Name : Arul Kumar ARK

Roll No. : 225229103

Lab : 7

Data Visualization in Seaborn

In [3]:

```
import pandas as pd
```

In [4]:

```
df = pd.read_csv('train_upvote_mini.csv')
df.head()
```

Out[4]:

	ID	Tag	Reputation	Answers	Username	Views	Upvotes
0	52664	a	3942.0	2.0	155623	7855.0	42.0
1	327662	a	26046.0	12.0	21781	55801.0	1175.0
2	468453	c	1358.0	4.0	56177	8067.0	60.0
3	96996	a	264.0	3.0	168793	27064.0	9.0
4	131465	c	4271.0	4.0	112223	13986.0	83.0

In [5]:

```
df.shape
```

Out[5]:

```
(15440, 7)
```

In [6]:

```
df.dtypes
```

Out[6]:

```
ID              int64
Tag            object
Reputation    float64
Answers       float64
Username        int64
Views         float64
Upvotes       float64
dtype: object
```

In [7]:

```
df['Tag'].unique()
```

Out[7]:

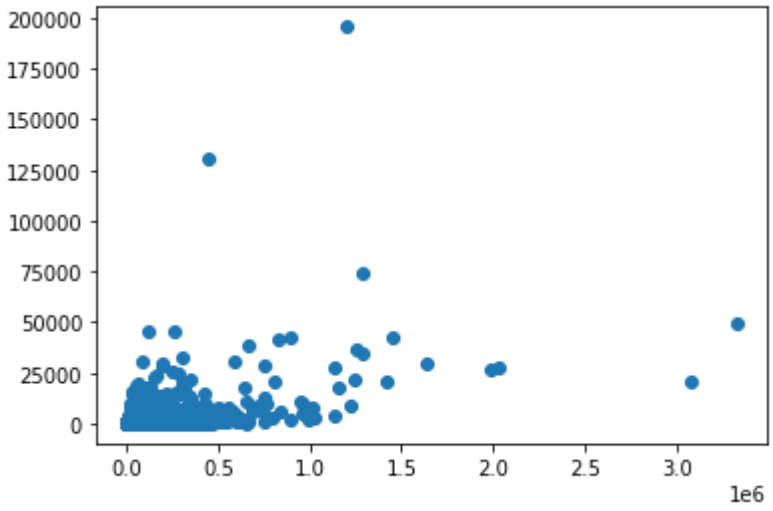
```
array(['a', 'c', 'r', 'j', 'p', 's', 'h', 'o', 'i', 'x'], dtype=object)
```

In [8]:

```
import matplotlib.pyplot as plt
plt.scatter(x=df['Views'], y=df['Upvotes'])
```

Out[8]:

```
<matplotlib.collections.PathCollection at 0x7f63c0b7a3d0>
```

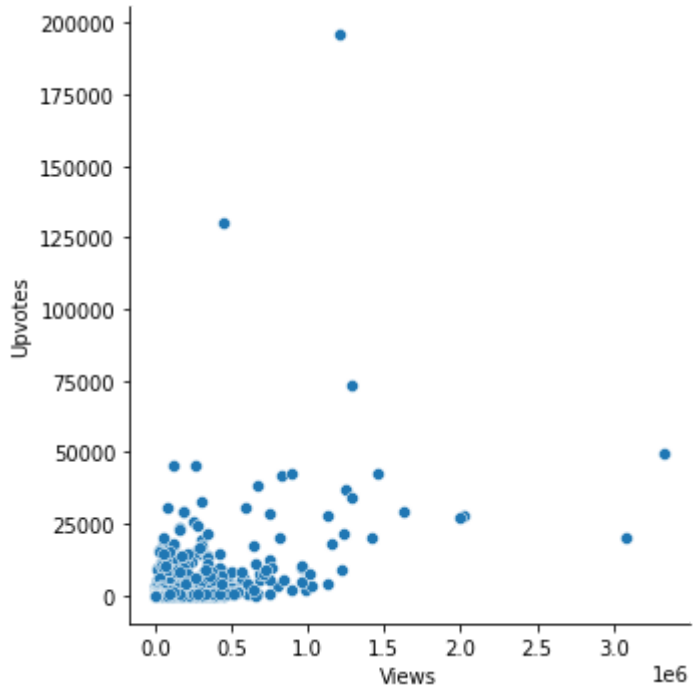


In [9]:

```
import seaborn as sns
sns.relplot(data=df,x="Views",y="Upvotes")
```

Out[9]:

```
<seaborn.axisgrid.FacetGrid at 0x7f63b6e28f40>
```

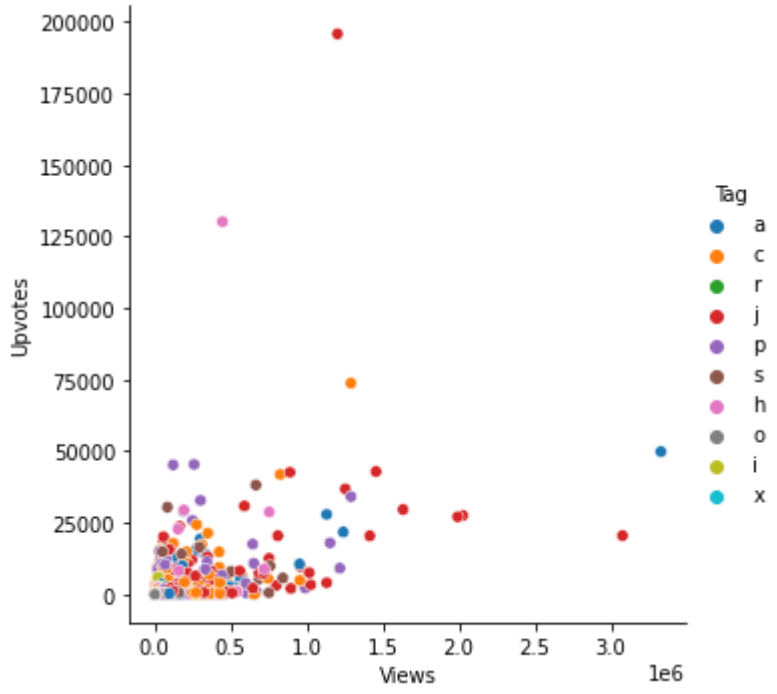


In [10]:

```
sns.relplot(data=df,x='Views',y='Upvotes',hue='Tag')
```

Out[10]:

```
<seaborn.axisgrid.FacetGrid at 0x7f63be33feb0>
```

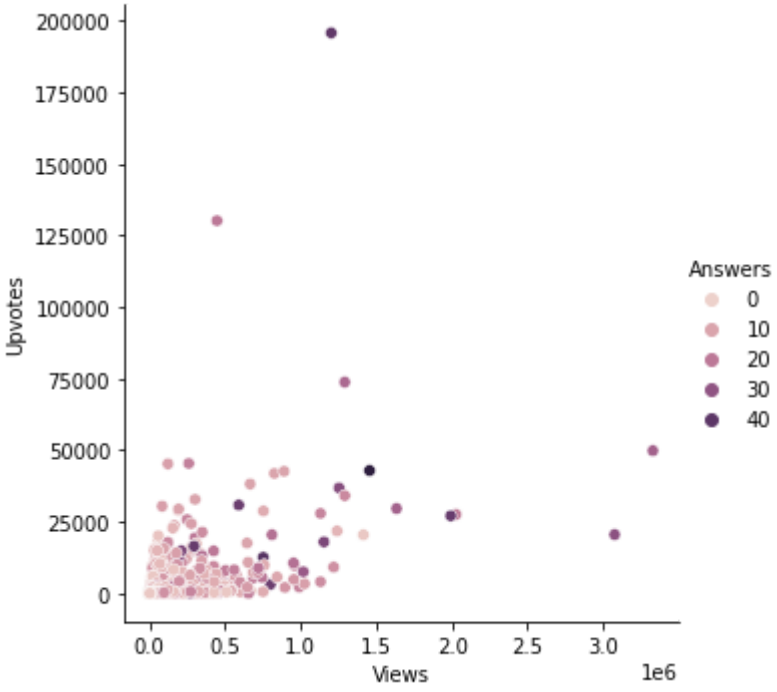


In [11]:

```
sns.relplot(data=df,x='Views',y='Upvotes',hue='Answers')
```

Out[11]:

```
<seaborn.axisgrid.FacetGrid at 0x7f63b6b44910>
```

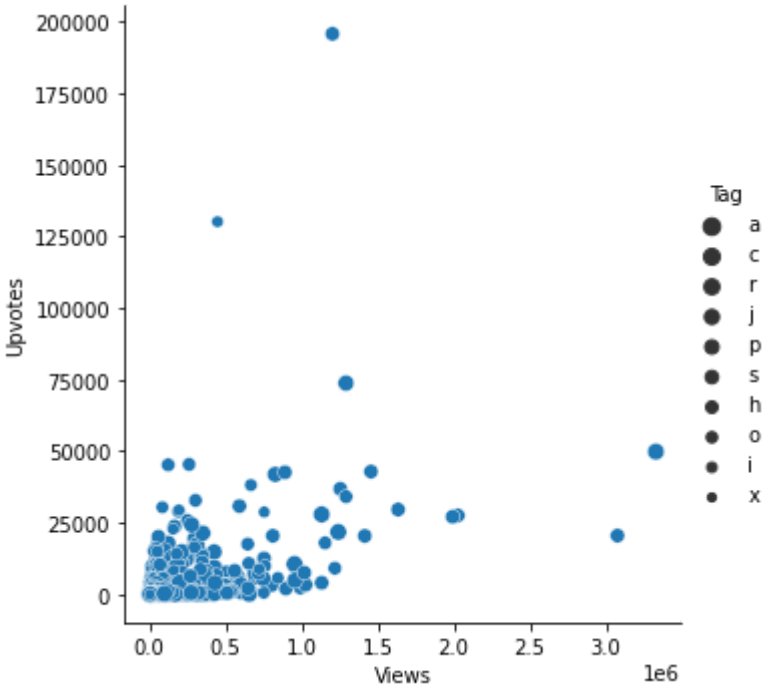


In [12]:

```
sns.relplot(data=df,x='Views',y='Upvotes',size='Tag')
```

Out[12]:

```
<seaborn.axisgrid.FacetGrid at 0x7f63b6b8e7c0>
```

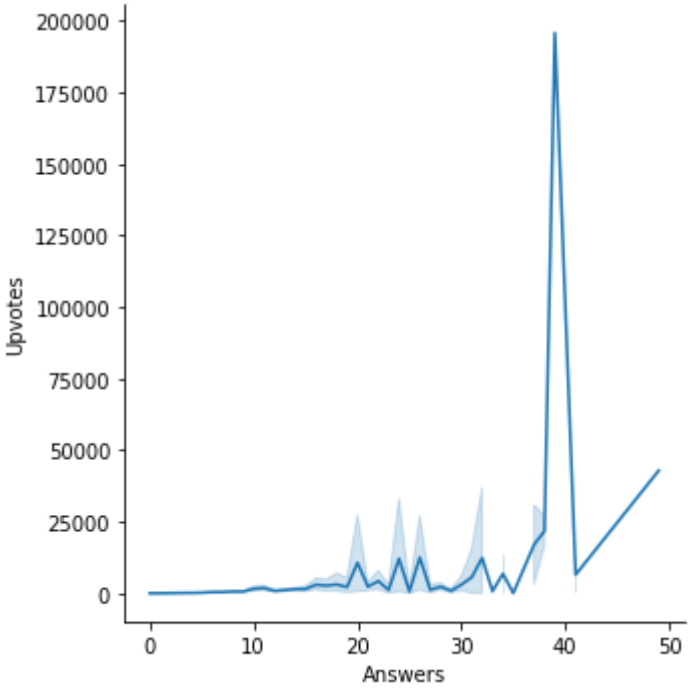


In [13]:

```
sns.relplot(data=df, x="Answers", y="Upvotes", kind='line')
```

Out[13]:

```
<seaborn.axisgrid.FacetGrid at 0x7f63b6aa32e0>
```

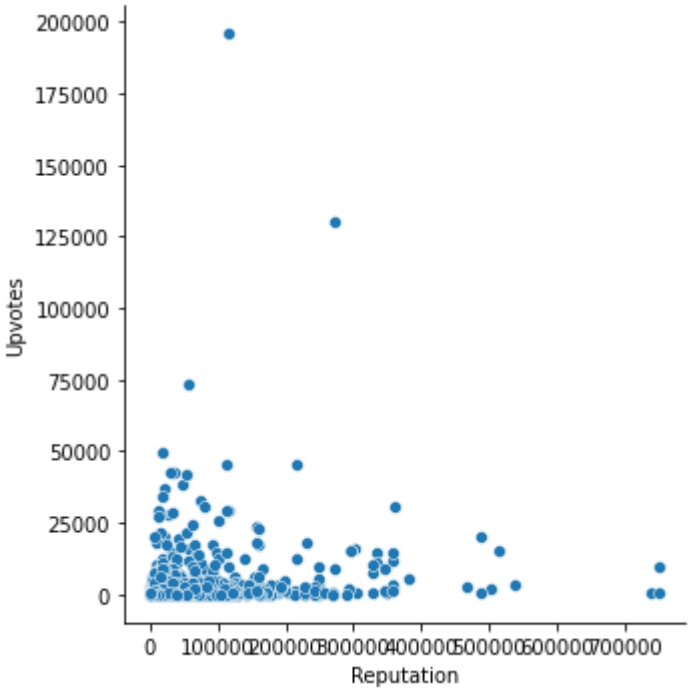


In [15]:

```
sns.relplot(data=df, x="Reputation", y="Upvotes")
```

Out[15]:

```
<seaborn.axisgrid.FacetGrid at 0x7f63be3565e0>
```



In [ ]:

```
```

2. Visualizating Categorial *Data*

Various Categorial Plots in Seaborn

Jitter Plot

In [18]:

```
df2=pd.read_csv('train_hr_mini.csv')
```

In [19]:

```
df2.head()
```

Out[19]:

	employee_id	department	region	education	gender	recruitment_channel	no_of_trainings
0	65438	Sales & Marketing	region_7	Master's & above	f	sourcing	1
1	65141	Operations	region_22	Bachelor's	m	other	1
2	7513	Sales & Marketing	region_19	Bachelor's	m	sourcing	1
3	2542	Sales & Marketing	region_23	Bachelor's	m	other	2
4	48945	Technology	region_26	Bachelor's	m	other	1

In [20]:

```
df2.shape
```

Out[20]:

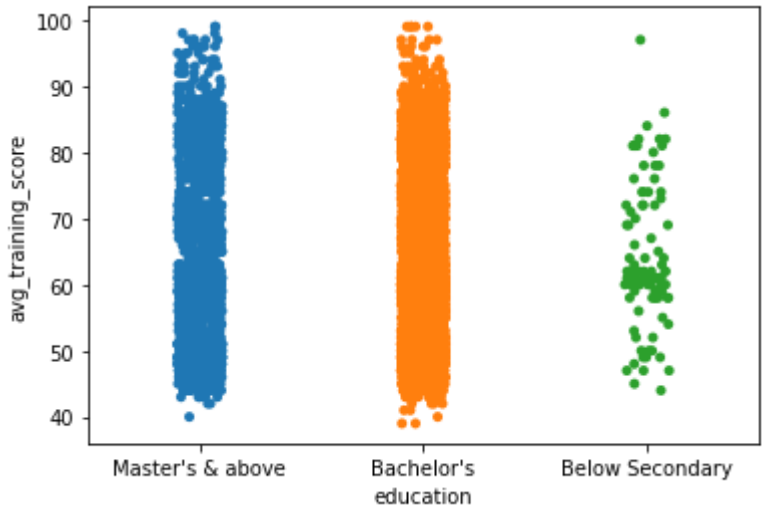
```
(6397, 14)
```

In [21]:

```
sns.stripplot(data=df2, x="education", y="avg_training_score", jitter=True)
```

Out[21]:

```
<AxesSubplot:xlabel='education', ylabel='avg_training_score'>
```

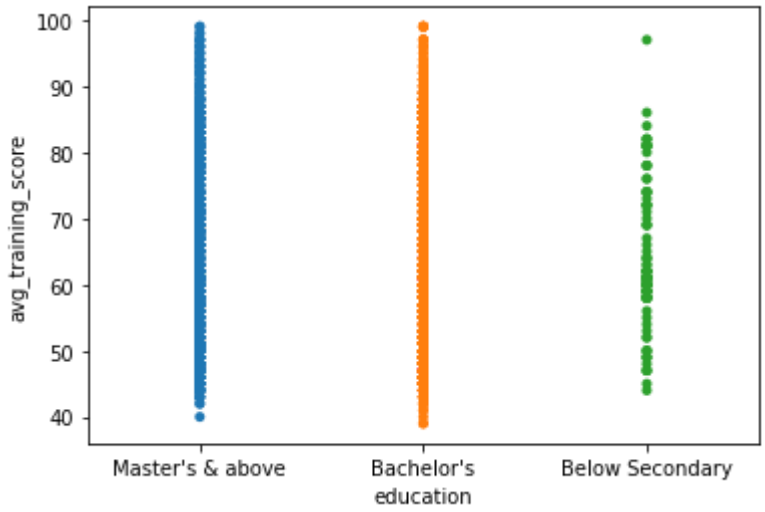


In [22]:

```
sns.stripplot(data=df2, x="education", y="avg_training_score", jitter=False)
```

Out[22]:

```
<AxesSubplot:xlabel='education', ylabel='avg_training_score'>
```



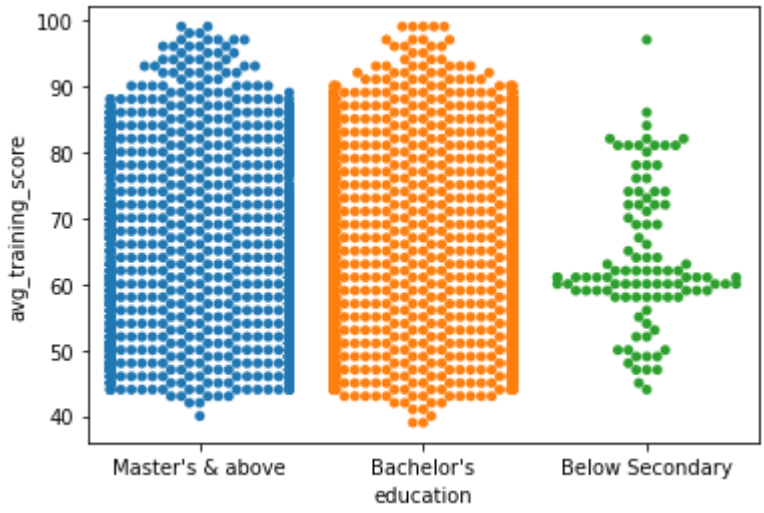
In [23]:

```
sns.swarmplot(data=df2,x='education',y='avg_training_score')
```

```
/usr/local/lib/python3.9/dist-packages/seaborn/categorical.py:1296: UserWarn
ing: 74.2% of the points cannot be placed; you may want to decrease the size
of the markers or use stripplot.
  warnings.warn(msg, UserWarning)
/usr/local/lib/python3.9/dist-packages/seaborn/categorical.py:1296: UserWarn
ing: 88.1% of the points cannot be placed; you may want to decrease the size
of the markers or use stripplot.
  warnings.warn(msg, UserWarning)
```

Out[23]:

```
<AxesSubplot:xlabel='education', ylabel='avg_training_score'>
```



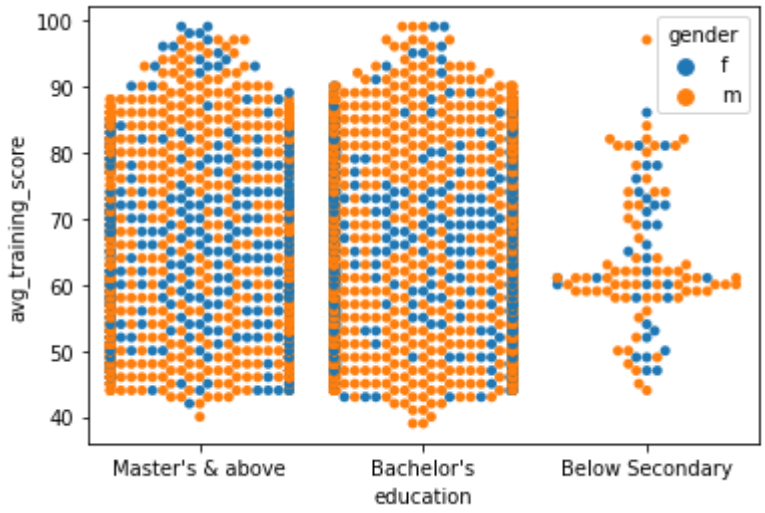
In [24]:

```
sns.swarmplot(data=df2,x="education",y="avg_training_score",hue ='gender')
```

```
/usr/local/lib/python3.9/dist-packages/seaborn/categorical.py:1296: UserWarn
ing: 74.2% of the points cannot be placed; you may want to decrease the size
of the markers or use stripplot.
  warnings.warn(msg, UserWarning)
/usr/local/lib/python3.9/dist-packages/seaborn/categorical.py:1296: UserWarn
ing: 88.1% of the points cannot be placed; you may want to decrease the size
of the markers or use stripplot.
  warnings.warn(msg, UserWarning)
```

Out[24]:

```
<AxesSubplot:xlabel='education', ylabel='avg_training_score'>
```



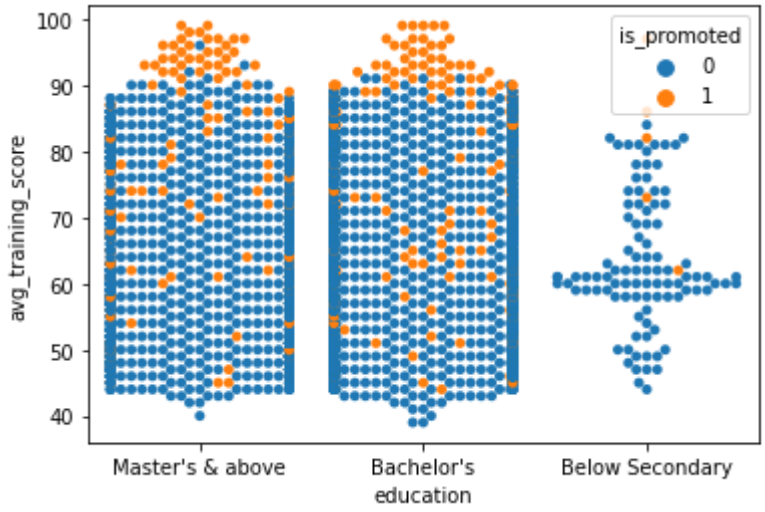
In [25]:

```
sns.swarmplot(data=df2,x="education",y="avg_training_score",hue ='is_promoted')
```

```
/usr/local/lib/python3.9/dist-packages/seaborn/categorical.py:1296: UserWarn
ing: 74.2% of the points cannot be placed; you may want to decrease the size
of the markers or use stripplot.
  warnings.warn(msg, UserWarning)
/usr/local/lib/python3.9/dist-packages/seaborn/categorical.py:1296: UserWarn
ing: 88.1% of the points cannot be placed; you may want to decrease the size
of the markers or use stripplot.
  warnings.warn(msg, UserWarning)
```

Out[25]:

```
<AxesSubplot:xlabel='education', ylabel='avg_training_score'>
```

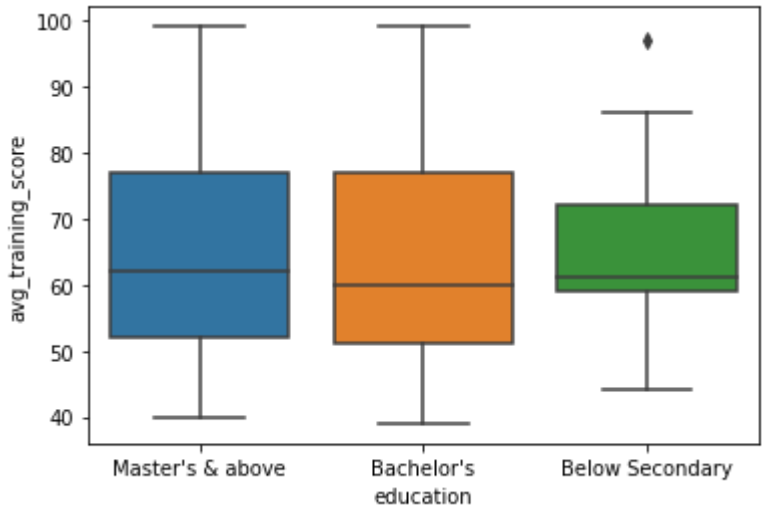


In [26]:

```
sns.boxplot(data=df2,x="education",y="avg_training_score")
```

Out[26]:

```
<AxesSubplot:xlabel='education', ylabel='avg_training_score'>
```

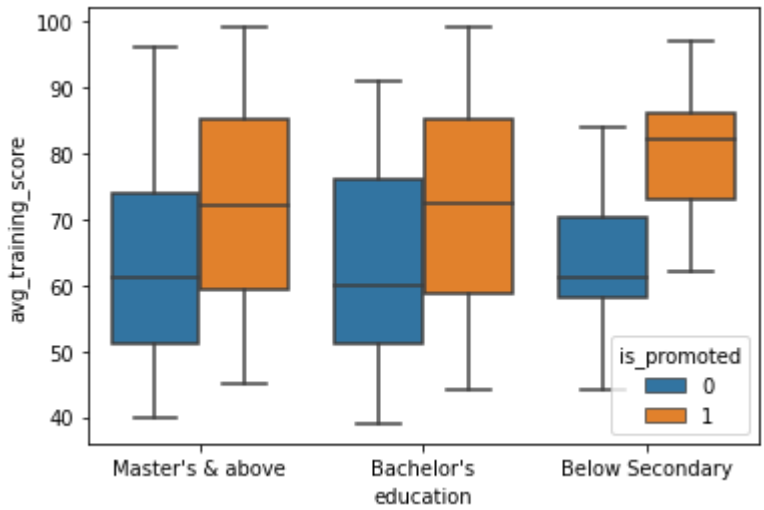


In [27]:

```
sns.boxplot(data=df2,x="education",y="avg_training_score",hue='is_promoted')
```

Out[27]:

```
<AxesSubplot:xlabel='education', ylabel='avg_training_score'>
```

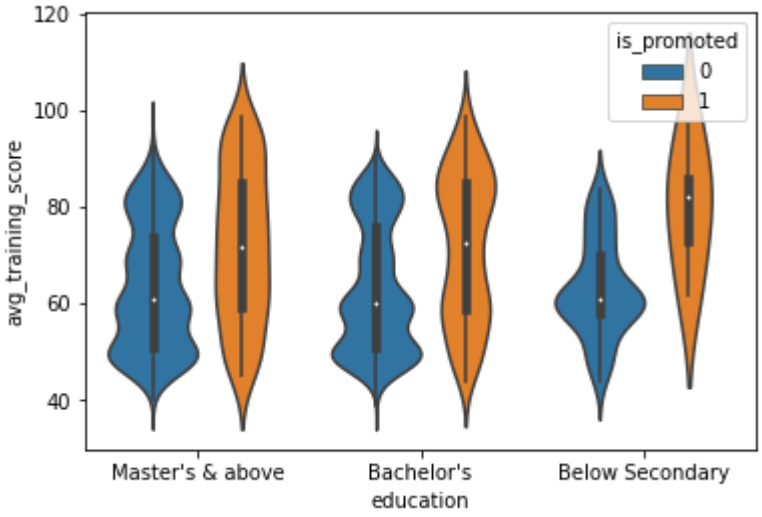


In [28]:

```
sns.violinplot(data=df2,x="education",y="avg_training_score",hue='is_promoted')
```

Out[28]:

```
<AxesSubplot:xlabel='education', ylabel='avg_training_score'>
```

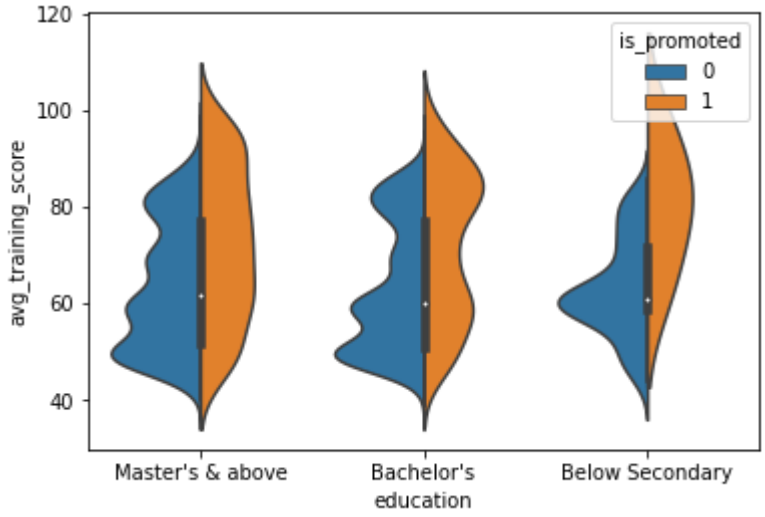


In [29]:

```
sns.violinplot(data=df2,x="education",y="avg_training_score",hue='is_promoted',split=True)
```

Out[29]:

```
<AxesSubplot:xlabel='education', ylabel='avg_training_score'>
```

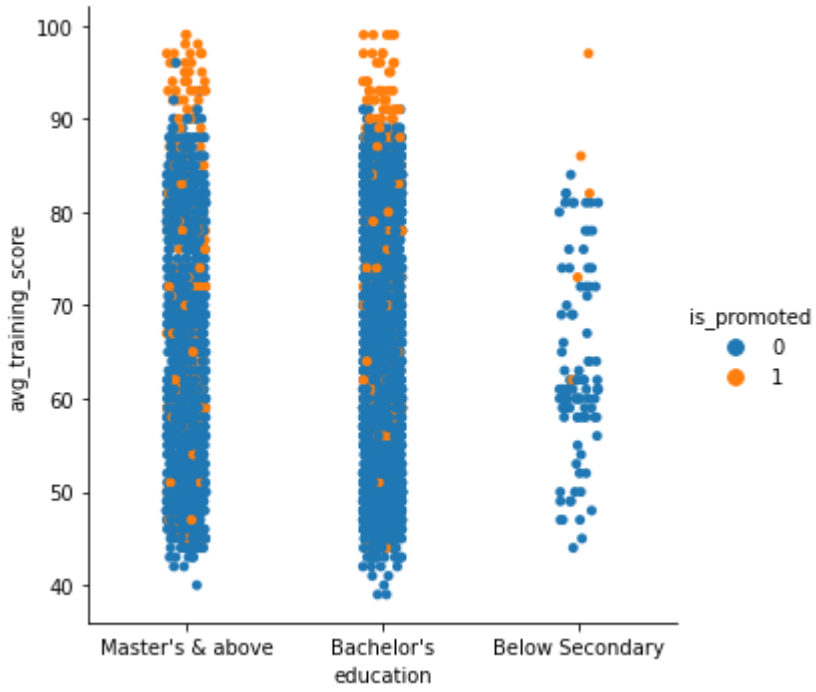


In [30]:

```
sns.catplot(data=df2,x="education",y="avg_training_score",hue ='is_promoted')
```

Out[30]:

```
<seaborn.axisgrid.FacetGrid at 0x7f63b4c3e8e0>
```

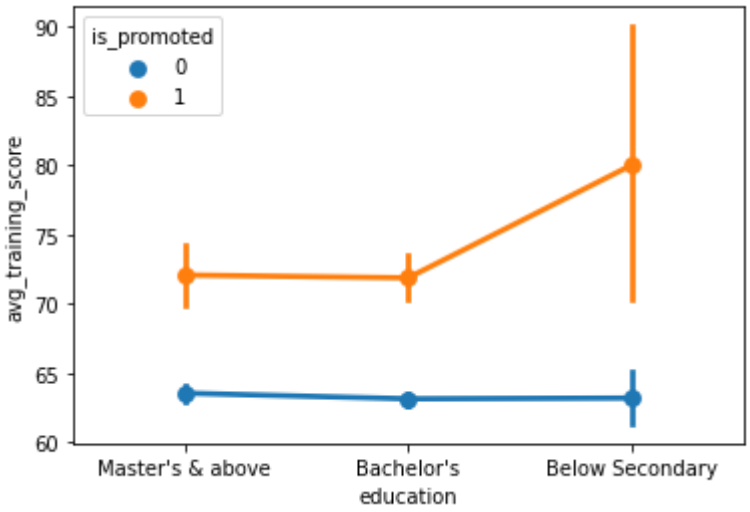


In [31]:

```
sns.pointplot(data=df2,x="education",y="avg_training_score",hue ='is_promoted')
```

Out[31]:

```
<AxesSubplot:xlabel='education', ylabel='avg_training_score'>
```



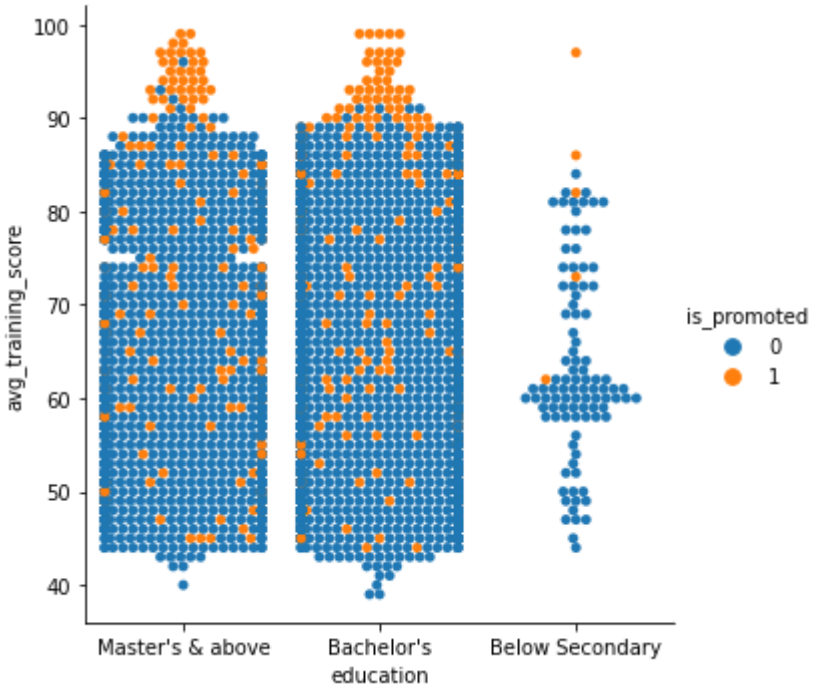
In [32]:

```
sns.catplot(data=df2,x="education",y="avg_training_score",hue ='is_promoted',kind='swarm')
```

```
/usr/local/lib/python3.9/dist-packages/seaborn/categorical.py:1296: UserWarn
ing: 56.8% of the points cannot be placed; you may want to decrease the size
of the markers or use stripplot.
  warnings.warn(msg, UserWarning)
/usr/local/lib/python3.9/dist-packages/seaborn/categorical.py:1296: UserWarn
ing: 81.5% of the points cannot be placed; you may want to decrease the size
of the markers or use stripplot.
  warnings.warn(msg, UserWarning)
```

Out[32]:

```
<seaborn.axisgrid.FacetGrid at 0x7f63b2a23730>
```



In [ ]:

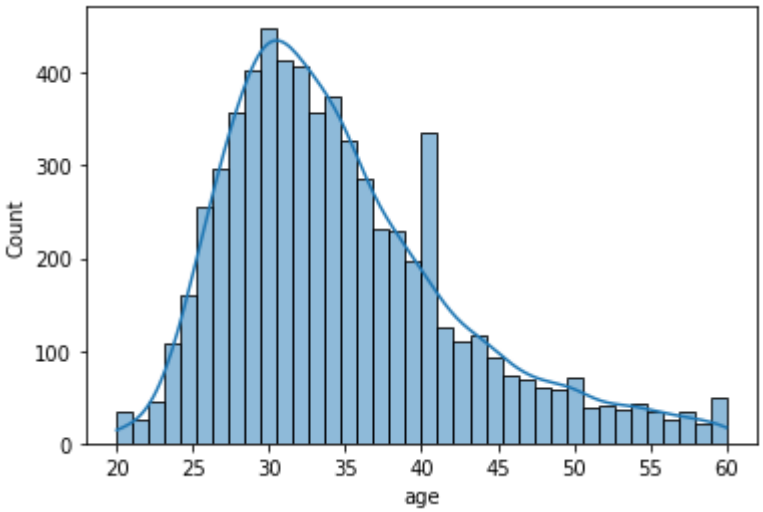
Visualizing the Distribution of Data

In [33]:

```
sns.histplot(x='age',data=df2,kde=True)
```

Out[33]:

```
<AxesSubplot:xlabel='age', ylabel='Count'>
```

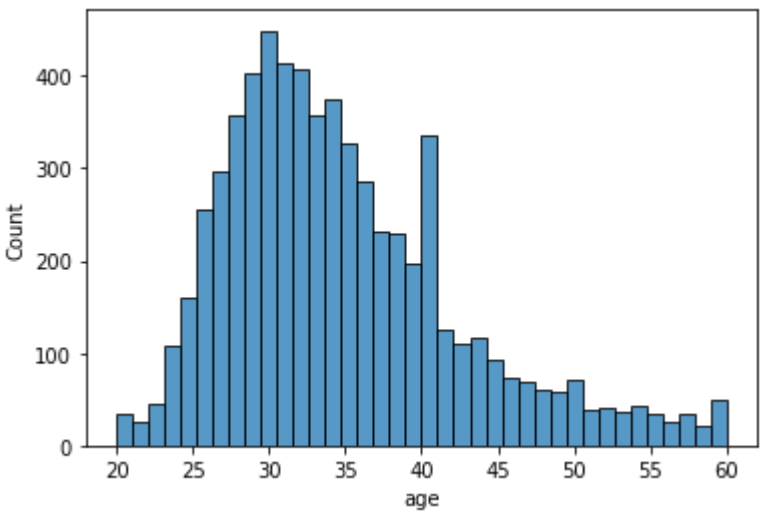


In [34]:

```
sns.histplot(x='age',data=df2)
```

Out[34]:

```
<AxesSubplot:xlabel='age', ylabel='Count'>
```

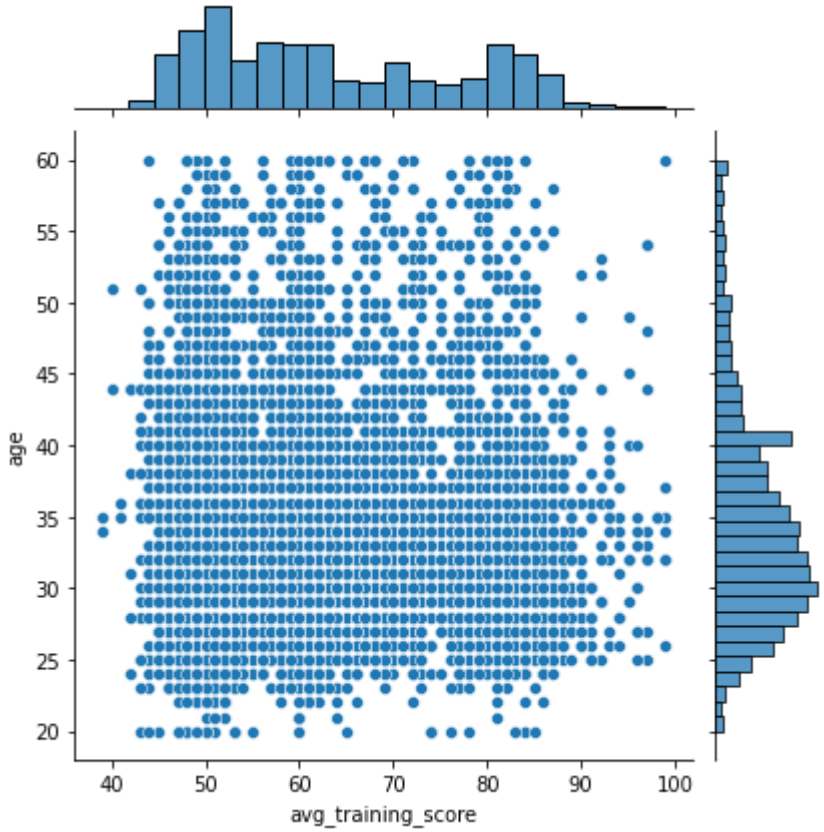


In [35]:

```
sns.jointplot(x='avg_training_score',y='age',data=df2)
```

Out[35]:

```
<seaborn.axisgrid.JointGrid at 0x7f63b29200a0>
```

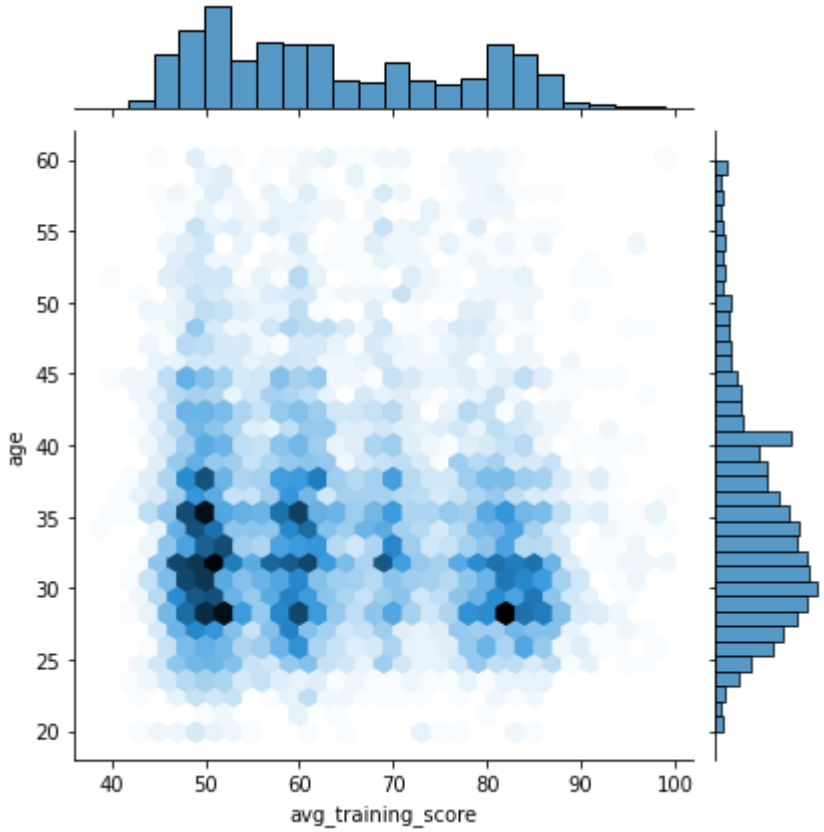


In [36]:

```
sns.jointplot(x='avg_training_score',y='age',kind='hex',data=df2)
```

Out[36]:

```
<seaborn.axisgrid.JointGrid at 0x7f63b6e4cc40>
```

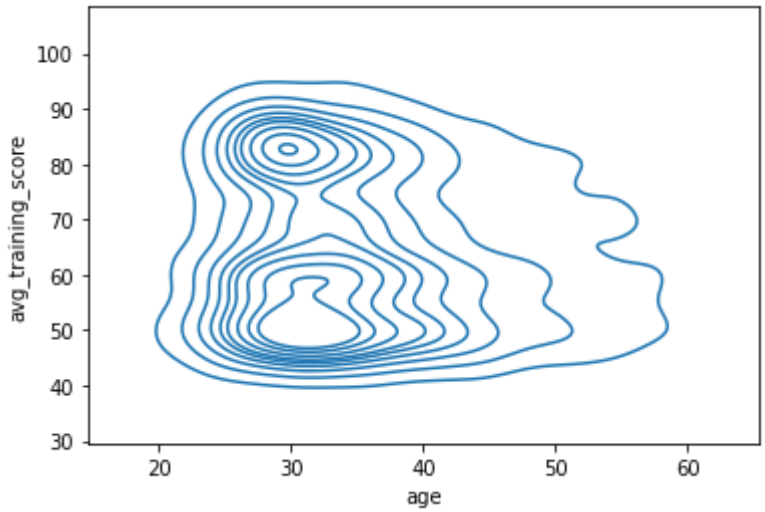


In [37]:

```
sns.kdeplot(x='age',y='avg_training_score',data=df2)
```

Out[37]:

```
<AxesSubplot:xlabel='age', ylabel='avg_training_score'>
```

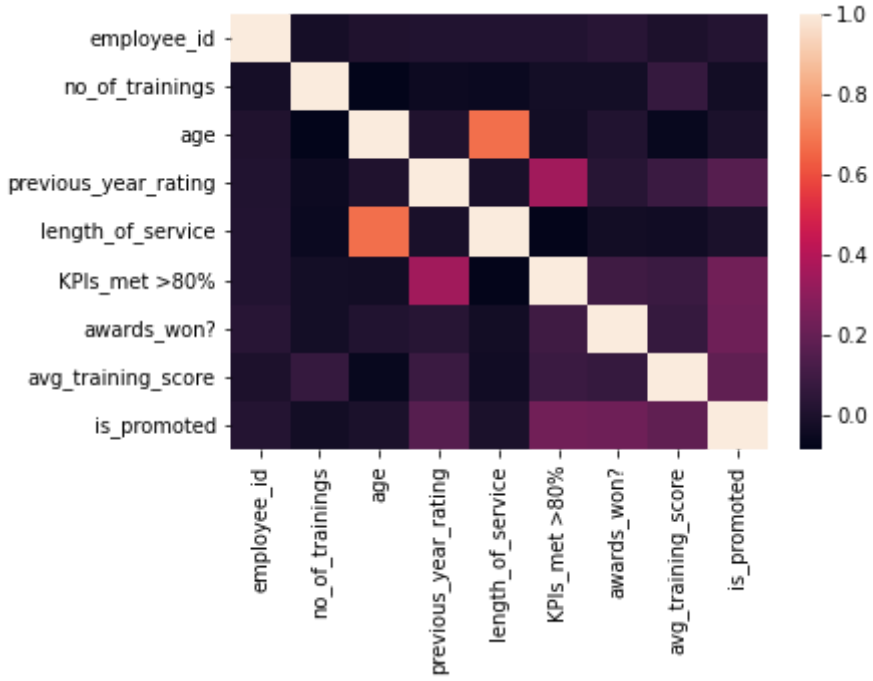


In [38]:

```
sns.heatmap(df2.corr())
```

Out[38]:

<AxesSubplot:>

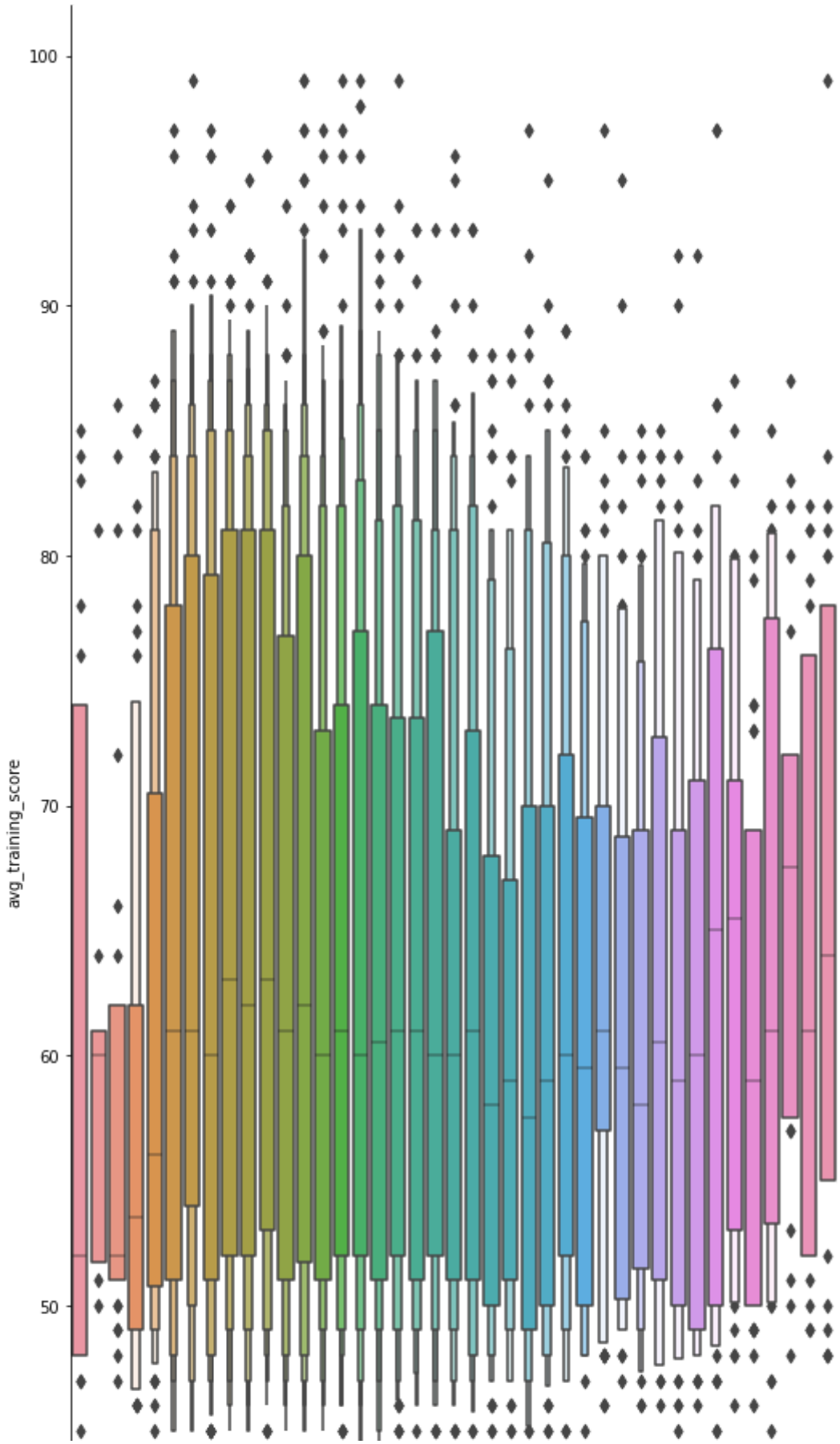


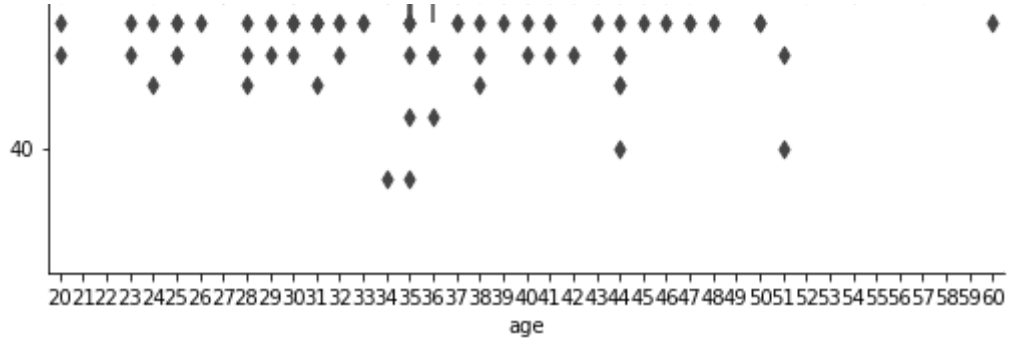
In [39]:

```
sns.catplot(x='age',y='avg_training_score',data=df2,kind='boxen',height=15,aspect=.5)
```

Out[39]:

```
<seaborn.axisgrid.FacetGrid at 0x7f63b2461ac0>
```



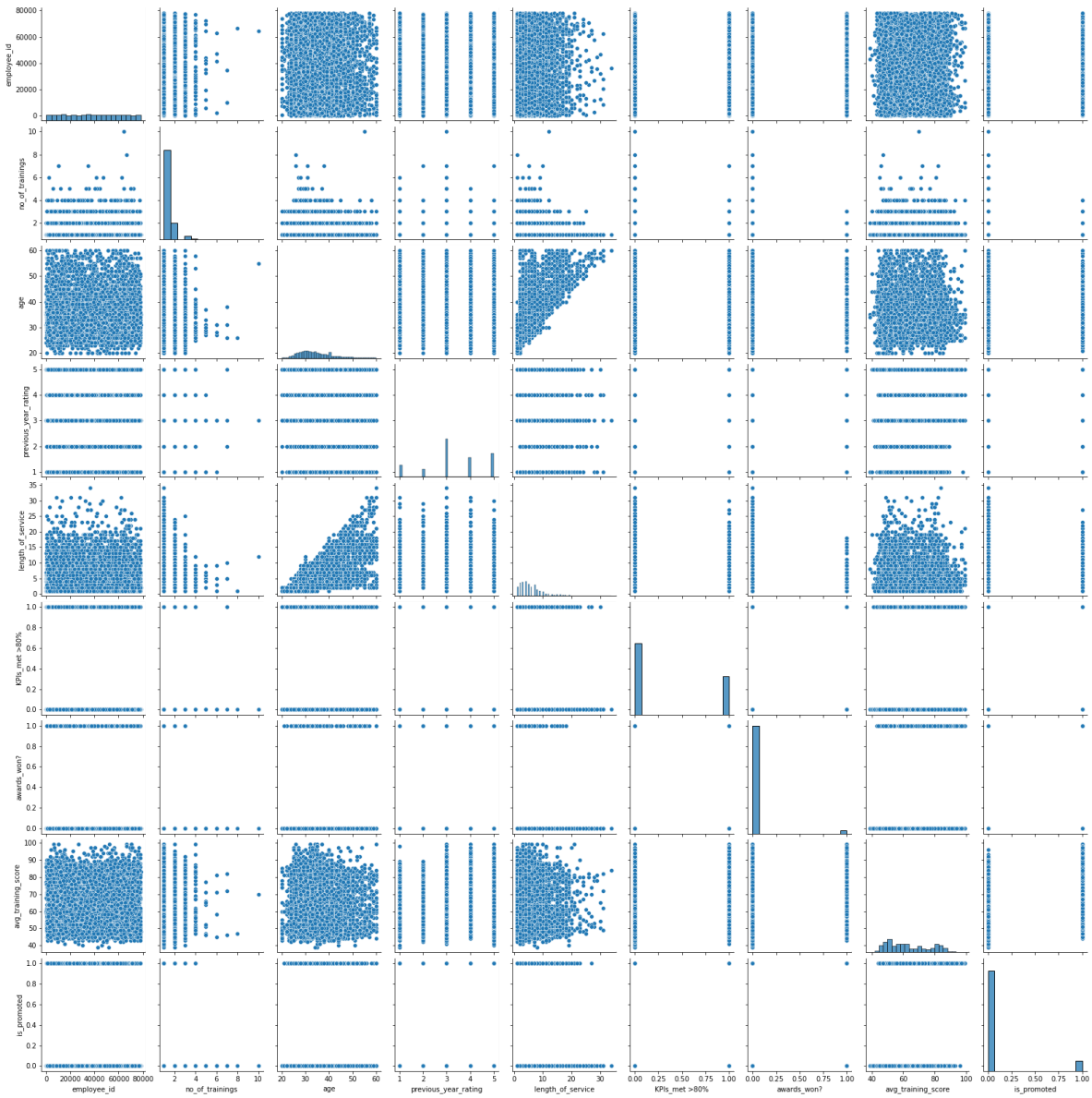


In [40]:

```
sns.pairplot(df2)
```

Out[40]:

```
<seaborn.axisgrid.PairGrid at 0x7f63b2116fa0>
```



In [ ]: