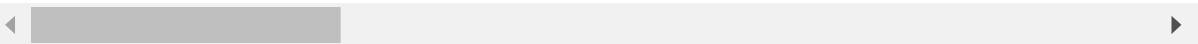In [2]:

```python
import pandas as pd
df = pd.read_csv('E:\EmployeeAttrition.csv')
print(df.shape)
df.head(5)
```

(1470, 35)

Out[2]:

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	Edu
0	41	Yes	Travel_Rarely	1102	Sales	1	2	L
1	49	No	Travel_Frequently	279	Research & Development	8	1	L
2	37	Yes	Travel_Rarely	1373	Research & Development	2	2	
3	33	No	Travel_Frequently	1392	Research & Development	3	4	L
4	27	No	Travel_Rarely	591	Research & Development	2	1	

5 rows × 35 columns

In [3]:

```python
df.describe()
```

Out[3]:

	Age	DailyRate	DistanceFromHome	Education	EmployeeCount	EmployeeN
count	1470.000000	1470.000000	1470.000000	1470.000000	1470.0	1470.0
mean	36.923810	802.485714	9.192517	2.912925	1.0	1024.8
std	9.135373	403.509100	8.106864	1.024165	0.0	602.0
min	18.000000	102.000000	1.000000	1.000000	1.0	1.0
25%	30.000000	465.000000	2.000000	2.000000	1.0	491.2
50%	36.000000	802.000000	7.000000	3.000000	1.0	1020.5
75%	43.000000	1157.000000	14.000000	4.000000	1.0	1555.7
max	60.000000	1499.000000	29.000000	5.000000	1.0	2068.0

8 rows × 26 columns

In [4]:

```python
print(len(df))
df.isnull().sum()
```

1470

Out[4]:

```
Age                        0
Attrition                  0
BusinessTravel             0
DailyRate                  0
Department                 0
DistanceFromHome           0
Education                  0
EducationField             0
EmployeeCount              0
EmployeeNumber             0
EnvironmentSatisfaction    0
Gender                     0
HourlyRate                 0
JobInvolvement             0
JobLevel                   0
JobRole                    0
JobSatisfaction            0
MaritalStatus              0
MonthlyIncome              0
MonthlyRate                0
NumCompaniesWorked         0
Over18                     0
OverTime                   0
PercentSalaryHike          0
PerformanceRating          0
RelationshipSatisfaction   0
StandardHours              0
StockOptionLevel           0
TotalWorkingYears          0
TrainingTimesLastYear      0
WorkLifeBalance            0
YearsAtCompany             0
YearsInCurrentRole         0
YearsSinceLastPromotion    0
YearsWithCurrManager       0
dtype: int64
```
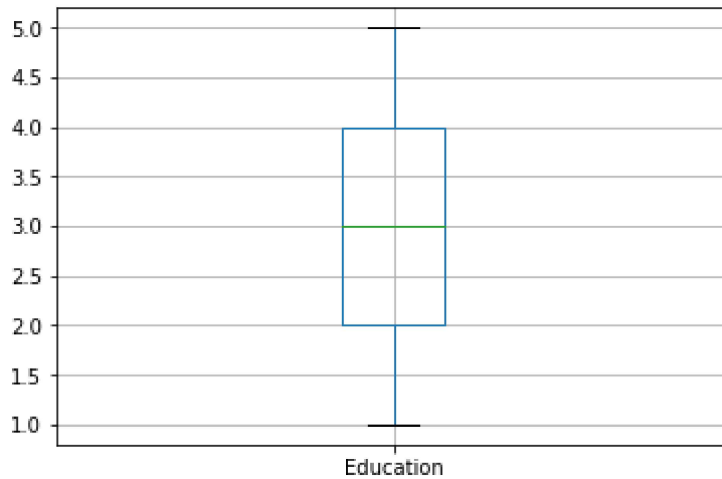
In [6]:

```python
import matplotlib.pyplot as plt
df.boxplot('Education')
```
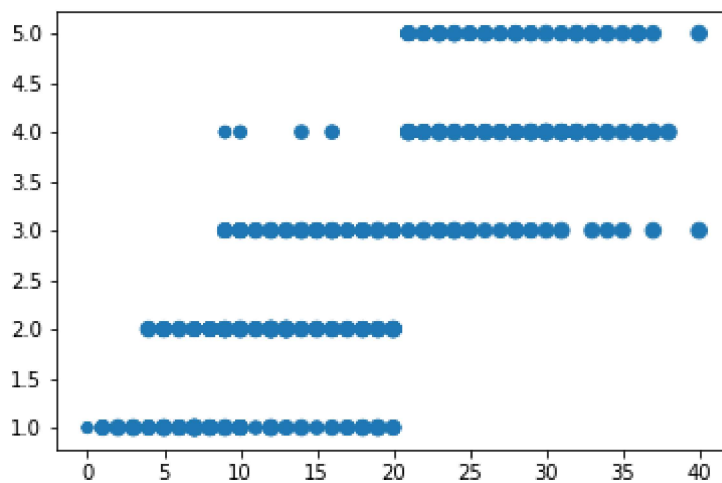
Out[6]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x21996ac3e80>
```


In [7]:

```python
plt.scatter(df.TotalWorkingYears,df.JobLevel,s=df.Age)
```

Out[7]:

```
<matplotlib.collections.PathCollection at 0x21997166a90>
```


In [ ]: