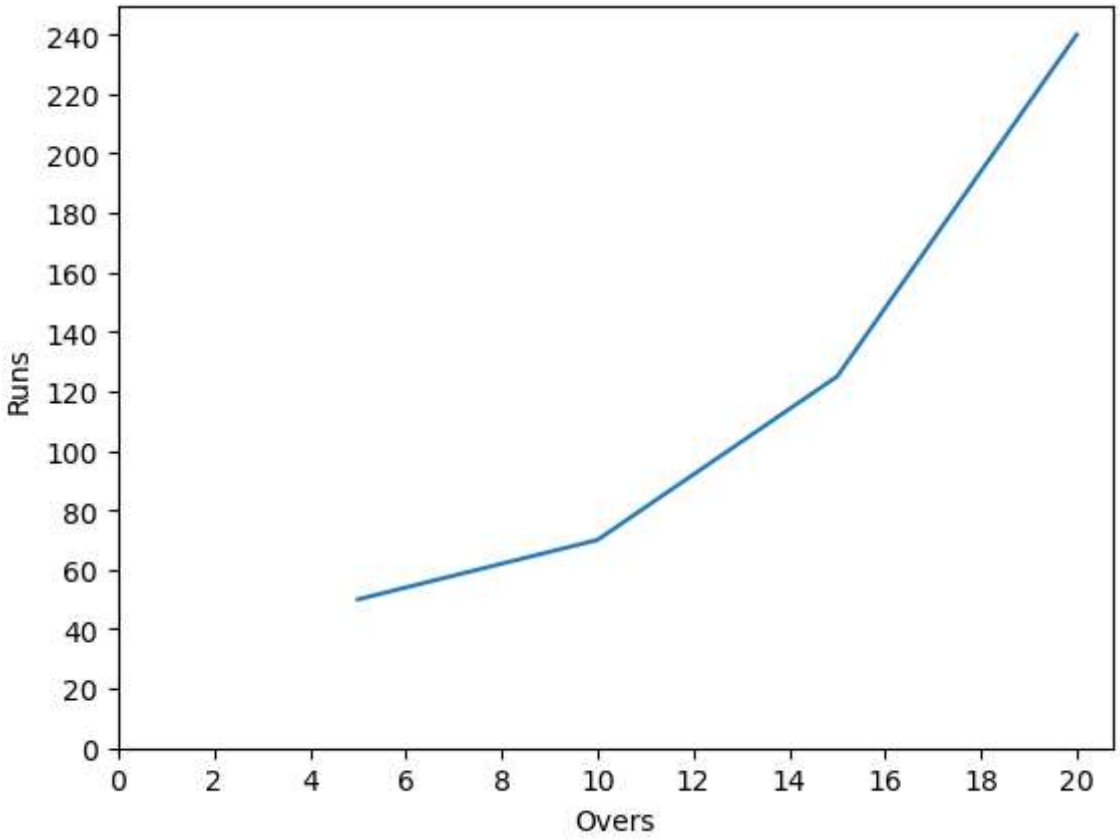
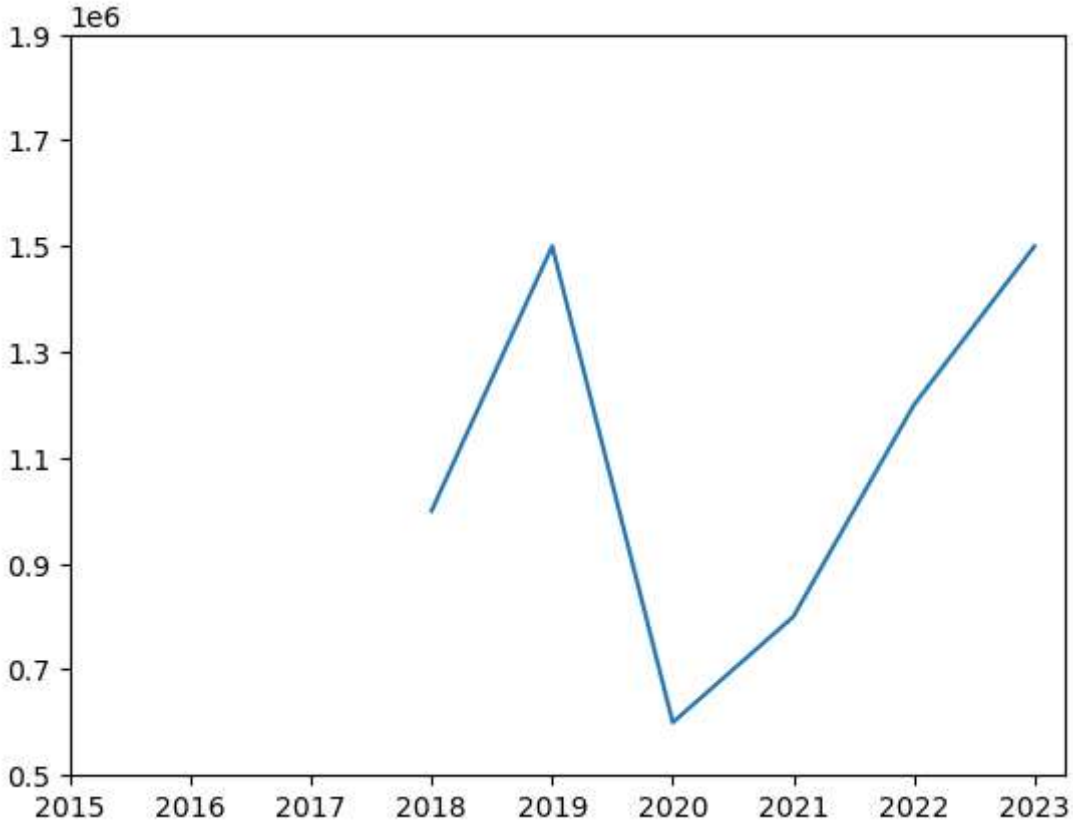
In [39]:
```
import matplotlib.pyplot as plt
import pandas as pd
```

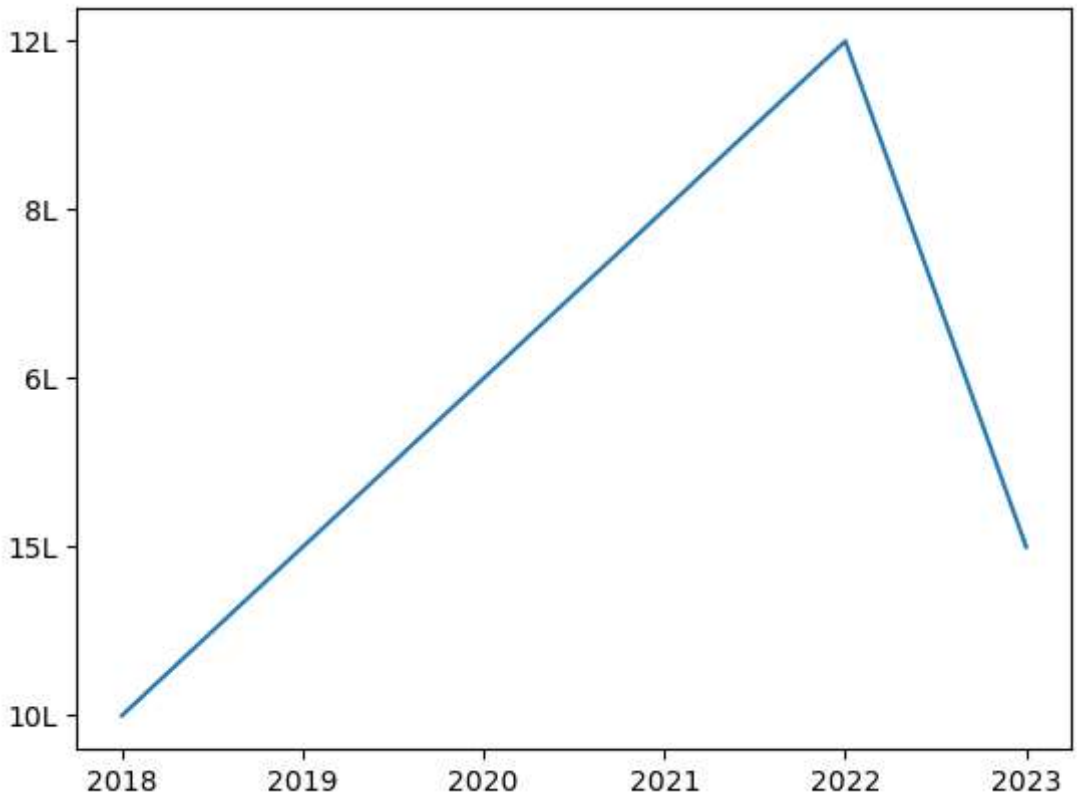
In [40]:
```
over=[5,10,15,20]
runs=[50,70,125,240]
plt.plot(over,runs)
plt.xticks(range(0,22,2))
plt.yticks(range(0,250,20))
plt.xlabel("Overs")
plt.ylabel('Runs')
plt.show()
```



In [41]:
```
years=[2018,2019,2020,2021,2022,2023]
sales=[1000000,1500000,600000,800000,1200000,1500000]
plt.plot(years,sales)
plt.xticks(range(2015,2024))
plt.yticks(range(500000,2000000,200000))
plt.show()
```

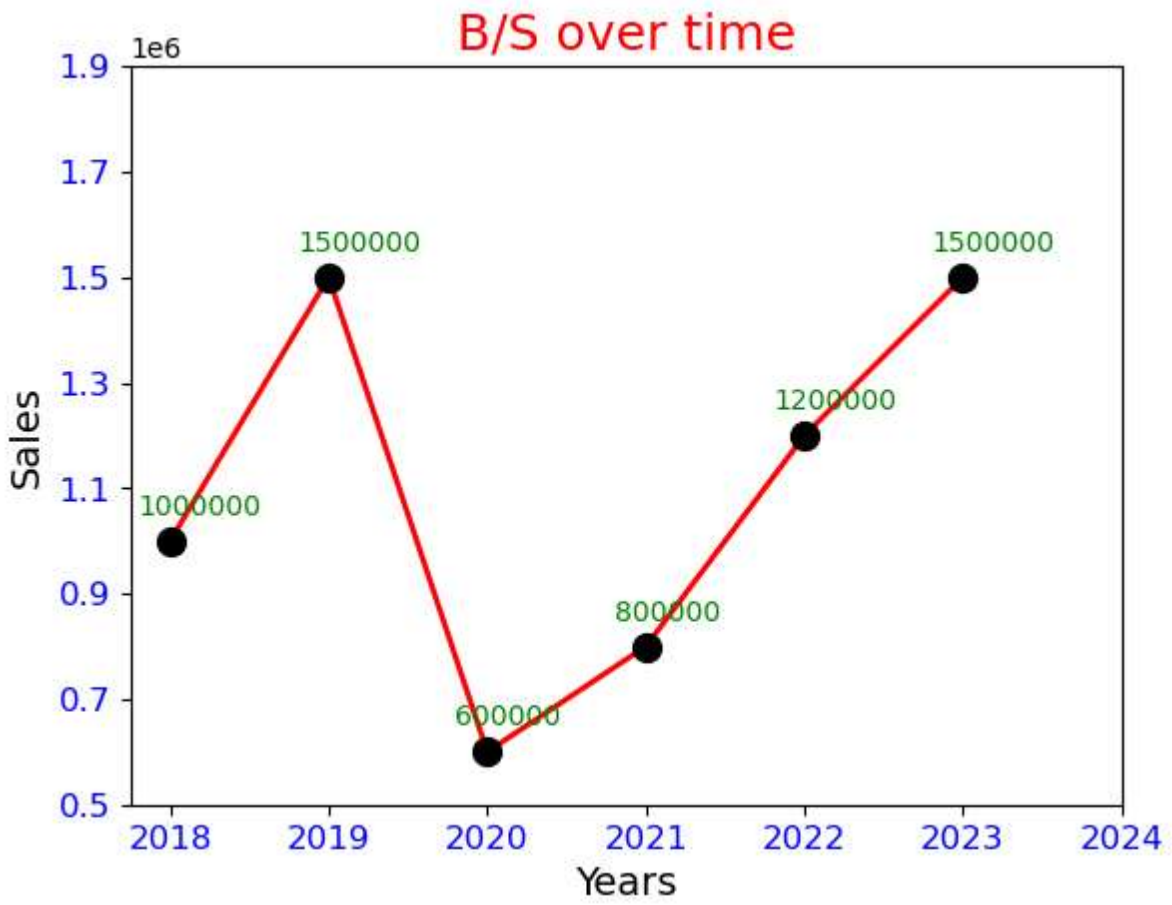


In [42]:
```
years=[2018,2019,2020,2021,2022,2023]
sales=['10L','15L','6L','8L','12L','15L']
plt.plot(years,sales)
plt.show()
```



In [43]:
```
years=[2018,2019,2020,2021,2022,2023]
sales=[1000000,1500000,600000,800000,1200000,1500000]
plt.plot(years,sales,c='r',linewidth=2,linestyle='solid',marker='o',markerfacecolor
plt.xticks(range(2018,2025,1),c='b',size=12)
plt.yticks(range(500000,2000000,200000),c='b',size=12)
```

```
plt.xlabel('Years',size=14)
plt.ylabel('Sales',size=14)
plt.title('B/S over time',c='r',size=18)
for y,s in zip(years,sales):
    plt.text(y- .2,s+50000,s,c='g',size=10)
plt.show()
```



In [44]:
```
years=[2018,2019,2020,2021,2022,2023]
sales1=[1000000,1500000,600000,800000,1200000,1500000]
sales2=[900000,1500000,1600000,1800000,2200000,2500000]
plt.plot(years,sales1,c='r',linewidth=2,linestyle='solid')
plt.plot(years,sales2,c='b',linewidth=2,linestyle='solid')
plt.xticks(range(2018,2025,1),c='b',size=12)
plt.yticks(range(500000,3000000,200000),c='b',size=12)
plt.xlabel('Years',size=14)
plt.ylabel('Sales',size=14)
plt.title('B/S over time',c='r',size=18)
plt.legend(['ABC','XYZ'])
plt.grid(c='g')
plt.show()
```

In [45]:
```
years=[2018,2019,2020,2021,2022,2023]
sales1=[1000000,1500000,600000,800000,1200000,1500000]
sales2=[900000,1500000,1600000,1800000,2200000,2500000]
plt.plot(years,sales1,c='r',linewidth=2,linestyle='solid')
plt.plot(years,sales2,c='b',linewidth=2,linestyle='solid')
plt.xticks(range(2018,2025,1),c='b',size=12)
plt.yticks(range(500000,3000000,200000),c='b',size=12)
plt.xlabel('Years',size=14)
plt.ylabel('Sales',size=14)
plt.title('B/S over time',c='r',size=18)
plt.legend(['ABC','XYZ'])
plt.grid(c='g')
plt.savefig("g:/B_S.png")
```

In [47]:
```
df=pd.read_excel('g:/dataset/Sample - Superstore.xlsx')
df
```

Out[47]:

	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country/Region
0	1	US-2020-103800	2020-01-03	2020-01-07	Standard Class	DP-13000	Darren Powers	Consumer	United States
1	2	US-2020-112326	2020-01-04	2020-01-08	Standard Class	PO-19195	Phillina Ober	Home Office	United States
2	3	US-2020-112326	2020-01-04	2020-01-08	Standard Class	PO-19195	Phillina Ober	Home Office	United States
3	4	US-2020-112326	2020-01-04	2020-01-08	Standard Class	PO-19195	Phillina Ober	Home Office	United States
4	5	US-2020-141817	2020-01-05	2020-01-12	Standard Class	MB-18085	Mick Brown	Consumer	United States
...	...	...	...	...	...	...	...	...	...
10189	10190	US-2023-143259	2023-12-30	2024-01-03	Standard Class	PO-18865	Patrick O'Donnell	Consumer	United States
10190	10191	US-2023-115427	2023-12-30	2024-01-03	Standard Class	EB-13975	Erica Bern	Corporate	United States
10191	10192	US-2023-156720	2023-12-30	2024-01-03	Standard Class	JM-15580	Jill Matthias	Consumer	United States
10192	10193	US-2023-143259	2023-12-30	2024-01-03	Standard Class	PO-18865	Patrick O'Donnell	Consumer	United States
10193	10194	CA-2023-143500	2023-12-30	2024-01-03	Standard Class	HO-15230	Harry Olson	Consumer	Canada

10194 rows × 21 columns

In [48]:
```
df.dtypes
```

Out[48]:

```
Row ID                  int64
Order ID               object
Order Date     datetime64[ns]
Ship Date      datetime64[ns]
Ship Mode              object
Customer ID            object
Customer Name          object
Segment                object
Country/Region         object
City                   object
State/Province         object
Postal Code            object
Region                 object
Product ID             object
Category               object
Sub-Category           object
Product Name           object
Sales                 float64
Quantity                int64
Discount              float64
Profit                float64
dtype: object
```

In [50]:

```
df['year']=df['Order Date'].dt.year
df
```

Out[50]:

	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country/Region
0	1	US-2020-103800	2020-01-03	2020-01-07	Standard Class	DP-13000	Darren Powers	Consumer	United States
1	2	US-2020-112326	2020-01-04	2020-01-08	Standard Class	PO-19195	Phillina Ober	Home Office	United States
2	3	US-2020-112326	2020-01-04	2020-01-08	Standard Class	PO-19195	Phillina Ober	Home Office	United States
3	4	US-2020-112326	2020-01-04	2020-01-08	Standard Class	PO-19195	Phillina Ober	Home Office	United States
4	5	US-2020-141817	2020-01-05	2020-01-12	Standard Class	MB-18085	Mick Brown	Consumer	United States
...	...	...	...	...	...	...	...	...	...
10189	10190	US-2023-143259	2023-12-30	2024-01-03	Standard Class	PO-18865	Patrick O'Donnell	Consumer	United States
10190	10191	US-2023-115427	2023-12-30	2024-01-03	Standard Class	EB-13975	Erica Bern	Corporate	United States
10191	10192	US-2023-156720	2023-12-30	2024-01-03	Standard Class	JM-15580	Jill Matthias	Consumer	United States
10192	10193	US-2023-143259	2023-12-30	2024-01-03	Standard Class	PO-18865	Patrick O'Donnell	Consumer	United States
10193	10194	CA-2023-143500	2023-12-30	2024-01-03	Standard Class	HO-15230	Harry Olson	Consumer	Canada

10194 rows × 22 columns

In [51]:
```
res_df=df.groupby(by='year')[['Sales']].sum()
```

In [52]:
```
res_df
```

Out[52]:

year	Sales
2020	494040.2121
2021	472993.0310
2022	613933.5800
2023	745567.5312

In [53]:
```
res_df.index
```

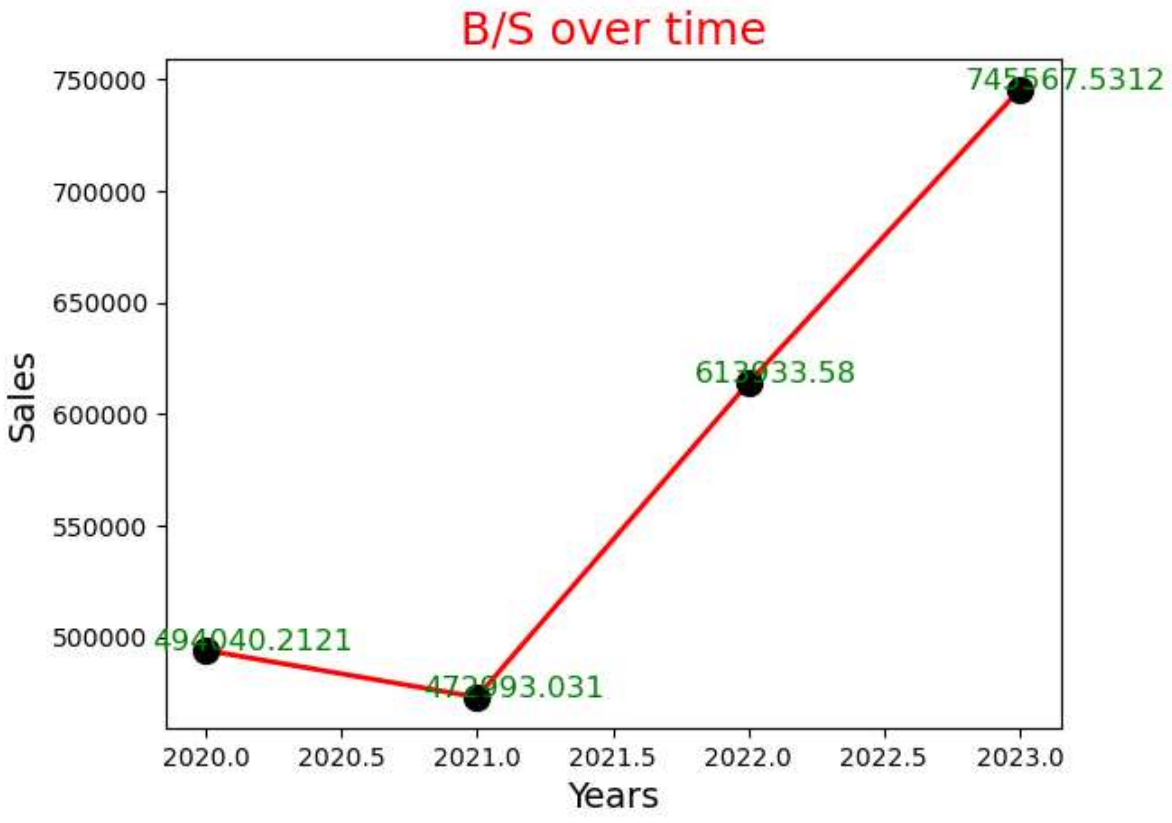
Out[53]:
```
Int64Index([2020, 2021, 2022, 2023], dtype='int64', name='year')
```

In [54]:
```
res_df.Sales.values
```

Out[54]:
```
array([494040.2121, 472993.031 , 613933.58  , 745567.5312])
```

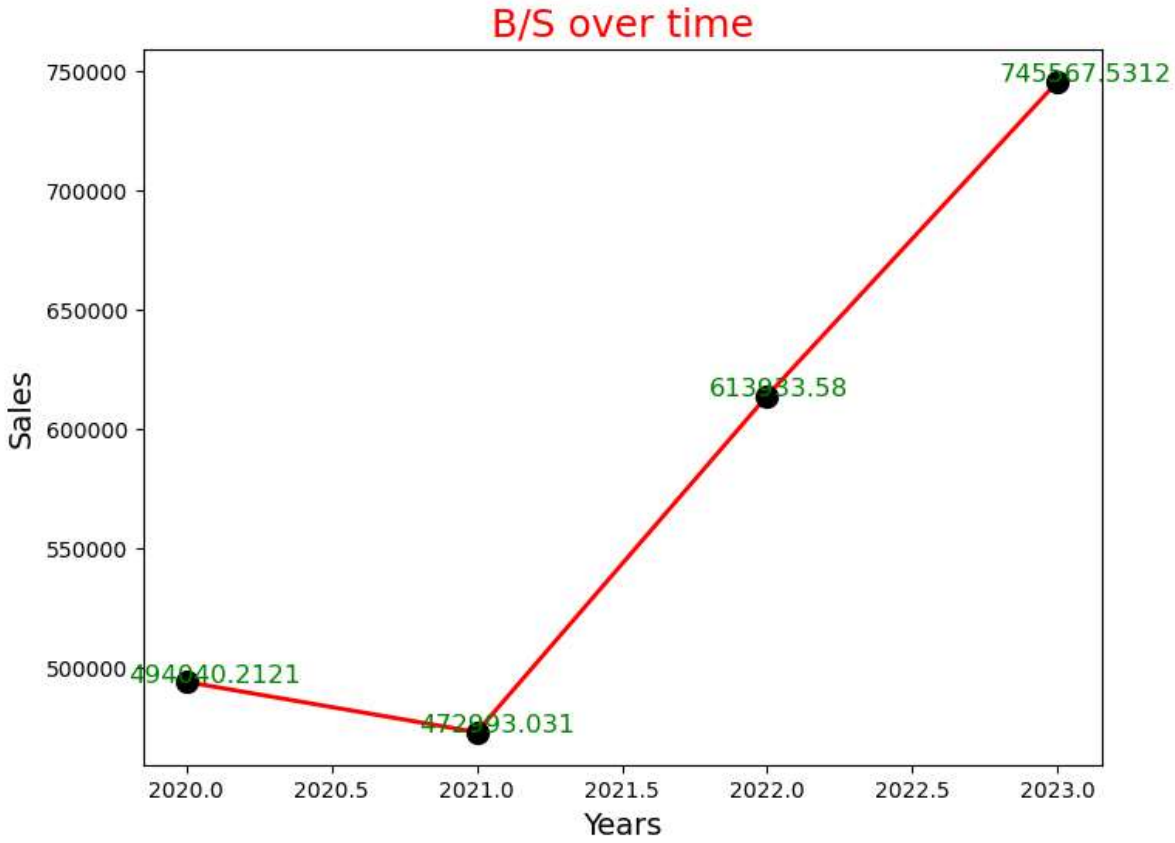
In [57]:
```
years=res_df.index
sales=res_df.Sales.values

plt.plot(years,sales,c='r',linewidth=2,linestyle='solid',marker='o',markerfacecolor
plt.xlabel('Years',size=14)
plt.ylabel('Sales',size=14)
plt.title('B/S over time',c='r',size=18)
for y,s in zip(years,sales):
    plt.text(y- .2,s,s,c='g',size=12)
plt.show()
```

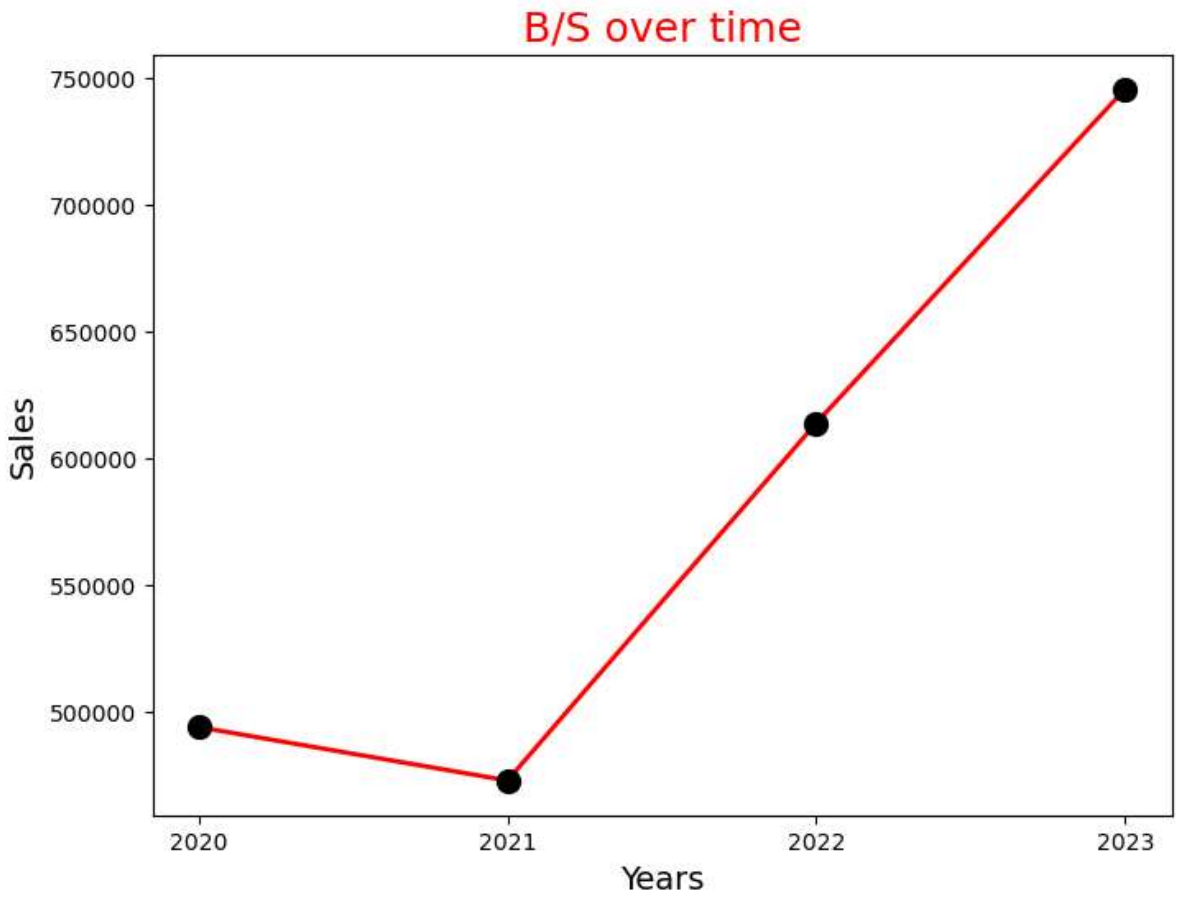


In [61]:
```
years=res_df.index
sales=res_df.Sales.values
plt.figure(figsize=(8,6))
plt.plot(years,sales,c='r',linewidth=2,linestyle='solid',marker='o',markerfacecolor
plt.xlabel('Years',size=14)
```

```
plt.ylabel('Sales',size=14)
plt.title('B/S over time',c='r',size=18)
for y,s in zip(years,sales):
    plt.text(y- .2,s,s,c='g',size=12)
plt.show()
```



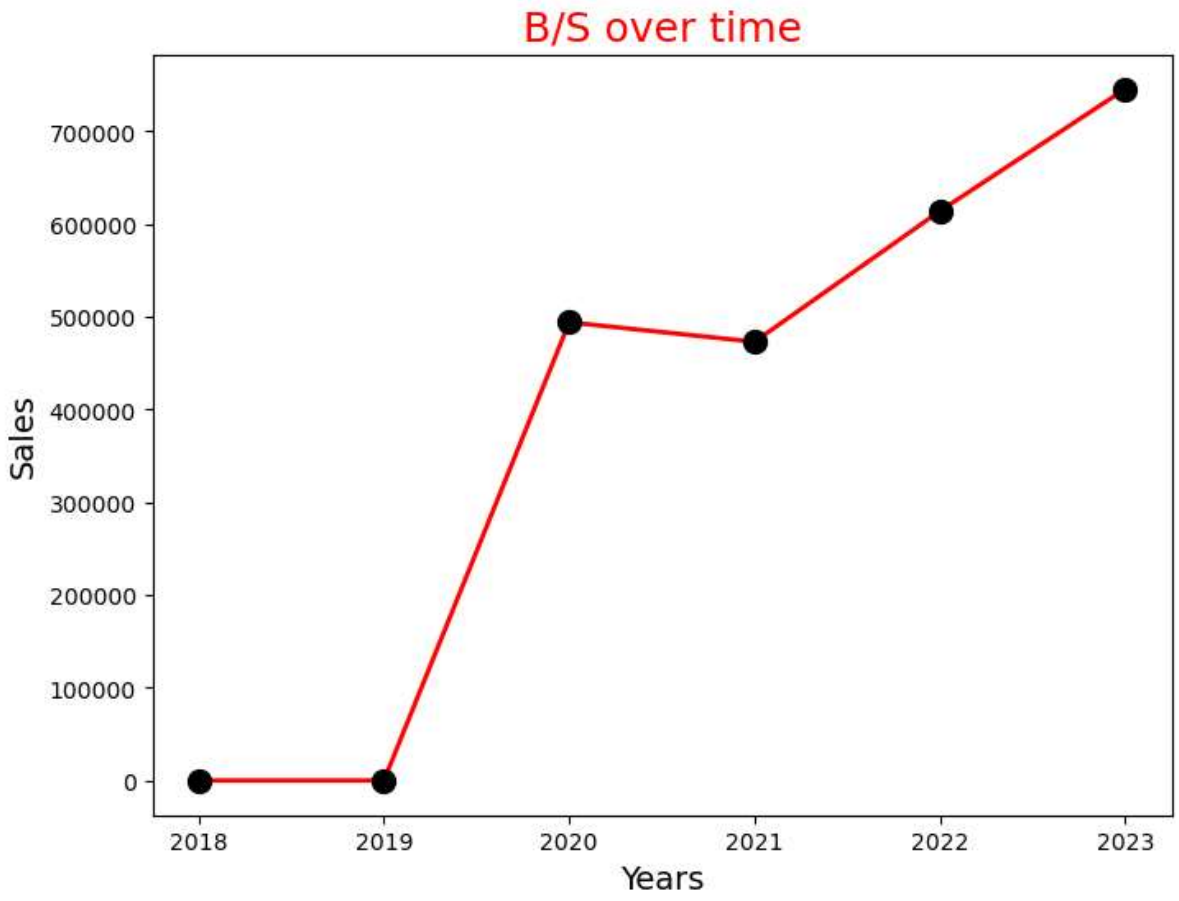
In [63]:
```
years=res_df.index
sales=res_df.Sales.values
plt.figure(figsize=(8,6))
plt.plot(years,sales,c='r',linewidth=2,linestyle='solid',marker='o',markerfacecolor
plt.xlabel('Years',size=14)
plt.ylabel('Sales',size=14)
plt.title('B/S over time',c='r',size=18)
plt.xticks(range(2020,2024,1))
plt.show()
```



In [68]:

```
import pandas as pd
df=pd.read_excel('g:/dataset/pdf_files/Sample - Superstore.xlsx')
df['year']=df['Order Date'].dt.year
res_df=df.groupby(by='year')[['Sales']].sum()

years=res_df.index
sales=res_df.Sales.values
plt.figure(figsize=(8,6))
plt.plot(years,sales,c='r',linewidth=2,linestyle='solid',marker='o',markerfacecolor
plt.xlabel('Years',size=14)
plt.ylabel('Sales',size=14)
plt.title('B/S over time',c='r',size=18)
plt.show()
```


B/S over time
Sales
Years
700000
600000
500000
400000
300000
200000
100000
0
2018
2019
2020
2021
2022
2023

In [ ]: