Player Contracts and Transactions (Not Used in 3.0)

The purpose of this notebook is to scrape and organize all player contract files.

```python
#Load
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 sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import BaggingClassifier, RandomForestClassifier
from matplotlib.offsetbox import OffsetImage, AnnotationBbox
from selenium import webdriver
from selenium.webdriver.common.keys import Keys
from selenium.webdriver.common.by import By
from bs4 import BeautifulSoup
import time
import requests      # to get images
import shutil        # to save files locally
import datetime
from scipy.stats import norm
from sklearn.preprocessing import OneHotEncoder, OrdinalEncoder
import os
```

```python
path = 'C:\\Users\\tmcro\\OneDrive\\0_NBA_Statistitcs_2021\\NBA Current Studies\\Fl
p = os.listdir(path)
pf = pd.DataFrame(p)
pf
```

Out[ ]:

	0
0	Aaron Brooks.xlsx
1	Aaron Gordon.xlsx
2	Aaron Harrison.xlsx
3	Aaron Henry.xlsx
4	Aaron Holiday.xlsx
...	...
995	Zion Williamson.xlsx
996	Zylan Cheatham.xlsx
997	Álex Abrines.xlsx
998	Ángel Delgado.xlsx
999	Ömer Aşık.xlsx

1000 rows × 1 columns

```
In [ ]: path = 'C:\\Users\\tmcro\\OneDrive\\0_NBA_Statistitcs_2021\\NBA Current Studies\\Fl
        p = os.listdir(path)
        pf2 = pd.DataFrame(p)
        pf2
```

Out[ ]:

	0
0	Aaron Brooks.xlsx
1	Aaron Gordon.xlsx
2	Aaron Harrison.xlsx
3	Aaron Henry.xlsx
4	Aaron Holiday.xlsx
...	...
521	Josh Richardson.xlsx
522	Josh Smith.xlsx
523	Joshua Primo.xlsx
524	José Calderón.xlsx
525	Jrue Holiday.xlsx

526 rows × 1 columns

```
In [ ]: cont = pd.read_excel('player_contracts\\Aaron Gordon.xlsx')
        cont
```

Out[ ]:	**Unnamed: 0**	**Aaron Gordon**
0	0	June 26, 2014: Drafted by the Orlando Magic in...
1	1	July 2, 2014: Signed a multi-year contract wit...
2	2	July 6, 2018: Signed as a free agent with the ...
3	3	March 25, 2021: Traded by the Orlando Magic wi...
4	4	September 27, 2021: Signed a contract extensio...

```
In [ ]: nmm = cont.columns[1]
```

```
In [ ]: cont = cont[nmm].str.split(':', expand=True)
```

```
In [ ]: cont.rename(columns = {0: 'Date', 1 : 'Transaction'}, inplace = True)
        cont
```

Out[ ]:	**Date**	**Transaction**
0	June 26, 2014	Drafted by the Orlando Magic in the 1st round...
1	July 2, 2014	Signed a multi-year contract with the Orlando...
2	July 6, 2018	Signed as a free agent with the Orlando Magic.
3	March 25, 2021	Traded by the Orlando Magic with Gary Clark t...
4	September 27, 2021	Signed a contract extension with the Denver N...

```
In [ ]: cont['Player'] = nmm
```

```
In [ ]: cont.to_excel('player_contracts_outputs\\' + str(nmm) + '.xlsx')
```

```
In [ ]: for n in pf[0]:
            contract = pd.read_excel('player_contracts/' + str(n))
            nm = contract.columns[1]
            contract = contract[nm].str.split(':', expand=True)
            contract.rename(columns = {0: 'Date', 1 : 'Transaction'}, inplace = True)
            contract['Player'] = nm
            contract.to_excel('player_contracts_outputs\\' + str(n))
```

```
In [ ]: path = 'C:\\Users\\tmcro\\OneDrive\\0_NBA_Statistitcs_2021\\NBA Current Studies\\Fl
        p = os.listdir(path)
        pf = pd.DataFrame(p)
        pf
```

	0
0	Aaron Brooks.xlsx
1	Aaron Gordon.xlsx
2	Aaron Harrison.xlsx
3	Aaron Henry.xlsx
4	Aaron Holiday.xlsx
...	...
995	Zion Williamson.xlsx
996	Zylan Cheatham.xlsx
997	Álex Abrines.xlsx
998	Ángel Delgado.xlsx
999	Ömer Aşık.xlsx

1000 rows × 1 columns

```python
# append the data
appended_contracts = []

for file in pf[0]:
    data = pd.read_excel('player_contracts_outputs\\' + str(file))
    dat = pd.DataFrame(data)
    appended_contracts.append(dat)

appended_contracts = pd.concat(appended_contracts)
appended_contracts
```

	Unnamed: 0	Date	Transaction	Player
0	0	June 28, 2007	Drafted by the Houston Rockets in the 1st rou...	Aaron Brooks
1	1	December 6, 2007	Assigned to the of the G-League.	Aaron Brooks
2	2	December 14, 2007	Recalled from the of the G-League.	Aaron Brooks
3	3	February 24, 2011	Traded by the Houston Rockets to the Phoenix ...	Aaron Brooks
4	4	July 16, 2012	Signed as a free agent with the Sacramento Ki...	Aaron Brooks
...	...	...	...	...
3	3	July 24, 2012	Signed as a free agent with the Houston Rockets.	Ömer Aşık
4	4	July 15, 2014	As part of a 3-team trade, traded by the Hous...	Ömer Aşık
5	5	July 9, 2015	Signed a multi-year contract with the New Orl...	Ömer Aşık
6	6	February 1, 2018	Traded by the New Orleans Pelicans with Tony ...	Ömer Aşık
7	7	October 21, 2018	Waived by the Chicago Bulls.	Ömer Aşık

7741 rows × 4 columns

```python
appended_contracts = appended_contracts.drop(columns = 'Unnamed: 0')
```

```python
appended_contracts
```

	Date	Transaction	Player
0	June 28, 2007	Drafted by the Houston Rockets in the 1st rou...	Aaron Brooks
1	December 6, 2007	Assigned to the of the G-League.	Aaron Brooks
2	December 14, 2007	Recalled from the of the G-League.	Aaron Brooks
3	February 24, 2011	Traded by the Houston Rockets to the Phoenix ...	Aaron Brooks
4	July 16, 2012	Signed as a free agent with the Sacramento Ki...	Aaron Brooks
...	...	...	...
3	July 24, 2012	Signed as a free agent with the Houston Rockets.	Ömer Aşık
4	July 15, 2014	As part of a 3-team trade, traded by the Hous...	Ömer Aşık
5	July 9, 2015	Signed a multi-year contract with the New Orl...	Ömer Aşık
6	February 1, 2018	Traded by the New Orleans Pelicans with Tony ...	Ömer Aşık
7	October 21, 2018	Waived by the Chicago Bulls.	Ömer Aşık

7741 rows × 3 columns

In [ ]:
```python
appended_contracts.to_excel('player_contracts_outputs\\all_contracts_df.xlsx')
```