In [370…

```
# ** CMSC 320 UMD**
# *Project 2*
## **Author : Franklin Nkokam Ngongang**
### Due date : March 26, 2021
### Email : fnkokamn@terpmail.umd.edu
### UID : 117548327
```

In [400…

```
import sqlite3
import pandas
from matplotlib import pyplot as plt
import os
```

In [401…

```
sqlite_file = 'lahman2014.sqlite'
backup_database_file = 'blabladb.sqlite'
conn = sqlite3.connect(sqlite_file)
```

Pre-required data. This part was added last. Due to running several queries. we found out that we will need access to a list of team ids for this project. We will therefore load the team ids into a list, and reuse that list everyt time.

In [402…

```
select_team_details_query = " SELECT DISTINCT T.name , T.teamId from Teams as
team_details_result = pandas.read_sql(select_team_details_query, conn)
```

In [403…

```
team_details_result.columns
```

Out[403…

```
Index(['name', 'teamID'], dtype='object')
```

In [404…

```
team_names = [] # saving the team names in a list
team_ids = []
for index, row in team_details_result.iterrows():
    team_id = str(row['teamID']) # the team ID should be in string format
    team_name = str(row['name']) # The name should also be in a string format
    if not(bool(str(team_id))) == False:
        team_ids.append(team_id)
        team_names.append(team_name)
```

Part 1 : Wrangling

We need to compute the total payroll and winning percentage for each team. The total payroll will be a sum of salaries in the salaries table, grouping by the teamId or grouping by yearID, depending if we want to view per team or per year. The winning percentage is as defined the number of wins divided by number of games played(not games played at home), all multiplied by 100. Now we already have the list of team ids saved in a list. We can find that total payroll and winning percentage for each team by looping through that list of eam ids

In [405…

```
working_query = "SELECT Teams.name as `TEAM NAME`, Salaries.teamID as `TEAM I
        "printf('%5.2f %',((SUM(Teams.W)*1.0)/(SUM(Teams.G)*1.0))*100.0) AS `
        "printf('$%,d',SUM(Salaries.salary)) AS `TOTAL PAYROLL` " + \
        "FROM Salaries " + \
        "LEFT JOIN Teams ON Salaries.teamID = Teams.teamID " + \
        "GROUP BY `TEAM ID` " + \
        "ORDER BY `TEAM NAME`,`TEAM ID` "
query_result = pandas.read_sql(working_query, conn)
#query_result
```

In [406…

```
query_result
```

Out[406…

	TEAM NAME	TEAM ID	WINNING PERCENTAGE	TOTAL PAYROLL
0	None	NYM	0.00 %	$54,806,990
1	None	SFG	0.00 %	$143,510,167
2	Anaheim Angels	ANA	51.23 %	$3,744,735,784
3	Arizona Diamondbacks	ARI	49.20 %	$20,569,578,876
4	Atlanta Braves	ATL	51.76 %	$92,264,392,416
5	Baltimore Orioles	BAL	51.23 %	$99,442,202,318
6	Boston Red Sox	BOS	51.43 %	$277,327,906,590
7	California Angels	CAL	48.24 %	$8,703,325,760
8	Chicago White Sox	CHA	50.18 %	$193,784,626,302
9	Chicago White Stockings	CHN	50.70 %	$258,516,830,785
10	Cincinnati Reds	CIN	50.14 %	$179,346,918,500
11	Cleveland Indians	CLE	50.66 %	$160,171,838,592
12	Colorado Rockies	COL	46.86 %	$28,466,034,058
13	Detroit Tigers	DET	50.56 %	$201,916,536,318
14	Florida Marlins	FLO	47.67 %	$12,825,304,095
15	Houston Astros	HOU	48.69 %	$74,628,937,321
16	Kansas City Royals	KCA	48.23 %	$54,163,480,528
17	Los Angeles Angels	LAA	53.55 %	$16,585,403,604
18	Los Angeles Dodgers	LAN	53.56 %	$127,552,702,071
19	Miami Marlins	MIA	42.80 %	$580,550,400
20	Milwaukee Brewers	MIL	47.26 %	$17,818,737,901
21	Milwaukee Brewers	ML4	48.47 %	$6,542,082,512
22	Minnesota Twins	MIN	49.60 %	$68,592,804,696
23	Montreal Expos	MON	48.32 %	$14,695,335,396
24	New York Mets	NYN	47.76 %	$104,184,073,243
25	New York Yankees	NYA	56.66 %	$367,709,365,408

	TEAM NAME	TEAM ID	WINNING PERCENTAGE	TOTAL PAYROLL
26	Oakland Athletics	OAK	52.20 %	$55,768,059,291
27	Philadelphia Quakers	PHI	47.02 %	$261,670,965,600
28	Pittsburgh Pirates	PIT	50.19 %	$127,302,043,904
29	San Diego Padres	SDN	46.37 %	$54,164,998,252
30	San Francisco Giants	SFN	51.89 %	$96,685,846,497
31	Seattle Mariners	SEA	46.78 %	$63,579,450,378
32	St. Louis Cardinals	SLN	50.76 %	$218,402,653,836
33	Tampa Bay Devil Rays	TBA	46.22 %	$13,057,884,875
34	Texas Rangers	TEX	49.07 %	$71,813,891,414
35	Toronto Blue Jays	TOR	49.47 %	$59,693,263,414

Explanation of the working_query

The first line selects the tteam id found in the Salaries table. The next line calculates the percentage wins. Here we multiply the sums by 1.0 to make sqlite convert it to floating points number, else the division operator will return zero for every sum of wins over sum of total games played, since the division operator returns zero if an numerator is divided by an denominator when the numerator is less than the denominator. The next line calculates the total salary paid to players for that particular team for a specific year. It uses the printf function to format the data to 5 spots(including the dot character) : two decimal places, and two places before the decimal like the C printf function would do. It's just a semi replica of the c printf function. The next line tells sqlite from what tables to pull the data from, both from Salaries and Teams table. The next line specifies the mathematical joining clause. In this instancd the left join clause.

Dealing with the empty values

Let's examine the query above, more precisely the left join clause : SELECT Teams.name as `TEAM NAME` , Salaries.teamID as `TEAM ID` , printf('%5.2f %',((SUM(Teams.W)*1.0)/(SUM(Teams.G)*1.0))*100.0) AS `WINNING PERCENTAGE` , printf('$%,d',SUM(Salaries.salary)) AS `TOTAL PAYROLL`
FROM Salaries
LEFT JOIN Teams ON Salaries.teamID = teams.teamID
GROUP BY `TEAM ID`
ORDER BY `TEAM ID`

The left join line tells sqlite to pull out results from the Salaries tables, and the teams table, for each row such that the team id for that row in the Salaries table is being foreign-keyed by a team id from the Teams table. So what this means is that the we want the results to be in such a way that we only see results when the team id is found in the Salaries table. Now in our case, every row in the Salaries table has a team id that comes from the Teams table. Now some teams might not have any rows in the salaries table. So doing a left join on the Salaries table instead will bring out some null values. But doing a left join on the Teams table will not bring out any nulls unless

there are team ids in the Salaries table that have no teams in the Teams table, which does not make sense.
To find out where the empty sets come from, we decided to search for any team in the Salaries table that does not have a row in the Teams table. Thinking about that, it's not supposed to be possible. But in this case it is. Let's have a look at that. First let's select team ids in the salary table that are not present in the teams table

In [407…

```
surprise_query = "SELECT DISTINCT teamID FROM Salaries " + \
                 "WHERE teamID NOT IN (" + \
                 "SELECT DISTINCT teamID FROM Teams"  + \
                 " ) "
surprise_result = pandas.read_sql(surprise_query, conn)
surprise_result
```

Out[407…

	teamID
0	NYM
1	SFG

So there are two teams rows in the Salaries table not found in the Teams table. Let's find their total payroll

In [408…

```
further_query = "SELECT teamID, " + \
        "printf('$%,d',SUM(salary)) AS `TOTAL PAYROLL` " + \
        "FROM Salaries " + \
        "WHERE (teamID = 'NYM' OR teamID ='SFG') " + \
        "GROUP BY teamID " + \
        "ORDER BY teamID "
further_result = pandas.read_sql(further_query, conn)
further_result
```

Out[408…

	teamID	TOTAL PAYROLL
0	NYM	$54,806,990
1	SFG	$143,510,167

This results matches the two None teamId starting rows we have in the beginning of this part. An alternative will be to left join on Teams table instead.

In [409…

```
good_query = "SELECT Teams.name as `TEAM NAME`, Salaries.teamID as `TEAM ID`,
        "printf('%5.2f%',((SUM(Teams.W)*1.0)/(SUM(Teams.G)*1.0))*100.0) AS `W
        "printf('$%,d',SUM(Salaries.salary)) AS `TOTAL PAYROLL` " + \
        "FROM Teams " + \
        "LEFT JOIN Salaries ON Salaries.teamID = Teams.teamID " + \
        "GROUP BY `TEAM ID` " + \
        "ORDER BY `TEAM NAME`,`TEAM ID` "
```

In [410…

```
good_result = pandas.read_sql(good_query, conn)
good_result
```

Out[410…

	TEAM NAME	TEAM ID	WINNING PERCENTAGE	TOTAL PAYROLL
0	Anaheim Angels	ANA	51.23%	$3,744,735,784
1	Arizona Diamondbacks	ARI	49.20%	$20,569,578,876
2	Atlanta Braves	ATL	51.76%	$92,264,392,416
3	Baltimore Orioles	BAL	51.23%	$99,442,202,318
4	Boston Americans	BOS	51.43%	$277,327,906,590
5	Boston Red Stockings	None	48.06%	$0
6	California Angels	CAL	48.24%	$8,703,325,760
7	Chicago White Sox	CHA	50.18%	$193,784,626,302
8	Chicago White Stockings	CHN	50.70%	$258,516,830,785
9	Cincinnati Reds	CIN	50.14%	$179,346,918,500
10	Cleveland Blues	CLE	50.66%	$160,171,838,592
11	Colorado Rockies	COL	46.86%	$28,466,034,058
12	Detroit Tigers	DET	50.56%	$201,916,536,318
13	Florida Marlins	FLO	47.67%	$12,825,304,095
14	Houston Colt .45's	HOU	48.69%	$74,628,937,321
15	Kansas City Royals	KCA	48.23%	$54,163,480,528
16	Los Angeles Angels	LAA	53.55%	$16,585,403,604
17	Los Angeles Dodgers	LAN	53.56%	$127,552,702,071
18	Miami Marlins	MIA	42.80%	$580,550,400
19	Milwaukee Brewers	MIL	47.26%	$17,818,737,901
20	Milwaukee Brewers	ML4	48.47%	$6,542,082,512
21	Minnesota Twins	MIN	49.60%	$68,592,804,696
22	Montreal Expos	MON	48.32%	$14,695,335,396
23	New York Highlanders	NYA	56.66%	$367,709,365,408
24	New York Mets	NYN	47.76%	$104,184,073,243
25	Oakland Athletics	OAK	52.20%	$55,768,059,291
26	Philadelphia Quakers	PHI	47.02%	$261,670,965,600
27	Pittsburg Alleghenys	PIT	50.19%	$127,302,043,904
28	San Diego Padres	SDN	46.37%	$54,164,998,252
29	San Francisco Giants	SFN	51.89%	$96,685,846,497
30	Seattle Mariners	SEA	46.78%	$63,579,450,378
31	St. Louis Browns	SLN	50.76%	$218,402,653,836
32	Tampa Bay Devil Rays	TBA	46.22%	$13,057,884,875
33	Texas Rangers	TEX	49.07%	$71,813,891,414

	TEAM NAME	TEAM ID	WINNING PERCENTAGE	TOTAL PAYROLL
34	Toronto Blue Jays	TOR	49.47%	$59,693,263,414

With this query above, we get a result without None values.

Part 2 : Exploratory Data Analysis

Payroll Distribution

Distribution of payrolls conditioned on time between the years 1990 and 2014

For this section, we want to see how the payrolls distribute over time between 1990 and 2014. For that we will use a similar query to the query above.

In [411…

```
payroll_query =  \
        "SELECT Salaries.teamID as `TEAM ID`, " + \
        "Salaries.yearID AS `YEAR`, " + \
        "SUM(Salaries.salary) AS `TOTAL SALARY PAID` " + \
        "FROM Salaries " + \
        "WHERE Salaries.yearID >= '1990' AND " + \
        "Salaries.yearID <= '2014' " + \
        "GROUP BY `TEAM ID`, `YEAR` " + \
        "ORDER BY `TEAM ID`, `YEAR` "
payroll_result = pandas.read_sql(payroll_query, conn)
payroll_result.head()
```

Out[411…

	TEAM ID	YEAR	TOTAL SALARY PAID
0	ANA	1997	31135472.0
1	ANA	1998	41281000.0
2	ANA	1999	55388166.0
3	ANA	2000	51464167.0
4	ANA	2001	47535167.0

In [412…

```
payroll_result.tail()
```

Out[412…

	TEAM ID	YEAR	TOTAL SALARY PAID
725	WAS	2010	61400000.0
726	WAS	2011	63856928.0
727	WAS	2012	80855143.0
728	WAS	2013	113703270.0
729	WAS	2014	131983680.0

Above is what the results will look like. So we need to plot the results for all the years. We could do it as a query, then plot it. This will require manipulating the dataframe to pick the salaries paid per team for each year. That is tedious, when sql can do the job for us in case we have the team IDs in a list. So a better way is to have the list of teamIds and names in a python list, then make an sql query selection based on each value for that list, then plot the result for each team, since each value in the list will correspond to a unique team. Luckily at the begnining of this document, we selected the list of team ids and team names and stored them in a list.

In [413…

```
# saving the result of running the read_sql function in a list
result_frames_objects = []

for index in range(len(team_ids)):
    team_id = str(team_ids[index]) # the team ID should be in string format
    team_name = str(team_names[index]) # The name should also be in a string
    the_query = "SELECT SUM(salary) as `Total_Payroll`, " + \
                "AVG(salary) AS `Average_Payroll`, " + \
                "yearID as `Year` " + \
                "FROM Salaries " + \
                "WHERE teamID = '" + team_id + "' "  + \
                "AND yearID >= '1990' AND yearID <= '2014' " + \
                "GROUP BY yearID ORDER BY yearID "

    result_frames_objects.append(pandas.read_sql(the_query, conn))
```

Explaing the query above

We are selecting the total payrolls ad average payrolls for each year. That salary sum will be identified in the dataframe as Total_Payroll. The where clause identifies the team for which the selection should be made, and the yearId should be between 1990 and 2014,inclusive. The result will be saved in a list. We will iterate over that list to plot it.

In [414…

```
len(result_frames_objects) # checking if the list is not empty
```

Out[414… 185

In [415…

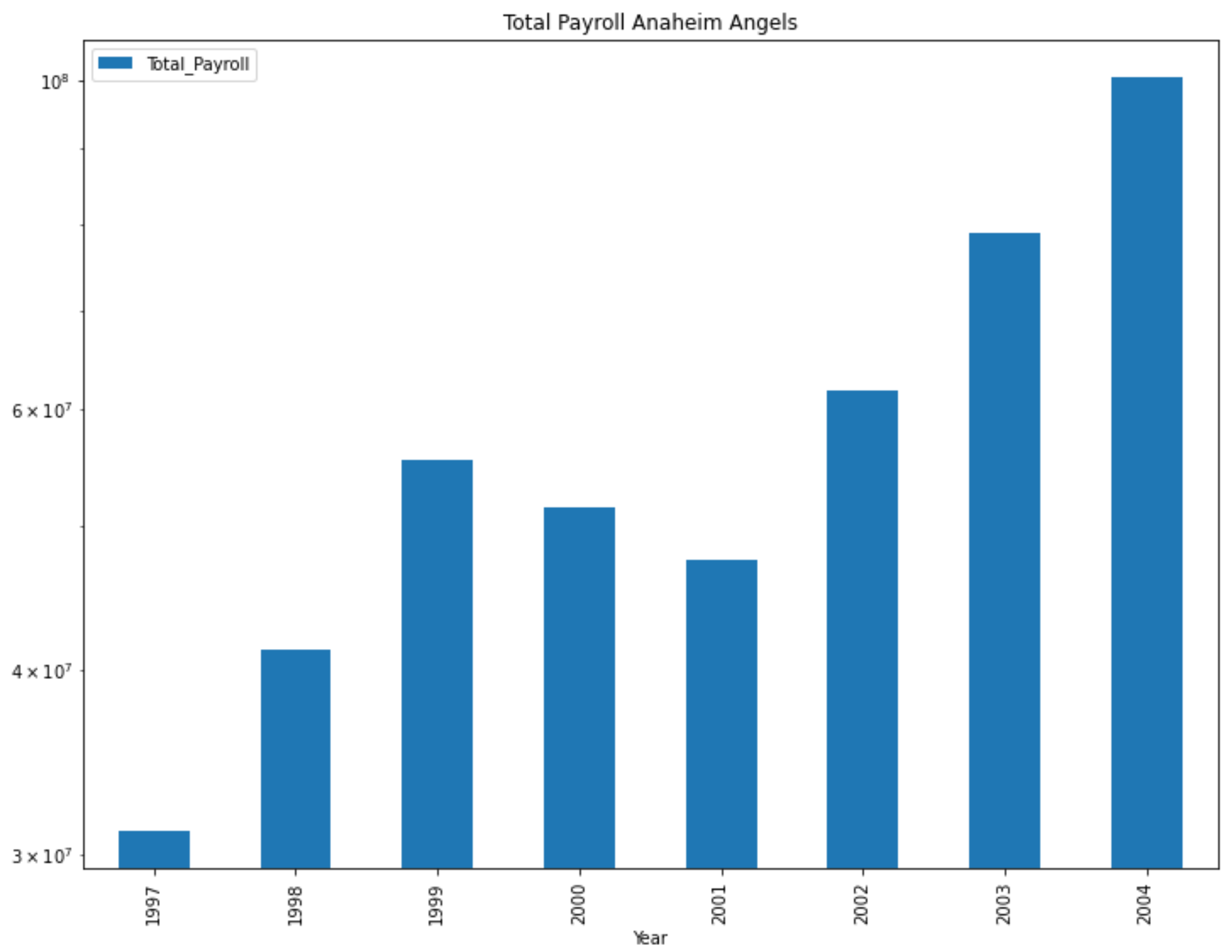
```
len(result_frames_objects) # checking if the list is not empty
```

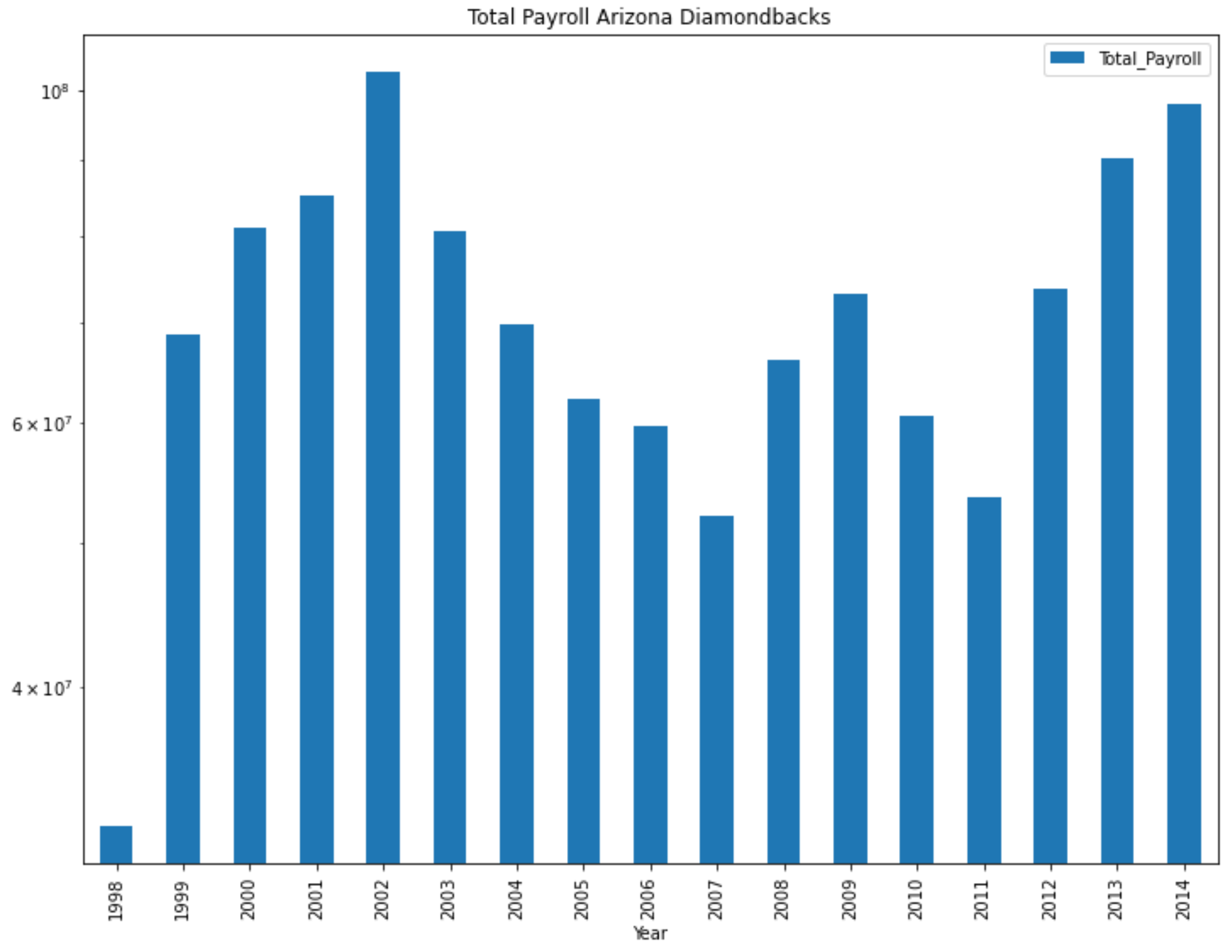
Out[415… 185

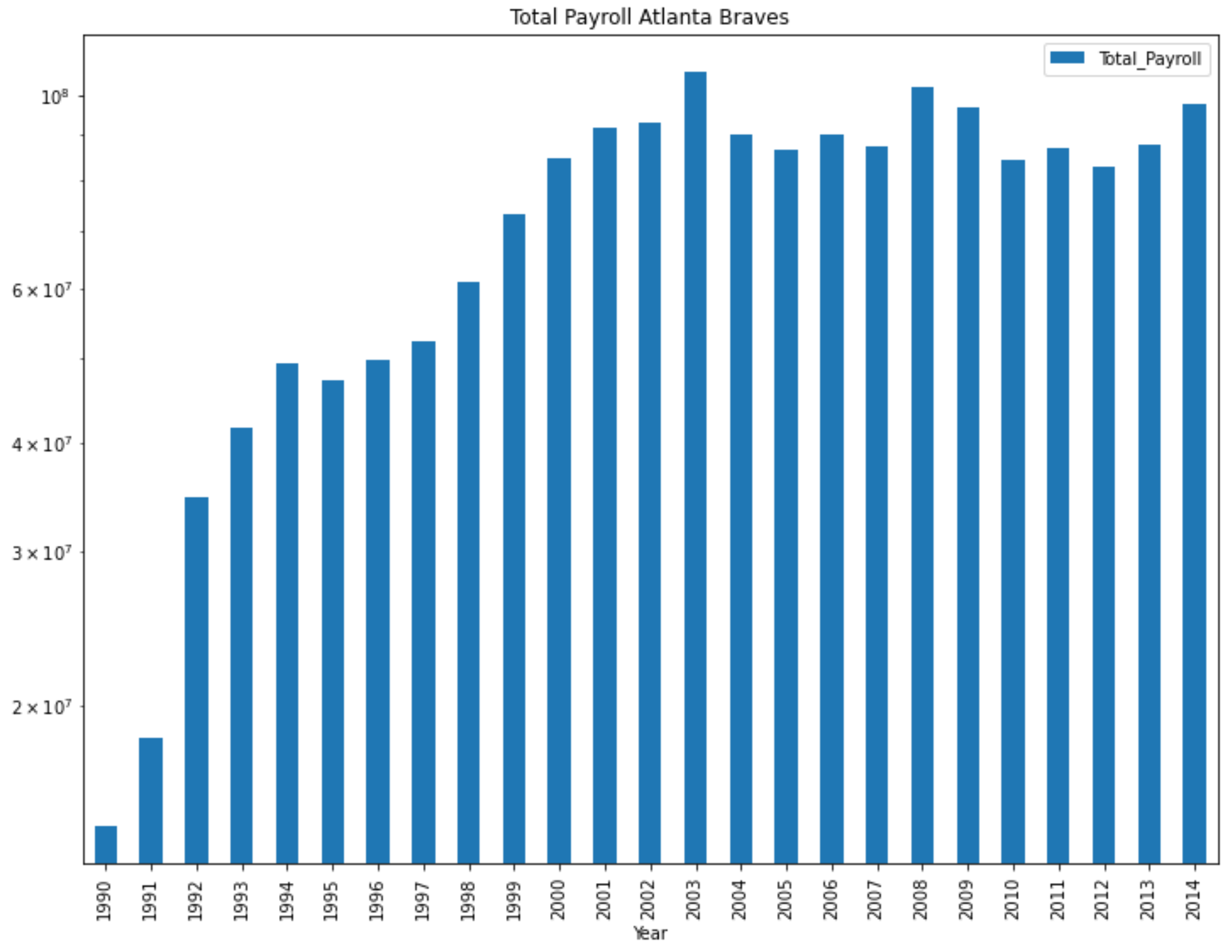
Now let's have fun plotting the results we just got. Plotting these results generate warnings because more than 20 figures are generated. The solution is to save the plots as images and just show the images after closing the opened figures

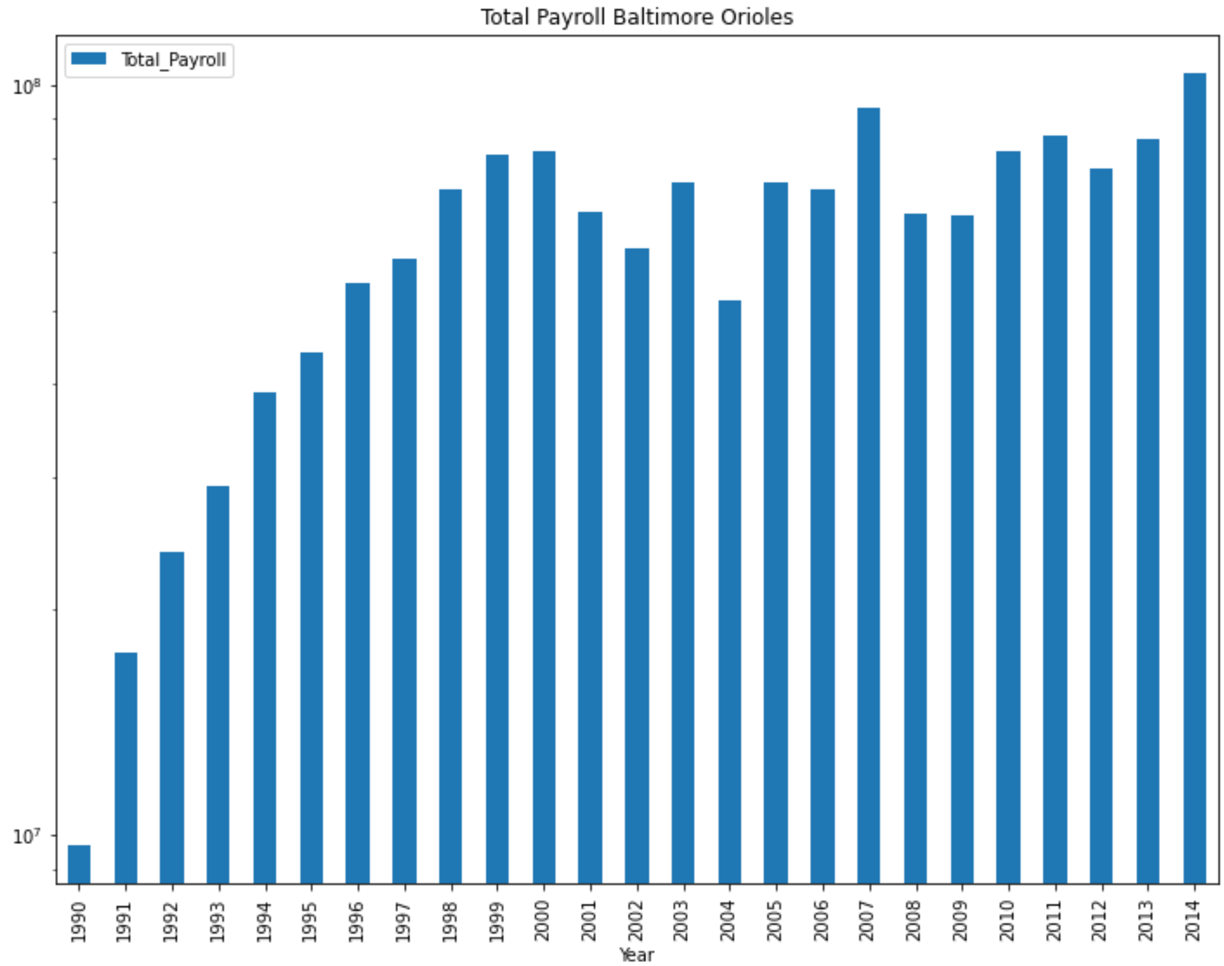
In [416…

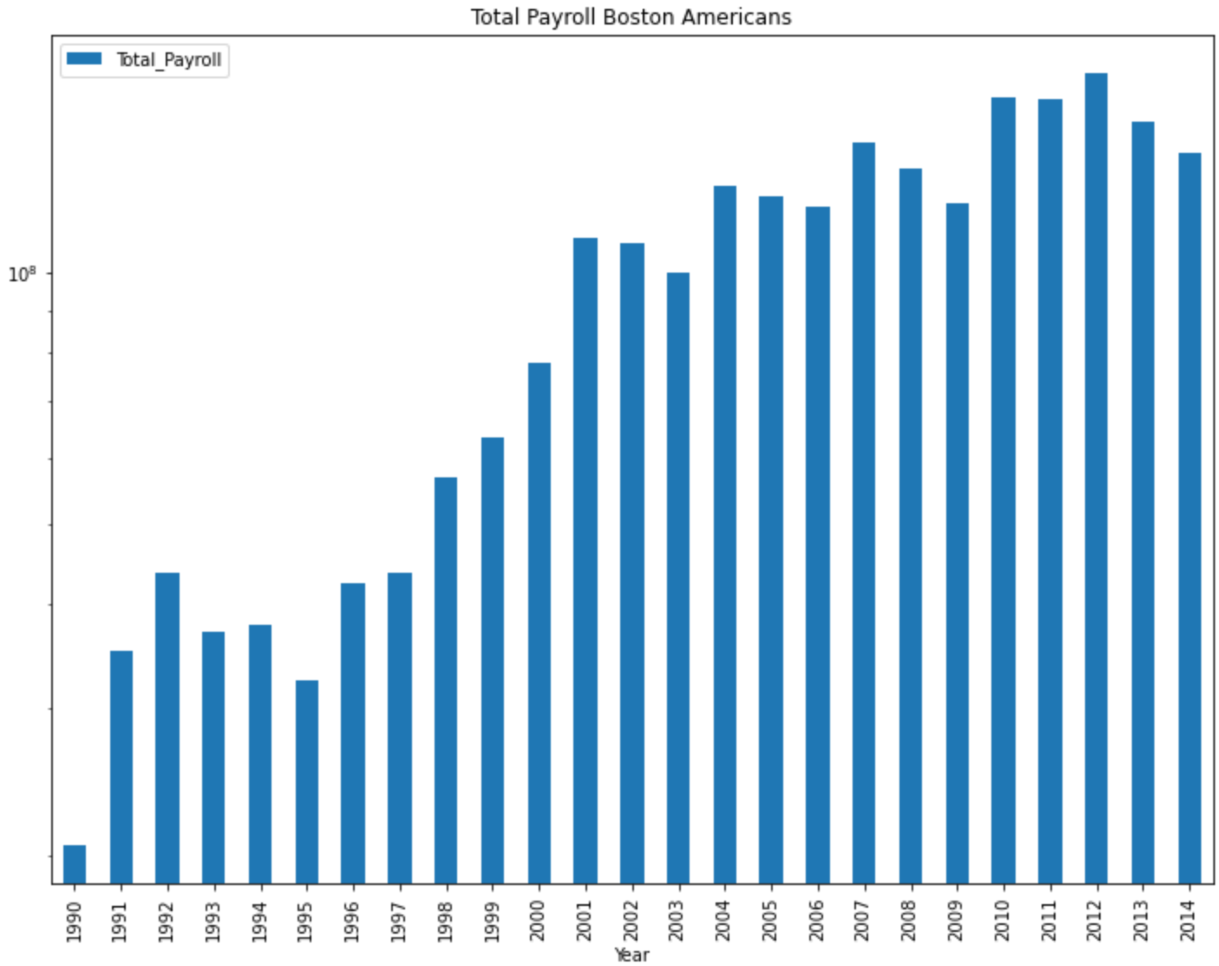
```
num_results = len(result_frames_objects)
#prevents warning after 20 figures shown
plt.rcParams.update({'figure.max_open_warning':0})
figsize = (12,9)
for i in range(num_results):
    dfr = result_frames_objects[i]
    team_name = team_names[i]
    team_id = team_ids[i]
    #we only plot if the data frame in
    #question is not empty
    if not dfr.empty:
        dfr.plot(x='Year', y='Total_Payroll', logy=True, \
                  title = 'Total Payroll ' + team_name, kind='bar',figsize=fi

        #fig = plt.figure((i+1))
        #ax = fig.add_subplot(111)
        #image_name = 'figs/graph_{}_{}.png'.format(team_id, i)

        #fig.savefig(image_name)
        #plt.close(fig)
```

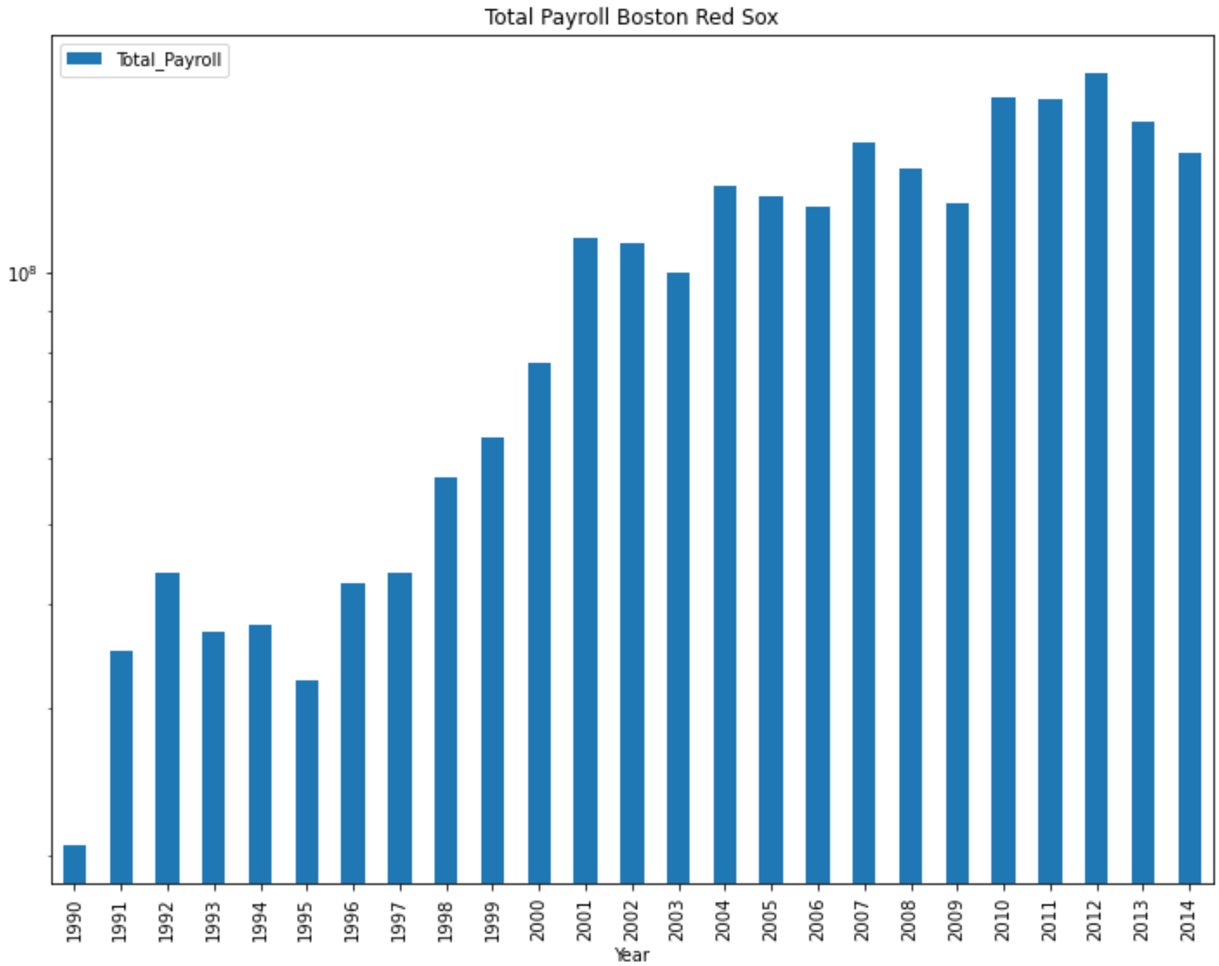


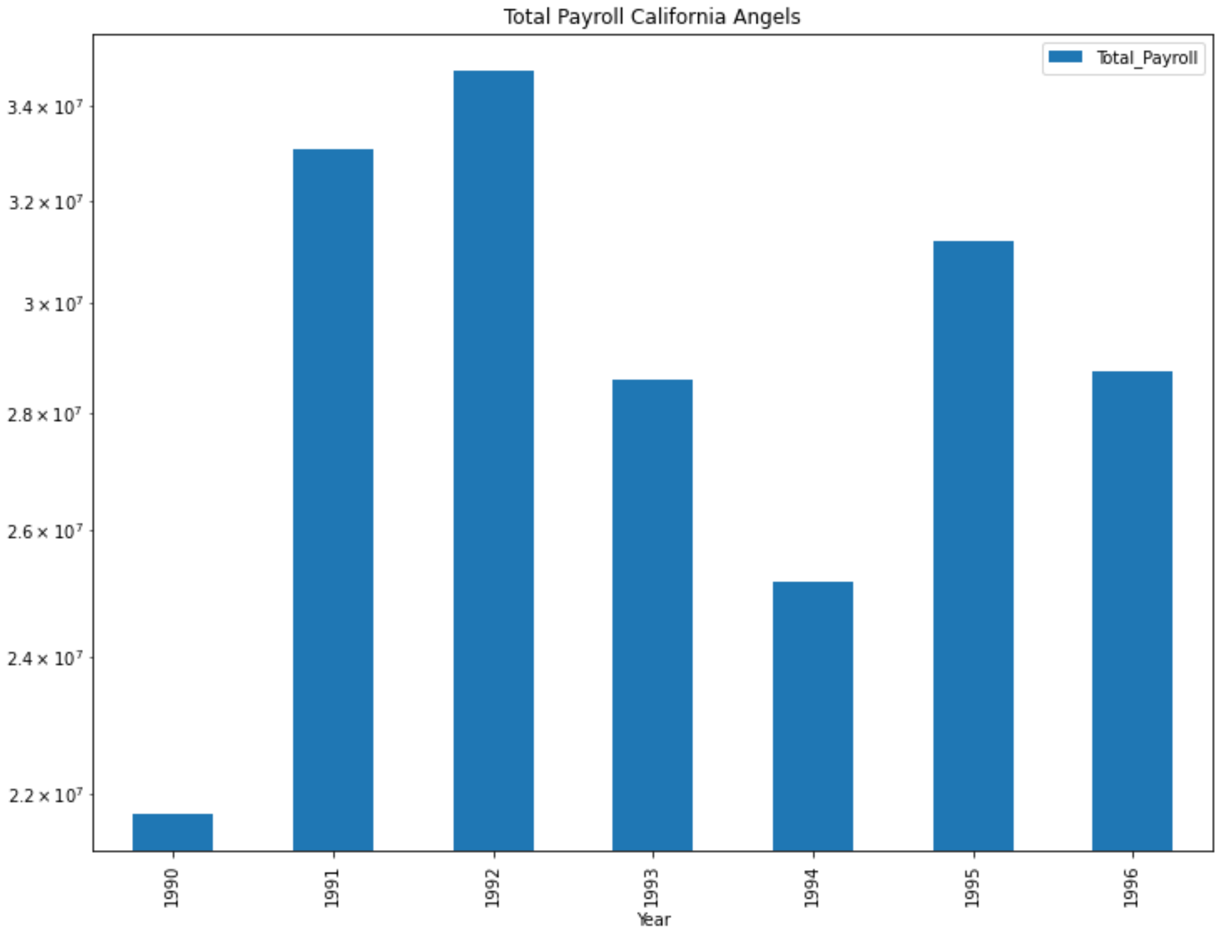

Total Payroll Arizona Diamondbacks
Total_Payroll
10^8
6×10^7
4×10^7
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

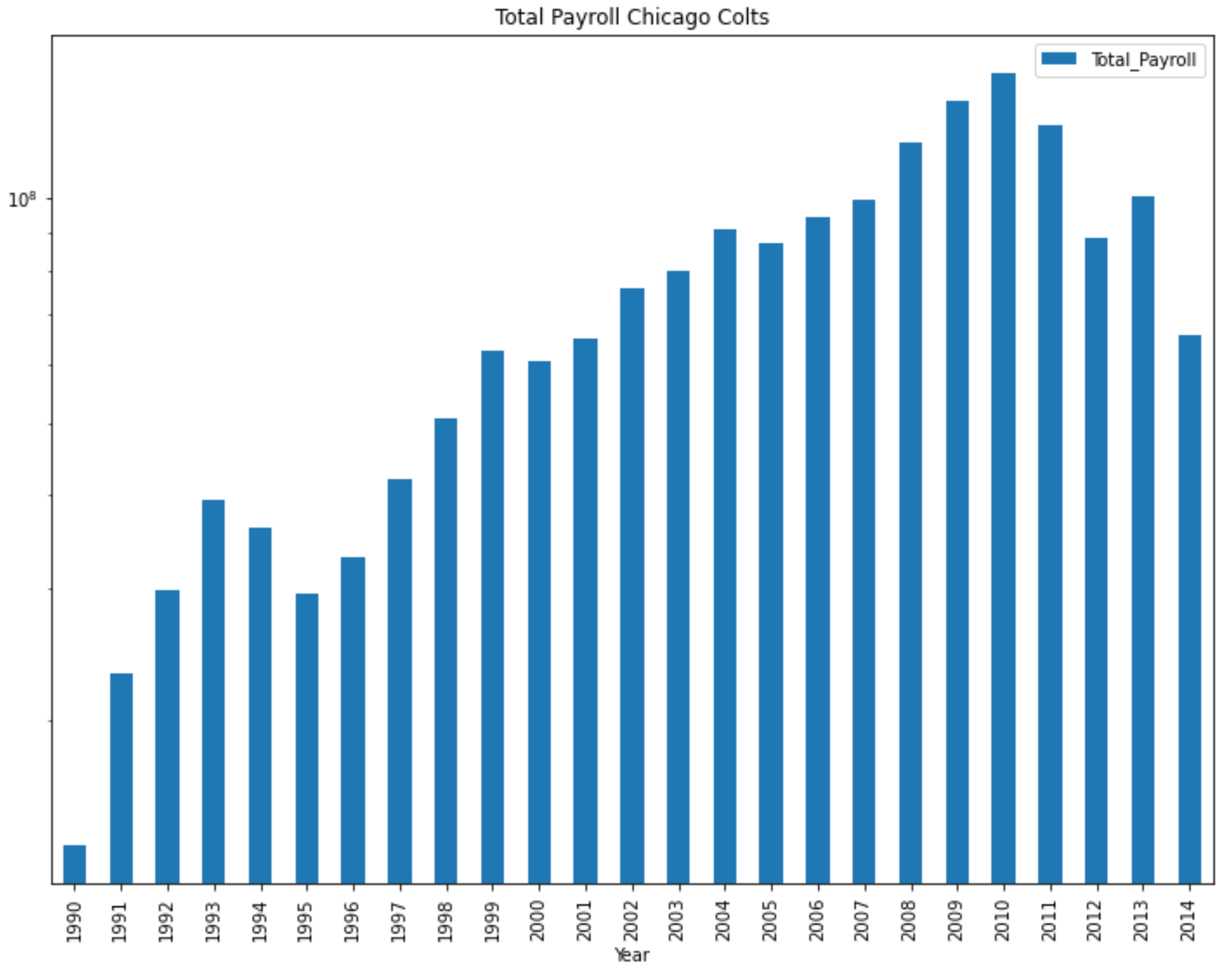

Total Payroll Atlanta Braves
Total_Payroll
10^8
6×10^7
4×10^7
3×10^7
2×10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

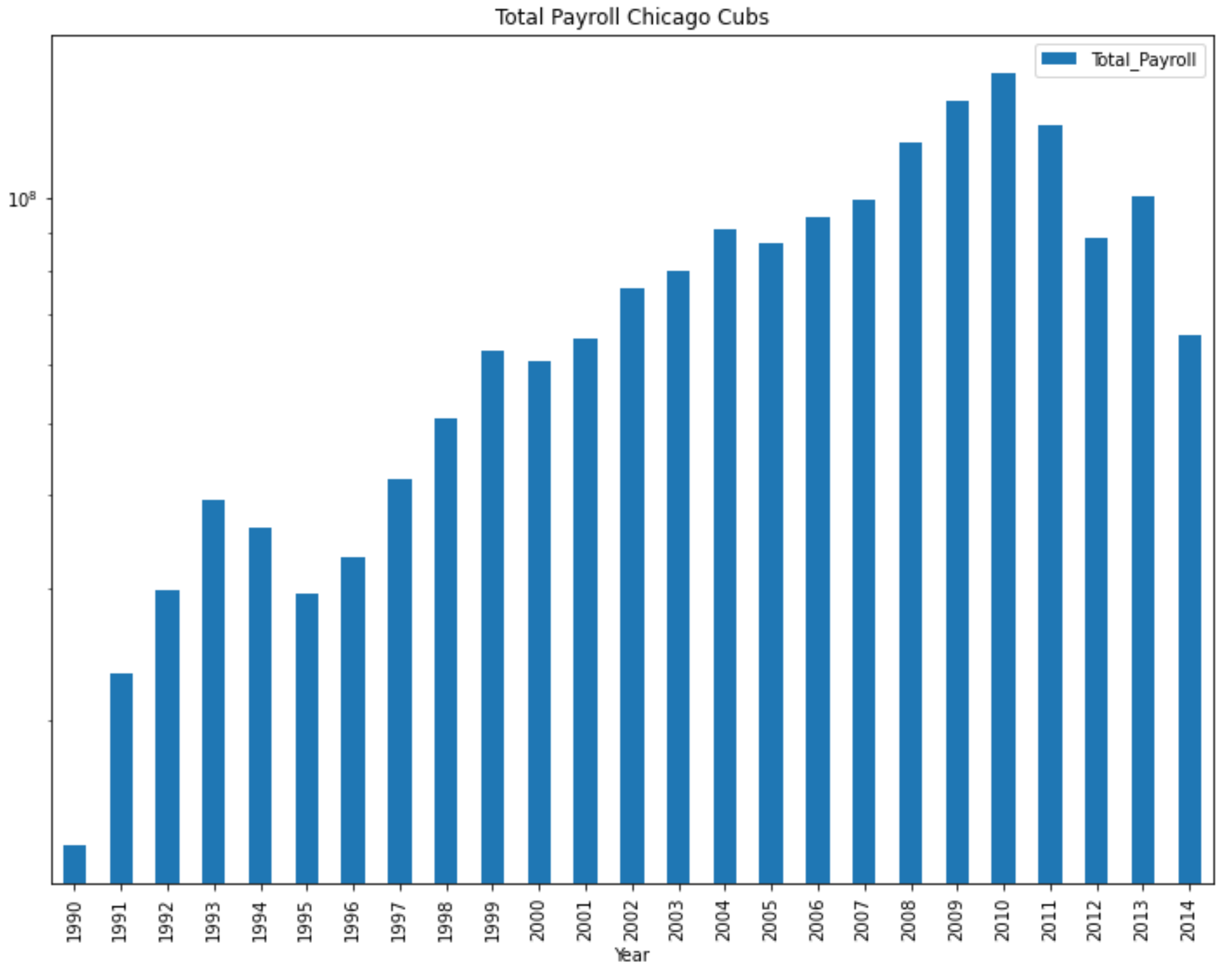

Total Payroll Baltimore Orioles
Total_Payroll
10^8
10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

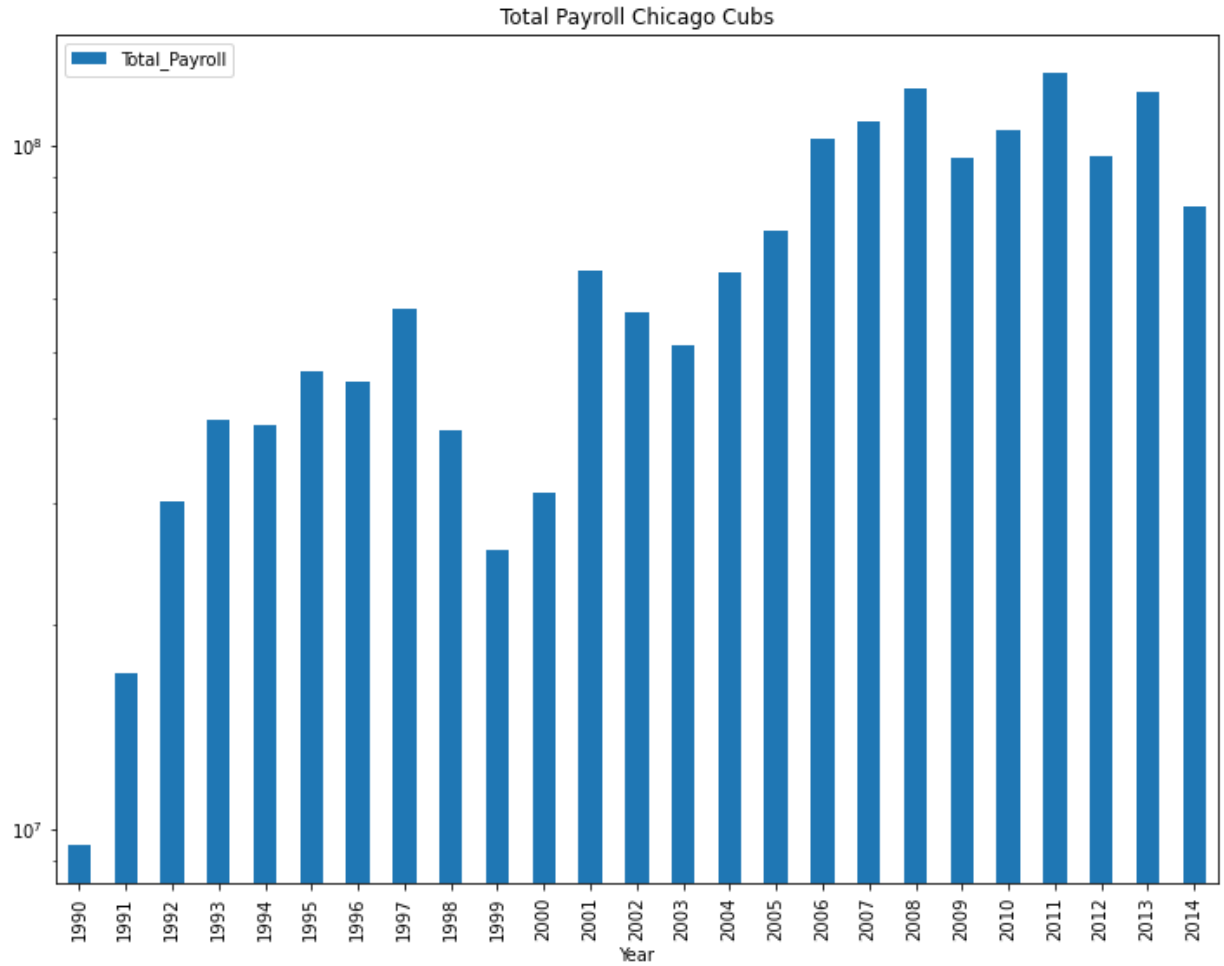

Total Payroll Boston Americans
Total_Payroll
10^8
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

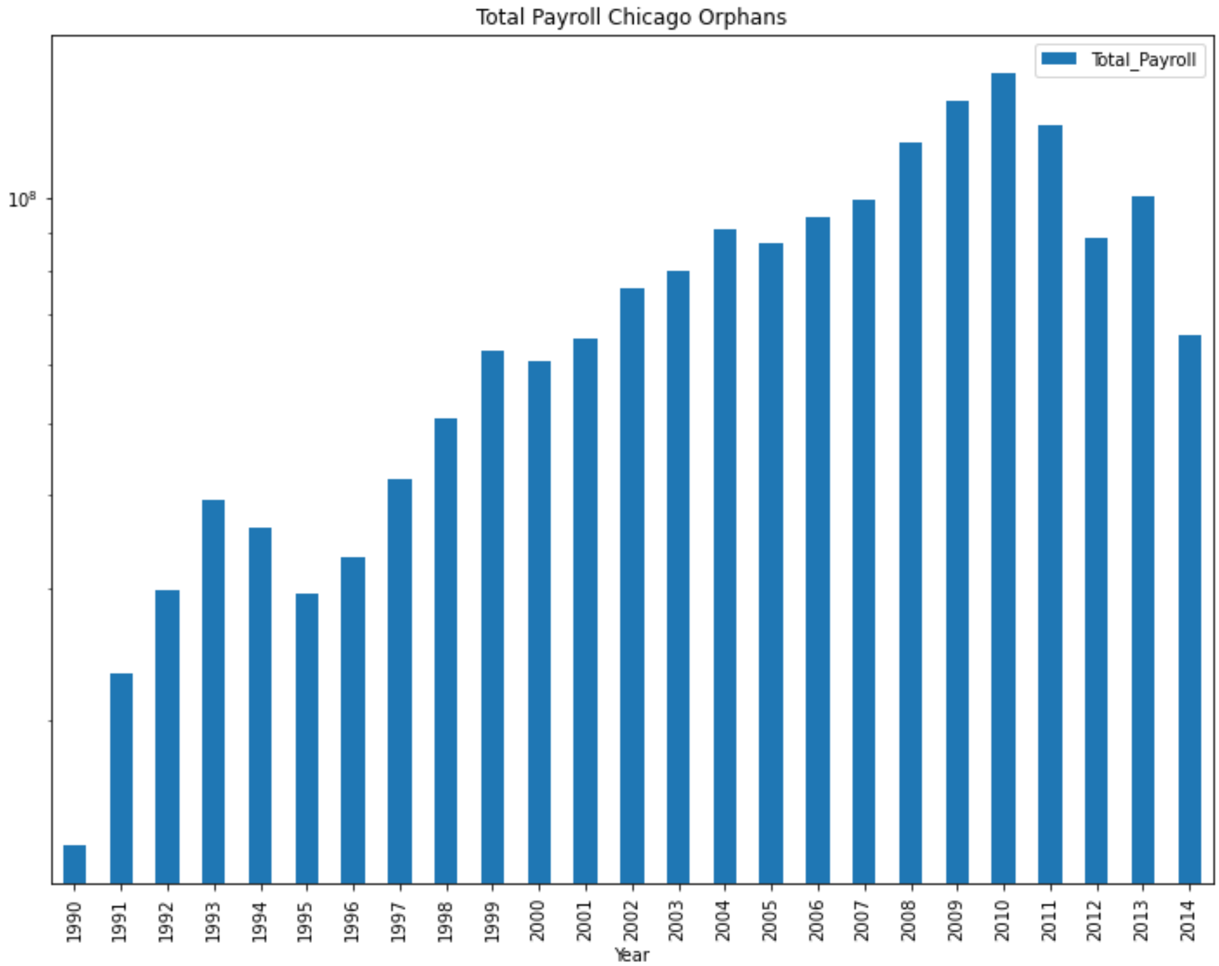

Total Payroll Boston Red Sox
Total_Payroll
10^8
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

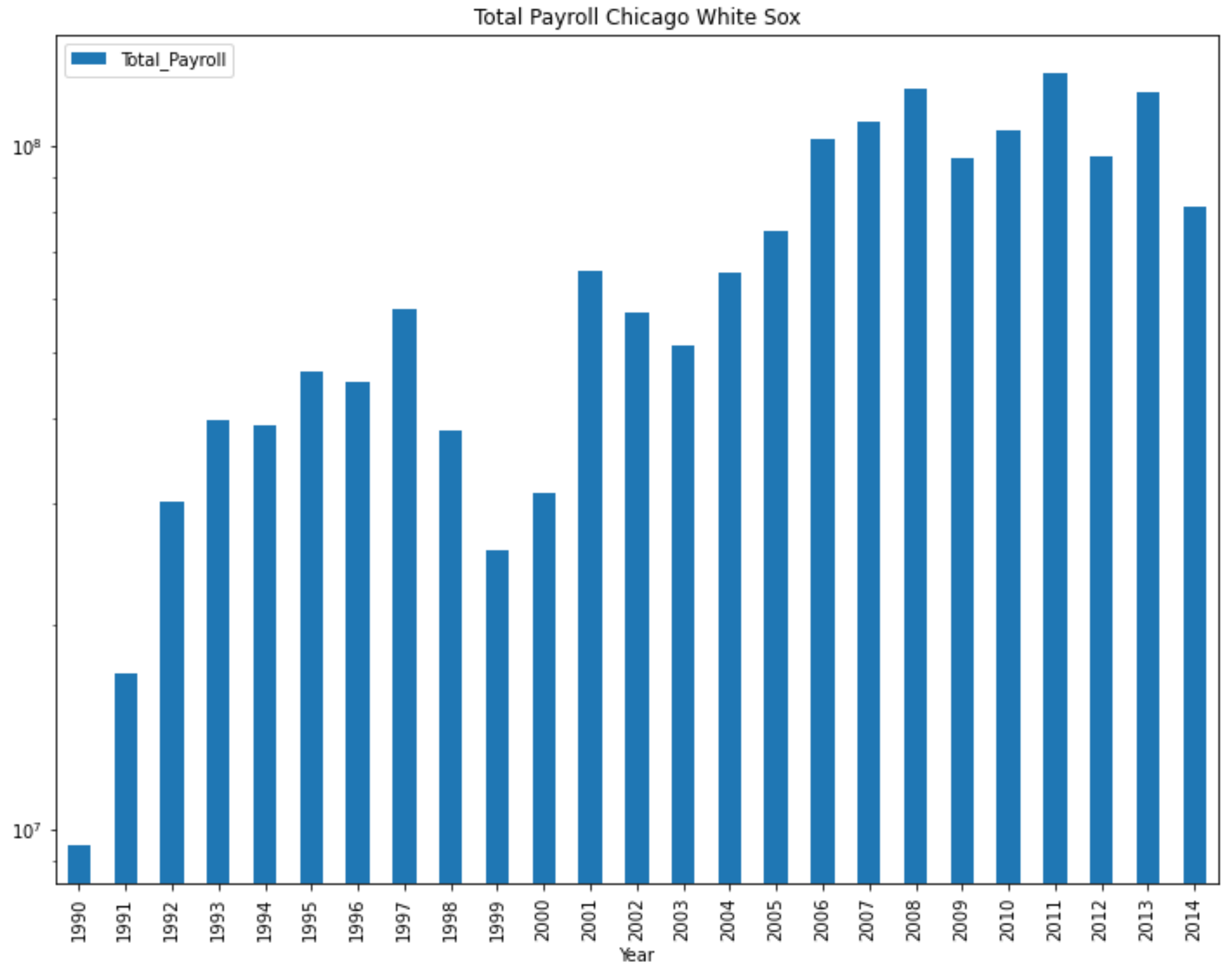

Total Payroll California Angels
Total_Payroll
3.4 × 10⁷
3.2 × 10⁷
3 × 10⁷
2.8 × 10⁷
2.6 × 10⁷
2.4 × 10⁷
2.2 × 10⁷
1990
1991
1992
1993
1994
1995
1996
Year

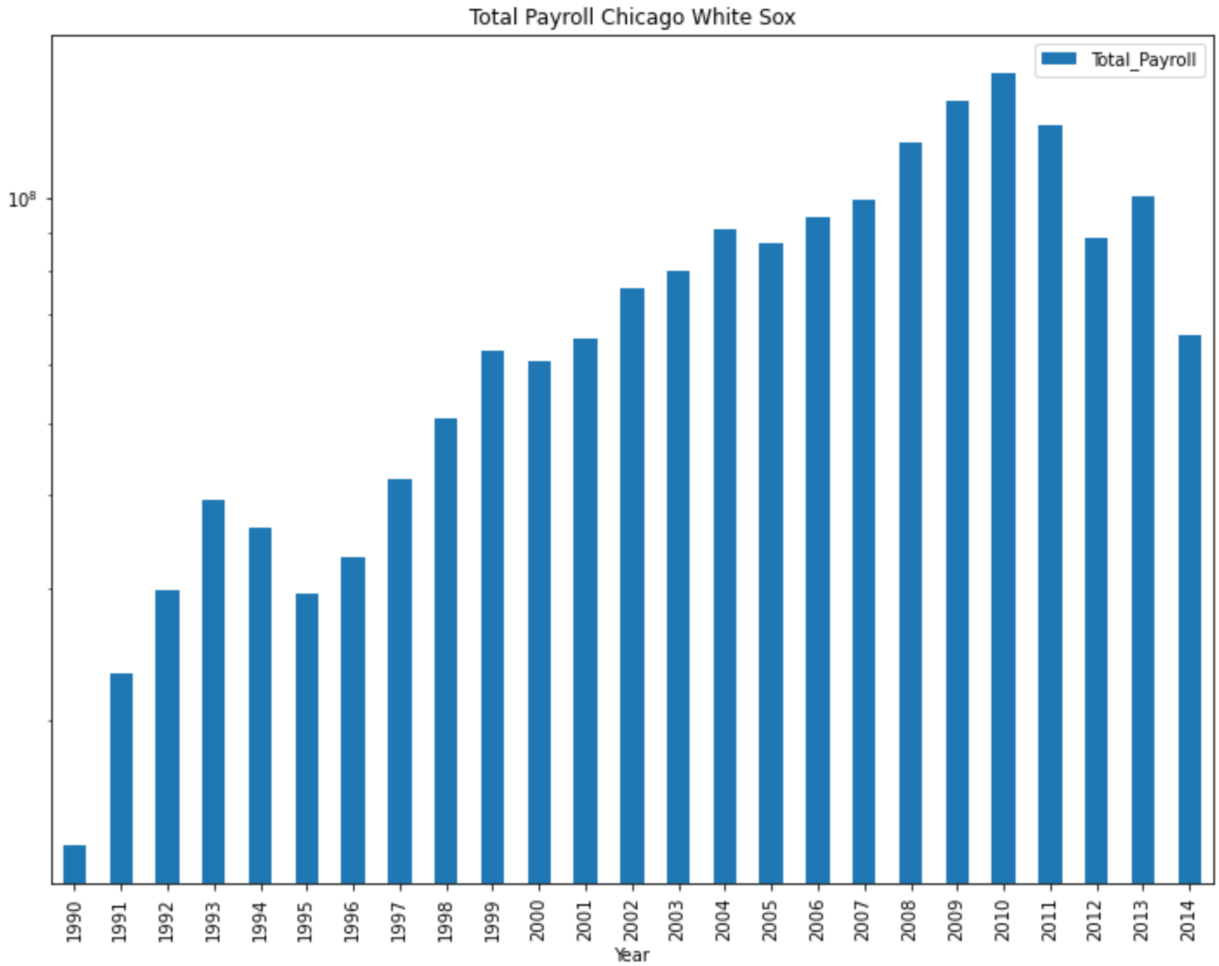

Total Payroll Chicago Colts
Total_Payroll
10^8
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

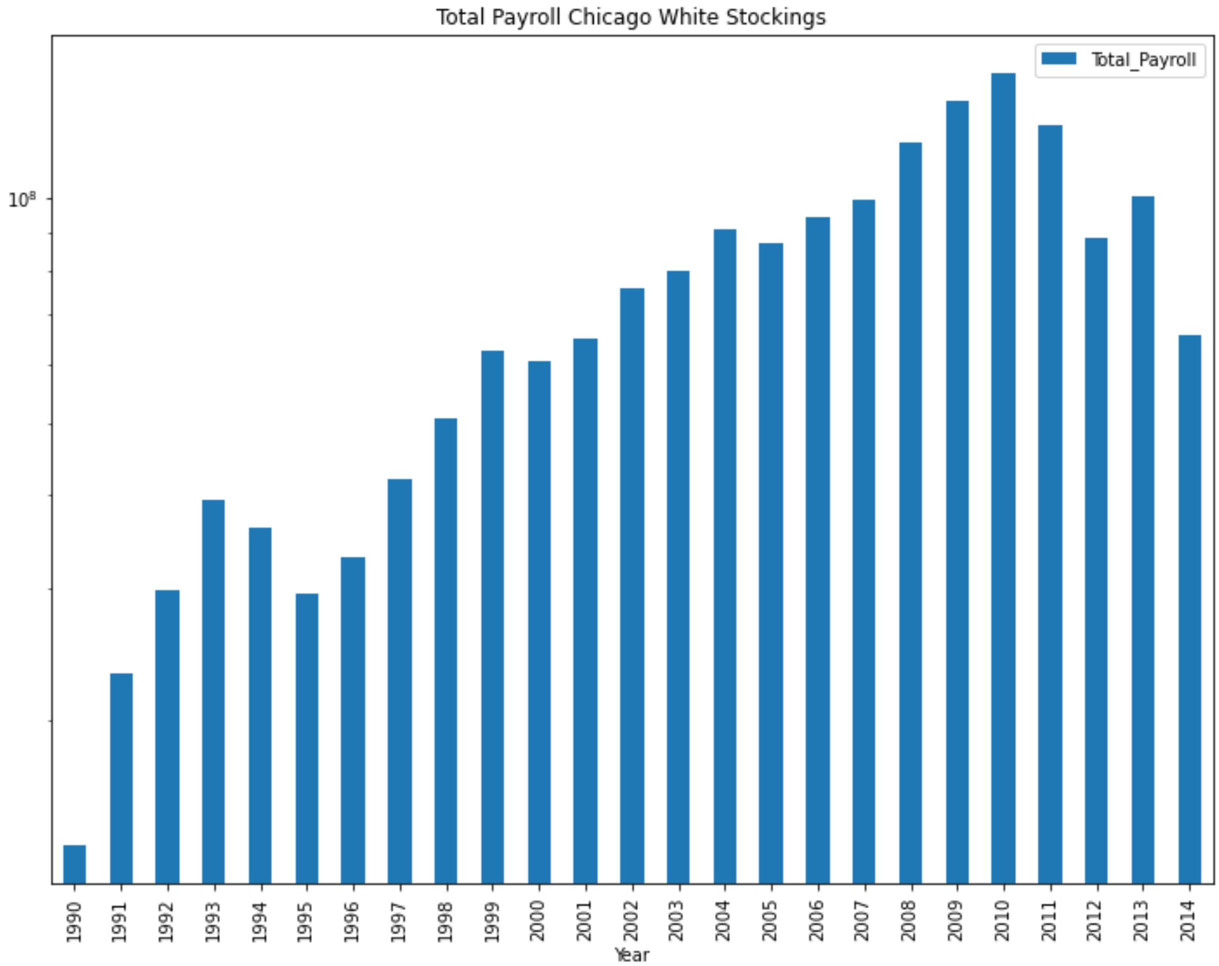

Total Payroll Chicago Cubs
Total_Payroll
10^8
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

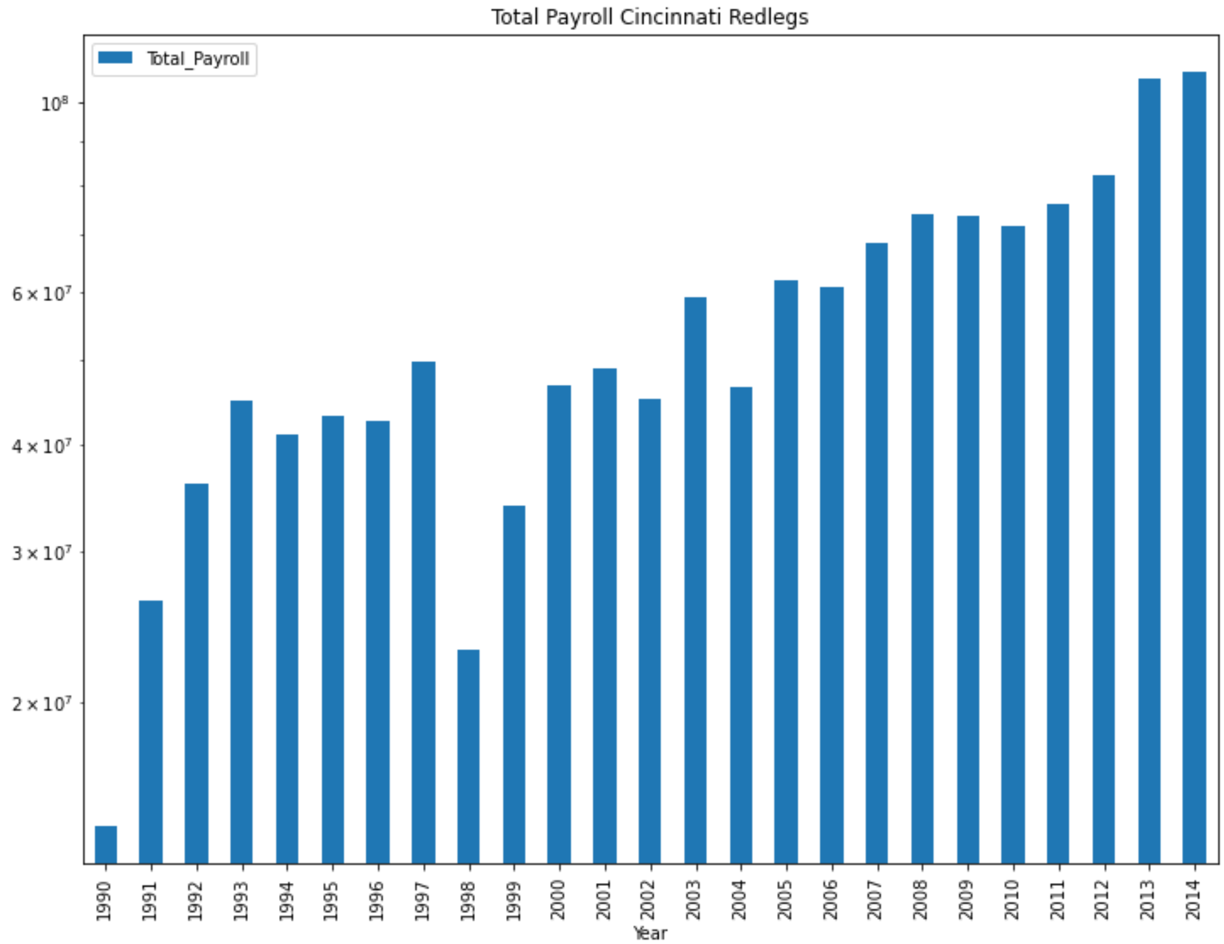

Total Payroll Chicago Cubs
Total_Payroll
10^8
10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

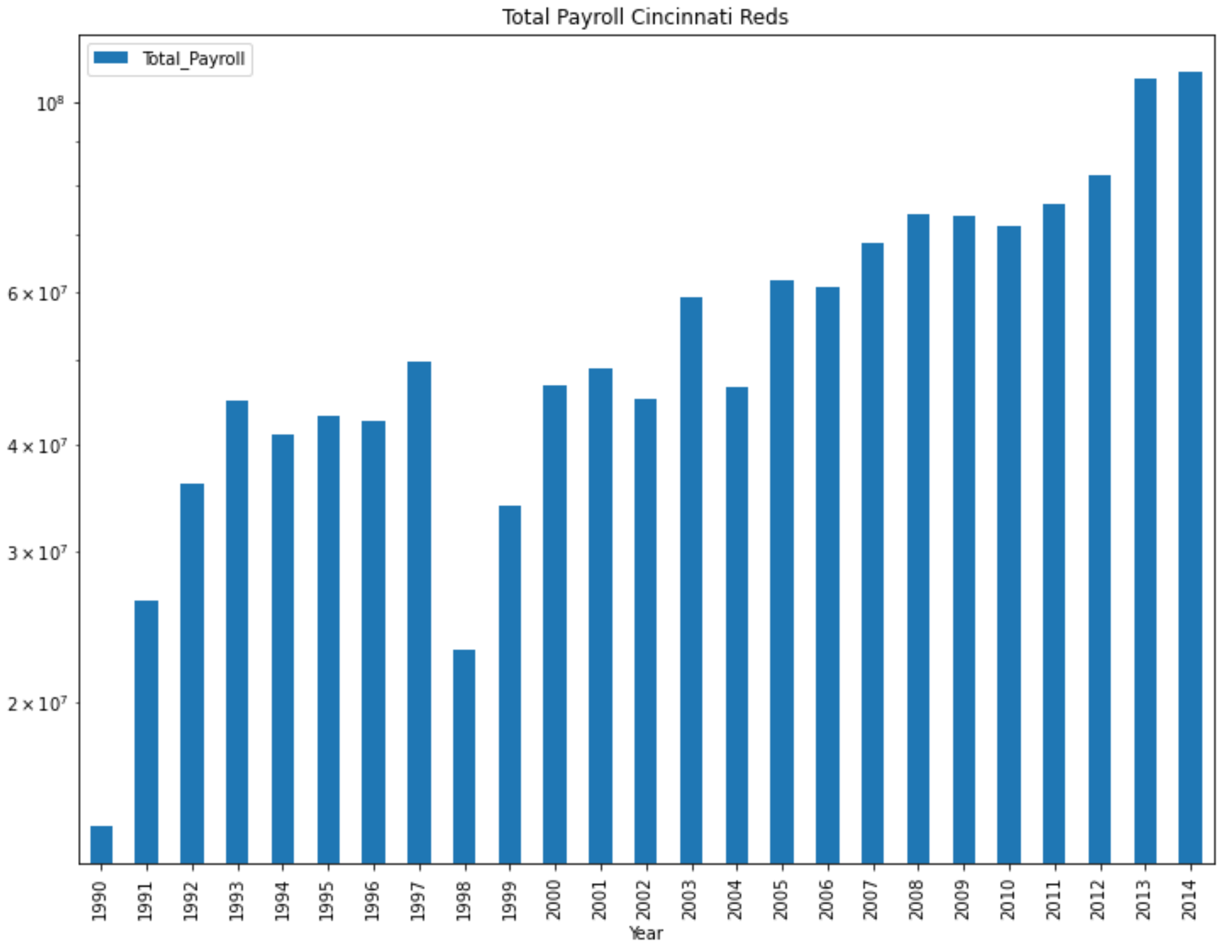

Total Payroll Chicago Orphans
Total_Payroll
10^8
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

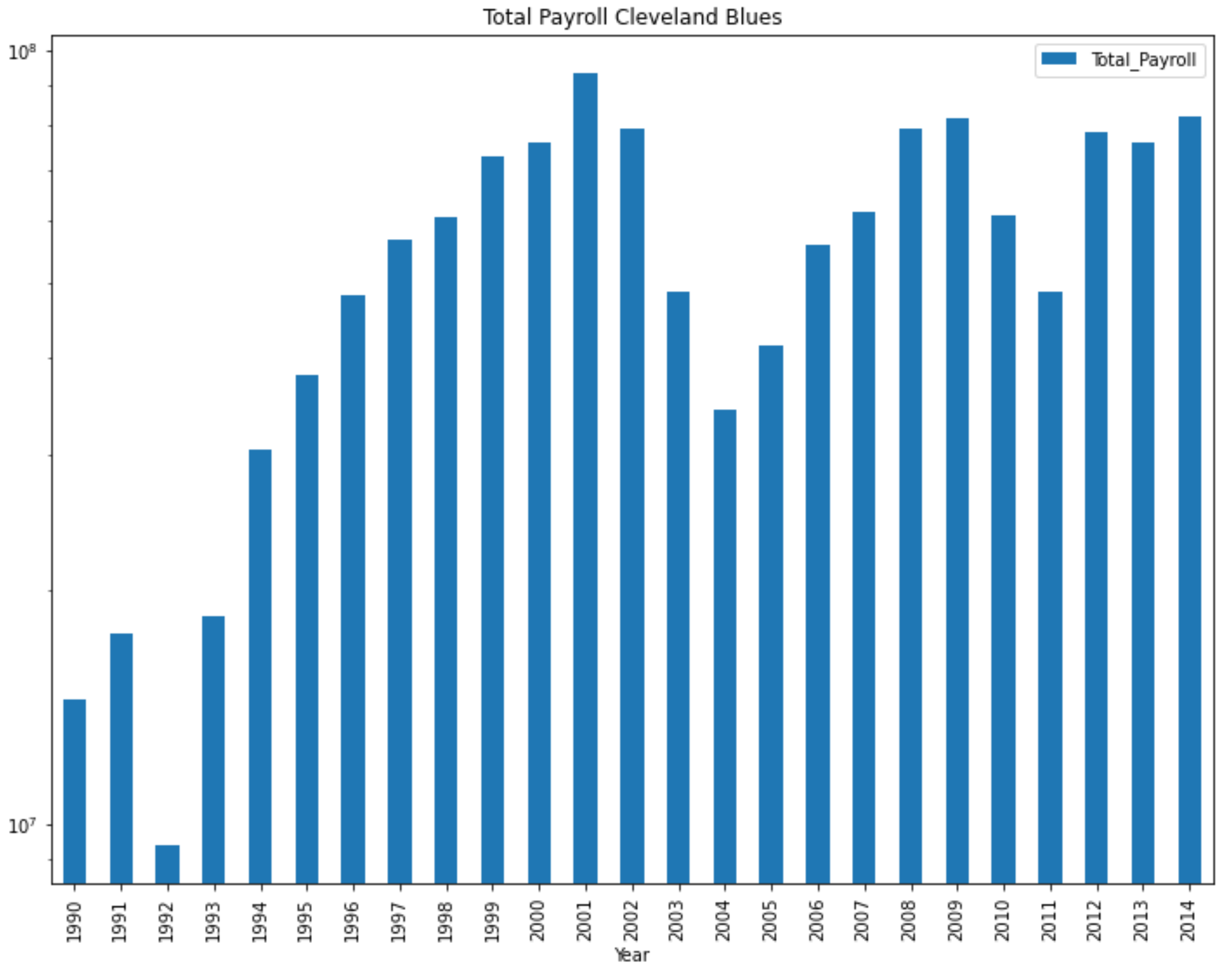

Total Payroll Chicago White Sox
Total_Payroll
10^8
10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

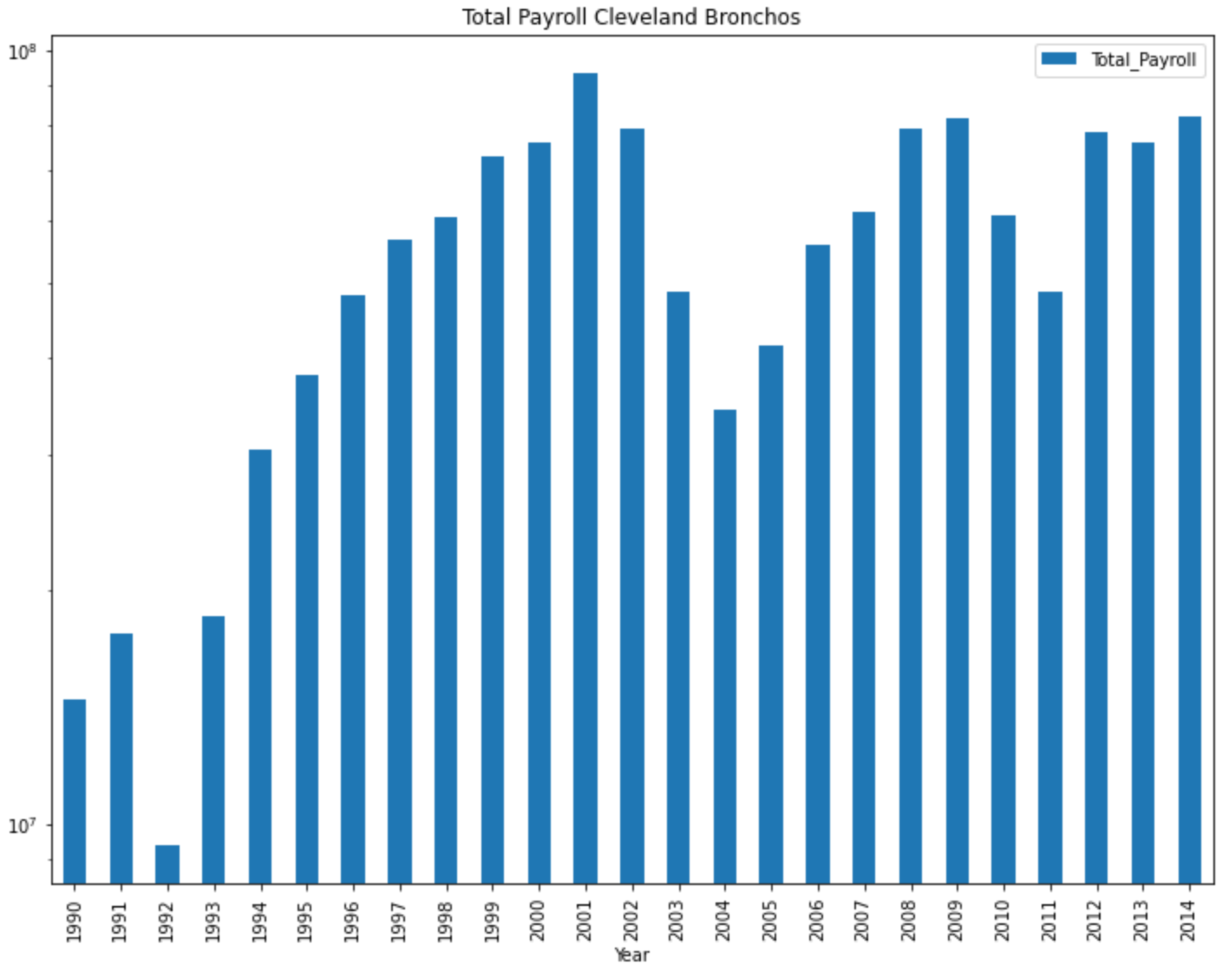

Total Payroll Chicago White Sox
Total_Payroll
10^8
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

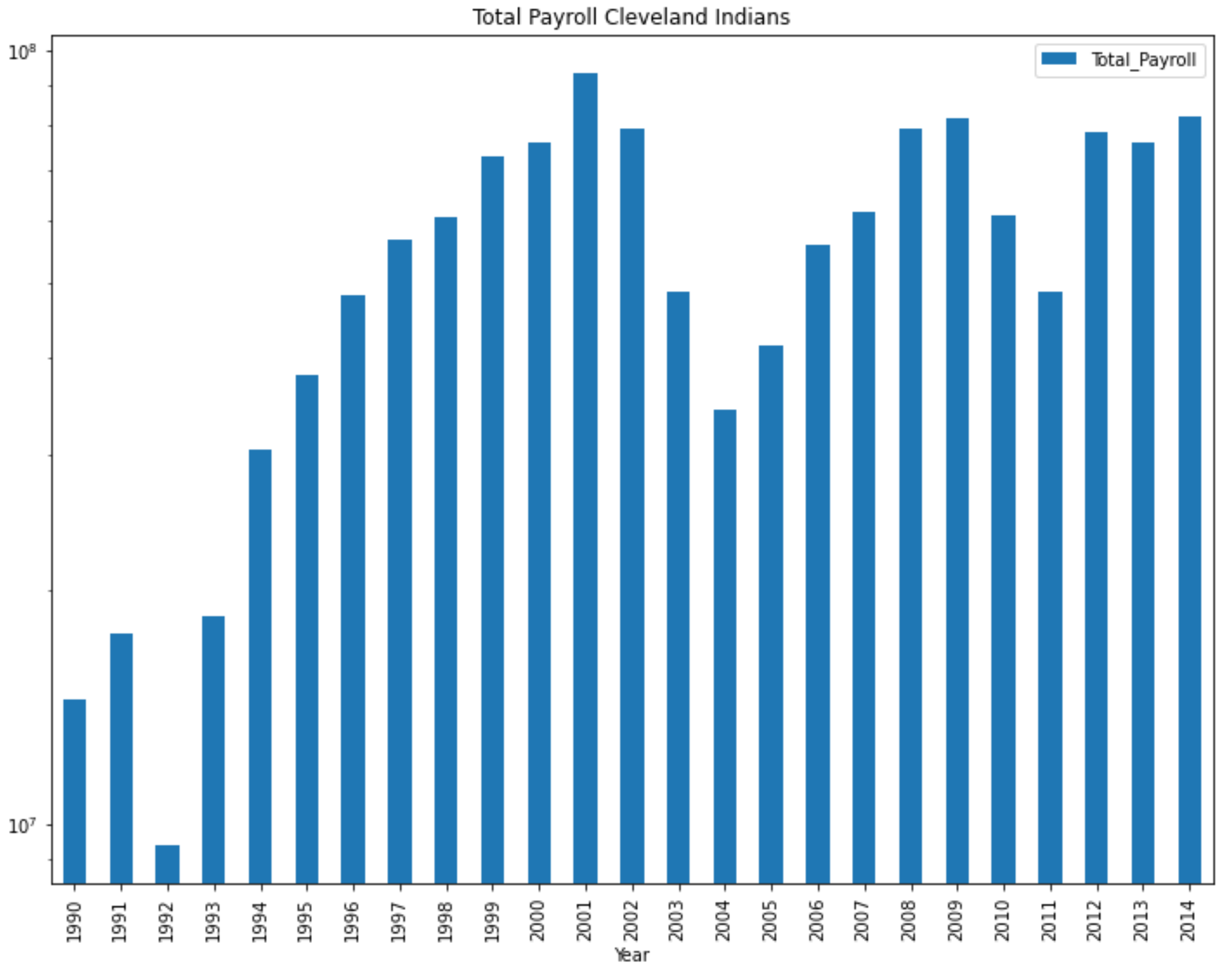

Total Payroll Chicago White Stockings
Total_Payroll
10^8
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014
Year

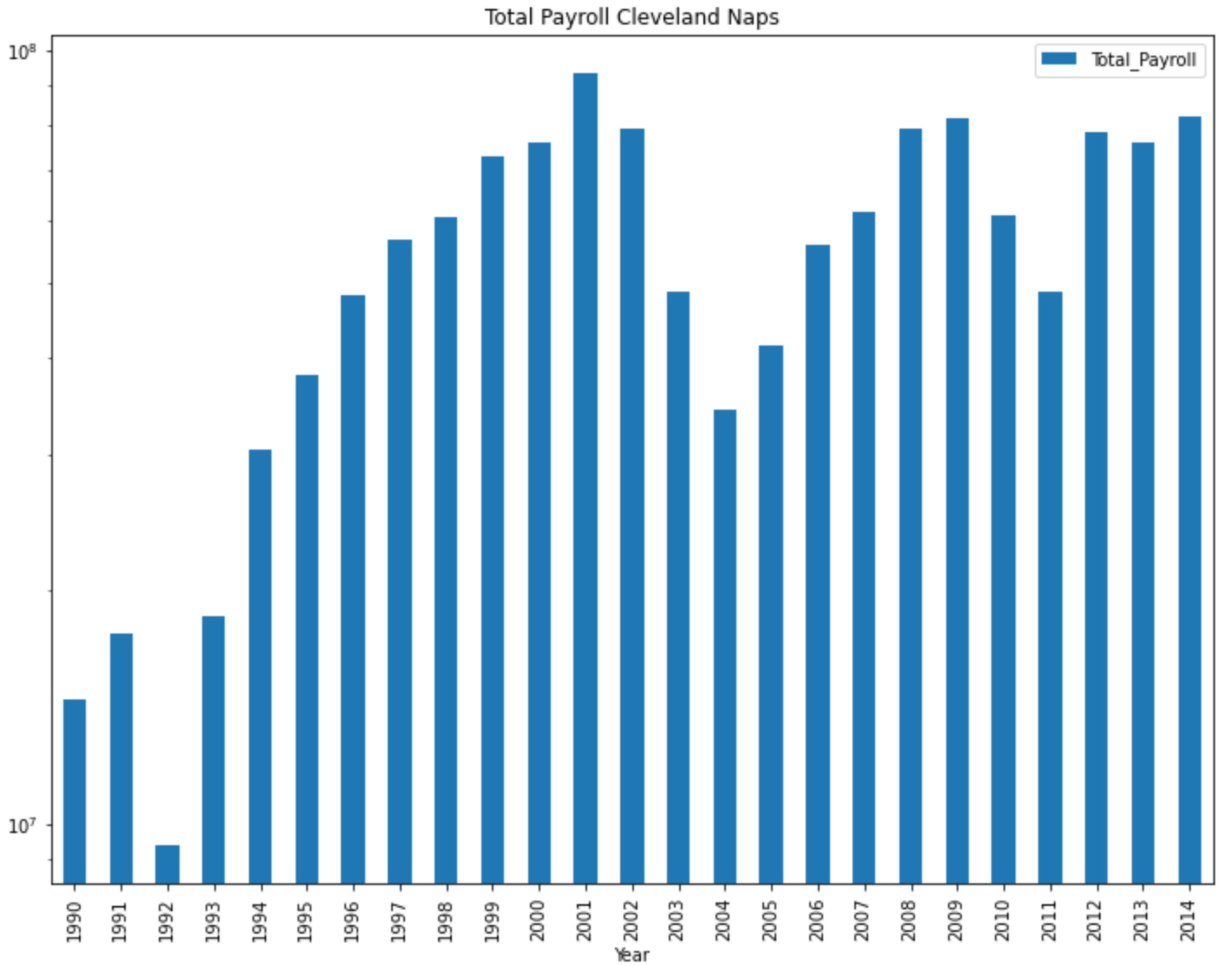

Total Payroll Cincinnati Redlegs
Total_Payroll
10^8
6×10^7
4×10^7
3×10^7
2×10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

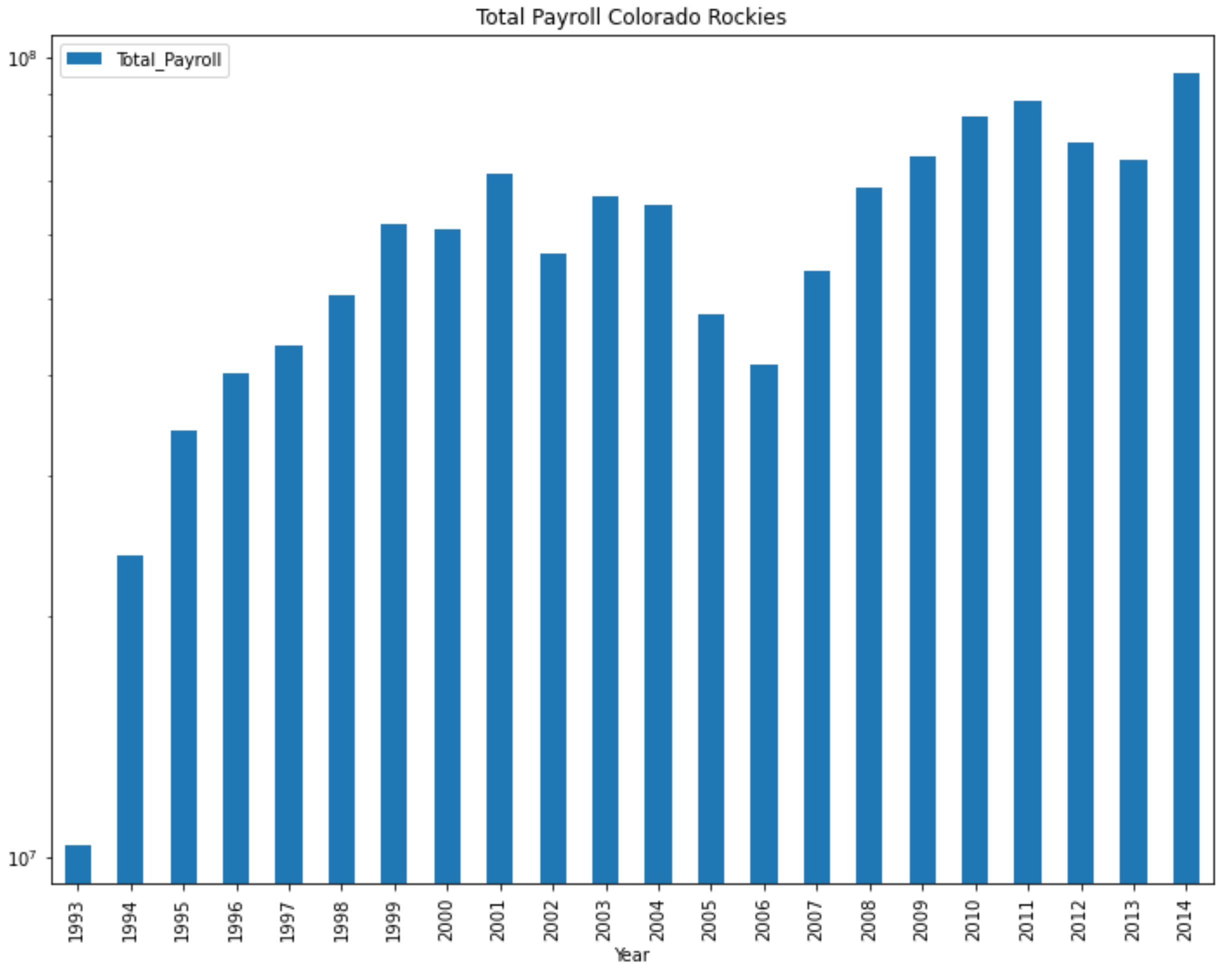

Total Payroll Cincinnati Reds
Total_Payroll
10^8
6×10^7
4×10^7
3×10^7
2×10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

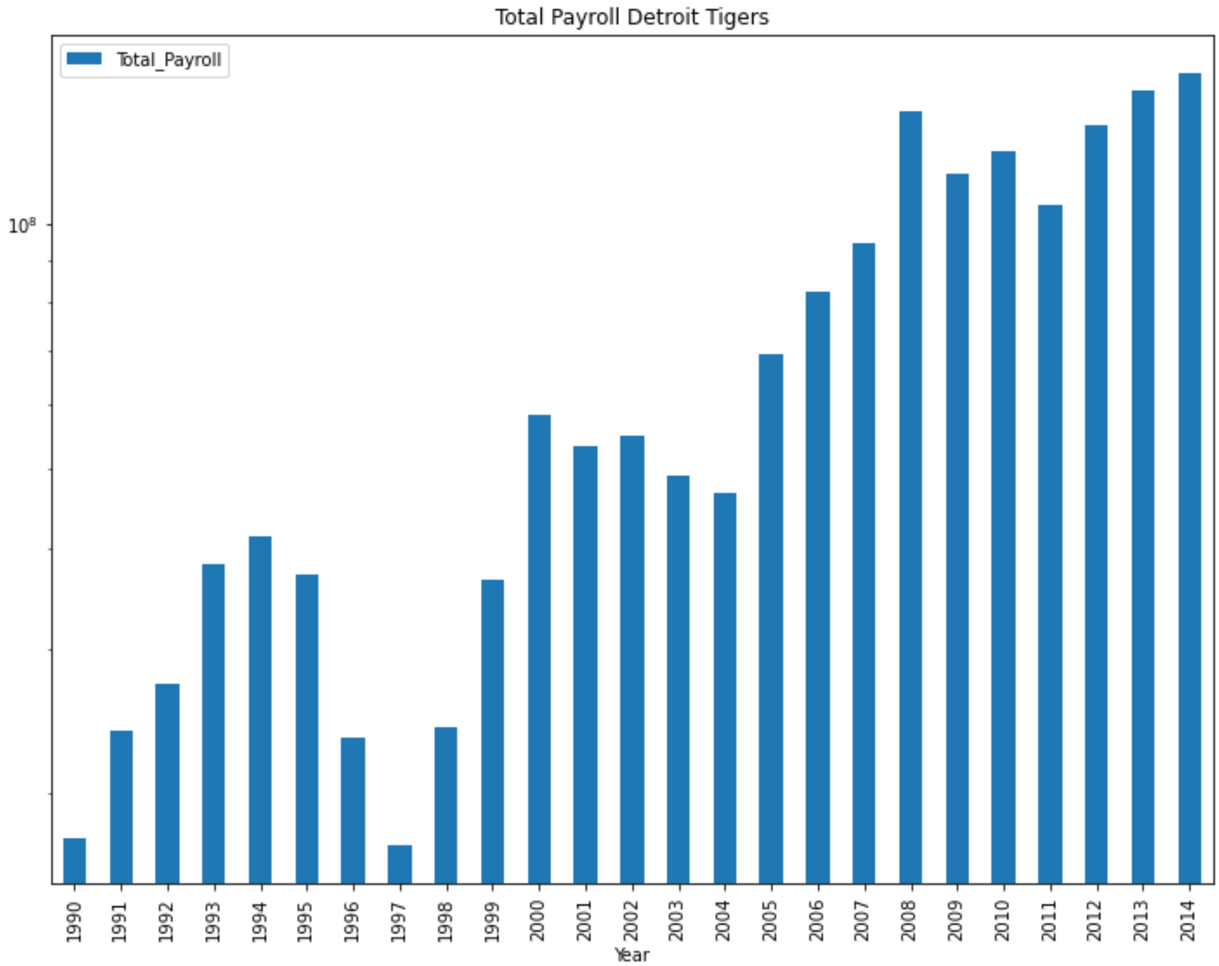

Total Payroll Cleveland Blues
Total_Payroll
10^8
10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

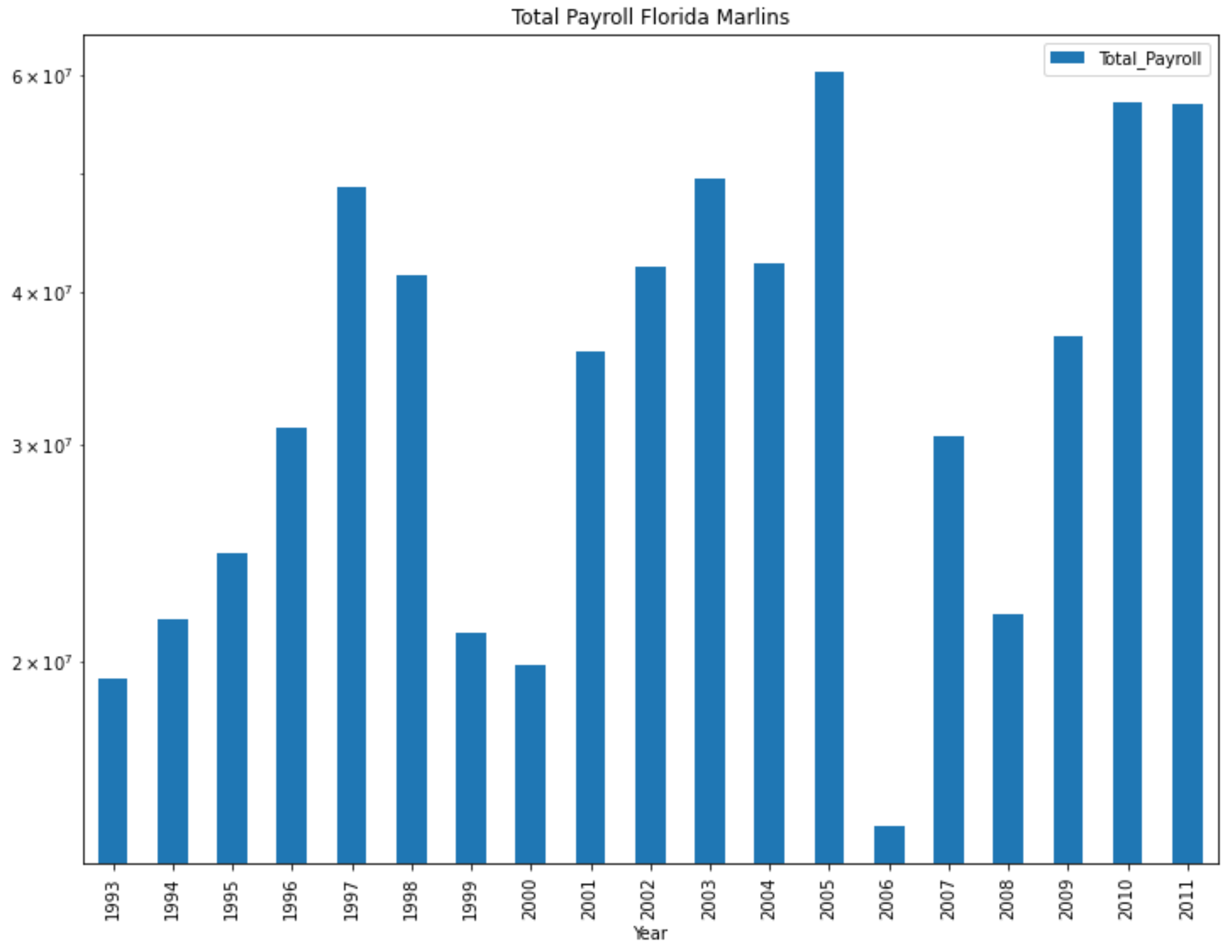

Total Payroll Cleveland Bronchos
Total_Payroll
10^8
10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

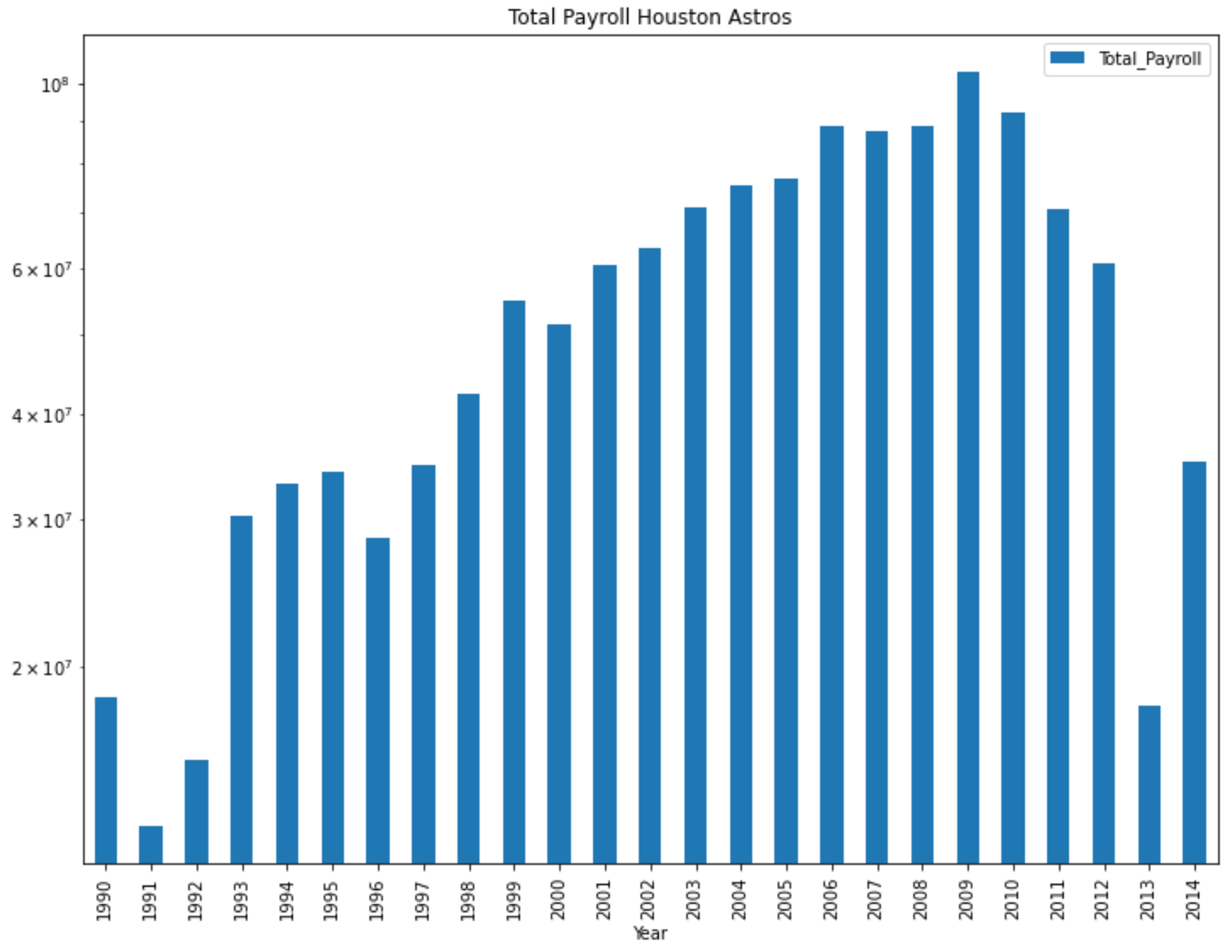

Total Payroll Cleveland Indians
Total_Payroll
10^8
10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

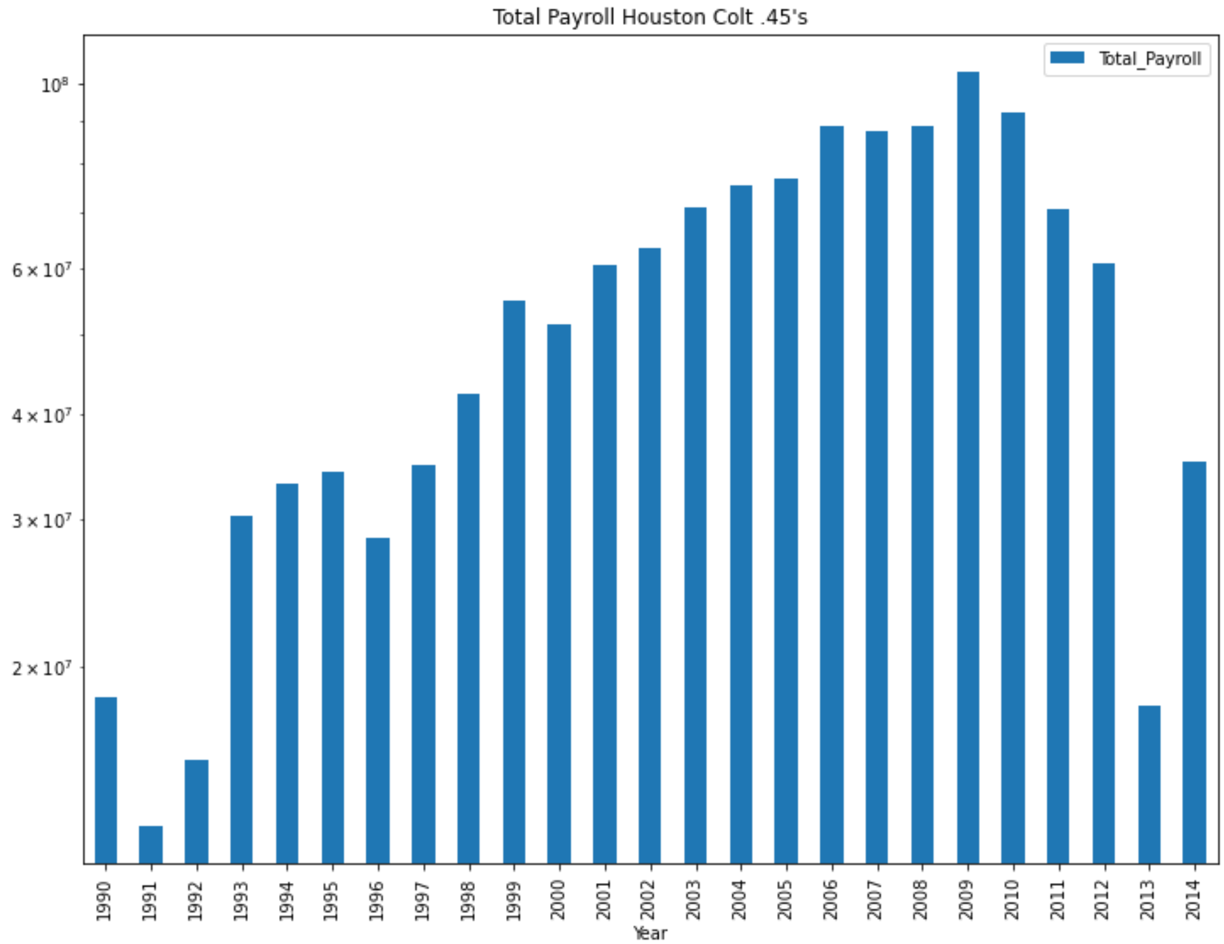

Total Payroll Cleveland Naps
Total_Payroll
10^8
10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

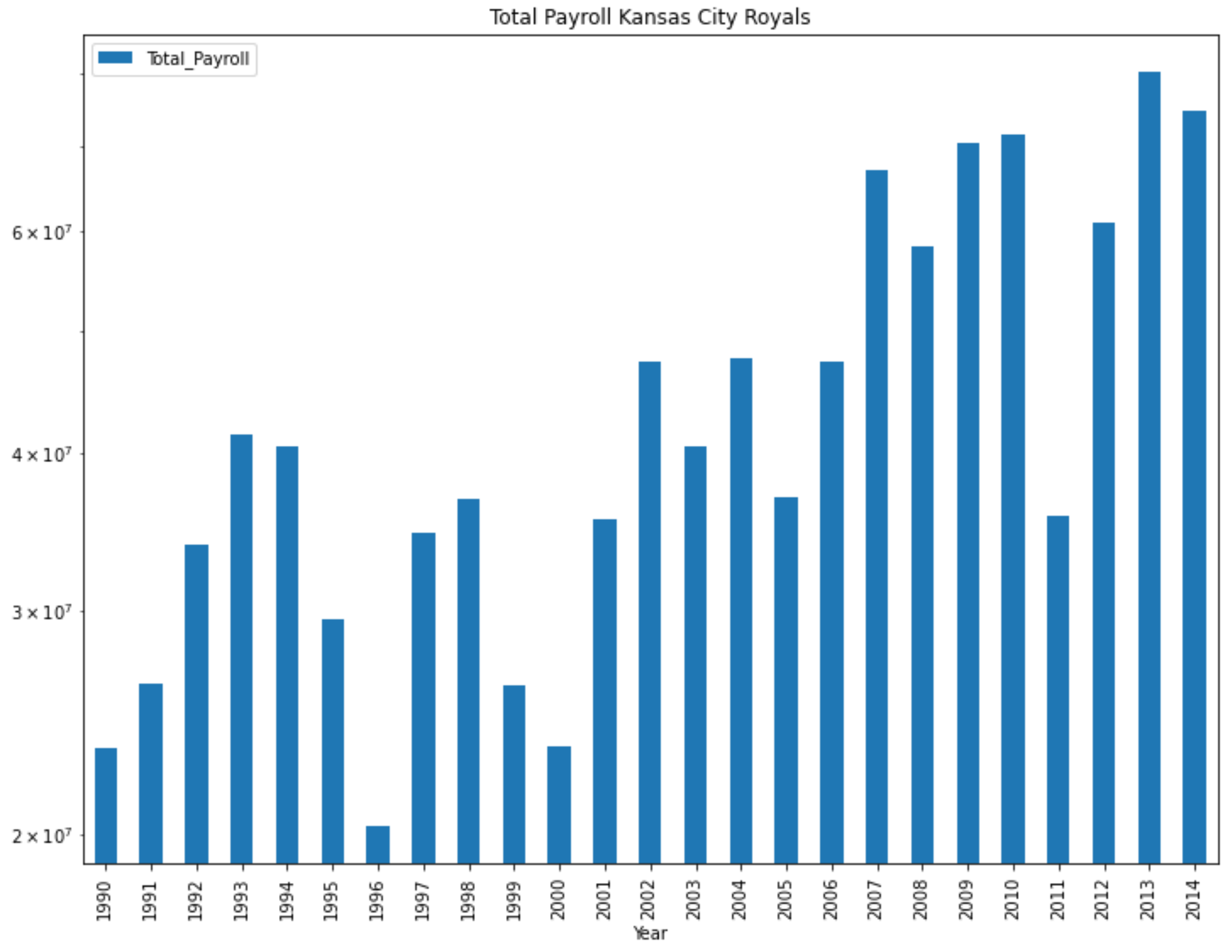

Total Payroll Colorado Rockies
Total_Payroll
10^8
10^7
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

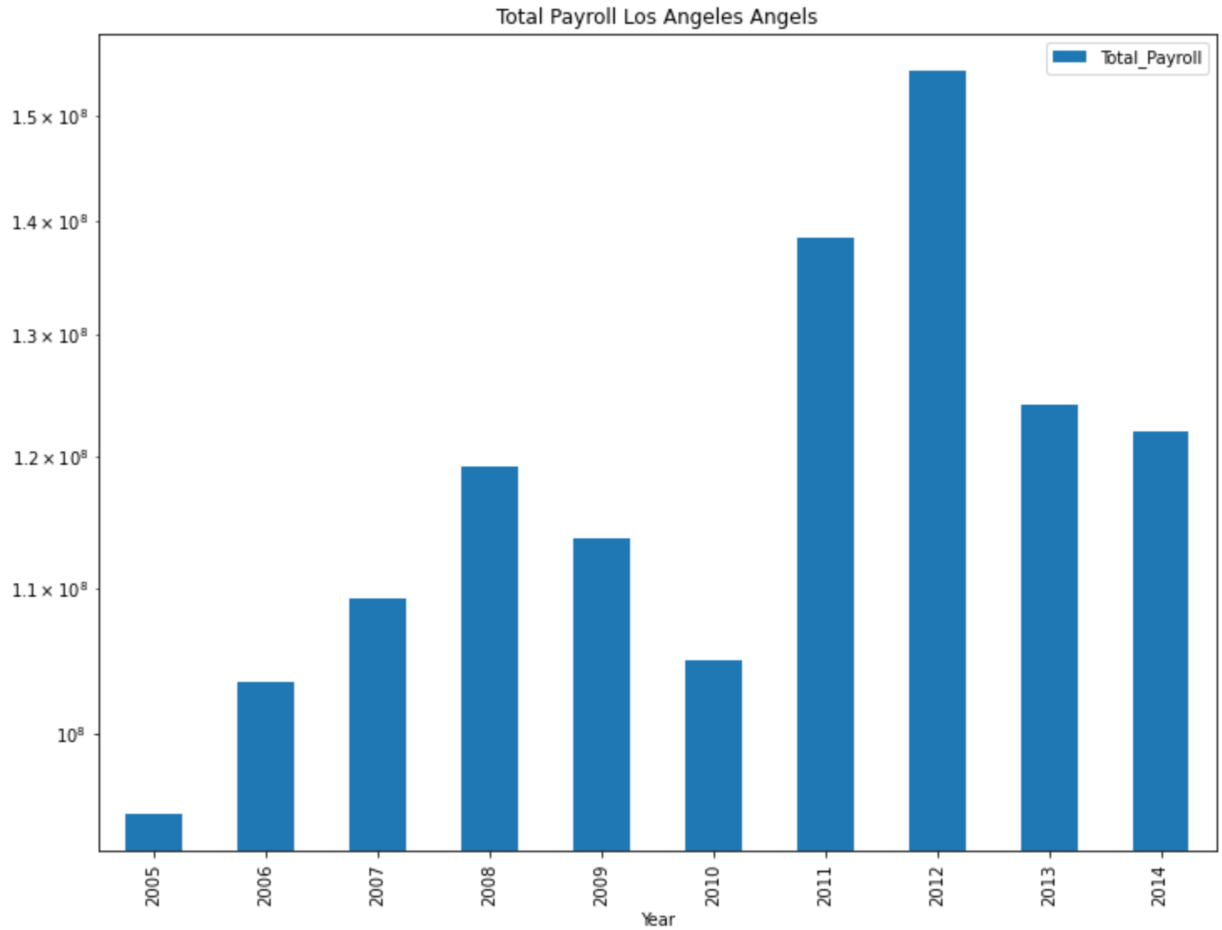

Total Payroll Detroit Tigers
Total_Payroll
10^8
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

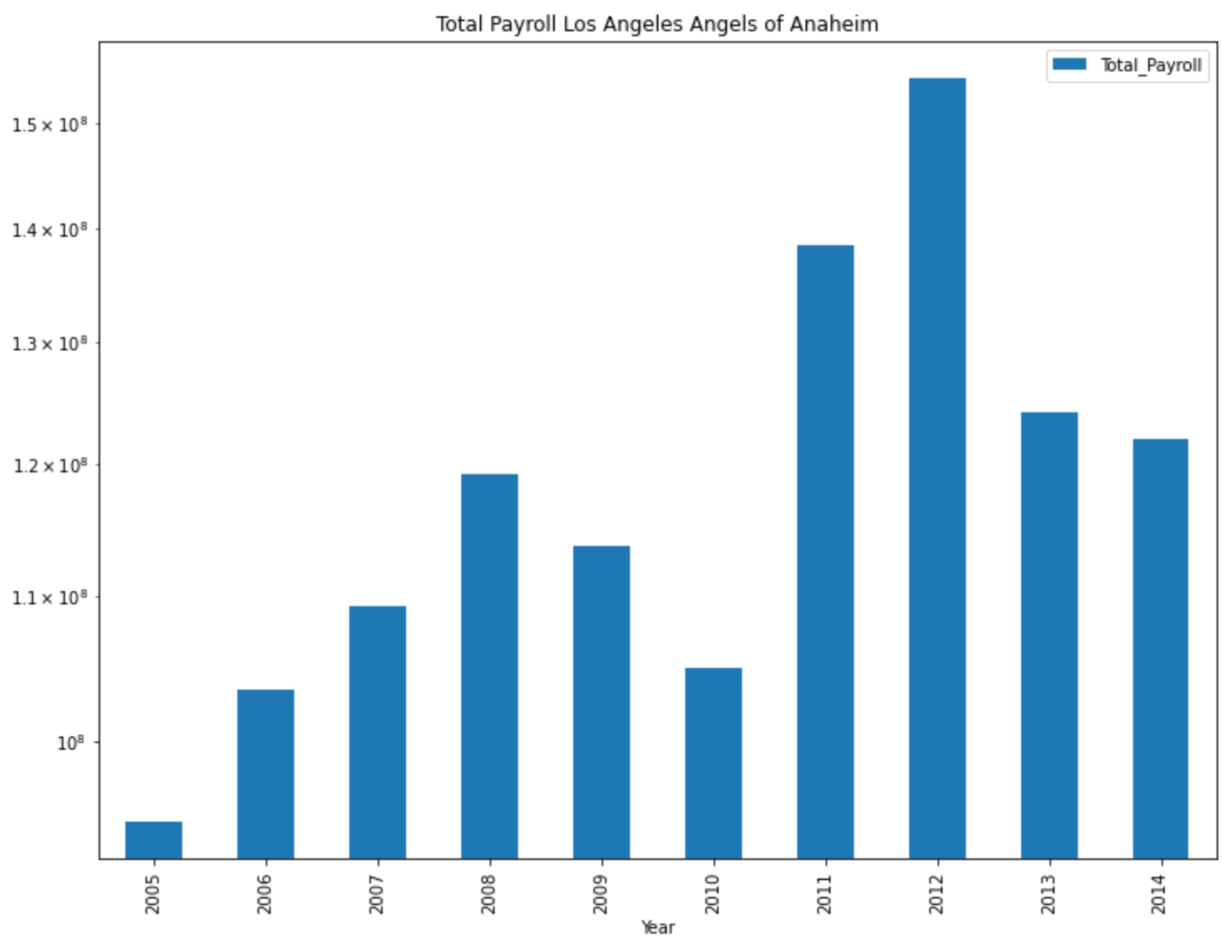

Total Payroll Florida Marlins
Total_Payroll
6×10^7
4×10^7
3×10^7
2×10^7
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
Year

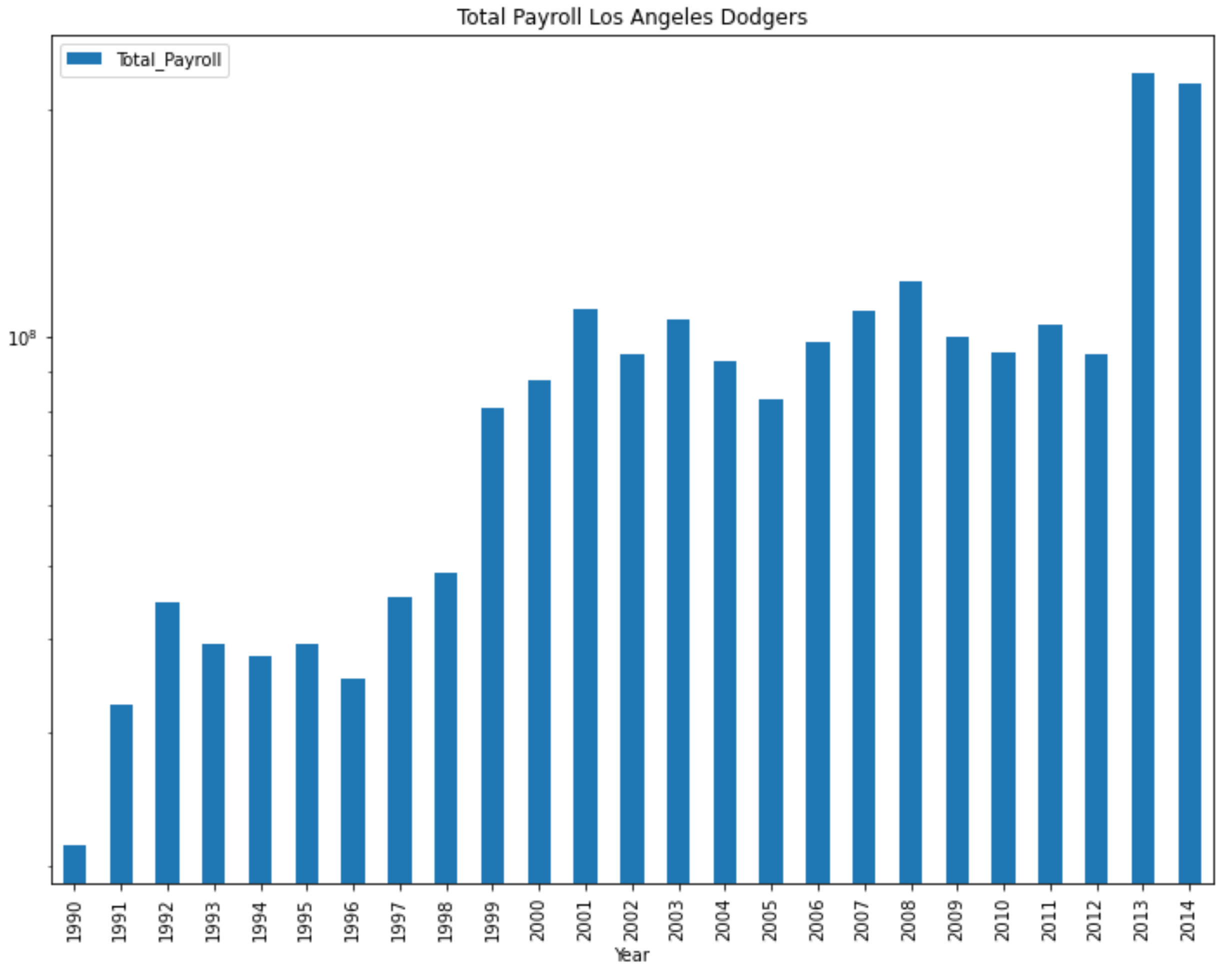

Total Payroll Houston Astros
Total_Payroll
10^8
6×10^7
4×10^7
3×10^7
2×10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

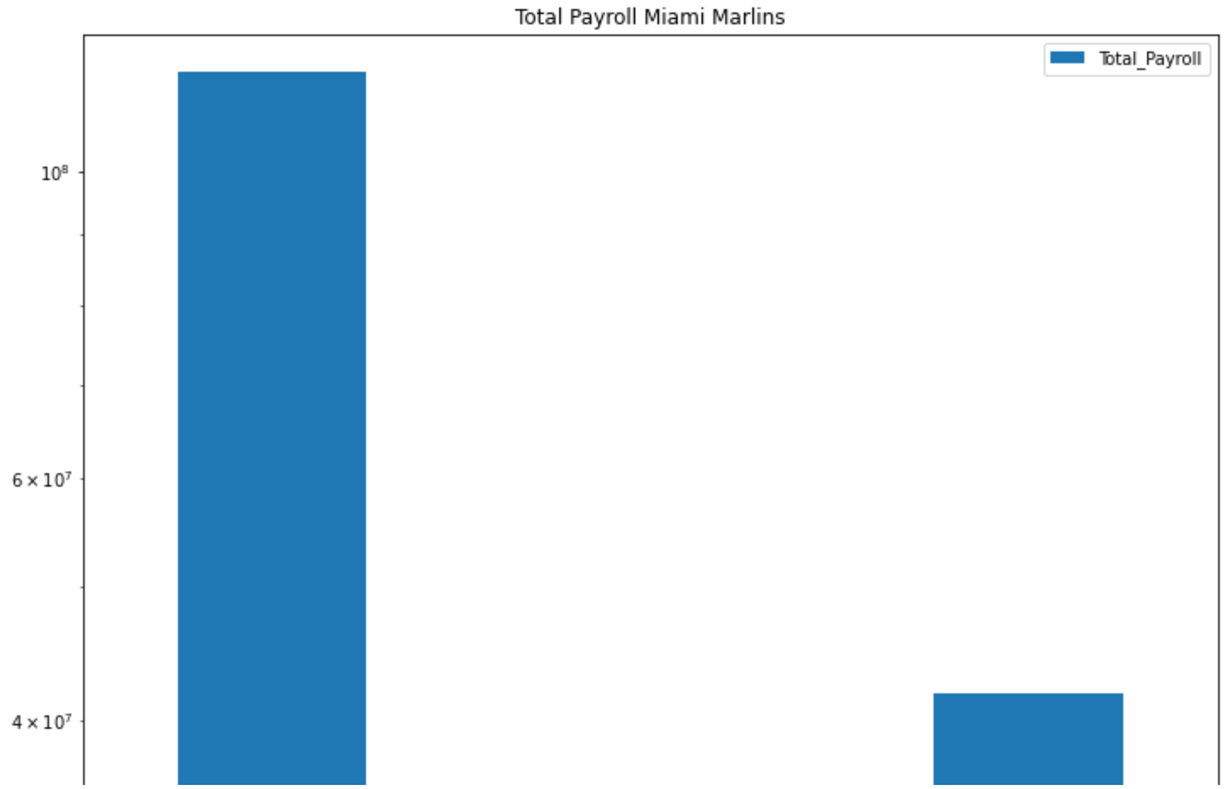

Total Payroll Houston Colt .45's
Total_Payroll
10^8
6×10^7
4×10^7
3×10^7
2×10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year


Total Payroll Kansas City Royals
Total_Payroll
6×10^7
4×10^7
3×10^7
2×10^7
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year


Total Payroll Los Angeles Angels
Total_Payroll
1.5×10^8
1.4×10^8
1.3×10^8
1.2×10^8
1.1×10^8
10^8
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year


Total Payroll Los Angeles Angels of Anaheim
Total_Payroll
1.5×10^8
1.4×10^8
1.3×10^8
1.2×10^8
1.1×10^8
10^8
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year


Total Payroll Los Angeles Dodgers
Total_Payroll
10^8
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year



Observations about the plots :

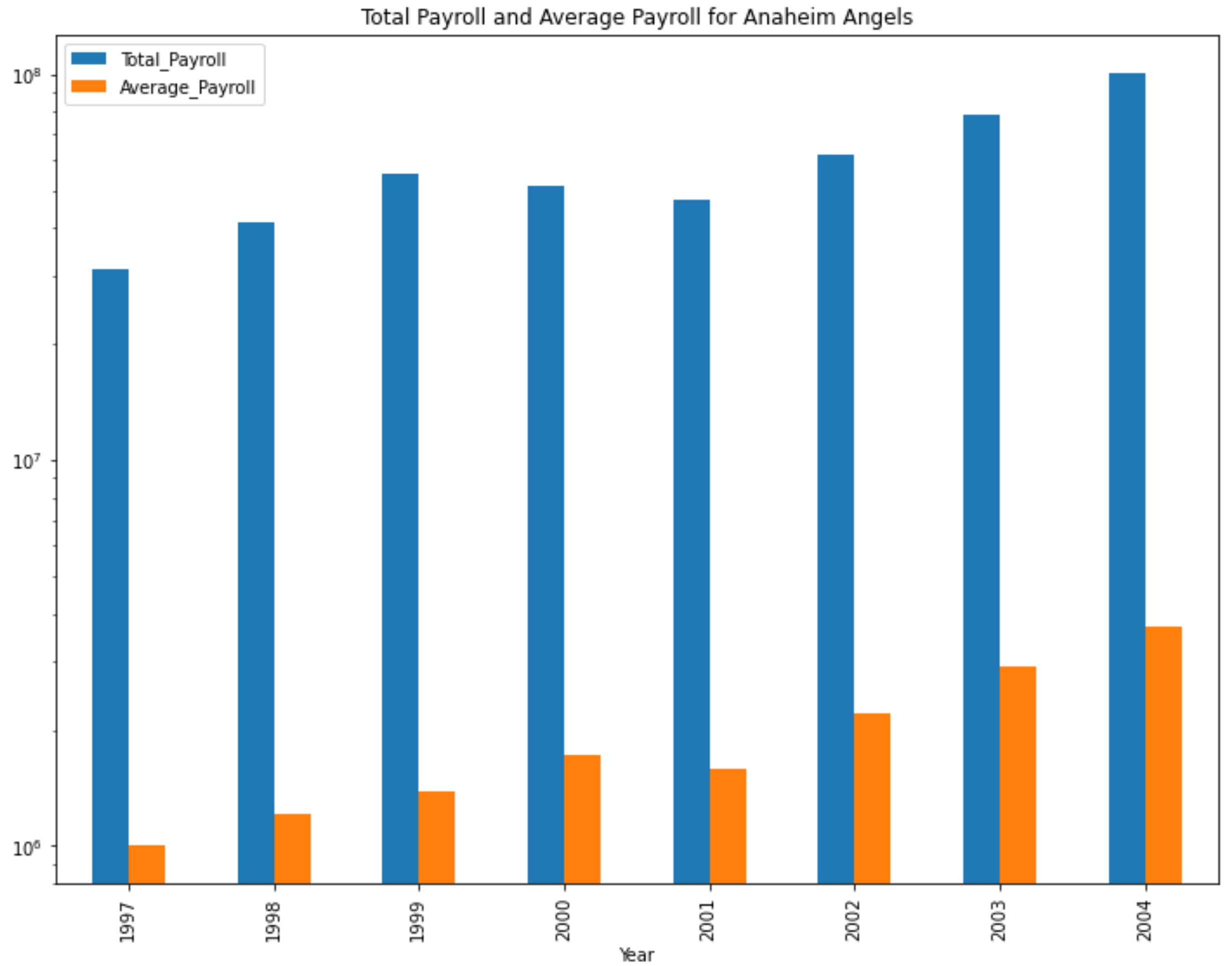
Over time, the payroll has increased for most of the teams.

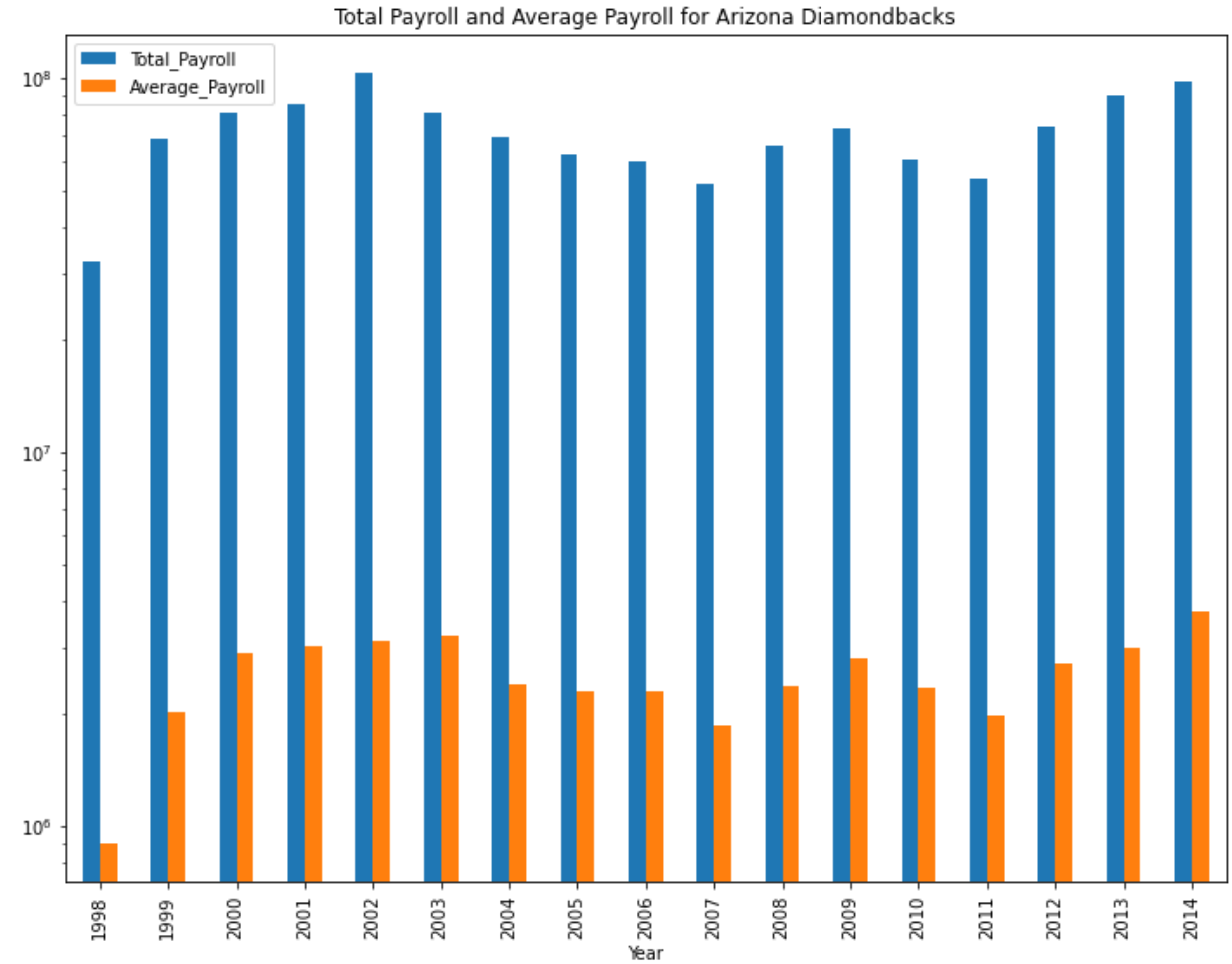
The average and total payroll have increased steadily over time.

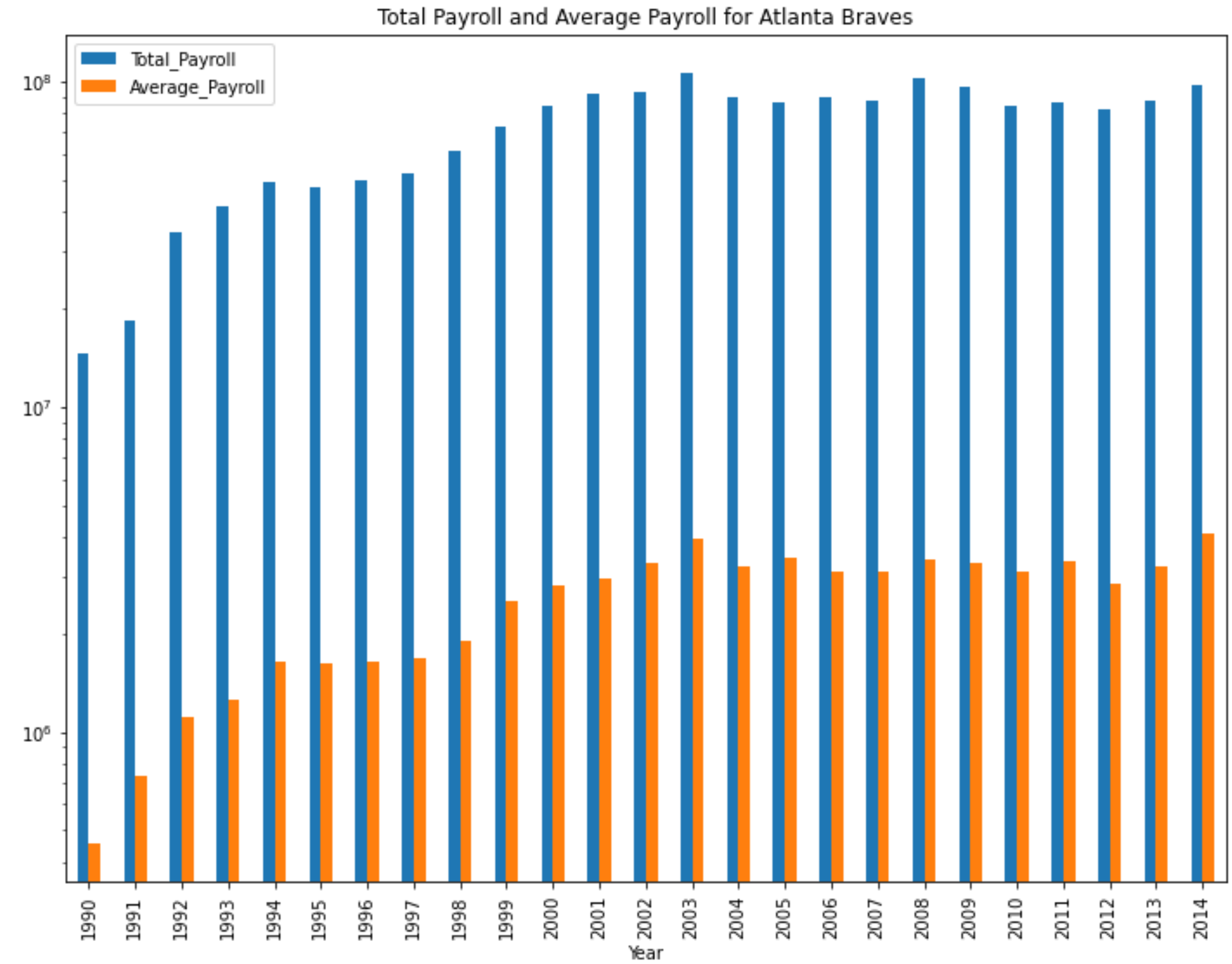
Here is a plot of the sums and the average over the years for each team. The bar plots below show an increasing trend.

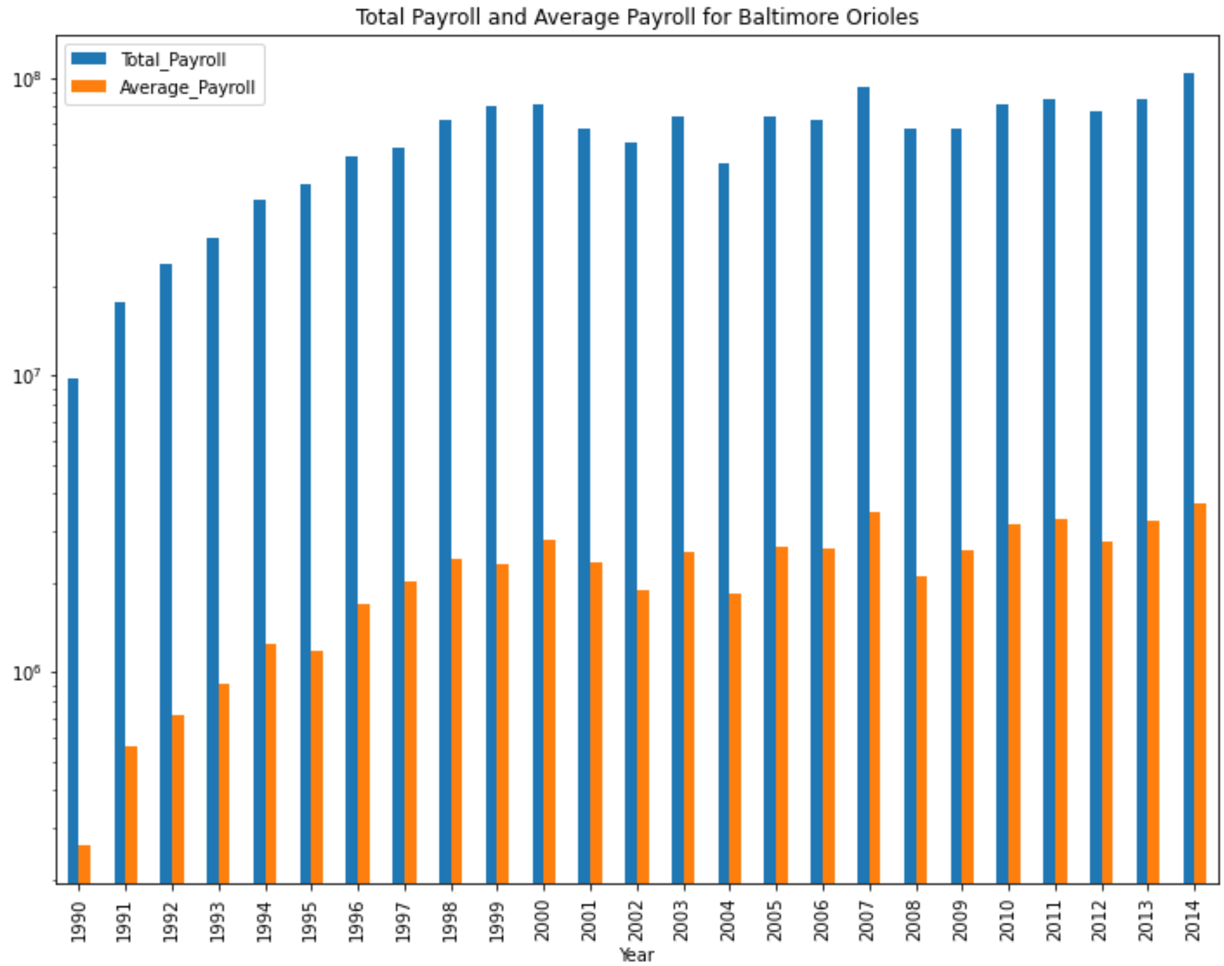
In [417…

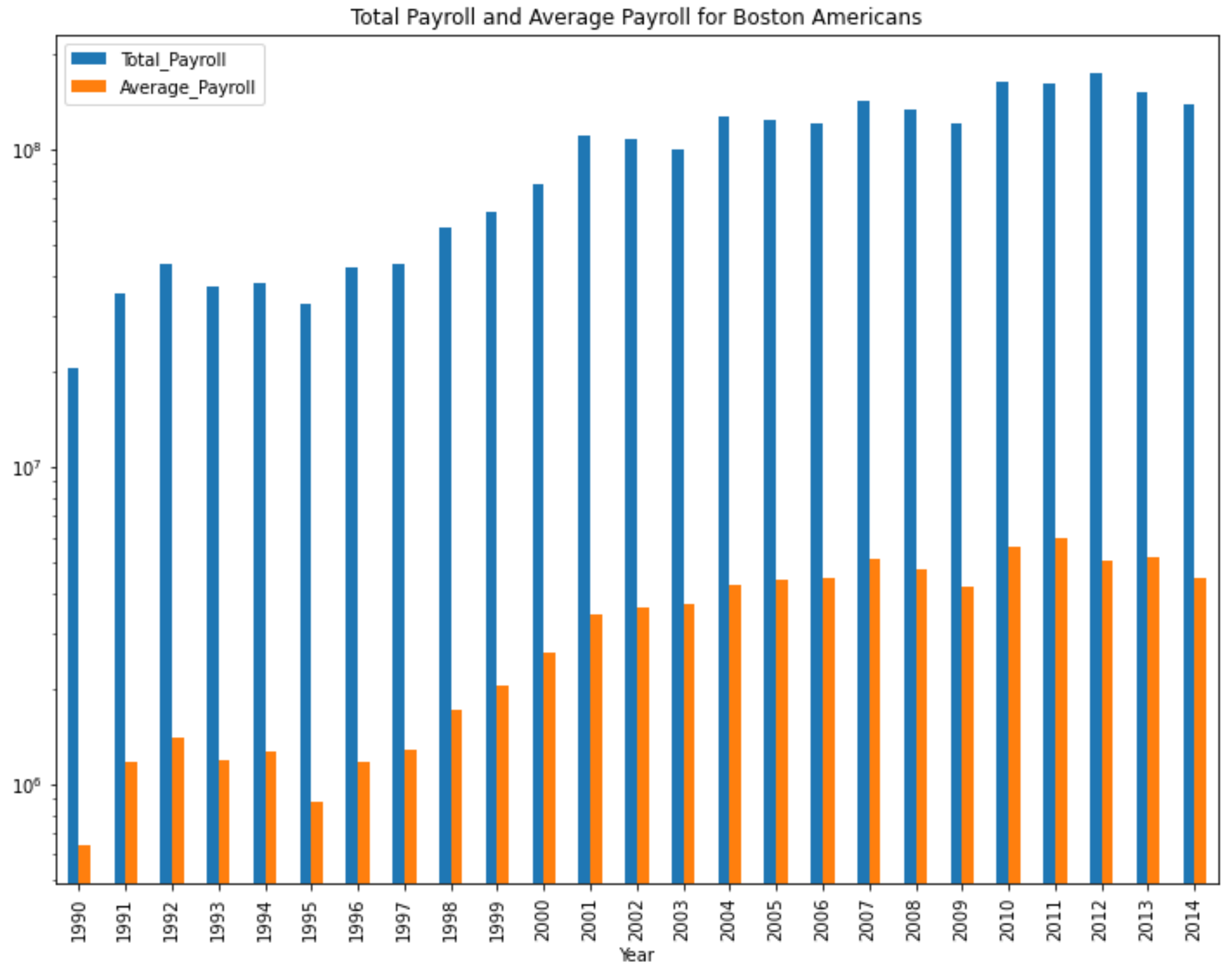
```
num_results = len(result_frames_objects)
#prevents warning after 20 figures shown
plt.rcParams.update({'figure.max_open_warning':0})
figsize = (12,9)
for i in range(num_results):
    dfr = result_frames_objects[i]
    team_name = team_names[i]
    team_id = team_ids[i]
    #we only plot if the data frame in
    #question is not empty
    if not dfr.empty:
        dfr.plot(x='Year', y=['Total_Payroll','Average_Payroll'], logy=True,
                 title = 'Total Payroll and Average Payroll for ' + team_nam
```

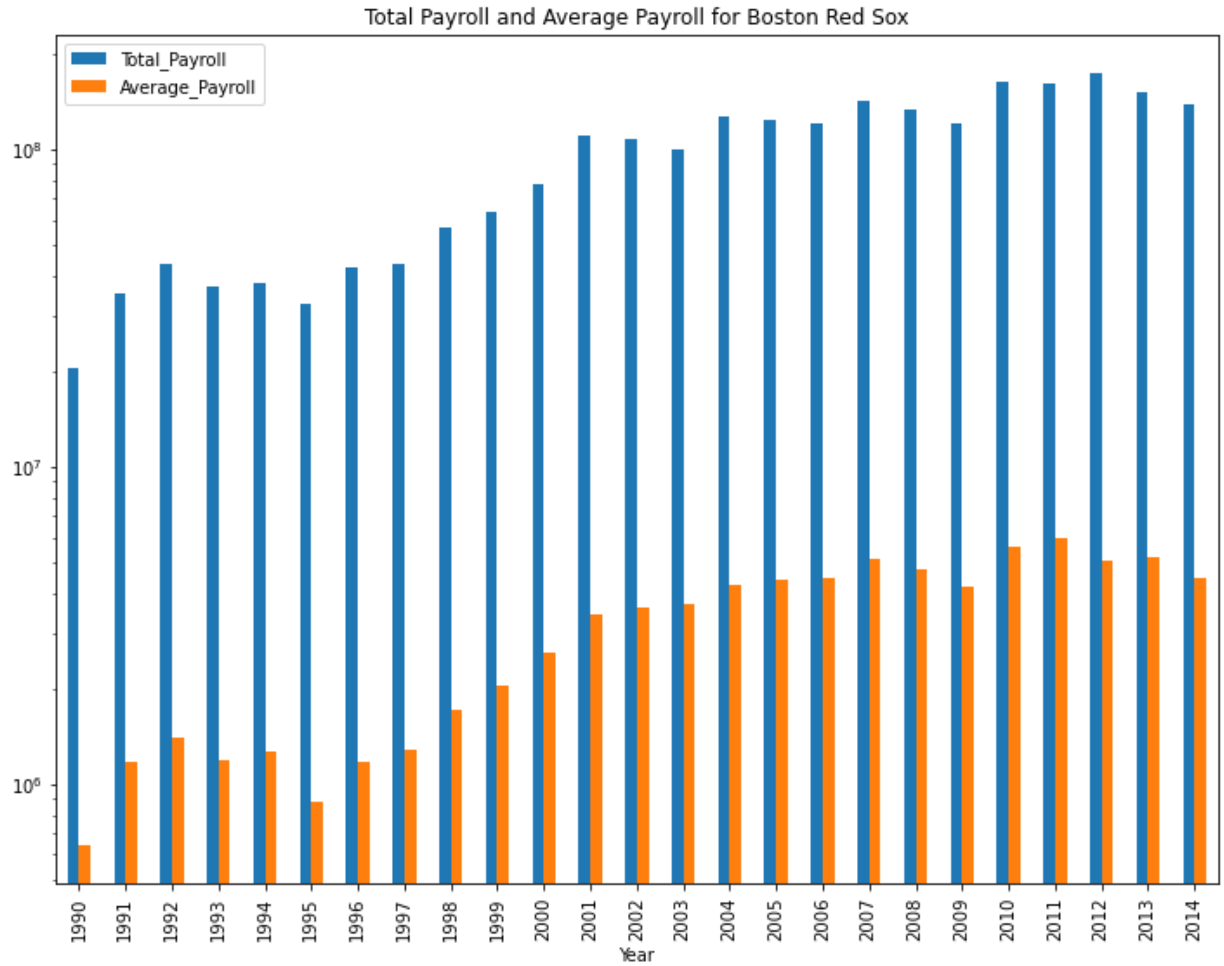

Total Payroll and Average Payroll for Anaheim Angels
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1997
1998
1999
2000
2001
2002
2003
2004
Year

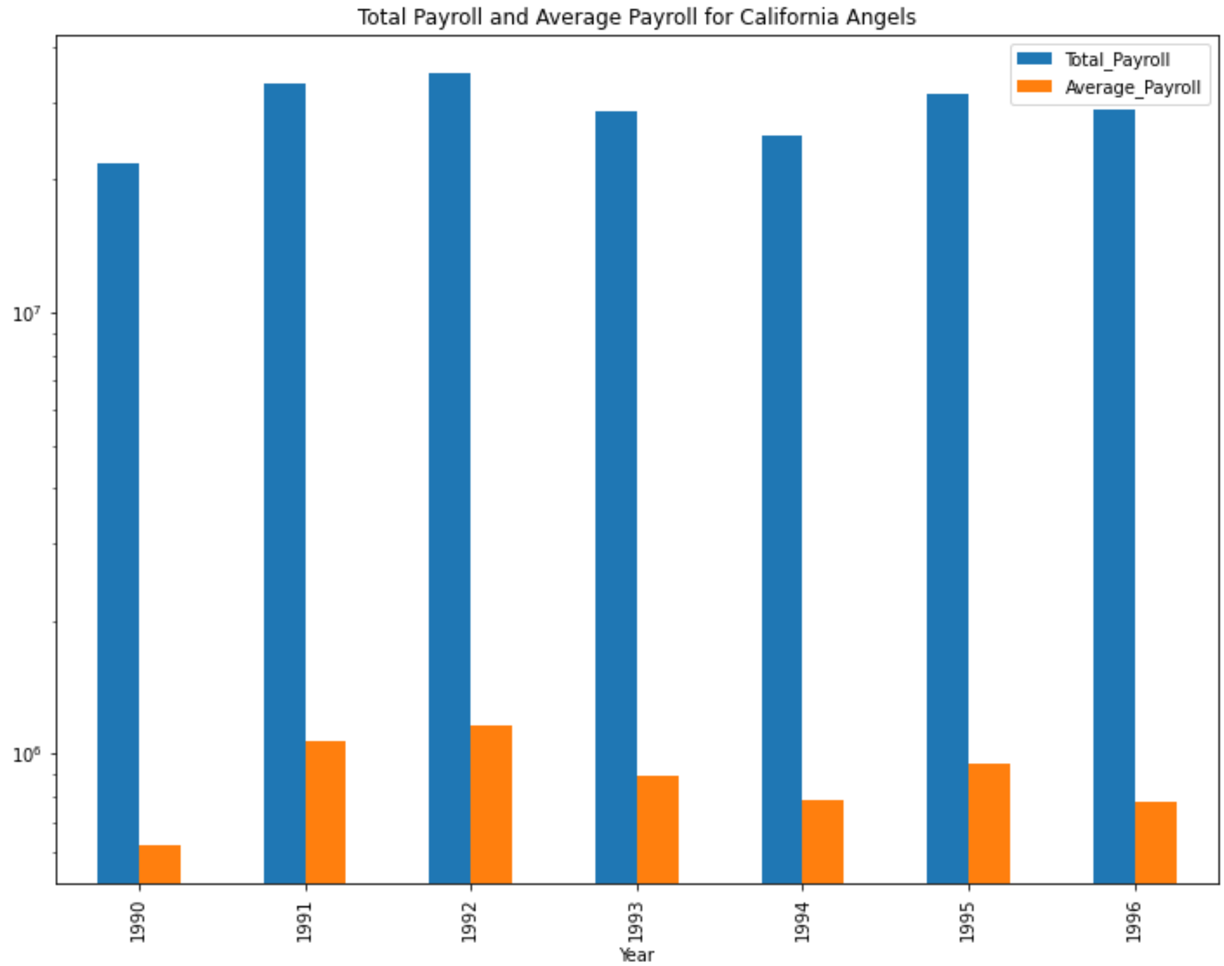

Total Payroll and Average Payroll for Arizona Diamondbacks
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

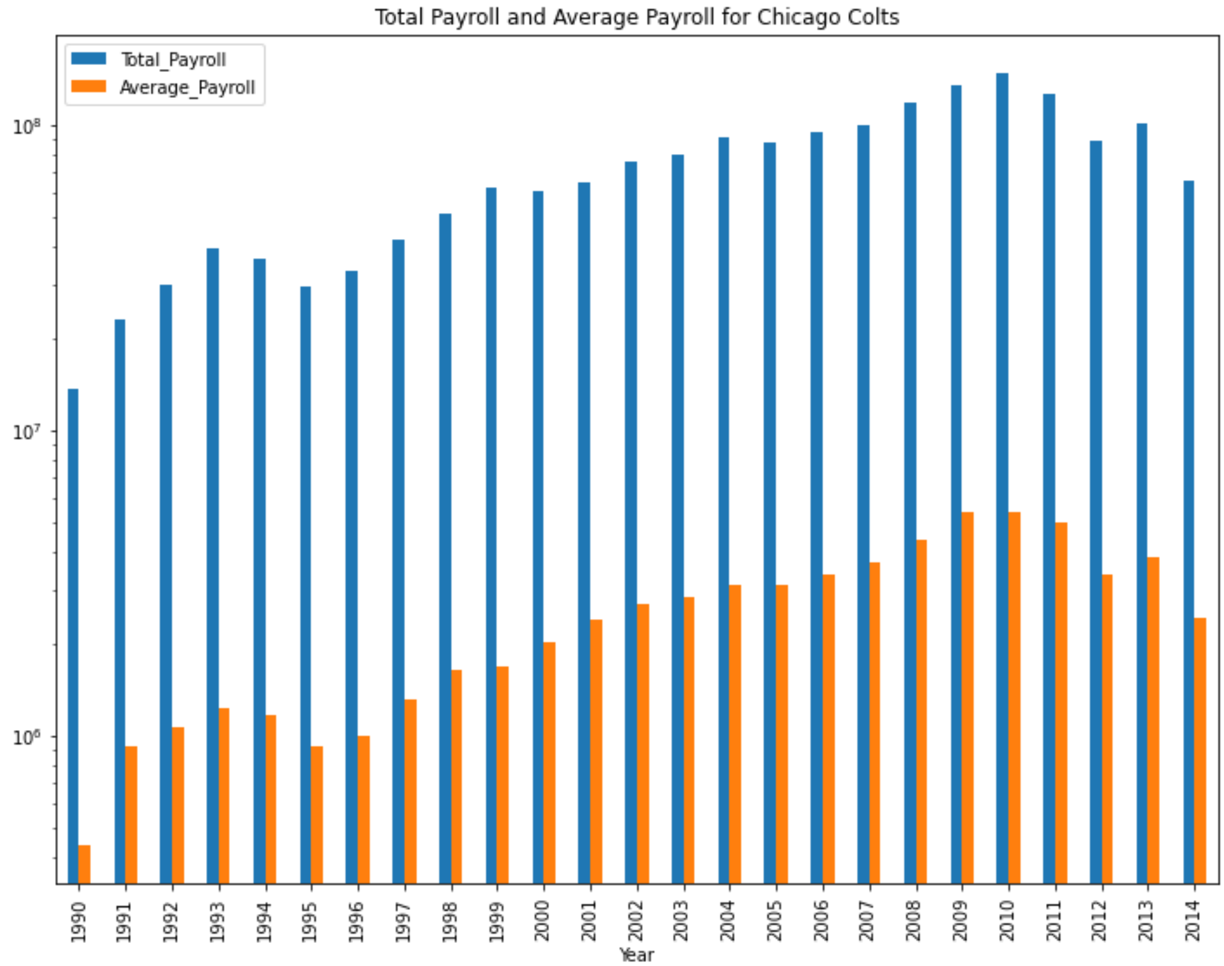

Total Payroll and Average Payroll for Atlanta Braves
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

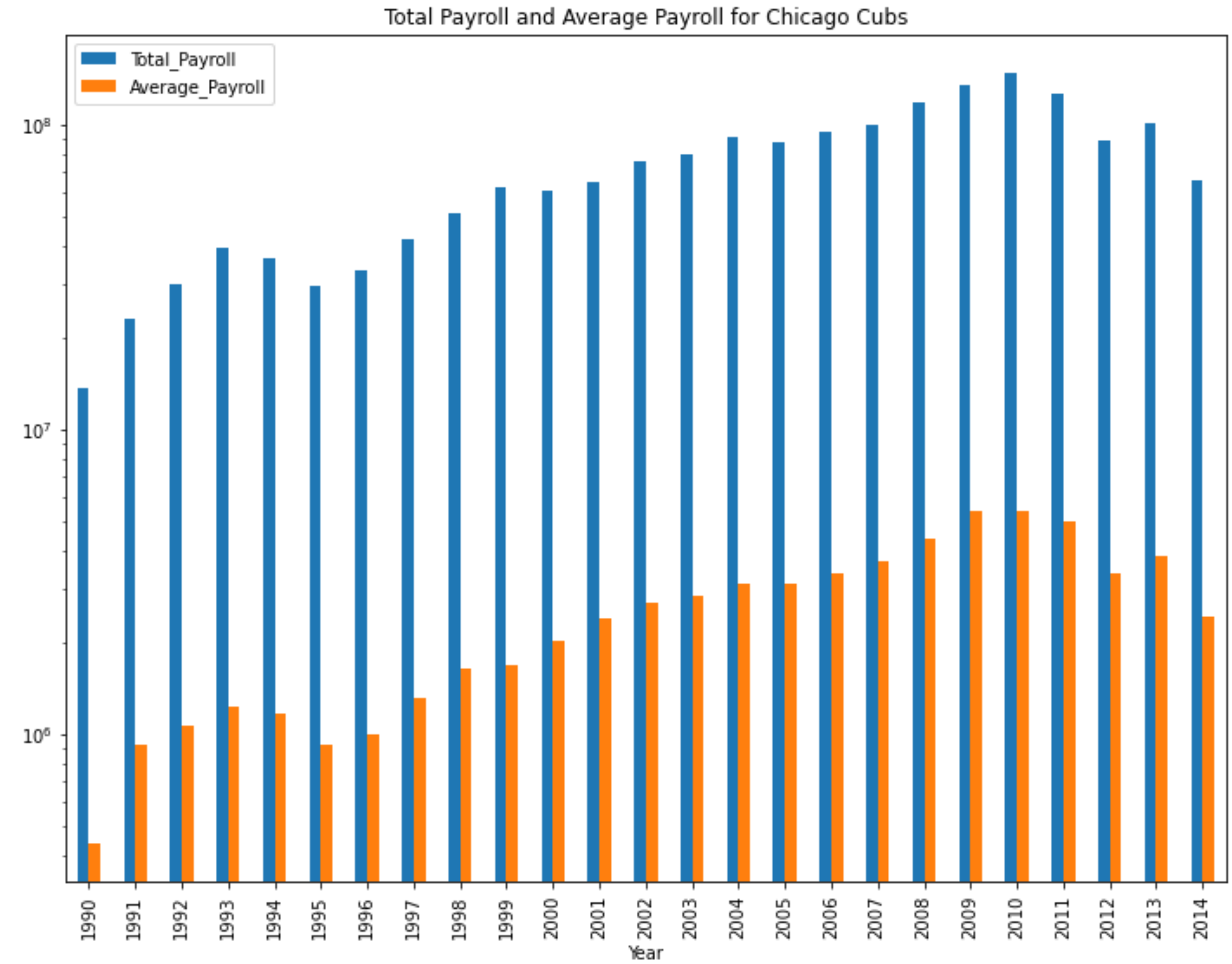

Total Payroll and Average Payroll for Baltimore Orioles
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

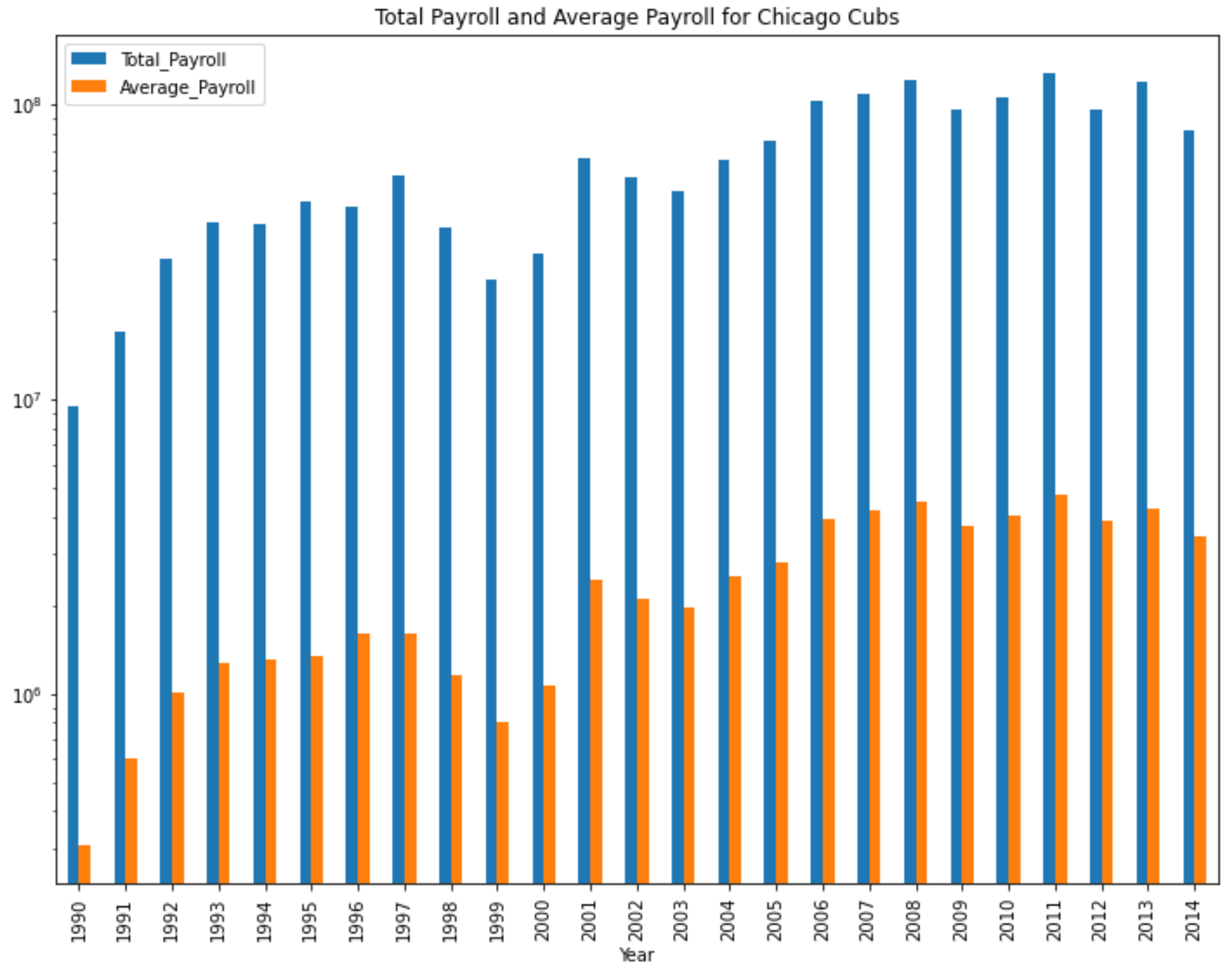

Total Payroll and Average Payroll for Boston Americans
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

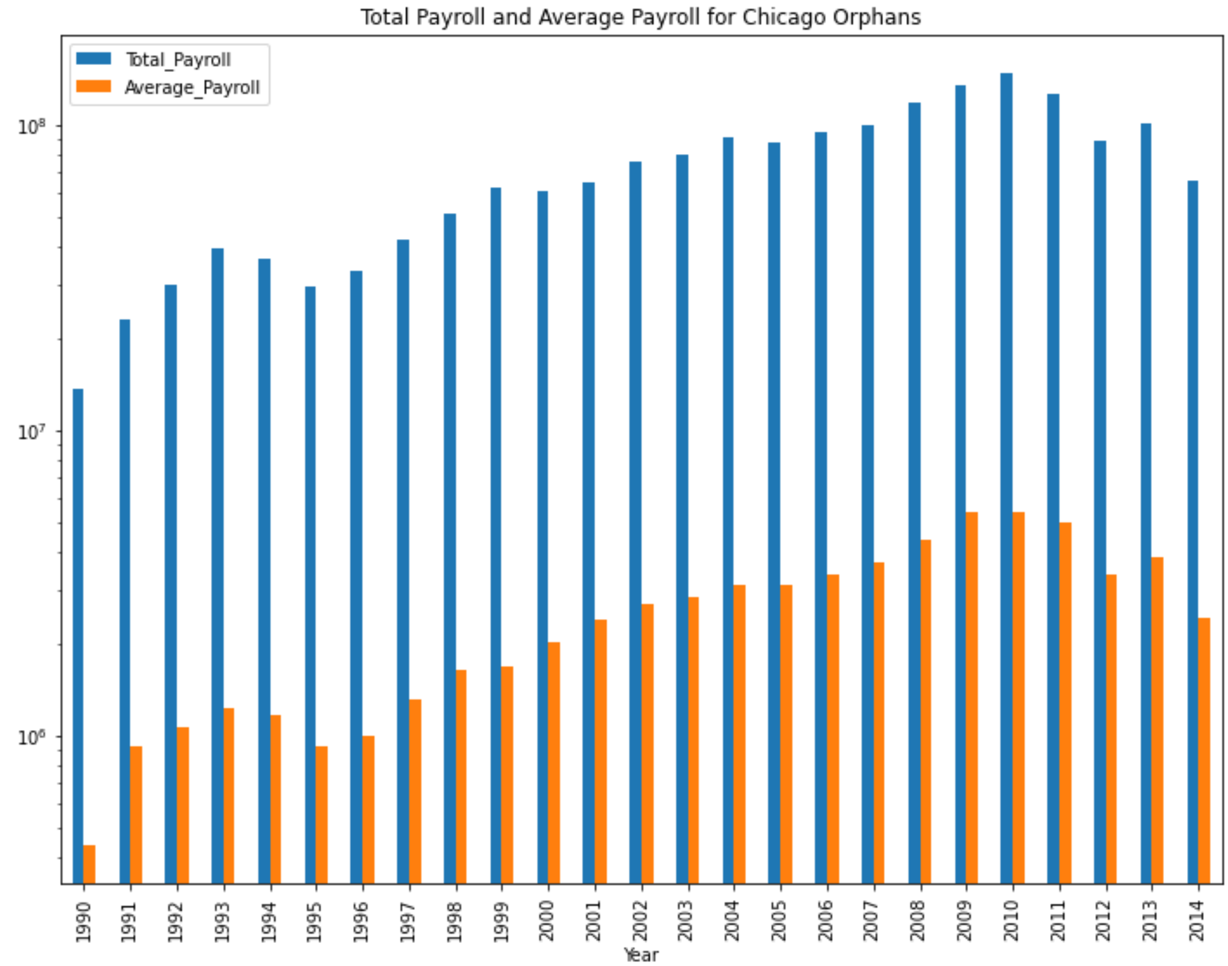

Total Payroll and Average Payroll for Boston Red Sox
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

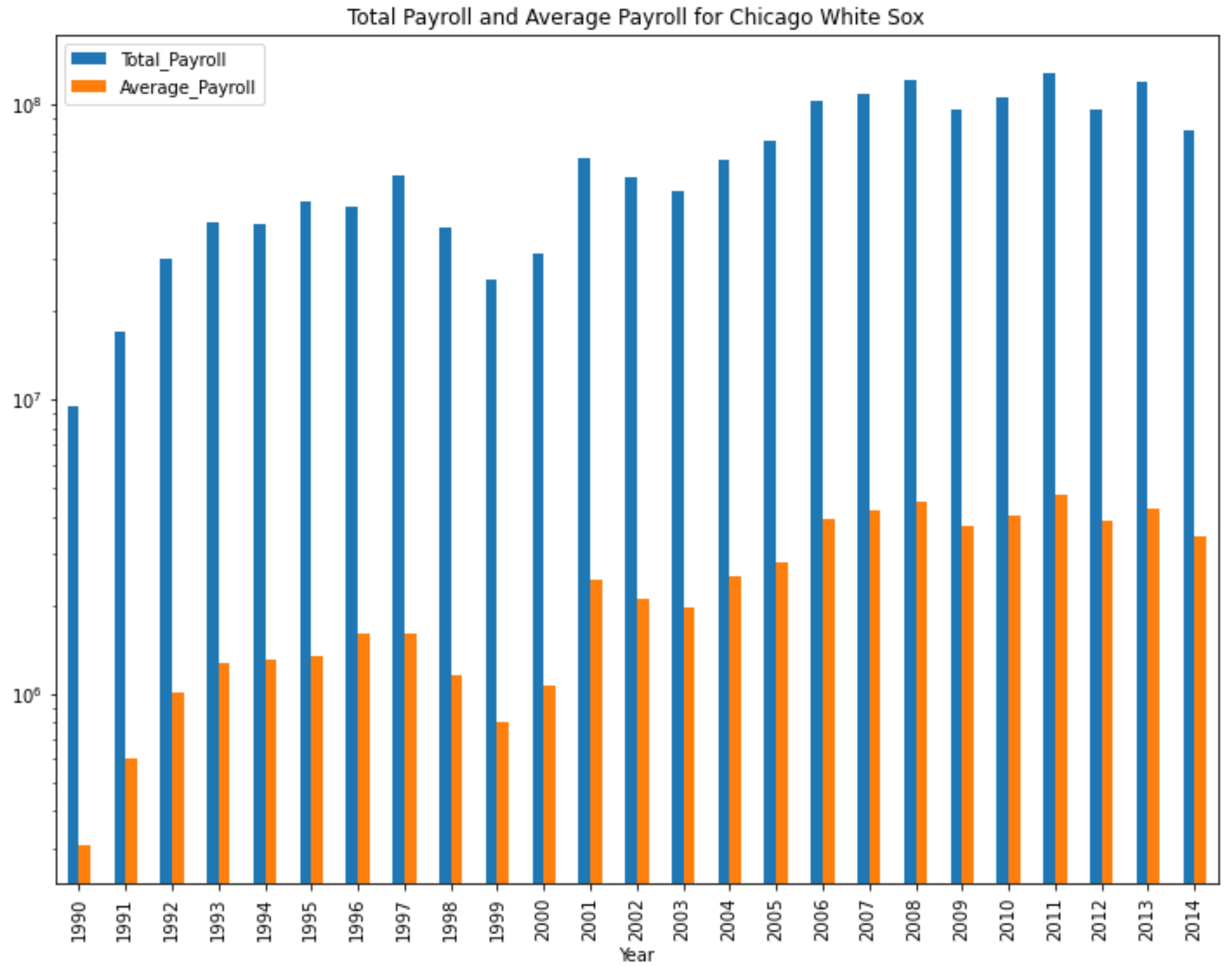

Total Payroll and Average Payroll for California Angels
Total_Payroll
Average_Payroll
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
Year

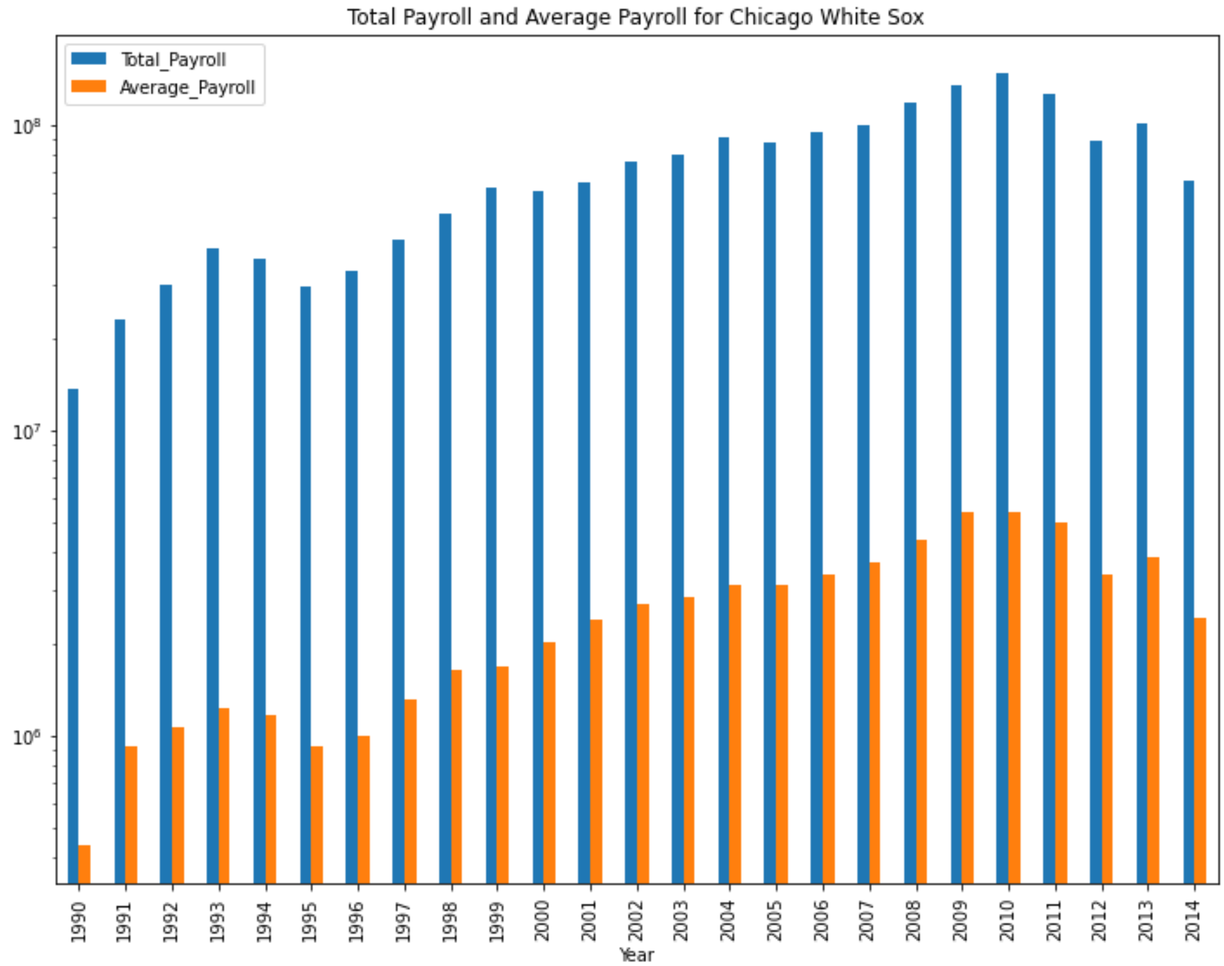

Total Payroll and Average Payroll for Chicago Colts
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014
Year

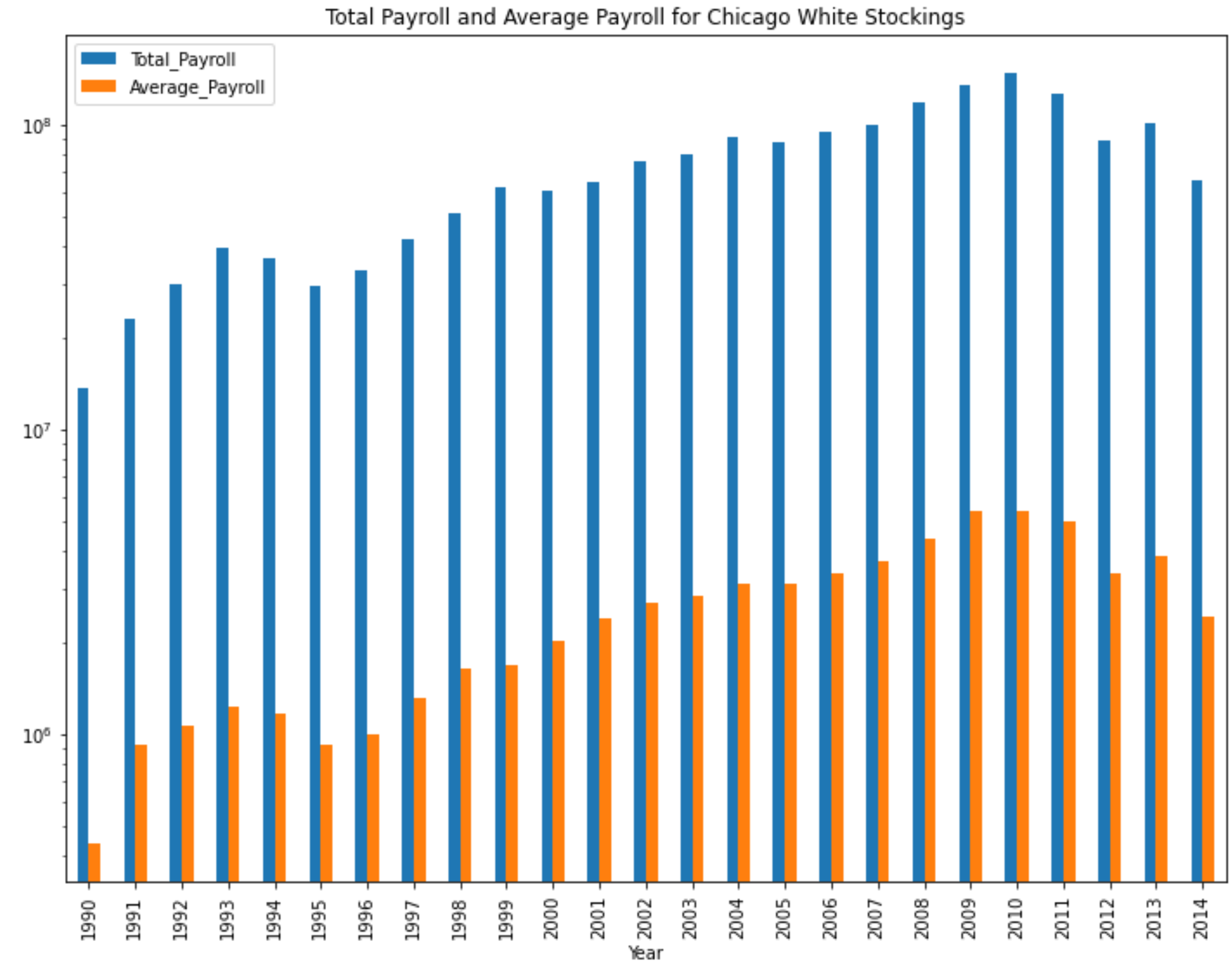

Total Payroll and Average Payroll for Chicago Cubs
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

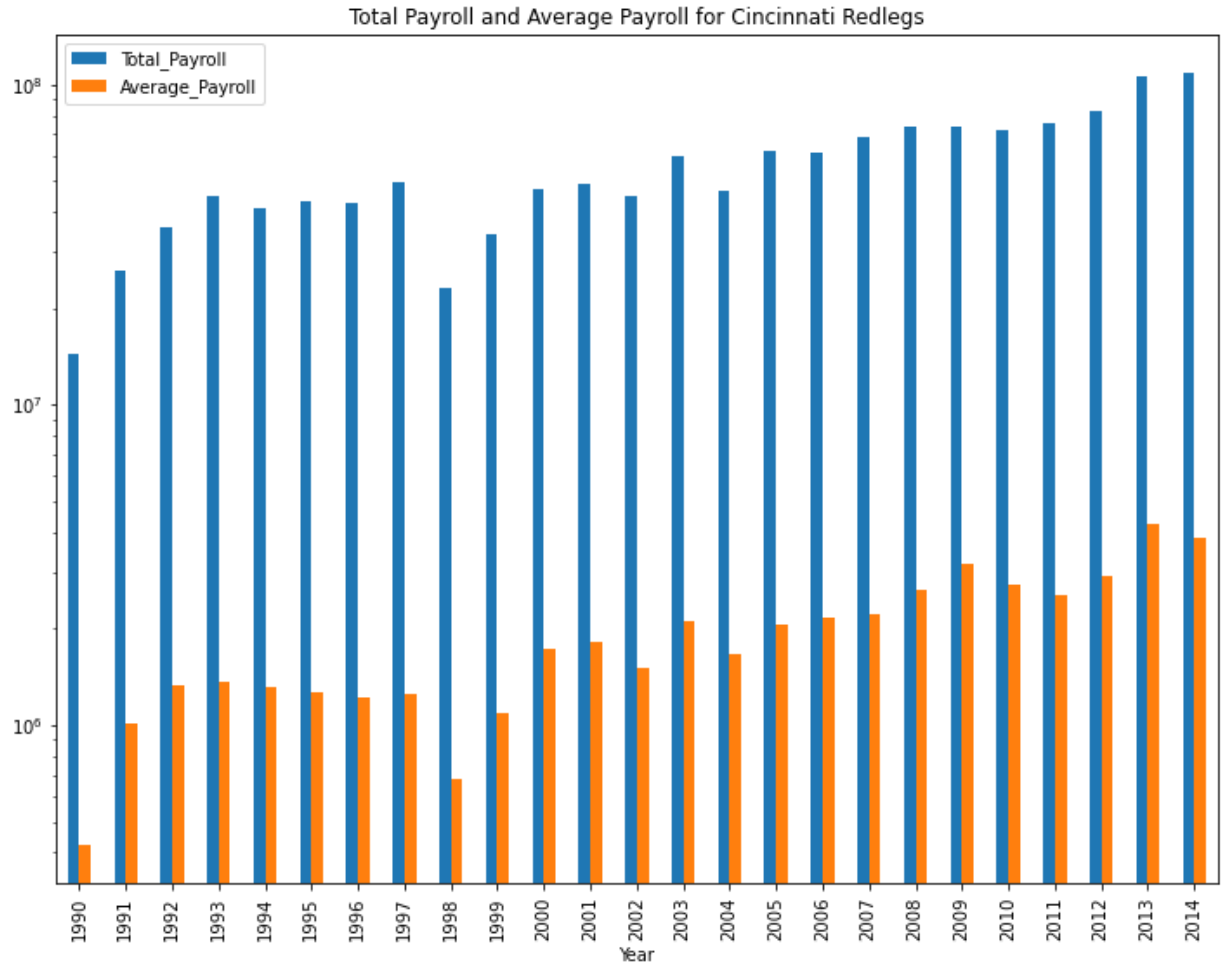

Total Payroll and Average Payroll for Chicago Cubs
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

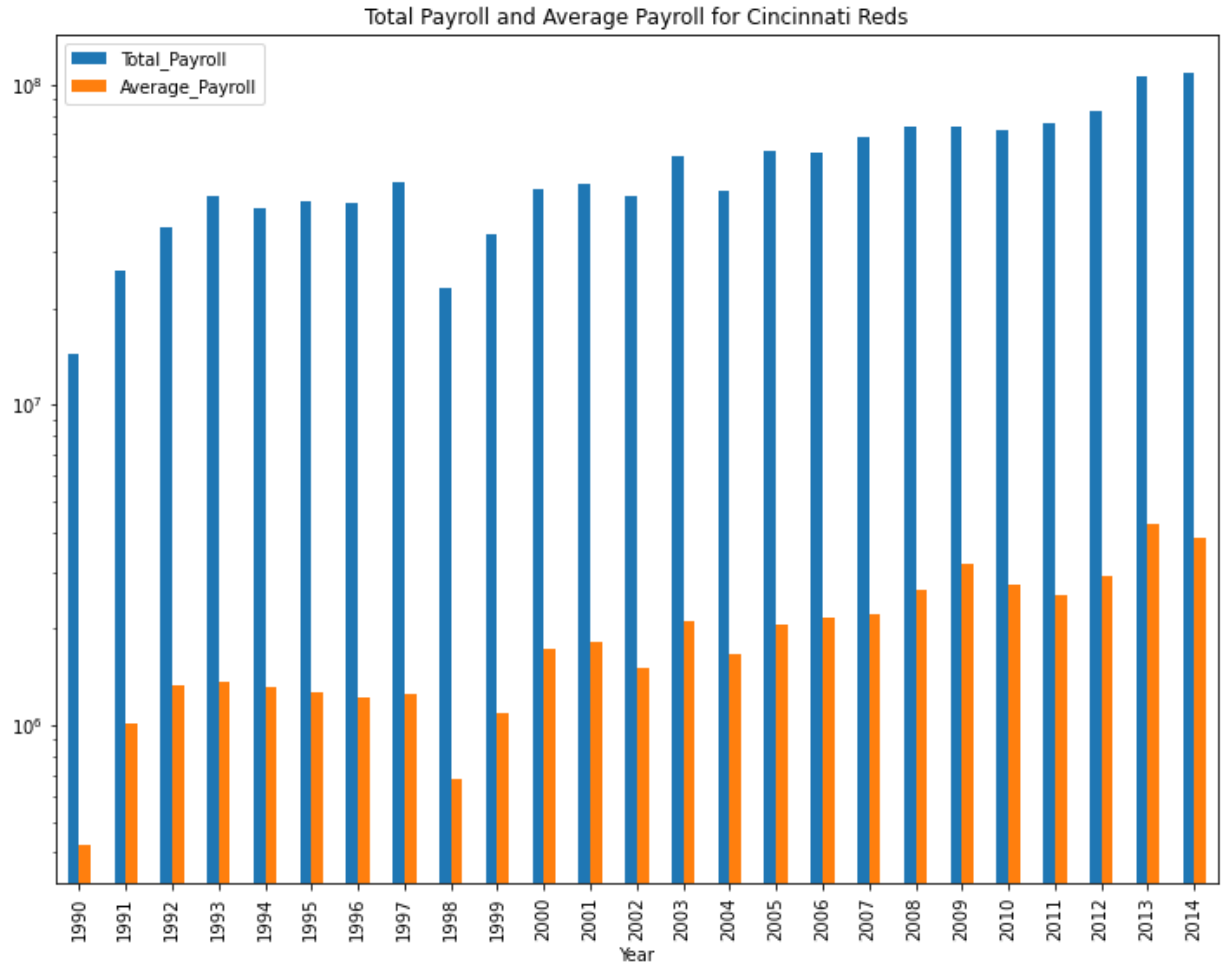

Total Payroll and Average Payroll for Chicago Orphans
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

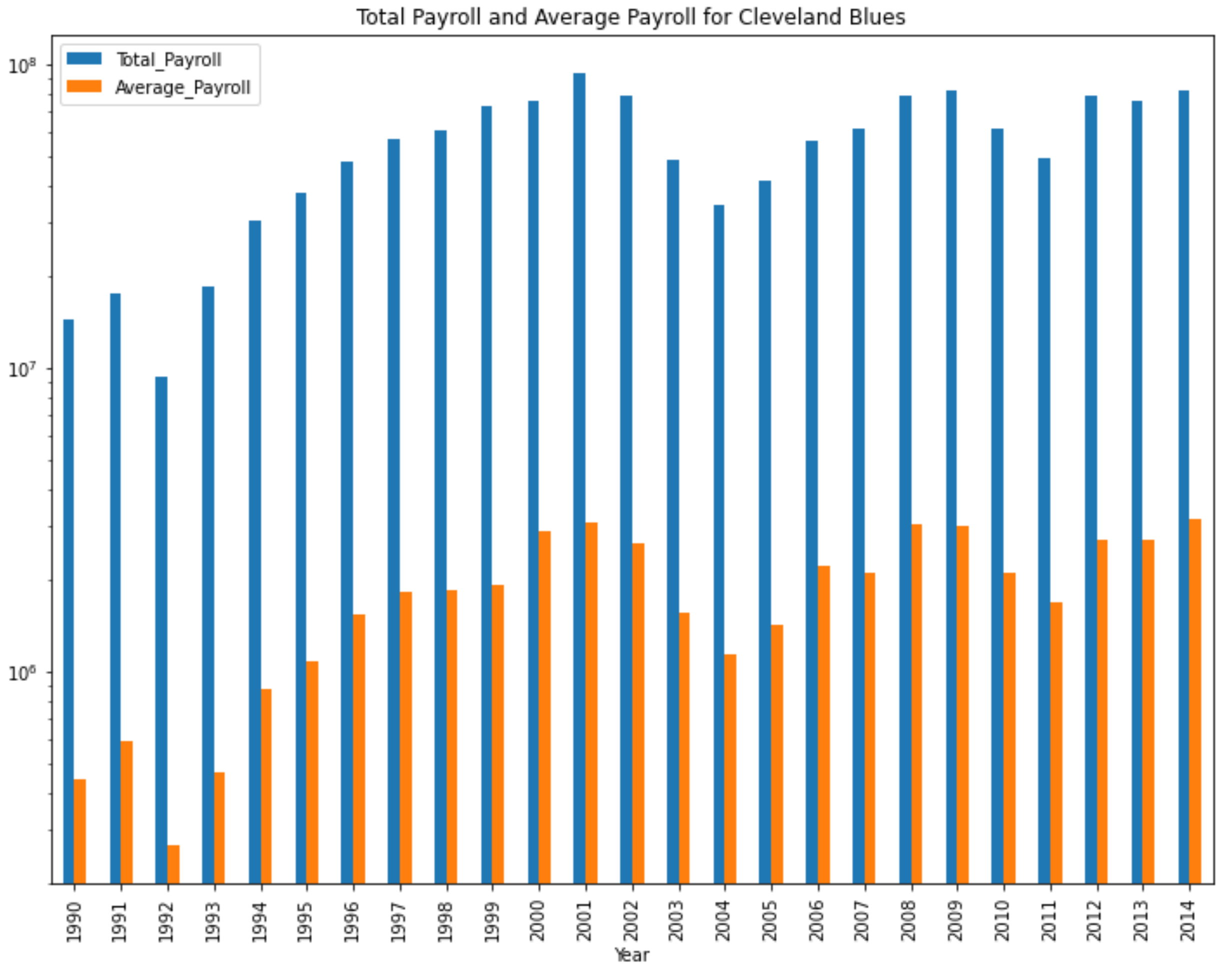

Total Payroll and Average Payroll for Chicago White Sox
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

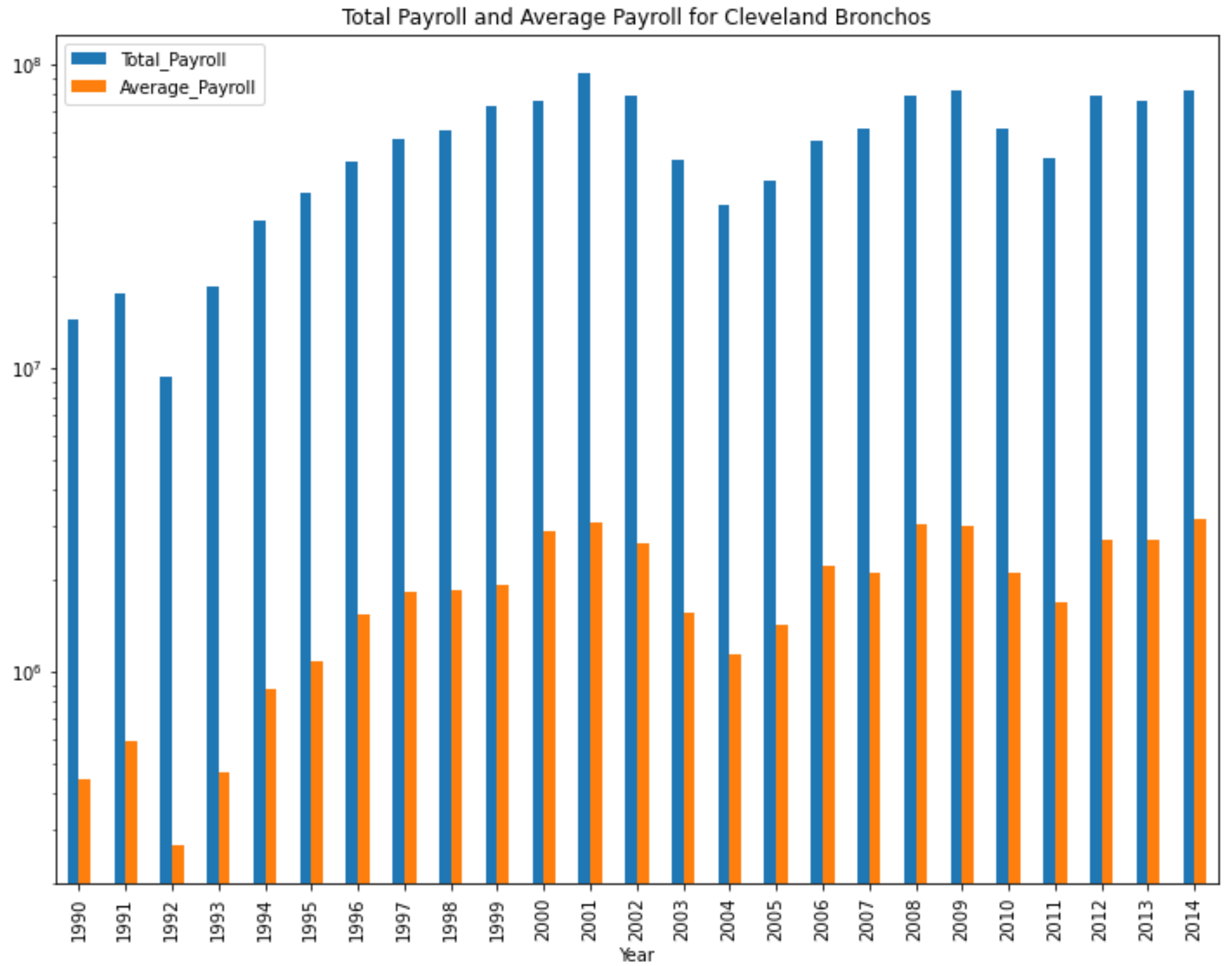

Total Payroll and Average Payroll for Chicago White Sox
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

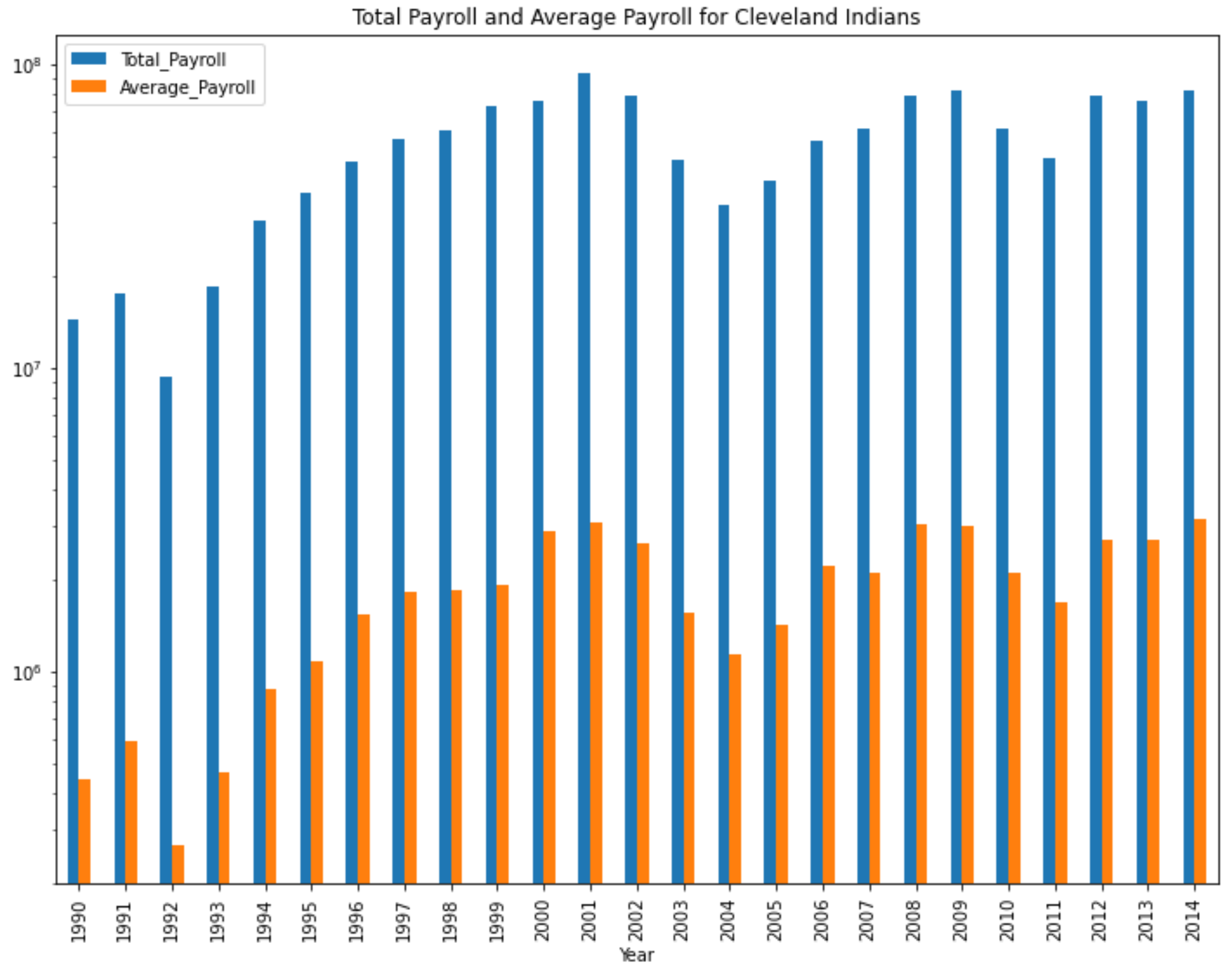

Total Payroll and Average Payroll for Chicago White Stockings
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

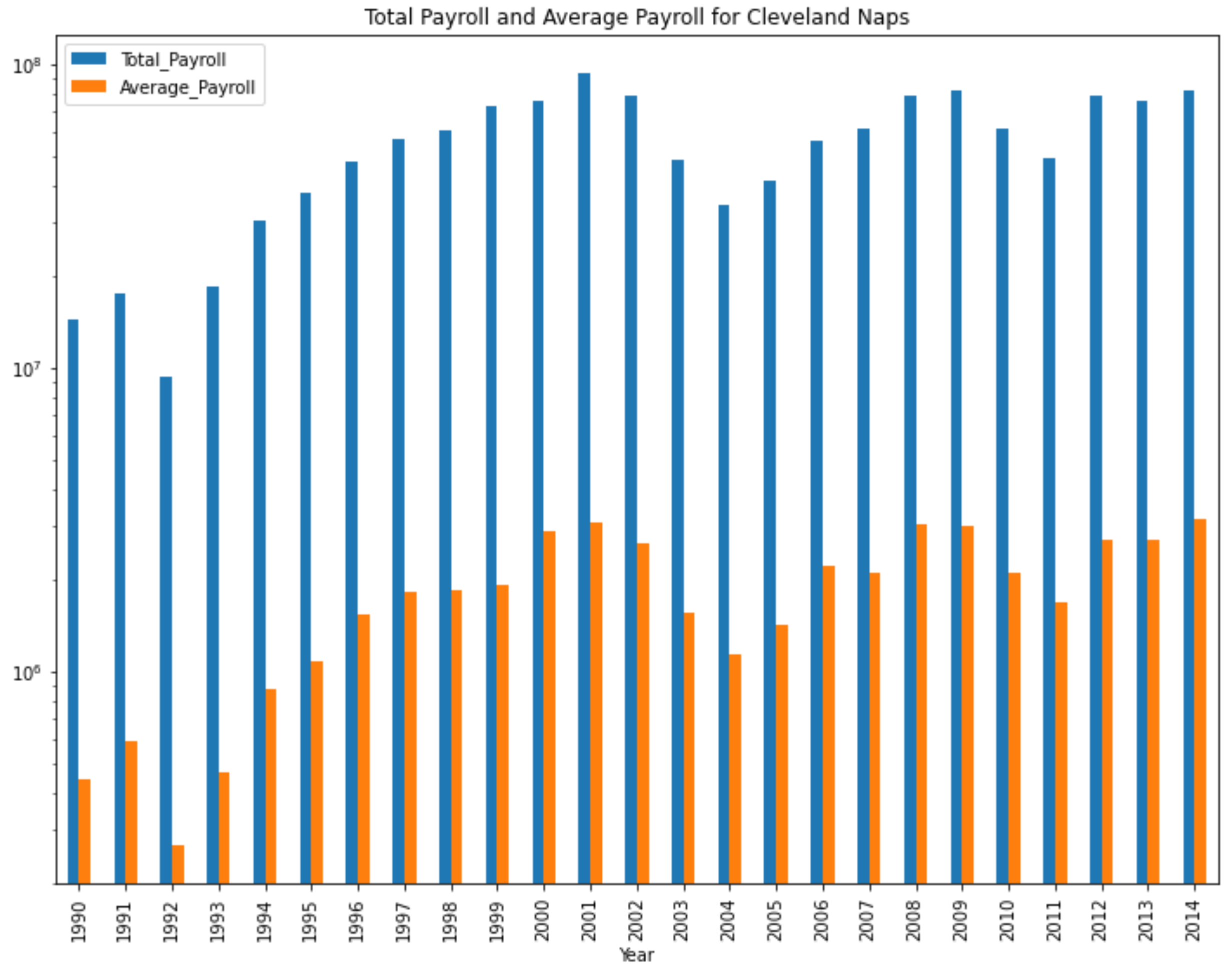

Total Payroll and Average Payroll for Cincinnati Redlegs
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

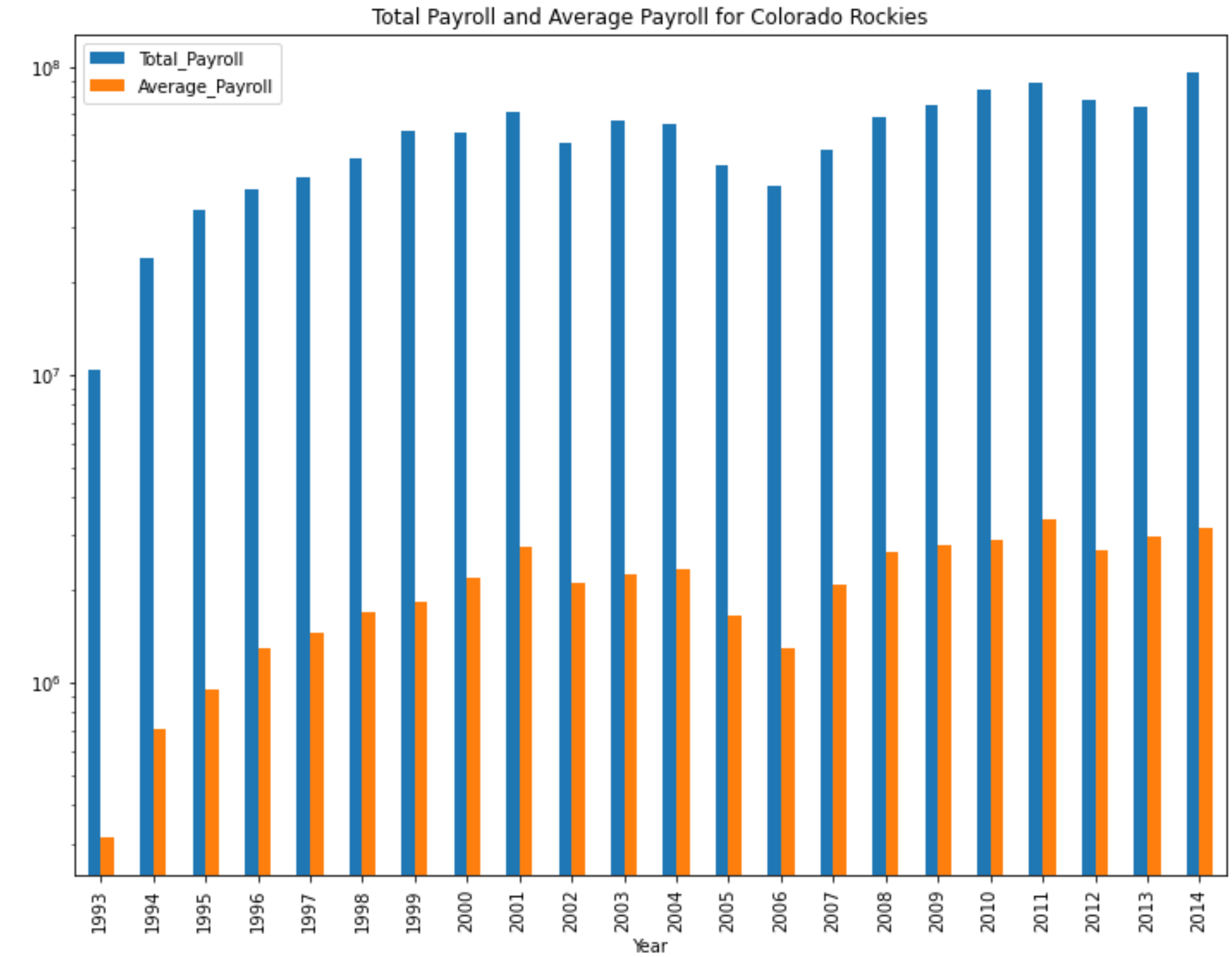

Total Payroll and Average Payroll for Cincinnati Reds
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

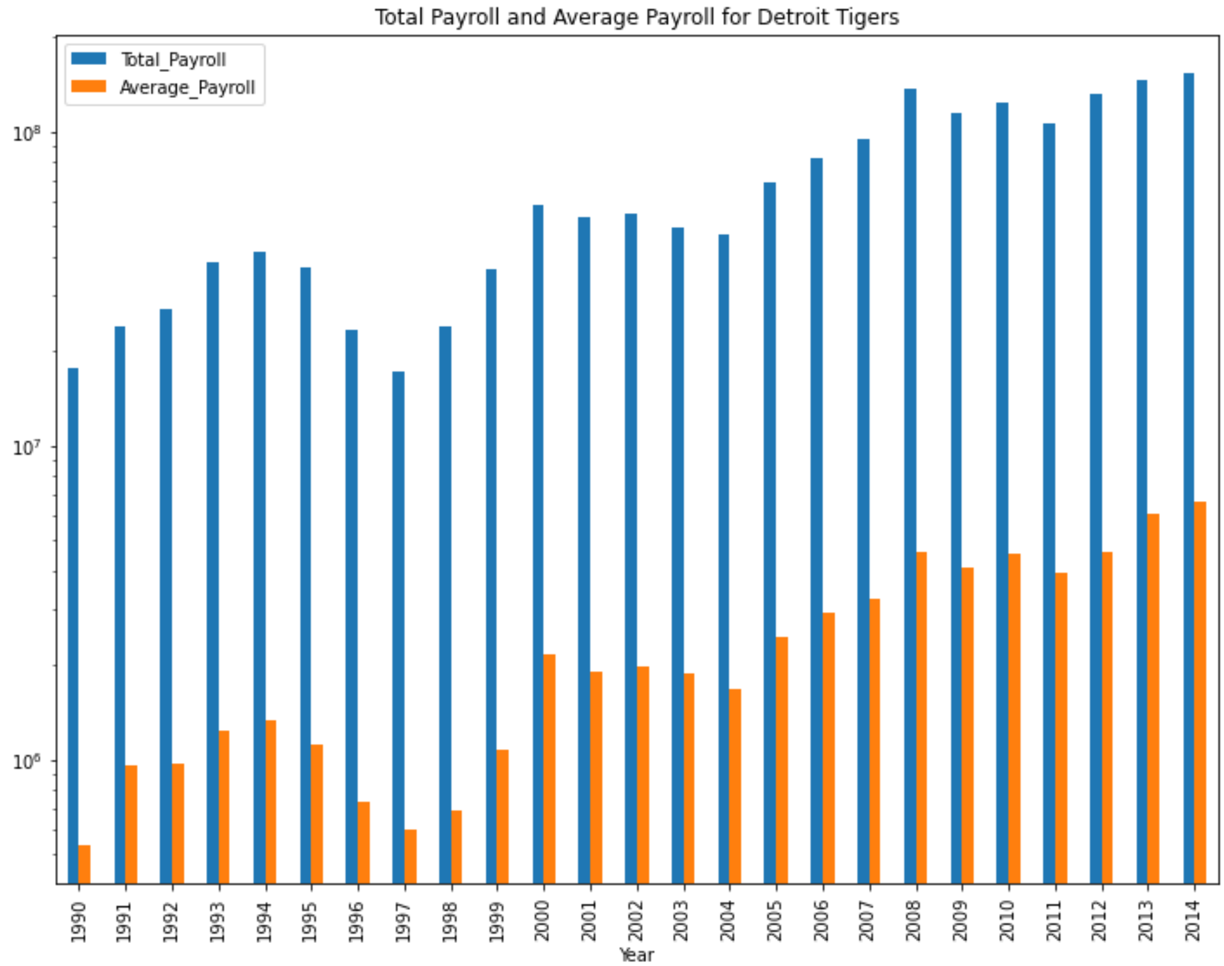

Total Payroll and Average Payroll for Cleveland Blues
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

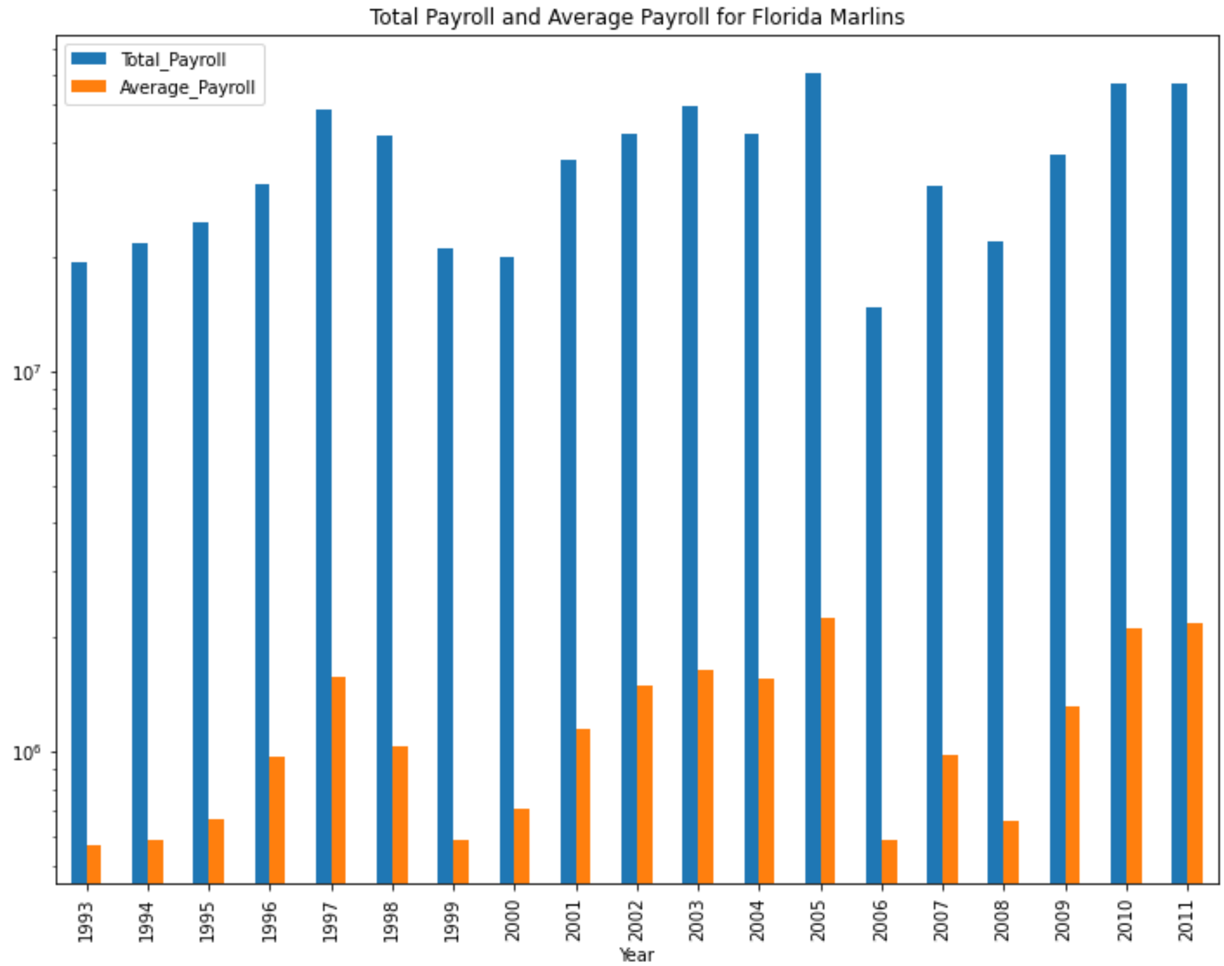

Total Payroll and Average Payroll for Cleveland Bronchos
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

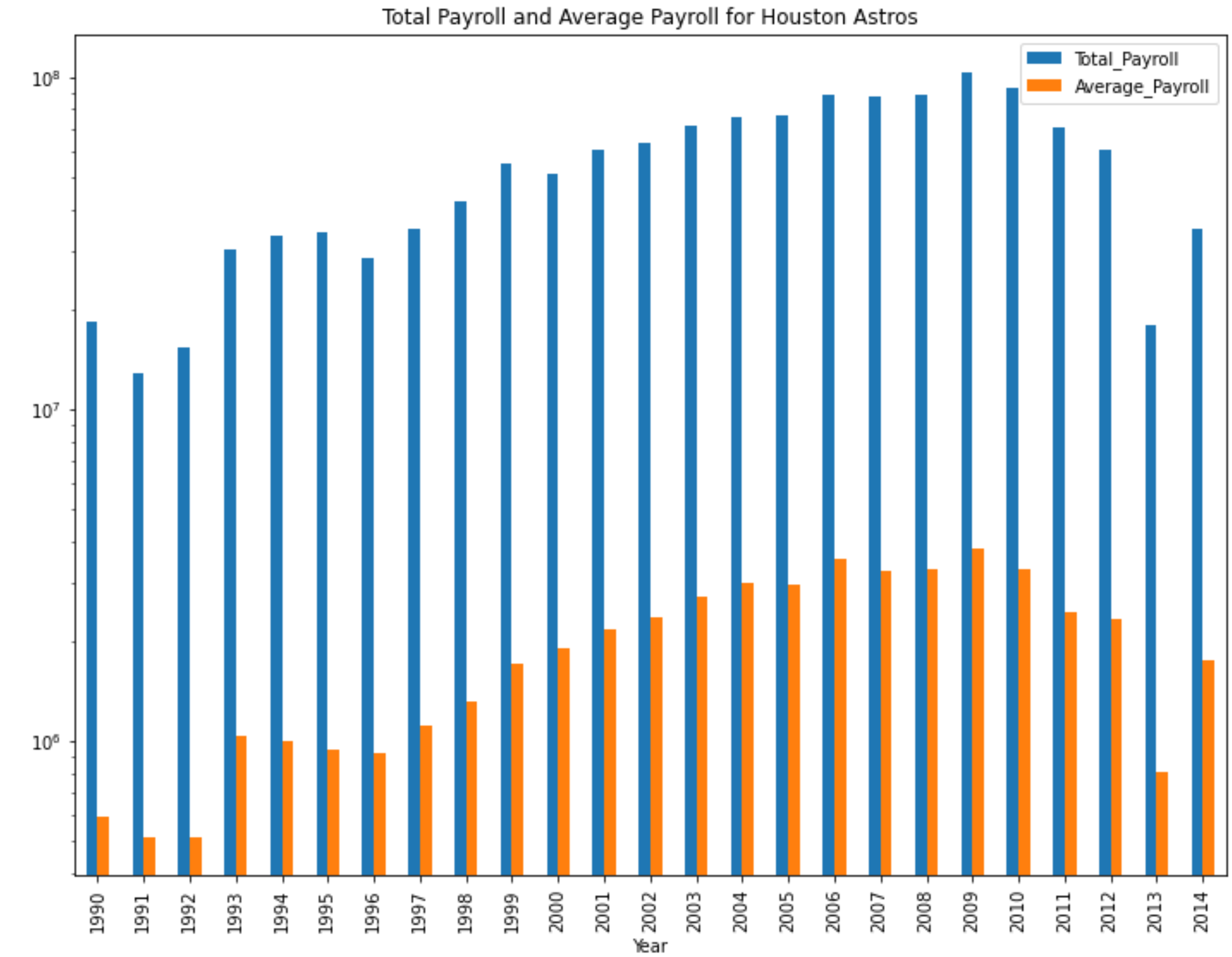

Total Payroll and Average Payroll for Cleveland Indians
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

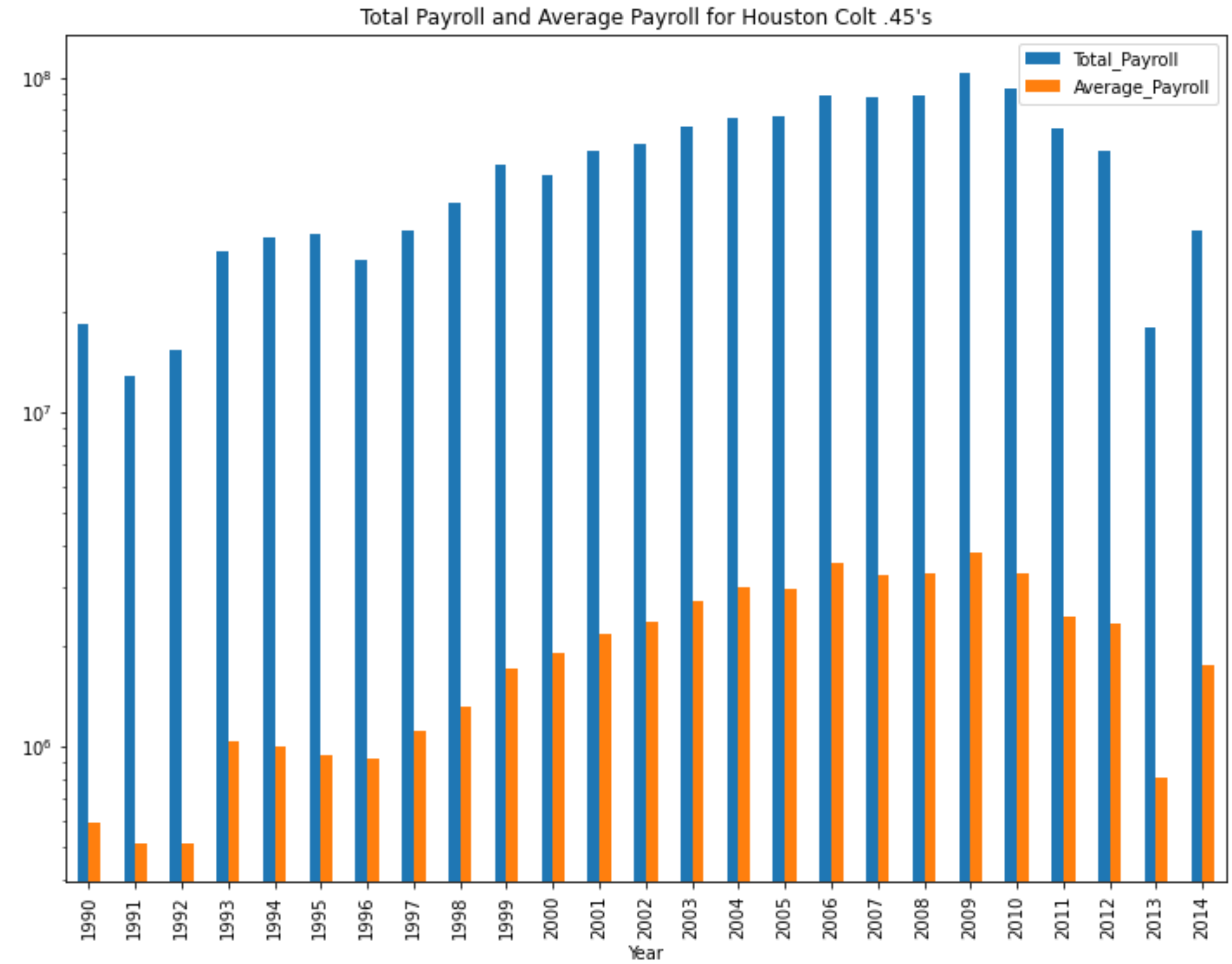

Total Payroll and Average Payroll for Cleveland Naps
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

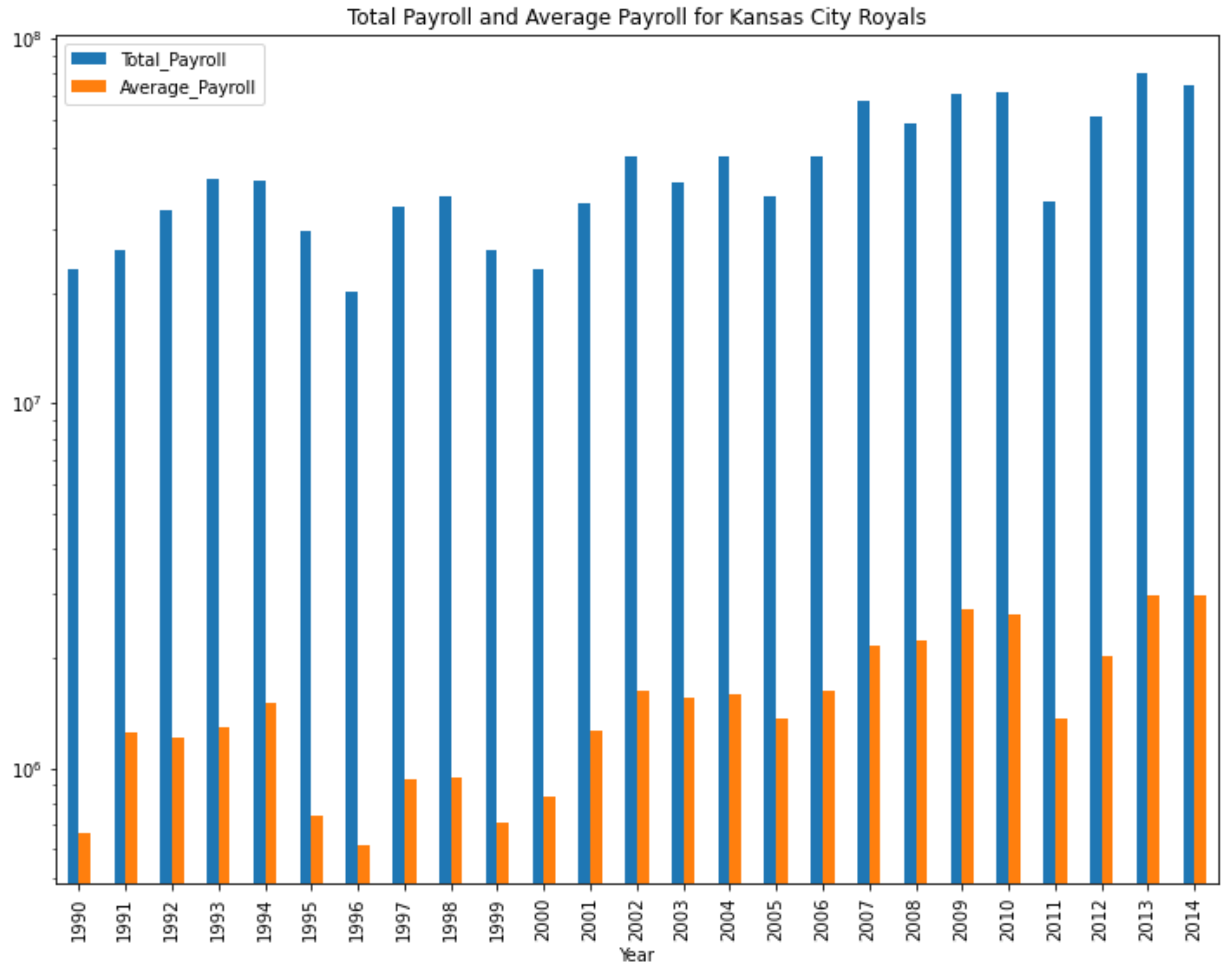

Total Payroll and Average Payroll for Colorado Rockies
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

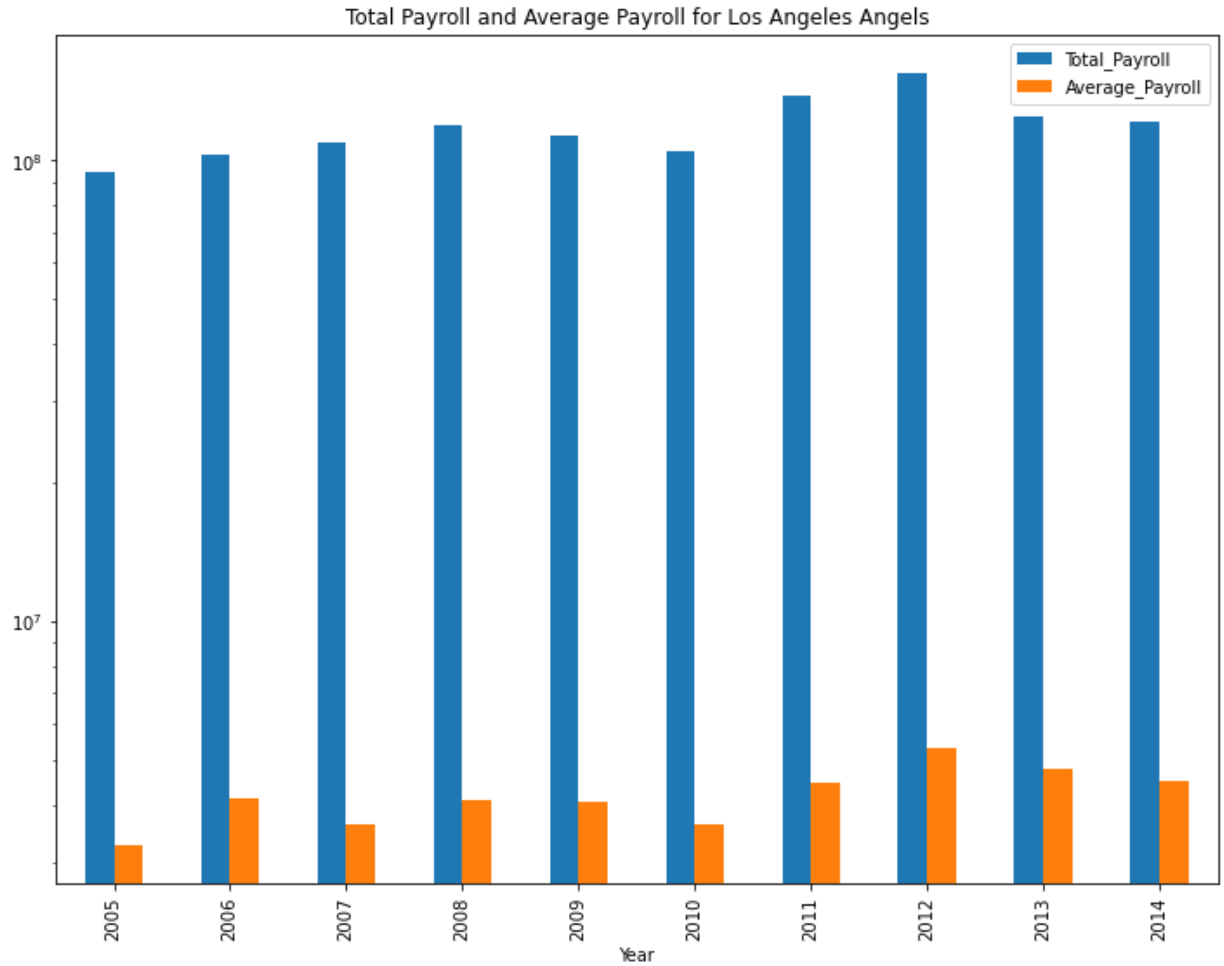

Total Payroll and Average Payroll for Detroit Tigers
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year

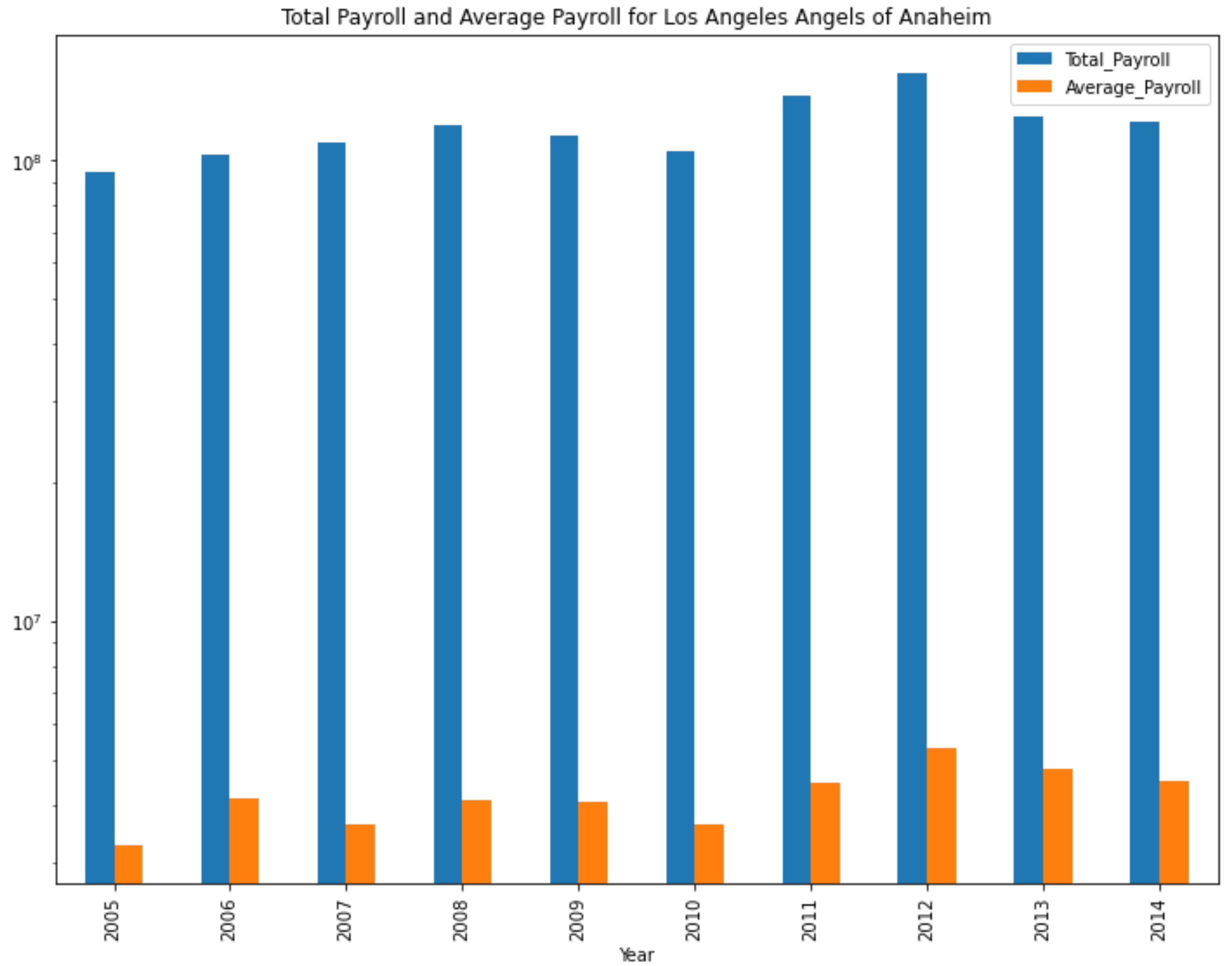

Total Payroll and Average Payroll for Florida Marlins
Total_Payroll
Average_Payroll
10^7
10^6
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
Year

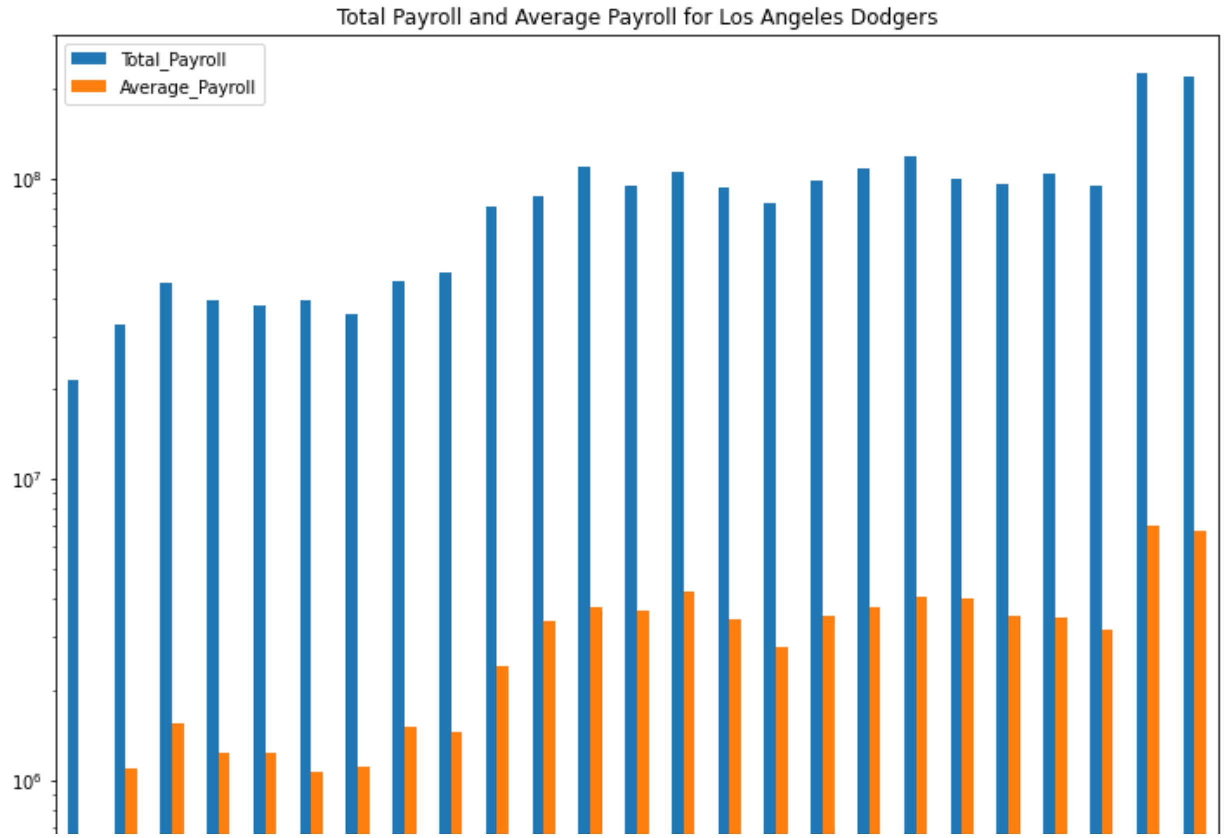

Total Payroll and Average Payroll for Houston Astros
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year


Total Payroll and Average Payroll for Houston Colt .45's
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year


Total Payroll and Average Payroll for Kansas City Royals
Total_Payroll
Average_Payroll
10^8
10^7
10^6
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year


Total Payroll and Average Payroll for Los Angeles Angels
Total_Payroll
Average_Payroll
10^8
10^7
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year


Total Payroll and Average Payroll for Los Angeles Angels of Anaheim
Total_Payroll
Average_Payroll
10^8
10^7
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Year



The total payroll seems to be very high compared to the average payroll. However, the average here shows the averag salary paid to each player.

Correlation between payroll and winning percentage

We will first select the percentage winnings for each team for each year. Then we select the average and total payroll for each team over the years, and then join the results together.

In [418…

```
winning_query = "SELECT teamID, yearID , " + \
        "(SUM(W)*1.0/(SUM(G)*1.0))*100.0 AS `PERCENTAGE WIN` " + \
        "FROM Teams " + \
        "GROUP BY teamID, yearID " + \
        "ORDER BY teamID, yearID "
winning_percentage_result = pandas.read_sql(winning_query, conn)
winning_percentage_result
```

Out[418…

	teamID	yearID	PERCENTAGE WIN
0	ALT	1884	24.000000
1	ANA	1997	51.851852
2	ANA	1998	52.469136
3	ANA	1999	43.209877
4	ANA	2000	50.617284
...	...	...	...
2770	WS8	1887	36.507937
2771	WS8	1888	35.294118

	teamID	yearID	PERCENTAGE WIN
2772	WS8	1889	32.283465
2773	WS9	1891	31.654676
2774	WSU	1884	41.228070

2775 rows × 3 columns

In [419…

```
some_results = []
for index in range(len(team_ids)) :
    team_id = team_ids[index]
    some_query = " SELECT Salaries.yearID, " + \
        "((COUNT(Teams.W)*1.0)/(COUNT(Teams.W)*1.0 + COUNT(Teams.L)*1.0))*100
        "printf('%,d',AVG(Salaries.salary)) AS `AVERAGE PAYROLL` " + \
        "FROM Salaries, Teams " + \
        "WHERE Salaries.teamID = '" + team_id + "' " + \
        "AND Teams.teamID = '" + team_id + "' " + \
        "AND Teams.teamID = Salaries.teamID " + \
        "GROUP BY Salaries.yearID " + \
        "ORDER BY Salaries.teamID, Salaries.yearID "
    v = pandas.read_sql(some_query, conn)
    some_results.append(v)

some_results
```

Out[419…

```
[Empty DataFrame
 Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
 Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
 0    1997                     50.0       1,004,370
 1    1998                     50.0       1,214,147
 2    1999                     50.0       1,384,704
 3    2000                     50.0       1,715,472
 4    2001                     50.0       1,584,505
 5    2002                     50.0       2,204,345
 6    2003                     50.0       2,927,098
 7    2004                     50.0       3,723,506,
     yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
 0     1998                     50.0         898,527
 1     1999                     50.0       2,020,705
 2     2000                     50.0       2,893,851
 3     2001                     50.0       3,038,678
 4     2002                     50.0       3,115,757
 5     2003                     50.0       3,226,280
 6     2004                     50.0       2,406,232
 7     2005                     50.0       2,308,487
 8     2006                     50.0       2,295,547
 9     2007                     50.0       1,859,555
 10    2008                     50.0       2,364,382
 11    2009                     50.0       2,812,141
 12    2010                     50.0       2,335,314
 13    2011                     50.0       1,986,660
 14    2012                     50.0       2,733,512
 15    2013                     50.0       3,004,400
 16    2014                     50.0       3,763,903,
     yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
 0     1985                     50.0         673,045
```

```
1     1986                     50.0         589,751
2     1987                     50.0         517,017
3     1988                     50.0         438,902
4     1989                     50.0         370,411
5     1990                     50.0         454,859
6     1991                     50.0         736,140
7     1992                     50.0       1,116,946
8     1993                     50.0       1,261,861
9     1994                     50.0       1,646,117
10    1995                     50.0       1,628,808
11    1996                     50.0       1,656,616
12    1997                     50.0       1,686,403
13    1998                     50.0       1,912,062
14    1999                     50.0       2,522,068
15    2000                     50.0       2,817,927
16    2001                     50.0       2,965,682
17    2002                     50.0       3,316,798
18    2003                     50.0       3,934,950
19    2004                     50.0       3,220,803
20    2005                     50.0       3,458,292
21    2006                     50.0       3,108,857
22    2007                     50.0       3,117,529
23    2008                     50.0       3,412,189
24    2009                     50.0       3,335,385
25    2010                     50.0       3,126,802
26    2011                     50.0       3,346,257
27    2012                     50.0       2,856,204
28    2013                     50.0       3,254,500
29    2014                     50.0       4,067,041,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         525,486
1     1986                     50.0         448,319
2     1987                     50.0         463,342
3     1988                     50.0         501,187
4     1989                     50.0         318,275
5     1990                     50.0         261,623
6     1991                     50.0         565,129
7     1992                     50.0         720,626
8     1993                     50.0         909,265
```

```
9     1994                        50.0       1,253,218
10    1995                        50.0       1,187,635
11    1996                        50.0       1,702,822
12    1997                        50.0       2,017,806
13    1998                        50.0       2,411,854
14    1999                        50.0       2,303,024
15    2000                        50.0       2,808,532
16    2001                        50.0       2,331,018
17    2002                        50.0       1,890,421
18    2003                        50.0       2,547,500
19    2004                        50.0       1,843,690
20    2005                        50.0       2,639,797
21    2006                        50.0       2,592,342
22    2007                        50.0       3,450,918
23    2008                        50.0       2,099,882
24    2009                        50.0       2,580,833
25    2010                        50.0       3,138,942
26    2011                        50.0       3,280,924
27    2012                        50.0       2,762,642
28    2013                        50.0       3,245,897
29    2014                        50.0       3,693,428,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                        50.0         435,902
1     1986                        50.0         496,628
2     1987                        50.0         676,277
3     1988                        50.0         579,003
4     1989                        50.0         672,374
5     1990                        50.0         642,447
6     1991                        50.0       1,172,250
7     1992                        50.0       1,406,793
8     1993                        50.0       1,197,438
9     1994                        50.0       1,261,969
10    1995                        50.0         877,176
11    1996                        50.0       1,177,597
12    1997                        50.0       1,281,139
13    1998                        50.0       1,719,909
14    1999                        50.0       2,048,306
15    2000                        50.0       2,598,011
16    2001                        50.0       3,438,619
17    2002                        50.0       3,612,202
18    2003                        50.0       3,701,722
19    2004                        50.0       4,243,283
20    2005                        50.0       4,410,897
21    2006                        50.0       4,448,141
22    2007                        50.0       5,108,079
23    2008                        50.0       4,763,929
24    2009                        50.0       4,184,344
25    2010                        50.0       5,601,632
26    2011                        50.0       5,991,202
27    2012                        50.0       5,093,724
28    2013                        50.0       5,225,172
29    2014                        50.0       4,484,513,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
```

```
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         435,902
1     1986                     50.0         496,628
2     1987                     50.0         676,277
3     1988                     50.0         579,003
4     1989                     50.0         672,374
5     1990                     50.0         642,447
6     1991                     50.0       1,172,250
7     1992                     50.0       1,406,793
8     1993                     50.0       1,197,438
9     1994                     50.0       1,261,969
10    1995                     50.0         877,176
11    1996                     50.0       1,177,597
12    1997                     50.0       1,281,139
13    1998                     50.0       1,719,909
14    1999                     50.0       2,048,306
15    2000                     50.0       2,598,011
16    2001                     50.0       3,438,619
17    2002                     50.0       3,612,202
18    2003                     50.0       3,701,722
19    2004                     50.0       4,243,283
20    2005                     50.0       4,410,897
21    2006                     50.0       4,448,141
22    2007                     50.0       5,108,079
23    2008                     50.0       4,763,929
24    2009                     50.0       4,184,344
25    2010                     50.0       5,601,632
26    2011                     50.0       5,991,202
27    2012                     50.0       5,093,724
28    2013                     50.0       5,225,172
29    2014                     50.0       4,484,513,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
```

```
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         515,281
1     1986                     50.0         497,491
2     1987                     50.0         475,685
3     1988                     50.0         426,692
4     1989                     50.0         580,685
5     1990                     50.0         620,571
6     1991                     50.0       1,066,451
7     1992                     50.0       1,158,311
8     1993                     50.0         893,385
9     1994                     50.0         786,131
10    1995                     50.0         946,156
11    1996                     50.0         776,702,
```

```
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         577,405
1     1986                     50.0         555,102
2     1987                     50.0         550,307
3     1988                     50.0         524,767
4     1989                     50.0         444,500
5     1990                     50.0         439,483
6     1991                     50.0         927,026
7     1992                     50.0       1,065,345
8     1993                     50.0       1,230,833
9     1994                     50.0       1,170,559
10    1995                     50.0         922,057
11    1996                     50.0       1,002,454
12    1997                     50.0       1,317,354
13    1998                     50.0       1,639,935
14    1999                     50.0       1,684,945
15    2000                     50.0       2,017,977
16    2001                     50.0       2,396,882
17    2002                     50.0       2,703,244
18    2003                     50.0       2,852,440
19    2004                     50.0       3,122,758
20    2005                     50.0       3,108,319
21    2006                     50.0       3,372,303
22    2007                     50.0       3,691,493
23    2008                     50.0       4,383,179
24    2009                     50.0       5,392,360
25    2010                     50.0       5,429,962
26    2011                     50.0       5,001,893
27    2012                     50.0       3,392,193
28    2013                     50.0       3,867,989
29    2014                     50.0       2,426,759,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         577,405
1     1986                     50.0         555,102
2     1987                     50.0         550,307
3     1988                     50.0         524,767
4     1989                     50.0         444,500
5     1990                     50.0         439,483
6     1991                     50.0         927,026
7     1992                     50.0       1,065,345
8     1993                     50.0       1,230,833
9     1994                     50.0       1,170,559
10    1995                     50.0         922,057
11    1996                     50.0       1,002,454
12    1997                     50.0       1,317,354
13    1998                     50.0       1,639,935
14    1999                     50.0       1,684,945
15    2000                     50.0       2,017,977
16    2001                     50.0       2,396,882
17    2002                     50.0       2,703,244
18    2003                     50.0       2,852,440
19    2004                     50.0       3,122,758
20    2005                     50.0       3,108,319
21    2006                     50.0       3,372,303
22    2007                     50.0       3,691,493
23    2008                     50.0       4,383,179
24    2009                     50.0       5,392,360
```

```
25    2010                     50.0       5,429,962
26    2011                     50.0       5,001,893
27    2012                     50.0       3,392,193
28    2013                     50.0       3,867,989
29    2014                     50.0       2,426,759,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         468,865
1     1986                     50.0         336,090
2     1987                     50.0         443,410
3     1988                     50.0         266,250
4     1989                     50.0         290,616
5     1990                     50.0         306,177
6     1991                     50.0         604,273
7     1992                     50.0       1,005,361
8     1993                     50.0       1,280,521
9     1994                     50.0       1,306,127
10    1995                     50.0       1,341,750
11    1996                     50.0       1,612,125
12    1997                     50.0       1,603,888
13    1998                     50.0       1,161,666
14    1999                     50.0         800,625
15    2000                     50.0       1,073,568
16    2001                     50.0       2,431,617
17    2002                     50.0       2,113,067
18    2003                     50.0       1,961,923
19    2004                     50.0       2,508,173
20    2005                     50.0       2,784,370
21    2006                     50.0       3,951,948
22    2007                     50.0       4,179,685
23    2008                     50.0       4,488,493
24    2009                     50.0       3,694,942
25    2010                     50.0       4,058,846
26    2011                     50.0       4,732,925
27    2012                     50.0       3,876,780
28    2013                     50.0       4,288,045
29    2014                     50.0       3,409,604,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         577,405
1     1986                     50.0         555,102
2     1987                     50.0         550,307
3     1988                     50.0         524,767
4     1989                     50.0         444,500
5     1990                     50.0         439,483
6     1991                     50.0         927,026
7     1992                     50.0       1,065,345
8     1993                     50.0       1,230,833
9     1994                     50.0       1,170,559
10    1995                     50.0         922,057
11    1996                     50.0       1,002,454
12    1997                     50.0       1,317,354
13    1998                     50.0       1,639,935
14    1999                     50.0       1,684,945
15    2000                     50.0       2,017,977
16    2001                     50.0       2,396,882
17    2002                     50.0       2,703,244
18    2003                     50.0       2,852,440
19    2004                     50.0       3,122,758
20    2005                     50.0       3,108,319
21    2006                     50.0       3,372,303
22    2007                     50.0       3,691,493
```

```
23    2008                        50.0      4,383,179
24    2009                        50.0      5,392,360
25    2010                        50.0      5,429,962
26    2011                        50.0      5,001,893
27    2012                        50.0      3,392,193
28    2013                        50.0      3,867,989
29    2014                        50.0      2,426,759,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         468,865
1     1986                     50.0         336,090
2     1987                     50.0         443,410
3     1988                     50.0         266,250
4     1989                     50.0         290,616
5     1990                     50.0         306,177
6     1991                     50.0         604,273
7     1992                     50.0       1,005,361
8     1993                     50.0       1,280,521
9     1994                     50.0       1,306,127
10    1995                     50.0       1,341,750
11    1996                     50.0       1,612,125
12    1997                     50.0       1,603,888
13    1998                     50.0       1,161,666
14    1999                     50.0         800,625
15    2000                     50.0       1,073,568
16    2001                     50.0       2,431,617
17    2002                     50.0       2,113,067
18    2003                     50.0       1,961,923
19    2004                     50.0       2,508,173
20    2005                     50.0       2,784,370
21    2006                     50.0       3,951,948
22    2007                     50.0       4,179,685
23    2008                     50.0       4,488,493
24    2009                     50.0       3,694,942
25    2010                     50.0       4,058,846
26    2011                     50.0       4,732,925
27    2012                     50.0       3,876,780
28    2013                     50.0       4,288,045
29    2014                     50.0       3,409,604,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         577,405
1     1986                     50.0         555,102
2     1987                     50.0         550,307
3     1988                     50.0         524,767
4     1989                     50.0         444,500
5     1990                     50.0         439,483
6     1991                     50.0         927,026
7     1992                     50.0       1,065,345
8     1993                     50.0       1,230,833
9     1994                     50.0       1,170,559
10    1995                     50.0         922,057
11    1996                     50.0       1,002,454
12    1997                     50.0       1,317,354
13    1998                     50.0       1,639,935
14    1999                     50.0       1,684,945
```

```
15    2000                         50.0       2,017,977
16    2001                         50.0       2,396,882
17    2002                         50.0       2,703,244
18    2003                         50.0       2,852,440
19    2004                         50.0       3,122,758
20    2005                         50.0       3,108,319
21    2006                         50.0       3,372,303
22    2007                         50.0       3,691,493
23    2008                         50.0       4,383,179
24    2009                         50.0       5,392,360
25    2010                         50.0       5,429,962
26    2011                         50.0       5,001,893
27    2012                         50.0       3,392,193
28    2013                         50.0       3,867,989
29    2014                         50.0       2,426,759,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                         50.0         577,405
1     1986                         50.0         555,102
2     1987                         50.0         550,307
3     1988                         50.0         524,767
4     1989                         50.0         444,500
5     1990                         50.0         439,483
6     1991                         50.0         927,026
7     1992                         50.0       1,065,345
8     1993                         50.0       1,230,833
9     1994                         50.0       1,170,559
10    1995                         50.0         922,057
11    1996                         50.0       1,002,454
12    1997                         50.0       1,317,354
13    1998                         50.0       1,639,935
14    1999                         50.0       1,684,945
15    2000                         50.0       2,017,977
16    2001                         50.0       2,396,882
17    2002                         50.0       2,703,244
18    2003                         50.0       2,852,440
19    2004                         50.0       3,122,758
20    2005                         50.0       3,108,319
21    2006                         50.0       3,372,303
22    2007                         50.0       3,691,493
23    2008                         50.0       4,383,179
24    2009                         50.0       5,392,360
25    2010                         50.0       5,429,962
26    2011                         50.0       5,001,893
27    2012                         50.0       3,392,193
28    2013                         50.0       3,867,989
29    2014                         50.0       2,426,759,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
```

```
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         379,996
1     1986                     50.0         396,879
2     1987                     50.0         356,980
3     1988                     50.0         355,536
4     1989                     50.0         381,793
5     1990                     50.0         422,647
6     1991                     50.0       1,011,743
7     1992                     50.0       1,330,796
8     1993                     50.0       1,359,989
9     1994                     50.0       1,321,349
10    1995                     50.0       1,268,960
11    1996                     50.0       1,215,038
12    1997                     50.0       1,244,200
13    1998                     50.0         676,617
14    1999                     50.0       1,095,572
15    2000                     50.0       1,735,822
16    2001                     50.0       1,814,296
17    2002                     50.0       1,501,679
18    2003                     50.0       2,119,845
19    2004                     50.0       1,664,830
20    2005                     50.0       2,063,086
21    2006                     50.0       2,175,339
22    2007                     50.0       2,210,483
23    2008                     50.0       2,647,060
24    2009                     50.0       3,198,195
25    2010                     50.0       2,760,059
26    2011                     50.0       2,531,571
27    2012                     50.0       2,935,843
28    2013                     50.0       4,256,178
29    2014                     50.0       3,864,910,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         379,996
1     1986                     50.0         396,879
2     1987                     50.0         356,980
3     1988                     50.0         355,536
4     1989                     50.0         381,793
5     1990                     50.0         422,647
6     1991                     50.0       1,011,743
7     1992                     50.0       1,330,796
8     1993                     50.0       1,359,989
9     1994                     50.0       1,321,349
10    1995                     50.0       1,268,960
11    1996                     50.0       1,215,038
12    1997                     50.0       1,244,200
13    1998                     50.0         676,617
14    1999                     50.0       1,095,572
15    2000                     50.0       1,735,822
16    2001                     50.0       1,814,296
17    2002                     50.0       1,501,679
18    2003                     50.0       2,119,845
19    2004                     50.0       1,664,830
20    2005                     50.0       2,063,086
```

```
21    2006                          50.0       2,175,339
22    2007                          50.0       2,210,483
23    2008                          50.0       2,647,060
24    2009                          50.0       3,198,195
25    2010                          50.0       2,760,059
26    2011                          50.0       2,531,571
27    2012                          50.0       2,935,843
28    2013                          50.0       4,256,178
29    2014                          50.0       3,864,910,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         327,583
1     1986                     50.0         251,919
2     1987                     50.0         304,062
3     1988                     50.0         406,204
4     1989                     50.0         336,833
5     1990                     50.0         439,000
6     1991                     50.0         587,833
7     1992                     50.0         267,801
8     1993                     50.0         464,025
9     1994                     50.0         871,157
10    1995                     50.0       1,083,938
11    1996                     50.0       1,551,850
12    1997                     50.0       1,832,337
13    1998                     50.0       1,842,429
14    1999                     50.0       1,920,485
15    2000                     50.0       2,918,491
16    2001                     50.0       3,105,066
17    2002                     50.0       2,630,314
18    2003                     50.0       1,567,252
19    2004                     50.0       1,143,976
20    2005                     50.0       1,431,120
21    2006                     50.0       2,241,260
22    2007                     50.0       2,126,664
23    2008                     50.0       3,037,310
24    2009                     50.0       3,021,450
25    2010                     50.0       2,110,481
26    2011                     50.0       1,681,950
27    2012                     50.0       2,704,493
28    2013                     50.0       2,706,135
29    2014                     50.0       3,159,688,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         327,583
1     1986                     50.0         251,919
2     1987                     50.0         304,062
3     1988                     50.0         406,204
4     1989                     50.0         336,833
5     1990                     50.0         439,000
6     1991                     50.0         587,833
7     1992                     50.0         267,801
8     1993                     50.0         464,025
9     1994                     50.0         871,157
10    1995                     50.0       1,083,938
11    1996                     50.0       1,551,850
12    1997                     50.0       1,832,337
```

```
13    1998                      50.0       1,842,429
14    1999                      50.0       1,920,485
15    2000                      50.0       2,918,491
16    2001                      50.0       3,105,066
17    2002                      50.0       2,630,314
18    2003                      50.0       1,567,252
19    2004                      50.0       1,143,976
20    2005                      50.0       1,431,120
21    2006                      50.0       2,241,260
22    2007                      50.0       2,126,664
23    2008                      50.0       3,037,310
24    2009                      50.0       3,021,450
25    2010                      50.0       2,110,481
26    2011                      50.0       1,681,950
27    2012                      50.0       2,704,493
28    2013                      50.0       2,706,135
29    2014                      50.0       3,159,688,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                      50.0         327,583
1     1986                      50.0         251,919
2     1987                      50.0         304,062
3     1988                      50.0         406,204
4     1989                      50.0         336,833
5     1990                      50.0         439,000
6     1991                      50.0         587,833
7     1992                      50.0         267,801
8     1993                      50.0         464,025
9     1994                      50.0         871,157
10    1995                      50.0       1,083,938
11    1996                      50.0       1,551,850
12    1997                      50.0       1,832,337
13    1998                      50.0       1,842,429
14    1999                      50.0       1,920,485
15    2000                      50.0       2,918,491
16    2001                      50.0       3,105,066
17    2002                      50.0       2,630,314
18    2003                      50.0       1,567,252
19    2004                      50.0       1,143,976
20    2005                      50.0       1,431,120
21    2006                      50.0       2,241,260
22    2007                      50.0       2,126,664
23    2008                      50.0       3,037,310
24    2009                      50.0       3,021,450
25    2010                      50.0       2,110,481
26    2011                      50.0       1,681,950
27    2012                      50.0       2,704,493
28    2013                      50.0       2,706,135
29    2014                      50.0       3,159,688,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                      50.0         327,583
1     1986                      50.0         251,919
2     1987                      50.0         304,062
3     1988                      50.0         406,204
4     1989                      50.0         336,833
```

```
5     1990                        50.0         439,000
6     1991                        50.0         587,833
7     1992                        50.0         267,801
8     1993                        50.0         464,025
9     1994                        50.0         871,157
10    1995                        50.0       1,083,938
11    1996                        50.0       1,551,850
12    1997                        50.0       1,832,337
13    1998                        50.0       1,842,429
14    1999                        50.0       1,920,485
15    2000                        50.0       2,918,491
16    2001                        50.0       3,105,066
17    2002                        50.0       2,630,314
18    2003                        50.0       1,567,252
19    2004                        50.0       1,143,976
20    2005                        50.0       1,431,120
21    2006                        50.0       2,241,260
22    2007                        50.0       2,126,664
23    2008                        50.0       3,037,310
24    2009                        50.0       3,021,450
25    2010                        50.0       2,110,481
26    2011                        50.0       1,681,950
27    2012                        50.0       2,704,493
28    2013                        50.0       2,706,135
29    2014                        50.0       3,159,688,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1993                        50.0         313,742
1     1994                        50.0         702,568
2     1995                        50.0         948,742
3     1996                        50.0       1,296,123
4     1997                        50.0       1,451,988
5     1998                        50.0       1,682,821
6     1999                        50.0       1,821,642
7     2000                        50.0       2,182,542
8     2001                        50.0       2,751,589
9     2002                        50.0       2,105,594
10    2003                        50.0       2,239,322
11    2004                        50.0       2,337,327
12    2005                        50.0       1,649,620
13    2006                        50.0       1,288,531
14    2007                        50.0       2,078,500
15    2008                        50.0       2,640,596
16    2009                        50.0       2,785,222
17    2010                        50.0       2,904,379
18    2011                        50.0       3,390,310
19    2012                        50.0       2,692,054
20    2013                        50.0       2,976,362
21    2014                        50.0       3,180,116,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                        50.0         517,407
1     1986                        50.0         456,878
```

```
2     1987                       50.0         527,081
3     1988                       50.0         559,546
4     1989                       50.0         582,554
5     1990                       50.0         533,128
6     1991                       50.0         953,533
7     1992                       50.0         975,815
8     1993                       50.0       1,230,650
9     1994                       50.0       1,336,983
10    1995                       50.0       1,122,550
11    1996                       50.0         732,437
12    1997                       50.0         595,586
13    1998                       50.0         687,571
14    1999                       50.0       1,073,225
15    2000                       50.0       2,157,969
16    2001                       50.0       1,907,720
17    2002                       50.0       1,966,000
18    2003                       50.0       1,891,076
19    2004                       50.0       1,672,571
20    2005                       50.0       2,467,571
21    2006                       50.0       2,950,459
22    2007                       50.0       3,268,978
23    2008                       50.0       4,589,506
24    2009                       50.0       4,110,183
25    2010                       50.0       4,550,552
26    2011                       50.0       3,914,823
27    2012                       50.0       4,562,068
28    2013                       50.0       6,082,895
29    2014                       50.0       6,645,891,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1993                       50.0         568,545
1     1994                       50.0         584,675
2     1995                       50.0         662,588
3     1996                       50.0         969,453
4     1997                       50.0       1,570,725
5     1998                       50.0       1,033,066
6     1999                       50.0         585,694
7     2000                       50.0         709,714
8     2001                       50.0       1,153,629
9     2002                       50.0       1,499,282
10    2003                       50.0       1,648,333
11    2004                       50.0       1,560,853
12    2005                       50.0       2,237,364
13    2006                       50.0         586,860
14    2007                       50.0         984,096
15    2008                       50.0         660,954
16    2009                       50.0       1,315,500
17    2010                       50.0       2,112,211
18    2011                       50.0       2,190,153,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
```

```
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         499,652
1     1986                     50.0         411,386
2     1987                     50.0         450,298
3     1988                     50.0         472,544
4     1989                     50.0         556,648
5     1990                     50.0         591,290
6     1991                     50.0         514,100
7     1992                     50.0         513,583
8     1993                     50.0       1,041,741
9     1994                     50.0       1,003,818
10    1995                     50.0         949,162
11    1996                     50.0         918,935
12    1997                     50.0       1,121,854
13    1998                     50.0       1,324,187
14    1999                     50.0       1,716,062
15    2000                     50.0       1,899,596
16    2001                     50.0       2,164,738
17    2002                     50.0       2,349,941
18    2003                     50.0       2,732,307
19    2004                     50.0       3,015,880
20    2005                     50.0       2,953,038
21    2006                     50.0       3,547,777
22    2007                     50.0       3,250,333
23    2008                     50.0       3,293,719
24    2009                     50.0       3,814,682
25    2010                     50.0       3,298,410
26    2011                     50.0       2,437,724
27    2012                     50.0       2,332,730
28    2013                     50.0         813,213
29    2014                     50.0       1,755,815,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         499,652
1     1986                     50.0         411,386
2     1987                     50.0         450,298
3     1988                     50.0         472,544
4     1989                     50.0         556,648
5     1990                     50.0         591,290
6     1991                     50.0         514,100
7     1992                     50.0         513,583
8     1993                     50.0       1,041,741
9     1994                     50.0       1,003,818
10    1995                     50.0         949,162
11    1996                     50.0         918,935
12    1997                     50.0       1,121,854
13    1998                     50.0       1,324,187
14    1999                     50.0       1,716,062
15    2000                     50.0       1,899,596
16    2001                     50.0       2,164,738
17    2002                     50.0       2,349,941
18    2003                     50.0       2,732,307
19    2004                     50.0       3,015,880
20    2005                     50.0       2,953,038
21    2006                     50.0       3,547,777
22    2007                     50.0       3,250,333
23    2008                     50.0       3,293,719
24    2009                     50.0       3,814,682
```

```
25    2010                        50.0       3,298,410
26    2011                        50.0       2,437,724
27    2012                        50.0       2,332,730
28    2013                        50.0         813,213
29    2014                        50.0       1,755,815,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                        50.0         423,689
1     1986                        50.0         465,846
2     1987                        50.0         438,076
3     1988                        50.0         559,867
4     1989                        50.0         691,984
5     1990                        50.0         667,459
6     1991                        50.0       1,253,325
7     1992                        50.0       1,210,494
8     1993                        50.0       1,292,067
9     1994                        50.0       1,501,530
10    1995                        50.0         738,320
11    1996                        50.0         614,583
12    1997                        50.0         936,621
13    1998                        50.0         945,192
14    1999                        50.0         708,783
15    2000                        50.0         836,892
16    2001                        50.0       1,265,089
17    2002                        50.0       1,629,551
18    2003                        50.0       1,558,384
19    2004                        50.0       1,586,966
20    2005                        50.0       1,365,962
21    2006                        50.0       1,630,827
22    2007                        50.0       2,165,048
23    2008                        50.0       2,240,211
24    2009                        50.0       2,712,282
25    2010                        50.0       2,644,637
26    2011                        50.0       1,373,538
```

```
27    2012                      50.0       2,030,540
28    2013                      50.0       2,966,360
29    2014                      50.0       2,983,763,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
   yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0    2005                     50.0       3,271,304
1    2006                     50.0       4,138,880
2    2007                     50.0       3,641,711
3    2008                     50.0       4,110,908
4    2009                     50.0       4,061,035
5    2010                     50.0       3,619,443
6    2011                     50.0       4,469,134
7    2012                     50.0       5,327,074
8    2013                     50.0       4,775,951
9    2014                     50.0       4,518,083,
   yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0    2005                     50.0       3,271,304
1    2006                     50.0       4,138,880
2    2007                     50.0       3,641,711
3    2008                     50.0       4,110,908
4    2009                     50.0       4,061,035
5    2010                     50.0       3,619,443
6    2011                     50.0       4,469,134
7    2012                     50.0       5,327,074
8    2013                     50.0       4,775,951
9    2014                     50.0       4,518,083,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         476,865
1     1986                     50.0         466,055
2     1987                     50.0         488,407
3     1988                     50.0         561,683
4     1989                     50.0         679,727
5     1990                     50.0         609,105
6     1991                     50.0       1,093,022
7     1992                     50.0       1,544,419
8     1993                     50.0       1,229,124
9     1994                     50.0       1,225,806
10    1995                     50.0       1,061,437
11    1996                     50.0       1,104,843
12    1997                     50.0       1,512,676
13    1998                     50.0       1,435,882
14    1999                     50.0       2,378,307
15    2000                     50.0       3,381,703
16    2001                     50.0       3,762,274
17    2002                     50.0       3,648,113
18    2003                     50.0       4,222,904
19    2004                     50.0       3,440,814
20    2005                     50.0       2,767,966
21    2006                     50.0       3,515,970
22    2007                     50.0       3,739,811
23    2008                     50.0       4,089,259
24    2009                     50.0       4,016,583
25    2010                     50.0       3,531,778
26    2011                     50.0       3,472,966
27    2012                     50.0       3,171,452
28    2013                     50.0       6,980,068
29    2014                     50.0       6,781,706,
Empty DataFrame
```

```
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
   yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0    2012                     50.0       4,373,259
1    2013                     50.0       1,400,079
2    2014                     50.0       1,549,514,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         593,900
1     1986                     50.0         355,130
2     1987                     50.0         260,472
3     1988                     50.0         381,909
4     1989                     50.0         384,433
5     1990                     50.0         532,950
6     1991                     50.0       1,005,021
7     1992                     50.0       1,148,654
8     1993                     50.0         721,419
9     1994                     50.0         785,500
10    1995                     50.0         508,537
11    1996                     50.0         639,117
12    1997                     50.0         695,745,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1998                     50.0         997,497
1     1999                     50.0       1,355,543
2     2000                     50.0       1,140,791
3     2001                     50.0       1,567,386
4     2002                     50.0       1,734,063
5     2003                     50.0       1,400,931
6     2004                     50.0       1,101,140
7     2005                     50.0       1,597,393
8     2006                     50.0       2,214,166
9     2007                     50.0       2,629,129
10    2008                     50.0       2,790,948
11    2009                     50.0       3,083,942
12    2010                     50.0       2,796,837
13    2011                     50.0       2,849,911
14    2012                     50.0       3,755,920
```

```
15    2013                    50.0      3,077,881
16    2014                    50.0      3,748,777,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                    50.0        303,411
1     1986                    50.0        397,643
2     1987                    50.0        710,833
3     1988                    50.0        541,855
4     1989                    50.0        621,266
5     1990                    50.0        405,611
6     1991                    50.0        973,409
7     1992                    50.0      1,000,994
8     1993                    50.0        855,088
9     1994                    50.0        947,950
10    1995                    50.0        770,015
11    1996                    50.0        700,515
12    1997                    50.0      1,099,112
13    1998                    50.0        963,017
14    1999                    50.0        733,017
15    2000                    50.0        635,365
16    2001                    50.0        893,703
17    2002                    50.0      1,497,222
18    2003                    50.0      2,134,807
19    2004                    50.0      2,060,961
20    2005                    50.0      2,080,962
21    2006                    50.0      2,438,307
22    2007                    50.0      2,551,410
23    2008                    50.0      2,277,310
24    2009                    50.0      2,251,698
25    2010                    50.0      3,484,255
26    2011                    50.0      4,509,480
27    2012                    50.0      3,484,629
28    2013                    50.0      2,790,277
29    2014                    50.0      3,102,314,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                    50.0        473,508
1     1986                    50.0        346,987
2     1987                    50.0        289,252
3     1988                    50.0        384,133
4     1989                    50.0        493,121
5     1990                    50.0        535,044
6     1991                    50.0        631,313
7     1992                    50.0        586,012
8     1993                    50.0        484,598
9     1994                    50.0        682,071
10    1995                    50.0        374,666
11    1996                    50.0        524,661
12    1997                    50.0        622,435
13    1998                    50.0        322,469
14    1999                    50.0        511,514
15    2000                    50.0      1,137,735
16    2001                    50.0      1,134,177
17    2002                    50.0      1,381,089
18    2003                    50.0      1,998,019
19    2004                    50.0      1,410,258,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
```

```
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         711,910
1     1986                     50.0         660,509
2     1987                     50.0         488,563
3     1988                     50.0         648,038
4     1989                     50.0         552,076
5     1990                     50.0         633,706
6     1991                     50.0         976,577
7     1992                     50.0       1,137,676
8     1993                     50.0       1,291,663
9     1994                     50.0       1,576,942
10    1995                     50.0       1,437,495
11    1996                     50.0       1,593,876
12    1997                     50.0       2,146,260
13    1998                     50.0       2,087,714
14    1999                     50.0       2,990,839
15    2000                     50.0       3,297,795
16    2001                     50.0       3,622,165
17    2002                     50.0       4,342,364
18    2003                     50.0       5,455,350
19    2004                     50.0       6,351,515
20    2005                     50.0       8,011,800
21    2006                     50.0       6,952,252
22    2007                     50.0       6,759,251
23    2008                     50.0       6,929,892
24    2009                     50.0       7,748,045
25    2010                     50.0       8,253,335
26    2011                     50.0       6,975,000
27    2012                     50.0       6,776,630
28    2013                     50.0       7,483,189
29    2014                     50.0       8,230,996,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         515,941
1     1986                     50.0         549,775
2     1987                     50.0         553,868
3     1988                     50.0         610,772
4     1989                     50.0         710,181
5     1990                     50.0         678,838
6     1991                     50.0       1,253,461
7     1992                     50.0       1,715,461
8     1993                     50.0       1,301,455
9     1994                     50.0         998,599
10    1995                     50.0         674,999
11    1996                     50.0         741,803
12    1997                     50.0       1,137,154
13    1998                     50.0       1,578,121
14    1999                     50.0       2,099,744
15    2000                     50.0       3,180,391
```

```
16    2001                         50.0       3,212,911
17    2002                         50.0       3,639,753
18    2003                         50.0       4,174,158
19    2004                         50.0       3,452,177
20    2005                         50.0       3,752,067
21    2006                         50.0       3,743,887
22    2007                         50.0       3,841,055
23    2008                         50.0       4,593,112
24    2009                         50.0       5,334,785
25    2010                         50.0       4,800,819
26    2011                         50.0       4,401,752
27    2012                         50.0       3,457,554
28    2013                         50.0       1,648,278
29    2014                         50.0      10,250,000,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                         50.0         711,910
1     1986                         50.0         660,509
2     1987                         50.0         488,563
3     1988                         50.0         648,038
4     1989                         50.0         552,076
5     1990                         50.0         633,706
6     1991                         50.0         976,577
7     1992                         50.0       1,137,676
8     1993                         50.0       1,291,663
9     1994                         50.0       1,576,942
10    1995                         50.0       1,437,495
11    1996                         50.0       1,593,876
12    1997                         50.0       2,146,260
13    1998                         50.0       2,087,714
14    1999                         50.0       2,990,839
15    2000                         50.0       3,297,795
16    2001                         50.0       3,622,165
17    2002                         50.0       4,342,364
18    2003                         50.0       5,455,350
19    2004                         50.0       6,351,515
20    2005                         50.0       8,011,800
21    2006                         50.0       6,952,252
22    2007                         50.0       6,759,251
23    2008                         50.0       6,929,892
24    2009                         50.0       7,748,045
25    2010                         50.0       8,253,335
26    2011                         50.0       6,975,000
27    2012                         50.0       6,776,630
28    2013                         50.0       7,483,189
29    2014                         50.0       8,230,996,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                         50.0         431,362
1     1986                         50.0         315,465
2     1987                         50.0         449,263
3     1988                         50.0         421,304
4     1989                         50.0         624,522
```

```
5     1990                       50.0         584,926
6     1991                       50.0       1,423,044
7     1992                       50.0       1,282,343
8     1993                       50.0       1,050,342
9     1994                       50.0       1,035,530
10    1995                       50.0       1,019,979
11    1996                       50.0         574,135
12    1997                       50.0         615,858
13    1998                       50.0         608,657
14    1999                       50.0         814,394
15    2000                       50.0       1,184,123
16    2001                       50.0       1,252,250
17    2002                       50.0       1,481,635
18    2003                       50.0       1,933,109
19    2004                       50.0       2,122,345
20    2005                       50.0       2,131,760
21    2006                       50.0       2,489,723
22    2007                       50.0       2,834,533
23    2008                       50.0       1,713,111
24    2009                       50.0       2,292,962
25    2010                       50.0       1,726,715
26    2011                       50.0       2,376,303
27    2012                       50.0       1,845,750
28    2013                       50.0       1,939,758
29    2014                       50.0       2,784,938,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                       50.0         532,892
1     1986                       50.0         373,876
2     1987                       50.0         479,759
3     1988                       50.0         532,230
4     1989                       50.0         407,846
5     1990                       50.0         439,122
6     1991                       50.0         749,577
7     1992                       50.0         812,794
8     1993                       50.0         891,822
9     1994                       50.0         987,468
10    1995                       50.0         783,485
11    1996                       50.0         798,011
12    1997                       50.0         916,412
13    1998                       50.0       1,037,071
14    1999                       50.0         932,132
15    2000                       50.0       1,631,310
16    2001                       50.0       1,602,455
17    2002                       50.0       2,069,821
18    2003                       50.0       2,440,689
```

```
19    2004                     50.0      3,573,814
20    2005                     50.0      3,673,923
21    2006                     50.0      3,269,382
22    2007                     50.0      2,980,940
23    2008                     50.0      3,495,710
24    2009                     50.0      4,185,335
25    2010                     50.0      5,068,870
26    2011                     50.0      5,765,879
27    2012                     50.0      5,817,964
28    2013                     50.0      6,533,199
29    2014                     50.0      5,654,530,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         532,892
1     1986                     50.0         373,876
2     1987                     50.0         479,759
3     1988                     50.0         532,230
4     1989                     50.0         407,846
5     1990                     50.0         439,122
6     1991                     50.0         749,577
7     1992                     50.0         812,794
8     1993                     50.0         891,822
9     1994                     50.0         987,468
10    1995                     50.0         783,485
11    1996                     50.0         798,011
12    1997                     50.0         916,412
13    1998                     50.0       1,037,071
14    1999                     50.0         932,132
15    2000                     50.0       1,631,310
16    2001                     50.0       1,602,455
17    2002                     50.0       2,069,821
18    2003                     50.0       2,440,689
19    2004                     50.0       3,573,814
20    2005                     50.0       3,673,923
21    2006                     50.0       3,269,382
22    2007                     50.0       2,980,940
23    2008                     50.0       3,495,710
24    2009                     50.0       4,185,335
25    2010                     50.0       5,068,870
26    2011                     50.0       5,765,879
27    2012                     50.0       5,817,964
28    2013                     50.0       6,533,199
29    2014                     50.0       5,654,530,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         532,892
1     1986                     50.0         373,876
2     1987                     50.0         479,759
3     1988                     50.0         532,230
4     1989                     50.0         407,846
5     1990                     50.0         439,122
6     1991                     50.0         749,577
7     1992                     50.0         812,794
8     1993                     50.0         891,822
9     1994                     50.0         987,468
10    1995                     50.0         783,485
```

```
11    1996                      50.0         798,011
12    1997                      50.0         916,412
13    1998                      50.0       1,037,071
14    1999                      50.0         932,132
15    2000                      50.0       1,631,310
16    2001                      50.0       1,602,455
17    2002                      50.0       2,069,821
18    2003                      50.0       2,440,689
19    2004                      50.0       3,573,814
20    2005                      50.0       3,673,923
21    2006                      50.0       3,269,382
22    2007                      50.0       2,980,940
23    2008                      50.0       3,495,710
24    2009                      50.0       4,185,335
25    2010                      50.0       5,068,870
26    2011                      50.0       5,765,879
27    2012                      50.0       5,817,964
28    2013                      50.0       6,533,199
29    2014                      50.0       5,654,530,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                      50.0         485,657
1     1986                      50.0         373,913
2     1987                      50.0         306,080
3     1988                      50.0         222,166
4     1989                      50.0         439,224
5     1990                      50.0         432,111
6     1991                      50.0         875,358
7     1992                      50.0       1,212,291
8     1993                      50.0         752,195
9     1994                      50.0         712,272
10    1995                      50.0         556,222
11    1996                      50.0         657,642
12    1997                      50.0         307,761
13    1998                      50.0         430,428
14    1999                      50.0         649,938
15    2000                      50.0       1,112,628
16    2001                      50.0       1,863,252
17    2002                      50.0       1,459,434
18    2003                      50.0       1,957,586
19    2004                      50.0       1,193,627
20    2005                      50.0       1,361,892
21    2006                      50.0       1,668,491
22    2007                      50.0       1,427,327
23    2008                      50.0       1,872,683
24    2009                      50.0       1,872,807
25    2010                      50.0       1,294,185
26    2011                      50.0       1,553,344
27    2012                      50.0       2,248,285
28    2013                      50.0       2,752,214
29    2014                      50.0       2,756,357,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
```

```
0     1985                    50.0         485,657
1     1986                    50.0         373,913
2     1987                    50.0         306,080
3     1988                    50.0         222,166
4     1989                    50.0         439,224
5     1990                    50.0         432,111
6     1991                    50.0         875,358
7     1992                    50.0       1,212,291
8     1993                    50.0         752,195
9     1994                    50.0         712,272
10    1995                    50.0         556,222
11    1996                    50.0         657,642
12    1997                    50.0         307,761
13    1998                    50.0         430,428
14    1999                    50.0         649,938
15    2000                    50.0       1,112,628
16    2001                    50.0       1,863,252
17    2002                    50.0       1,459,434
18    2003                    50.0       1,957,586
19    2004                    50.0       1,193,627
20    2005                    50.0       1,361,892
21    2006                    50.0       1,668,491
22    2007                    50.0       1,427,327
23    2008                    50.0       1,872,683
24    2009                    50.0       1,872,807
25    2010                    50.0       1,294,185
26    2011                    50.0       1,553,344
27    2012                    50.0       2,248,285
28    2013                    50.0       2,752,214
29    2014                    50.0       2,756,357,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                    50.0         501,662
1     1986                    50.0         455,227
2     1987                    50.0         409,844
3     1988                    50.0         354,111
4     1989                    50.0         525,740
5     1990                    50.0         549,635
6     1991                    50.0         738,333
7     1992                    50.0         959,077
8     1993                    50.0         750,333
9     1994                    50.0         426,180
10    1995                    50.0         851,043
11    1996                    50.0         944,939
12    1997                    50.0       1,288,402
13    1998                    50.0       1,562,050
```

```
14    1999                     50.0       1,508,126
15    2000                     50.0       1,827,366
16    2001                     50.0       1,399,386
17    2002                     50.0       1,428,448
18    2003                     50.0       1,507,000
19    2004                     50.0       2,130,185
20    2005                     50.0       2,260,386
21    2006                     50.0       2,496,290
22    2007                     50.0       2,235,021
23    2008                     50.0       2,376,697
24    2009                     50.0       1,604,951
25    2010                     50.0       1,453,819
26    2011                     50.0       1,479,649
27    2012                     50.0       1,973,025
28    2013                     50.0       2,342,339
29    2014                     50.0       2,703,060,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         411,085
1     1986                     50.0         319,535
2     1987                     50.0         316,956
3     1988                     50.0         495,200
4     1989                     50.0         534,386
5     1990                     50.0         568,686
6     1991                     50.0       1,191,064
7     1992                     50.0       1,326,526
8     1993                     50.0       1,130,645
9     1994                     50.0       1,332,458
10    1995                     50.0         934,943
11    1996                     50.0       1,061,277
12    1997                     50.0       1,112,261
13    1998                     50.0       1,520,208
14    1999                     50.0       1,606,726
15    2000                     50.0       2,066,839
16    2001                     50.0       2,343,709
17    2002                     50.0       2,899,993
18    2003                     50.0       3,186,621
19    2004                     50.0       2,645,779
20    2005                     50.0       3,469,211
21    2006                     50.0       3,463,708
22    2007                     50.0       3,469,963
23    2008                     50.0       2,641,189
24    2009                     50.0       2,965,230
25    2010                     50.0       3,522,904
26    2011                     50.0       4,377,716
27    2012                     50.0       3,920,689
28    2013                     50.0       5,006,440
29    2014                     50.0      20,000,000,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                     50.0         256,277
1     1986                     50.0         229,165
2     1987                     50.0         251,500
3     1988                     50.0         282,401
4     1989                     50.0         407,479
5     1990                     50.0         369,225
6     1991                     50.0         603,532
7     1992                     50.0         799,304
8     1993                     50.0         990,797
9     1994                     50.0         885,712
10    1995                     50.0       1,042,323
11    1996                     50.0       1,215,544
```

```
12    1997                        50.0       1,298,145
13    1998                        50.0       1,423,343
14    1999                        50.0       1,503,472
15    2000                        50.0       2,265,961
16    2001                        50.0       2,668,601
17    2002                        50.0       3,211,306
18    2003                        50.0       3,220,709
19    2004                        50.0       2,911,279
20    2005                        50.0       2,742,322
21    2006                        50.0       3,257,771
22    2007                        50.0       3,942,993
23    2008                        50.0       4,525,633
24    2009                        50.0       3,532,291
25    2010                        50.0       3,089,642
26    2011                        50.0       2,969,331
27    2012                        50.0       2,927,789
28    2013                        50.0       2,846,347
29    2014                        50.0       3,701,244,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                        50.0         472,683
1     1986                        50.0         365,741
2     1987                        50.0         379,290
3     1988                        50.0         477,037
4     1989                        50.0         574,244
5     1990                        50.0         586,380
6     1991                        50.0         753,793
7     1992                        50.0         951,166
8     1993                        50.0         778,911
9     1994                        50.0         975,853
10    1995                        50.0         976,342
11    1996                        50.0       1,220,292
12    1997                        50.0       1,262,685
13    1998                        50.0       1,562,072
14    1999                        50.0       1,382,727
15    2000                        50.0       2,276,069
16    2001                        50.0       2,617,944
17    2002                        50.0       2,871,572
18    2003                        50.0       2,702,795
19    2004                        50.0       3,201,089
20    2005                        50.0       3,542,570
21    2006                        50.0       3,292,273
22    2007                        50.0       3,224,529
23    2008                        50.0       3,018,922
24    2009                        50.0       3,278,829
25    2010                        50.0       3,741,630
```

```
26    2011                        50.0      3,904,947
27    2012                        50.0      3,939,316
28    2013                        50.0      3,295,003
29    2014                        50.0      4,310,464,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                        50.0        472,683
1     1986                        50.0        365,741
2     1987                        50.0        379,290
3     1988                        50.0        477,037
4     1989                        50.0        574,244
5     1990                        50.0        586,380
6     1991                        50.0        753,793
7     1992                        50.0        951,166
8     1993                        50.0        778,911
9     1994                        50.0        975,853
10    1995                        50.0        976,342
11    1996                        50.0      1,220,292
12    1997                        50.0      1,262,685
13    1998                        50.0      1,562,072
14    1999                        50.0      1,382,727
15    2000                        50.0      2,276,069
16    2001                        50.0      2,617,944
17    2002                        50.0      2,871,572
18    2003                        50.0      2,702,795
19    2004                        50.0      3,201,089
20    2005                        50.0      3,542,570
21    2006                        50.0      3,292,273
22    2007                        50.0      3,224,529
23    2008                        50.0      3,018,922
24    2009                        50.0      3,278,829
25    2010                        50.0      3,741,630
26    2011                        50.0      3,904,947
27    2012                        50.0      3,939,316
28    2013                        50.0      3,295,003
29    2014                        50.0      4,310,464,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                        50.0        472,683
1     1986                        50.0        365,741
2     1987                        50.0        379,290
3     1988                        50.0        477,037
4     1989                        50.0        574,244
5     1990                        50.0        586,380
6     1991                        50.0        753,793
7     1992                        50.0        951,166
8     1993                        50.0        778,911
9     1994                        50.0        975,853
10    1995                        50.0        976,342
11    1996                        50.0      1,220,292
12    1997                        50.0      1,262,685
13    1998                        50.0      1,562,072
14    1999                        50.0      1,382,727
```

```
15    2000                        50.0       2,276,069
16    2001                        50.0       2,617,944
17    2002                        50.0       2,871,572
18    2003                        50.0       2,702,795
19    2004                        50.0       3,201,089
20    2005                        50.0       3,542,570
21    2006                        50.0       3,292,273
22    2007                        50.0       3,224,529
23    2008                        50.0       3,018,922
24    2009                        50.0       3,278,829
25    2010                        50.0       3,741,630
26    2011                        50.0       3,904,947
27    2012                        50.0       3,939,316
28    2013                        50.0       3,295,003
29    2014                        50.0       4,310,464,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1998                        50.0         826,666
1     1999                        50.0       1,022,894
2     2000                        50.0       2,024,681
3     2001                        50.0       2,110,370
4     2002                        50.0       1,227,857
5     2003                        50.0         785,200
6     2004                        50.0       1,094,691
7     2005                        50.0       1,023,416
8     2006                        50.0       1,293,258
9     2007                        50.0         893,462
10    2008                        50.0       1,460,686
11    2009                        50.0       2,183,208
12    2010                        50.0       2,663,832
13    2011                        50.0       1,578,983
14    2012                        50.0       2,291,910
15    2013                        50.0       2,302,403
16    2014                        50.0       2,907,564,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1998                        50.0         826,666
1     1999                        50.0       1,022,894
2     2000                        50.0       2,024,681
3     2001                        50.0       2,110,370
4     2002                        50.0       1,227,857
5     2003                        50.0         785,200
6     2004                        50.0       1,094,691
7     2005                        50.0       1,023,416
8     2006                        50.0       1,293,258
9     2007                        50.0         893,462
10    2008                        50.0       1,460,686
```

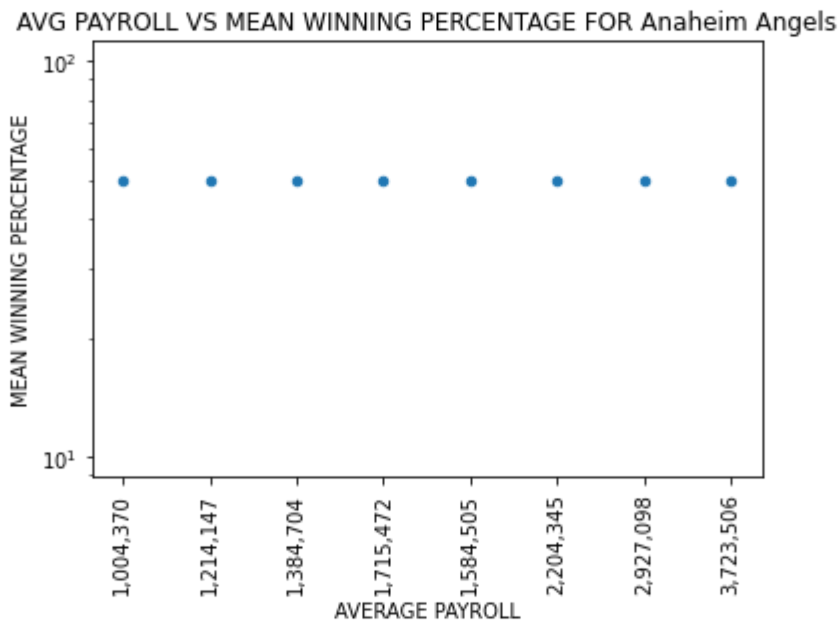
```
11    2009                      50.0      2,183,208
12    2010                      50.0      2,663,832
13    2011                      50.0      1,578,983
14    2012                      50.0      2,291,910
15    2013                      50.0      2,302,403
16    2014                      50.0      2,907,564,
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                      50.0        383,825
1     1986                      50.0        259,350
2     1987                      50.0        220,000
3     1988                      50.0        242,824
4     1989                      50.0        396,459
5     1990                      50.0        402,010
6     1991                      50.0        867,833
7     1992                      50.0        971,876
8     1993                      50.0        957,288
9     1994                      50.0        999,199
10    1995                      50.0      1,017,101
11    1996                      50.0      1,183,076
12    1997                      50.0      1,444,563
13    1998                      50.0      1,885,736
14    1999                      50.0      2,556,997
15    2000                      50.0      2,722,920
16    2001                      50.0      2,859,145
17    2002                      50.0      3,768,790
18    2003                      50.0      3,449,722
19    2004                      50.0      1,898,290
20    2005                      50.0      1,801,580
21    2006                      50.0      2,200,924
22    2007                      50.0      2,439,952
23    2008                      50.0      2,334,907
24    2009                      50.0      2,350,993
25    2010                      50.0      1,905,191
26    2011                      50.0      3,182,733
27    2012                      50.0      4,635,037
28    2013                      50.0      3,880,089
29    2014                      50.0      4,677,294,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
    yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0     1985                      50.0        440,627
1     1986                      50.0        467,075
2     1987                      50.0        455,630
3     1988                      50.0        470,816
4     1989                      50.0        580,773
5     1990                      50.0        657,660
6     1991                      50.0        796,096
7     1992                      50.0      1,244,129
8     1993                      50.0      1,432,702
9     1994                      50.0      1,447,788
10    1995                      50.0      1,533,030
11    1996                      50.0        895,608
12    1997                      50.0      1,426,661
13    1998                      50.0      1,605,500
14    1999                      50.0      1,420,135
15    2000                      50.0      1,793,533
```

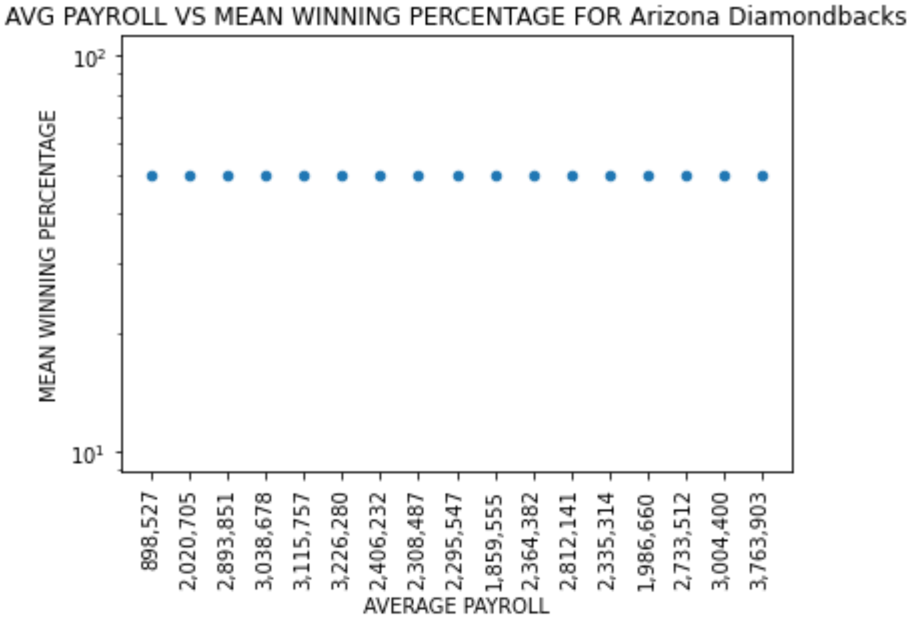
```
16    2001                         50.0       2,746,285
17    2002                         50.0       2,650,494
18    2003                         50.0       1,898,851
19    2004                         50.0       1,923,730
20    2005                         50.0       1,758,442
21    2006                         50.0       2,744,807
22    2007                         50.0       3,034,918
23    2008                         50.0       3,621,996
24    2009                         50.0       2,876,367
25    2010                         50.0       2,074,466
26    2011                         50.0       2,018,316
27    2012                         50.0       2,778,118
28    2013                         50.0       4,073,809
29    2014                         50.0       4,396,804,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
   yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0    2005                         50.0       1,619,383
1    2006                         50.0       2,036,870
2    2007                         50.0       1,319,553
3    2008                         50.0       1,895,206
4    2009                         50.0       2,140,285
5    2010                         50.0       2,046,666
6    2011                         50.0       2,201,963
7    2012                         50.0       2,695,171
8    2013                         50.0       4,548,130
9    2014                         50.0       4,399,456,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
   yearID  MEAN WINNING PERCENTAGE AVERAGE PAYROLL
0    2005                         50.0       1,619,383
1    2006                         50.0       2,036,870
2    2007                         50.0       1,319,553
3    2008                         50.0       1,895,206
4    2009                         50.0       2,140,285
5    2010                         50.0       2,046,666
6    2011                         50.0       2,201,963
```

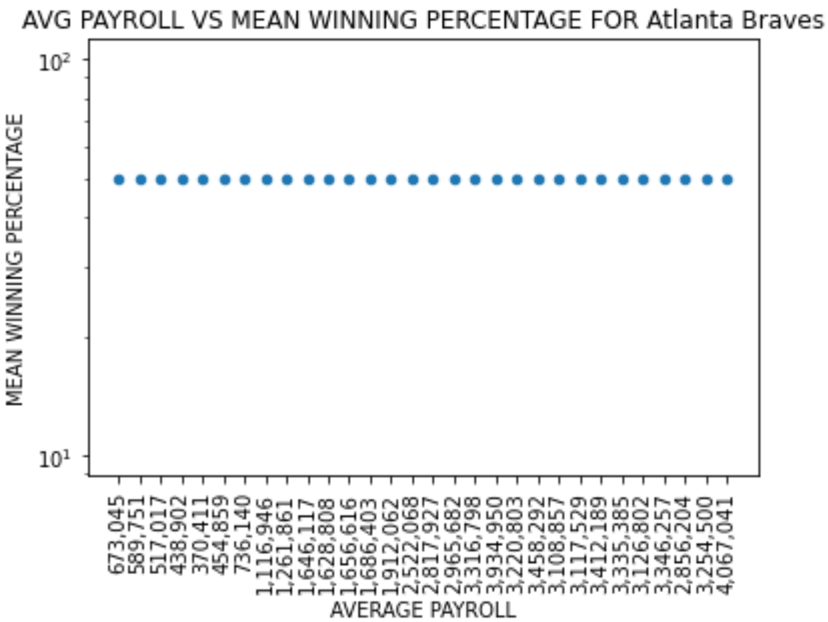
```
7    2012                         50.0       2,695,171
8    2013                         50.0       4,548,130
9    2014                         50.0       4,399,456,
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
Index: [],
Empty DataFrame
Columns: [yearID, MEAN WINNING PERCENTAGE, AVERAGE PAYROLL]
```

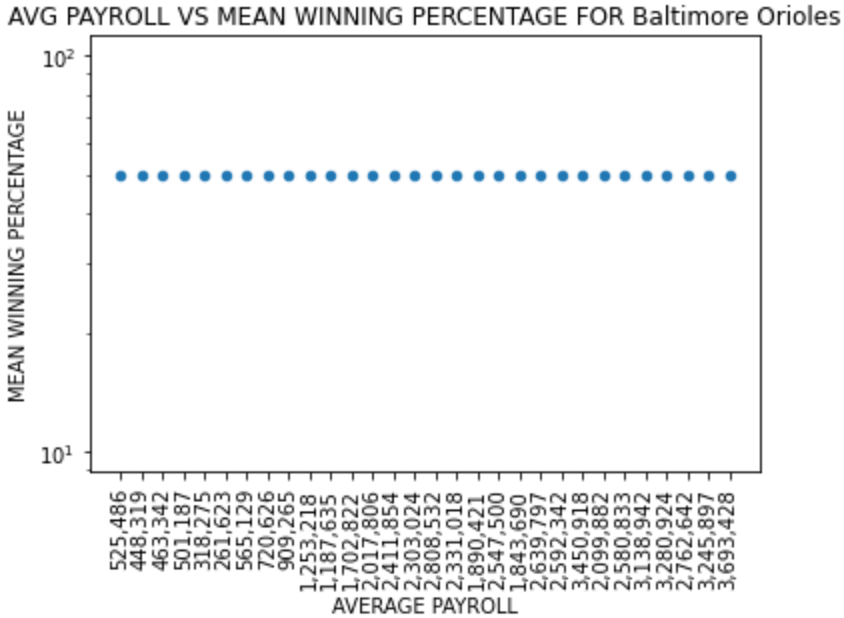
In [420…

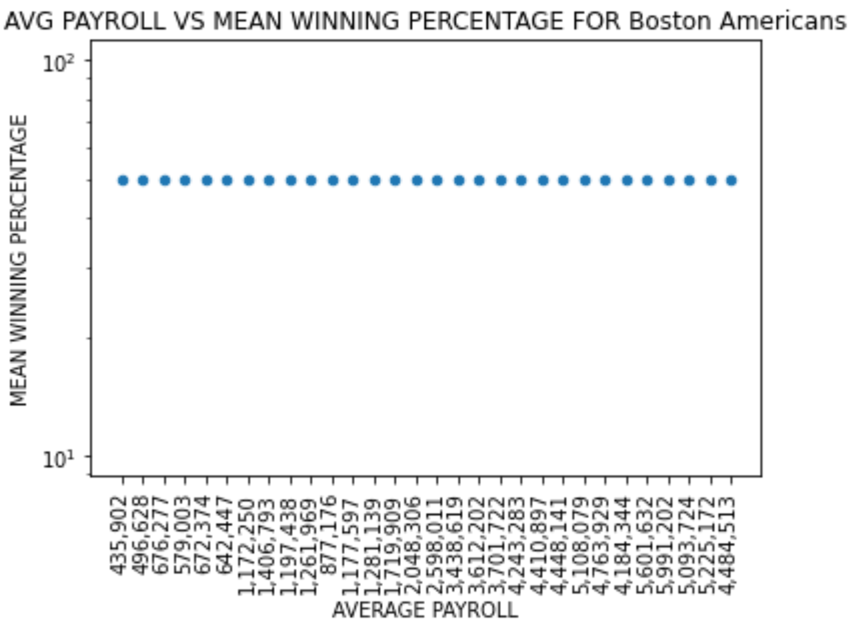
```
for index in range(len(some_results)):
    r = some_results[index]
    team_id = team_names[index]
    if r.empty == False:
        r.plot(x='AVERAGE PAYROLL', y='MEAN WINNING PERCENTAGE', rot=90, \
                logy=True, kind='scatter', title = 'AVG PAYROLL VS MEAN WINNI
```

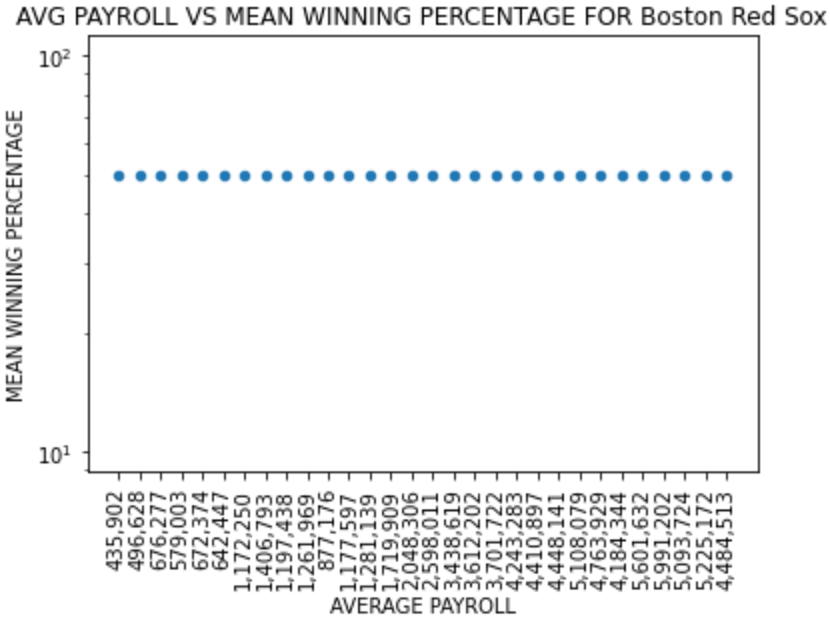


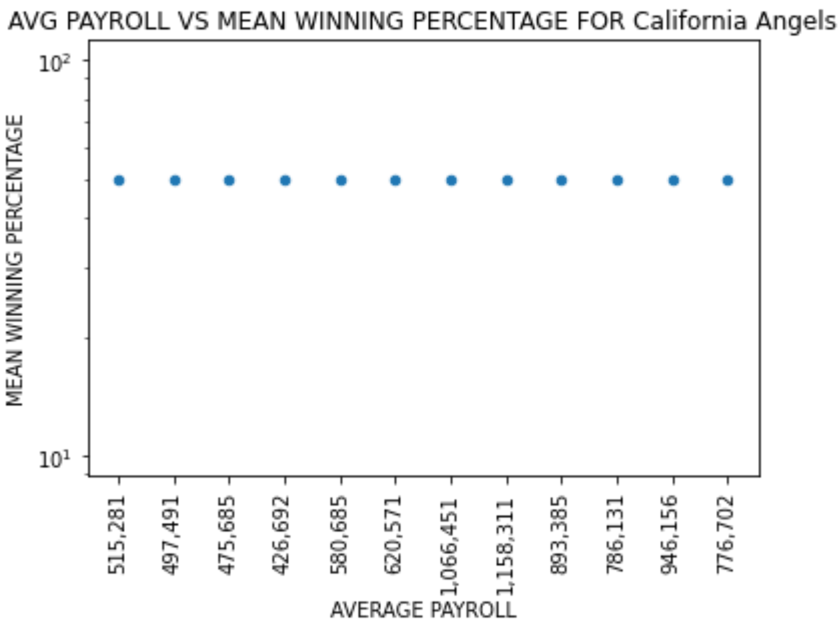

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Arizona Diamondbacks
MEAN WINNING PERCENTAGE
10^2
10^1
898,527
2,020,705
2,893,851
3,038,678
3,115,757
3,226,280
2,406,232
2,308,487
2,295,547
1,859,555
2,364,382
2,812,141
2,335,314
1,986,660
2,733,512
3,004,400
3,763,903
AVERAGE PAYROLL

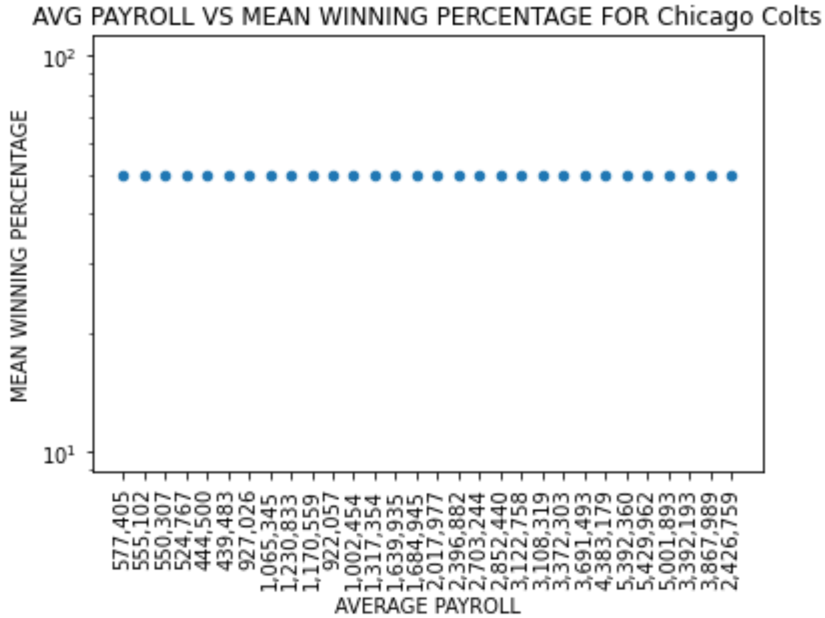

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Atlanta Braves
MEAN WINNING PERCENTAGE
10^2
10^1
673,045
589,751
517,017
438,902
370,411
454,859
736,140
1,116,946
1,261,861
1,646,117
1,628,808
1,656,616
1,686,403
1,912,062
2,522,068
2,817,927
2,965,682
3,316,798
3,934,950
3,220,803
3,458,292
3,108,857
3,117,529
3,412,189
3,335,385
3,126,802
3,346,257
2,856,204
3,254,500
4,067,041
AVERAGE PAYROLL

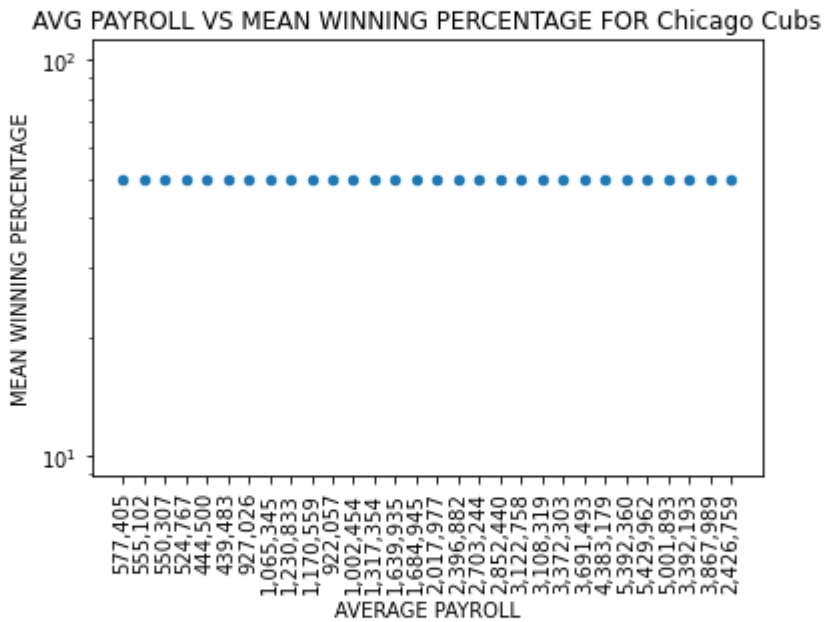

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Baltimore Orioles
MEAN WINNING PERCENTAGE
10^2
10^1
525,486
448,319
463,342
501,187
318,275
261,623
565,129
720,626
909,265
1,253,218
1,187,635
1,702,822
2,017,806
2,411,854
2,303,024
2,808,532
2,331,018
1,890,421
2,547,500
1,843,690
2,639,797
2,592,342
3,450,918
2,099,882
2,580,833
3,138,942
3,280,924
2,762,642
3,245,897
3,693,428
AVERAGE PAYROLL

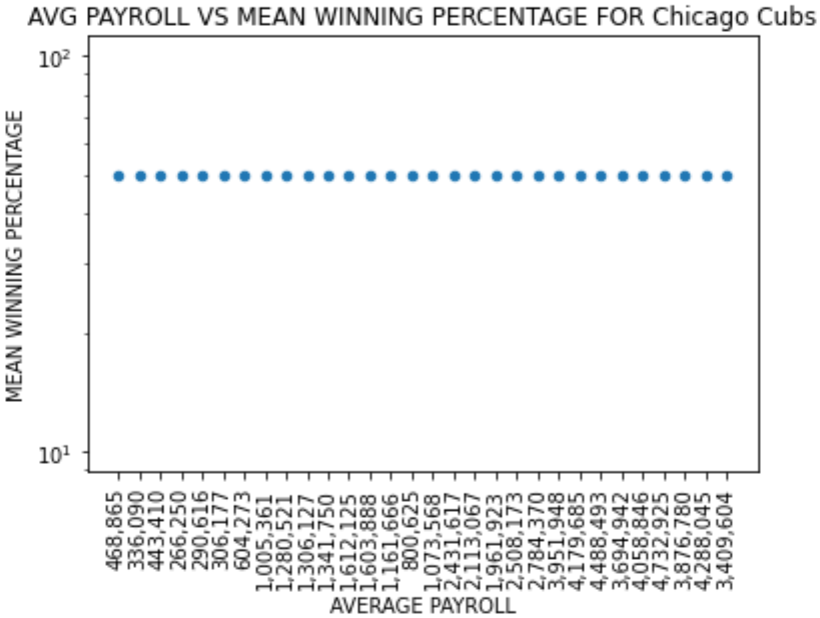

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Boston Americans
MEAN WINNING PERCENTAGE
10^2
10^1
435,902
496,628
676,277
579,003
672,374
642,447
1,172,250
1,406,793
1,197,438
1,261,969
877,176
1,177,597
1,281,139
1,719,909
2,048,306
2,598,011
3,438,619
3,612,202
3,701,722
4,243,283
4,410,897
4,448,141
5,108,079
4,763,929
4,184,344
5,601,632
5,991,202
5,093,724
5,225,172
4,484,513
AVERAGE PAYROLL

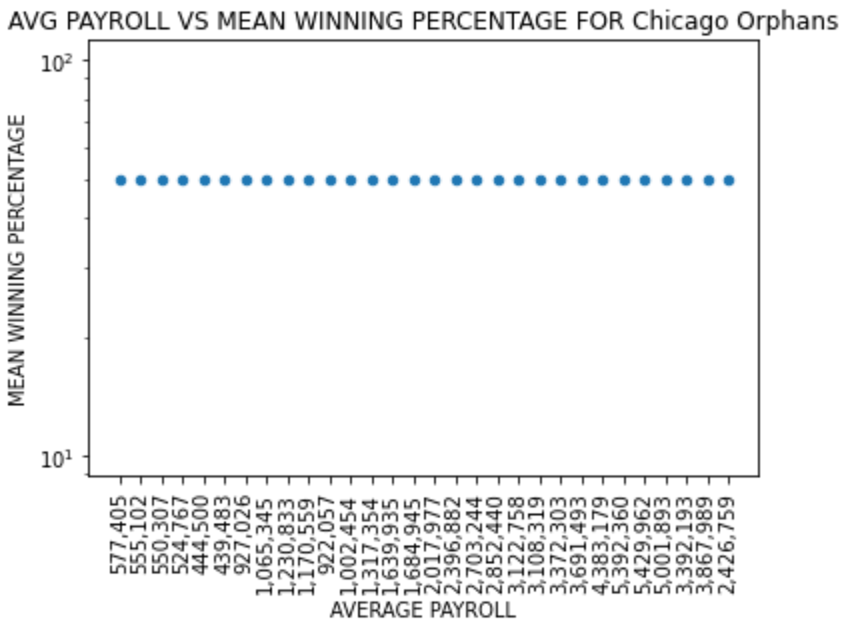

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Boston Red Sox
MEAN WINNING PERCENTAGE
AVERAGE PAYROLL
435,902
496,628
676,277
579,003
672,374
642,447
1,172,250
1,406,793
1,197,438
1,261,969
877,176
1,177,597
1,281,139
1,719,909
2,048,306
2,598,011
3,438,619
3,612,202
3,701,722
4,243,283
4,410,897
4,448,141
5,108,079
4,763,929
4,184,344
5,601,632
5,991,202
5,093,724
5,225,172
4,484,513

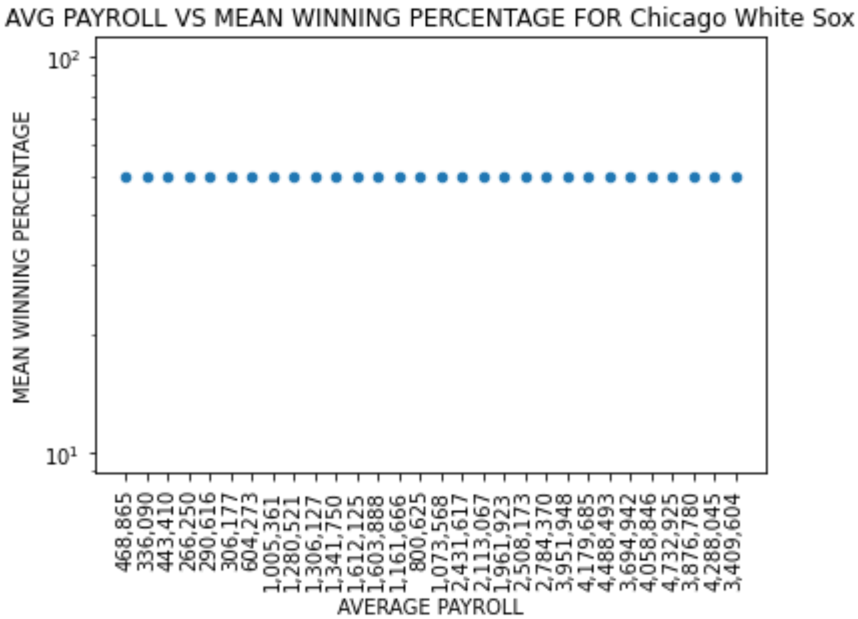

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR California Angels
MEAN WINNING PERCENTAGE
AVERAGE PAYROLL
515,281
497,491
475,685
426,692
580,685
620,571
1,066,451
1,158,311
893,385
786,131
946,156
776,702

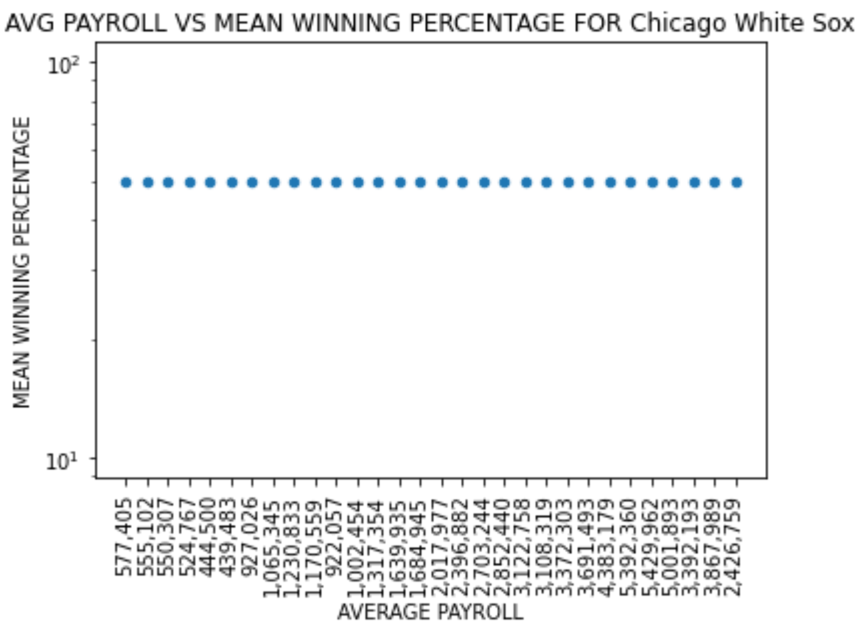

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Chicago Colts
MEAN WINNING PERCENTAGE
10^2
10^1
577,405
555,102
550,307
524,767
444,500
439,483
927,026
1,065,345
1,230,833
1,170,559
922,057
1,002,454
1,317,354
1,639,935
1,684,945
2,017,977
2,396,882
2,703,244
2,852,440
3,122,758
3,108,319
3,372,303
3,691,493
4,383,179
5,392,360
5,429,962
5,001,893
3,392,193
3,867,989
2,426,759
AVERAGE PAYROLL

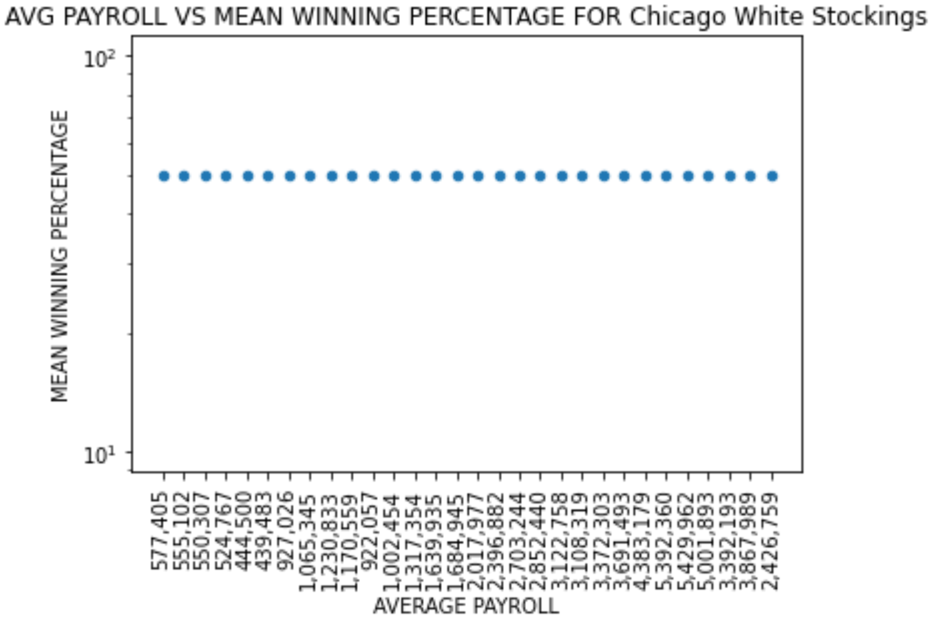

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Chicago Cubs
MEAN WINNING PERCENTAGE
10^2
10^1
577,405
555,102
550,307
524,767
444,500
439,483
927,026
1,065,345
1,230,833
1,170,559
922,057
1,002,454
1,317,354
1,639,935
1,684,945
2,017,977
2,396,882
2,703,244
2,852,440
3,122,758
3,108,319
3,372,303
3,691,493
4,383,179
5,392,360
5,429,962
5,001,893
3,392,193
3,867,989
2,426,759
AVERAGE PAYROLL

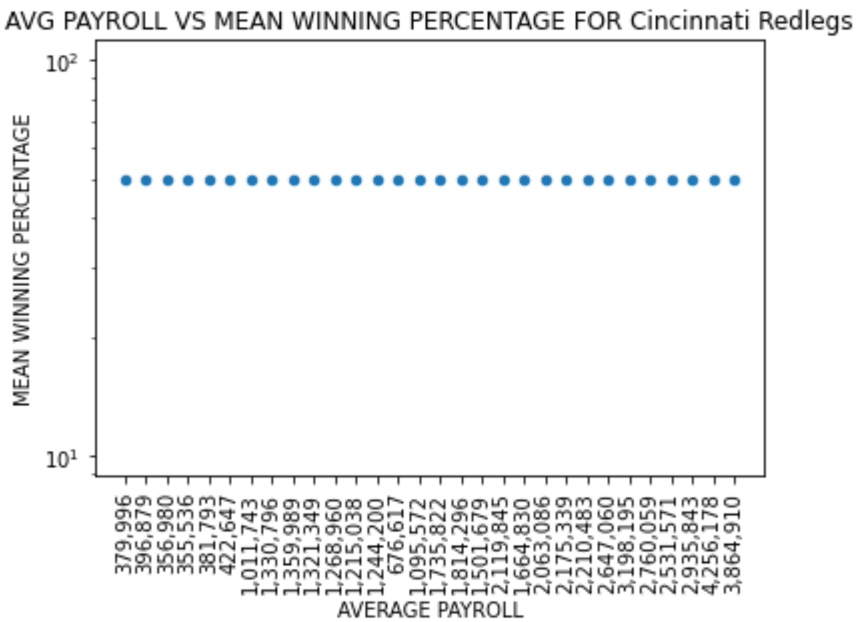

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Chicago Cubs
MEAN WINNING PERCENTAGE
AVERAGE PAYROLL
468,865
336,090
443,410
266,250
290,616
306,177
604,273
1,005,361
1,280,521
1,306,127
1,341,750
1,612,125
1,603,888
1,161,666
800,625
1,073,568
2,431,617
2,113,067
1,961,923
2,508,173
2,784,370
3,951,948
4,179,685
4,488,493
3,694,942
4,058,846
4,732,925
3,876,780
4,288,045
3,409,604

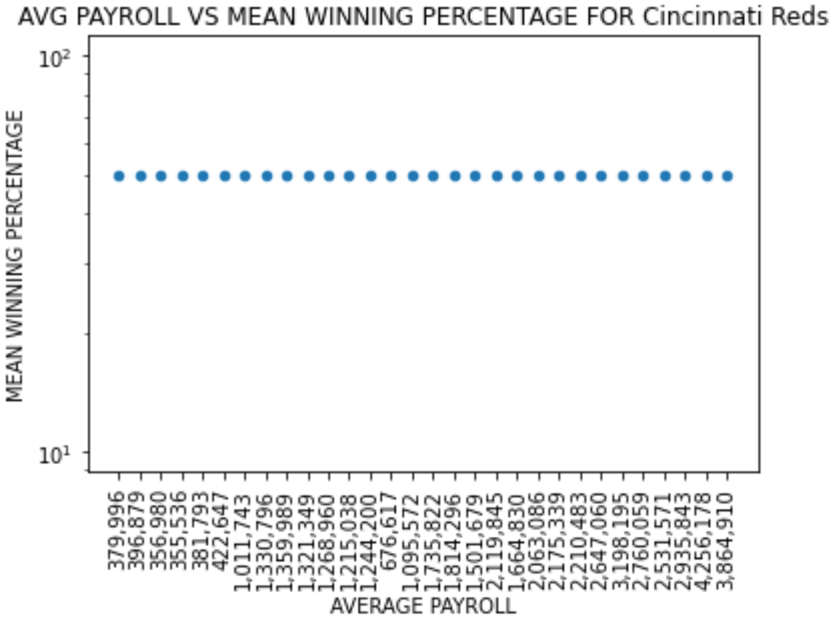

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Chicago Orphans
MEAN WINNING PERCENTAGE
AVERAGE PAYROLL
577,405
555,102
550,307
524,767
444,500
439,483
927,026
1,065,345
1,230,833
1,170,559
922,057
1,002,454
1,317,354
1,639,935
1,684,945
2,017,977
2,396,882
2,703,244
2,852,440
3,122,758
3,108,319
3,372,303
3,691,493
4,383,179
5,392,360
5,429,962
5,001,893
3,392,193
3,867,989
2,426,759

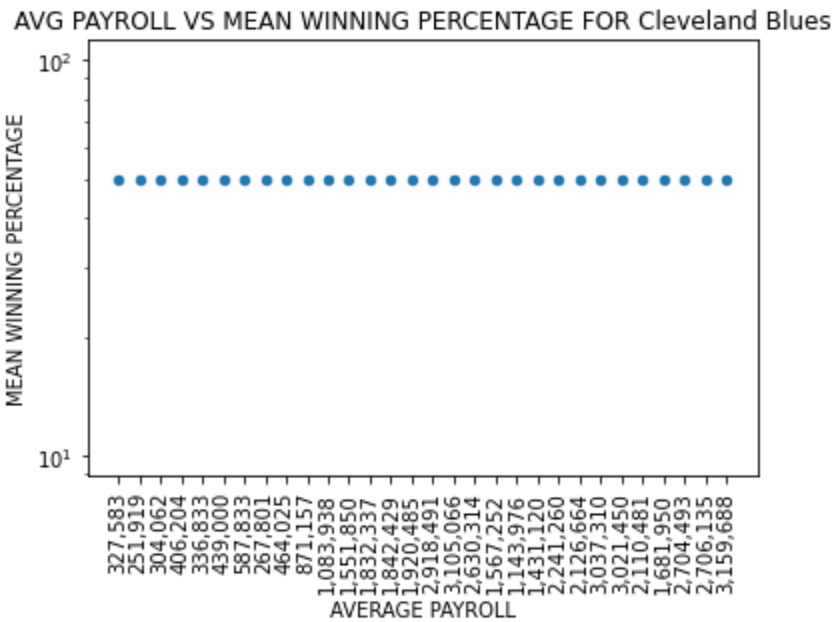

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Chicago White Sox
10^2
10^1
MEAN WINNING PERCENTAGE
468,865
336,090
443,410
266,250
290,616
306,177
604,273
1,005,361
1,280,521
1,306,127
1,341,750
1,612,125
1,603,888
1,161,666
800,625
1,073,568
2,431,617
2,113,067
1,961,923
2,508,173
2,784,370
3,951,948
4,179,685
4,488,493
3,694,942
4,058,846
4,732,925
3,876,780
4,288,045
3,409,604
AVERAGE PAYROLL

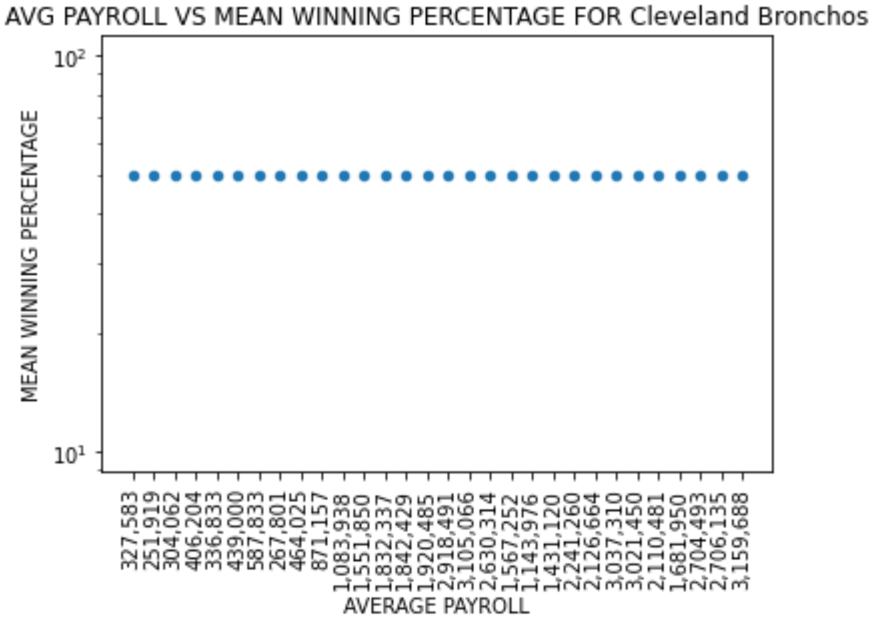

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Chicago White Sox
10^2
10^1
MEAN WINNING PERCENTAGE
577,405
555,102
550,307
524,767
444,500
439,483
927,026
1,065,345
1,230,833
1,170,559
922,057
1,002,454
1,317,354
1,639,935
1,684,945
2,017,977
2,396,882
2,703,244
2,852,440
3,122,758
3,108,319
3,372,303
3,691,493
4,383,179
5,392,360
5,429,962
5,001,893
3,392,193
3,867,989
2,426,759
AVERAGE PAYROLL

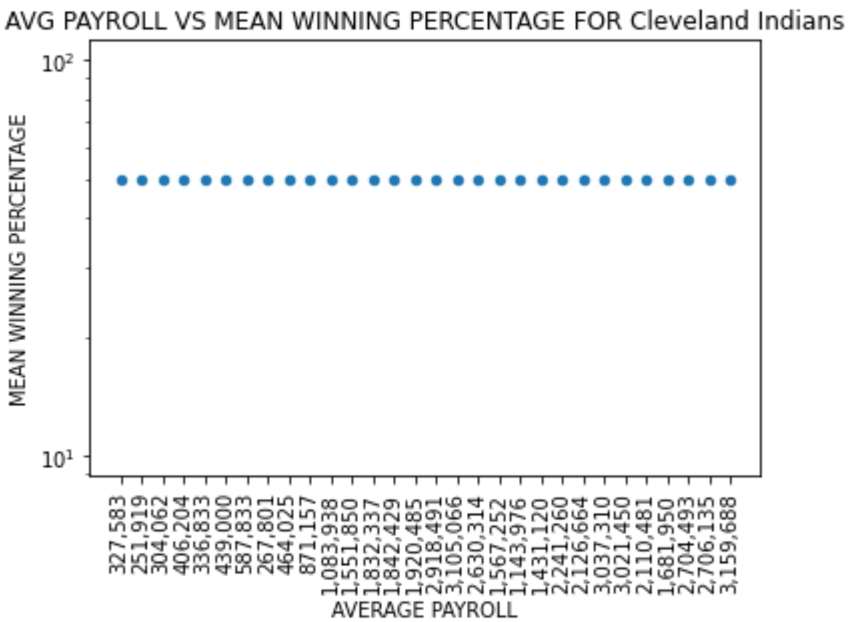

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Chicago White Stockings
10^2
10^1
MEAN WINNING PERCENTAGE
577,405
555,102
550,307
524,767
444,500
439,483
927,026
1,065,345
1,230,833
1,170,559
922,057
1,002,454
1,317,354
1,639,935
1,684,945
2,017,977
2,396,882
2,703,244
2,852,440
3,122,758
3,108,319
3,372,303
3,691,493
4,383,179
5,392,360
5,429,962
5,001,893
3,392,193
3,867,989
2,426,759
AVERAGE PAYROLL

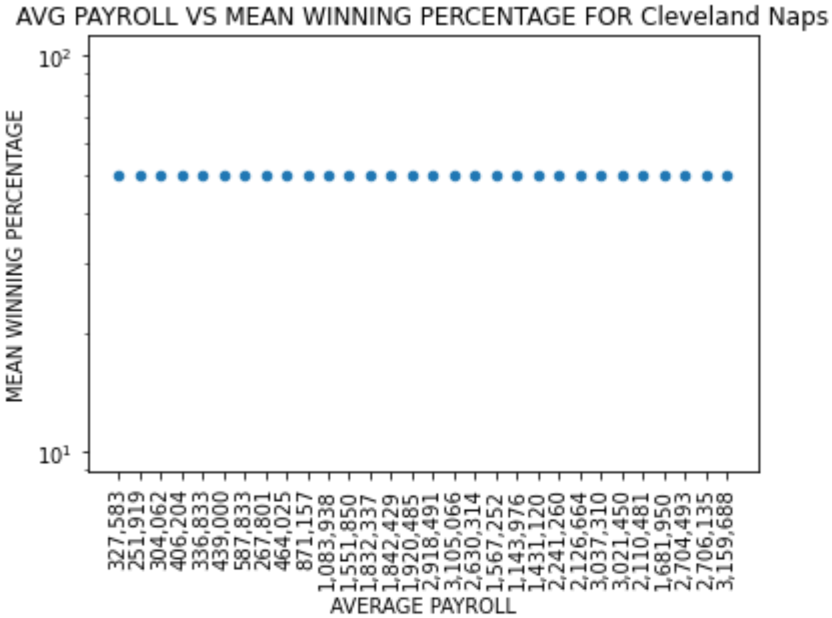

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Cincinnati Redlegs
10^2
10^1
MEAN WINNING PERCENTAGE
379,996
396,879
356,980
355,536
381,793
422,647
1,011,743
1,330,796
1,359,989
1,321,349
1,268,960
1,215,038
1,244,200
676,617
1,095,572
1,735,822
1,814,296
1,501,679
2,119,845
1,664,830
2,063,086
2,175,339
2,210,483
2,647,060
3,198,195
2,760,059
2,531,571
2,935,843
4,256,178
3,864,910
AVERAGE PAYROLL

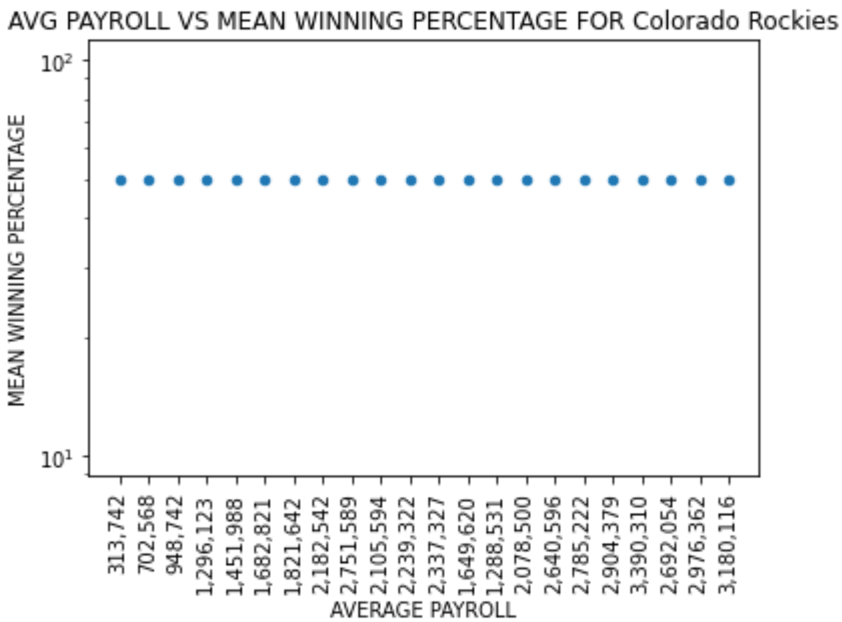

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Cincinnati Reds
MEAN WINNING PERCENTAGE
AVERAGE PAYROLL
379,996
396,879
356,980
355,536
381,793
422,647
1,011,743
1,330,796
1,359,989
1,321,349
1,268,960
1,215,038
1,244,200
676,617
1,095,572
1,735,822
1,814,296
1,501,679
2,119,845
1,664,830
2,063,086
2,175,339
2,210,483
2,647,060
3,198,195
2,760,059
2,531,571
2,935,843
4,256,178
3,864,910

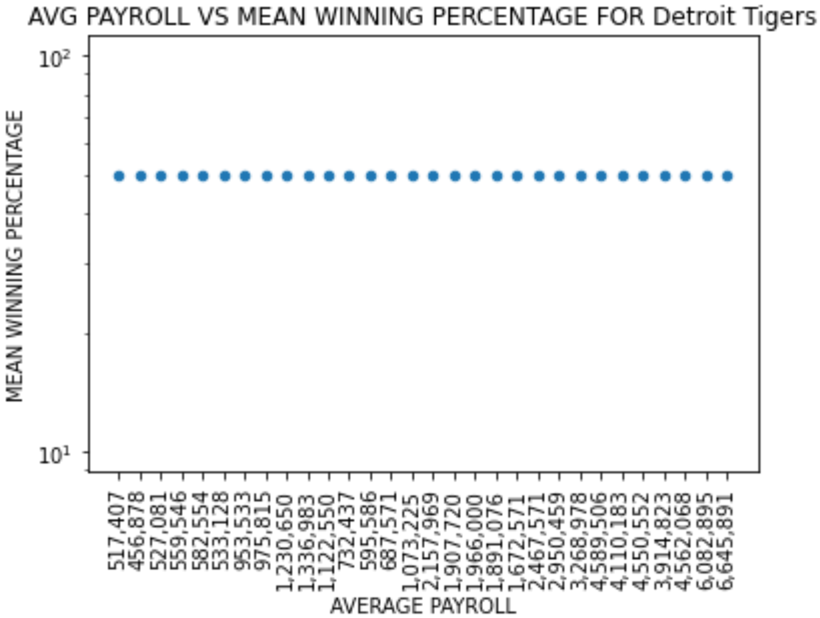

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Cleveland Blues
MEAN WINNING PERCENTAGE
AVERAGE PAYROLL
327,583
251,919
304,062
406,204
336,833
439,000
587,833
267,801
464,025
871,157
1,083,938
1,551,850
1,832,337
1,842,429
1,920,485
2,918,491
3,105,066
2,630,314
1,567,252
1,143,976
1,431,120
2,241,260
2,126,664
3,037,310
3,021,450
2,110,481
1,681,950
2,704,493
2,706,135
3,159,688

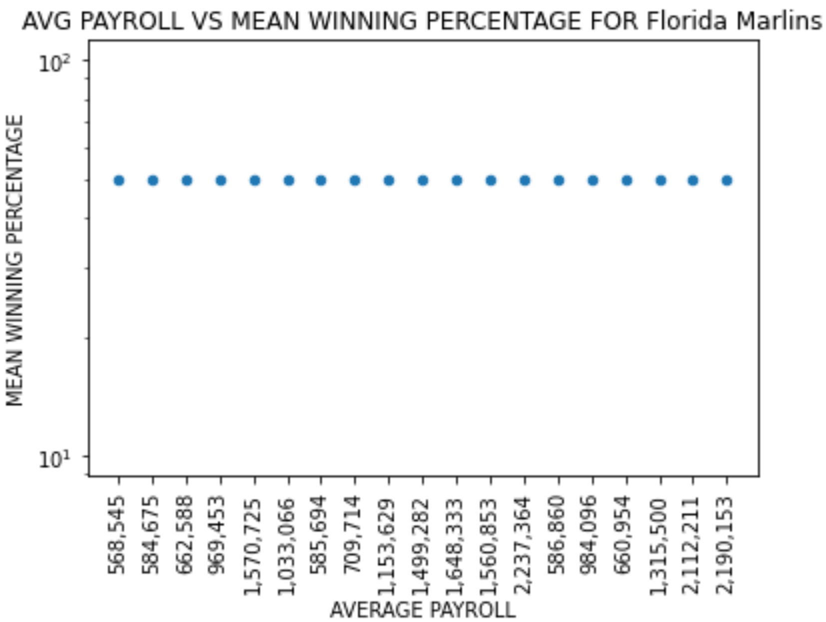

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Cleveland Bronchos
MEAN WINNING PERCENTAGE
10^2
10^1
327,583
251,919
304,062
406,204
336,833
439,000
587,833
267,801
464,025
871,157
1,083,938
1,551,850
1,832,337
1,842,429
1,920,485
2,918,491
3,105,066
2,630,314
1,567,252
1,143,976
1,431,120
2,241,260
2,126,664
3,037,310
3,021,450
2,110,481
1,681,950
2,704,493
2,706,135
3,159,688
AVERAGE PAYROLL

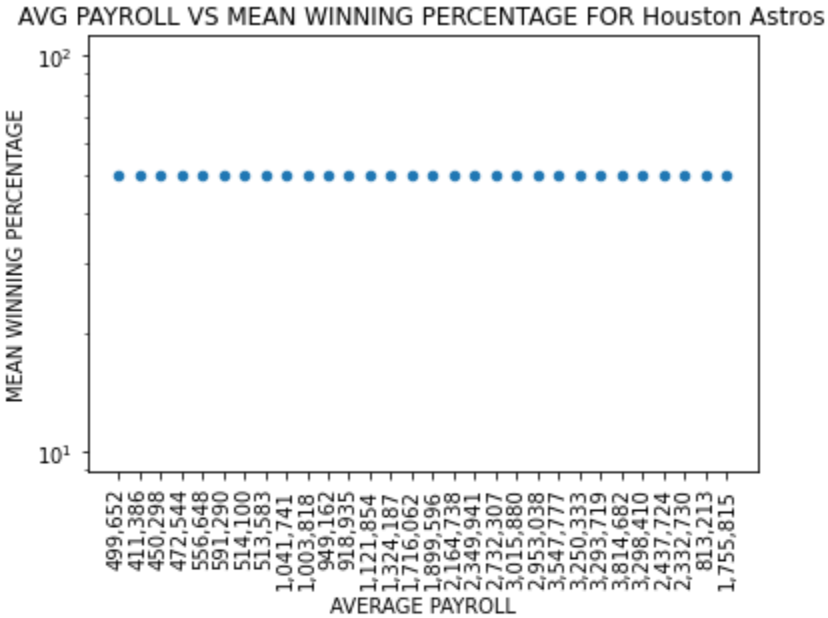

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Cleveland Indians
MEAN WINNING PERCENTAGE
10^2
10^1
327,583
251,919
304,062
406,204
336,833
439,000
587,833
267,801
464,025
871,157
1,083,938
1,551,850
1,832,337
1,842,429
1,920,485
2,918,491
3,105,066
2,630,314
1,567,252
1,143,976
1,431,120
2,241,260
2,126,664
3,037,310
3,021,450
2,110,481
1,681,950
2,704,493
2,706,135
3,159,688
AVERAGE PAYROLL

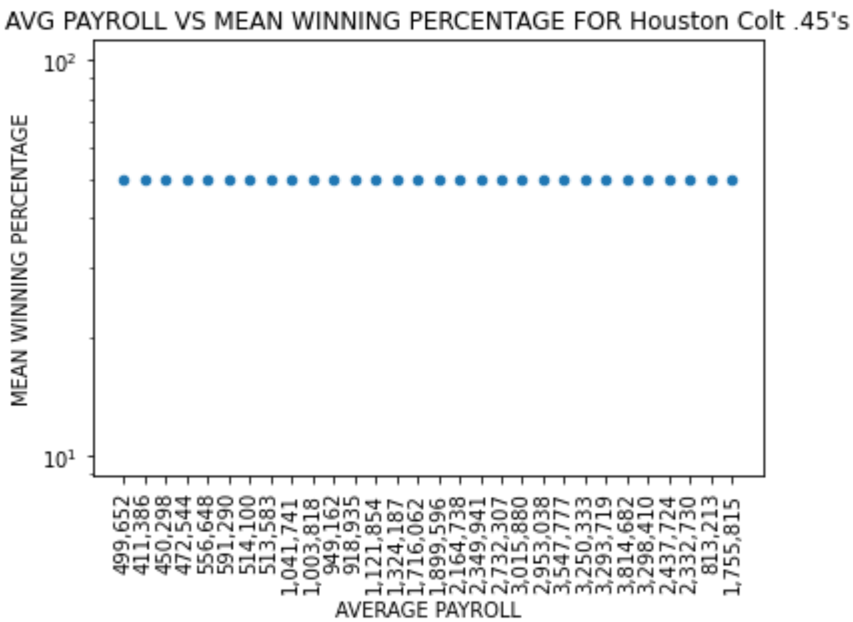

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Cleveland Naps
MEAN WINNING PERCENTAGE
AVERAGE PAYROLL


AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Colorado Rockies
MEAN WINNING PERCENTAGE
AVERAGE PAYROLL

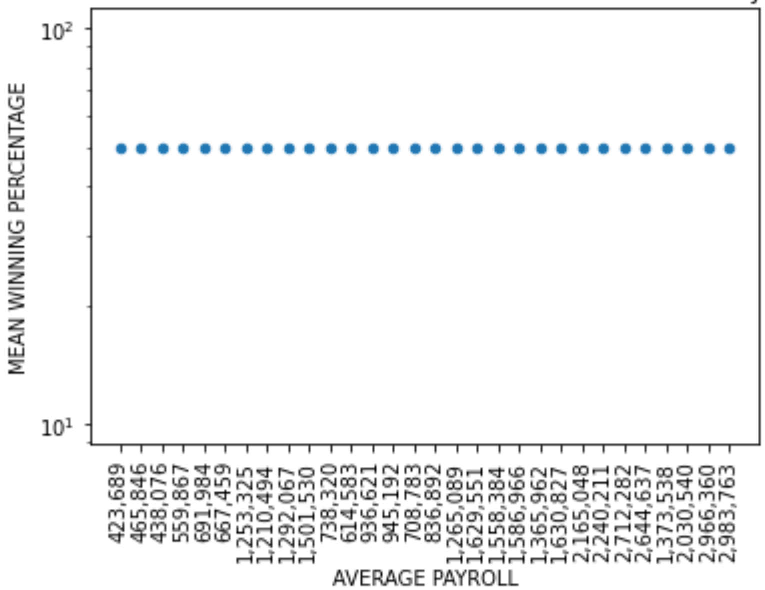

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Detroit Tigers
MEAN WINNING PERCENTAGE
10^2
10^1
517,407
456,878
527,081
559,546
582,554
533,128
953,533
975,815
1,230,650
1,336,983
1,122,550
732,437
595,586
687,571
1,073,225
2,157,969
1,907,720
1,966,000
1,891,076
1,672,571
2,467,571
2,950,459
3,268,978
4,589,506
4,110,183
4,550,552
3,914,823
4,562,068
6,082,895
6,645,891
AVERAGE PAYROLL


AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Florida Marlins
MEAN WINNING PERCENTAGE
10^2
10^1
568,545
584,675
662,588
969,453
1,570,725
1,033,066
585,694
709,714
1,153,629
1,499,282
1,648,333
1,560,853
2,237,364
586,860
984,096
660,954
1,315,500
2,112,211
2,190,153
AVERAGE PAYROLL

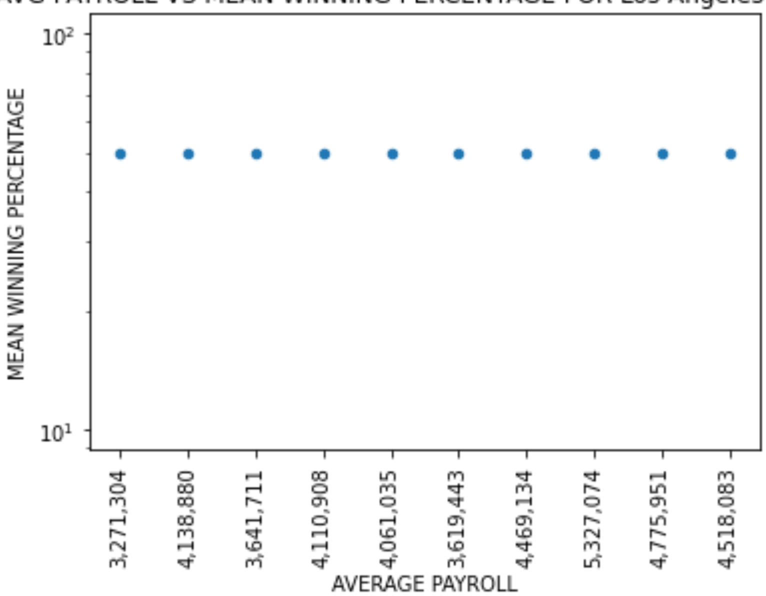

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Houston Astros
MEAN WINNING PERCENTAGE
AVERAGE PAYROLL

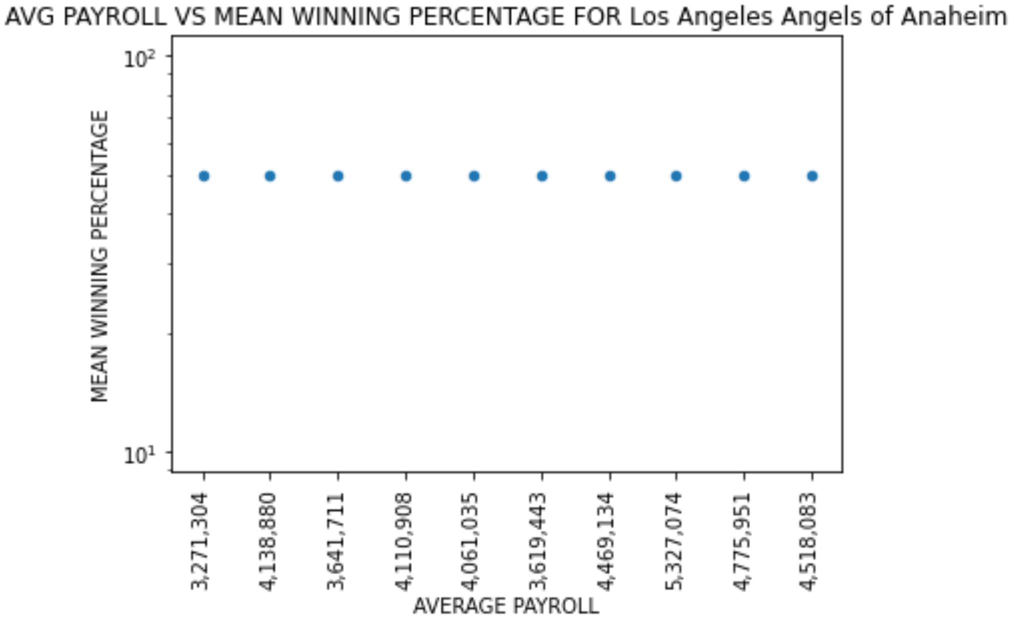

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Houston Colt .45's
MEAN WINNING PERCENTAGE
AVERAGE PAYROLL

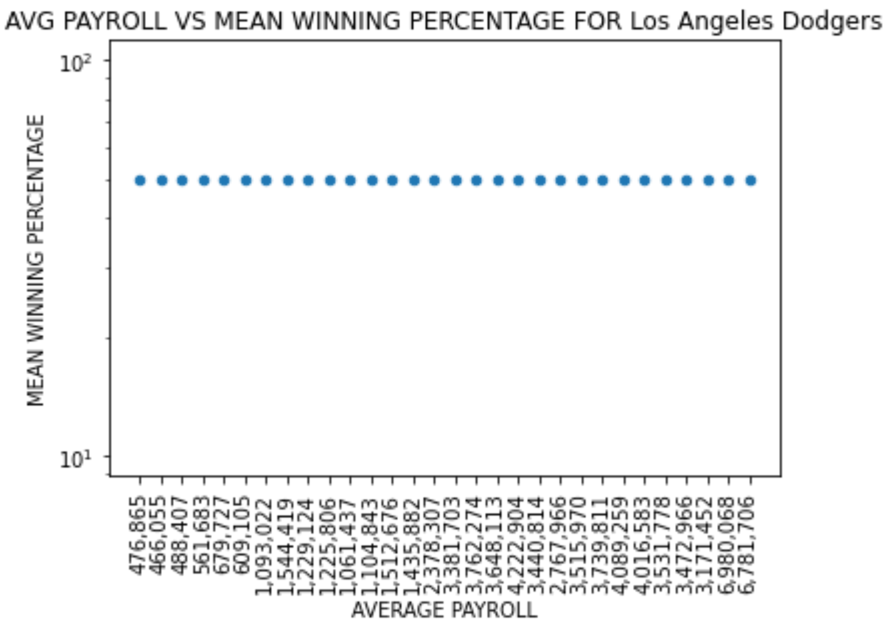
AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Kansas City Royals

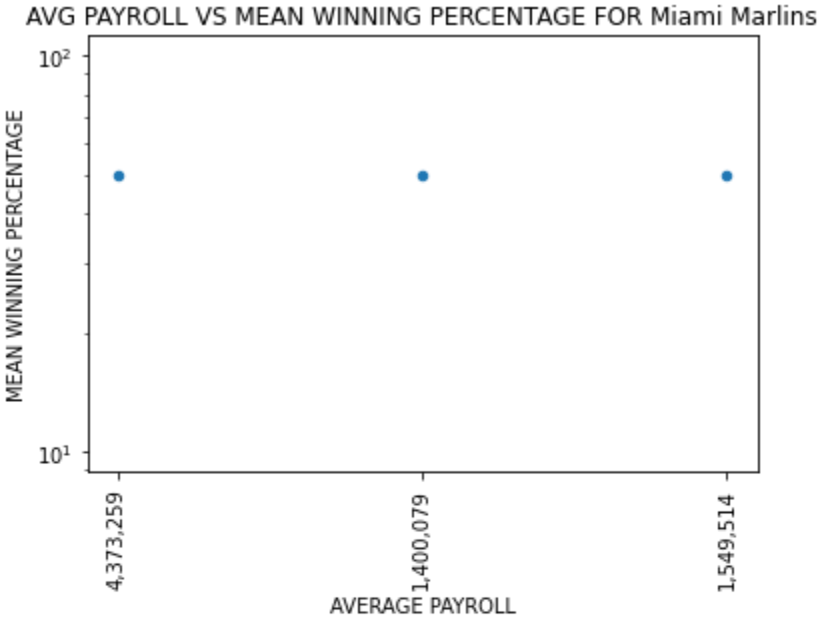


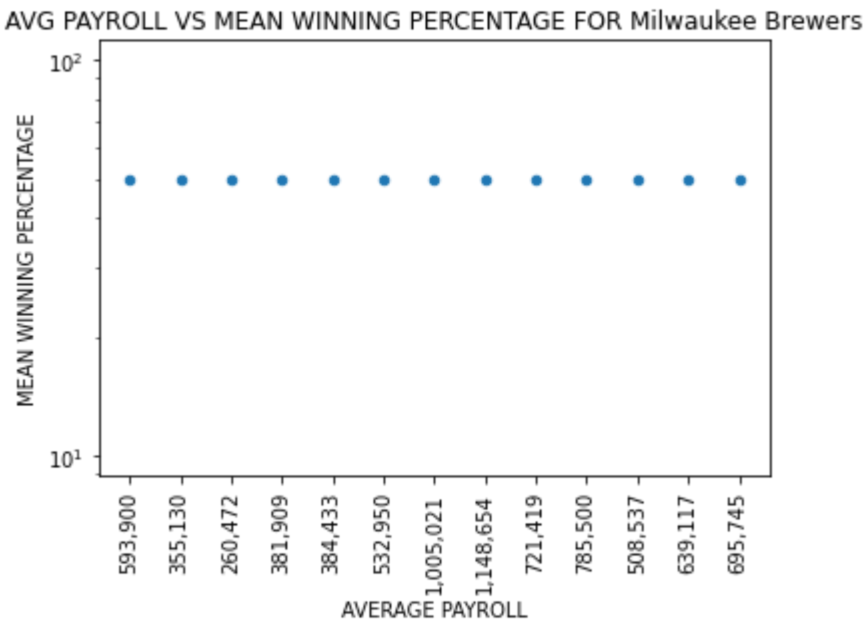
AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Los Angeles Angels

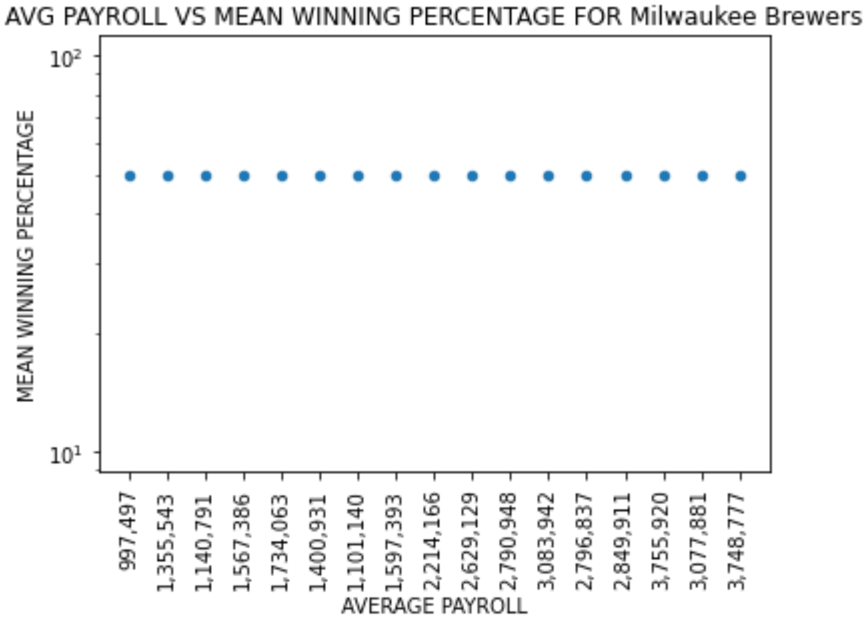


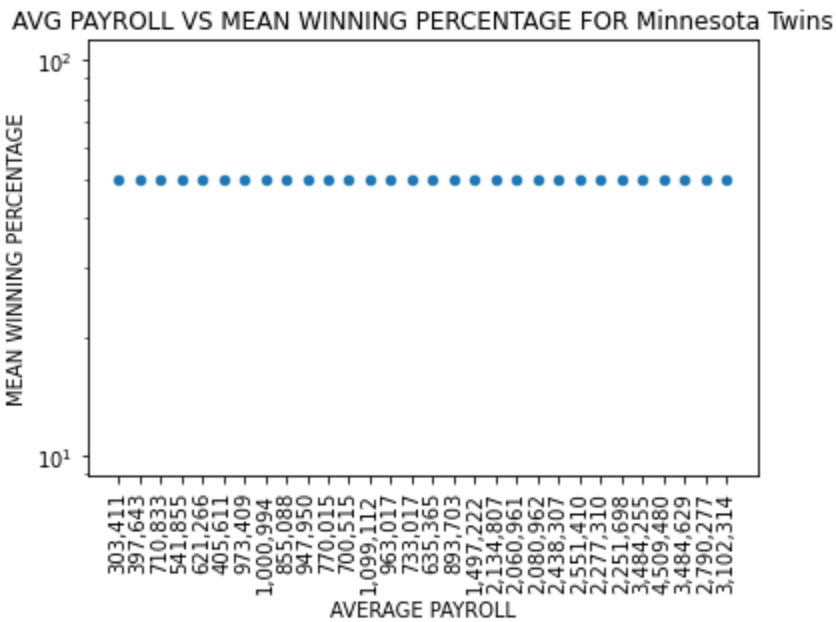

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Los Angeles Angels of Anaheim
10^2
10^1
MEAN WINNING PERCENTAGE
3,271,304
4,138,880
3,641,711
4,110,908
4,061,035
3,619,443
4,469,134
5,327,074
4,775,951
4,518,083
AVERAGE PAYROLL

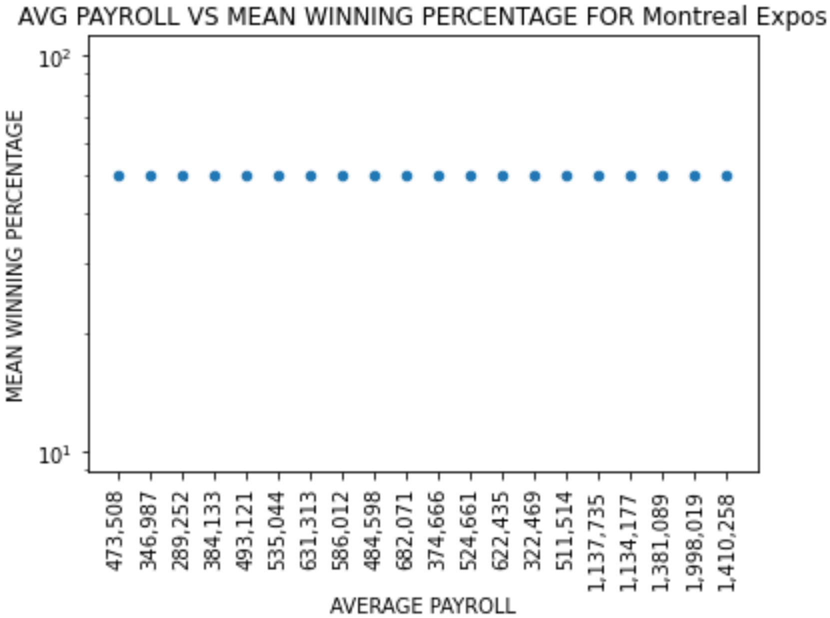

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Los Angeles Dodgers
10^2
10^1
MEAN WINNING PERCENTAGE
476,865
466,055
488,407
561,683
679,727
609,105
1,093,022
1,544,419
1,229,124
1,225,806
1,061,437
1,104,843
1,512,676
1,435,882
2,378,307
3,381,703
3,762,274
3,648,113
4,222,904
3,440,814
2,767,966
3,515,970
3,739,811
4,089,259
4,016,583
3,531,778
3,472,966
3,171,452
6,980,068
6,781,706
AVERAGE PAYROLL

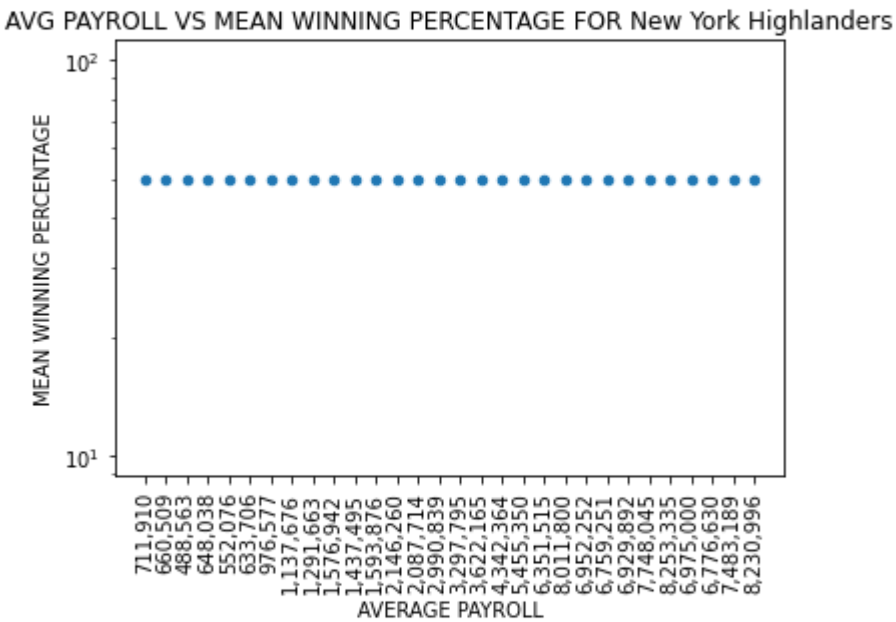

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Miami Marlins
MEAN WINNING PERCENTAGE
10^2
10^1
4,373,259
1,400,079
1,549,514
AVERAGE PAYROLL

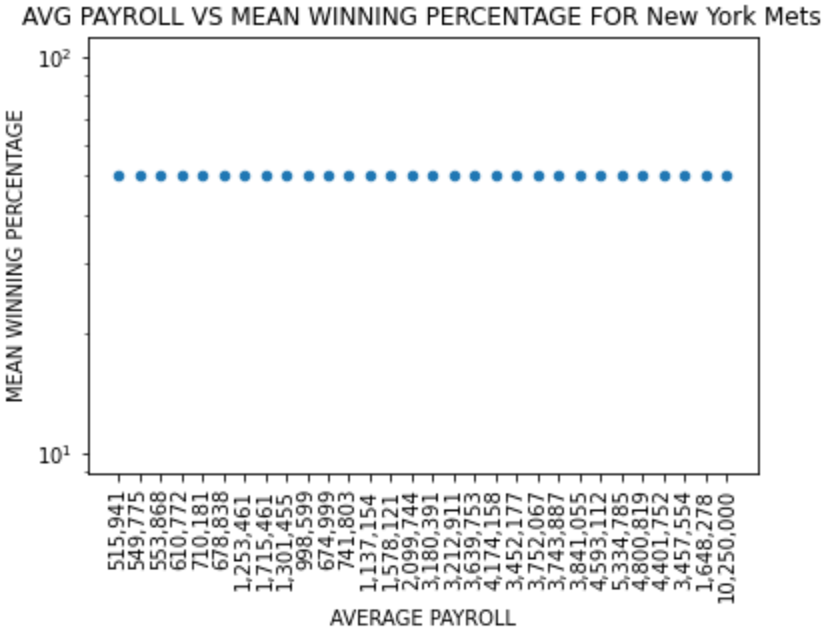

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Milwaukee Brewers
MEAN WINNING PERCENTAGE
10^2
10^1
593,900
355,130
260,472
381,909
384,433
532,950
1,005,021
1,148,654
721,419
785,500
508,537
639,117
695,745
AVERAGE PAYROLL

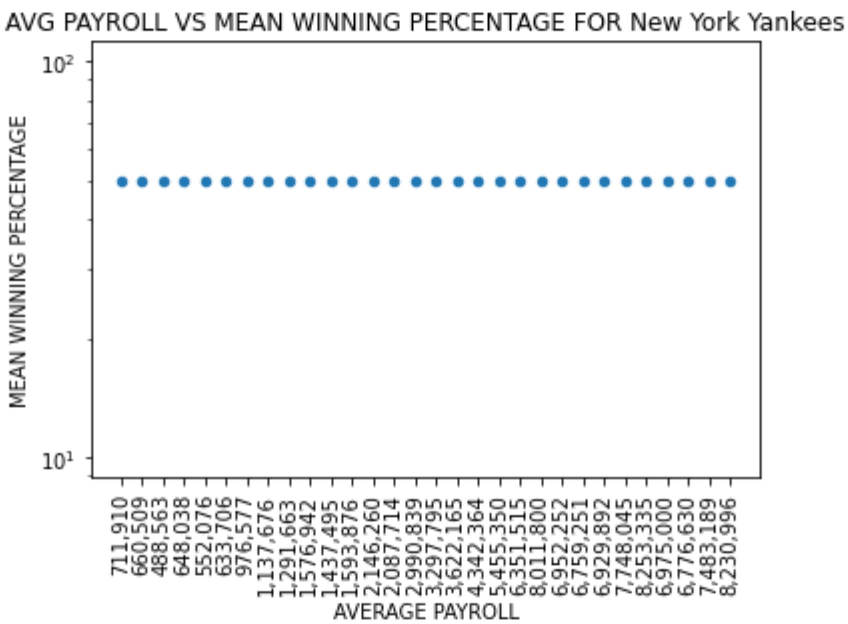

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Milwaukee Brewers
10^2
10^1
MEAN WINNING PERCENTAGE
997,497
1,355,543
1,140,791
1,567,386
1,734,063
1,400,931
1,101,140
1,597,393
2,214,166
2,629,129
2,790,948
3,083,942
2,796,837
2,849,911
3,755,920
3,077,881
3,748,777
AVERAGE PAYROLL

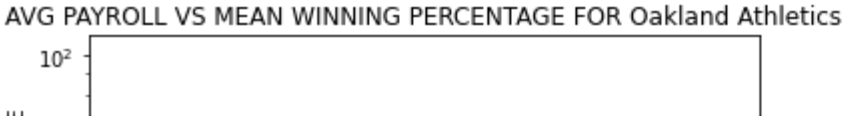

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Minnesota Twins
10^2
10^1
MEAN WINNING PERCENTAGE
303,411
397,643
710,833
541,855
621,266
405,611
973,409
1,000,994
855,088
947,950
770,015
700,515
1,099,112
963,017
733,017
635,365
893,703
1,497,222
2,134,807
2,060,961
2,080,962
2,438,307
2,551,410
2,277,310
2,251,698
3,484,255
4,509,480
3,484,629
2,790,277
3,102,314
AVERAGE PAYROLL


AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR Montreal Expos
MEAN WINNING PERCENTAGE
10^2
10^1
473,508
346,987
289,252
384,133
493,121
535,044
631,313
586,012
484,598
682,071
374,666
524,661
622,435
322,469
511,514
1,137,735
1,134,177
1,381,089
1,998,019
1,410,258
AVERAGE PAYROLL


AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR New York Highlanders
MEAN WINNING PERCENTAGE
10^2
10^1
711,910
660,509
488,563
648,038
552,076
633,706
976,577
1,137,676
1,291,663
1,576,942
1,437,495
1,593,876
2,146,260
2,087,714
2,990,839
3,297,795
3,622,165
4,342,364
5,455,350
6,351,515
8,011,800
6,952,252
6,759,251
6,929,892
7,748,045
8,253,335
6,975,000
6,776,630
7,483,189
8,230,996
AVERAGE PAYROLL

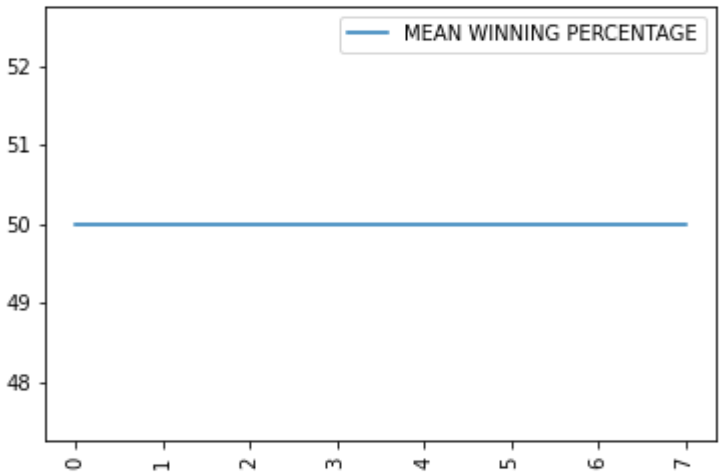

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR New York Mets
MEAN WINNING PERCENTAGE
AVERAGE PAYROLL

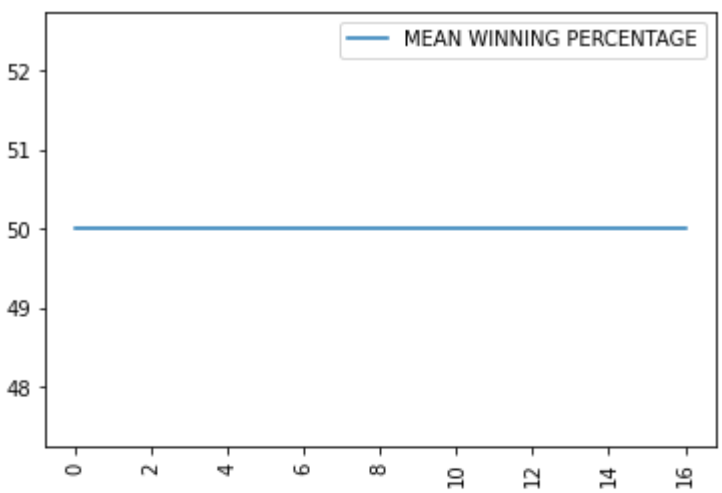

AVG PAYROLL VS MEAN WINNING PERCENTAGE FOR New York Yankees
MEAN WINNING PERCENTAGE
AVERAGE PAYROLL

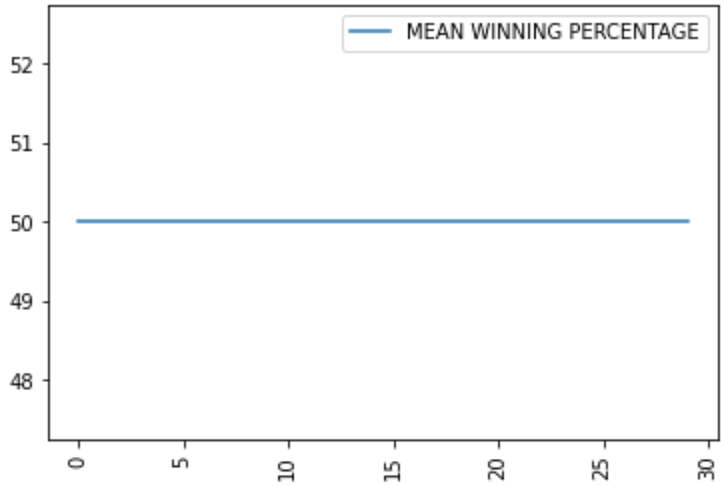


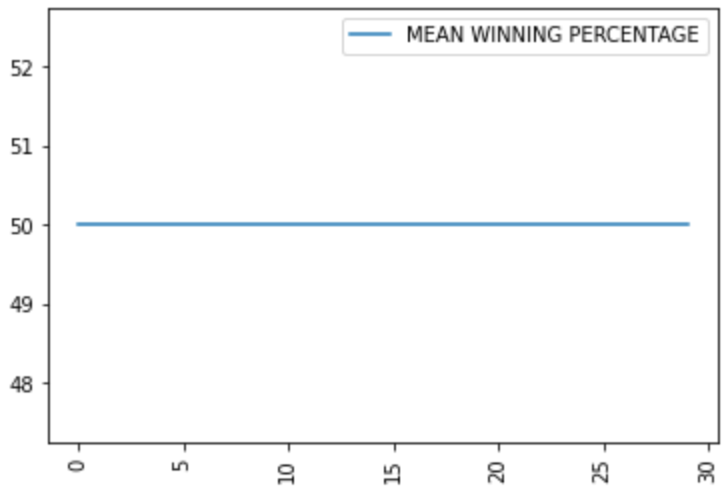
In [421…

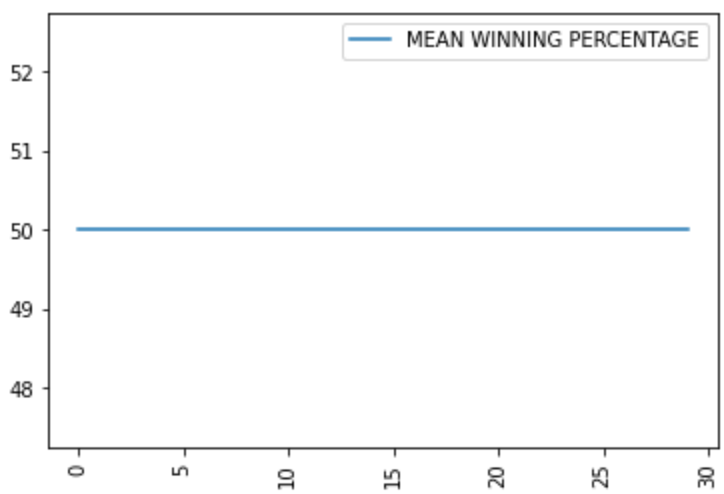
```
for r in some_results:
    if not r.empty:
        u = pandas.cut(r['yearID'], bins=5)
        r['Time Periods'] = u
        a = r.groupby(['Time Periods'])[['AVERAGE PAYROLL','MEAN WINNING PERC

        r.plot(y=['AVERAGE PAYROLL','MEAN WINNING PERCENTAGE'], rot=90 )
```

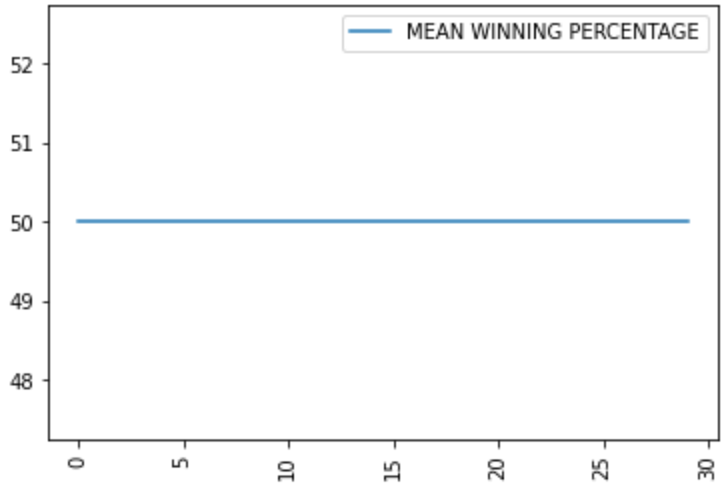


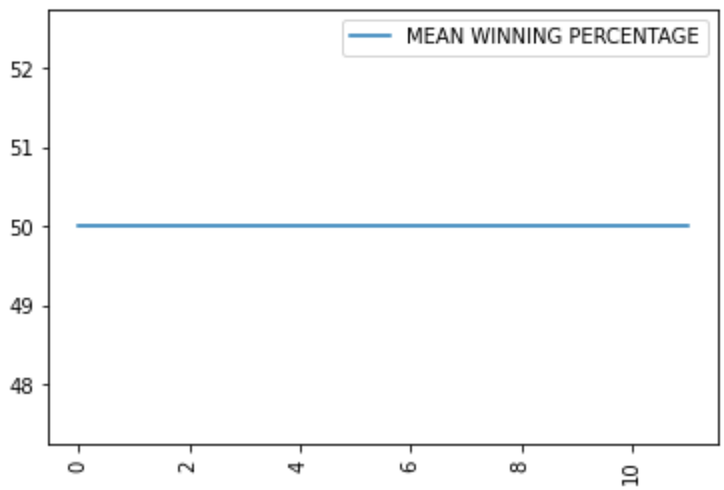


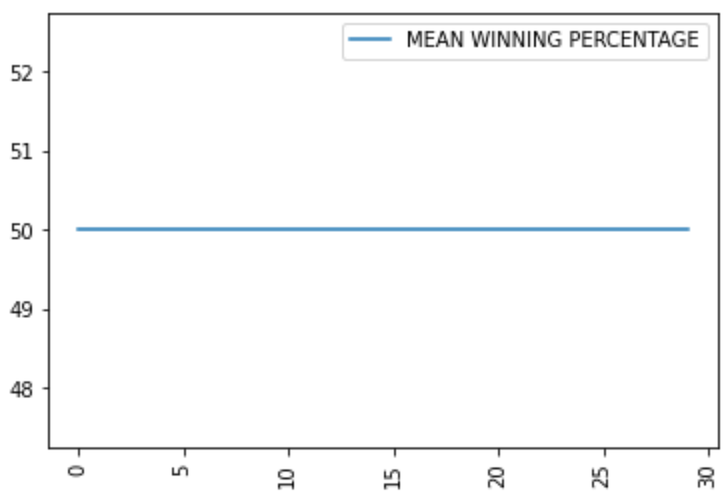

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

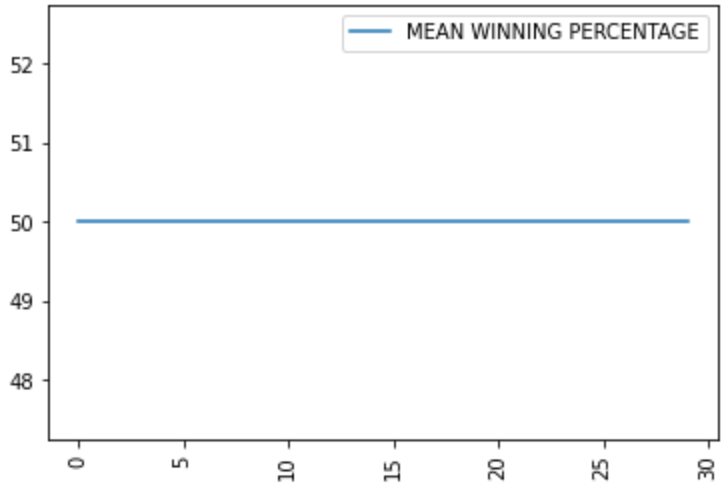

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

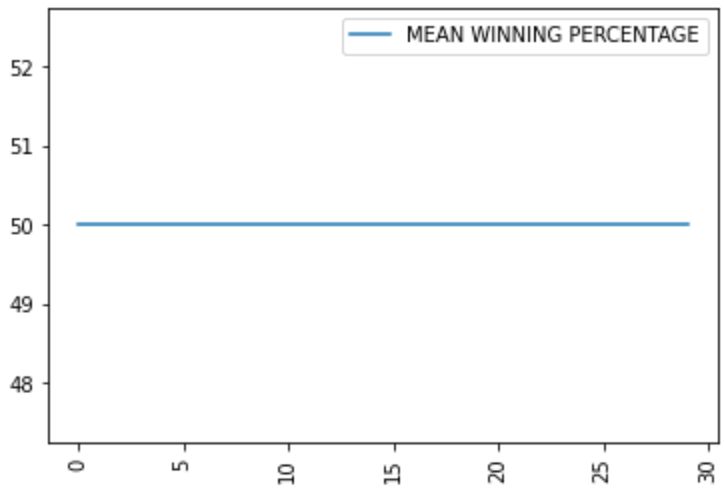

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

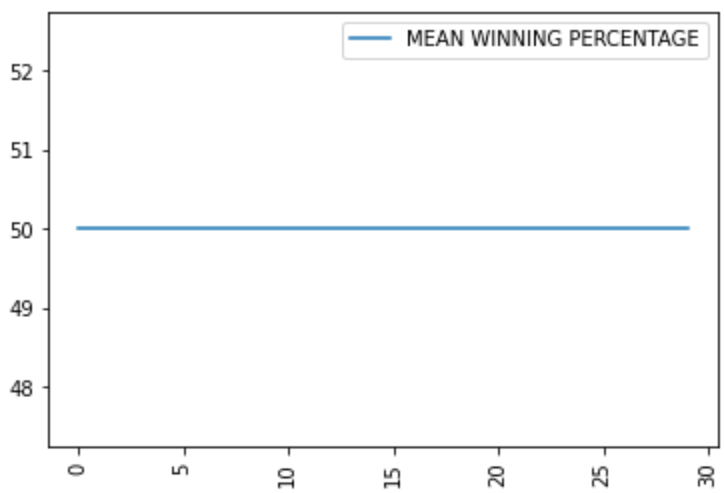

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

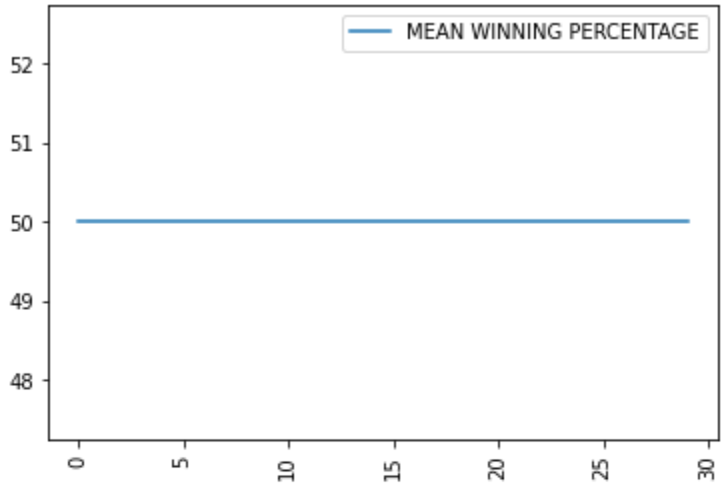

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
2
4
6
8
10

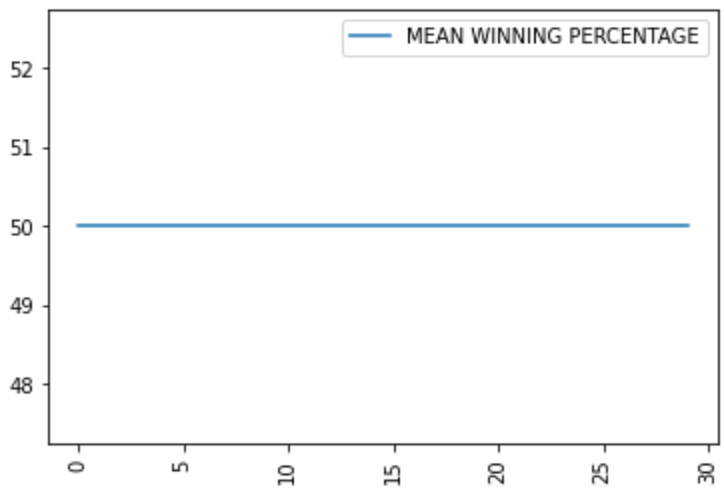

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

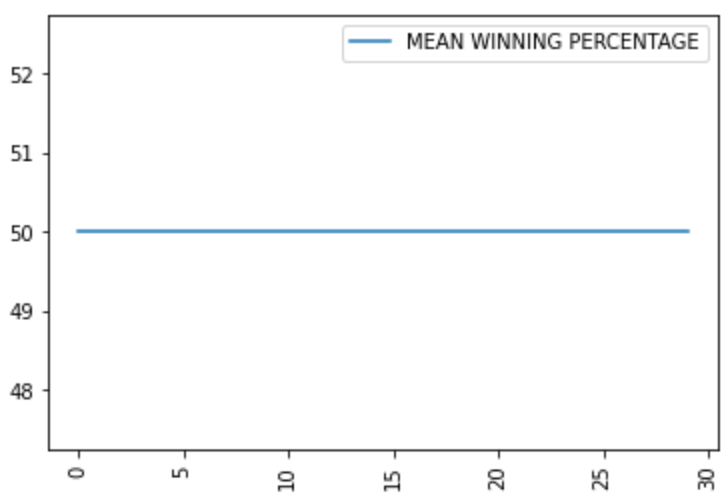

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

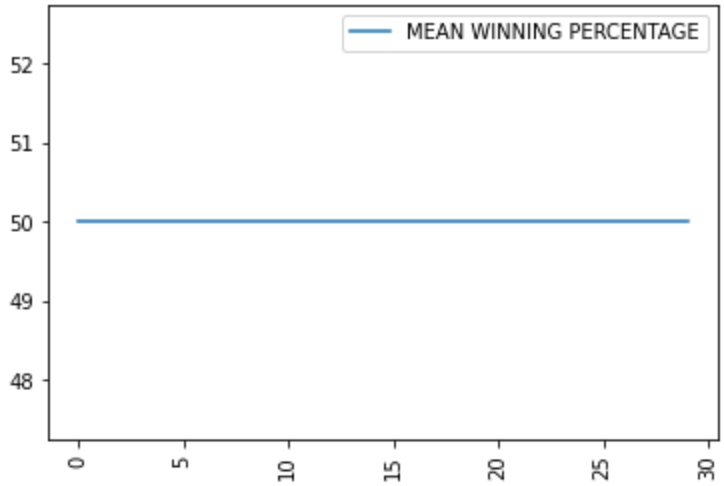

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

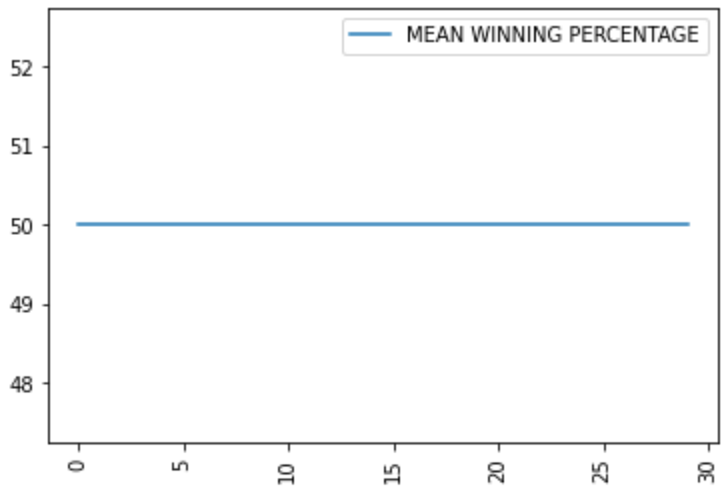

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

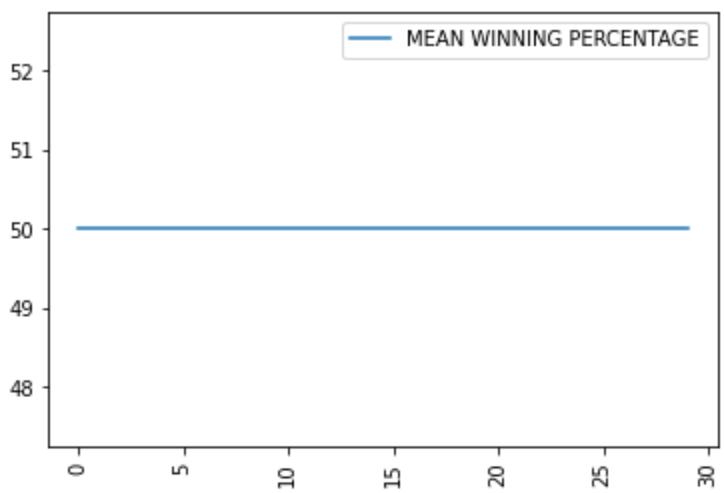

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

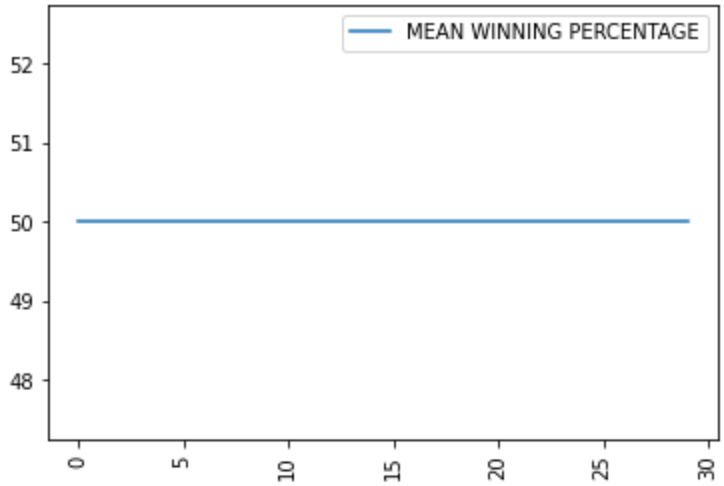

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

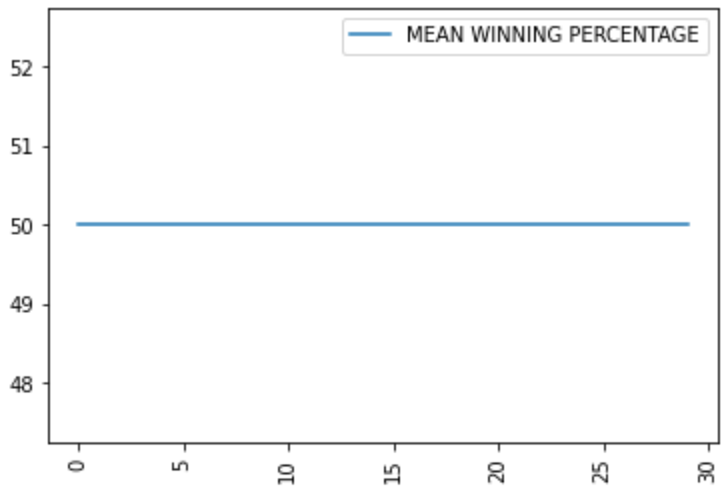

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

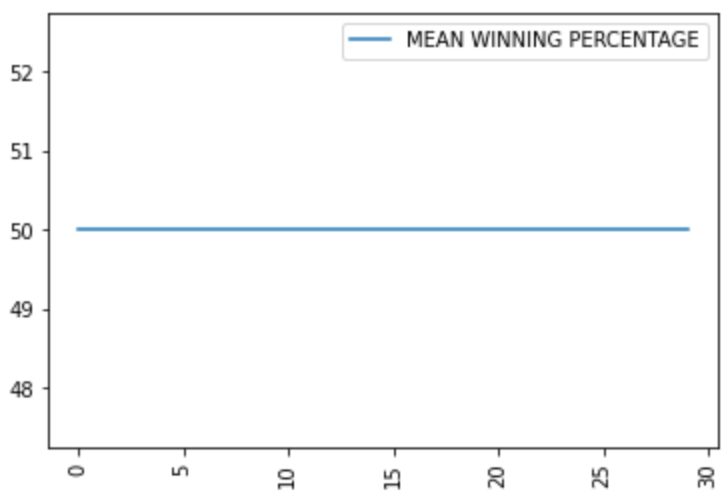

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

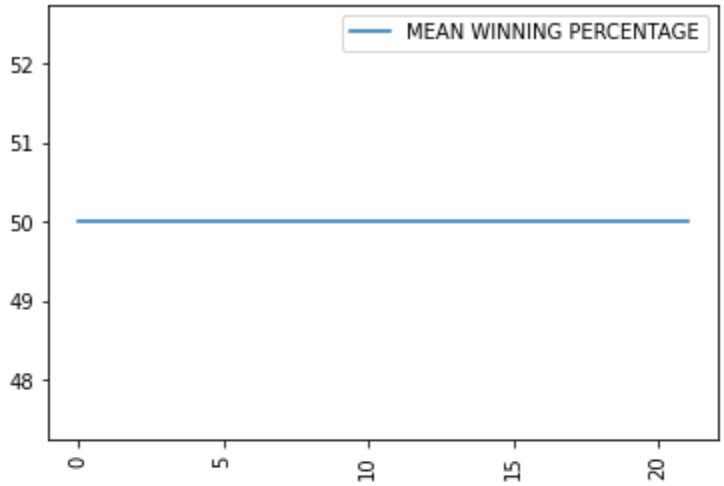

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

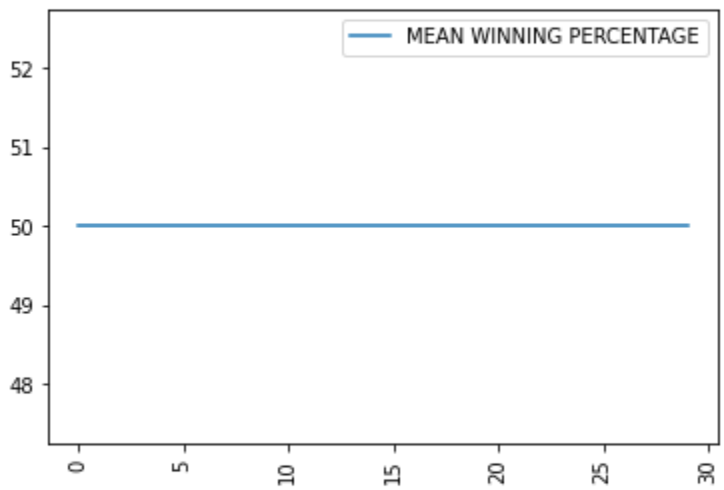

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

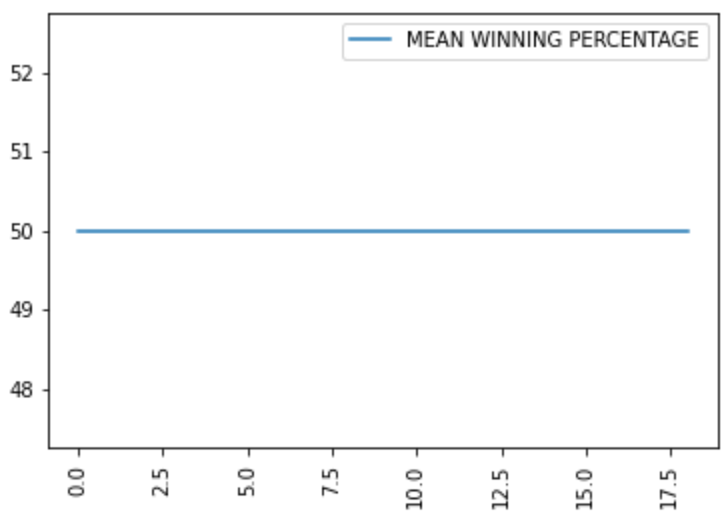

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

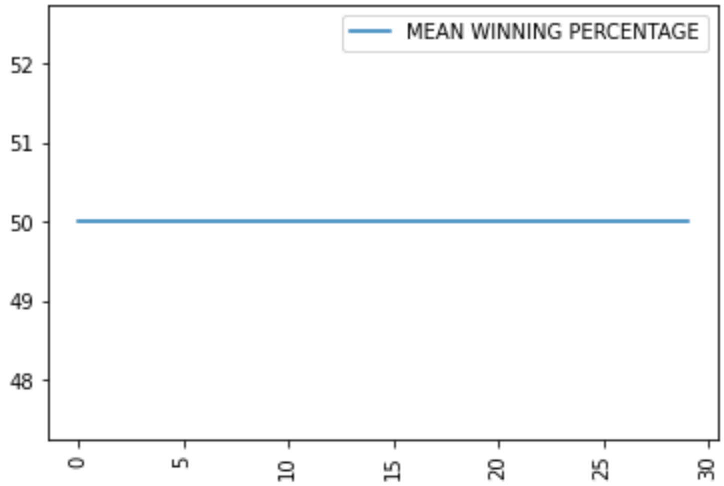

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

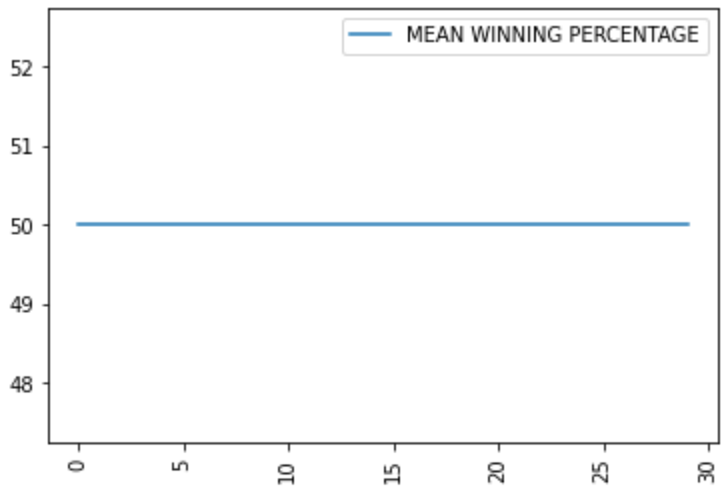

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

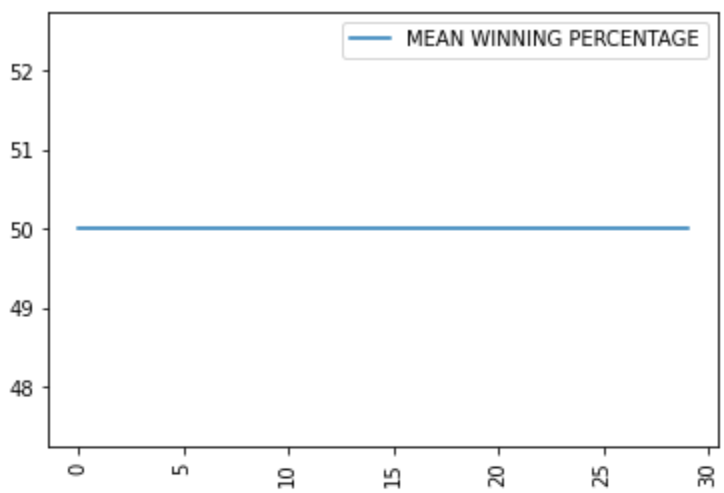

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20

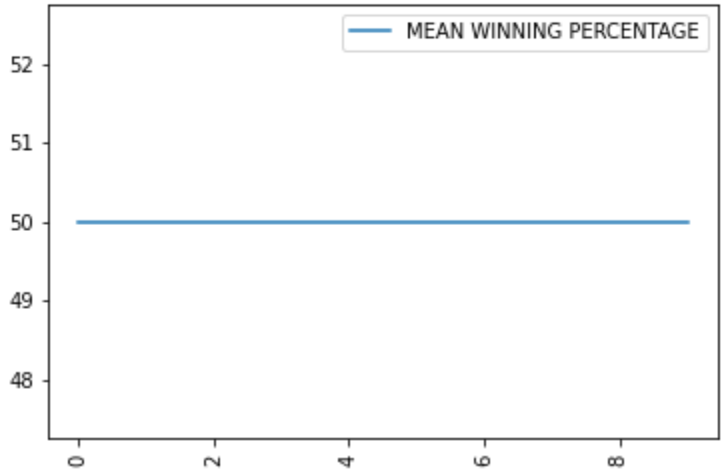

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

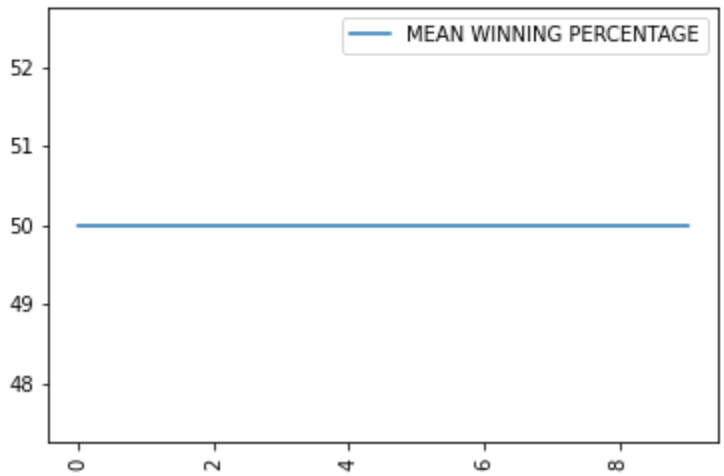

MEAN WINNING PERCENTAGE
52
51
50
49
48
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5

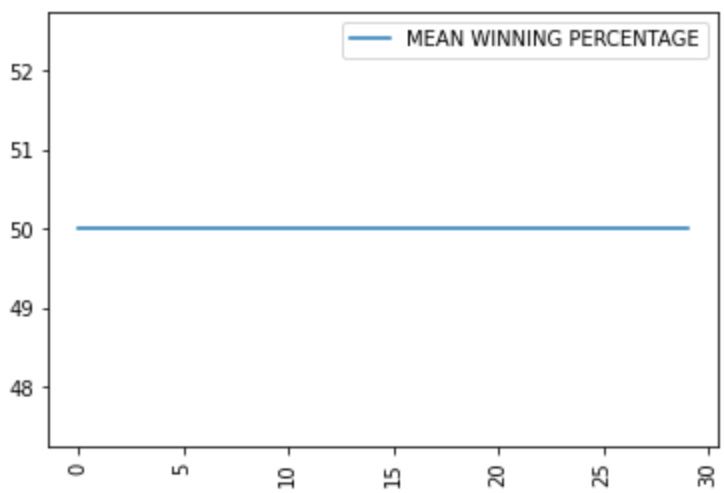

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

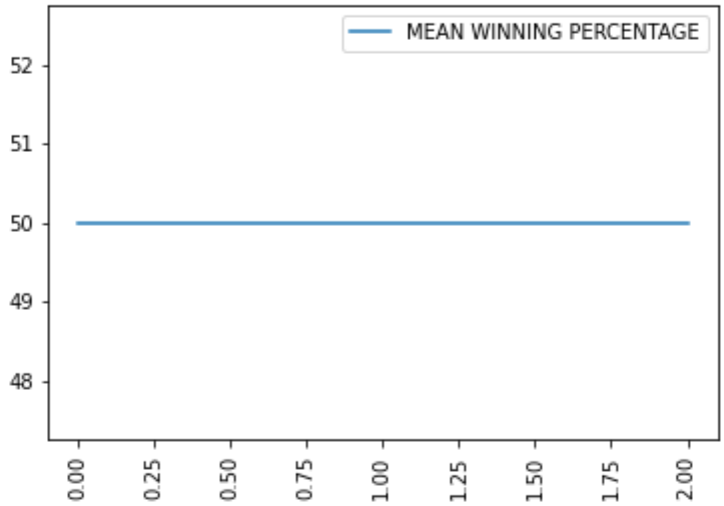

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

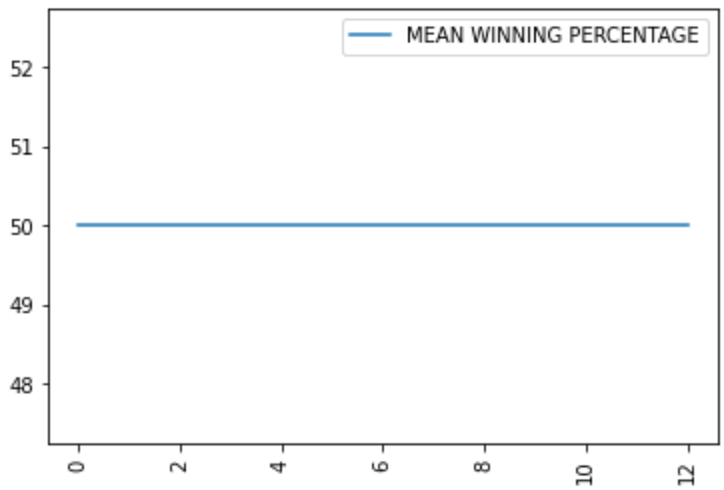

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

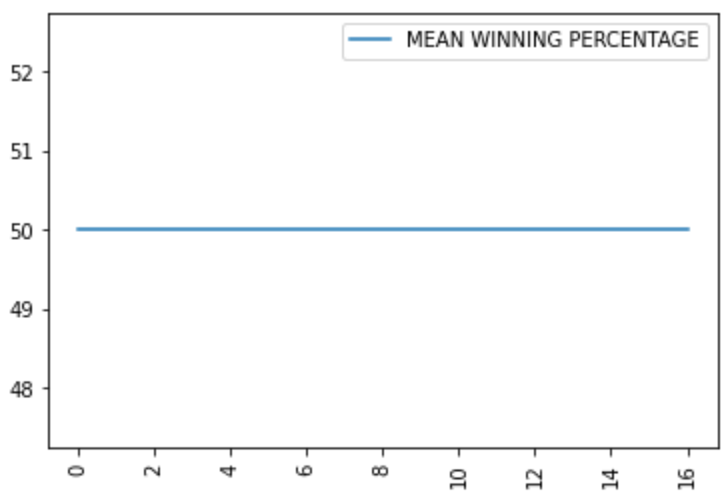

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
2
4
6
8

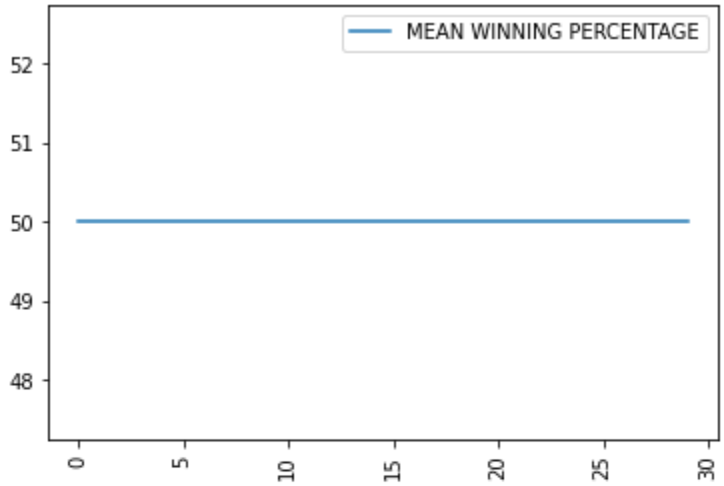

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
2
4
6
8

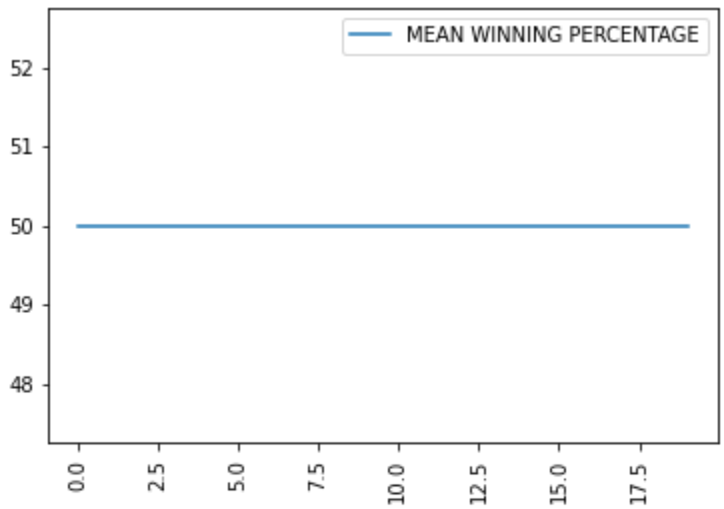

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

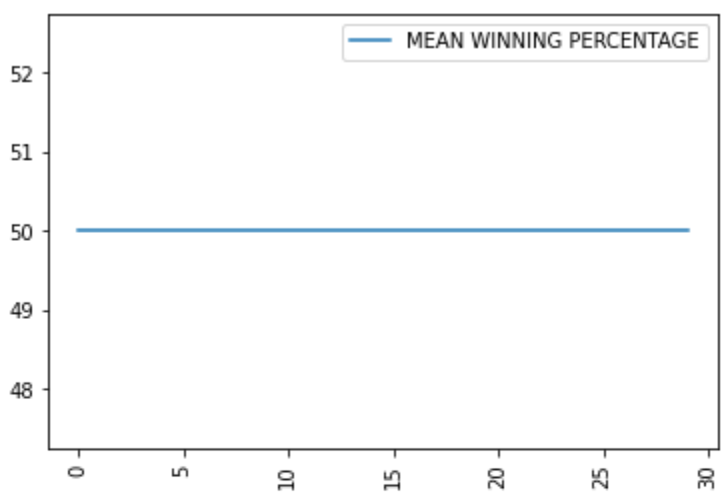

MEAN WINNING PERCENTAGE
52
51
50
49
48
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00

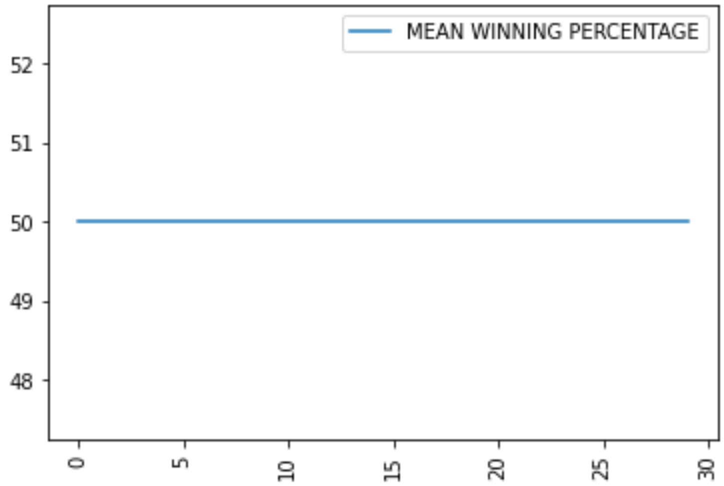

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
2
4
6
8
10
12

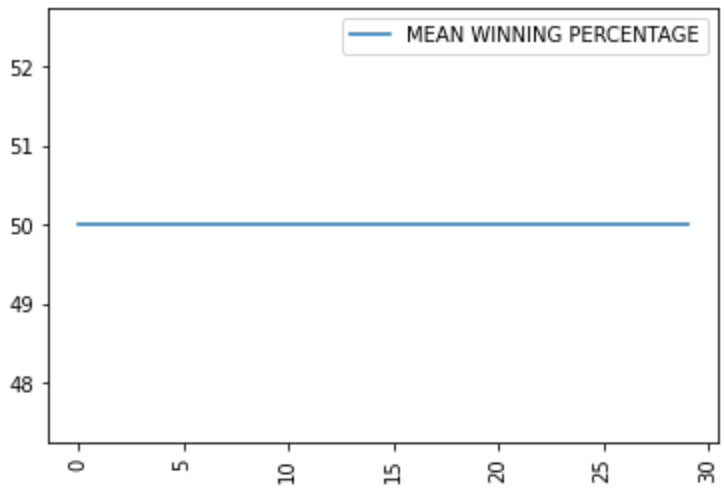

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
2
4
6
8
10
12
14
16

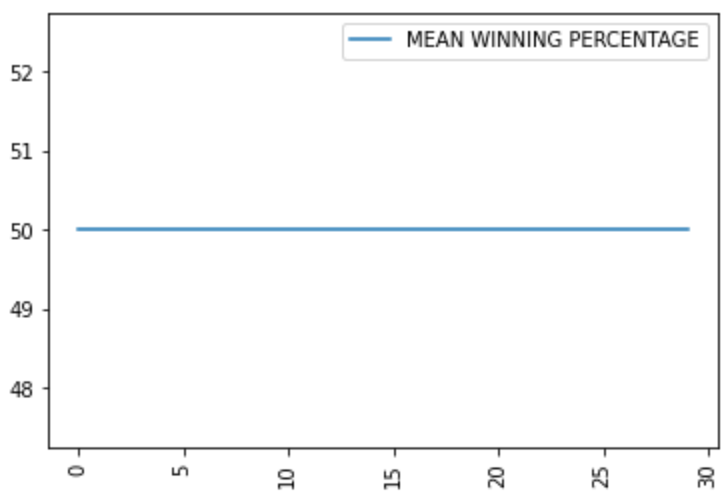

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

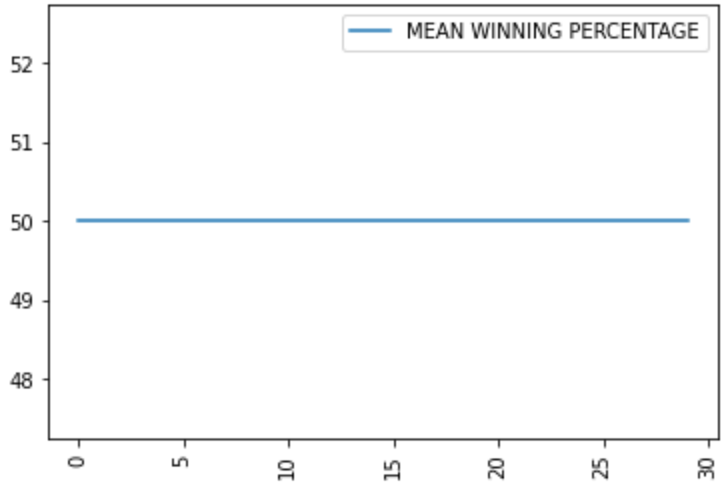

MEAN WINNING PERCENTAGE
52
51
50
49
48
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5

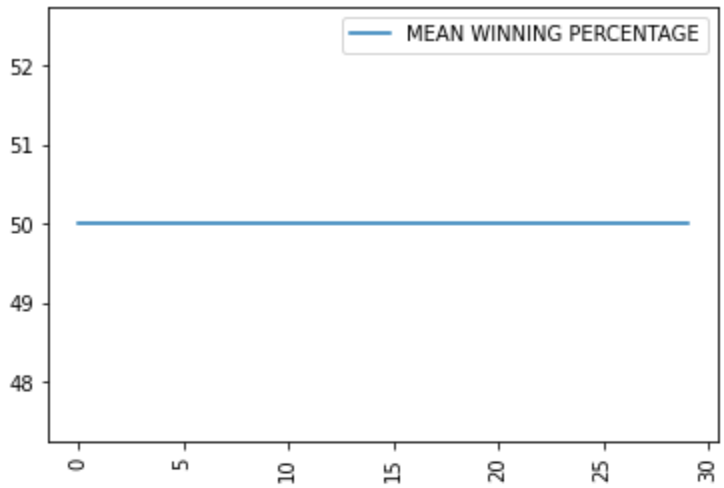

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

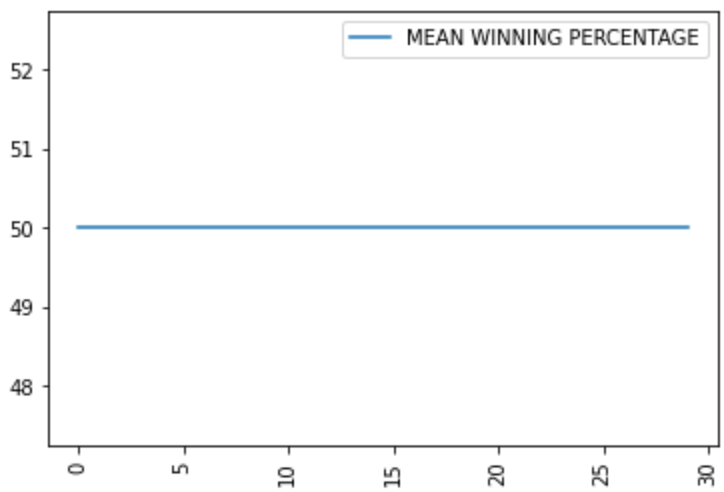

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

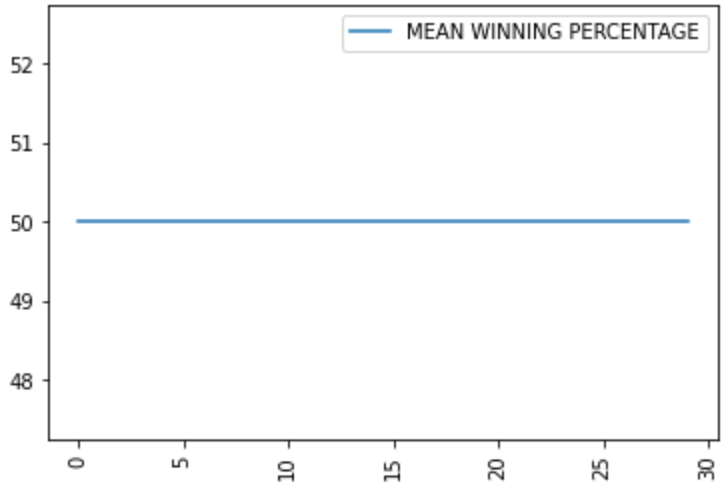

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

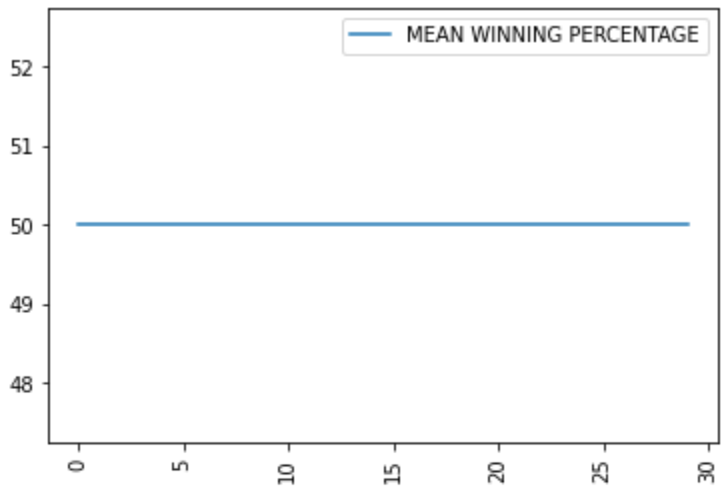

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

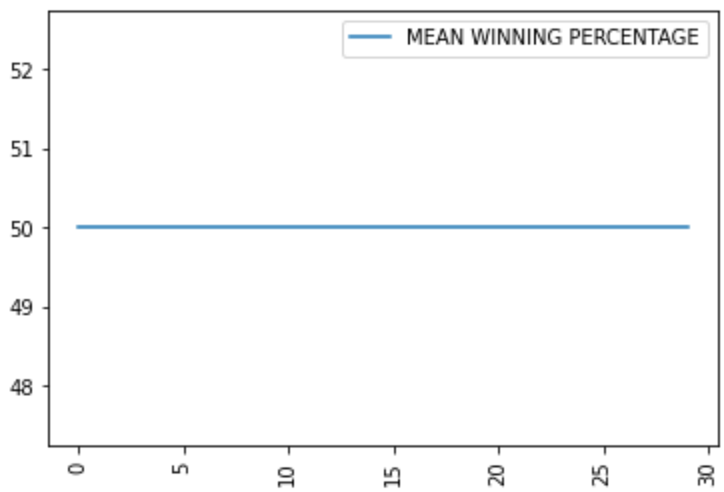

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

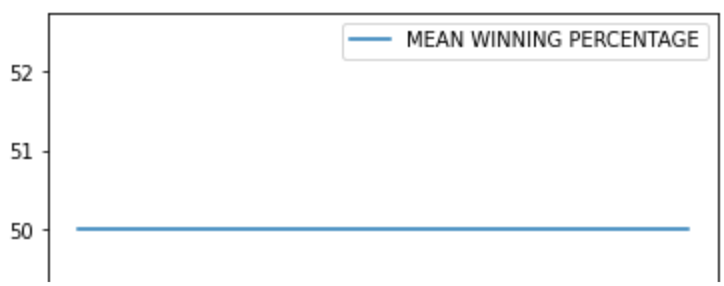

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30


MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30


MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30

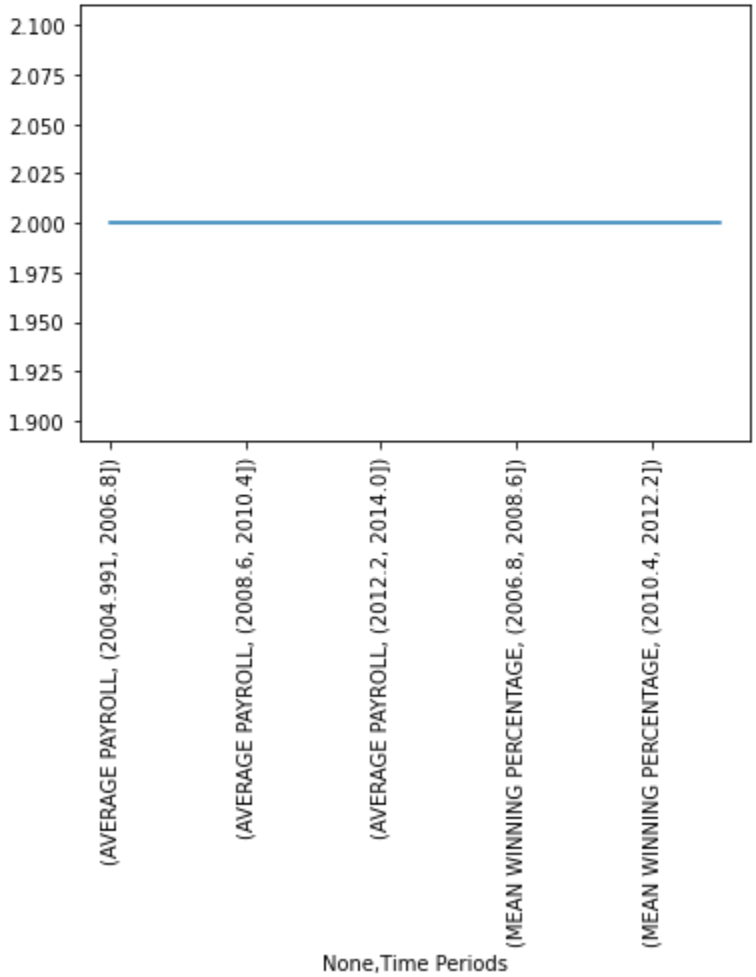

MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30


MEAN WINNING PERCENTAGE
52
51
50
49
48
0
5
10
15
20
25
30


MEAN WINNING PERCENTAGE
52
51
50

In [422…
```
a.count().unstack().plot(x='Average Payroll', y='Time Periods', rot=90)
```

Out[422… <AxesSubplot:xlabel='None,Time Periods'>



Extra Credit

Below we wll be creating a new sqlite database that has the newest available data (up to 2020)

Some imports we need

In [429…
```
from os import path
import requests
```

We will try to re create the database with the most up to date files. We will first download the zip file using python. Then we will extract the zip file into a newly created folder.

The url containing the link to download the csv file can be seen beow

In [430…
```
url_csv_zip = "https://github.com/chadwickbureau/baseballdatabank/archive/mas
```

Now let us download the zip file

In [435…
```
# download the file contents in binary format
r = requests.get(url_csv_zip)
zipfile_name = "uptodate_csv_files.zip"
# open method to open a file on your system and write the contents
with open(zipfile_name, "wb") as script:
    script.write(r.content)
```

We can see that a file called uptodate_csv_files.zip has just been created in the project folder(the same folder as this script rnning.)

In [436…
```
zipfile_exists = False
if path.exists(zipfile_name):
    zipfile_exists = True
    print("Zip file was downloaded successfully")
```

```
Zip file was downloaded successfully
```

Now we create a new directory to extract the zip file in it.

In [437…
```
new_directory = "uptodate_csv_files"
read_mode = 0o777 # mode for read write execute
path_to_extract = os.path.join(os.getcwd(), new_directory)
```

In [438…
```
os.mkdir(path_to_extract, read_mode)
```

In [439…
```
import zipfile
with zipfile.ZipFile(zipfile_name,"r") as zip_variable:
    zip_variable.extractall(path_to_extract)
```

The above code extracts the downloaded zip file. There are two to three directories that are created during the extract. We will skip these directories and go to the files that are listed in the core directory.

In [440…
```
root_dir = os.getcwd()
csv_locations = new_directory + "/baseballdatabank-master/core"
new_table_names_csv = os.listdir(csv_locations)
```

The list new_table_names_csv contains all the csv files in the core directory. Each csv file corresponds to one table in the new database. One of the files in the list is a txt file called readme2014.txt. We have to remove that from the list.

In [441…
```
# sort the list of table names
new_table_names_csv.sort()
```

In [442…
```
len(new_table_names_csv)
```

Out[442… 28

In [443…
```
new_table_names_csv
```

Out[443…
```
['AllstarFull.csv',
 'Appearances.csv',
 'AwardsManagers.csv',
 'AwardsPlayers.csv',
 'AwardsShareManagers.csv',
 'AwardsSharePlayers.csv',
 'Batting.csv',
 'BattingPost.csv',
 'CollegePlaying.csv',
 'Fielding.csv',
 'FieldingOF.csv',
 'FieldingOFsplit.csv',
 'FieldingPost.csv',
 'HallOfFame.csv',
 'HomeGames.csv',
 'Managers.csv',
 'ManagersHalf.csv',
 'Parks.csv',
 'People.csv',
 'Pitching.csv',
 'PitchingPost.csv',
 'Salaries.csv',
 'Schools.csv',
 'SeriesPost.csv',
 'Teams.csv',
 'TeamsFranchises.csv',
 'TeamsHalf.csv',
 'readme2014.txt']
```

Removing readme2014.txt from the list

In [444…
```
file_to_remove = 'readme2014.txt'
if file_to_remove in new_table_names_csv:
    new_table_names_csv.remove(file_to_remove)
new_table_names_csv
```

Out[444…
```
['AllstarFull.csv',
 'Appearances.csv',
 'AwardsManagers.csv',
 'AwardsPlayers.csv',
 'AwardsShareManagers.csv',
 'AwardsSharePlayers.csv',
 'Batting.csv',
 'BattingPost.csv',
 'CollegePlaying.csv',
 'Fielding.csv',
 'FieldingOF.csv',
 'FieldingOFsplit.csv',
 'FieldingPost.csv',
 'HallOfFame.csv',
```

```
 'HomeGames.csv',
 'Managers.csv',
 'ManagersHalf.csv',
 'Parks.csv',
 'People.csv',
 'Pitching.csv',
 'PitchingPost.csv',
 'Salaries.csv',
 'Schools.csv',
 'SeriesPost.csv',
 'Teams.csv',
 'TeamsFranchises.csv',
 'TeamsHalf.csv']
```

Now let's get the list of tables in the old database we started this script with, to make sure that we have the same table names, and table structure. If we have the same names of tables and table structure for each table, then we can use the create table statement of the old database to create a new database.

In [445…

```
show_tables_query = "SELECT name FROM sqlite_master WHERE type='table';"
old_db_tables_results = pandas.read_sql(show_tables_query, conn)
old_db_tables_results.sort_values(by=['name'])
```

Out[445…

	name
0	AllstarFull
1	Appearances
2	AwardsManagers
3	AwardsPlayers
4	AwardsShareManagers
5	AwardsSharePlayers
6	Batting
7	BattingPost
8	CollegePlaying
9	Fielding
10	FieldingOF
11	FieldingPost
12	HallOfFame
13	Managers
14	ManagersHalf
15	Master
16	Pitching
17	PitchingPost
18	Salaries
19	Schools

	name
20	SeriesPost
21	Teams
22	TeamsFranchises

The above shows 24 tables. So the old database has 24 tables

In [446…

```
len(new_table_names_csv)
```

Out[446… 27

Remove the csv extensions on the list of tables to have exactly the table names

The new databse seems to have 3 more tables than the old one. Now let's extact the table name, by removing the extension for each of the csv file names.

In [447…

```
new_db_table_names = []
for f in new_table_names_csv:
    t = os.path.splitext(f)
    new_db_table_names.append(t[0])
```

So we need to compare the list of tables in the new database to create with the list of tables in the old database. For that we will use the difference function of a set. We start by transforming each list to a set and comparing both

In [448…

```
old_db_table_names = old_db_tables_results['name'].tolist()
```

Now let's have a look at the content of the new database table not found in the old databse table.

In [449…

```
set(new_db_table_names).difference(set(old_db_table_names))
```

Out[449… {'FieldingOFsplit', 'HomeGames', 'Parks', 'People'}

Now let's have a look at the content of the old database table not found in the new databse table.

In [450…

```
set(old_db_table_names).difference(set(new_db_table_names))
```

Out[450… {'Master'}

Remove the Master table because it is not part of the new database

In [451…

```
old_db_table_names.remove('Master')
```

We can see that the new databse is different than the old database. The old database has a Master table that the new databse doe not have. The new database has 4 tables that are not found in the old database. The easiest way to do it is to manually create tables for each csv file.

It will be tedious but we will get there. We will create the new database, and we will create each table in the new downloaded extracted unzipped csv files.

In [452…

```
new_2020_sqlite_file = 'lahman2020.sqlite'
new_db_conn = sqlite3.connect(new_2020_sqlite_file)
```

See the content of the old table names. It does not contain the Master table anymore.

Now for each of the old tables found in the new database, generate the schema and use that schema to create the table in the new database.

In [453…

```
q = """
        SELECT sql
        FROM sqlite_master
            WHERE sql NOT NULL AND
            type == 'table'
    """
create_stmts = pandas.read_sql_query(q, conn)
```

In [454…

```
lst_queries = create_stmts['sql'].tolist()
```

In [455…

```
l = []
for q in lst_queries:
    l.append(q.replace("\n",""))
```

In [456…

```
cursor = new_db_conn.cursor()

for t in l:
    cursor.execute(t)
```

Check to see if the tables where created. If they were then basically the database will have tables sql schema in the sqlite master table.

In [457…

```
check_create = """
        SELECT sql
        FROM sqlite_master
            WHERE sql NOT NULL AND
            type == 'table'
    """
new_create_stmts = pandas.read_sql_query(check_create, new_db_conn)
```

In [458…

```
new_create_stmts['sql']
```

Out[458…

```
0     CREATE TABLE AllstarFull (playerID TEXT,yearID...
1     CREATE TABLE Appearances (yearID INTEGER,teamI...
2     CREATE TABLE AwardsManagers (playerID TEXT,awa...
3     CREATE TABLE AwardsPlayers (playerID TEXT,awar...
4     CREATE TABLE AwardsShareManagers (awardID TEXT...
5     CREATE TABLE AwardsSharePlayers (awardID TEXT,...
6     CREATE TABLE Batting (playerID TEXT,yearID INT...
```

```
7     CREATE TABLE BattingPost (yearID INTEGER,round...
8     CREATE TABLE CollegePlaying (playerID TEXT,sch...
9     CREATE TABLE Fielding (playerID TEXT,yearID IN...
10    CREATE TABLE FieldingOF (playerID TEXT,yearID ...
11    CREATE TABLE FieldingPost (playerID TEXT,yearI...
12    CREATE TABLE HallOfFame (playerID TEXT,yearid ...
13    CREATE TABLE Managers (playerID TEXT,yearID IN...
14    CREATE TABLE ManagersHalf (playerID TEXT,yearI...
15    CREATE TABLE Master (playerID TEXT,birthYear I...
16    CREATE TABLE Pitching (playerID TEXT,yearID IN...
17    CREATE TABLE PitchingPost (playerID TEXT,yearI...
18    CREATE TABLE Salaries (yearID INTEGER,teamID T...
19    CREATE TABLE Schools (schoolID TEXT,name_full ...
20    CREATE TABLE SeriesPost (yearID INTEGER,round ...
21    CREATE TABLE Teams (yearID INTEGER,lgID TEXT,t...
22    CREATE TABLE TeamsFranchises (franchID TEXT,fr...
23    CREATE TABLE TeamsHalf (yearID INTEGER,lgID TE...
```

Now for the four new tables, let's add them to the new database. Here they are below :

In [459…

```
set_difference = set(new_db_table_names).difference(set(old_db_table_names))
```

In [460…

```
lst_difference = list(set_difference)
lst_difference
```

Out[460…

```
['People', 'FieldingOFsplit', 'Parks', 'HomeGames']
```

We will use the function found at

https://www.geeksforgeeks.org/get-column-names-from-csv-using-python/

In [461…

```
def list_columns(file_path_name):
    import csv

    # opening the csv file by specifying
    # the location
    # with the variable name as csv_file
    with open(file_path_name) as csv_file:

        # creating an object of csv reader
        # with the delimiter as ,
        csv_reader = csv.reader(csv_file, delimiter = ',')

        # list to store the names of columns
        list_of_column_names = []

        # loop to iterate thorugh the rows of csv
        for row in csv_reader:

            # adding the first row
            list_of_column_names.append(row)

            # breaking the loop after the
            # first iteration itself
            break

        return list_of_column_names[0]
```

Now let's just create the 4 missing tables in the new databse. For that we need to open the file locations and list the columns for each file

In [462…

```
# create a ew dictionary
d = dict()
extension = ".csv"
for file_name in lst_difference :
    file_path_name = csv_locations + "/" + file_name + extension
    list_of_column_names = list_columns(file_path_name)
    d[file_name + extension] = list_of_column_names
```

In [463…

```
d.keys()
```

Out[463…

```
dict_keys(['People.csv', 'FieldingOFsplit.csv', 'Parks.csv', 'HomeGames.csv
'])
```

From the above we can generate the create statements query

In [464…

```
d['People.csv']
```

Out[464…

```
['playerID',
 'birthYear',
 'birthMonth',
 'birthDay',
 'birthCountry',
 'birthState',
 'birthCity',
```

```
 'deathYear',
 'deathMonth',
 'deathDay',
 'deathCountry',
 'deathState',
 'deathCity',
 'nameFirst',
 'nameLast',
 'nameGiven',
 'weight',
 'height',
 'bats',
 'throws',
 'debut',
 'finalGame',
 'retroID',
 'bbrefID']
```

Well just found out that the People table is basically the Master table in the old databse. Simply rename Master to People.

In [465…

```
rename_master_query = "ALTER TABLE Master RENAME TO People;"
nb_executes = pandas.read_sql_query(rename_master_query, new_db_conn)
```

In [466…

```
nb_executes
```

Out[466…

1

Operation was a success

Create table for FieldingOfSplit

dict_keys(['People.csv', 'FieldingOFsplit.csv', 'Parks.csv', 'HomeGames.csv'])

In [467…

```
d['FieldingOFsplit.csv']
```

Out[467…

```
['playerID',
 'yearID',
 'stint',
 'teamID',
 'lgID',
 'POS',
 'G',
 'GS',
 'InnOuts',
 'PO',
 'A',
 'E',
 'DP',
 'PB',
 'WP',
 'SB',
 'CS',
 'ZR']
```

Based on the abpve, we can generate the following query.

In [468…

```
drop_first = "DROP TABLE IF EXISTS FieldingOFsplit ;";
FieldingOFsplit_query = """
    CREATE TABLE IF NOT EXISTS FieldingOFsplit(
        playerID TEXT PRIMARY KEY,
        yearID INTEGER,
        stint INTEGER ,
        teamID TEXT ,
        lgID TEXT ,
        POS TEXT,
        G INTEGER,
        GS INTEGER,
        InnOuts INTEGER,
        PO INTEGER,
        A INTEGER,
        E INTEGER,
        DP INTEGER,
        PB TEXT,
        WP TEXT,
        SB TEXT,
        CS TEXT,
        ZR TEXT

    );

    """
```

Create the above query

In [469…

```
success = cursor.execute(drop_first)
success
```

Out[469… <sqlite3.Cursor at 0x7f72ff38e3b0>

In [470…

```
success = cursor.execute(FieldingOFsplit_query)
success
```

Out[470… <sqlite3.Cursor at 0x7f72ff38e3b0>

Check the creation by looking into the schema

In [471…

```
check_create = """
        SELECT sql
        FROM sqlite_master
            WHERE sql NOT NULL AND
            type == 'table'
    """
new_create_stmts = pandas.read_sql_query(check_create, new_db_conn)
new_create_stmts['sql']
```

Out[471…

```
0     CREATE TABLE AllstarFull (playerID TEXT,yearID...
1     CREATE TABLE Appearances (yearID INTEGER,teamI...
2     CREATE TABLE AwardsManagers (playerID TEXT,awa...
3     CREATE TABLE AwardsPlayers (playerID TEXT,awar...
```

```
4     CREATE TABLE AwardsShareManagers (awardID TEXT...
5     CREATE TABLE AwardsSharePlayers (awardID TEXT,...
6     CREATE TABLE Batting (playerID TEXT,yearID INT...
7     CREATE TABLE BattingPost (yearID INTEGER,round...
8     CREATE TABLE CollegePlaying (playerID TEXT,sch...
9     CREATE TABLE Fielding (playerID TEXT,yearID IN...
10    CREATE TABLE FieldingOF (playerID TEXT,yearID ...
11    CREATE TABLE FieldingPost (playerID TEXT,yearI...
12    CREATE TABLE HallOfFame (playerID TEXT,yearid ...
13    CREATE TABLE Managers (playerID TEXT,yearID IN...
14    CREATE TABLE ManagersHalf (playerID TEXT,yearI...
15    CREATE TABLE "People" (playerID TEXT,birthYear...
16    CREATE TABLE Pitching (playerID TEXT,yearID IN...
17    CREATE TABLE PitchingPost (playerID TEXT,yearI...
18    CREATE TABLE Salaries (yearID INTEGER,teamID T...
19    CREATE TABLE Schools (schoolID TEXT,name_full ...
20    CREATE TABLE SeriesPost (yearID INTEGER,round ...
21    CREATE TABLE Teams (yearID INTEGER,lgID TEXT,t...
22    CREATE TABLE TeamsFranchises (franchID TEXT,fr...
23    CREATE TABLE TeamsHalf (yearID INTEGER,lgID TE...
24    CREATE TABLE FieldingOFsplit(\n        playerI...
```

In [472…

```python
new_create_stmts['sql']
```

Out[472…

```
0     CREATE TABLE AllstarFull (playerID TEXT,yearID...
1     CREATE TABLE Appearances (yearID INTEGER,teamI...
2     CREATE TABLE AwardsManagers (playerID TEXT,awa...
3     CREATE TABLE AwardsPlayers (playerID TEXT,awar...
4     CREATE TABLE AwardsShareManagers (awardID TEXT...
5     CREATE TABLE AwardsSharePlayers (awardID TEXT,...
6     CREATE TABLE Batting (playerID TEXT,yearID INT...
7     CREATE TABLE BattingPost (yearID INTEGER,round...
8     CREATE TABLE CollegePlaying (playerID TEXT,sch...
9     CREATE TABLE Fielding (playerID TEXT,yearID IN...
10    CREATE TABLE FieldingOF (playerID TEXT,yearID ...
11    CREATE TABLE FieldingPost (playerID TEXT,yearI...
12    CREATE TABLE HallOfFame (playerID TEXT,yearid ...
13    CREATE TABLE Managers (playerID TEXT,yearID IN...
14    CREATE TABLE ManagersHalf (playerID TEXT,yearI...
15    CREATE TABLE "People" (playerID TEXT,birthYear...
16    CREATE TABLE Pitching (playerID TEXT,yearID IN...
17    CREATE TABLE PitchingPost (playerID TEXT,yearI...
18    CREATE TABLE Salaries (yearID INTEGER,teamID T...
19    CREATE TABLE Schools (schoolID TEXT,name_full ...
20    CREATE TABLE SeriesPost (yearID INTEGER,round ...
21    CREATE TABLE Teams (yearID INTEGER,lgID TEXT,t...
22    CREATE TABLE TeamsFranchises (franchID TEXT,fr...
23    CREATE TABLE TeamsHalf (yearID INTEGER,lgID TE...
24    CREATE TABLE FieldingOFsplit(\n        playerI...
Name: sql, dtype: object
```

Creating Parks table

In [473…

```python
d['Parks.csv']
```

Out[473…

```
['park.key', 'park.name', 'park.alias', 'city', 'state', 'country']
```

In [474…

```
drop_first = "DROP TABLE IF EXISTS Parks ;";
Parks_query = """
    CREATE TABLE IF NOT EXISTS Parks(
        id TEXT PRIMARY KEY,
        name TEXT,
        alias TEXT,
        city TEXT,
        state TEXT,
        country TEXT
    );

    """
```

In [475…

```
success = cursor.execute(drop_first)
success
```

Out[475… <sqlite3.Cursor at 0x7f72ff38e3b0>

In [476…

```
success = cursor.execute(Parks_query)
success
```

Out[476… <sqlite3.Cursor at 0x7f72ff38e3b0>

Now the whole new database has been created. Now we can insert the values into the tables

Open each csv file and insert their values in the table

In [477…

```
import csv
```

In [478…

```
# Batting generates an error. so remove it and deal with it seperately
new_table_names_csv.remove("Batting.csv")
for file_name in new_table_names_csv:
    file_path_name = csv_locations + "/" + file_name
    table_name = os.path.splitext(file_name)[0]
    with open(file_path_name, newline='') as f:
        reader = csv.reader(f)
        data = list(reader)
        k = len(data)
        if k > 1:
            for i in range(1,len(data)):
                words = data[i]
                values =','.join(f'\'{w}\'' for w in words)
                if values:
                    delete_query = " DELETE FROM " + table_name + " WHERE TRU
                    #','.join(f'"{w}"' for w in words)
                    query = "INSERT INTO "+ table_name  + \
                    " VALUES (" + values + " ) ; "
                    c = cursor.execute(delete_query)
                    c = cursor.execute(query)
```

```
---------------------------------------------------------------------------
OperationalError                          Traceback (most recent call last)
```

```
<ipython-input-478-5f9bf4a19ab4> in <module>
     17                         query = "INSERT INTO "+ table_name  + \
     18                         " VALUES (" + values + " ) ; "
---> 19                         c = cursor.execute(delete_query)
     20                         c = cursor.execute(query)

OperationalError: no such table: HomeGames
```

In [479…

```
check_create = """
        SELECT sql
        FROM sqlite_master
            WHERE sql NOT NULL AND
            type == 'table'
    """
new_create_stmts = pandas.read_sql_query(check_create, new_db_conn)
new_create_stmts
```

Out[479…

	sql
0	CREATE TABLE AllstarFull (playerID TEXT,yearID...
1	CREATE TABLE Appearances (yearID INTEGER,teamI...
2	CREATE TABLE AwardsManagers (playerID TEXT,awa...
3	CREATE TABLE AwardsPlayers (playerID TEXT,awar...
4	CREATE TABLE AwardsShareManagers (awardID TEXT...
5	CREATE TABLE AwardsSharePlayers (awardID TEXT,...
6	CREATE TABLE Batting (playerID TEXT,yearID INT...
7	CREATE TABLE BattingPost (yearID INTEGER,round...
8	CREATE TABLE CollegePlaying (playerID TEXT,sch...
9	CREATE TABLE Fielding (playerID TEXT,yearID IN...
10	CREATE TABLE FieldingOF (playerID TEXT,yearID ...
11	CREATE TABLE FieldingPost (playerID TEXT,yearI...
12	CREATE TABLE HallOfFame (playerID TEXT,yearid ...
13	CREATE TABLE Managers (playerID TEXT,yearID IN...
14	CREATE TABLE ManagersHalf (playerID TEXT,yearI...
15	CREATE TABLE "People" (playerID TEXT,birthYear...
16	CREATE TABLE Pitching (playerID TEXT,yearID IN...
17	CREATE TABLE PitchingPost (playerID TEXT,yearI...
18	CREATE TABLE Salaries (yearID INTEGER,teamID T...
19	CREATE TABLE Schools (schoolID TEXT,name_full ...
20	CREATE TABLE SeriesPost (yearID INTEGER,round ...
21	CREATE TABLE Teams (yearID INTEGER,lgID TEXT,t...
22	CREATE TABLE TeamsFranchises (franchID TEXT,fr...
23	CREATE TABLE TeamsHalf (yearID INTEGER,lgID TE...

	sql
24	CREATE TABLE FieldingOFsplit(\n playerI...

In [ ]: