In [13]:
```
import pandas as pd
```

In [15]:
```
df=pd.read_csv('D:\Project\Game Sales\Game.csv')
df.head(5)
```

Out[15]:

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Oth
0	1	Wii Sports	Wii	2006.0	Sports	Nintendo	41.49	29.02	3.77	
1	2	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	6.81	
2	3	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.85	12.88	3.79	
3	4	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.75	11.01	3.28	
4	5	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	

In [16]:
```
df.describe()
```

Out[16]:

	Rank	Year	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Glol
count	16598.000000	16327.000000	16598.000000	16598.000000	16598.000000	16598.000000	1659
mean	8300.605254	2006.406443	0.264667	0.146652	0.077782	0.048063	
std	4791.853933	5.828981	0.816683	0.505351	0.309291	0.188588	
min	1.000000	1980.000000	0.000000	0.000000	0.000000	0.000000	
25%	4151.250000	2003.000000	0.000000	0.000000	0.000000	0.000000	
50%	8300.500000	2007.000000	0.080000	0.020000	0.000000	0.010000	
75%	12449.750000	2010.000000	0.240000	0.110000	0.040000	0.040000	

	Rank	Year	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Glol
max	16600.000000	2020.000000	41.490000	29.020000	10.220000	10.570000	8

In [18]:
```
df.info
```

Out[18]:
```
<bound method DataFrame.info of         Rank                                     Name Platform  \
0          1                                 Wii Sports      Wii
1          2                          Super Mario Bros.      NES
2          3                             Mario Kart Wii      Wii
3          4                          Wii Sports Resort      Wii
4          5                    Pokemon Red/Pokemon Blue       GB
...      ...                                        ...      ...
16593  16596          Woody Woodpecker in Crazy Castle 5      GBA
16594  16597               Men in Black II: Alien Escape       GC
16595  16598  SCORE International Baja 1000: The Official Game      PS2
16596  16599                                Know How 2       DS
16597  16600                          Spirits & Spells      GBA

          Year         Genre   Publisher  NA_Sales  EU_Sales  JP_Sales
\
0       2006.0        Sports    Nintendo     41.49     29.02      3.77
1       1985.0      Platform    Nintendo     29.08      3.58      6.81
2       2008.0        Racing    Nintendo     15.85     12.88      3.79
```

```
3      2009.0         Sports    Nintendo     15.75     11.01      3.28

4      1996.0   Role-Playing    Nintendo     11.27      8.89     10.22

...       ...            ...         ...       ...       ...       ...

16593  2002.0       Platform       Kemco      0.01      0.00      0.00

16594  2003.0        Shooter  Infogrames      0.01      0.00      0.00

16595  2008.0         Racing  Activision      0.00      0.00      0.00

16596  2010.0         Puzzle    7G//AMES      0.00      0.01      0.00

16597  2003.0       Platform     Wanadoo      0.01      0.00      0.00


       Other_Sales  Global_Sales
0             8.46         82.74
1             0.77         40.24
2             3.31         35.82
3             2.96         33.00
4             1.00         31.37
...            ...           ...
16593         0.00          0.01
16594         0.00          0.01
16595         0.00          0.01
16596         0.00          0.01
16597         0.00          0.01

[16598 rows x 11 columns]>
```

In [19]:
```
df.shape
```

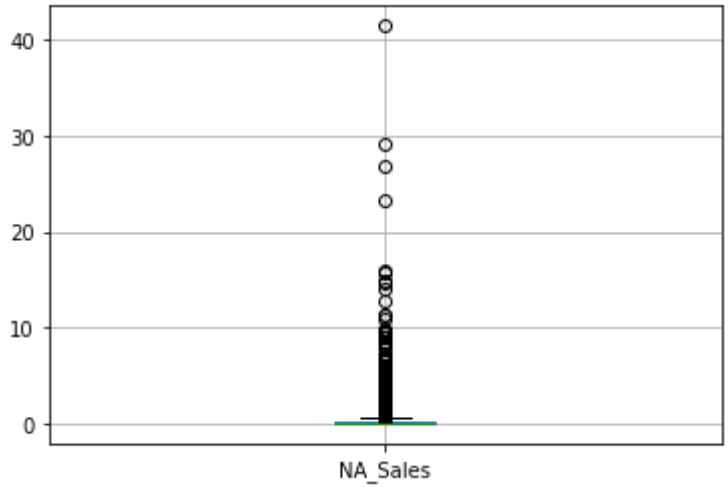
Out[19]:
```
(16598, 11)
```

In [21]:
```
df.isnull().sum()
```

Out[21]:
```
Rank            0
Name            0
Platform        0
Year          271
Genre           0
Publisher      58
NA_Sales        0
EU_Sales        0
JP_Sales        0
Other_Sales     0
Global_Sales    0
dtype: int64
```

In [25]:
```
df.boxplot(['NA_Sales'])
```

Out[25]:
```
<matplotlib.axes._subplots.AxesSubplot at 0xd4dff58>
```



In [39]:
```
df[df.NA_Sales>40]
```

Out[39]:

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales
0	1	Wii Sports	Wii	2006.0	Sports	Nintendo	41.49	29.02	3.77	8.46

In [41]:
```
df.drop([0],inplace=True)
```

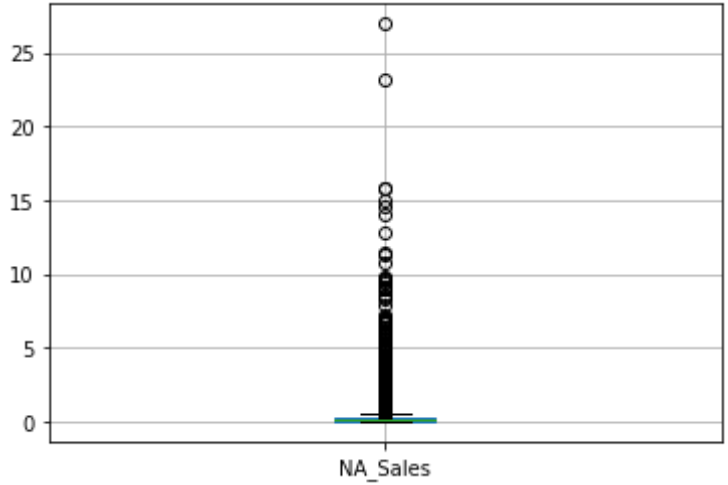
In [42]:
```
df[:5]
```

Out[42]:

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Oth
2	3	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.85	12.88	3.79	
3	4	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.75	11.01	3.28	
4	5	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	
5	6	Tetris	GB	1989.0	Puzzle	Nintendo	23.20	2.26	4.22	
6	7	New Super Mario Bros.	DS	2006.0	Platform	Nintendo	11.38	9.23	6.50	

In [43]:
```
df.boxplot(['NA_Sales'])
```

Out[43]:
```
<matplotlib.axes._subplots.AxesSubplot at 0xfd79778>
```



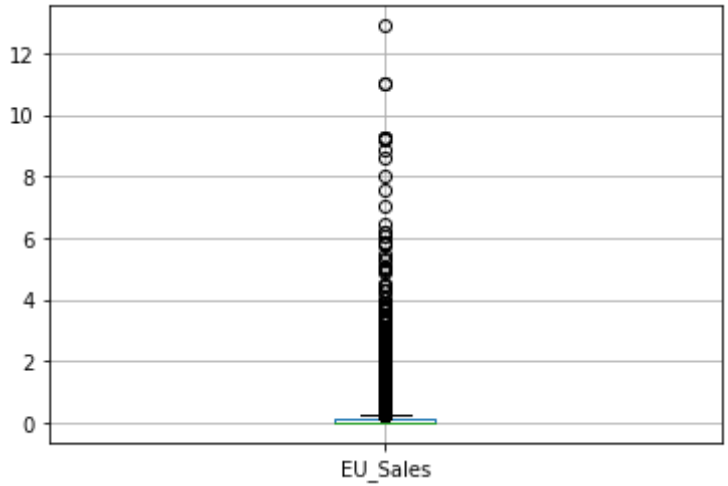
In [44]:
```
df.head(5)
```

Out[44]:

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Oth
2	3	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.85	12.88	3.79	
3	4	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.75	11.01	3.28	
4	5	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	
5	6	Tetris	GB	1989.0	Puzzle	Nintendo	23.20	2.26	4.22	
6	7	New Super Mario Bros.	DS	2006.0	Platform	Nintendo	11.38	9.23	6.50	

In [45]:
```
df.boxplot(['EU_Sales'])
```

Out[45]:
```
<matplotlib.axes._subplots.AxesSubplot at 0xffedc28>
```



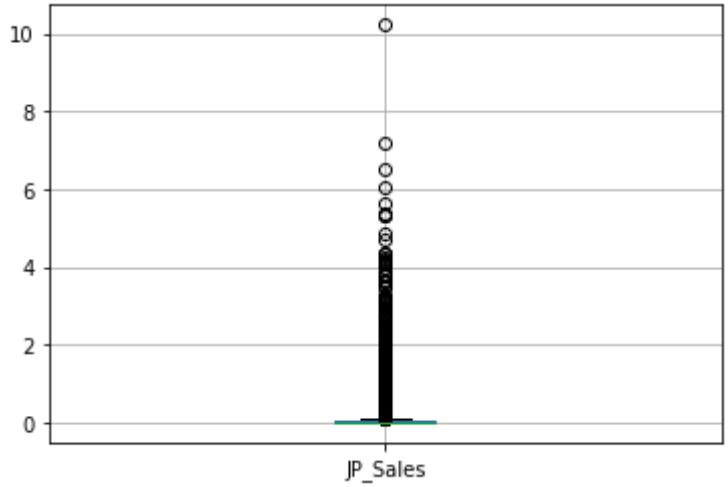
In [46]:
```
df[df.EU_Sales>12]
```

Out[46]:

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales
2	3	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.85	12.88	3.79	3.31

In [47]:
```
df.boxplot(['JP_Sales'])
```

Out[47]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x100f9fd0>
```



In [48]:
```
df[df.JP_Sales>10]
```

Out[48]:

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Othe
4	5	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	

In [49]:
```
df.drop([4],inplace=True)
```

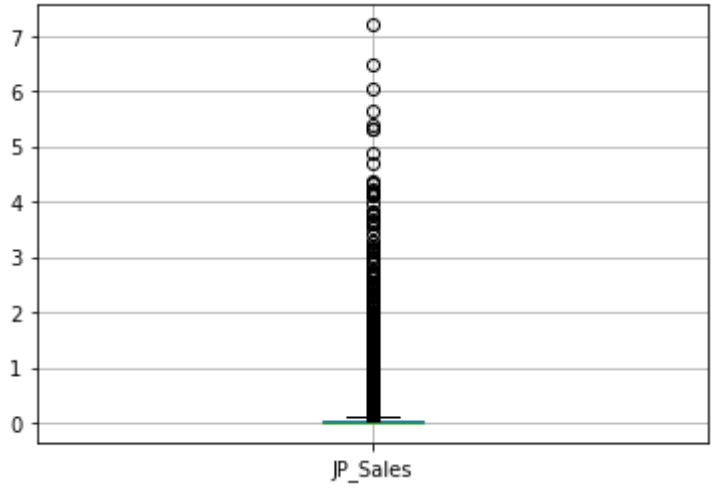
In [50]:
```
df[3:6]
```

Out[50]:

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sale
6	7	New Super Mario Bros.	DS	2006.0	Platform	Nintendo	11.38	9.23	6.50	2.9
7	8	Wii Play	Wii	2006.0	Misc	Nintendo	14.03	9.20	2.93	2.8
8	9	New Super Mario Bros. Wii	Wii	2009.0	Platform	Nintendo	14.59	7.06	4.70	2.2

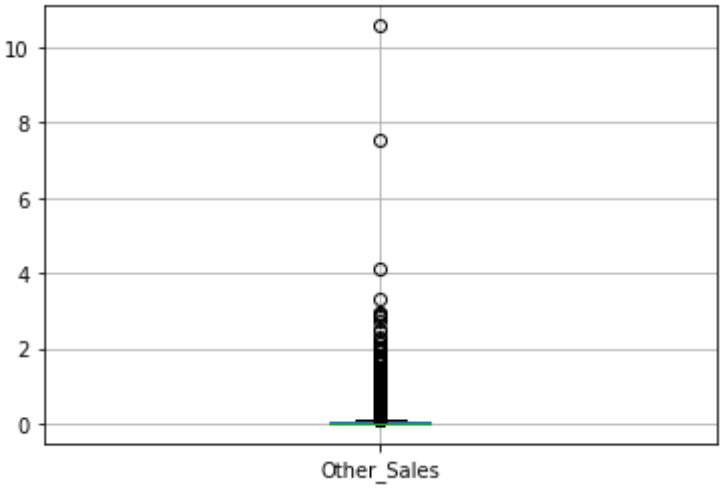
In [51]:
```
df.boxplot(['JP_Sales'])
```

Out[51]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x101259d0>
```



In [52]:
```
df.boxplot(['Other_Sales'])
```

Out[52]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x10237880>
```



In [53]:
```
df[df.Other_Sales>6]
```

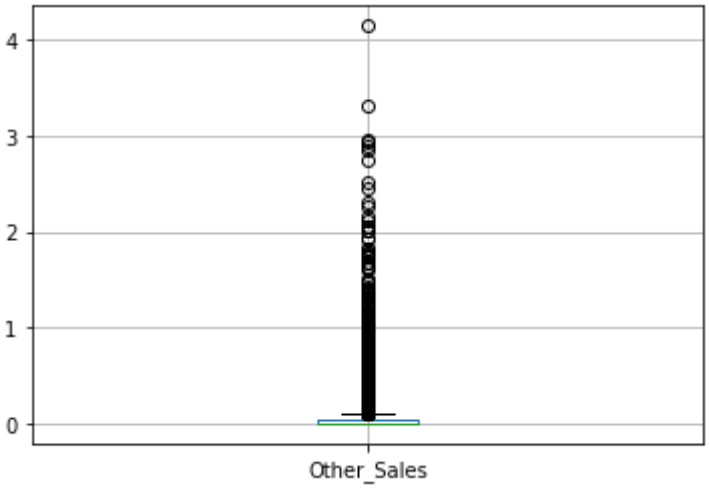
Out[53]:

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other
17	18	Grand Theft Auto: San Andreas	PS2	2004.0	Action	Take-Two Interactive	9.43	0.40	0.41	
47	48	Gran Turismo 4	PS2	2004.0	Racing	Sony Computer Entertainment	3.01	0.01	1.10	

In [54]:
```
df.drop([17,47],inplace=True)
```

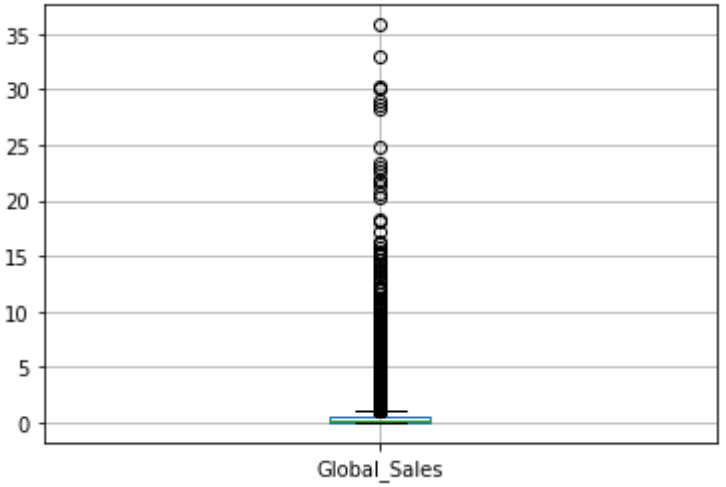
In [55]:
```
df.boxplot(['Other_Sales'])
```

Out[55]:
```
<matplotlib.axes._subplots.AxesSubplot at 0xe1f1f40>
```



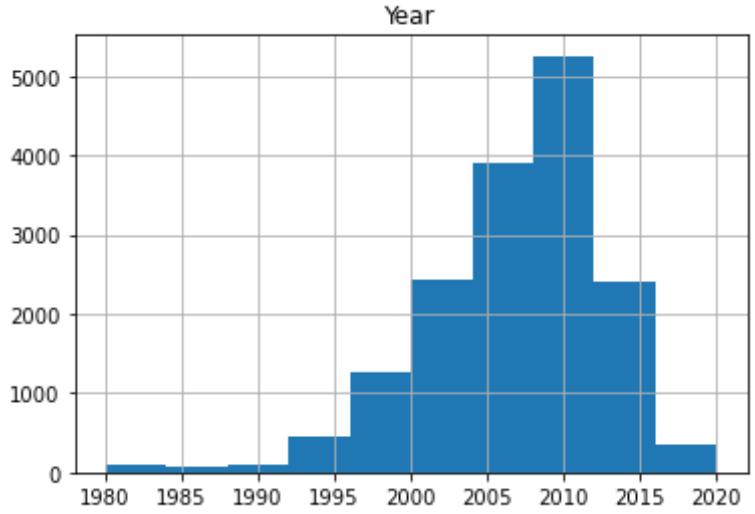
In [56]:
```
df.boxplot(['Global_Sales'])
```

Out[56]:
```
<matplotlib.axes._subplots.AxesSubplot at 0xfb76700>
```



In [59]:
```
df.hist(['Year'])
```

Out[59]:
```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x10576688>]],
      dtype=object)
```



In [62]:
```
path=('D:\Project\Game Sales\Game.csv')
df.to_csv(path,index=False)
```

In [ ]:
```

```