

```
In [19]: from matplotlib.colors import ListedColormap
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
import numpy as np
```

```
In [2]: HR = pd.read_csv(r"C:\Users\Admin\Downloads\Project 3 - HR Analytics-20230916T131952Z-001\Project 3 - HR Analytics\HR-Employee-At...
```

In [3]: HR

Out[3]:

	S/N	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EmployeeCount	...	RelationshipSatisfac
0	1	41	Yes	Travel_Rarely	1102	Sales		1	2	Life Sciences	1	...
1	2	49	No	Travel_Frequently	279	Research & Development		8	1	Life Sciences	1	...
2	3	37	Yes	Travel_Rarely	1373	Research & Development		2	2	Other	1	...
3	4	33	No	Travel_Frequently	1392	Research & Development		3	4	Life Sciences	1	...
4	5	27	No	Travel_Rarely	591	Research & Development		2	1	Medical	1	...
...	...	...	...	...	...	...	...	...	...	...	...	...
1465	1466	36	No	Travel_Frequently	884	Research & Development		23	2	Medical	1	...
1466	1467	39	No	Travel_Rarely	613	Research & Development		6	1	Medical	1	...

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

Out[4]:
```

	S/N	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EmployeeCount	...	RelationshipSatisfaction	St
0	1	41	Yes	Travel_Rarely	1102	Sales		1	2	Life Sciences	1	...	1
1	2	49	No	Travel_Frequently	279	Research & Development		8	1	Life Sciences	1	...	4
2	3	37	Yes	Travel_Rarely	1373	Research & Development		2	2	Other	1	...	2
3	4	33	No	Travel_Frequently	1392	Research & Development		3	4	Life Sciences	1	...	3
4	5	27	No	Travel_Rarely	591	Research & Development		2	1	Medical	1	...	4

5 rows x 36 columns

```
In [5]: missing_data = HR.isna().sum()
```

```
In [6]: HR.drop('EmployeeCount', axis=1, inplace=True)
```

```
In [7]: HR.drop('StandardHours', axis=1, inplace=True)
```

```
In [8]: HR.drop('S/N', axis=1, inplace=True)
```

```
In [9]: HR
```

Out[9]:

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EmployeeNumber	EnvironmentSatisfaction	...	Pe
0	41	Yes	Travel_Rarely	1102	Sales		1	2	Life Sciences	1	2	...
1	49	No	Travel_Frequently	279	Research & Development		8	1	Life Sciences	2	3	...
2	37	Yes	Travel_Rarely	1373	Research & Development		2	2	Other	4	4	...
3	33	No	Travel_Frequently	1392	Research & Development		3	4	Life Sciences	5	4	...
4	27	No	Travel_Rarely	591	Research & Development		2	1	Medical	7	1	...
...	...	...	...	...	...	...	...	...	...	...	...	...
1465	36	No	Travel_Frequently	884	Research & Development		23	2	Medical	2061	3	...
1466	39	No	Travel_Rarely	613	Research & Development		6	1	Medical	2062	4	...
1467	27	No	Travel_Rarely	155	Research & Development		4	3	Life Sciences	2064	2	...
1468	49	No	Travel_Frequently	1023	Sales		2	3	Medical	2065	4	...
1469	34	No	Travel_Rarely	628	Research & Development		8	3	Medical	2068	2	...

1470 rows x 33 columns

```
In [10]: HR.corr()
```

Out[10]:

	Age	DailyRate	DistanceFromHome	Education	EmployeeNumber	EnvironmentSatisfaction	HourlyRate	JobInvolvement	JobLevel
Age	1.000000	0.010661	-0.001686	0.208034	-0.010145	0.010146	0.024287	0.029820	0.50960
DailyRate	0.010661	1.000000	-0.004985	-0.016806	-0.050990	0.018355	0.023381	0.046135	0.00296
DistanceFromHome	-0.001686	-0.004985	1.000000	0.021042	0.032916	-0.016075	0.031131	0.008783	0.00530
Education	0.208034	-0.016806	0.021042	1.000000	0.042070	-0.027128	0.016775	0.042438	0.10158
EmployeeNumber	-0.010145	-0.050990	0.032916	0.042070	1.000000	0.017621	0.035179	-0.006888	-0.01851
EnvironmentSatisfaction	0.010146	0.018355	-0.016075	-0.027128	0.017621	1.000000	-0.049857	-0.008278	0.00121
HourlyRate	0.024287	0.023381	0.031131	0.016775	0.035179	-0.049857	1.000000	0.042861	-0.02785
JobInvolvement	0.029820	0.046135	0.008783	0.042438	-0.006888	-0.008278	0.042861	1.000000	-0.01263
JobLevel	0.509604	0.002966	0.005303	0.101589	-0.018519	0.001212	-0.027853	-0.012630	1.00000
JobSatisfaction	-0.004892	0.030571	-0.003669	-0.011296	-0.046247	-0.006784	-0.071335	-0.021476	-0.00194
MonthlyIncome	0.497855	0.007707	-0.017014	0.094961	-0.014829	-0.006259	-0.015794	-0.015271	0.95030
MonthlyRate	0.028051	-0.032182	0.027473	-0.026084	0.012648	0.037600	-0.015297	-0.016322	0.03956
NumCompaniesWorked	0.299635	0.038153	-0.029251	0.126317	-0.001251	0.012594	0.022157	0.015012	0.14250
PercentSalaryHike	0.003634	0.022704	0.040235	-0.011111	-0.012944	-0.031701	-0.009062	-0.017205	-0.03473
PerformanceRating	0.001904	0.000473	0.027110	-0.024539	-0.020359	-0.029548	-0.002172	-0.029071	-0.02122
RelationshipSatisfaction	0.053535	0.007846	0.006557	-0.009118	-0.069861	0.007665	0.001330	0.034297	0.02164
StockOptionLevel	0.037510	0.042143	0.044872	0.018422	0.062227	0.003432	0.050263	0.021523	0.01398
TotalWorkingYears	0.680381	0.014515	0.004628	0.148280	-0.014365	-0.002693	-0.002334	-0.005533	0.78220
TrainingTimesLastYear	-0.019621	0.002453	-0.036942	-0.025100	0.023603	-0.019359	-0.008548	-0.015338	-0.01819
WorkLifeBalance	-0.021490	-0.037848	-0.026556	0.009819	0.010309	0.027627	-0.004607	-0.014617	0.03781
YearsAtCompany	0.311309	-0.034055	0.009508	0.069114	-0.011240	0.001458	-0.019582	-0.021355	0.53473
YearsInCurrentRole	0.212901	0.009932	0.018845	0.060236	-0.008416	0.018007	-0.024106	0.008717	0.38944
YearsSinceLastPromotion	0.216513	-0.033229	0.010029	0.054254	-0.009019	0.016194	-0.026716	-0.024184	0.35388
YearsWithCurrManager	0.202089	-0.026363	0.014406	0.069065	-0.009197	-0.004999	-0.020123	0.025976	0.37528

24 rows x 24 columns

```
In [ ]: 
```

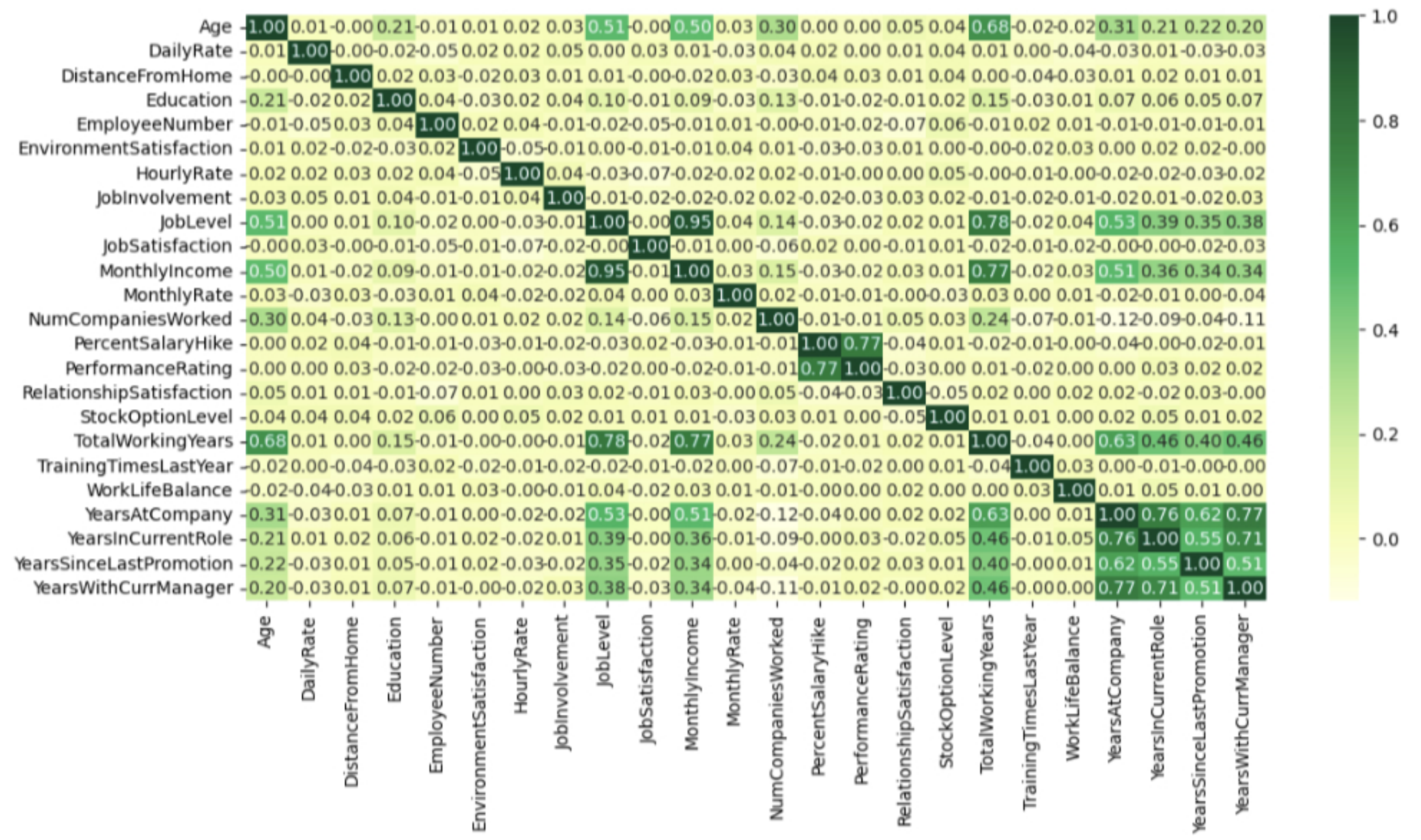
```
In [20]: sns.color_palette()
```



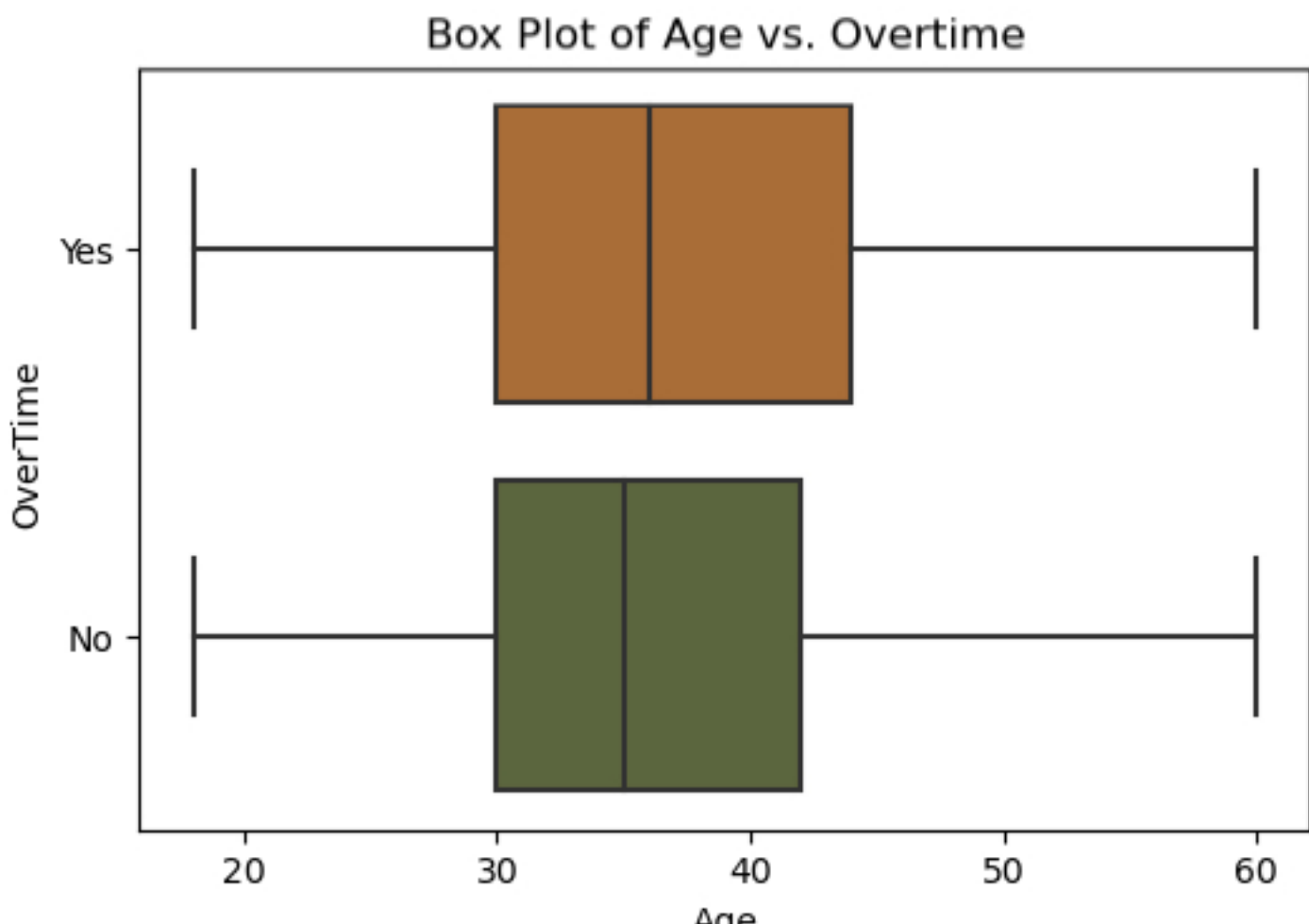
```
In [21]: gyr = ['#606c38', '#bc6c25']
sns.palplot(sns.color_palette(gyr))
```



```
In [43]: plt.figure(figsize=(13, 6))
sns.heatmap(HR.corr(), cmap='YlGn', annot=True, fmt='.2f')
plt.show()
```



```
In [57]: plt.figure(figsize=(6, 4))
custom_palette = ['#bc6c25', '#606c38']
sns.set_palette(custom_palette)
sns.boxplot(x='Age', y='OverTime', data=HR)
plt.title('Box Plot of Age vs. OverTime')
plt.show()
```



```
In [ ]: 
```

```
In [ ]: 
```