In [25]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [26]:

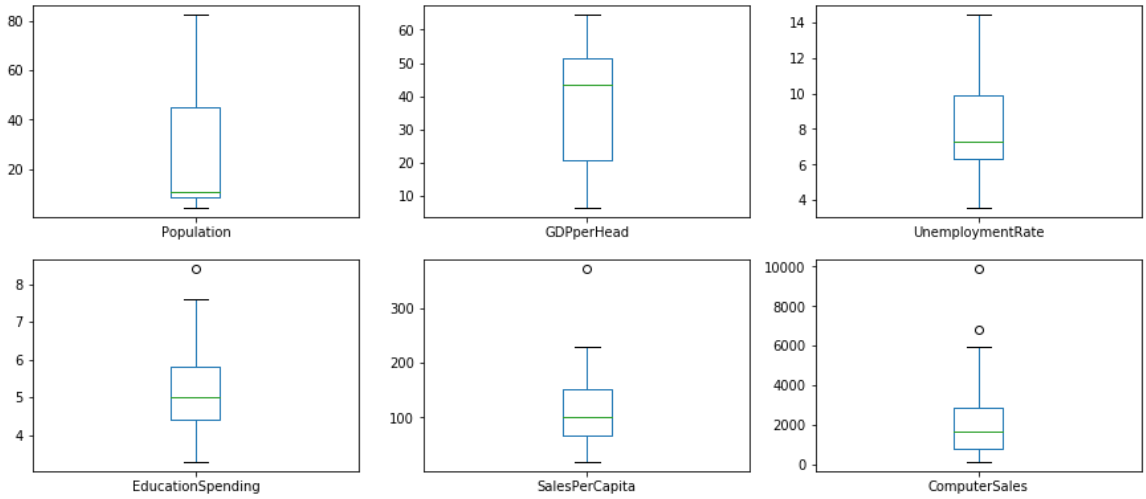
```
data = pd.read_csv("EuropeanSales.csv")
```

In [27]:

```
data1 = data.drop(["Country"],axis=1)
```

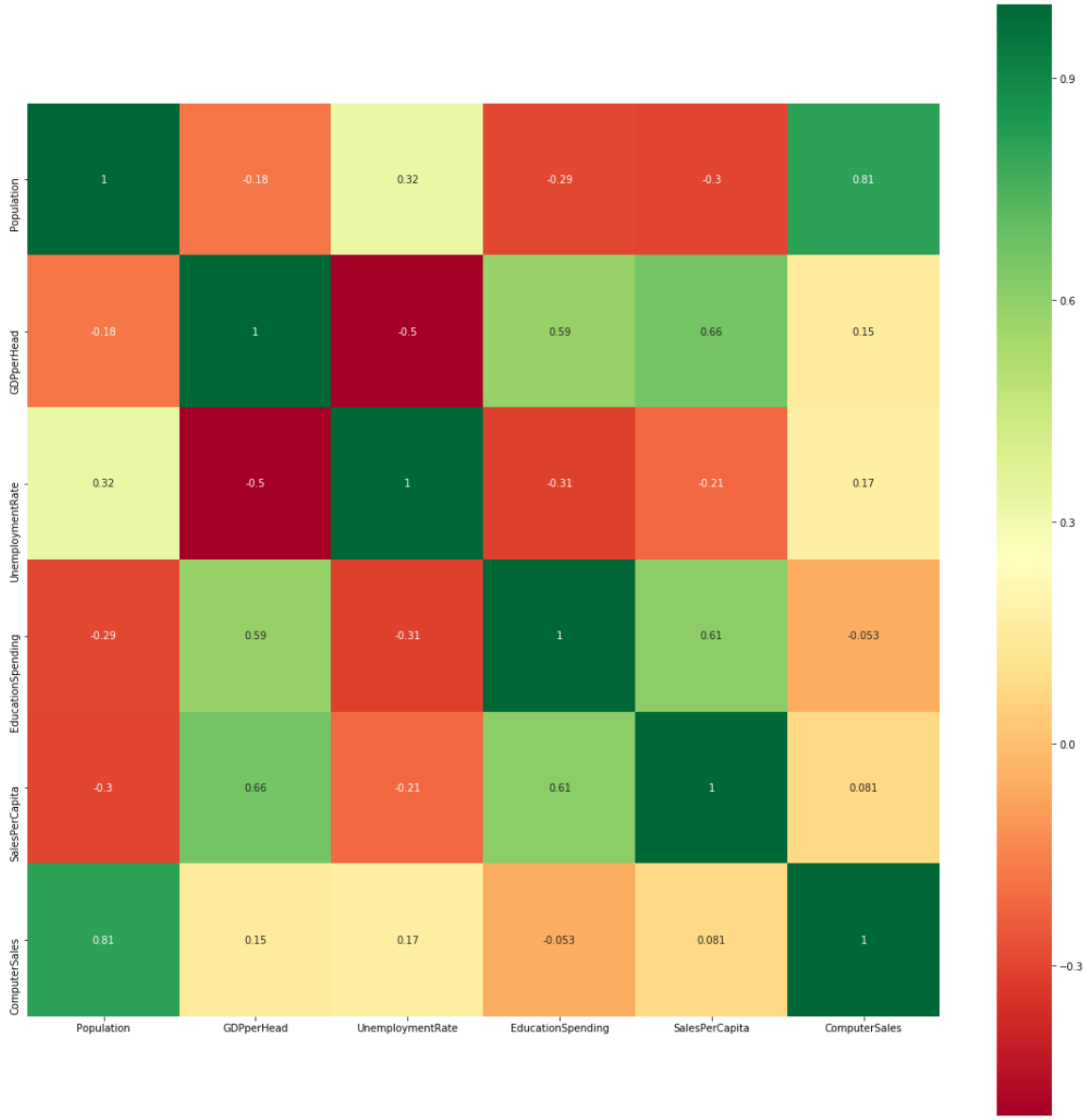
In [29]:

```
data1.plot(kind='box',figsize=(15,10),subplots=True,layout=(3,3))
plt.show()
```



In [30]:

```
plt.figure(figsize=(20,20))
p=sns.heatmap(data1.corr(), annot=True,cmap='RdYlGn',square=True)
```



In [36]:

```
data2 = data.drop(["Country","SalesPerCapita","SalesPerCapita","EducationSpending"],axis=1)
```

In [40]:

```
X = data2.drop('ComputerSales',axis=1)
y = data2['ComputerSales']
xc = st.add_constant(X)
lm = st.OLS(y,xc).fit()
```

```
C:\Users\User\Anaconda3\lib\site-packages\numpy\core\fromnumeric.py:2389:
FutureWarning: Method .ptp is deprecated and will be removed in a future v
ersion. Use numpy.ptp instead.
  return ptp(axis=axis, out=out, **kwargs)
```

In [41]:

```
lm.summary()
```

Out[41]:

OLS Regression Results

Dep. Variable:	ComputerSales	**R-squared:**	0.745
Model:	OLS	**Adj. R-squared:**	0.700
Method:	Least Squares	**F-statistic:**	16.53
Date:	Mon, 24 Feb 2020	**Prob (F-statistic):**	2.75e-05
Time:	22:14:47	**Log-Likelihood:**	-179.25
No. Observations:	21	**AIC:**	366.5
Df Residuals:	17	**BIC:**	370.7
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-1765.3933	1442.889	-1.224	0.238	-4809.623	1278.836
Population	80.7945	12.381	6.526	0.000	54.673	106.916
GDPperHead	44.7542	18.995	2.356	0.031	4.678	84.830
UnemploymentRate	54.8559	117.952	0.465	0.648	-194.000	303.712

Omnibus:	23.114	**Durbin-Watson:**	2.135
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	41.691
Skew:	1.783	**Prob(JB):**	8.85e-10
Kurtosis:	8.910	**Cond. No.**	242.

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

In [42]:

```
data3 = data.drop(["Country","ComputerSales","UnemploymentRate","Population"],axis=1)
```

In [43]:

```
X = data3.drop('SalesPerCapita',axis=1)
y = data3['SalesPerCapita']
xc = st.add_constant(X)
lm = st.OLS(y,xc).fit()
```

In [44]:

```
lm.summary()
```

Out[44]:

OLS Regression Results

Dep. Variable:	SalesPerCapita	**R-squared:**	0.514
Model:	OLS	**Adj. R-squared:**	0.460
Method:	Least Squares	**F-statistic:**	9.522
Date:	Mon, 24 Feb 2020	**Prob (F-statistic):**	0.00151
Time:	22:17:42	**Log-Likelihood:**	-113.46
No. Observations:	21	**AIC:**	232.9
Df Residuals:	18	**BIC:**	236.0
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-66.5294	54.849	-1.213	0.241	-181.762	48.703
GDPperHead	1.9555	0.861	2.272	0.036	0.147	3.764
EducationSpending	21.1470	12.645	1.672	0.112	-5.418	47.712

Omnibus:	31.301	**Durbin-Watson:**	2.671
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	68.425
Skew:	2.494	**Prob(JB):**	1.39e-15
Kurtosis:	10.302	**Cond. No.**	183.

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

In [ ]:

```

```