

In [2]:

```
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
import seaborn as sns
%matplotlib inline
```

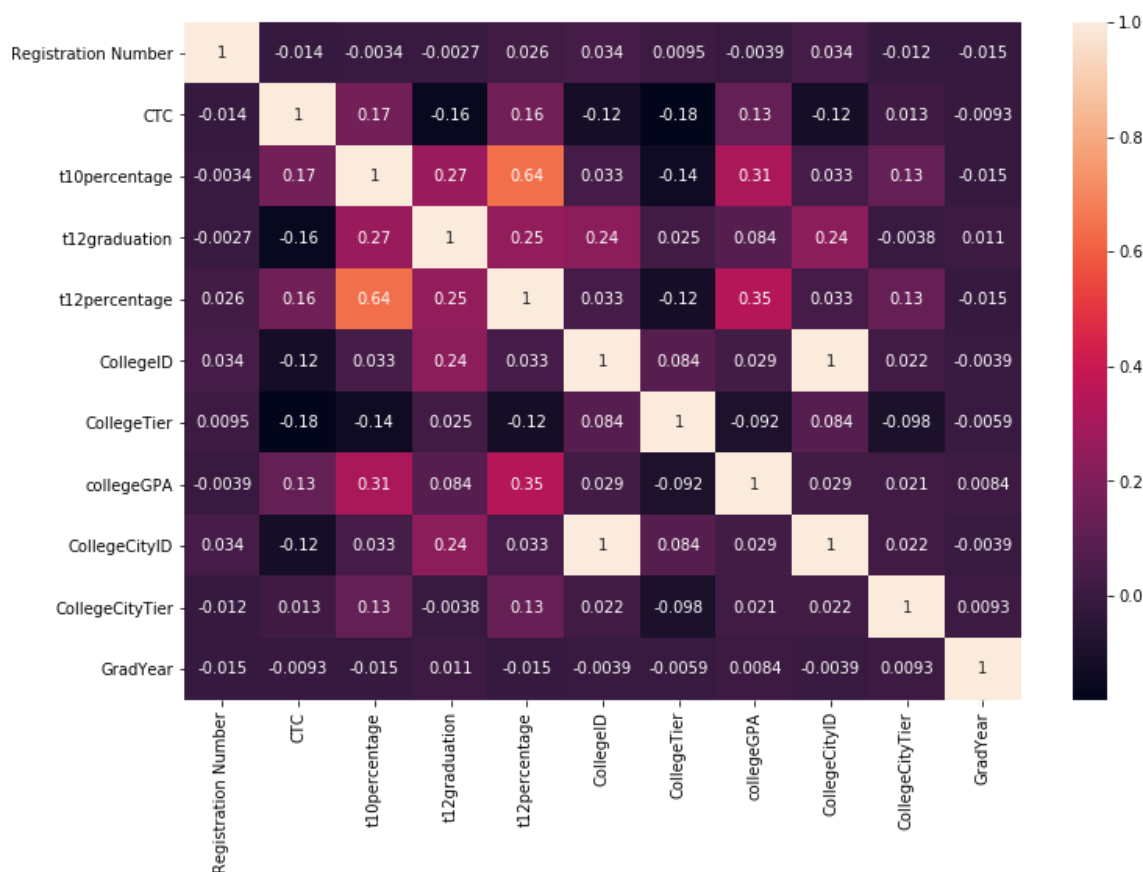
In [3]:

```
sal=pd.read_csv('G:\Data Analysis\Salary Data_Sharath.csv')
pt=pd.read_csv('G:\Data Analysis\Salary Data_Sharath.csv')
```

In [3]:

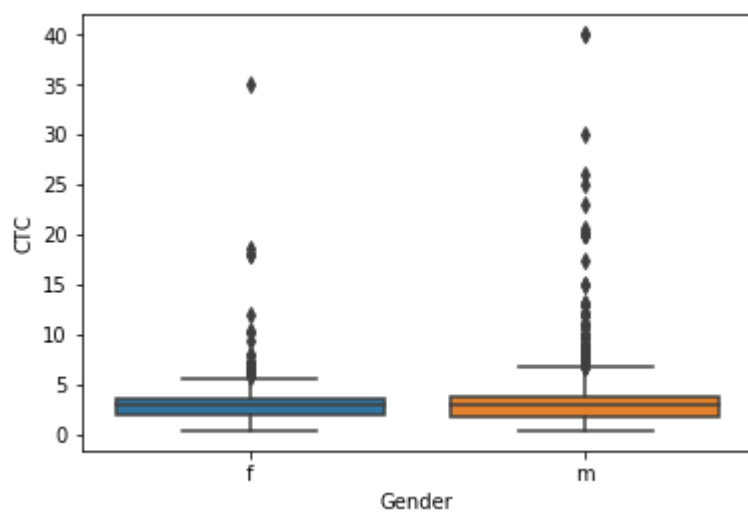
```
plt.figure(figsize=(12, 8))

sal_corr = sal.corr()
sns.heatmap(sal_corr,
            xticklabels = sal_corr.columns.values,
            yticklabels = sal_corr.columns.values,
            annot = True);
```



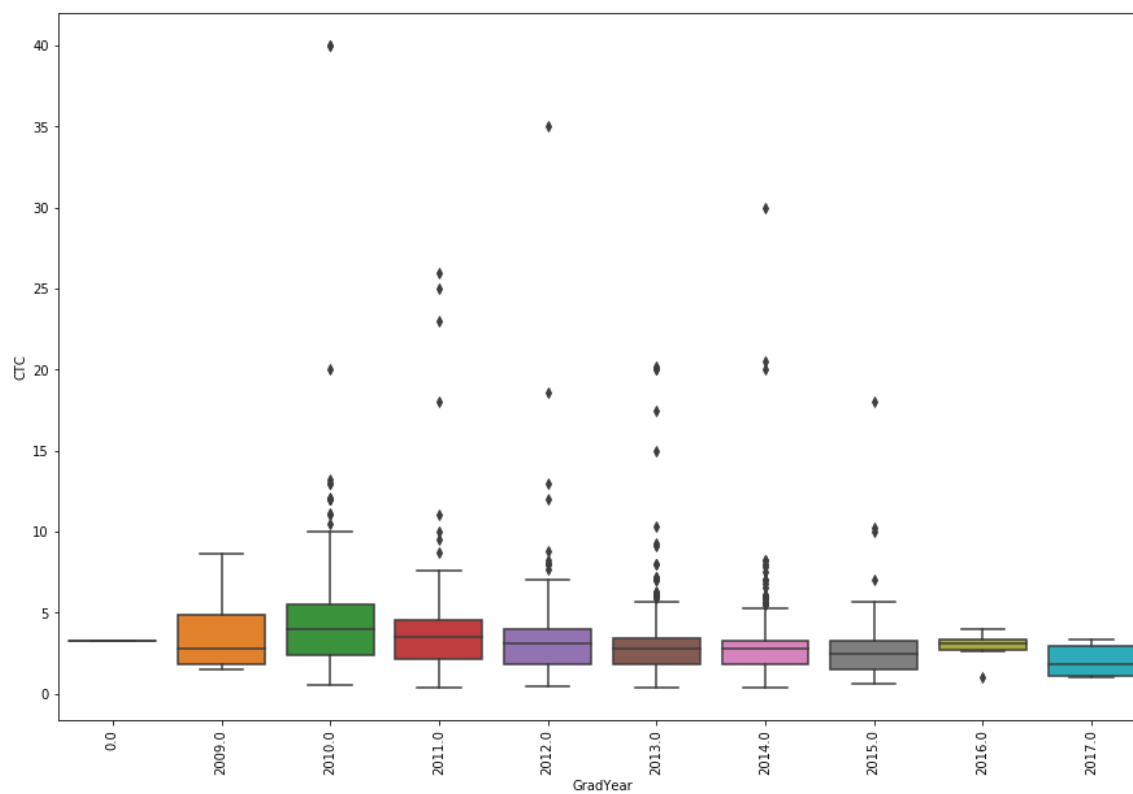
In [22]:

```
sns.boxplot(x='Gender',y='CTC',data=sal )  
plt.show()
```



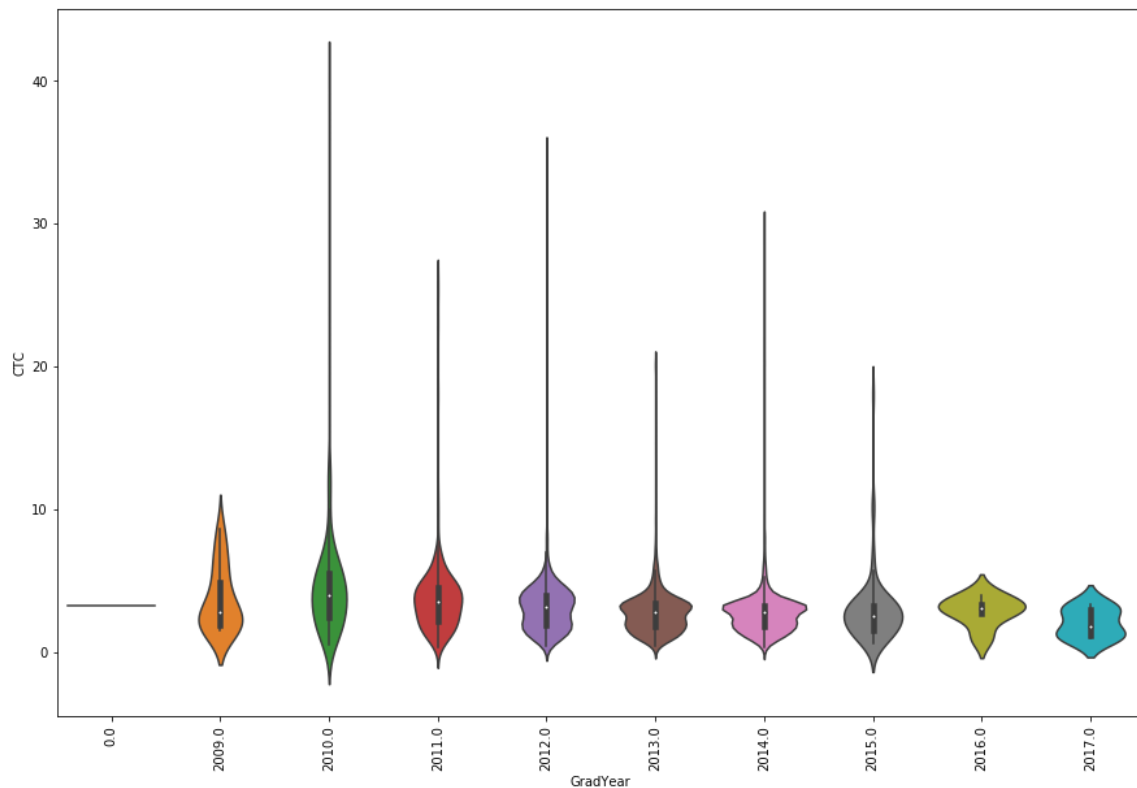
In [32]:

```
plt.figure(figsize=(15, 10))  
sns.boxplot(x='GradYear',y='CTC',data=sal )  
plt.xticks(rotation=90)  
plt.show()
```



In [34]:

```
plt.figure(figsize=(15, 10))
sns.violinplot(x='GradYear',y='CTC',data=sal )
plt.xticks(rotation=90)
plt.show()
```



In [8]:

```
sal=pd.read_csv('G:\Data Analysis\Salary Data_Sharath.csv')
pt=pd.read_csv('G:\Data Analysis\ptd.csv')
pt = pt.dropna(axis=1)
merged = sal.merge(pt, on='Registration Number')
merged.to_csv("G:\Data Analysis\output.csv", index=False)
```

In [9]:

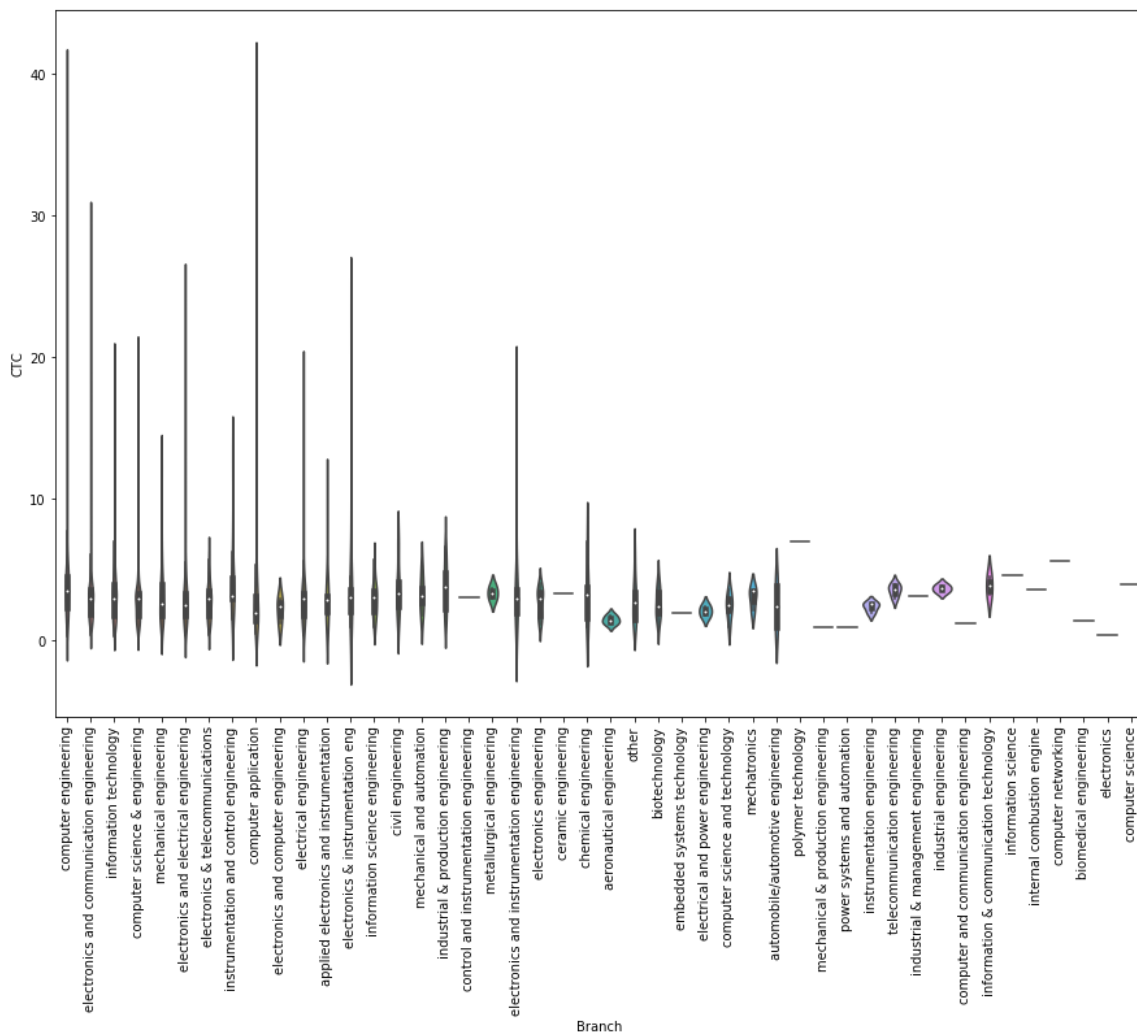
```
sal=pd.read_csv('G:\Data Analysis\output.csv')
pt=pd.read_csv('G:\Data Analysis\ent.csv')
pt = pt.dropna(axis=1)
merged = sal.merge(pt, on='Registration Number')
merged.to_csv("G:\Data Analysis\output.csv", index=False)
```

In [10]:

```
sal=pd.read_csv('G:\Data Analysis\output.csv')
```

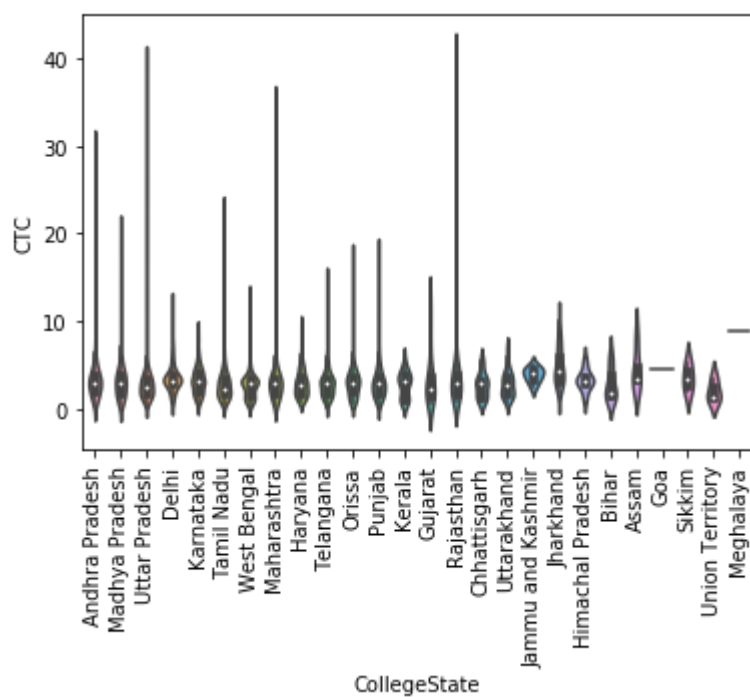
In [14]:

```
plt.figure(figsize=(15, 10))
sns.violinplot(x='Branch',y='CTC',data=sal )
plt.xticks(rotation=90)
plt.show()
```



In [23]:

```
#plt.figure(figsize=(75, 20))
sns.violinplot(x='CollegeState',y='CTC',data=sal )
plt.xticks(rotation=90)
plt.show()
```



In [37]:

```
sal.groupby(['Branch', 'CTC'], as_index=False).mean().groupby('Branch').mean().CTC.sort_
_values(ascending=False).head(20)
```

Out[37]:

Branch	
polymer technology	7.000000
computer networking	5.650000
computer engineering	5.057661
information science	4.600000
instrumentation and control engineering	4.289286
electronics and communication engineering	4.256780
computer application	4.081757
computer science	4.000000
information technology	3.920283
electronics & instrumentation eng	3.893478
information & communication technology	3.875000
computer science & engineering	3.873936
mechanical engineering	3.779722
civil engineering	3.762500
industrial & production engineering	3.738889
industrial engineering	3.700000
electronics and instrumentation engineering	3.692105
internal combustion engine	3.600000
telecommunication engineering	3.575000
applied electronics and instrumentation	3.418750

Name: CTC, dtype: float64

In [36]:

```
sal.groupby(['Branch', 'CTC'], as_index=False).max().groupby('Branch').max().CTC.sort_v
alues(ascending=False).head(20)
```

Out[36]:

Branch	
computer engineering	40.00
computer application	40.00
electronics and communication engineering	30.00
electronics and electrical engineering	25.00
electronics & instrumentation eng	23.00
computer science & engineering	20.50
information technology	20.00
electrical engineering	18.60
electronics and instrumentation engineering	17.45
mechanical engineering	13.00
instrumentation and control engineering	13.00
applied electronics and instrumentation	9.50
civil engineering	7.20
polymer technology	7.00
chemical engineering	7.00
industrial & production engineering	6.60
electronics & telecommunications	6.30
other	6.00
information science engineering	5.70
computer networking	5.65

Name: CTC, dtype: float64

In [35]:

```
sal.groupby(['Branch', 'CTC'], as_index=False).min().groupby('Branch').min().CTC.sort_values(ascending=False).head(20)
```

Out[35]:

Branch	
polymer technology	7.00
computer networking	5.65
information science	4.60
computer science	4.00
internal combustion engine	3.60
industrial engineering	3.50
ceramic engineering	3.35
information & communication technology	3.25
industrial & management engineering	3.20
control and instrumentation engineering	3.05
telecommunication engineering	3.00
metallurgical engineering	3.00
mechatronics	2.15
instrumentation engineering	2.00
embedded systems technology	2.00
electrical and power engineering	1.80
mechanical and automation	1.80
applied electronics and instrumentation	1.75
industrial & production engineering	1.70
instrumentation and control engineering	1.50

Name: CTC, dtype: float64

In []: