In [41]:

```
import numpy as np
import pandas as pd
pd.options.display.max_columns=100
from sklearn.feature_selection import SelectKBest, f_classif, mutual_info_cl
from sklearn.preprocessing import StandardScaler as SSc
from sklearn.preprocessing import MinMaxScaler as MMS
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline

from sklearn.decomposition import PCA

from IPython.core.display import display, HTML
display(HTML("<style>.container { width:95% !important; }</style>"))
```

In [28]:

```
data = pd.read_csv("Data-Prepped.csv",index_col=0)
data = data.astype(np.float32)
data.head()
```

Out[28]:

	Bronze	Silver	Gold	Platinum	Diamond	Master	GrandMaster	LeagueIndex	Age	HoursPerWe
0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	5.0	27.0	1
1	0.0	0.0	0.0	0.0	1.0	0.0	0.0	5.0	23.0	1
2	0.0	0.0	0.0	1.0	0.0	0.0	0.0	4.0	30.0	1
3	0.0	0.0	1.0	0.0	0.0	0.0	0.0	3.0	19.0	2
4	0.0	0.0	1.0	0.0	0.0	0.0	0.0	3.0	32.0	1

In [3]:
```
1 data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 3338 entries, 0 to 3337
Data columns (total 26 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Bronze                3338 non-null   float32
 1   Silver                3338 non-null   float32
 2   Gold                  3338 non-null   float32
 3   Platinum              3338 non-null   float32
 4   Diamond               3338 non-null   float32
 5   Master                3338 non-null   float32
 6   GrandMaster           3338 non-null   float32
 7   LeagueIndex           3338 non-null   float32
 8   Age                   3338 non-null   float32
 9   HoursPerWeek          3338 non-null   float32
 10  TotalHours            3338 non-null   float32
 11  APM                   3338 non-null   float32
 12  SelectByHotkeys       3338 non-null   float32
 13  AssignToHotkeys       3338 non-null   float32
 14  UniqueHotkeys         3338 non-null   float32
 15  MinimapAttacks        3338 non-null   float32
 16  MinimapRightClicks    3338 non-null   float32
 17  NumberOfPACs          3338 non-null   float32
 18  GapBetweenPACs        3338 non-null   float32
 19  ActionLatency         3338 non-null   float32
 20  ActionsInPAC          3338 non-null   float32
 21  TotalMapExplored      3338 non-null   float32
 22  WorkersMade           3338 non-null   float32
 23  UniqueUnitsMade       3338 non-null   float32
 24  ComplexUnitsMade      3338 non-null   float32
 25  ComplexAbilitiesUsed  3338 non-null   float32
dtypes: float32(26)
memory usage: 365.1 KB
```

In [4]:
```
inddatatxt = open("Data Explanation -- StarCraft II Ranked Players.txt")
print(inddatatxt.read())
var = []
for i in data:
    var.append(i)
```

```
Data project data categories

Data pulled from https://www.kaggle.com/danofer/skillcraft (https://www.kaggle.
com/danofer/skillcraft) which pulled the data from the University of California
Irvine's machine learning repository located at http://archive.ics.uci.edu/ml/d
atasets/SkillCraft1+Master+Table+Dataset (http://archive.ics.uci.edu/ml/dataset
s/SkillCraft1+Master+Table+Dataset)

This data is taken from replays of 1v1 ranked games of StarCraft II, a real tim
e strategy game in which two or more players compete by accruing resources, bui
lding an army and wiping out the enemy.
A player loses when all of their buildings are destroyed or when they surrende
r.
Players each start with one base hub and a number of workers located symmetrica
lly on a map. The map is littered with locations containing clusters of mineral
s and usually gas geysers.
These are the resources which the players will use to build buildings, construc
t new technologies and produce armies.



00 GameID                                #
01 LeagueIndex                  # Numerical list (1-8) based on categorical ran
k. Can be considered both categorical and numerical
02 Age                          # Age of player
03 HoursPerWeek                 # Hours player plays per week
04 TotalHours                   # Total hours player has played
05 APM                          # Average number of actions per minute (actions
are commands given to units or structures)

                # For the next few variables, one must understand control group
s. Control groups are hotkeys that a player can assign units to such that when
 the hotkey is selected, the units and buildings assigned become selected for q
uick use.

06 SelectByHotkeys                       # Number of units/buildings selected wi
th hotkeys per timestamp
07 AssignToHotkeys                       # Number of units/buildings assigned to
control groups per timestamp
08 UniqueHotkeys                         # Number of unique hotkeys used per tim
estamp
09 MinimapAttacks                        # Number of times a player uses the min
imap to order a unit to attack anything on the way to a location per timestamp
10 MinimapRightClicks           # Number of right clicks on the minimap per tim
estamp

                # For the next few variables, PACs (Perception Action Cycles) a
re when a player focuses on a certain area by moving their screen and then make
s an action.

11 NumberOfPACs                 # Number of PACs per timestamp
```

```
12 GapBetweenPACs                       # Average number of milliseconds betwee
n PACs
13 ActionLatency                        # Average time between focusing on an a
rea and making the first action of the current PAC
14 ActionsInPAC                 # Average number of actions within a PAC
15 TotalMapExplored             # The amount of the map explored by a player pe
r timestamp
16 WorkersMade                  # Number of worker units trained per timestamp
17 UniqueUnitsMade                      # Number of unique units made per times
tamp
18 ComplexUnitsMade             # Number of units that require direct controls
 to be effective trained per timestamp
19 ComplexAbilitiesUsed         # Number of times abilities from aforementioned
complex units are used per timestamp
```

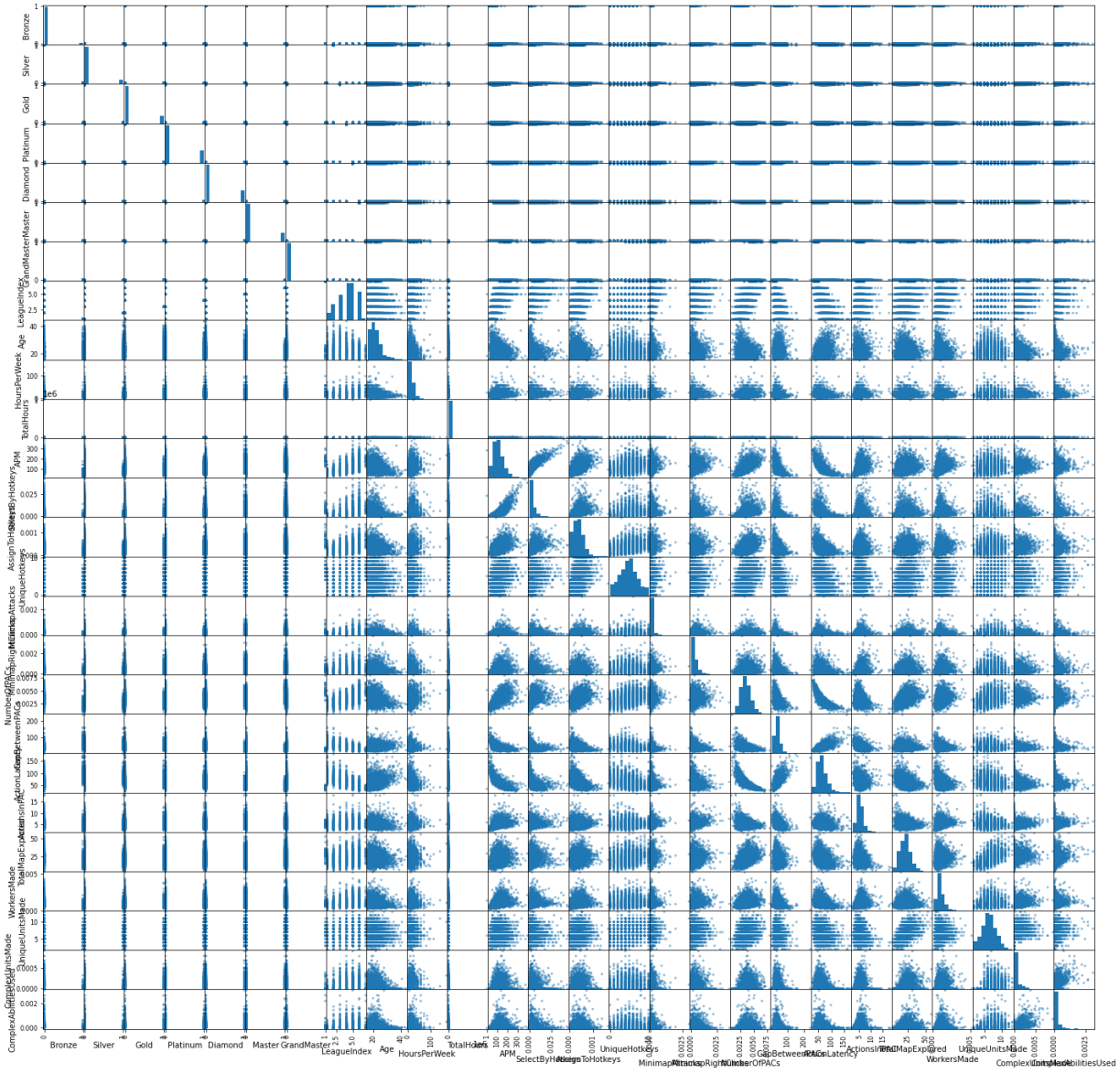
In [5]:
```
1 data.describe()
```

Out[5]:

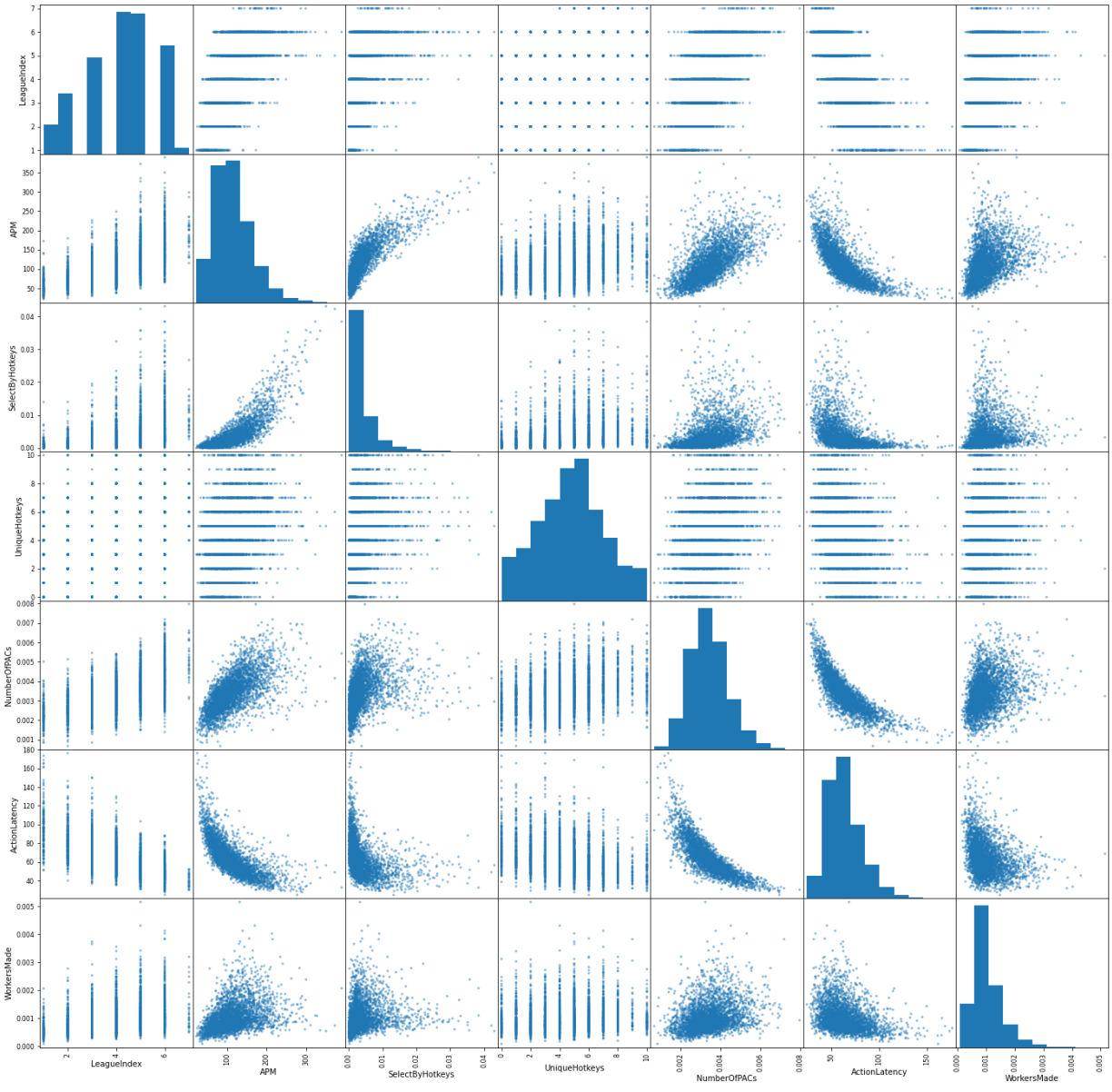
	Bronze	Silver	Gold	Platinum	Diamond	Master	GrandMaste
count	3338.000000	3338.000000	3338.000000	3338.000000	3338.000000	3338.000000	3338.000000
mean	0.050030	0.103954	0.165668	0.242960	0.240863	0.186040	0.01048
std	0.218039	0.305247	0.371838	0.428935	0.427671	0.389197	0.10187
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
50%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
75%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
max	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

In [6]:

```
pd.plotting.scatter_matrix(data,figsize=(24,24))
plt.show()
```



In [7]:
```
IDs = ['LeagueIndex','APM','SelectByHotkeys','UniqueHotkeys','NumberOfPACs',
pd.plotting.scatter_matrix(data[IDs],figsize=(24,24))
plt.show()
```



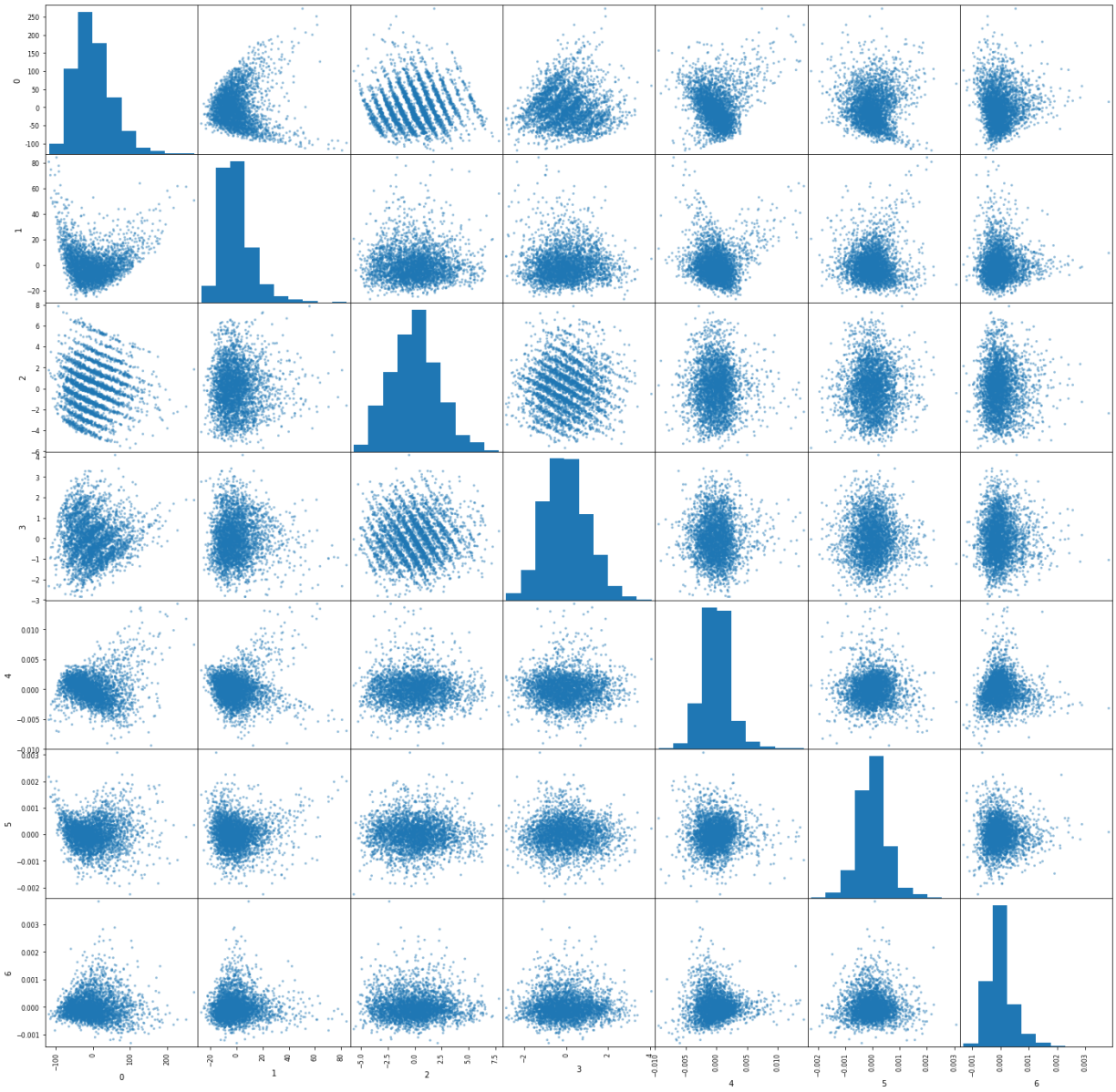
Examine variables from sklearn.decomposition.PCA

In [8]:
```
pca = PCA()
pca.fit(data[IDs])
data_transformed = pd.DataFrame(pca.transform(data[IDs]))
data_transformed.shape
```

Out[8]: (3338, 7)

In [9]:

```
1 pd.plotting.scatter_matrix(data_transformed,figsize=(24,24))
2 plt.show()
```



In [10]:

```
# Code for the following graphs based on https://www.kaggle.com/immodal/pred
leagueTitles = [       "Bronze","Silver","Gold",   "Platinum","Diamond","Mas
leagueIndexs = [            1.0,     2.0,   3.0,          4.0,      5.0,
colors       = ['darkgoldenrod','silver','gold','springgreen',   'aqua',  'b
leagueDict   = dict()
for i, ind in enumerate(leagueIndexs):
    leagueDict[ind] = leagueTitles[i]
leagueLabeled = data["LeagueIndex"].replace(leagueDict)

def violin_plot(y, title, w, h):
    plt.figure(figsize=(w,h))
    ax = sns.violinplot(x=leagueLabeled,y=y,palette=colors,order=leagueTitle
    ax.set(xlabel='League')
    ax.set_ylim(0,)
    plt.title(title)
    plt.show()

def density_plot(x, title, w, h):
    fig, (ax_box, ax_hist) = plt.subplots(2, sharex=True, gridspec_kw={"heig
    fig.set_size_inches(w,h)
    ax_box.set_xlim(0,x.max())
    ax_hist.set_xlim(0,x.max())
    sns.boxplot(x,ax=ax_box)
    sns.distplot(x,ax=ax_hist)
    ax_box.set(yticks=[])
    sns.despine(ax=ax_hist)
    sns.despine(ax=ax_box,left=True)
    ax_box.set_title(title)
    plt.show()

def plot(feature, fig_num):
    density_plot(data[feature],f"Figure {fig_num}: {feature} Distribution",
    violin_plot(data[feature],f"Figure {fig_num}: {feature} by League", 11,8
```

In [11]:
```
1 plot("APM",1)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

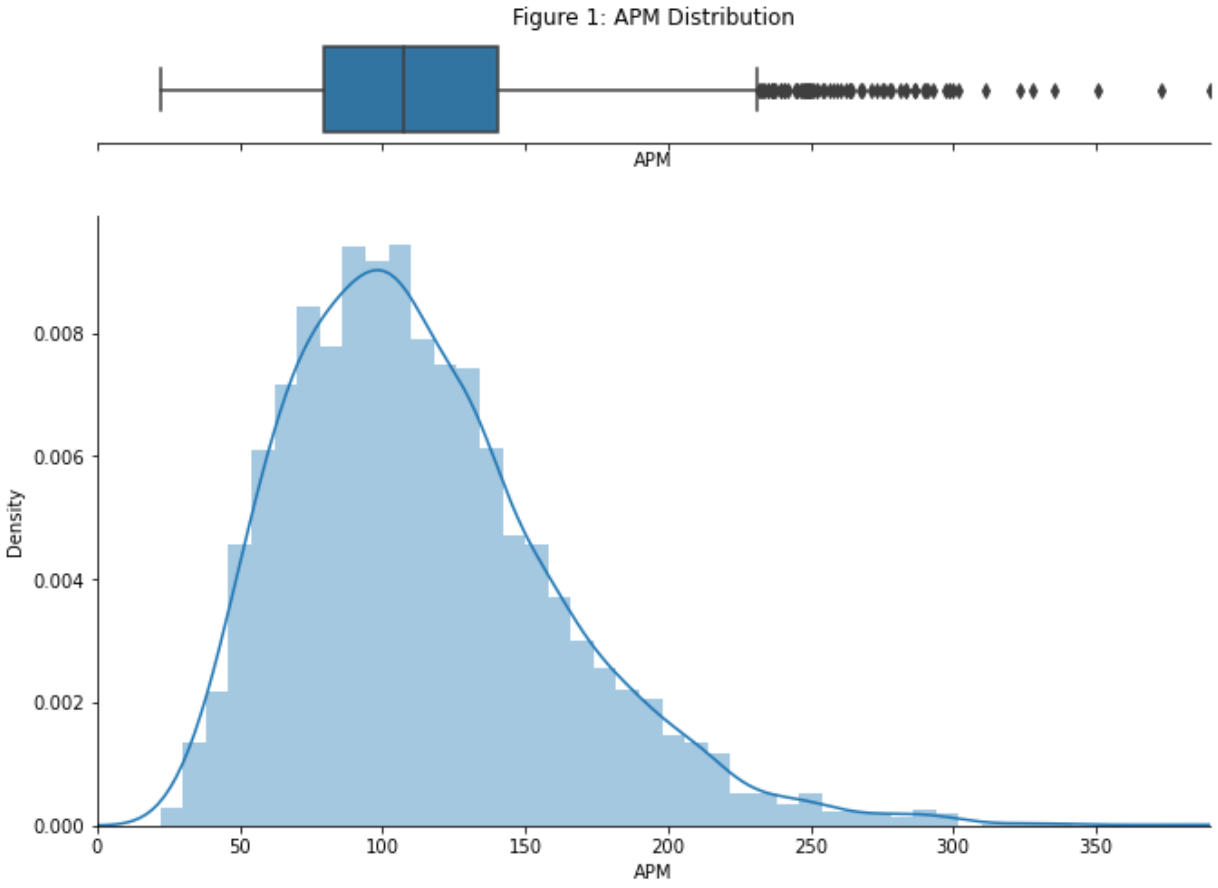
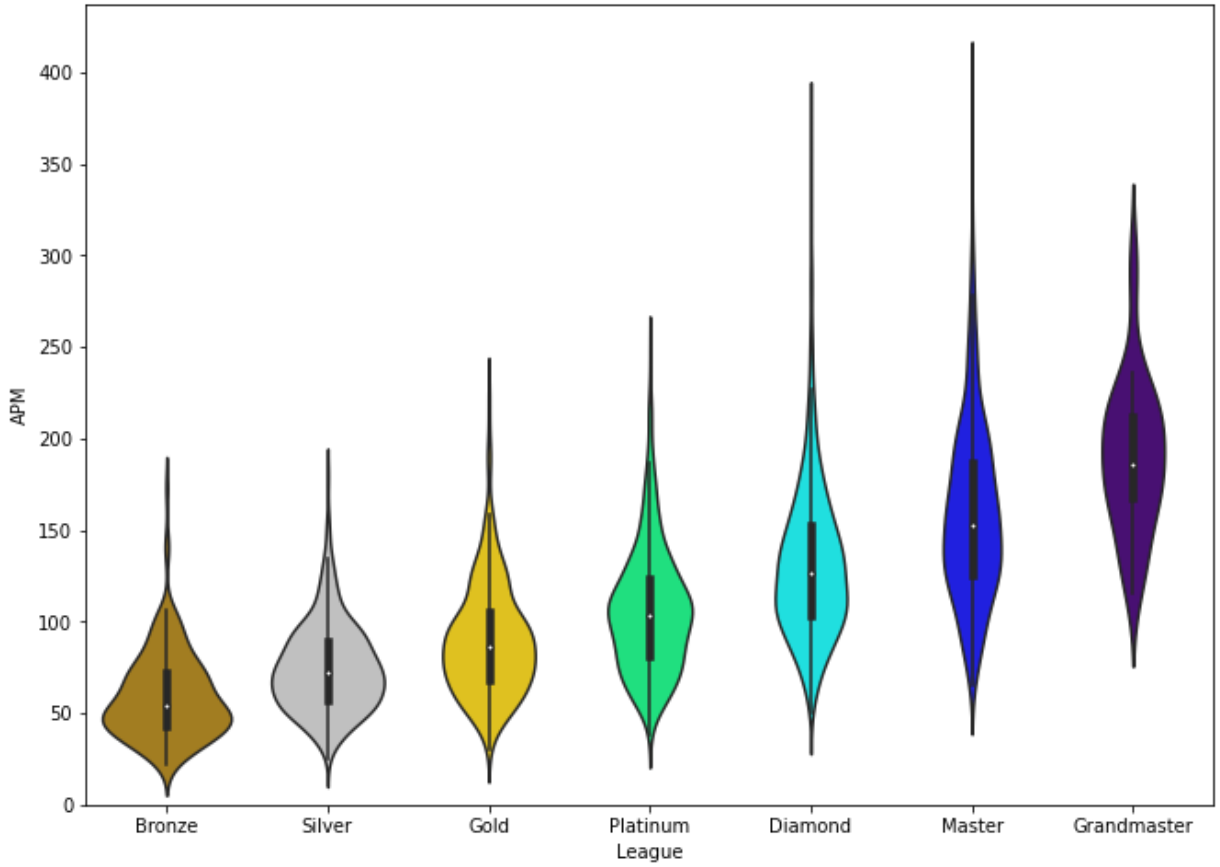


Figure 1: APM by League



In [12]:
```
1 plot("HoursPerWeek",2)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

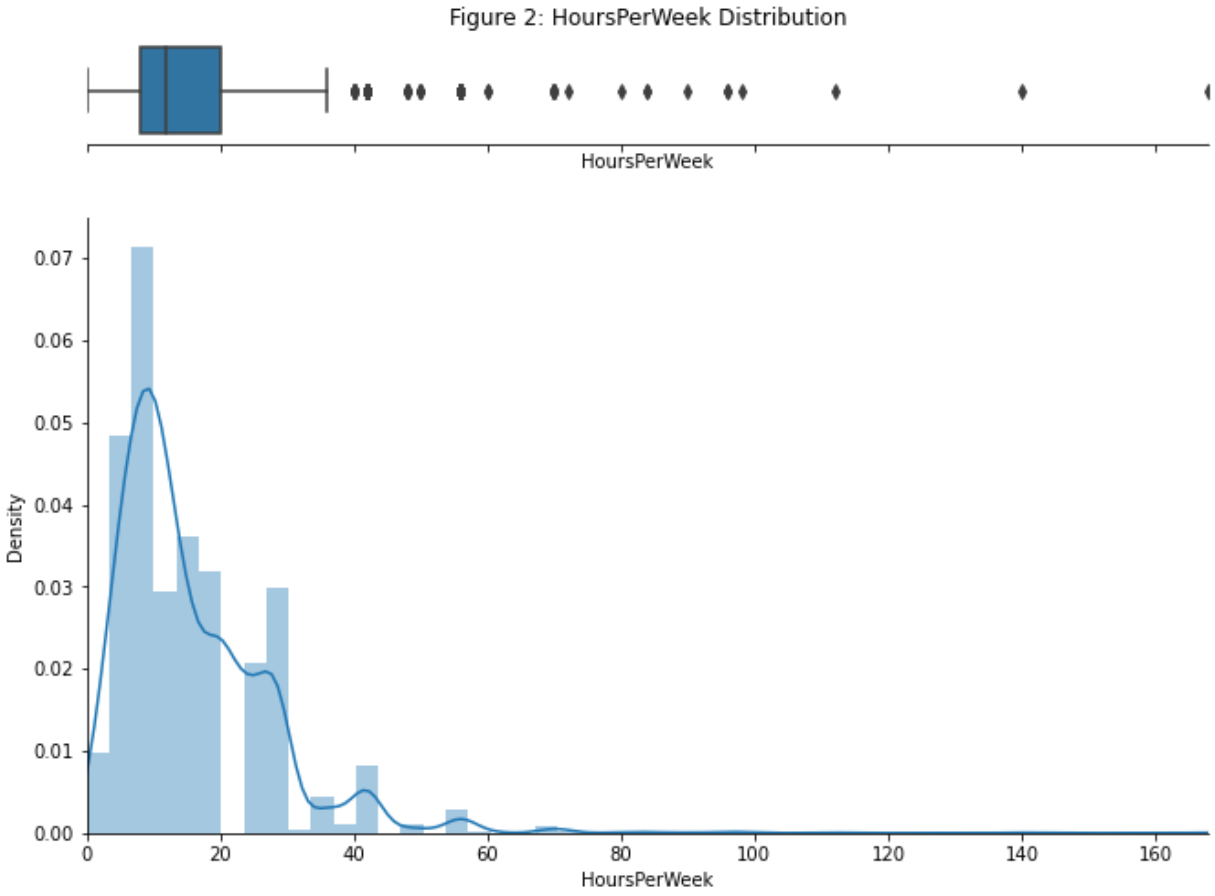
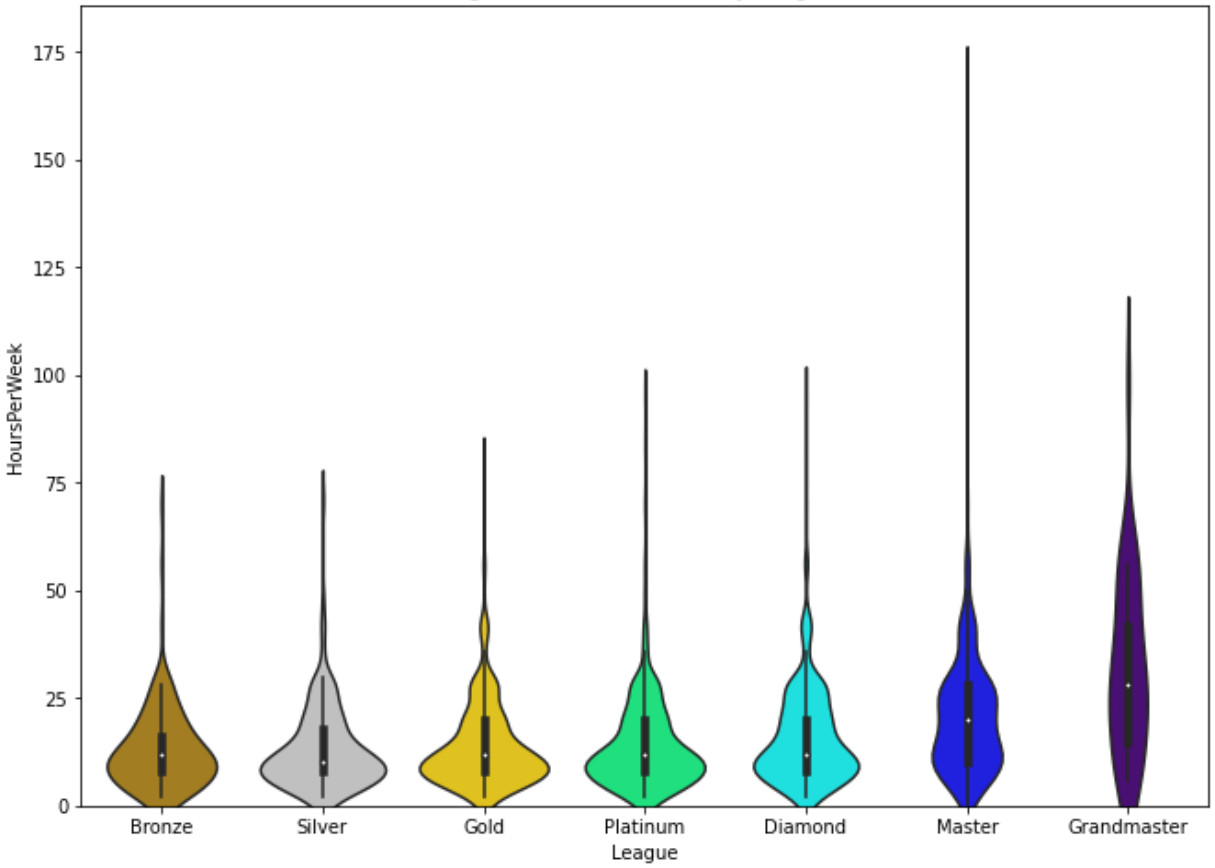
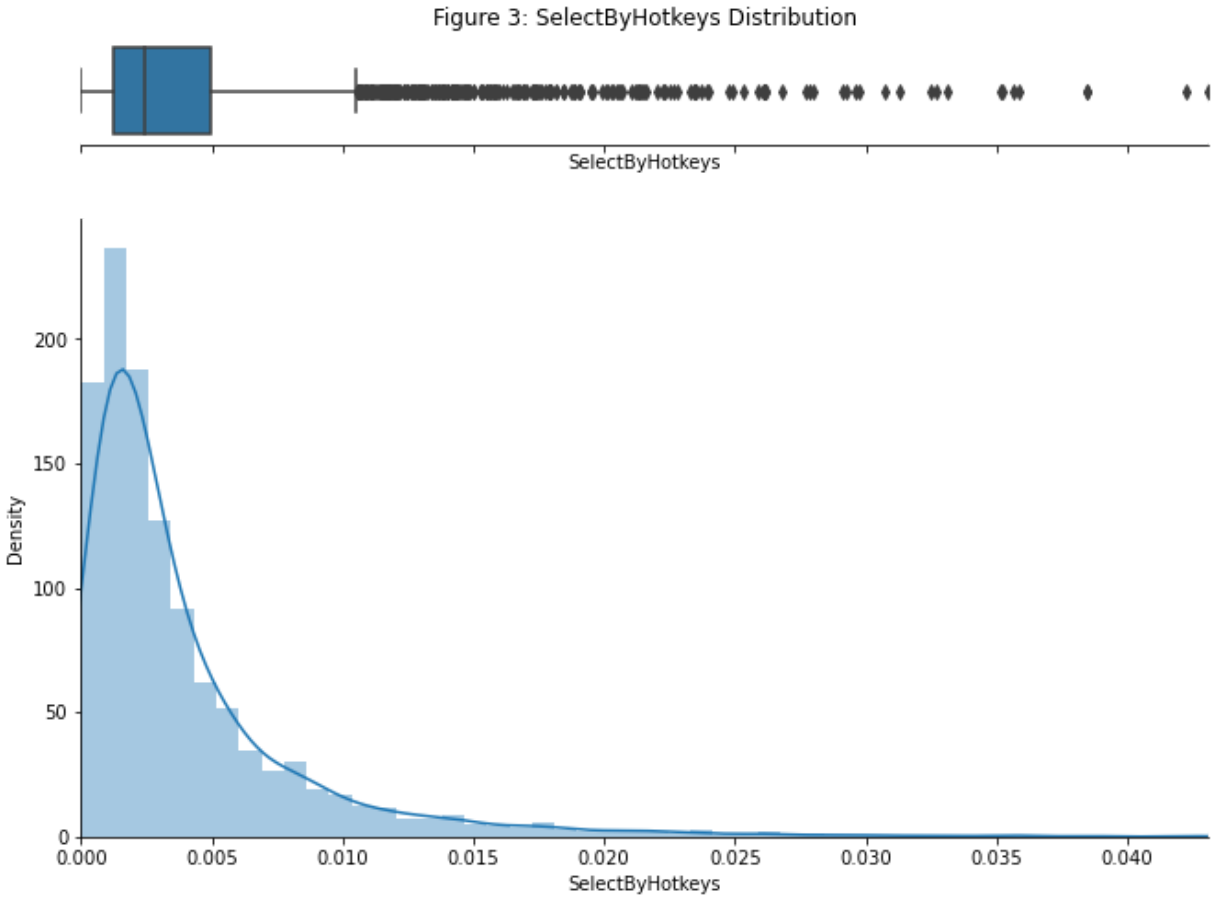


Figure 2: HoursPerWeek by League



In [13]:
```
1 plot("SelectByHotkeys",3)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```



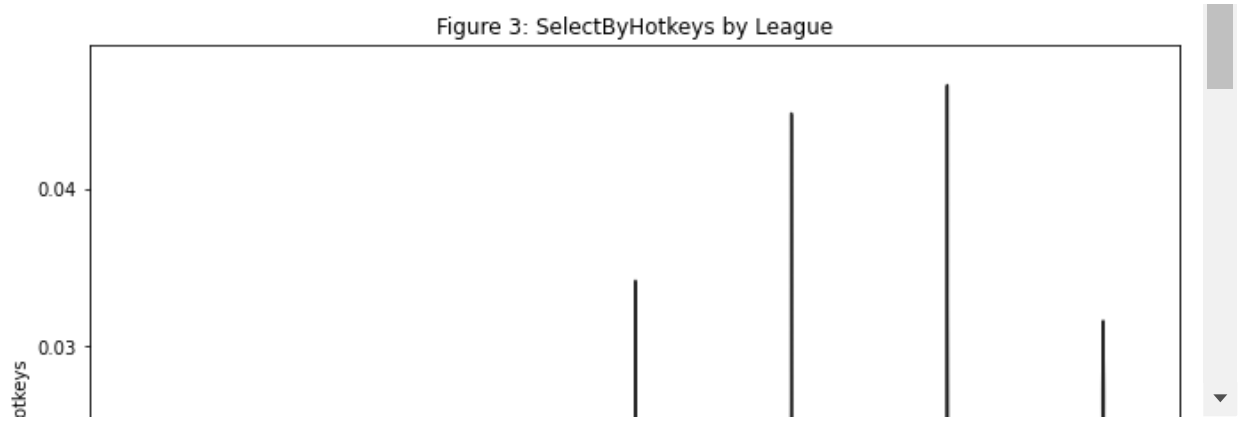
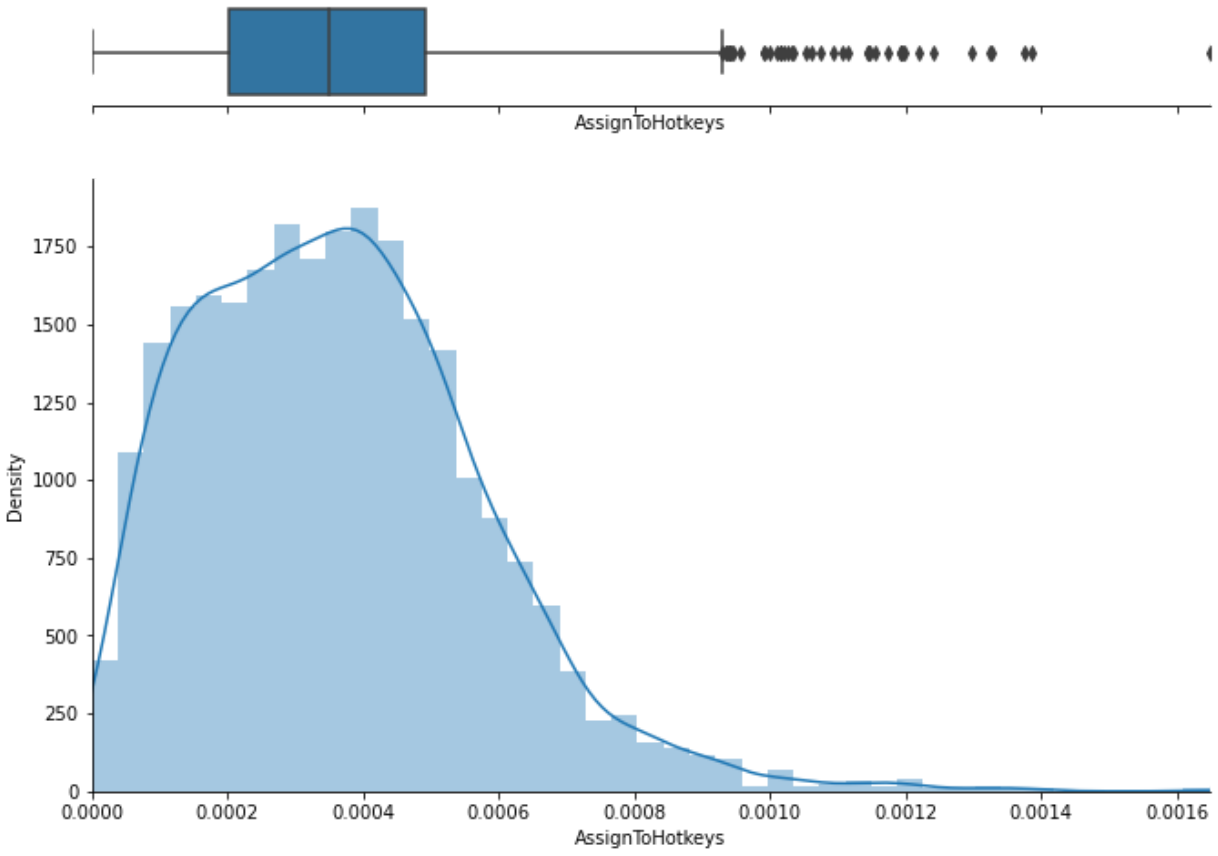


Figure 3: SelectByHotkeys by League
0.04
0.03
otkeys

In [14]:
```
1 plot("AssignToHotkeys",4)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Figure 4: AssignToHotkeys Distribution



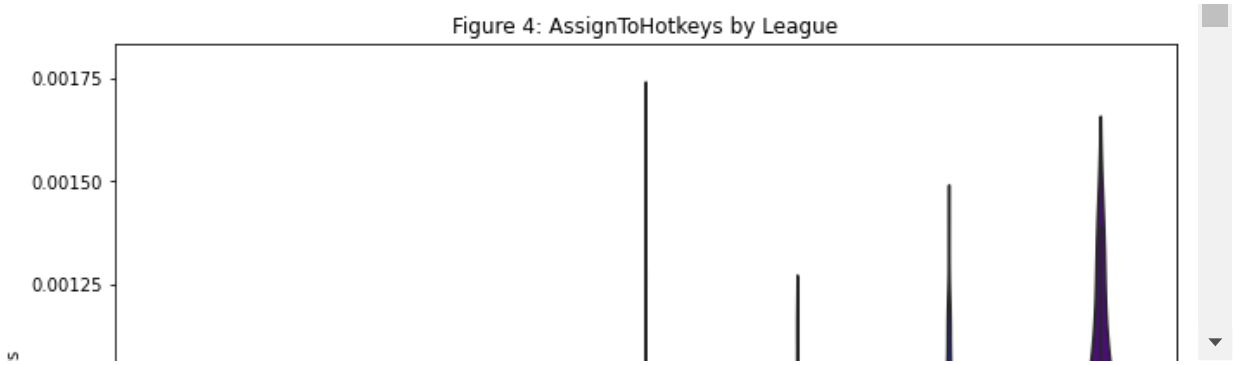


Figure 4: AssignToHotkeys by League
0.00175
0.00150
0.00125

In [15]:
```
1 plot("UniqueHotkeys",5)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Figure 5: UniqueHotkeys Distribution

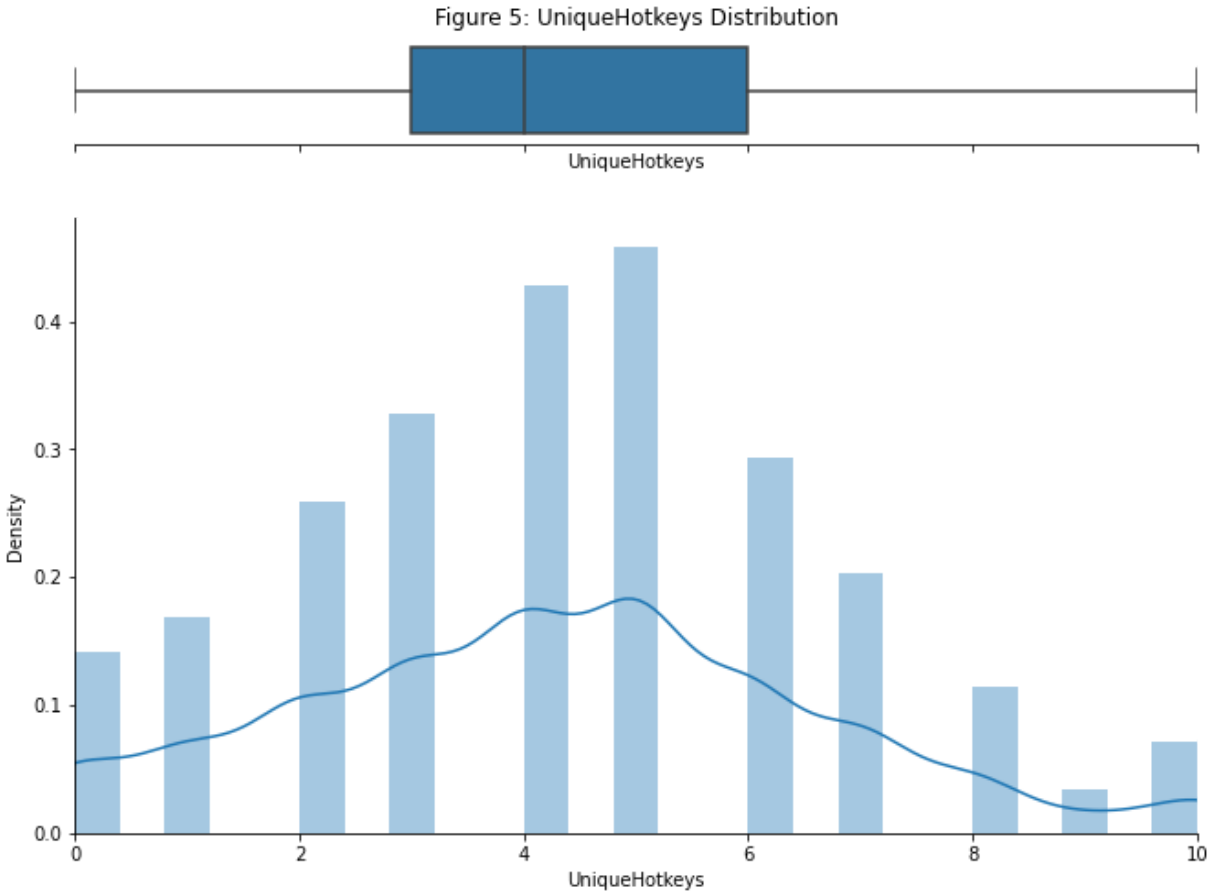
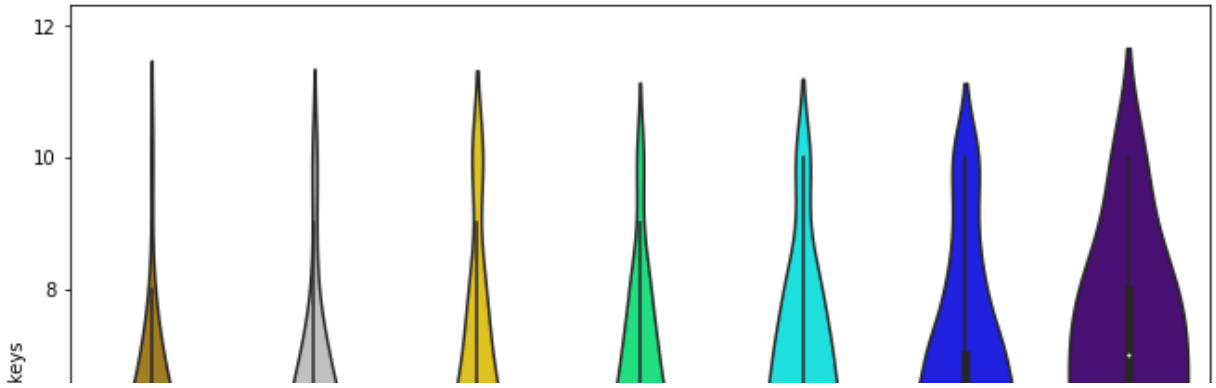


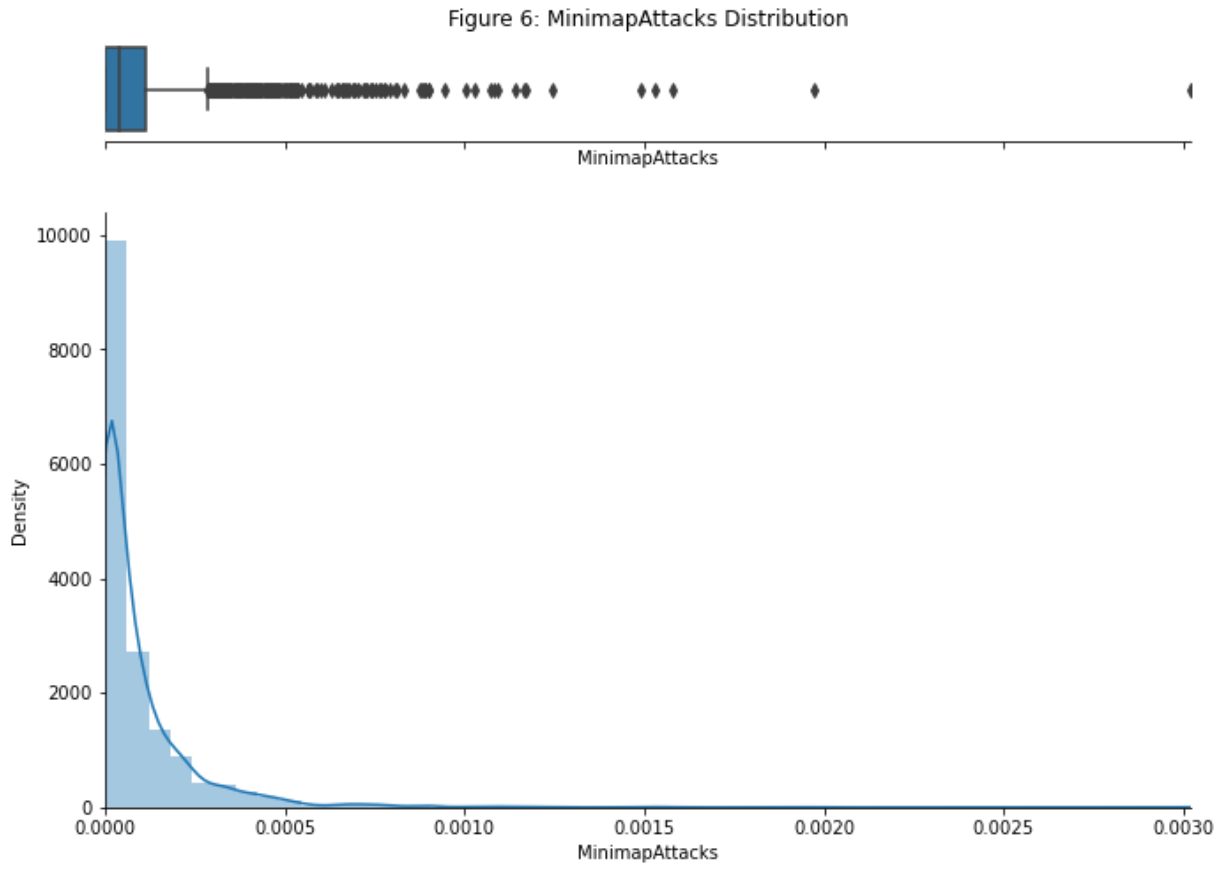
Figure 5: UniqueHotkeys by League



In [16]:
```
1 plot("MinimapAttacks",6)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Figure 6: MinimapAttacks Distribution



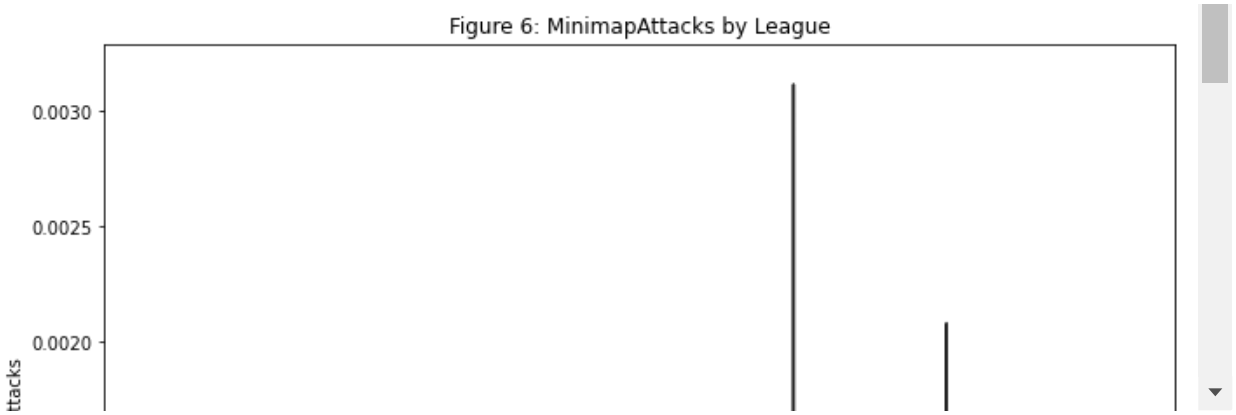


Figure 6: MinimapAttacks by League
0.0030
0.0025
0.0020
ttacks

In [17]:
```
1 plot("MinimapRightClicks",7)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Figure 7: MinimapRightClicks Distribution

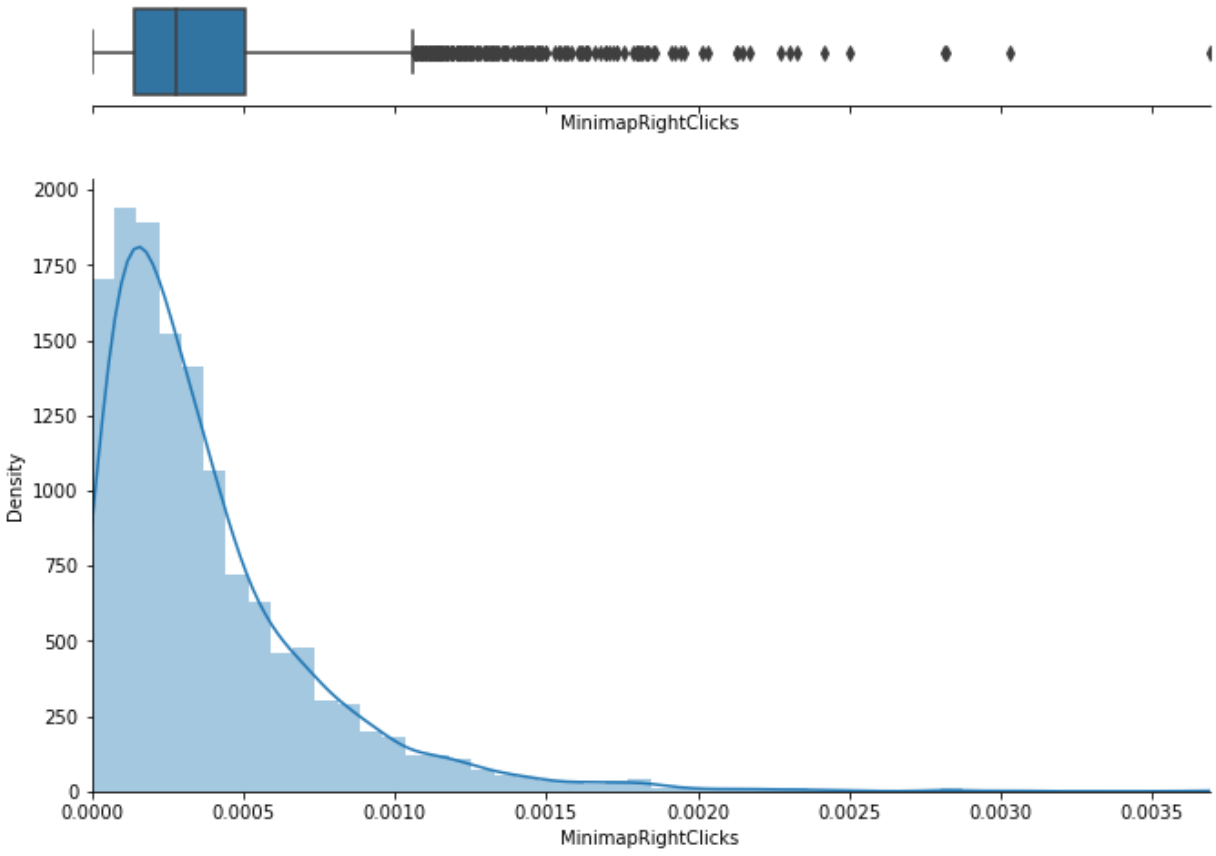
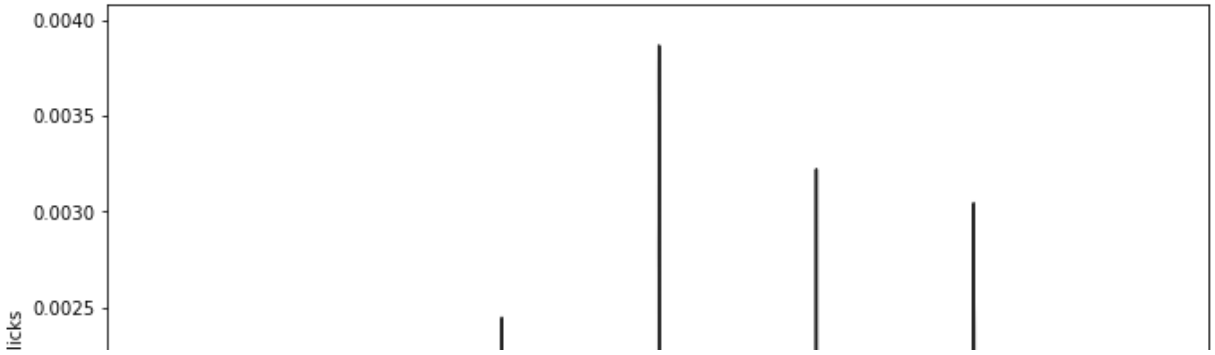


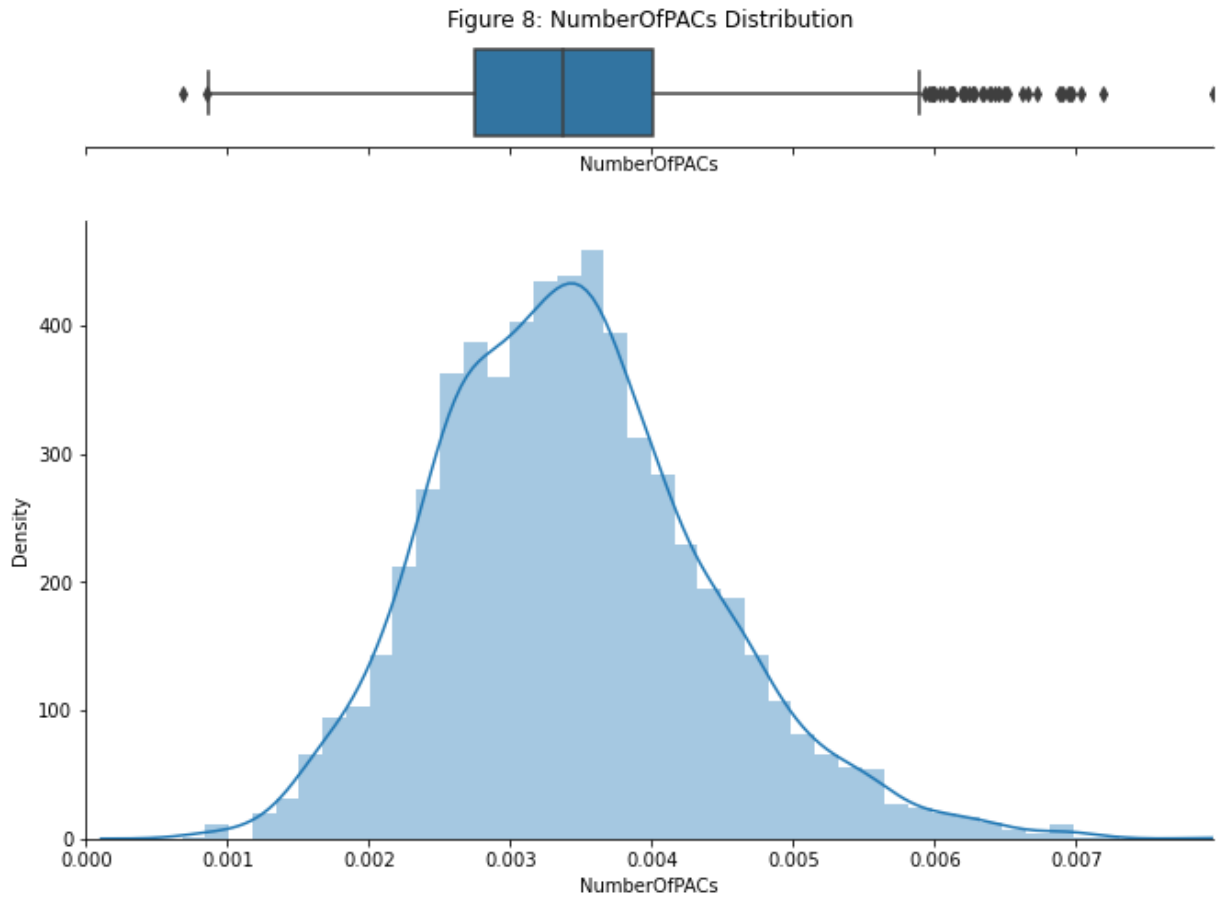
Figure 7: MinimapRightClicks by League



In [18]:
```
1 plot("NumberOfPACs",8)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Figure 8: NumberOfPACs Distribution



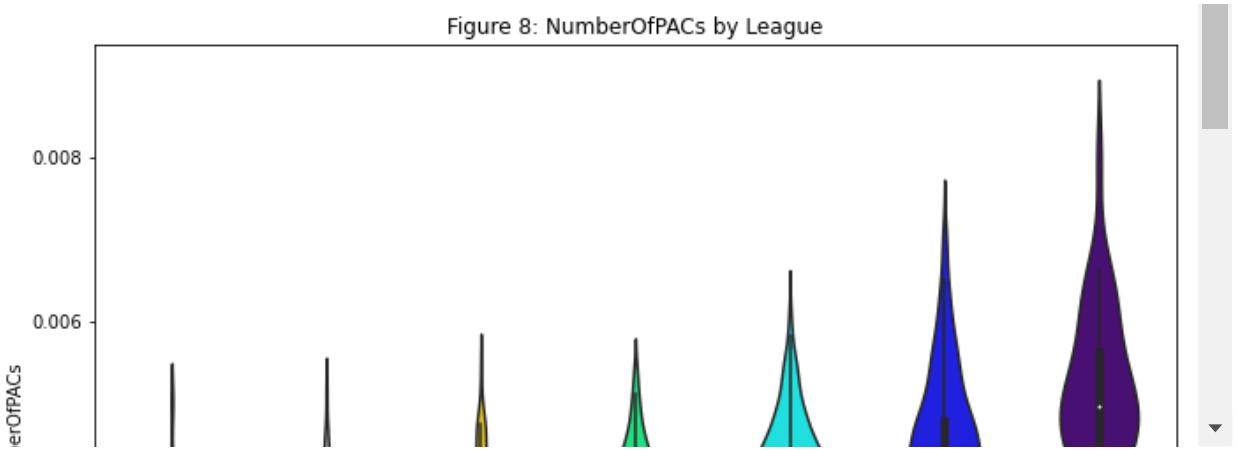


Figure 8: NumberOfPACs by League
0.008
0.006
erOfPACs

In [19]:
```
1 plot("GapBetweenPACs",9)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Figure 9: GapBetweenPACs Distribution

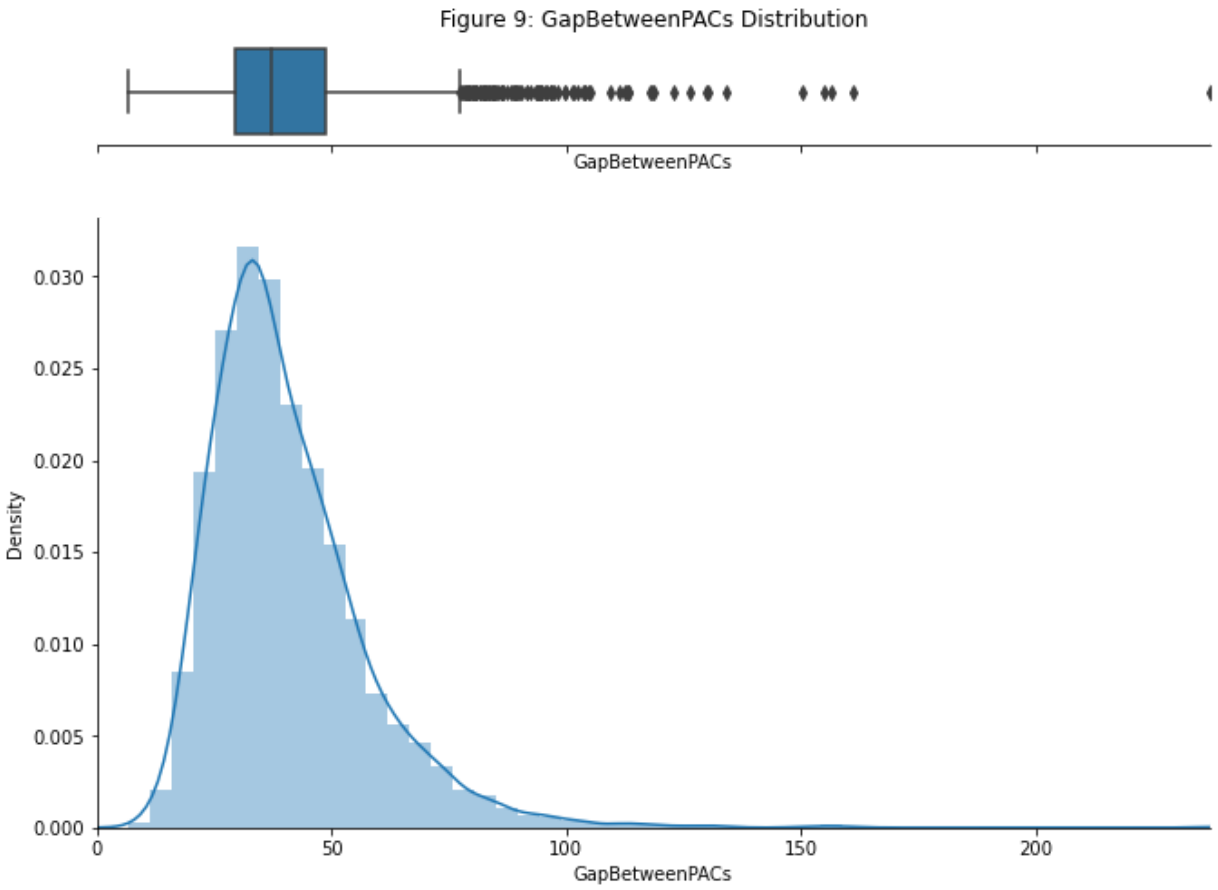
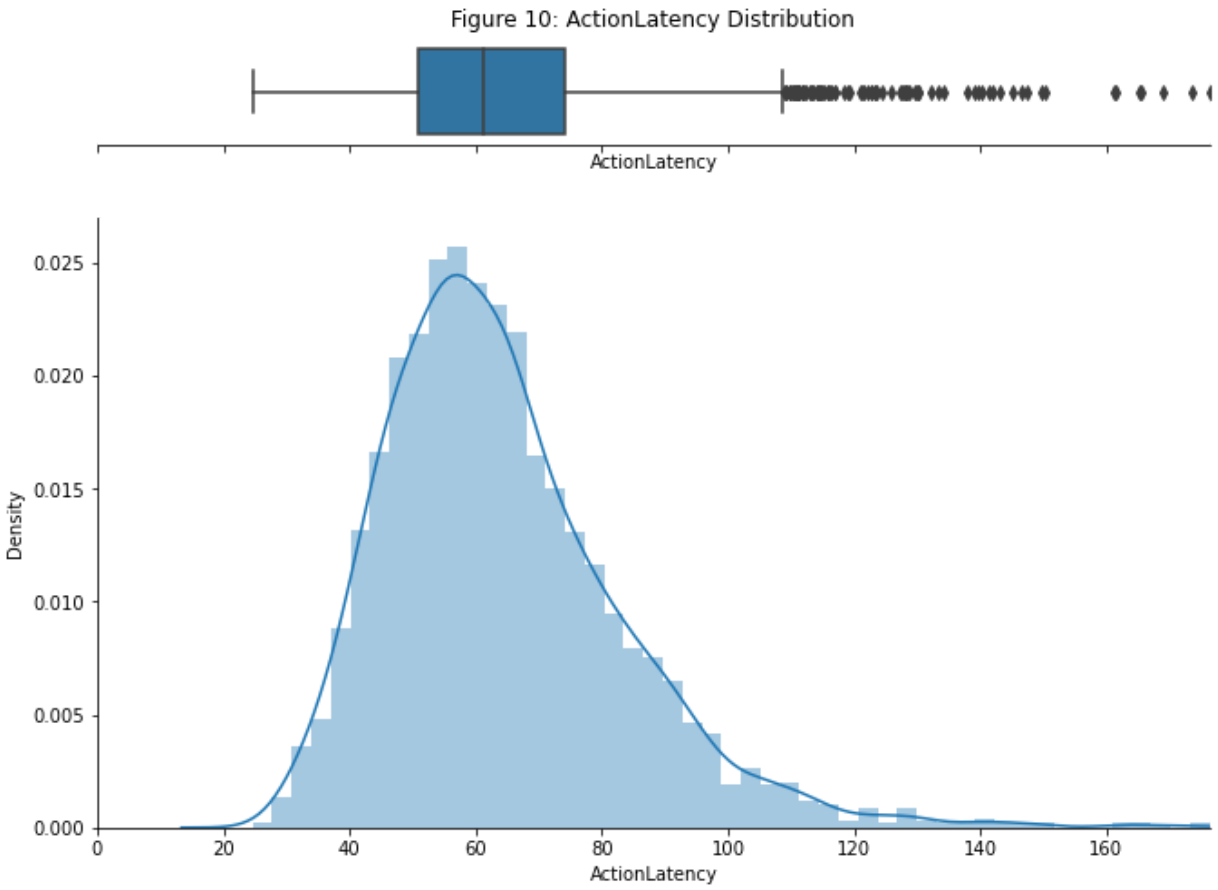


Figure 9: GapBetweenPACs by League

In [20]:
```
1 plot("ActionLatency",10)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Figure 10: ActionLatency Distribution



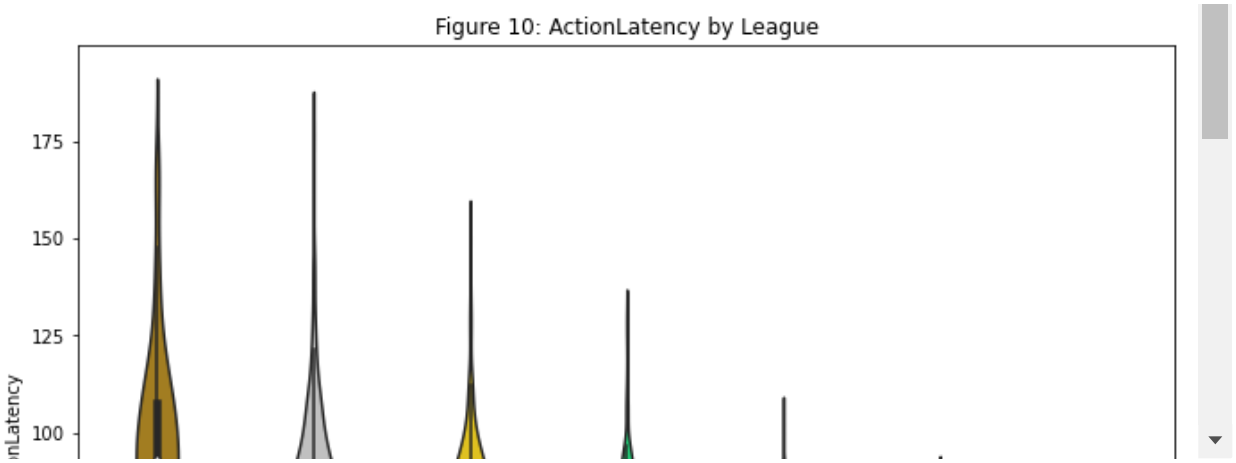


Figure 10: ActionLatency by League
175
150
125
100
nLatency

In [21]:
```
1 plot("ActionsInPAC",11)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Figure 11: ActionsInPAC Distribution

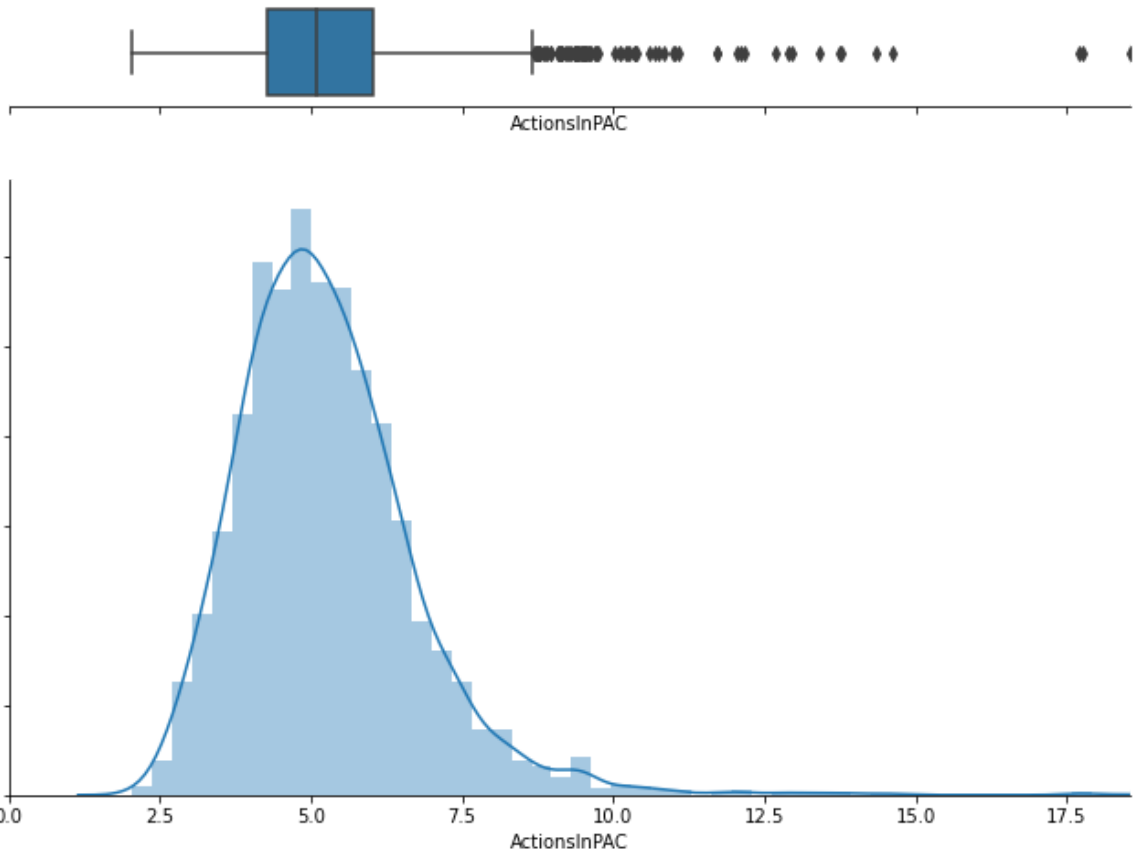
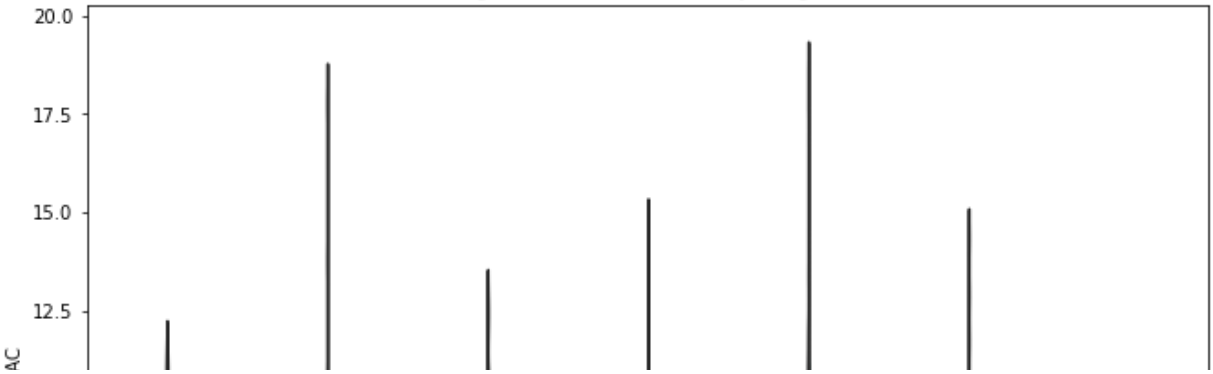


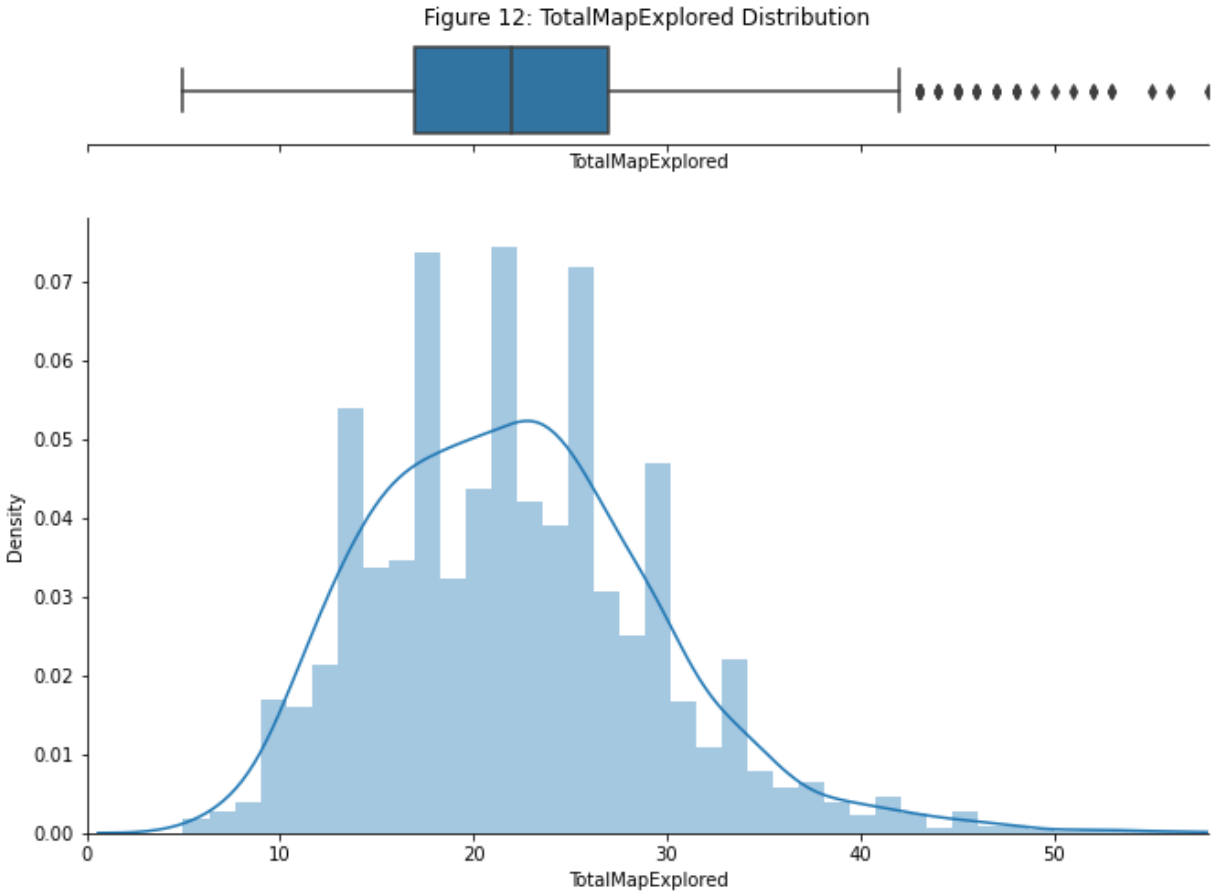
Figure 11: ActionsInPAC by League



In [22]:
```
1 plot("TotalMapExplored",12)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Figure 12: TotalMapExplored Distribution



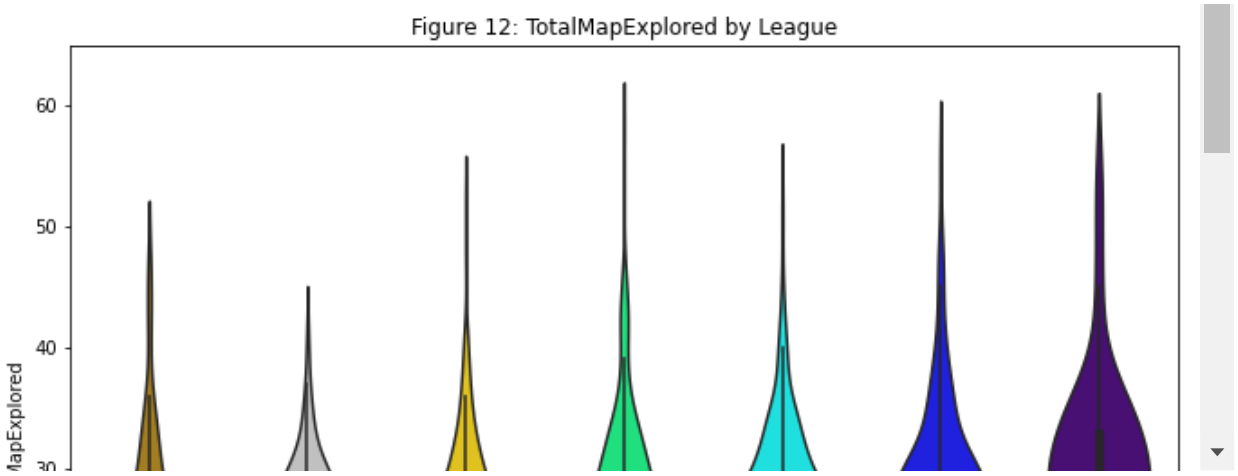


Figure 12: TotalMapExplored by League
60
50
40
MapExplored

In [23]:
```
1 plot("WorkersMade",13)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Figure 13: WorkersMade Distribution

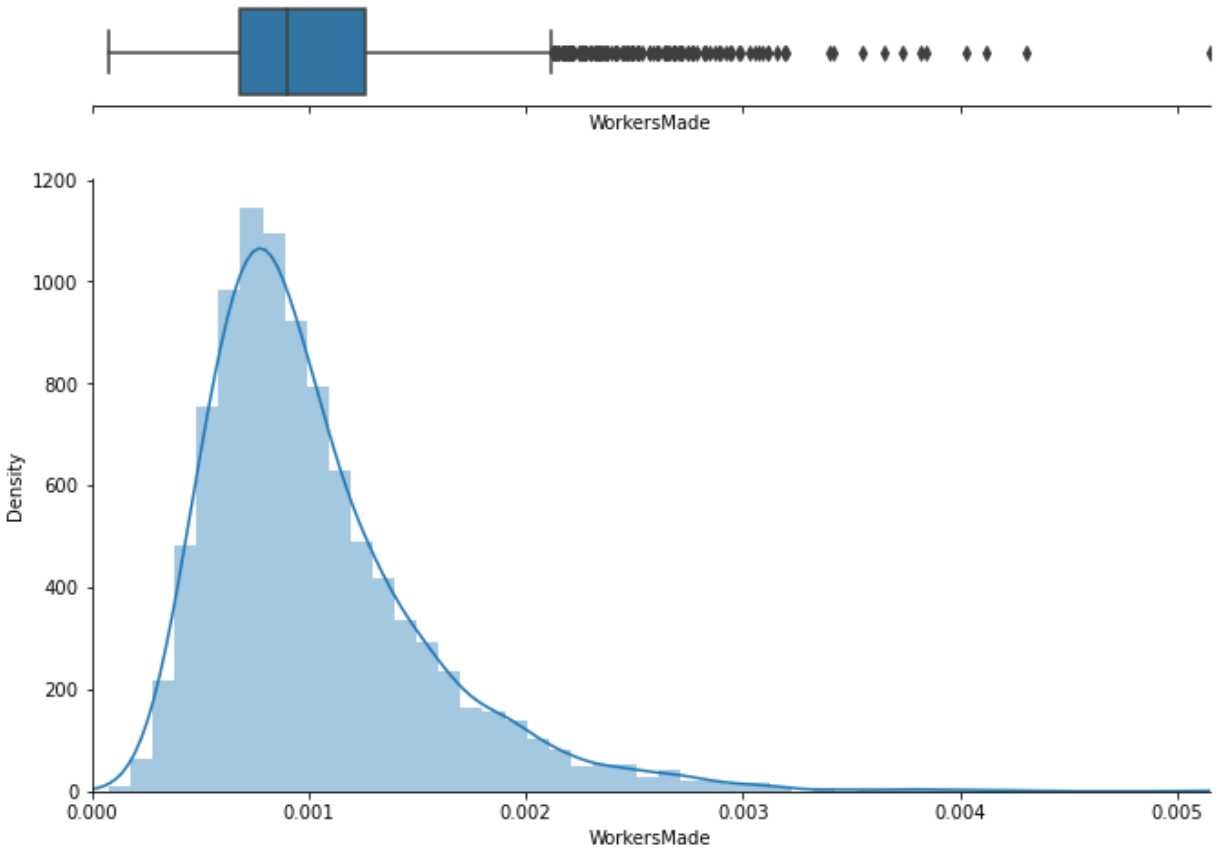
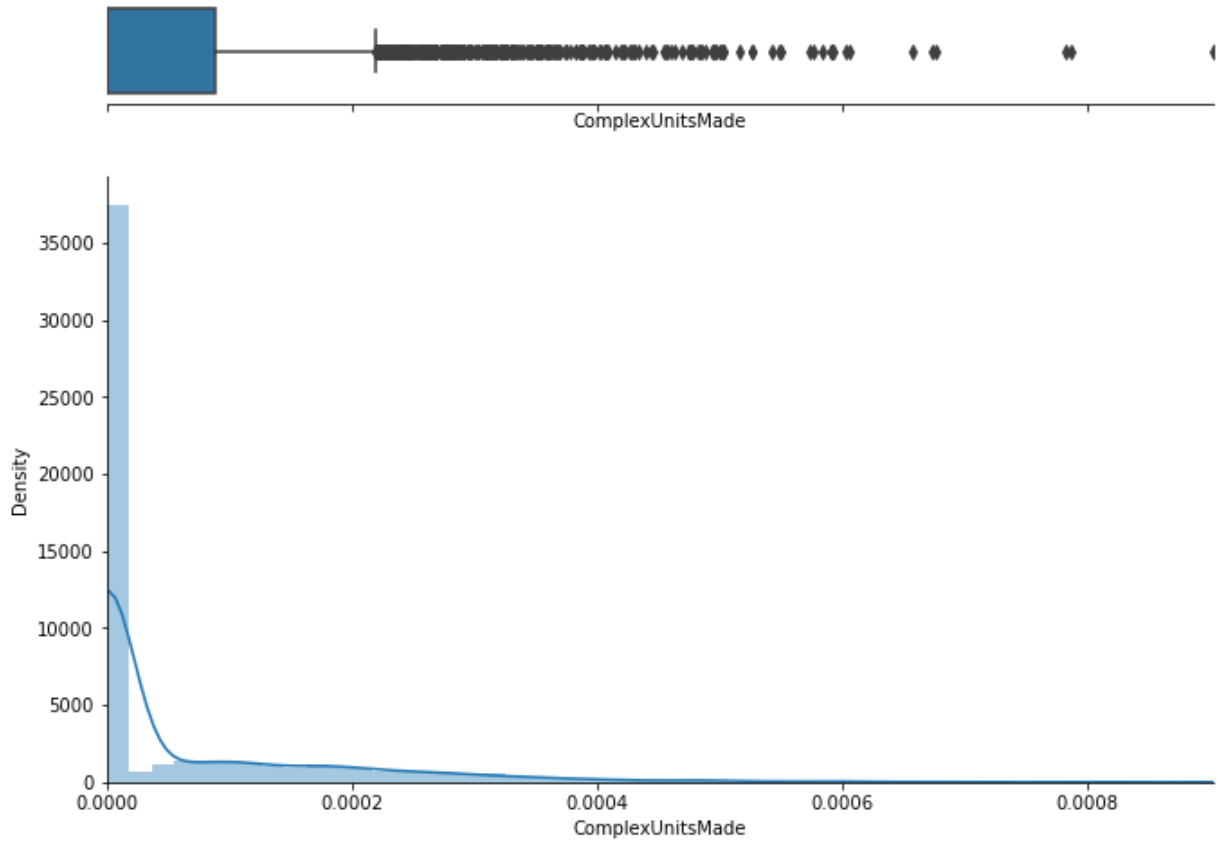


Figure 13: WorkersMade by League

In [24]:
```
1 plot("ComplexUnitsMade",14)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Figure 14: ComplexUnitsMade Distribution



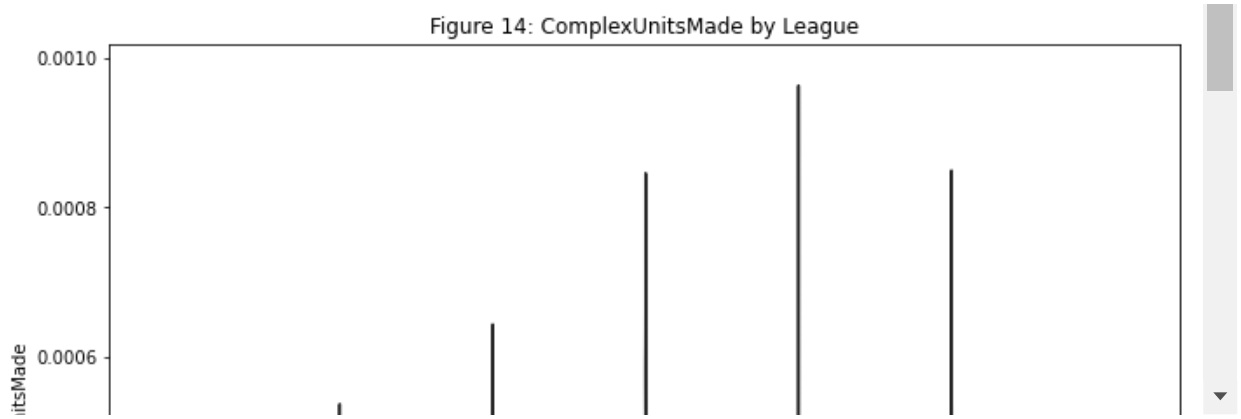


Figure 14: ComplexUnitsMade by League
0.0010
0.0008
0.0006
itsMade

In [25]:
```
1 plot("ComplexAbilitiesUsed",15)
```

```
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\Triplea657\anaconda3\envs\MSCS335\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Figure 15: ComplexAbilitiesUsed Distribution

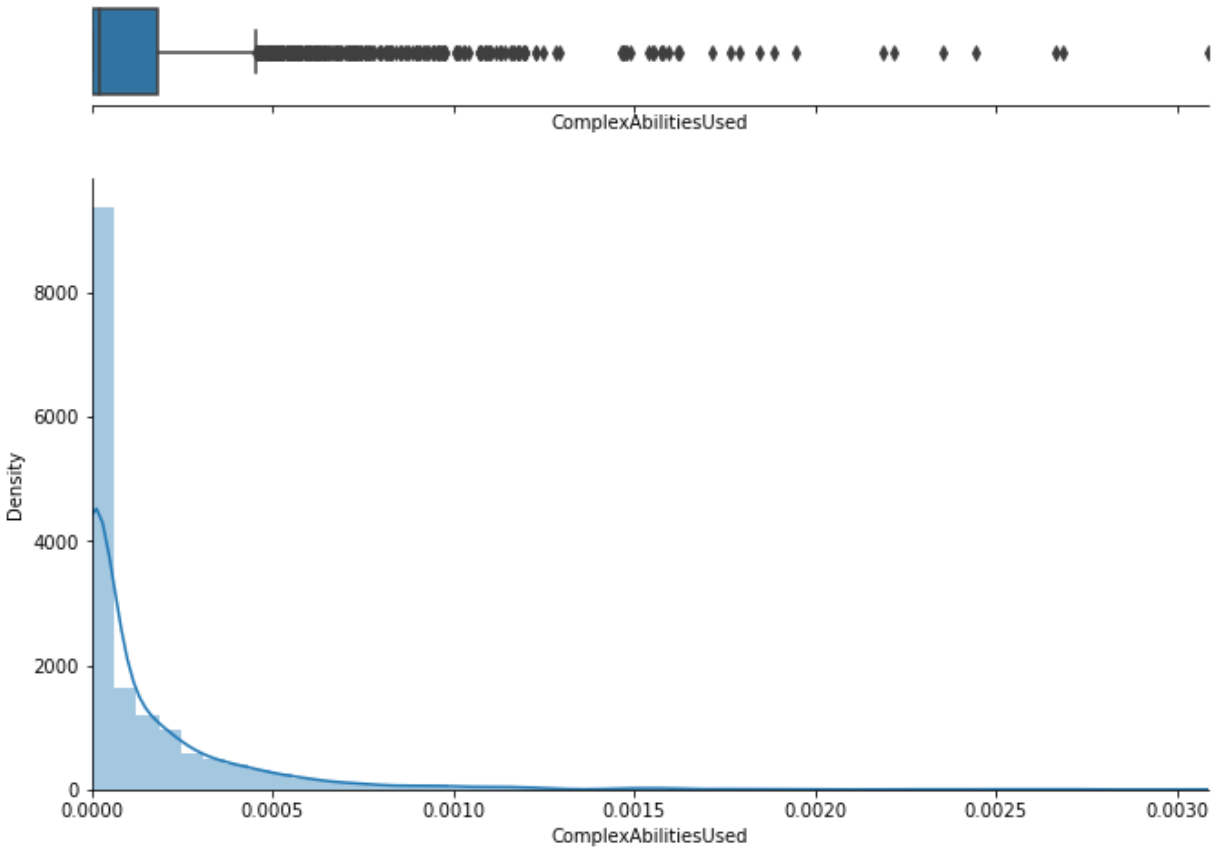


Figure 15: ComplexAbilitiesUsed by League