail Blazers	DET	POR	DET	POR
32	1989	Finals	Detroit Pistons	Los Angeles Lakers	LAL	DET	DET	LAL
33	1988	Finals	Los Angeles Lakers	Detroit Pistons	LAL	DET	LAL	DET

```
In [ ]: fin_df.to_csv('data/team/cleaned_finals_results.csv')
```

```
In [ ]: fin_df['predicted?'] = np.where(fin_df['Winner_abv'] == fin_df['Favorite'], 1, 0)
```

```
In [ ]: fin_df
```

	Out[]:	Yr	Series	Winning_team	Losing_team	Favorite	Underdog	Winner_abv	Loser_abv	predic
0	2021	Finals	Milwaukee Bucks	Phoenix Suns	PHO	MIL	MIL	MIL	None	
1	2020	Finals	Los Angeles Lakers	Miami Heat	LAL	MIA	LAL	MIA	MIA	
2	2019	Finals	Toronto Raptors	Golden State Warriors	GSW	TOR	TOR	GSW	GSW	
3	2018	Finals	Golden State Warriors	Cleveland Cavaliers	GSW	CLE	GSW	CLE	CLE	
4	2017	Finals	Golden State Warriors	Cleveland Cavaliers	GSW	CLE	GSW	CLE	CLE	
5	2016	Finals	Cleveland Cavaliers	Golden State Warriors	GSW	CLE	CLE	GSW	GSW	
6	2015	Finals	Golden State Warriors	Cleveland Cavaliers	GSW	CLE	GSW	CLE	CLE	
7	2014	Finals	San Antonio Spurs	Miami Heat	SAS	MIA	None	MIA	MIA	
8	2013	Finals	Miami Heat	San Antonio Spurs	MIA	SAS	MIA	None	None	
9	2012	Finals	Miami Heat	Oklahoma City Thunder	OKC	MIA	MIA	OKC	OKC	
10	2011	Finals	Dallas Mavericks	Miami Heat	MIA	DAL	DAL	MIA	MIA	
11	2010	Finals	Los Angeles Lakers	Boston Celtics	LAL	BOS	LAL	BOS	BOS	
12	2009	Finals	Los Angeles Lakers	Orlando Magic	LAL	ORL	LAL	ORL	ORL	
13	2008	Finals	Boston Celtics	Los Angeles Lakers	LAL	BOS	BOS	LAL	LAL	
14	2007	Finals	San Antonio Spurs	Cleveland Cavaliers	SAS	CLE	None	CLE	CLE	
15	2006	Finals	Miami Heat	Dallas Mavericks	DAL	MIA	MIA	DAL	DAL	
16	2005	Finals	San Antonio Spurs	Detroit Pistons	SAS	DET	None	DET	DET	
17	2004	Finals	Detroit Pistons	Los Angeles Lakers	LAL	DET	DET	LAL	LAL	
18	2003	Finals	San Antonio Spurs	New Jersey Nets	SAS	NJN	None	NJN	NJN	
19	2002	Finals	Los Angeles Lakers	New Jersey Nets	LAL	NJN	LAL	NJN	NJN	
20	2001	Finals	Los Angeles Lakers	Philadelphia 76ers	LAL	PHI	LAL	PHI	PHI	

	Yr	Series	Winning_team	Losing_team	Favorite	Underdog	Winner_abv	Loser_abv	predic
21	2000	Finals	Los Angeles Lakers	Indiana Pacers	LAL	IND	LAL	IND	
22	1999	Finals	San Antonio Spurs	New York Knicks	SAS	NYK	None	NYK	
23	1998	Finals	Chicago Bulls	Utah Jazz	CHI	UTA	CHI	None	
24	1997	Finals	Chicago Bulls	Utah Jazz	CHI	UTA	CHI	None	
25	1996	Finals	Chicago Bulls	Seattle SuperSonics	CHI	SEA	CHI	None	
26	1995	Finals	Houston Rockets	Orlando Magic	ORL	HOU	HOU	ORL	
27	1994	Finals	Houston Rockets	New York Knicks	HOU	NYK	HOU	NYK	
28	1993	Finals	Chicago Bulls	Phoenix Suns	CHI	PHO	CHI	None	
29	1992	Finals	Chicago Bulls	Portland Trail Blazers	CHI	POR	CHI	POR	
30	1991	Finals	Chicago Bulls	Los Angeles Lakers	CHI	LAL	CHI	LAL	
31	1990	Finals	Detroit Pistons	Portland Trail Blazers	DET	POR	DET	POR	
32	1989	Finals	Detroit Pistons	Los Angeles Lakers	LAL	DET	DET	LAL	
33	1988	Finals	Los Angeles Lakers	Detroit Pistons	LAL	DET	LAL	DET	

Note: Since 2003, the favorite (of the two finalists) has only won the finals 7/19 times. Interesting.