It is believed that discrimination in academia has been decreased over time. However, gender discrimination is truly a subtle form of prejudice that has a striking impact on all aspects of life. The level of gender discrimination has been analyzed in the five different departments of the Houston College of Medicine. Female doctors claimed that the College has engaged in a pattern and practice of discrimination against women in giving promotions and setting salaries. Data explorations helps to reveal hidden patterns in raw data and well understand the relation between features. In terms of different departments involved in this claim, analyzing each department is indispensable to perceive the level and type of gender gap in medical school.

In [1]:
```
import numpy as np
import pandas as pd

import seaborn as sns;sns.set()
from sklearn. model_selection import train_test_split

import matplotlib.pyplot as plt
from sklearn.model_selection import cross_val_score
```

In [2]:
```
Dis=pd.read_excel('C:\\Users\\noosh\\Documents\\Custom Office Templates\\Lawsuit.xlt')
```

In [3]:
```
Dis.drop('ID',1,inplace=True)
```

General outlook of data

In [4]:
```
pd.set_option('display.max_rows',500)
```

In [5]:
```
Dis.head()
```

Out[5]:

| | Dept | Gender | Clin | Cert | Prate | Exper | Rank | Sal94 | Sal95 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 0 | 0 | 7.4 | 9 | 3 | 77836 | 84612 |
| 1 | 1 | 1 | 0 | 0 | 6.7 | 10 | 2 | 69994 | 78497 |
| 2 | 1 | 1 | 0 | 0 | 8.1 | 6 | 1 | 62872 | 67756 |
| 3 | 1 | 1 | 1 | 1 | 5.1 | 27 | 3 | 155196 | 173220 |
| 4 | 1 | 1 | 0 | 0 | 7.0 | 10 | 3 | 89268 | 96099 |

In [6]:
```
Dis.tail()
```

Out[6]:

| | Dept | Gender | Clin | Cert | Prate | Exper | Rank | Sal94 | Sal95 |
|---|---|---|---|---|---|---|---|---|---|
| 256 | 6 | 0 | 1 | 1 | 2.9 | 8 | 2 | 297174 | 323699 |
| 257 | 6 | 0 | 1 | 1 | 2.6 | 9 | 3 | 308081 | 339664 |
| 258 | 6 | 0 | 1 | 1 | 2.8 | 8 | 1 | 211269 | 241577 |
| 259 | 6 | 0 | 1 | 1 | 2.4 | 2 | 1 | 210801 | 233905 |
| 260 | 6 | 0 | 1 | 1 | 3.2 | 5 | 1 | 244551 | 265502 |

In [7]:
```
Dis.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 261 entries, 0 to 260
Data columns (total 9 columns):
Dept      261 non-null int64
Gender    261 non-null int64
Clin      261 non-null int64
Cert      261 non-null int64
Prate     261 non-null float64
Exper     261 non-null int64
Rank      261 non-null int64
Sal94     261 non-null int64
Sal95     261 non-null int64
dtypes: float64(1), int64(8)
memory usage: 18.5 KB
```

Data Cleaning

In [8]:
```
Dum = pd.get_dummies(Dis. Clin, prefix='Clin').iloc[:]
Dis1= pd.concat([Dis,Dum],axis=1)
Dum1= pd.get_dummies(Dis1.Cert , prefix='Cert').iloc[:]
Dis2= pd.concat([Dis1,Dum1],axis=1)
Dum2= pd.get_dummies(Dis2.Rank, prefix='Rank').iloc[:]
DIS = pd.concat([Dis2,Dum2],axis=1)
```

In [9]:
```
DIS.head()
```

Out[9]:

| | Dept | Gender | Clin | Cert | Prate | Exper | Rank | Sal94 | Sal95 | Clin_0 | Clin_1 | Cert_0 | Cert_1 | Rank_1 | Rank_2 | Rank_3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 0 | 0 | 7.4 | 9 | 3 | 77836 | 84612 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| 1 | 1 | 1 | 0 | 0 | 6.7 | 10 | 2 | 69994 | 78497 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 2 | 1 | 1 | 0 | 0 | 8.1 | 6 | 1 | 62872 | 67756 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 3 | 1 | 1 | 1 | 1 | 5.1 | 27 | 3 | 155196 | 173220 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| 4 | 1 | 1 | 0 | 0 | 7.0 | 10 | 3 | 89268 | 96099 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |

In [10]:
```
DIS.rename(columns={'Clin_0':'Primarily_research_emphasis','Clin_1':'Primarily_clinical_emphasis','Cert_0':'not_certified','Cert_
```

In [11]:
```
DIS.drop(['Clin','Cert','Rank'],1,inplace=True)
```

In [12]:
```
DIS.head()
```

Out[12]:

| | Dept | Gender | Prate | Exper | Sal94 | Sal95 | Primarily_research_emphasis | Primarily_clinical_emphasis | not_certified | Board_certified | Assistant | Associate | fu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 7.4 | 9 | 77836 | 84612 | 1 | 0 | 1 | 0 | 0 | 0 | |
| 1 | 1 | 1 | 6.7 | 10 | 69994 | 78497 | 1 | 0 | 1 | 0 | 0 | 1 | |
| 2 | 1 | 1 | 8.1 | 6 | 62872 | 67756 | 1 | 0 | 1 | 0 | 1 | 0 | |
| 3 | 1 | 1 | 5.1 | 27 | 155196 | 173220 | 0 | 1 | 0 | 1 | 0 | 0 | |
| 4 | 1 | 1 | 7.0 | 10 | 89268 | 96099 | 1 | 0 | 1 | 0 | 0 | 0 | |

In [13]:
```
DIS.dtypes
```

Out[13]:
```
Dept                          int64
Gender                        int64
Prate                       float64
Exper                         int64
Sal94                         int64
Sal95                         int64
Primarily_research_emphasis   uint8
Primarily_clinical_emphasis   uint8
not_certified                 uint8
Board_certified               uint8
Assistant                     uint8
Associate                     uint8
full_professor                uint8
dtype: object
```

In [14]:
```
DIS=DIS.replace(',.','',regex=True).astype('float')
```

In [15]:
```
DIS.isnull().sum().reset_index()
```

Out[15]:

| | index | 0 |
|---|---|---|
| 0 | Dept | 0 |
| 1 | Gender | 0 |
| 2 | Prate | 0 |
| 3 | Exper | 0 |
| 4 | Sal94 | 0 |
| 5 | Sal95 | 0 |
| 6 | Primarily_research_emphasis | 0 |
| 7 | Primarily_clinical_emphasis | 0 |
| 8 | not_certified | 0 |
| 9 | Board_certified | 0 |
| 10 | Assistant | 0 |
| 11 | Associate | 0 |
| 12 | full_professor | 0 |

In [16]:
```
DIS[DIS['Exper']<30]
```

Out[16]:

| | Dept | Gender | Prate | Exper | Sal94 | Sal95 | Primarily_research_emphasis | Primarily_clinical_emphasis | not_certified | Board_certified | Assistant | Ass |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1.0 | 1.0 | 7.4 | 9.0 | 77836.0 | 84612.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | |
| 1 | 1.0 | 1.0 | 6.7 | 10.0 | 69994.0 | 78497.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | |
| 2 | 1.0 | 1.0 | 8.1 | 6.0 | 62872.0 | 67756.0 | 1.0 | 0.0 | 1.0 | 0.0 | 1.0 | |
| 3 | 1.0 | 1.0 | 5.1 | 27.0 | 155196.0 | 173220.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | |
| 4 | 1.0 | 1.0 | 7.0 | 10.0 | 89268.0 | 96099.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | |
| 5 | 1.0 | 1.0 | 7.7 | 10.0 | 79714.0 | 87531.0 | 1.0 | 0.0 | 0.0 | 1.0 | 0.0 | |
| 6 | 1.0 | 1.0 | 8.3 | 9.0 | 89781.0 | 99972.0 | 1.0 | 0.0 | 0.0 | 1.0 | 0.0 | |
| 7 | 1.0 | 1.0 | 5.3 | 11.0 | 151423.0 | 166601.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | |
| 8 | 1.0 | 1.0 | 7.7 | 18.0 | 81271.0 | 85437.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | |
| 9 | 1.0 | 1.0 | 7.5 | 9.0 | 83018.0 | 91369.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | |

1=Dept **Biochemistry/Molecular Biology** ,**Physiology** ,**Genetics** ,**Pediatrics** ,**Medicine** ,**Surgery**

2 Gender **1=Male, 0=Female**

3 Clin **1=Primarily clinical emphasis, 0=Primarily research emphasis**

4 Cert **1=Board certified, 0=not certified**

5 **Prate Publication rate** (# publications on cv)/(# years between CV date and MD date)

6 **Exper** # years since obtaining MD

7 Rank **1=Assistant, 2=Associate, 3=Full professor** (a proxy for productivity)

8 Sal94 Salary in academic year 1994

9 Sal95 Salary after increment to 1994

**Biochemistry**

In [17]:
```
Biochemistry=DIS.loc[1:49,]
```

In [18]:
```
Biochemistry[Biochemistry['Exper']<30].head()
```

Out[18]:

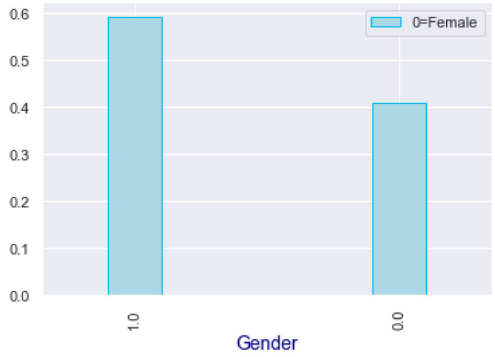
| | Dept | Gender | Prate | Exper | Sal94 | Sal95 | Primarily_research_emphasis | Primarily_clinical_emphasis | not_certified | Board_certified | Assistant | Associate |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.0 | 1.0 | 6.7 | 10.0 | 69994.0 | 78497.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | 1.0 |
| 2 | 1.0 | 1.0 | 8.1 | 6.0 | 62872.0 | 67756.0 | 1.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 |
| 3 | 1.0 | 1.0 | 5.1 | 27.0 | 155196.0 | 173220.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 4 | 1.0 | 1.0 | 7.0 | 10.0 | 89268.0 | 96099.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| 5 | 1.0 | 1.0 | 7.7 | 10.0 | 79714.0 | 87531.0 | 1.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |

In [19]:
```
Biochemistry.Gender.value_counts(normalize=True)
```

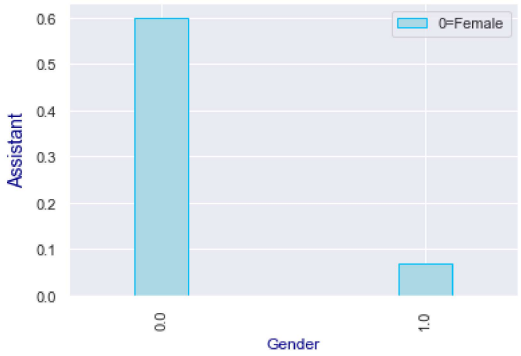
Out[19]:
```
1.0    0.591837
0.0    0.408163
Name: Gender, dtype: float64
```

In [20]:
```
Biochemistry.Gender.value_counts(normalize=True).plot(kind='bar',width=0.2, color='lightblue',edgecolor='deepskyblue',figsize=(6,
plt.xlabel('Gender',size=14,color='darkblue')
plt.legend(['0=Female'])
```

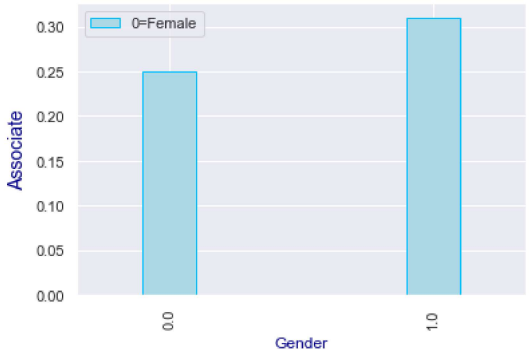
Out[20]:
```
<matplotlib.legend.Legend at 0x1e22e3a2f48>
```



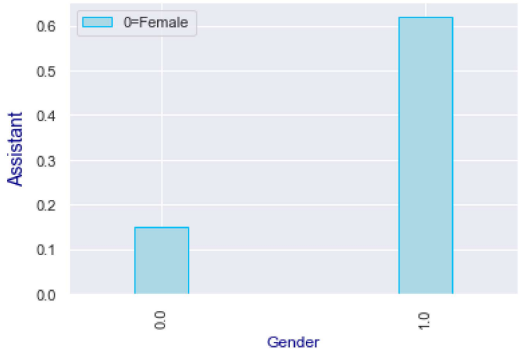

In [21]:
```
Biochemistry.groupby('Gender').Assistant.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deepskyblue',figsize=(6,4
plt.xlabel('Gender',size=12,labelpad=6,color='darkblue')
plt.ylabel('Assistant', size=14,labelpad=6,color='darkblue')
plt.legend(['0=Female'])
plt.show()
```



In [22]:
```
Biochemistry.groupby('Gender').Associate.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deepskyblue',figsize=(6,4
plt.xlabel('Gender',size=12,labelpad=6,color='darkblue')
plt.ylabel('Associate', size=14,labelpad=6,color='darkblue')
plt.legend(['0=Female'])
plt.show()
```

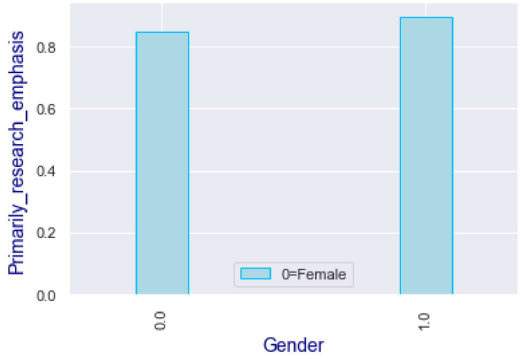



In [23]:
```
Biochemistry.groupby('Gender').full_professor.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deepskyblue',figsize
plt.xlabel('Gender',size=12,labelpad=6,color='darkblue')
plt.ylabel('Assistant', size=14,labelpad=6,color='darkblue')
plt.legend(['0=Female'])
plt.show()
```

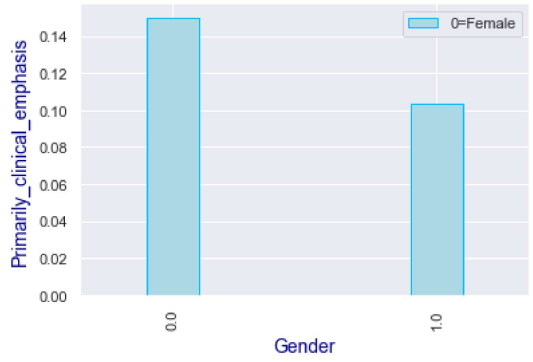



Funders of medical research the world over are increasingly seeking, in research assessment, to complement traditional output measures of scientific publications with more outcome-based indicators of societal and economic impact

In [24]:
```
Biochemistry.groupby('Gender').Primarily_research_emphasis.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deepsky
plt.xlabel('Gender',size=14,labelpad=6,color='darkblue')
plt.ylabel('Primarily_research_emphasis',size=14,labelpad=6,color='darkblue')
plt.legend(['0=Female'])
plt.show()
```



In [25]:
```
Biochemistry.groupby('Gender').Primarily_clinical_emphasis.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deepsky
plt.xlabel('Gender',size=14,labelpad=6,color='darkblue')
plt.ylabel('Primarily_clinical_emphasis',size=14,labelpad=6,color='darkblue')
plt.legend(['0=Female'])
plt.show()
```

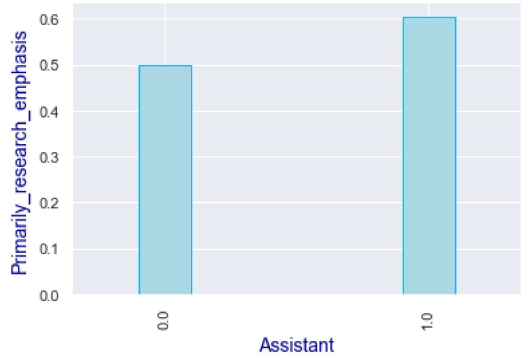



In [26]:
```
Biochemistry.Assistant.value_counts(normalize=True)
```

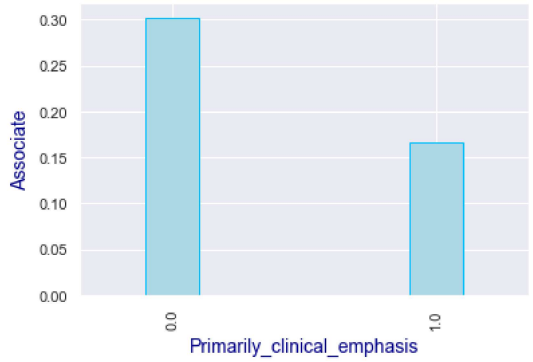
Out[26]:
```
0.0    0.714286
1.0    0.285714
Name: Assistant, dtype: float64
```

Funders of medical research the world over are increasingly seeking, in research assessment, to complement traditional output measures of scientific publications with more outcome-based indicators of societal and economic impact

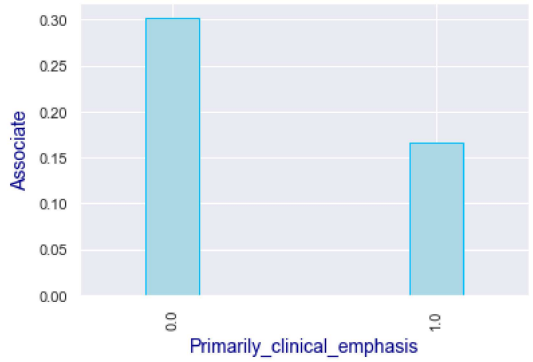
In [27]:
```
Biochemistry.groupby('Primarily_research_emphasis').Gender.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deepsky
plt.xlabel('Assistant',size=14,labelpad=6,color='darkblue')
plt.ylabel('Primarily_research_emphasis', size=14,labelpad=6,color='darkblue')
plt.show()
```



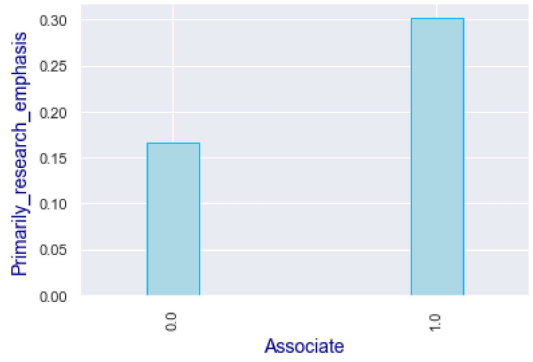

In [28]:
```
Biochemistry.groupby('Primarily_clinical_emphasis').Assistant.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deep
plt.xlabel('Primarily_clinical_emphasis',size=14,labelpad=6,color='darkblue')
plt.ylabel('Associate', size=14,labelpad=6,color='darkblue')
plt.show()
```



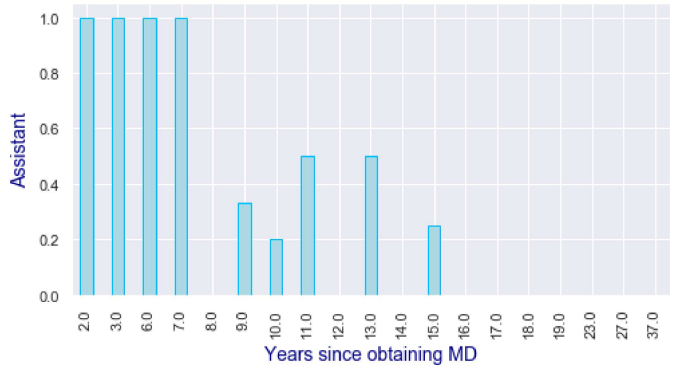
In [29]:
```
Biochemistry.groupby('Primarily_clinical_emphasis').Associate.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deeps
plt.xlabel('Primarily_clinical_emphasis',size=14,labelpad=6,color='darkblue')
plt.ylabel('Associate', size=14,labelpad=6,color='darkblue')
plt.show()
```




In [30]:
```
Biochemistry.groupby('Primarily_research_emphasis').Associate.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deeps
plt.xlabel('Associate',size=14,labelpad=6,color='darkblue')
plt.ylabel('Primarily_research_emphasis', size=14,labelpad=6,color='darkblue')
plt.show()
```

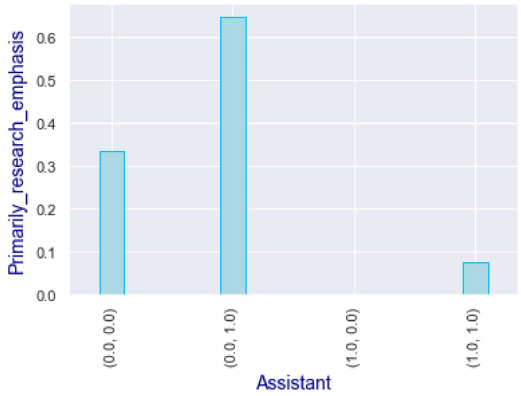



In [31]:
```
Biochemistry.groupby('Exper').Assistant.mean().plot(kind='bar',width=0.4,color='lightblue',edgecolor='deepskyblue',figsize=(8,4)
plt.xlabel('Years since obtaining MD',size=14,labelpad=6,color='darkblue')
plt.ylabel('Assistant', size=14,labelpad=6,color='darkblue')
plt.show()
```

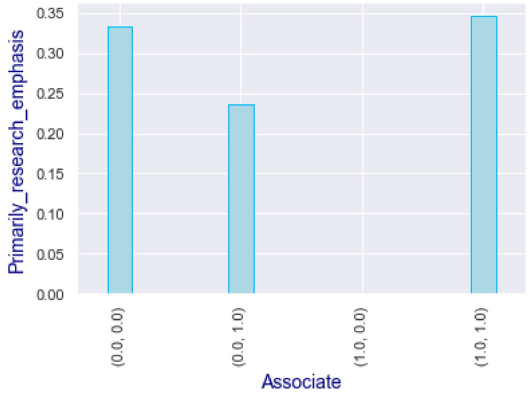



In [ ]:
```
Biochemistry.groupby('Exper').Associate.mean().plot(kind='bar',width=0.4,color='lightblue',edgecolor='deepskyblue',figsize=(6,4)
plt.xlabel('Years since obtaining MD',size=14,labelpad=6,color='darkblue')
plt.ylabel('Associate', size=14,labelpad=6,color='darkblue')
plt.show()
```

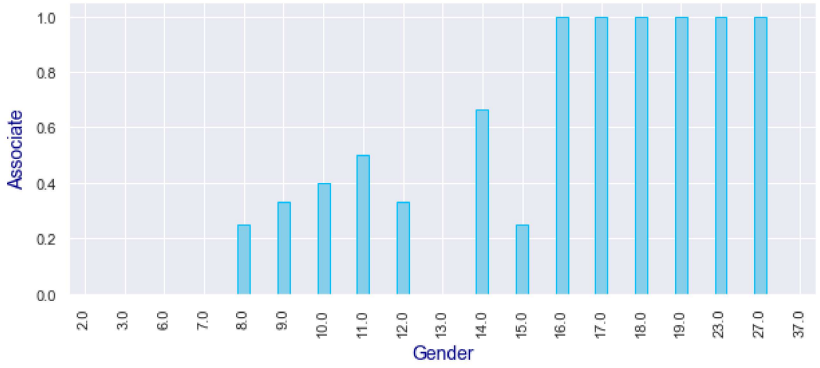
In [33]:
```
Biochemistry.groupby(['Gender','Primarily_research_emphasis']).Assistant.mean().plot(kind='bar',width=0.2,color='lightblue',edge
plt.xlabel('Assistant',size=14,labelpad=6,color='darkblue')
plt.ylabel('Primarily_research_emphasis', size=14,labelpad=6,color='darkblue')
plt.show()
```




In [34]:
```
Biochemistry.groupby(['Gender','Primarily_research_emphasis']).Associate.mean().plot(kind='bar',width=0.2,color='lightblue',edge
plt.xlabel('Associate',size=14,labelpad=6,color='darkblue')
plt.ylabel('Primarily_research_emphasis', size=14,labelpad=6,color='darkblue')
plt.show()
```

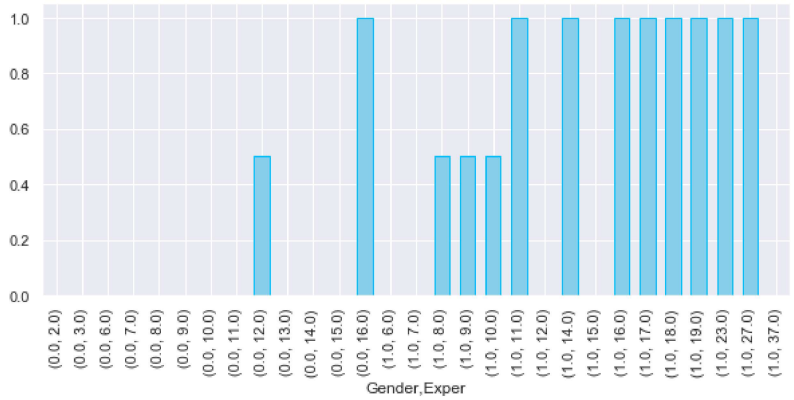



In [35]:
```
Biochemistry.groupby('Exper').full_professor.mean().plot(kind='bar',width=0.3,color='skyblue',edgecolor='deepskyblue',figsize=(1
plt.xlabel('Gender',size=14,labelpad=6,color='darkblue')
plt.ylabel('Associate', size=14,labelpad=6,color='darkblue')
plt.show()
```

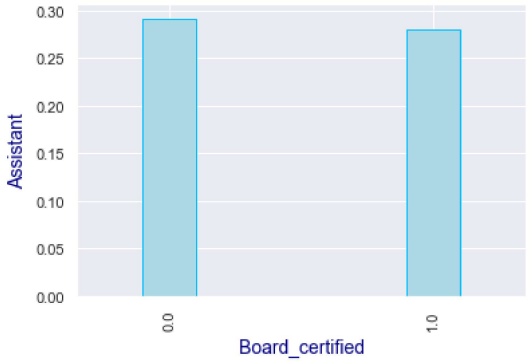


In [36]:
```
Biochemistry.groupby(['Gender','Exper']).full_professor.median().plot(kind='bar',width=0.6,color='skyblue',edgecolor='deepskyblu
```

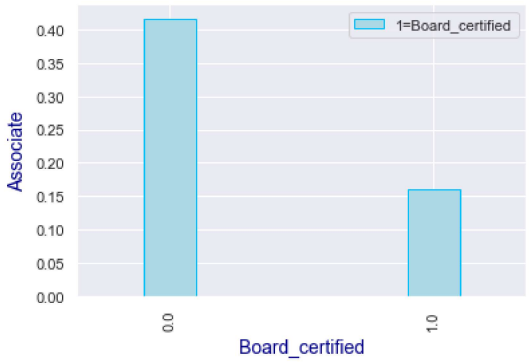
Out[36]: <matplotlib.axes._subplots.AxesSubplot at 0x1e22e8fc848>



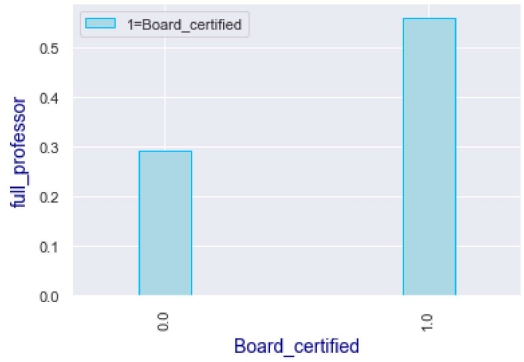

In [37]:
```
Biochemistry.groupby('Board_certified').Assistant.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deepskyblue',fig
plt.xlabel('Board_certified', size=14,labelpad=6,color='darkblue')
plt.ylabel('Assistant', size=14,labelpad=6,color='darkblue')
plt.show()
```



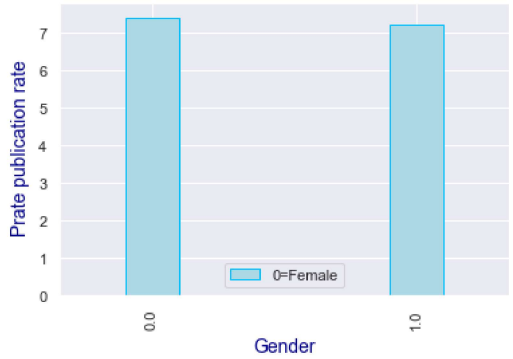

In [38]:
```
Biochemistry.groupby('Board_certified').Associate.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deepskyblue',fig
plt.xlabel('Board_certified',size=14,labelpad=6,color='darkblue')
plt.ylabel('Associate', size=14,labelpad=6,color='darkblue')
plt.legend(['1=Board_certified'])
plt.show()
```



In [39]:
```
Biochemistry.groupby('Board_certified').full_professor.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deepskyblue
plt.xlabel('Board_certified',size=14,labelpad=6,color='darkblue')
plt.ylabel('full_professor', size=14,labelpad=6,color='darkblue')
plt.legend(['1=Board_certified'])
plt.show()
```




In [40]:
```
Biochemistry.groupby('Gender').Prate.mean().plot(kind='bar',width=0.2,color='lightblue',edgecolor='deepskyblue',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='darkblue')
plt.ylabel(' Prate publication rate', size=14,labelpad=6,color='darkblue')
plt.legend(['0=Female'])
plt.show()
```




**Physiology**

In [41]:
```
Physiology=DIS.loc[50:89,]
```

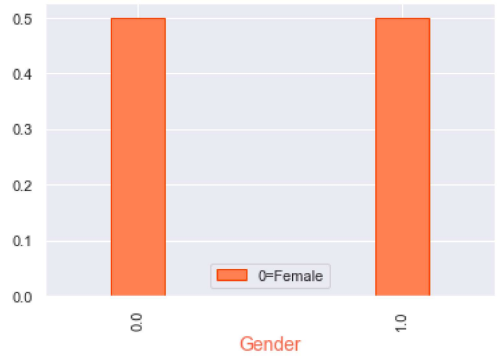
In [42]:
```
Physiology[Physiology['Exper']<30].head()
```

Out[42]:

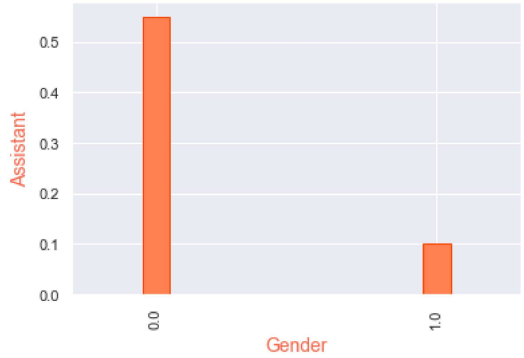
| | Dept | Gender | Prate | Exper | Sal94 | Sal95 | Primarily_research_emphasis | Primarily_clinical_emphasis | not_certified | Board_certified | Assistant | Associat |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 | 2.0 | 1.0 | 5.1 | 10.0 | 97707.0 | 101617.0 | 0.0 | 1.0 | 0.0 | 1.0 | 1.0 | 0. |
| 51 | 2.0 | 1.0 | 5.1 | 9.0 | 106130.0 | 118258.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0. |
| 52 | 2.0 | 1.0 | 6.8 | 8.0 | 81049.0 | 88505.0 | 1.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0. |
| 53 | 2.0 | 1.0 | 5.5 | 25.0 | 140315.0 | 150710.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0. |
| 54 | 2.0 | 1.0 | 7.1 | 11.0 | 84571.0 | 92879.0 | 1.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0. |

In [43]:
```
Physiology.Gender.value_counts(normalize=True).plot(kind='bar',width=0.2, color='coral',edgecolor='orangered',figsize=(6,4))
plt.xlabel('Gender',size=14,color='tomato')
plt.legend(['0=Female'])
```

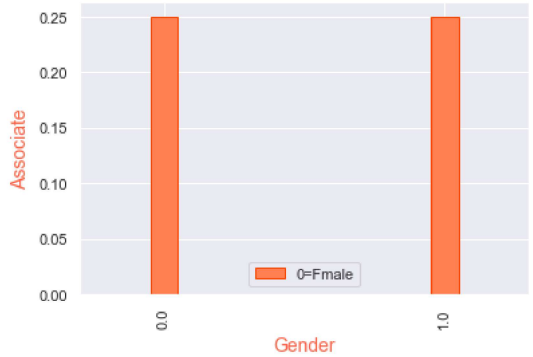
Out[43]: <matplotlib.legend.Legend at 0x1e22e74be48>



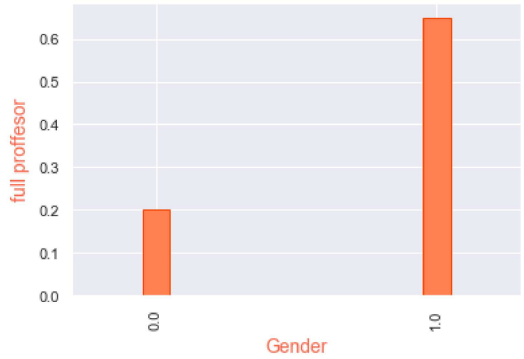

In [44]:
```
Physiology.groupby('Gender').Assistant.mean().plot(kind='bar',width=0.1,color='coral',edgecolor='orangered',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='tomato')
plt.ylabel('Assistant', size=14,labelpad=6,color='tomato')
plt.show()
```




In [45]:
```
Physiology.groupby('Gender').Associate.mean().plot(kind='bar',width=0.1,color='coral',edgecolor='orangered',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='tomato')
plt.ylabel('Associate', size=14,labelpad=6,color='tomato')
plt.legend(['0=Fmale'])
plt.show()
```



In [46]:
```
Physiology.groupby('Gender').full_professor.mean().plot(kind='bar',width=0.1,color='coral',edgecolor='orangered',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='tomato')
plt.ylabel('full proffesor', size=14,labelpad=6,color='tomato')
plt.show()
```

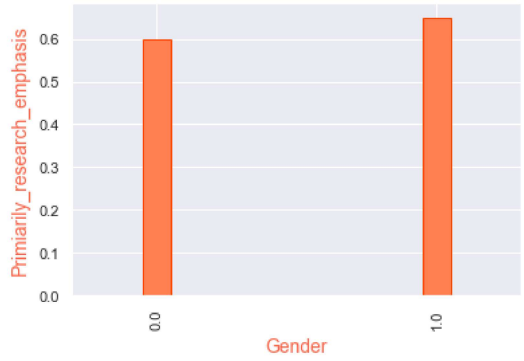



In [47]:
```
Physiology.groupby('Board_certified').Assistant.mean().plot(kind='bar',width=0.2,color='coral',edgecolor='orangered',figsize=(6,
plt.xlabel('Board_certified',size=14,labelpad=6,color='tomato')
plt.ylabel('Associate', size=14,labelpad=6,color='tomato')
plt.legend(['1=Board_certified'])
plt.show()
```

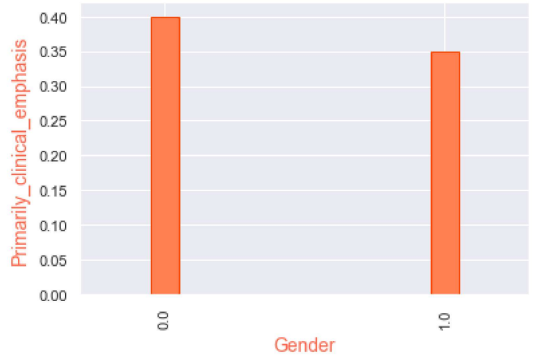

In [48]:
```
Physiology.groupby('Board_certified').Associate.mean().plot(kind='bar',width=0.2,color='coral',edgecolor='orangered',figsize=(6,
plt.xlabel('Board_certified',size=14,labelpad=6,color='tomato')
plt.ylabel('Associate', size=14,labelpad=6,color='tomato')
plt.legend(['1=Board_certified'])
plt.show()
```

In [49]:
```
Physiology.groupby('Board_certified').full_professor.mean().plot(kind='bar',width=0.2,color='coral',edgecolor='orangered',figsiz
plt.xlabel('Board_certified',size=14,labelpad=6,color='tomato')
plt.ylabel('full_professor', size=14,labelpad=6,color='tomato')
plt.legend(['1=Board_certified'])
plt.show()
```

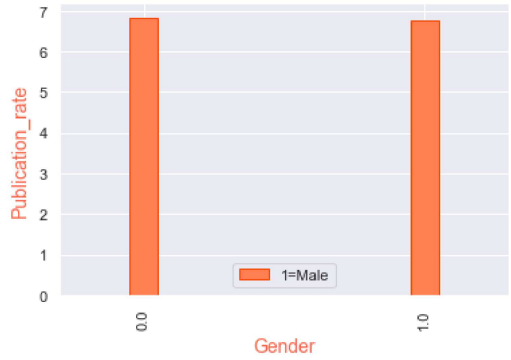

In [50]:
```
Physiology.groupby('Gender').Primarily_research_emphasis.mean().plot(kind='bar',width=0.1,color='coral',edgecolor='orangered',fi
plt.xlabel('Gender',size=14,labelpad=6,color='tomato')
plt.ylabel('Primiarily_research_emphasis', size=14,labelpad=6,color='tomato')
plt.show()
```



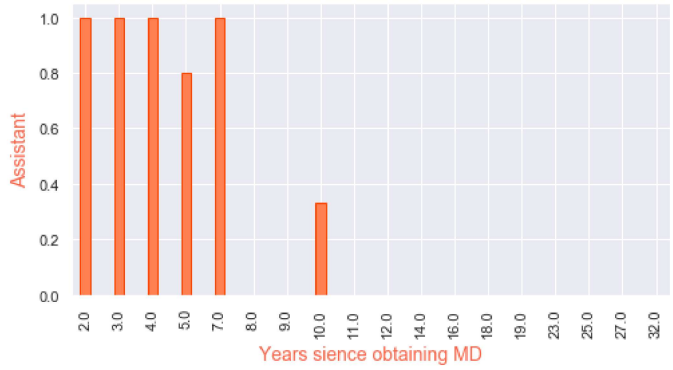

In [51]:
```
Physiology.groupby('Gender').Primarily_clinical_emphasis.mean().plot(kind='bar',width=0.1,color='coral',edgecolor='orangered',fi
plt.xlabel('Gender',size=14,labelpad=6,color='tomato')
plt.ylabel('Primarily_clinical_emphasis', size=14,labelpad=6,color='tomato')
plt.show()
```



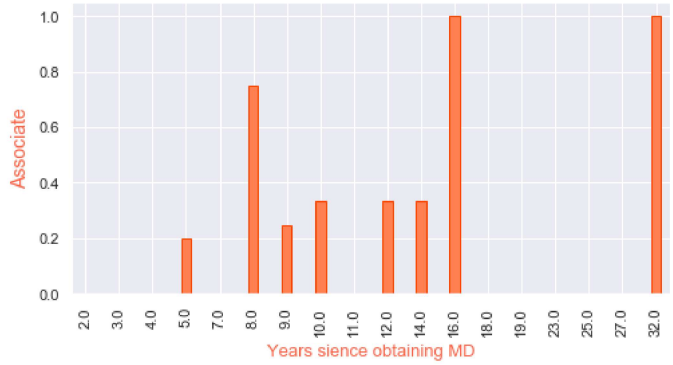
In [52]:
```
Physiology.groupby('Gender').Prate.mean().plot(kind='bar',width=0.1,color='coral',edgecolor='orangered',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='tomato')
plt.ylabel('Publication_rate', size=14,labelpad=6,color='tomato')
plt.legend(['1=Male'])
plt.show()
```



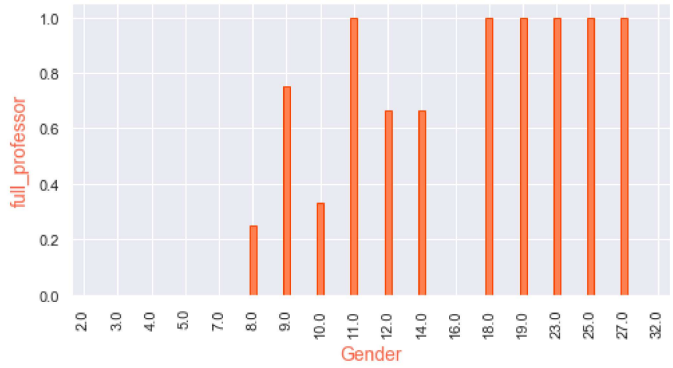

In [53]:
```
Physiology.groupby('Exper').Assistant.mean().plot(kind='bar',width=0.3,color='coral',edgecolor='orangered',figsize=(8,4))
plt.xlabel('Years sience obtaining MD',size=14,labelpad=6,color='tomato')
plt.ylabel('Assistant', size=14,labelpad=6,color='tomato')
plt.show()
```




In [54]:
```
Physiology.groupby('Exper').Associate.mean().plot(kind='bar',width=0.3,color='coral',edgecolor='orangered',figsize=(8,4))
plt.xlabel('Years sience obtaining MD',size=13,labelpad=6,color='tomato')
plt.ylabel('Associate', size=14,labelpad=6,color='tomato')
plt.show()
```

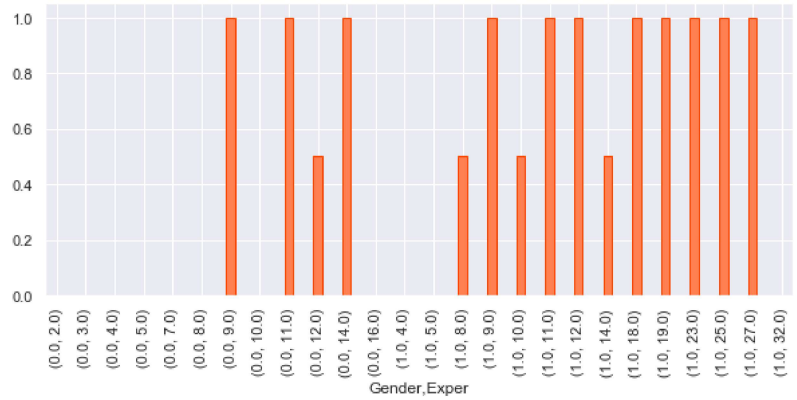


In [55]:
```
Physiology.groupby('Exper').full_professor.mean().plot(kind='bar',width=0.2,color='coral',edgecolor='orangered',figsize=(8,4))
plt.xlabel('Gender',size=14,labelpad=6,color='tomato')
plt.ylabel('full_professor', size=14,labelpad=6,color='tomato')
plt.show()
```




In [56]:
```
Physiology.groupby(['Gender','Exper']).full_professor.median().plot(kind='bar',width=0.3,color='coral',edgecolor='orangered',fig
```

Out[56]: <matplotlib.axes._subplots.AxesSubplot at 0x1e22e5a69c8>




**Genetics**

In [57]:
```
Genetics=DIS.loc[90:110,]
```

In [58]:
```
Genetics.head()
```

Out[58]:

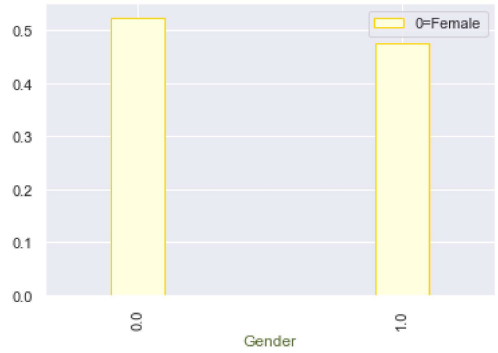
| | Dept | Gender | Prate | Exper | Sal94 | Sal95 | Primarily_research_emphasis | Primarily_clinical_emphasis | not_certified | Board_certified | Assistant | Associat |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 90 | 3.0 | 1.0 | 6.7 | 24.0 | 154326.0 | 165884.0 | 1.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0. |
| 91 | 3.0 | 1.0 | 5.9 | 11.0 | 87351.0 | 95429.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | 1. |
| 92 | 3.0 | 1.0 | 4.1 | 5.0 | 123859.0 | 130292.0 | 0.0 | 1.0 | 0.0 | 1.0 | 1.0 | 0. |
| 93 | 3.0 | 1.0 | 4.8 | 16.0 | 118121.0 | 130697.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | 0. |
| 94 | 3.0 | 1.0 | 4.3 | 11.0 | 174479.0 | 193342.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0. |

In [59]:
```
Genetics.Gender.value_counts(normalize=True)
```

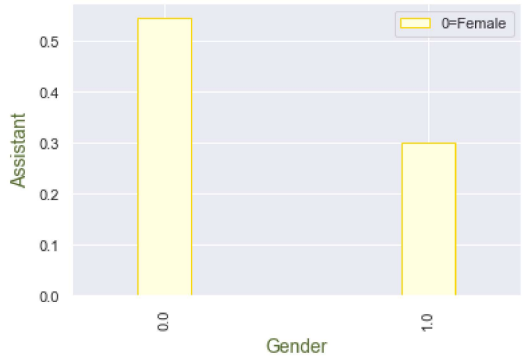
Out[59]:
```
0.0    0.52381
1.0    0.47619
Name: Gender, dtype: float64
```

In [60]:
```
Genetics.Gender.value_counts(normalize=True).plot(kind='bar',width=0.2, color='lightyellow',edgecolor='gold',figsize=(6,4))
plt.xlabel('Gender',size=12,color='darkolivegreen')
plt.legend(['0=Female'])
```

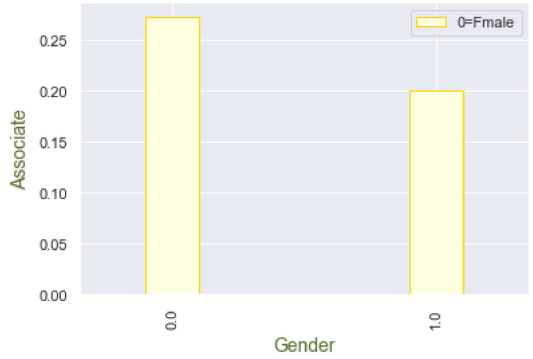
Out[60]: <matplotlib.legend.Legend at 0x1e22e7bdc48>



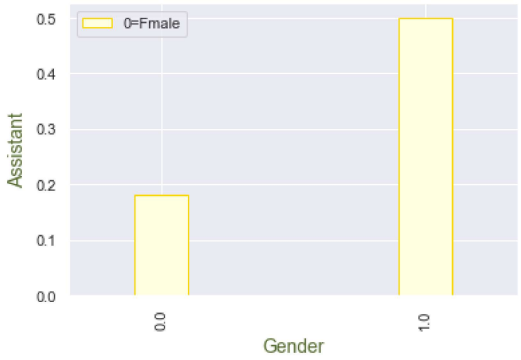

In [61]:
```
Genetics.groupby('Gender').Assistant.mean().plot(kind='bar',width=0.2,color='lightyellow',edgecolor='gold',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='darkolivegreen')
plt.ylabel('Assistant', size=14,labelpad=6,color='darkolivegreen')
plt.legend(['0=Female'])
plt.show()
```




In [62]:
```
Genetics.groupby('Gender').Associate.mean().plot(kind='bar',width=0.2,color='lightyellow',edgecolor='gold',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='darkolivegreen')
plt.ylabel('Associate', size=14,labelpad=6,color='darkolivegreen')
plt.legend(['0=Fmale'])
plt.show()
```

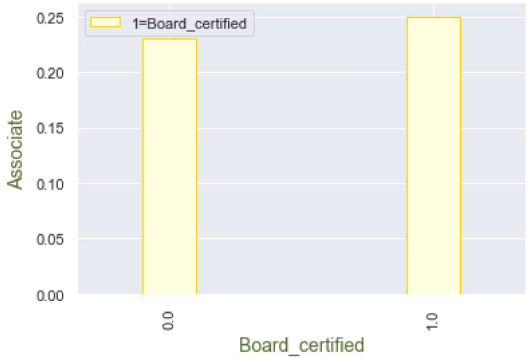


In [63]:
```
Genetics.groupby('Gender').full_professor.mean().plot(kind='bar',width=0.2,color='lightyellow',edgecolor='gold',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='darkolivegreen')
plt.ylabel('Assistant', size=14,labelpad=6,color='darkolivegreen')
plt.legend(['0=Fmale'])
plt.show()
```

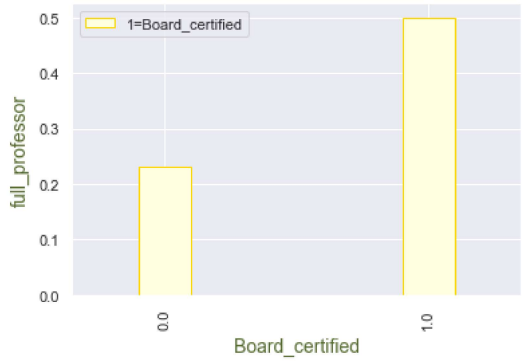



In [64]:
```
Genetics.groupby('Board_certified').Assistant.mean().plot(kind='bar',width=0.2,color='lightyellow',edgecolor='gold',figsize=(6,4
plt.xlabel('Board_certified',size=14,labelpad=6,color='darkolivegreen')
plt.ylabel('Assistant', size=14,labelpad=6,color='darkolivegreen')
plt.show()
```

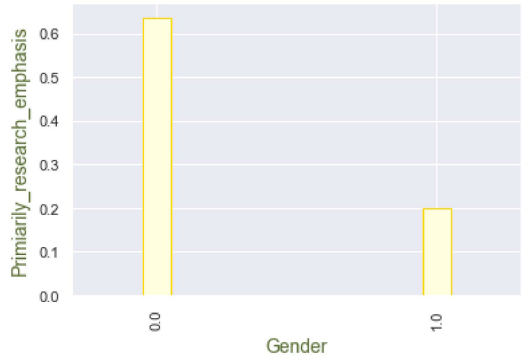

In [65]:
```
Genetics.groupby('Board_certified').Associate.mean().plot(kind='bar',width=0.2,color='lightyellow',edgecolor='gold',figsize=(6,4
plt.xlabel('Board_certified',size=14,labelpad=6,color='darkolivegreen')
plt.ylabel('Associate', size=14,labelpad=6,color='darkolivegreen')
plt.legend(['1=Board_certified'])
plt.show()
```



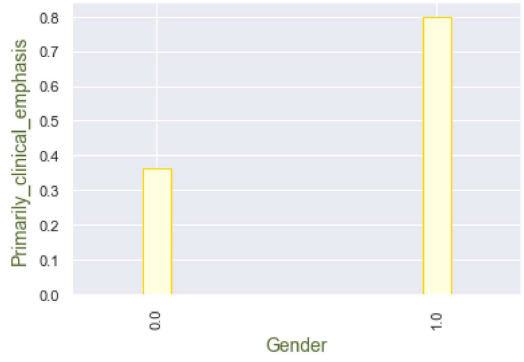
In [66]:
```
Genetics.groupby('Board_certified').full_professor.mean().plot(kind='bar',width=0.2,color='lightyellow',edgecolor='gold',figsize=
plt.xlabel('Board_certified',size=14,labelpad=6,color='darkolivegreen')
plt.ylabel('full_professor', size=14,labelpad=6,color='darkolivegreen')
plt.legend(['1=Board_certified'])
plt.show()
```



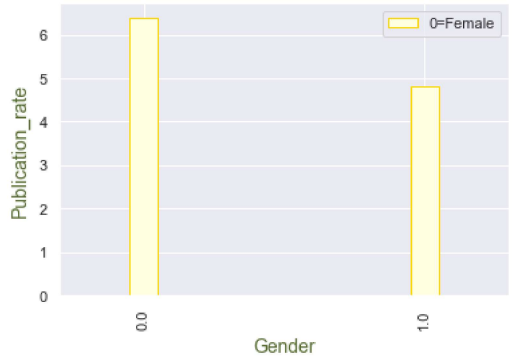

In [67]:
```
Genetics.groupby('Gender').Primarily_research_emphasis.mean().plot(kind='bar',width=0.1,color='lightyellow',edgecolor='gold',fig
plt.xlabel('Gender',size=14,labelpad=6,color='darkolivegreen')
plt.ylabel('Primiarily_research_emphasis', size=14,labelpad=6,color='darkolivegreen')
plt.show()
```



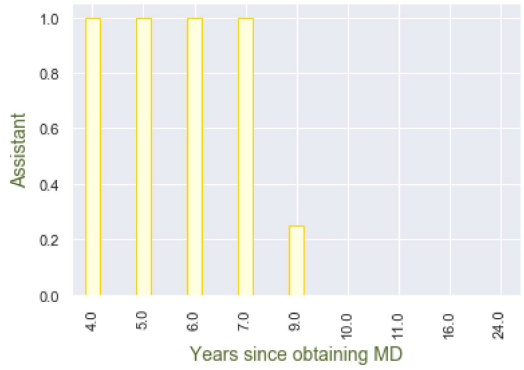

In [68]:
```
Genetics.groupby('Gender').Primarily_clinical_emphasis.mean().plot(kind='bar',width=0.1,color='lightyellow',edgecolor='gold',fig
plt.xlabel('Gender',size=14,labelpad=6,color='darkolivegreen')
plt.ylabel('Primarily_clinical_emphasis', size=14,labelpad=6,color='darkolivegreen')
plt.show()
```



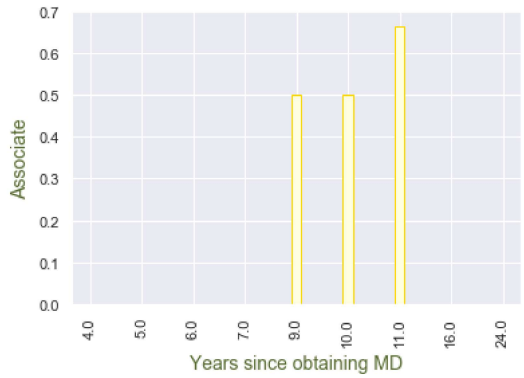
In [69]:
```
Genetics.groupby('Gender').Prate.mean().plot(kind='bar',width=0.1,color='lightyellow',edgecolor='gold',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='darkolivegreen')
plt.ylabel('Publication_rate', size=14,labelpad=6,color='darkolivegreen')
plt.legend(['0=Female'])
plt.show()
```



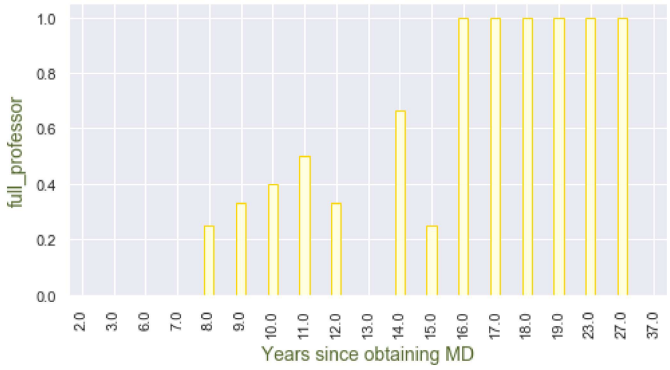

In [70]:
```
Genetics.groupby('Exper').Assistant.mean().plot(kind='bar',width=0.3,color='lightyellow',edgecolor='gold',figsize=(6,4))
plt.xlabel('Years since obtaining MD',size=14,labelpad=6,color='darkolivegreen')
plt.ylabel('Assistant', size=14,labelpad=6,color='darkolivegreen')
plt.show()
```




In [71]:
```
Genetics.groupby('Exper').Associate.mean().plot(kind='bar',width=0.2,color='lightyellow',edgecolor='gold',figsize=(6,4))
plt.xlabel('Years since obtaining MD',size=14,labelpad=6,color='darkolivegreen')
plt.ylabel('Associate', size=14,labelpad=6,color='darkolivegreen')
plt.show()
```



In [72]:
```
Biochemistry.groupby('Exper').full_professor.mean().plot(kind='bar',width=0.3,color='lightyellow',edgecolor='gold',figsize=(8,4))
plt.xlabel('Years since obtaining MD',size=14,labelpad=6,color='darkolivegreen')
plt.ylabel('full_professor', size=14,labelpad=6,color='darkolivegreen')
plt.show()
```




**Pediatrics**

In [73]:
```
Pediatrics=DIS.loc[111:140,]
```

In [74]:
```
Pediatrics.head()
```

Out[74]:

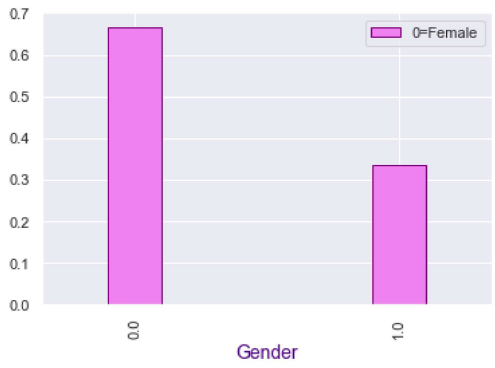
| | Dept | Gender | Prate | Exper | Sal94 | Sal95 | Primarily_research_emphasis | Primarily_clinical_emphasis | not_certified | Board_certified | Assistant | Associa |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 111 | 4.0 | 1.0 | 3.1 | 6.0 | 109599.0 | 121829.0 | 0.0 | 1.0 | 0.0 | 1.0 | 1.0 | 0 |
| 112 | 4.0 | 1.0 | 5.2 | 17.0 | 154159.0 | 169191.0 | 1.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0 |
| 113 | 4.0 | 1.0 | 3.3 | 10.0 | 145503.0 | 162253.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | 1 |
| 114 | 4.0 | 1.0 | 3.2 | 15.0 | 161383.0 | 181196.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | 1 |
| 115 | 4.0 | 1.0 | 4.8 | 10.0 | 113443.0 | 126189.0 | 1.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0 |

In [75]:
```
Pediatrics.Gender.value_counts(normalize=True)
```

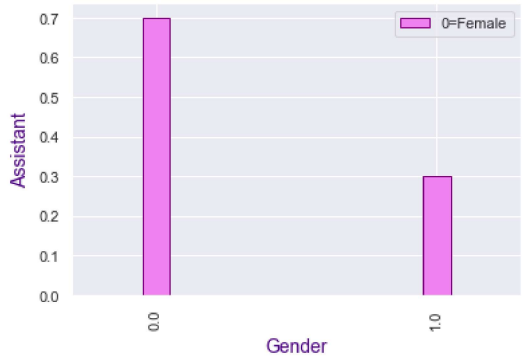
Out[75]:
```
0.0    0.666667
1.0    0.333333
Name: Gender, dtype: float64
```

In [76]:
```
Pediatrics.Gender.value_counts(normalize=True).plot(kind='bar',width=0.2, color='violet',edgecolor='purple',figsize=(6,4))
plt.xlabel('Gender',size=14,color='indigo')
plt.legend(['0=Female'])
```

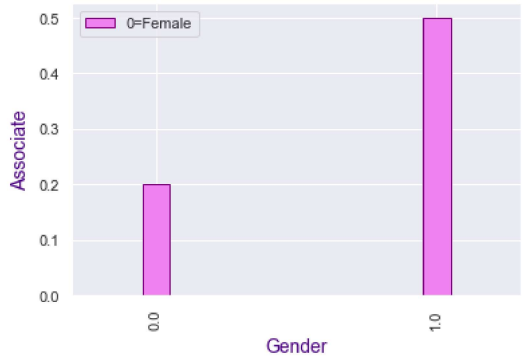
Out[76]:
```
<matplotlib.legend.Legend at 0x1e22e60a388>
```



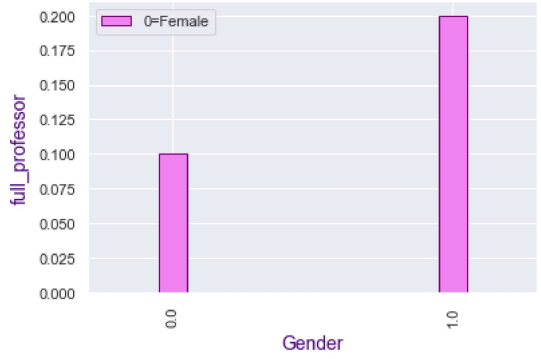
In [77]:
```
Pediatrics.groupby('Gender').Assistant.mean().plot(kind='bar',width=0.1,color='violet',edgecolor='purple',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='indigo')
plt.ylabel('Assistant', size=14,labelpad=6,color='indigo')
plt.legend(['0=Female'])
plt.show()
```



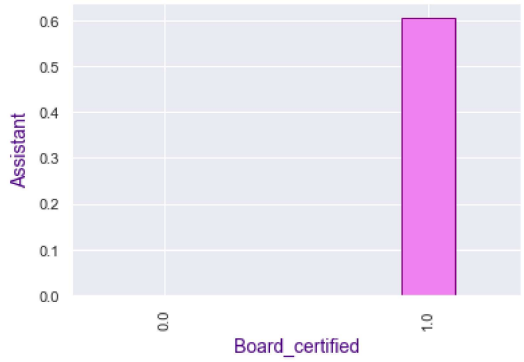

In [78]:
```
Pediatrics.groupby('Gender').Associate.mean().plot(kind='bar',width=0.1,color='violet',edgecolor='purple',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='indigo')
plt.ylabel('Associate', size=14,labelpad=6,color='indigo')
plt.legend(['0=Female'])
plt.show()
```



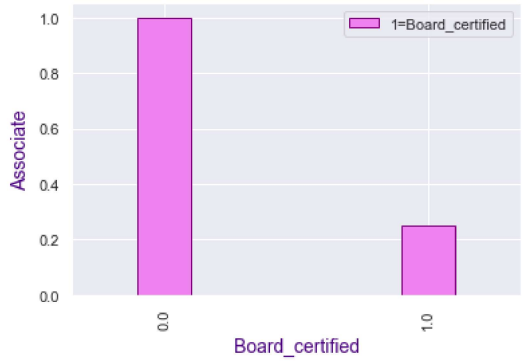

In [79]:
```
Pediatrics.groupby('Gender').full_professor.mean().plot(kind='bar',width=0.1,color='violet',edgecolor='purple',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='indigo')
plt.ylabel('full_professor', size=14,labelpad=6,color='indigo')
plt.legend(['0=Female'])
plt.show()
```



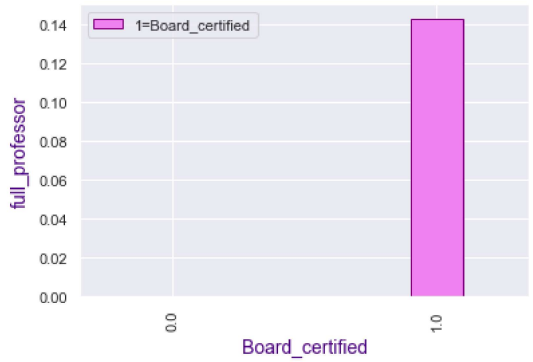
In [80]:
```
Pediatrics.groupby('Board_certified').Assistant.mean().plot(kind='bar',width=0.2,color='violet',edgecolor='purple',figsize=(6,4)
plt.xlabel('Board_certified',size=14,labelpad=6,color='indigo')
plt.ylabel('Assistant', size=14,labelpad=6,color='indigo')
plt.show()
```



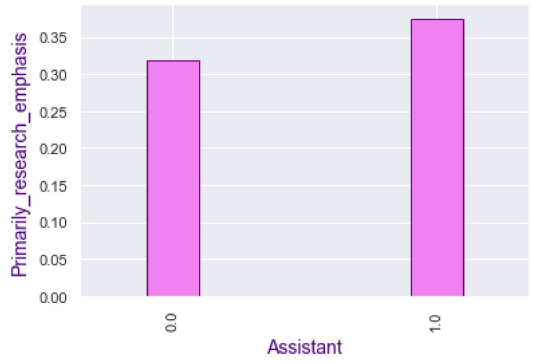

In [81]:
```
Pediatrics.groupby('Board_certified').Associate.mean().plot(kind='bar',width=0.2,color='violet',edgecolor='purple',figsize=(6,4)
plt.xlabel('Board_certified',size=14,labelpad=6,color='indigo')
plt.ylabel('Associate', size=14,labelpad=6,color='indigo')
plt.legend(['1=Board_certified'])
plt.show()
```



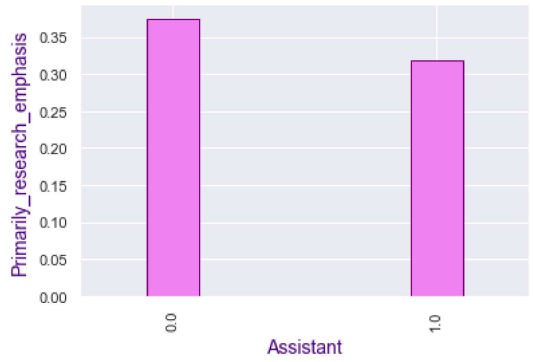

In [82]:
```
Pediatrics.groupby('Board_certified').full_professor.mean().plot(kind='bar',width=0.2,color='violet',edgecolor='purple',figsize=
plt.xlabel('Board_certified',size=14,labelpad=6,color='indigo')
plt.ylabel('full_professor', size=14,labelpad=6,color='indigo')
plt.legend(['1=Board_certified'])
plt.show()
```



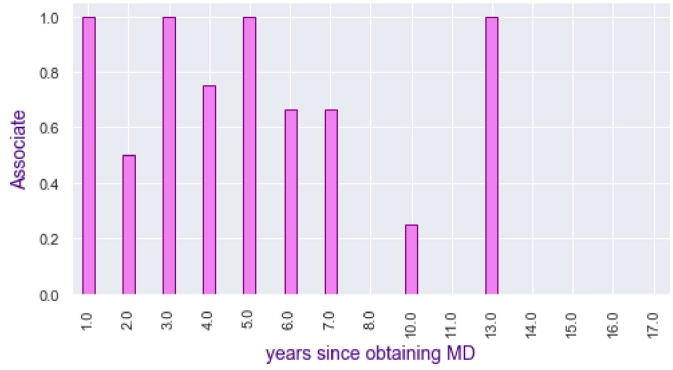
In [83]:
```
Pediatrics.groupby('Primarily_research_emphasis').Gender.mean().plot(kind='bar',width=0.2,color='violet',edgecolor='purple',figs
plt.xlabel('Assistant',size=14,labelpad=6,color='indigo')
plt.ylabel('Primarily_research_emphasis', size=14,labelpad=6,color='indigo')
plt.show()
```



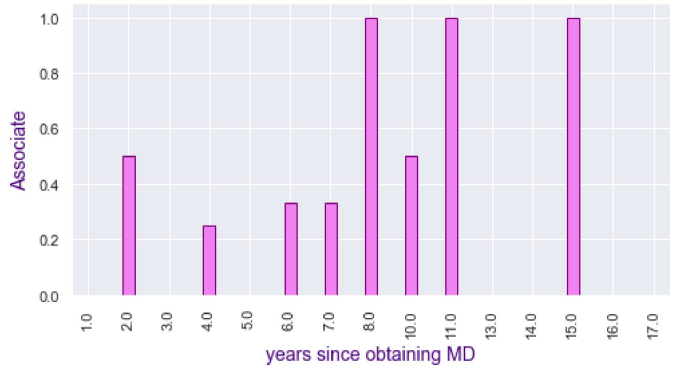

In [84]:
```
Pediatrics.groupby('Primarily_clinical_emphasis').Gender.mean().plot(kind='bar',width=0.2,color='violet',edgecolor='purple',figs
plt.xlabel('Assistant',size=14,labelpad=6,color='indigo')
plt.ylabel('Primarily_research_emphasis', size=14,labelpad=6,color='indigo')
plt.show()
```



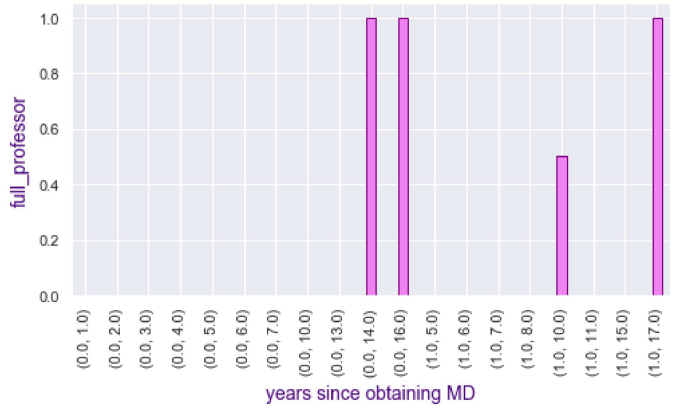

In [85]:
```
Pediatrics.groupby('Exper').Assistant.mean().plot(kind='bar',width=0.3,color='violet',edgecolor='purple',figsize=(8,4))
plt.xlabel('years since obtaining MD',size=14,labelpad=6,color='indigo')
plt.ylabel('Associate', size=14,labelpad=6,color='indigo')
plt.show()
```



In [86]:
```
Pediatrics.groupby('Exper').Associate.mean().plot(kind='bar',width=0.3,color='violet',edgecolor='purple',figsize=(8,4))
plt.xlabel('years since obtaining MD',size=14,labelpad=6,color='indigo')
plt.ylabel('Associate', size=14,labelpad=6,color='indigo')
plt.show()
```




In [87]:
```
Pediatrics.groupby(['Gender','Exper']).full_professor.mean().plot(kind='bar',width=0.3,color='violet',edgecolor='purple',figsize=
plt.xlabel('years since obtaining MD',size=14,labelpad=6,color='indigo')
plt.ylabel('full_professor', size=14,labelpad=6,color='indigo')
plt.show()
```




**Medicine**

In [88]:
```
Medicine=DIS.loc[141:220,]
```

In [89]:
```
Medicine[Medicine['Exper']<30].head()
```

Out[89]:

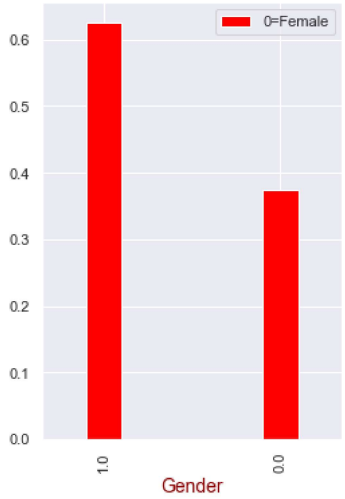
| | Dept | Gender | Prate | Exper | Sal94 | Sal95 | Primarily_research_emphasis | Primarily_clinical_emphasis | not_certified | Board_certified | Assistant | Associa |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 141 | 5.0 | 1.0 | 3.7 | 11.0 | 182945.0 | 198762.0 | 0.0 | 1.0 | 0.0 | 1.0 | 1.0 | ( |
| 142 | 5.0 | 1.0 | 3.2 | 13.0 | 163671.0 | 181774.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | ' |
| 143 | 5.0 | 1.0 | 4.3 | 9.0 | 204433.0 | 226581.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | ' |
| 144 | 5.0 | 1.0 | 3.2 | 4.0 | 135139.0 | 148117.0 | 0.0 | 1.0 | 0.0 | 1.0 | 1.0 | ( |
| 145 | 5.0 | 1.0 | 5.6 | 11.0 | 147213.0 | 155826.0 | 1.0 | 0.0 | 0.0 | 1.0 | 0.0 | ( |

In [90]:
```
Medicine.Gender.value_counts(normalize=True)
```

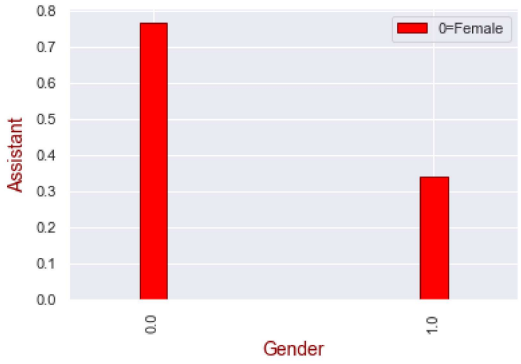
Out[90]:
```
1.0    0.625
0.0    0.375
Name: Gender, dtype: float64
```

In [91]:
```
Medicine.Gender.value_counts(normalize=True).plot(kind='bar',width=0.2, color='red',figsize=(4,6))
plt.xlabel('Gender',size=14,color='darkred')
plt.legend(['0=Female'])
```

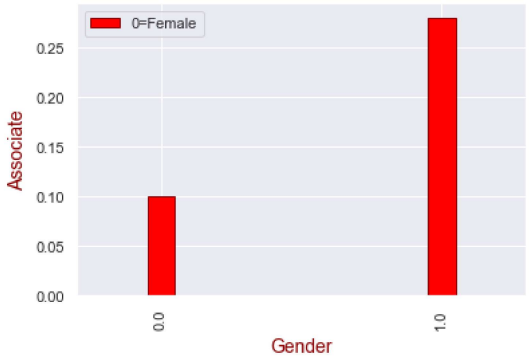
Out[91]: <matplotlib.legend.Legend at 0x1e22e821d48>



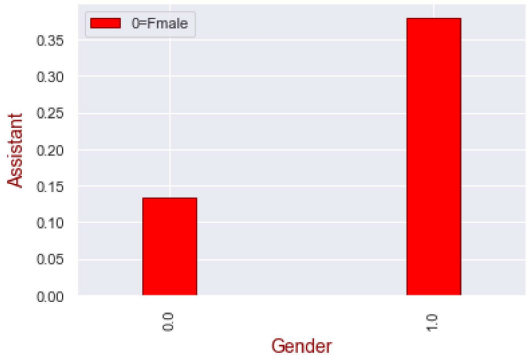

In [92]:
```
Medicine.groupby('Gender').Assistant.mean().plot(kind='bar',width=0.1,color='red',edgecolor='darkred',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='darkred')
plt.ylabel('Assistant', size=14,labelpad=6,color='darkred')
plt.legend(['0=Female'])
plt.show()
```



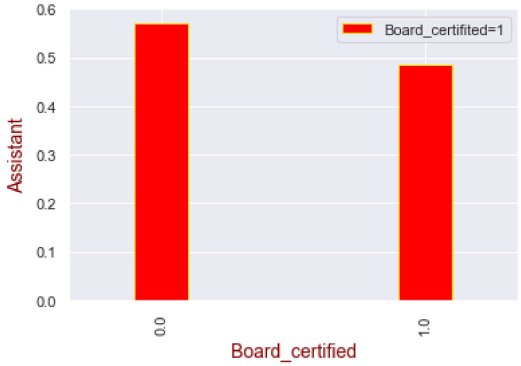

In [93]:
```
Medicine.groupby('Gender').Associate.mean().plot(kind='bar',width=0.1,color='red',edgecolor='darkred',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='darkred')
plt.ylabel('Associate', size=14,labelpad=6,color='darkred')
plt.legend(['0=Female'])
plt.show()
```



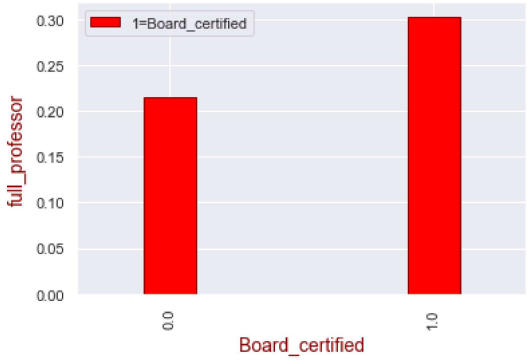
In [94]:
```
Medicine.groupby('Gender').full_professor.mean().plot(kind='bar',width=0.2,color='red',edgecolor='darkred',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='darkred')
plt.ylabel('Assistant', size=14,labelpad=6,color='darkred')
plt.legend(['0=Fmale'])
plt.show()
```



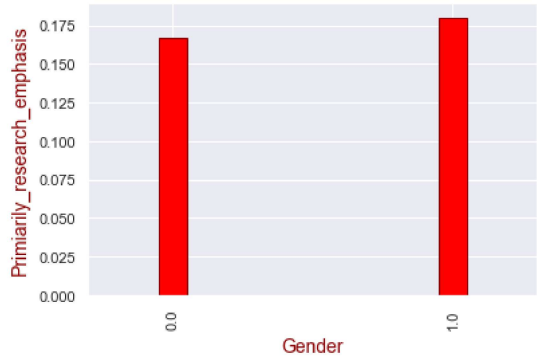

In [95]:
```
Medicine.groupby('Board_certified').Assistant.mean().plot(kind='bar',width=0.2,color='red',edgecolor='gold',figsize=(6,4))
plt.xlabel('Board_certified',size=14,labelpad=6,color='darkred')
plt.ylabel('Assistant', size=14,labelpad=6,color='darkred')
plt.legend(['Board_certifited=1'])
plt.show()
```



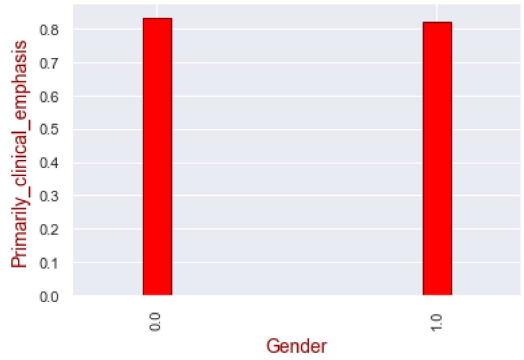

In [96]:
```
Medicine.groupby('Board_certified').full_professor.mean().plot(kind='bar',width=0.2,color='red',edgecolor='darkred',figsize=(6,4
plt.xlabel('Board_certified',size=14,labelpad=6,color='darkred')
plt.ylabel('full_professor', size=14,labelpad=6,color='darkred')
plt.legend(['1=Board_certified'])
plt.show()
```



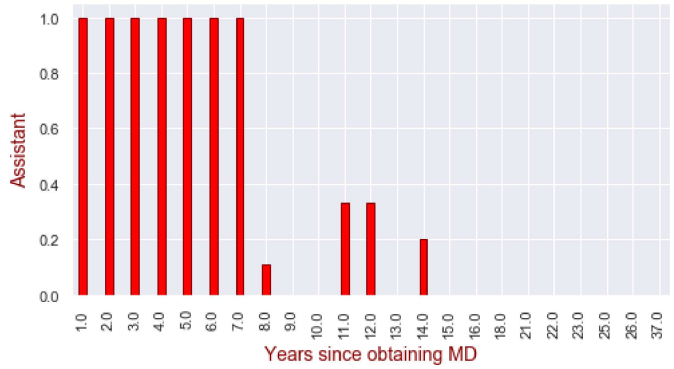
In [97]:
```
Medicine.groupby('Gender').Primarily_research_emphasis.mean().plot(kind='bar',width=0.1,color='red',edgecolor='darkred',figsize=
plt.xlabel('Gender',size=14,labelpad=6,color='darkred')
plt.ylabel('Primiarily_research_emphasis', size=14,labelpad=6,color='darkred')
plt.show()
```



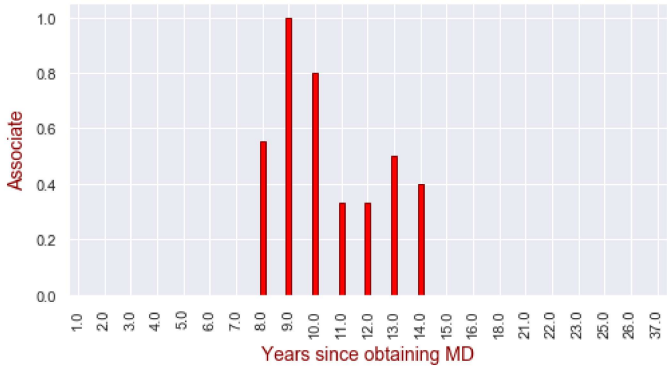

In [98]:
```
Medicine.groupby('Gender').Primarily_clinical_emphasis.mean().plot(kind='bar',width=0.1,color='red',edgecolor='darkred',figsize=
plt.xlabel('Gender',size=14,labelpad=6,color='darkred')
plt.ylabel('Primarily_clinical_emphasis', size=14,labelpad=6,color='darkred')
plt.show()
```



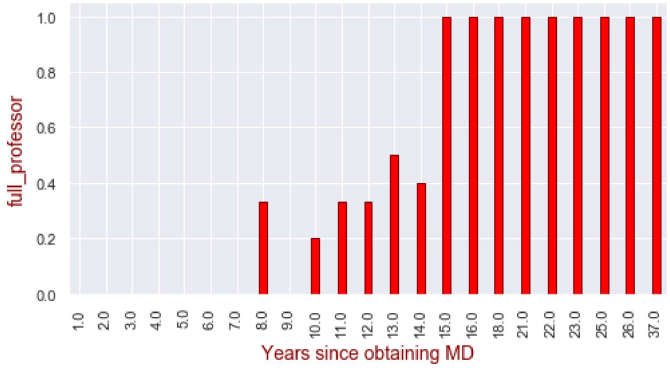

In [99]:
```
Medicine.groupby('Exper').Assistant.mean().plot(kind='bar',width=0.3,color='red',edgecolor='darkred',figsize=(8,4))
plt.xlabel('Years since obtaining MD',size=14,labelpad=6,color='darkred')
plt.ylabel('Assistant', size=14,labelpad=6,color='darkred')
plt.show()
```



In [100]:
```
Medicine.groupby('Exper').Associate.mean().plot(kind='bar',width=0.2,color='red',edgecolor='darkred',figsize=(8,4))
plt.xlabel('Years since obtaining MD',size=14,labelpad=6,color='darkred')
plt.ylabel('Associate', size=14,labelpad=6,color='darkred')
plt.show()
```




In [101]:
```
Medicine.groupby('Exper').full_professor.mean().plot(kind='bar',width=0.3,color='red',edgecolor='darkred',figsize=(8,4))
plt.xlabel('Years since obtaining MD',size=14,labelpad=6,color='darkred')
plt.ylabel('full_professor', size=14,labelpad=6,color='darkred')
plt.show()
```




**Surgery**

In [102]:
```
Surgery=DIS.loc[221:260,]
```

In [103]:
```
Surgery.head()
```

Out[103]:

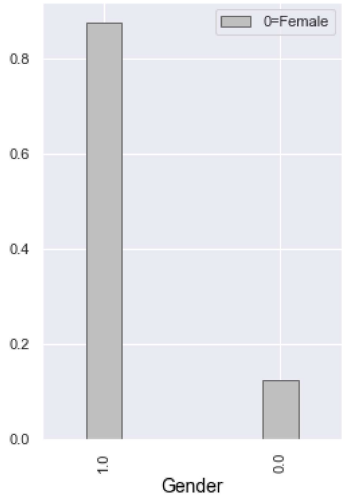
| | Dept | Gender | Prate | Exper | Sal94 | Sal95 | Primarily_research_emphasis | Primarily_clinical_emphasis | not_certified | Board_certified | Assistant | Associa |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 221 | 6.0 | 1.0 | 2.7 | 6.0 | 214219.0 | 236305.0 | 0.0 | 1.0 | 0.0 | 1.0 | 1.0 | 0 |
| 222 | 6.0 | 1.0 | 2.1 | 5.0 | 312966.0 | 336960.0 | 0.0 | 1.0 | 0.0 | 1.0 | 1.0 | 0 |
| 223 | 6.0 | 1.0 | 1.9 | 21.0 | 342834.0 | 367284.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | 1 |
| 224 | 6.0 | 1.0 | 2.1 | 16.0 | 312892.0 | 344801.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0 |
| 225 | 6.0 | 1.0 | 2.5 | 13.0 | 317399.0 | 345204.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0 |

In [104]:
```
Surgery.Gender.value_counts(normalize=True)
```

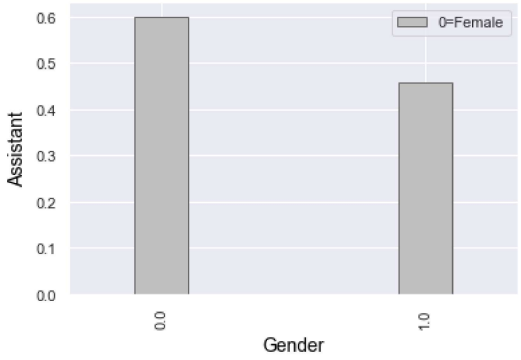
Out[104]:
```
1.0    0.875
0.0    0.125
Name: Gender, dtype: float64
```

In [105]:
```
Surgery.Gender.value_counts(normalize=True).plot(kind='bar',width=0.2, color='silver',edgecolor='dimgray',figsize=(4,6))
plt.xlabel('Gender',size=14,color='black')
plt.legend(['0=Female'])
```

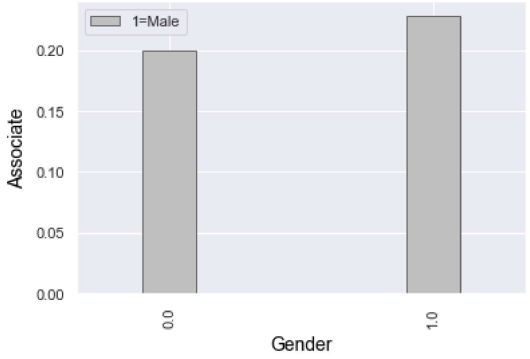
Out[105]: <matplotlib.legend.Legend at 0x1e22e61f688>



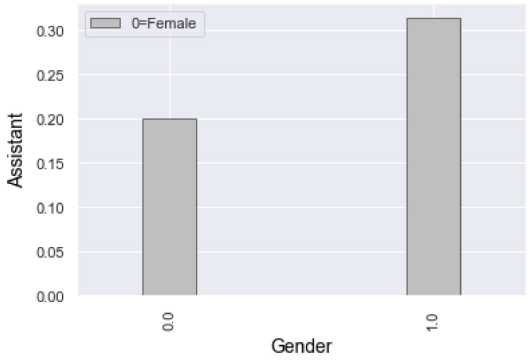

In [106]:
```
Surgery.groupby('Gender').Assistant.mean().plot(kind='bar',width=0.2,color='silver',edgecolor='dimgray',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='black')
plt.ylabel('Assistant', size=14,labelpad=6,color='black')
plt.legend(['0=Female'])
plt.show()
```



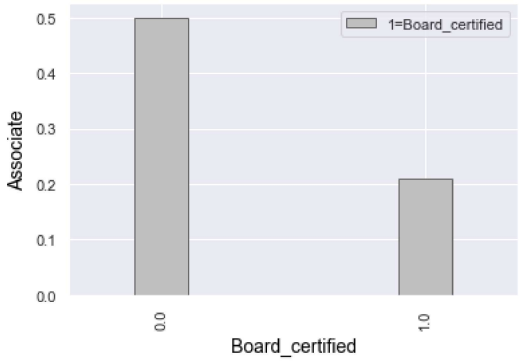

In [107]:
```
Surgery.groupby('Gender').Associate.mean().plot(kind='bar',width=0.2,color='silver',edgecolor='dimgray',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='black')
plt.ylabel('Associate', size=14,labelpad=6,color='black')
plt.legend(['1=Male'])
plt.show()
```



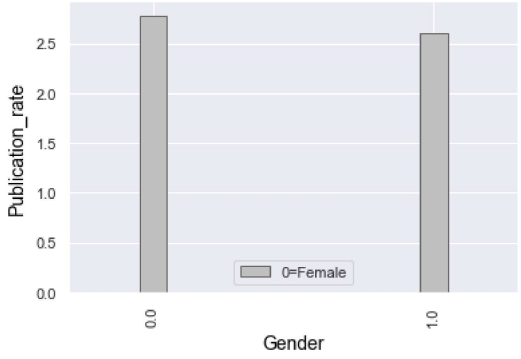
In [108]:
```
Surgery.groupby('Gender').full_professor.mean().plot(kind='bar',width=0.2,color='silver',edgecolor='dimgray',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='black')
plt.ylabel('Assistant', size=14,labelpad=6,color='black')
plt.legend(['0=Female'])
plt.show()
```



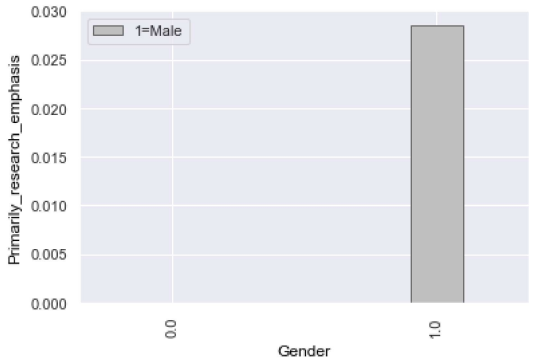

In [109]:
```
Surgery.groupby('Board_certified').Associate.mean().plot(kind='bar',width=0.2,color='silver',edgecolor='dimgray',figsize=(6,4))
plt.xlabel('Board_certified',size=14,labelpad=6,color='black')
plt.ylabel('Associate', size=14,labelpad=6,color='black')
plt.legend(['1=Board_certified'])
plt.show()
```



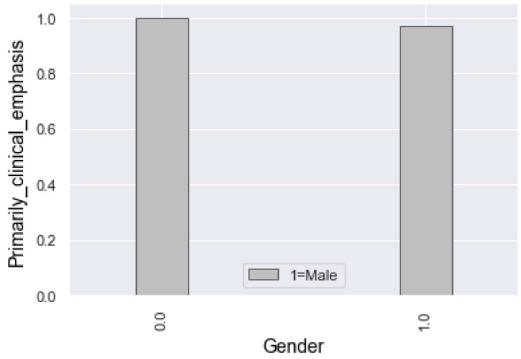

In [110]:
```
Surgery.groupby('Gender').Prate.mean().plot(kind='bar',width=0.1,color='silver',edgecolor='dimgray',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='black')
plt.ylabel('Publication_rate', size=14,labelpad=6,color='black')
plt.legend(['0=Female'])
plt.show()
```



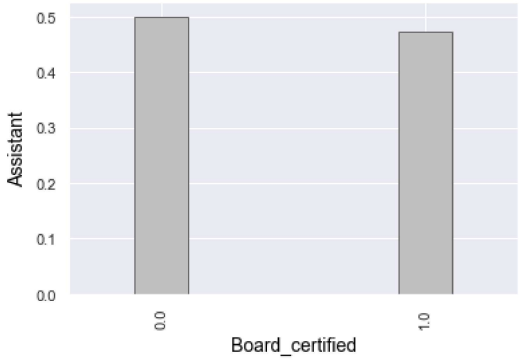
In [111]:
```
Surgery.groupby('Gender').Primarily_research_emphasis.mean().plot(kind='bar',width=0.2,color='silver',edgecolor='dimgray',figsiz
plt.xlabel('Gender',size=12,labelpad=6,color='black')
plt.ylabel('Primarily_research_emphasis',size=12,labelpad=6,color='black')
plt.legend(['1=Male','0=female'])
plt.show()
```




In [112]:
```
Surgery.groupby('Gender').Primarily_clinical_emphasis.mean().plot(kind='bar',width=0.2,color='silver',edgecolor='dimgray',figsiz
plt.xlabel('Gender',size=14,labelpad=6,color='black')
plt.ylabel('Primarily_clinical_emphasis',size=14,labelpad=6,color='black')
plt.legend(['1=Male','0=female'])
plt.show()
```




In [113]:
```
Surgery.groupby('Board_certified').Assistant.mean().plot(kind='bar',width=0.2,color='silver',edgecolor='dimgray',figsize=(6,4))
plt.xlabel('Board_certified',size=14,labelpad=6,color='black')
plt.ylabel('Assistant', size=14,labelpad=6,color='black')
plt.show()
```



In [114]:
```
Surgery.groupby('Gender').full_professor.mean().plot(kind='bar',width=0.2,color='silver',edgecolor='dimgray',figsize=(6,4))
plt.xlabel('Gender',size=14,labelpad=6,color='black')
plt.ylabel('Assistant', size=14,labelpad=6,color='black')
plt.legend(['1=Male'])
plt.show()
```