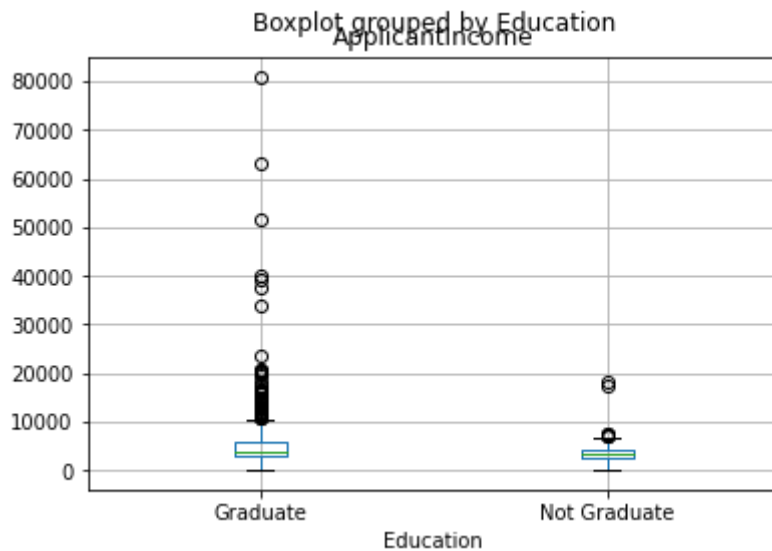
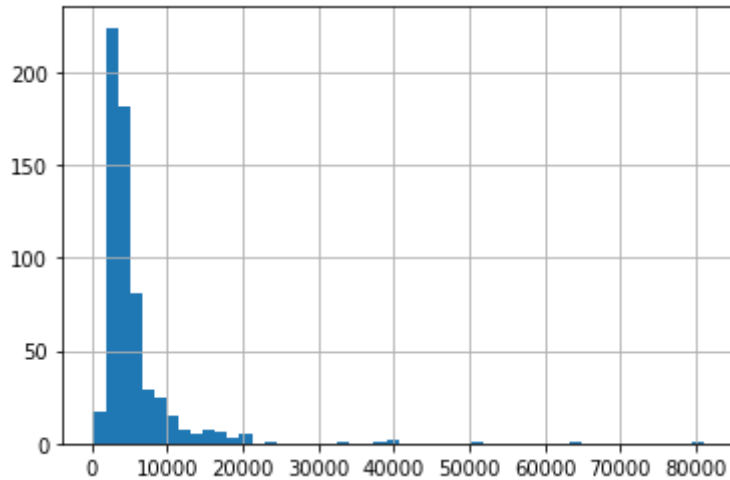


## Visualize Applicant Income

```
In [3]: df['ApplicantIncome'].hist(bins=50)  
df.boxplot(column='ApplicantIncome', by = 'Education')
```

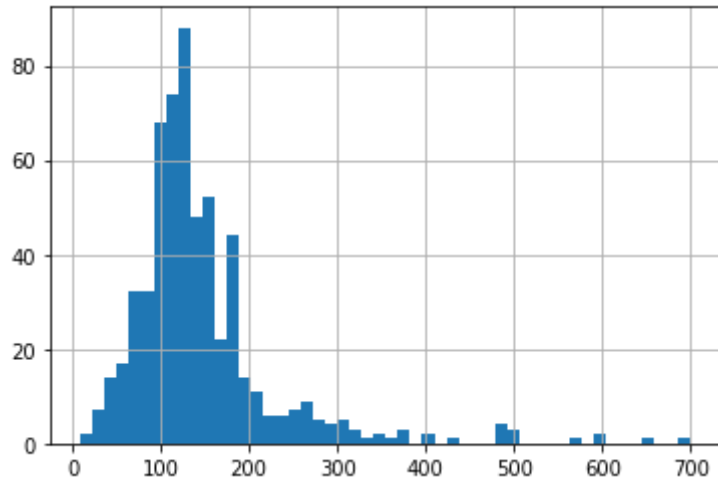
```
Out[3]: <matplotlib.axes._subplots.AxesSubplot at 0x7f41c656d7b8>
```



## Visualize Loan Amount

```
In [4]: df['LoanAmount'].hist(bins=50)
```

```
Out[4]: <matplotlib.axes._subplots.AxesSubplot at 0x7f41c5fc1320>
```



## Data Manipulation

```
In [5]: temp1 = df['Credit_History'].value_counts(ascending=True)
temp2 = df.pivot_table(values='Loan_Status',index=['Credit_History'],aggfunc=lambda x: x.map({'Y':1,'N':0}).mean())
print('Frequency Table for Credit History:')
print(temp1)

print('\nProbability of getting loan for each Credit History class:')
print(temp2)

temp3 = df['Married'].value_counts(ascending=True)
temp4 = df.pivot_table(values='Loan_Status',index=['Married'],aggfunc=lambda x: x.map({'Y':1,'N':0}).mean())
print('Frequency Table for Married History:')
print(temp3)
```

Frequency Table for Credit History:

0.0	89
-----	----

1.0	475
-----	-----

Name: Credit\_History, dtype: int64

Probability of getting loan for each Credit History class:

	Loan_Status
--	-------------

Credit_History	
----------------	--

0.0	0.078652
-----	----------

1.0	0.795789
-----	----------

Frequency Table for Married History:

No	213
----	-----

Yes	398
-----	-----

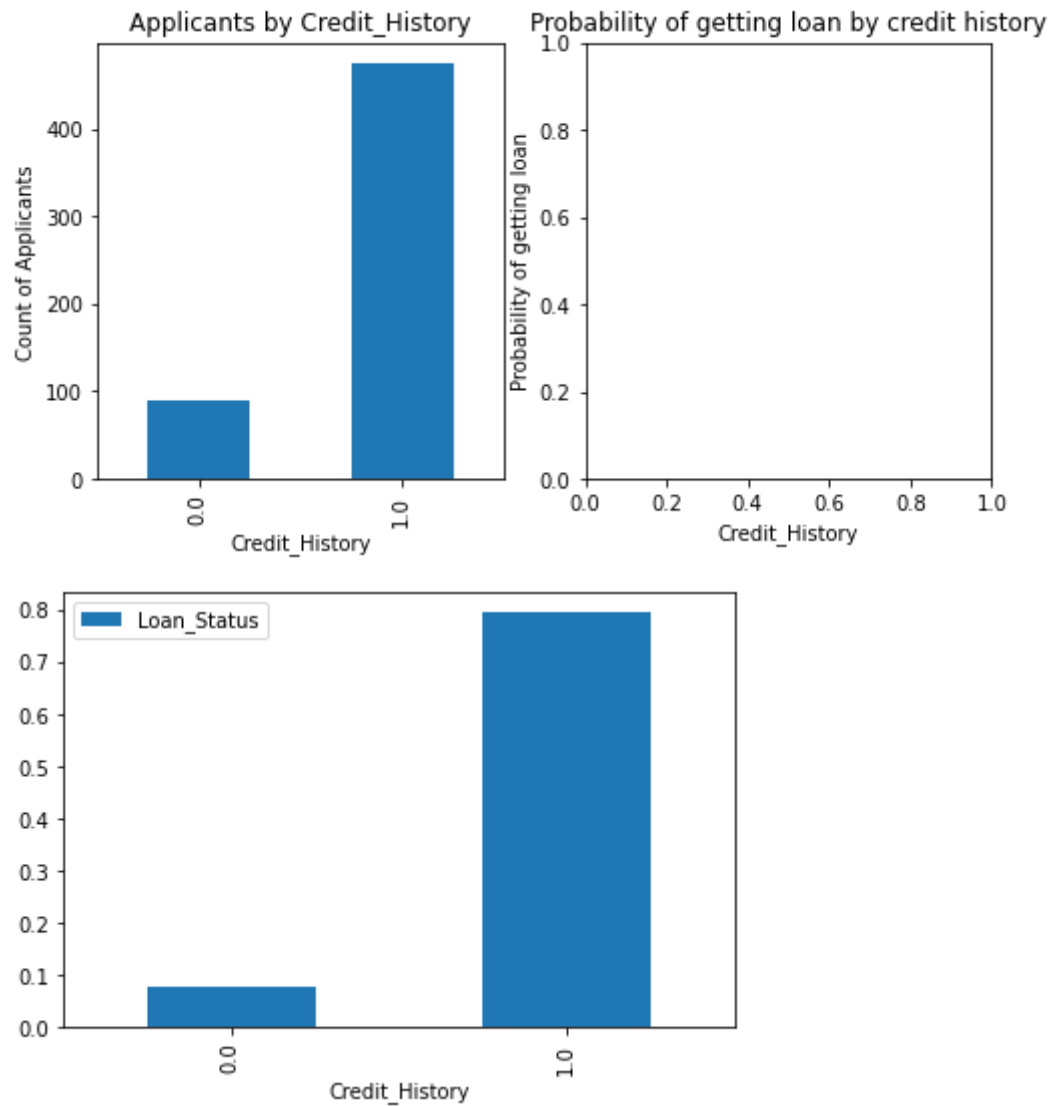
Name: Married, dtype: int64

## Visualize chances of getting loan

```
In [6]: fig = plt.figure(figsize=(8,4))
ax1 = fig.add_subplot(121)
ax1.set_xlabel('Credit_History')
ax1.set_ylabel('Count of Applicants')
ax1.set_title("Applicants by Credit_History")
temp1.plot(kind='bar')

ax2 = fig.add_subplot(122)
temp2.plot(kind = 'bar')
ax2.set_xlabel('Credit_History')
ax2.set_ylabel('Probability of getting loan')
ax2.set_title("Probability of getting loan by credit history")
```

Out[6]: Text(0.5, 1.0, 'Probability of getting loan by credit history')

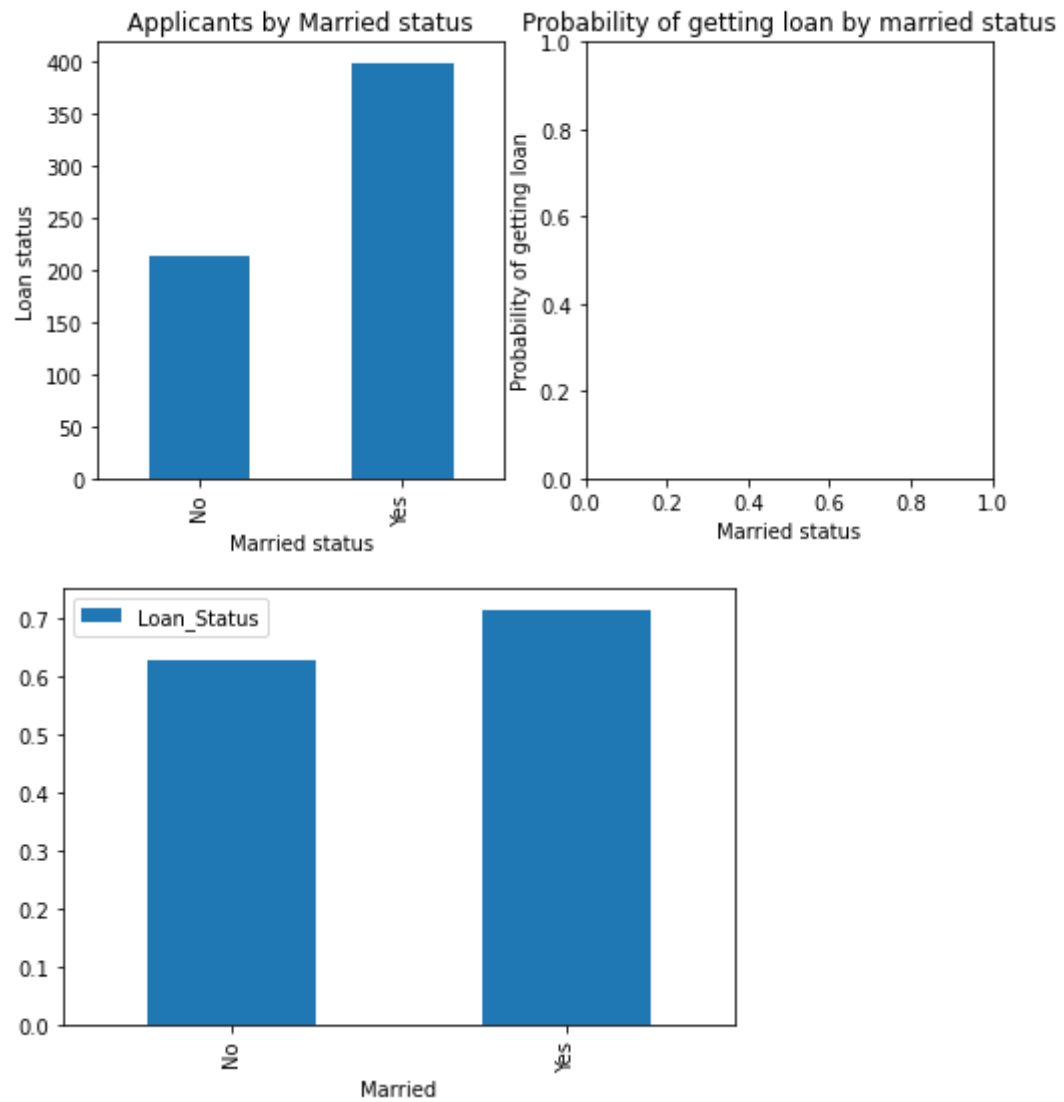


**Visualize chances of getting loan -- Married**

```
In [7]: fig = plt.figure(figsize=(8,4))
ax1 = fig.add_subplot(121)
ax1.set_xlabel('Married status')
ax1.set_ylabel('Loan status')
ax1.set_title("Applicants by Married status")
temp3.plot(kind='bar')

ax2 = fig.add_subplot(122)
temp4.plot(kind = 'bar')
ax2.set_xlabel('Married status')
ax2.set_ylabel('Probability of getting loan')
ax2.set_title("Probability of getting loan by married status")
```

Out[7]: Text(0.5, 1.0, 'Probability of getting loan by married status')



**Check missing values**

```
In [8]: df.apply(lambda x: sum(x.isnull()), axis=0)
```

```
Out[8]: Loan_ID          0  
Gender          13  
Married         3  
Dependents      15  
Education       0  
Self_Employed  32  
ApplicantIncome  0  
CoapplicantIncome  0  
LoanAmount      22  
Loan_Amount_Term 14  
Credit_History  50  
Property_Area    0  
Loan_Status      0  
dtype: int64
```

```
In [ ]:
```