main

September 22, 2024

[1]:
```
%load_ext autoreload
%autoreload 2
from mylib.lib import *
# print(mylib.lib.__file__)


example_csv = "https://projects.fivethirtyeight.com/nba-model/2023/
 ↪latest_RAPTOR_by_player.csv"
```

[2]:
```
general_df = load_and_preprocess(example_csv)
# print(general_df)
# print(general_df.shape)
assert general_df is not None
assert general_df.shape == (541, 21)
#  print(f"The table we got is glimpsed like: {general_df.head()}")
print(f"MP stands for Minutes Played, which represents the total number of␣
 ↪minutes a player was on the court during the game or season. ")
print(f"Poss stands for Possessions, which is an estimate of the number of␣
 ↪possessions a player was directly involved in during a game or season.")

general_df.head()
```

```
MP stands for Minutes Played, which represents the total number of minutes a
player was on the court during the game or season.
Poss stands for Possessions, which is an estimate of the number of possessions a
player was directly involved in during a game or season.
```

[2]:

	player_name	player_id	season	poss	mp	raptor_box_offense	raptor_box_defense	raptor_box_total	raptor_onoff_offense
0	Precious Achiuwa	achiupr01	2023	2328	1140	-2.442459	0.605968	-1.836491	1.583854
1	Steven Adams	adamsst01	2023	2391	1133	-0.539149	4.174473	3.635324	4.196019
2	Bam Adebayo	adebaba01	2023	6933	3448	-1.905568	2.343398	0.437830	1.584151
3	Ochai Agbaji	agbajoc01	2023	2604	1209	-1.323938			
4	Santi Aldama	aldamsa01	2023	3824	1783	-1.022699			

```
3          -2.198396         -3.522334              0.637793
4          -0.407462         -1.430161             -0.169196

   raptor_onoff_defense  …  raptor_offense  raptor_defense  raptor_total  \
0             -3.756532  …       -1.771180       -0.283673     -2.054852
1              0.186159  …        0.416279        3.550985      3.967264
2              2.634869  …       -1.279140        2.528053      1.248914
3             -0.569349  …       -1.010696       -1.964927     -2.975622
4             -1.337527  …       -0.907095       -0.675487     -1.582582

   war_total  war_reg_season  war_playoffs  predator_offense  \
0   0.400488        0.400488      0.000000         -1.867986
1   3.892111        3.892111      0.000000          0.025097
2   7.035148        5.692164      1.342984         -0.809324
3  -0.139469       -0.139469      0.000000         -1.145865
4   1.058655        1.159520     -0.100866         -1.245487

   predator_defense  predator_total  pace_impact
0         -0.182806       -2.050792    -0.956065
1          3.570024        3.595120     0.237905
2          2.539526        1.730202    -0.443789
3         -2.506387       -3.652252     0.216216
4         -0.165475       -1.410962    -0.035690

[5 rows x 21 columns]
```

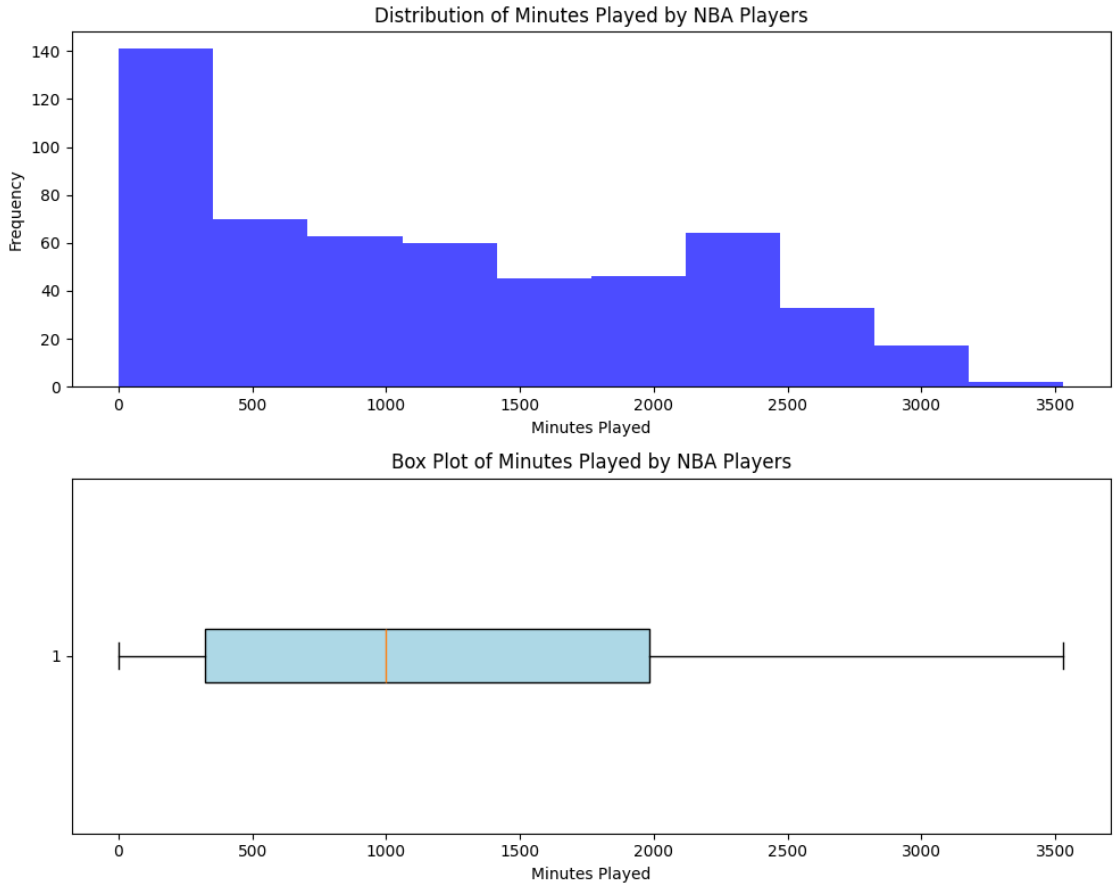
[3]:
```python
mean_test = process_mean(general_df, "mp")
quantile_test = process_quantile(general_df, "mp", 0.25)
median_test = process_median(general_df, "mp")
std_test = process_std(general_df, "mp")
describe_test = general_df.describe()

assert describe_test.loc["mean", "mp"] == mean_test
assert describe_test.loc["std", "mp"] == std_test
assert describe_test.loc["25%", "mp"] == quantile_test
# assert median_test == 996.0
```

[4]:
```python
print(f"MP stands for Minutes Played, which represents the total number of␣
 ↪minutes a player was on the court during the game or season. ")
print(f"Mean of MP is {mean_test}")
print(f"Median of MP is {median_test}")
print(f"Quantiles of MP is {quantile_test}")
NBA_bar_mp(general_df, True)
```

```
MP stands for Minutes Played, which represents the total number of minutes a
player was on the court during the game or season.
Mean of MP is 1174.4158964879853
Median of MP is 996.0
```

```
Quantiles of MP is 323.0
HERE
```

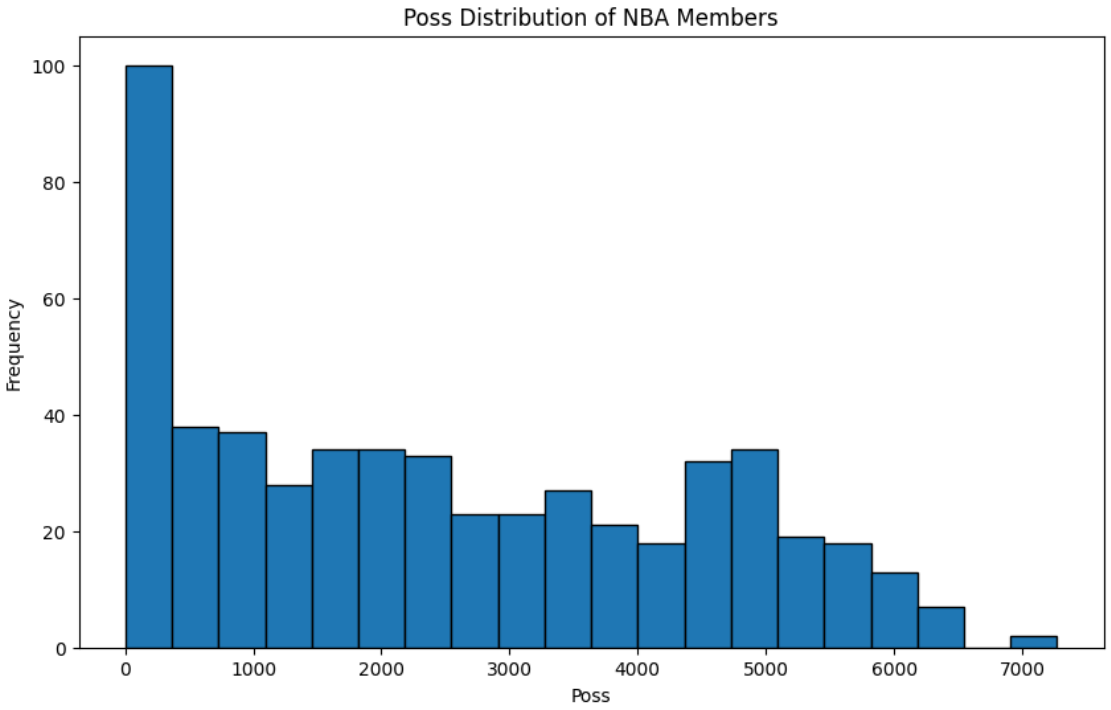


[5]:
```python
mean_test = process_mean(general_df, "poss")
quantile_test = process_quantile(general_df, "poss", 0.25)
median_test = process_median(general_df, "poss")
std_test = process_std(general_df, "poss")
describe_test = general_df.describe()

assert describe_test.loc["mean", "poss"] == mean_test
assert describe_test.loc["std", "poss"] == std_test
assert describe_test.loc["25%", "poss"] == quantile_test
```

[6]:
```python
print(f"Poss stands for Possessions, which is an estimate of the number of␣
  ↪possessions a player was directly involved in during a game or season.")
print(f"Mean of Possessions is {mean_test}")
print(f"Median of Possessions is {median_test}")
print(f"Quantiles of Possessions is {quantile_test}")
NBA_histogram_poss(general_df, True)
```

```
Poss stands for Possessions, which is an estimate of the number of possessions a
player was directly involved in during a game or season.
Mean of Possessions is 2463.776340110906
Median of Possessions is 2138.0
Quantiles of Possessions is 699.0
```



[7]:
```
from IPython.display import Markdown, display

# Read the markdown file and display it in the notebook
with open('nbval_results.md', 'r') as f:
    content = f.read()
display(Markdown(content))
```

1 nbval Test Results

```
============================= test session starts ==============================
platform darwin -- Python 3.12.3, pytest-8.3.3, pluggy-1.5.0
rootdir: /Users/momowang/Documents/Coding/DEindividual1/DEIndividual1
plugins: typeguard-4.3.0, anyio-4.4.0, nbval-0.11.0
collected 6 items

main.ipynb ......                                                        [100%]

============================== 6 passed in 2.20s ===============================
```

[ ]: