In [ ]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
import os
import random
import shutil
import plotly
os.chdir('C:\\Users\\Travis\\OneDrive\\Data Science\\Personal_Projects\\Sports\\NBA
import datetime
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
from scipy.stats import norm
import winsound
import warnings
warnings.filterwarnings('ignore')
from selenium.common.exceptions import WebDriverException
```

In [ ]:

```
url = 'https://www.oddsshark.com/nba/odds'

driver = webdriver.Chrome()
driver.get(url)

# Get the html of the page
soup = BeautifulSoup(driver.page_source, 'html.parser')

data = soup.find(class_= 'block block-oddsshark-data-blocks block-odds-block')
single_games = data.find_all(class_ = 'odds--group__event-container basketball')
big_odds_df = pd.DataFrame()
for game in single_games:
    # get teams
    teams = game.find(class_ = 'odds--group__event-participants').text.strip()
    teams = teams.replace('vs', '')
    teams = teams.replace(' ET', '')
    teams = teams.replace('  ', ' ')
    # get team 1
    # find index of first number
    index = 0
    for char in teams:
        if char.isdigit():
            break
        index += 1
    team1 = teams[:index].strip()
    # get time... from index of first number through "pm"
    pm = teams.find('pm')
    time = teams[index:pm]
    # get team 2
    team2 = teams[pm+2:].strip()
    # get odds
    spreads = game.find_all('div', 'odds-spread')

    odds_df = pd.DataFrame()

    n = 1
    for spread in spreads:
        spread = spread.text.replace('- ', '').replace(' - ', '').strip()[:-2][:-4]
        # opening spread
        if n == 1:
            opening_spread_t1 = spread
            n += 1
        elif n == 2:
            opening_spread_t2 = spread
            n += 1
        elif n == 3:
            current_spread_t1 = spread
            n += 1
        elif n == 4:
            current_spread_t2 = spread
            n += 1

    moneylines = game.find_all('div', 'odds-moneyline above-tablet-only')
    n = 1
    for moneyline in moneylines:
        moneyline = moneyline.text.replace('- ', '').replace(' - ', '').strip()
```

```
        # opening moneyline
        if n == 1:
            opening_moneyline_t1 = moneyline
            n += 1
        elif n == 2:
            opening_moneyline_t2 = moneyline
            n += 1
        elif n == 3:
            current_moneyline_t1 = moneyline
            n += 1
        elif n == 4:
            current_moneyline_t2 = moneyline
            n += 1

            # add to dataframe
            odds_df = odds_df.append({'team1': team1, 'team2': team2, 'time': time,

    big_odds_df = big_odds_df.append(odds_df, ignore_index=True)
big_odds_df
```

Out[ ]:

	team1	team2	time	opening_spread_t1	opening_spread_t2	current_spread_t1	currer
0	Detroit	Philadelphia	7:00	+9	-9	+13.5	
1	Oklahoma City	Miami	7:30	+7	-7	+2	
2	Charlotte	Toronto	7:30	+7	-7	+7.5	
3	Cleveland	Utah	9:00	-3.5	+3.5	-2.5	
4	Phoenix	Golden State	10:00	+4	-4	+12	
5	Orlando	Portland	10:00	+6.5	-6.5	+8	
6	Chicago	Washington	7:00	+1	-1		
7	Minnesota	Detroit	7:00				
8	New Orleans	Boston	7:30	+8.5	-8.5	+9	
9	Indiana	New York	7:30	+5	-5		
10	Milwaukee	Atlanta	7:30	-1	+1	Ev	
11	San Antonio	Memphis	8:00				
12	Phoenix	Denver	10:00				

In [ ]:

```
# save to csv? No?
```