Importing libraries

```
In [67]: import numpy as np
         import pandas as pd
         import seaborn as sns
         import matplotlib.pyplot as plt
         %matplotlib inline
         alpha_color = 0.5
```

Loading Dataset which is about health centers distribution throughout the country

```
In [3]: data=pd.read_csv('data/Distribution-of-HF-by-region-for-the-year-2014.csv')
```

Displaying the first five rows

```
In [4]: data.head()
```

Out[4]:

	Column	Dispensary Government	Dispensary Faith Based Organizations	Dispensary Parastatal	Dispensary Private	Health Centre Government	Health Centre Faith Based Organizations	Centr
0	Arusha	157	46	7	78	27	7	
1	Dar es salaam	107	23	13	246	4	9	
2	Dodoma	274	25	3	15	30	6	
3	Geita	93	16	1	20	18	1	
4	Iringa	151	38	1	14	17	4	

```
In [ ]:
```

Finding the total of each Dispenaries and hospitals for Government and Private

```
In [5]: data['Dispensary Private'].sum()
```

Out[5]: 716

```
In [6]: data['Dispensary Government'].sum()
```

Out[6]: 4502

```
In [7]: data['Health Centre Government'].sum()
```

Out[7]: 484

```
In [8]: data['Health CentrePrivate'].sum()
```

Out[8]: 79

```
In [9]: data['Hospital Government'].sum()
```

Out[9]: 129

```
In [10]: data['Hospital Private'].sum()
```

Out[10]: 34

From the outputs above shows that, there is a large number of Government hospitals and dispensries compared to Private hospitals and dispensaries

Displaying the list of all columns contained in the dataset

```
In [12]: data.columns
```

Out[12]: Index(['Column', 'Dispensary Government',
        'Dispensary Faith Based Organizations', 'Dispensary Parastatal',
        'Dispensary Private', 'Health Centre Government',
        'Health Centre Faith Based Organizations', 'Health CentreParastatal',
        'Health CentrePrivate', 'Hospital Government',
        'Hospital Faith Based Organizations', 'Hospital Parastatal',
        'Hospital Private'],
       dtype='object')

The length of dataset

```
In [13]: len(data)
```

Out[13]: 25

The shape of dataset contains 25 rows and 13 columns

```
In [14]: data.shape
```

Out[14]: (25, 13)

Adding the new column called TOTAL and placing total of each row in the TOTAL column

```
In [61]: data['TOTAL']=data['Dispensary Government']+data['Dispensary Faith Based Organiza
```

In [62]: `data.head()`

Out[62]:

...ry ...nt	Dispensary Faith Based Organizations	Dispensary Parastatal	Dispensary Private	Health Centre Government	Health Centre Faith Based Organizations	Health CentreParastatal	Healt CentrePrivat
57	46	7	78	27	7	0	1
07	23	13	246	4	9	4	2
74	25	3	15	30	6	1	
93	16	1	20	18	1	0	
51	38	1	14	17	4	1	

◄ ▬▬▬▬▬▬▬▬▬▬▬▬▬ ►

To remove the added column called TOTAL

In [63]:
```
data.drop('TOTAL', axis=1, inplace=True)
data.head()
```

Out[63]:

Dispensary overnment	Dispensary Faith Based Organizations	Dispensary Parastatal	Dispensary Private	Health Centre Government	Health Centre Faith Based Organizations	Health CentreParastatal	Cen
157	46	7	78	27	7	0	
107	23	13	246	4	9	4	
274	25	3	15	30	6	1	
93	16	1	20	18	1	0	
151	38	1	14	17	4	1	

◄ ▬▬▬▬▬▬▬▬▬▬▬▬▬ ►

Displaying the first row to the last one

In [55]: `data[0:]`

Out[55]:

Dispensary overnment	Dispensary Faith Based Organizations	Dispensary Parastatal	Dispensary Private	Health Centre Government	Health Centre Faith Based Organizations	Health CentreParastatal	Cen
157	46	7	78	27	7	0	
107	23	13	246	4	9	4	
274	25	3	15	30	6	1	
93	16	1	20	18	1	0	
151	38	1	14	17	4	1	
212	26	4	13	21	9	0	
48	5	0	6	10	3	0	
193	16	4	9	23	4	0	
200	88	6	41	28	10	1	
177	6	0	3	16	1	0	
124	17	1	15	15	6	0	
172	24	5	21	27	8	1	
304	29	10	32	24	9	0	
229	49	8	32	30	10	2	
161	13	11	10	15	3	0	
224	17	12	51	29	7	1	
173	22	2	6	10	10	0	
205	21	6	21	18	4	0	
160	12	3	9	13	8	0	
191	33	0	9	21	7	0	
137	15	2	19	15	3	0	
157	14	1	10	12	1	0	
160	22	2	2	15	1	0	
236	30	1	6	17	4	0	
257	19	13	28	29	6	1	

The general information about the dataset

In [19]: `data.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 25 entries, 0 to 24
Data columns (total 13 columns):
Column                                     25 non-null object
Dispensary Government                      25 non-null int64
Dispensary Faith Based Organizations       25 non-null int64
Dispensary Parastatal                      25 non-null int64
Dispensary Private                         25 non-null int64
Health Centre Government                    25 non-null int64
Health Centre Faith Based Organizations    25 non-null int64
Health CentreParastatal                    25 non-null int64
Health CentrePrivate                       25 non-null int64
Hospital Government                         25 non-null int64
Hospital Faith Based Organizations         25 non-null int64
Hospital Parastatal                        25 non-null int64
Hospital Private                           25 non-null int64
dtypes: int64(12), object(1)
memory usage: 2.6+ KB
```

This shows the basic statistics for each row

In [20]: `data.describe()`

Out[20]:

	Dispensary Government	Dispensary Faith Based Organizations	Dispensary Parastatal	Dispensary Private	Health Centre Government	Health Centre Faith Based Organizations	CentrePara
count	25.00000	25.000000	25.000000	25.000000	25.000000	25.000000	25.0(
mean	180.08000	25.040000	4.640000	28.640000	19.360000	5.640000	0.4
std	57.17001	16.981559	4.290299	48.403581	7.221957	3.067029	0.9
min	48.00000	5.000000	0.000000	2.000000	4.000000	1.000000	0.0(
25%	157.00000	16.000000	1.000000	9.000000	15.000000	3.000000	0.0(
50%	173.00000	22.000000	3.000000	15.000000	18.000000	6.000000	0.0(
75%	212.00000	29.000000	7.000000	28.000000	27.000000	8.000000	1.0(
max	304.00000	88.000000	13.000000	246.000000	30.000000	10.000000	4.0(

The data of Arusha region

```
In [21]: data[data["Column"] == "Arusha"].mean()
```

```
Out[21]: Dispensary Government                              157.0
         Dispensary Faith Based Organizations               46.0
         Dispensary Parastatal                               7.0
         Dispensary Private                                 78.0
         Health Centre Government                            27.0
         Health Centre Faith Based Organizations             7.0
         Health CentreParastatal                             0.0
         Health CentrePrivate                               12.0
         Hospital Government                                 8.0
         Hospital Faith Based Organizations                  5.0
         Hospital Parastatal                                 0.0
         Hospital Private                                    1.0
         dtype: float64
```

The region with minimum number of Dispensaries and Hospitals

```
In [64]: data.apply(np.min)
```

```
Out[64]: Column                                           Arusha
         Dispensary Government                               48
         Dispensary Faith Based Organizations                5
         Dispensary Parastatal                               0
         Dispensary Private                                  2
         Health Centre Government                            4
         Health Centre Faith Based Organizations             1
         Health CentreParastatal                             0
         Health CentrePrivate                                0
         Hospital Government                                 1
         Hospital Faith Based Organizations                  0
         Hospital Parastatal                                 0
         Hospital Private                                    0
         dtype: object
```

The region with maximun number of Dispensaries and Hospitals

```
In [65]: data.apply(np.max)
```

```
Out[65]: Column                                            Tanga
         Dispensary Government                              304
         Dispensary Faith Based Organizations               88
         Dispensary Parastatal                              13
         Dispensary Private                                246
         Health Centre Government                           30
         Health Centre Faith Based Organizations            10
         Health CentreParastatal                             4
         Health CentrePrivate                               21
         Hospital Government                                11
         Hospital Faith Based Organizations                  8
         Hospital Parastatal                                 4
         Hospital Private                                    20
         dtype: object
```

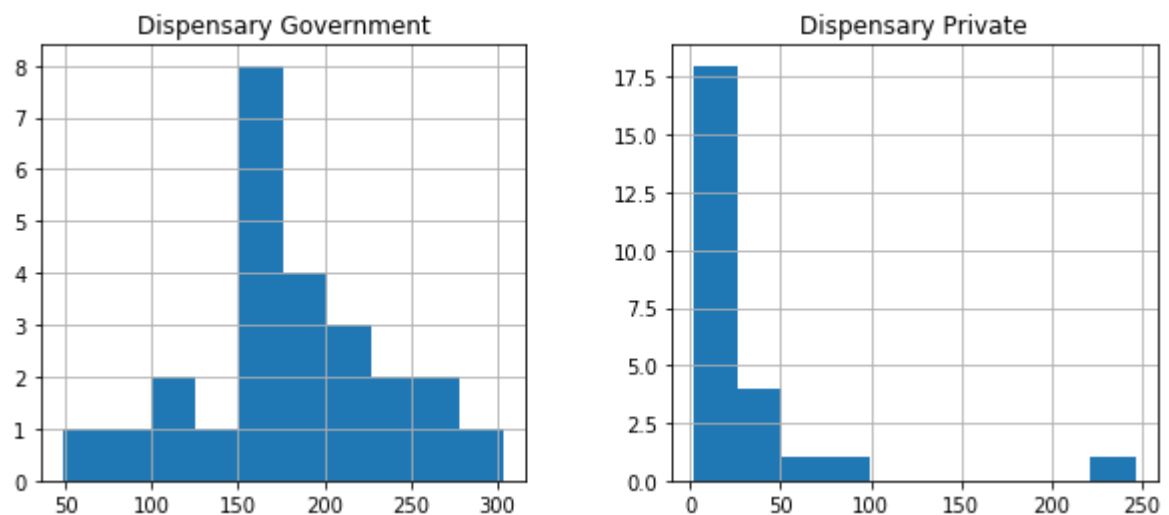**The proportion of Hospital Faith Based Organizations in the dataset**

In [24]: `data['Hospital Faith Based Organizations'].value_counts(normalize=True)`

Out[24]:
```
1    0.24
5    0.20
3    0.12
0    0.12
6    0.08
4    0.08
2    0.08
8    0.04
7    0.04
Name: Hospital Faith Based Organizations, dtype: float64
```
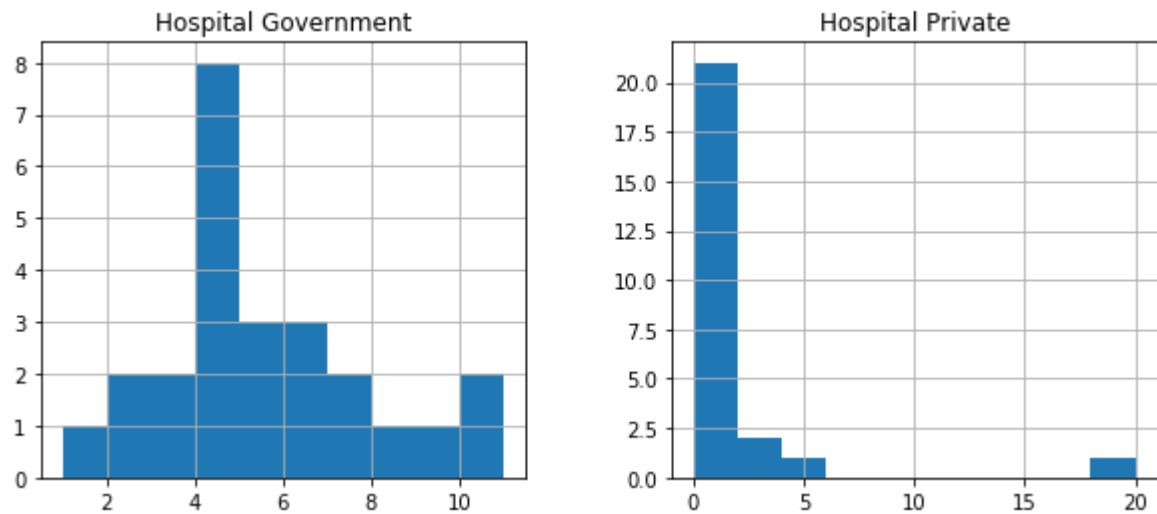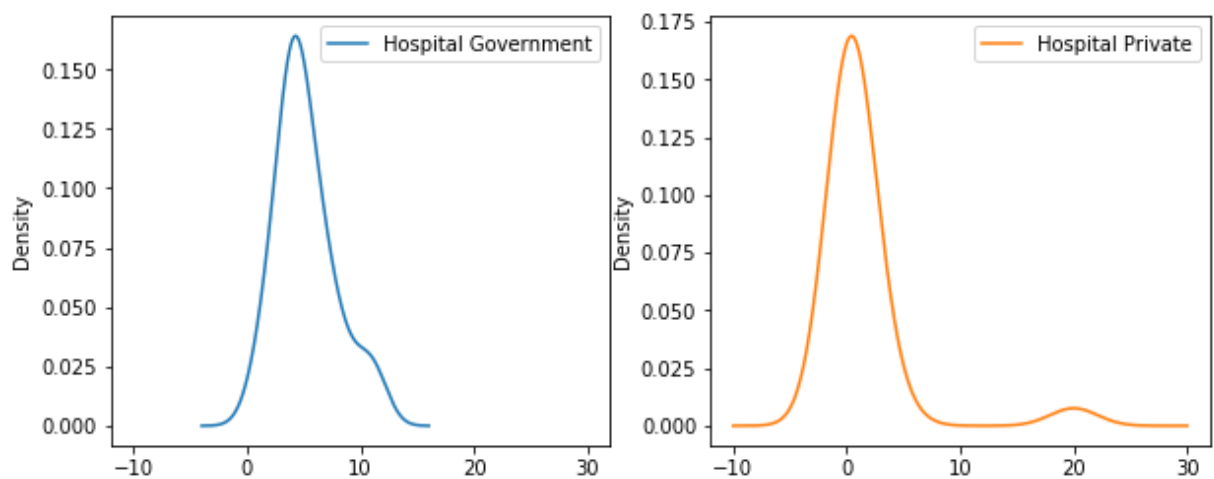
Visualization

In [25]:
```
features = ['Dispensary Government', 'Dispensary Private']
data[features].hist(figsize=(10, 4));
```
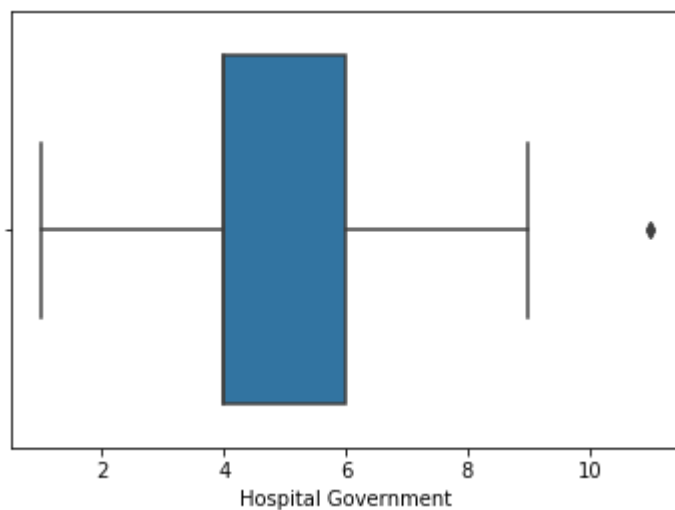
```
In [26]: features = ['Hospital Government', 'Hospital Private']
         data[features].hist(figsize=(10, 4));
```



```
In [27]: data[features].plot(kind='density', subplots=True, layout=(1, 2), figsize=(10, 4)
```



```
In [28]: sns.boxplot(x='Hospital Government', data=data);
```

In [29]: ```
sns.boxplot(x='Hospital Private', data=data);
```



In [30]: ```
sns.boxplot(x='Dispensary Government', data=data);
```
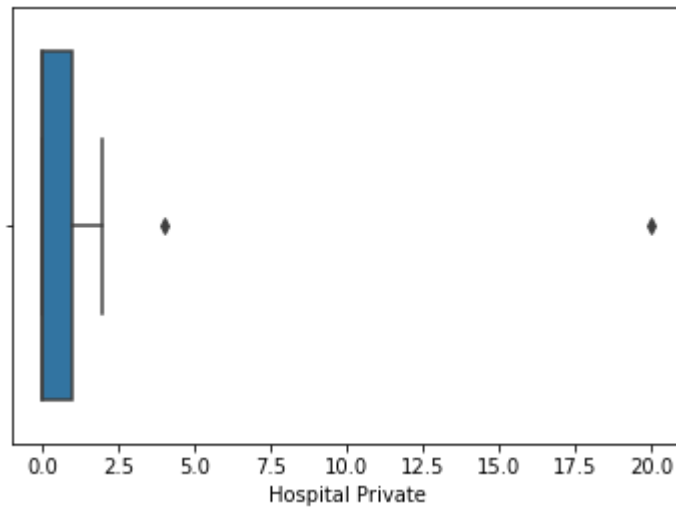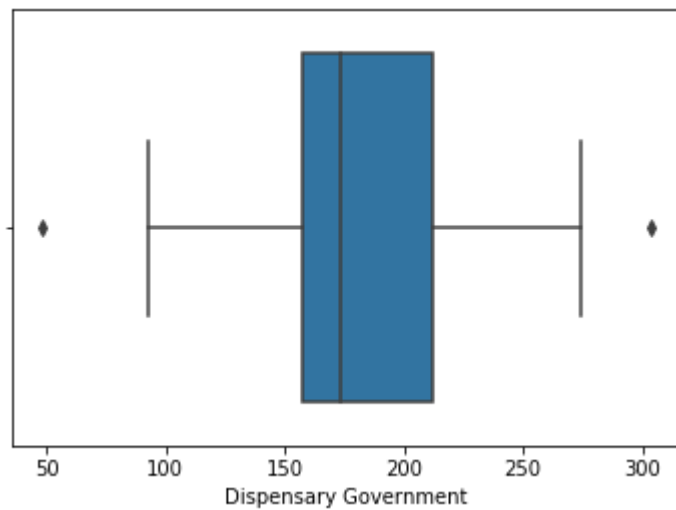

In [31]: ```
sns.boxplot(x='Dispensary Private', data=data);
```

In [32]:
```python
_, axes = plt.subplots(1, 2, sharey=True, figsize=(6, 4))
sns.boxplot(data=data['Hospital Private'], ax=axes[0]);
sns.violinplot(data=data['Hospital Government'], ax=axes[1]);
```
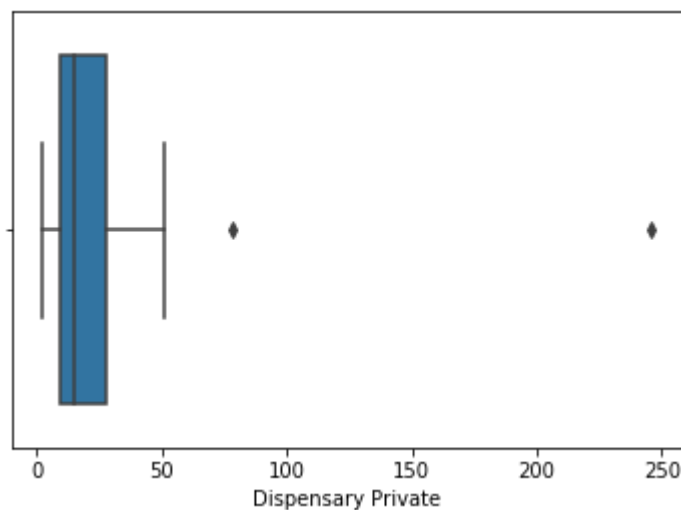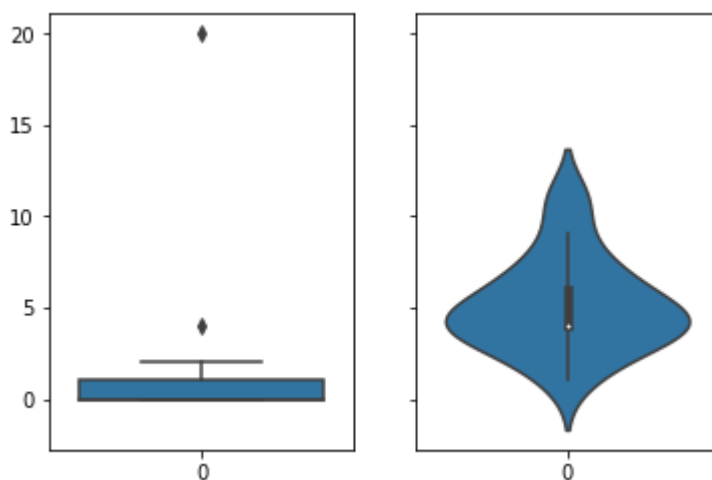

In [33]:
```python
sns.countplot(x='Hospital Private', data=data);
```
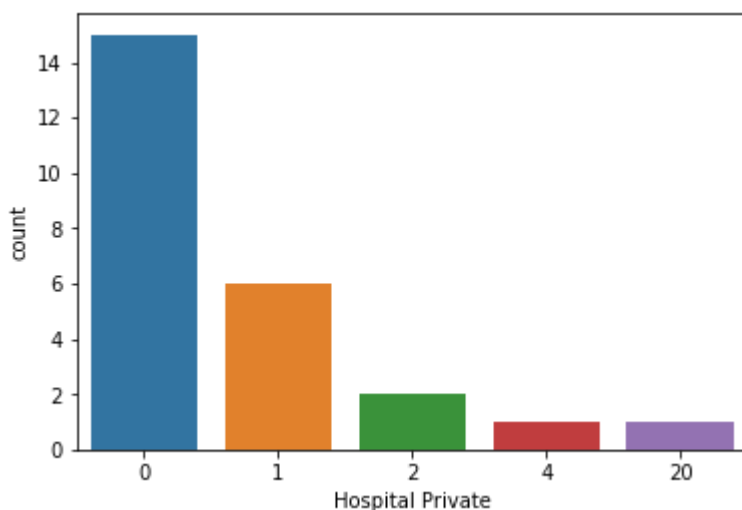

In [52]:
```python
sns.countplot(x='Hospital Government', data=data);
```
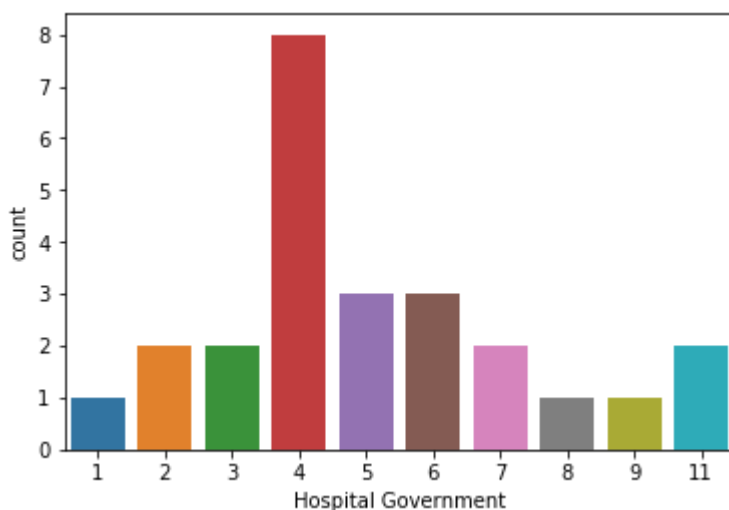
In [34]: `data[["Hospital Government", "Hospital Private"]].plot()`

Out[34]: `<matplotlib.axes._subplots.AxesSubplot at 0x2361e9878d0>`
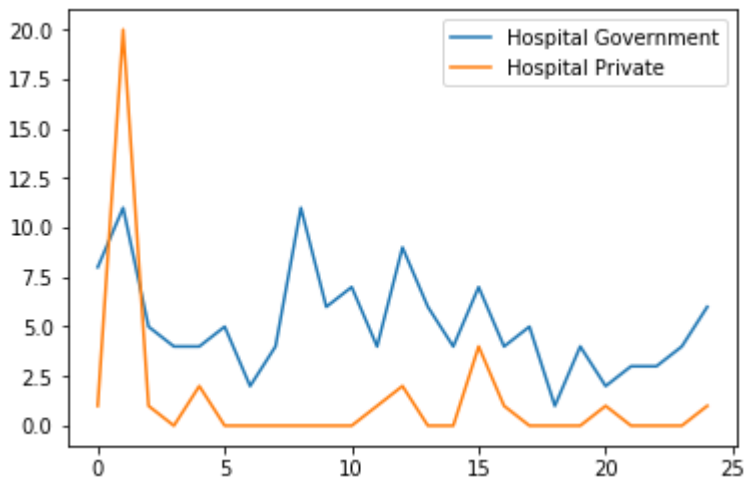
In [35]: `data[["Dispensary Government", "Dispensary Private", "Hospital Private", "Hospita`

Out[35]: `<matplotlib.axes._subplots.AxesSubplot at 0x2361e9e2e80>`
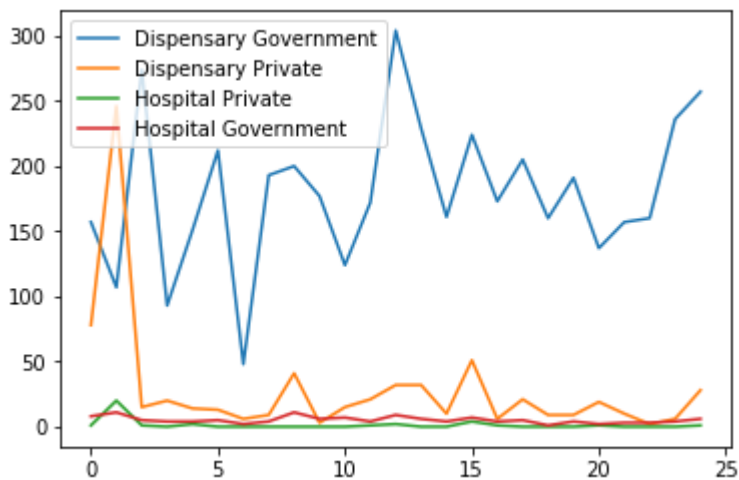
In [41]: `plt.scatter(data['Dispensary Government'], data['Dispensary Private']);`
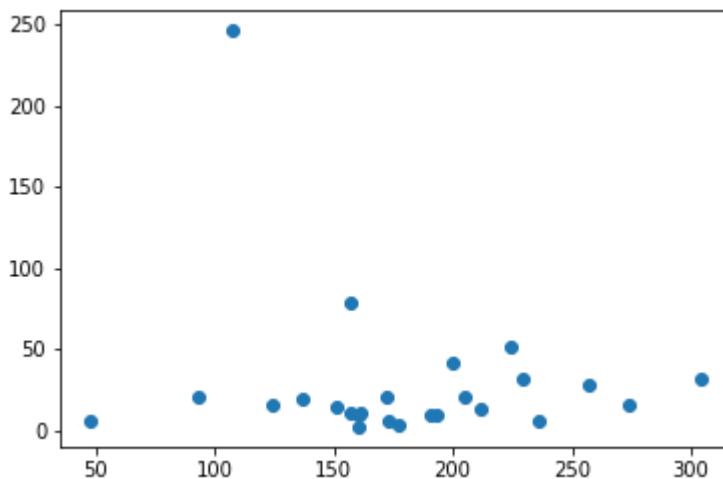
In [75]: `data['Hospital Private'].value_counts().sort_index().plot(kind='bar', alpha=alpha`
`plt.title("The graph of Hospital Private")`

Out[75]: `Text(0.5,1,'The graph of Hospital Private')`
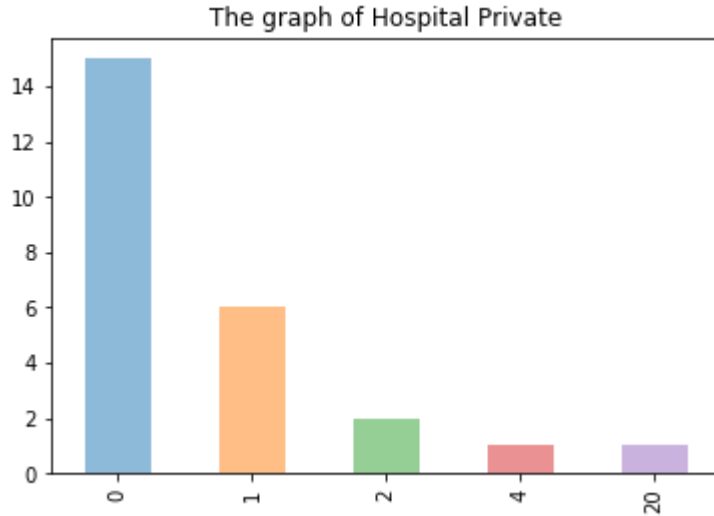
In [74]: `data['Hospital Government'].value_counts().sort_index().plot(kind='bar', alpha=al`
`plt.title("The graph of Hospital Government")`

Out[74]: `Text(0.5,1,'The graph of Hospital Government')`

In [ ]: