```
import numpy as np
import pandas as pd
import seaborn as sns
import plotly.express as px
import matplotlib.pyplot as plt
%matplotlib inline

from urllib.request import urlretrieve

datas= pd.read_csv('Talents for Startup.csv6')
datas.head()
```

```
   Present   Absent    E  Present (total)  Absent (total)  E(total)
Weeks
0        10        5   8            191.0            91.0       8.0
weekOne
1        43       28   0              NaN             NaN       NaN
weekTwo
2        69       27   0              NaN             NaN       NaN
weekThree
3        53       22   0              NaN             NaN       NaN
weekFour
4        16        9   0              NaN             NaN       NaN
weekFive
```

```
datas.isnull().sum()
```

```
Present             0
Absent              0
 E                  0
Present (total)     4
Absent (total)      4
E(total)            4
Weeks               0
dtype: int64
```

```
datas.info()
```

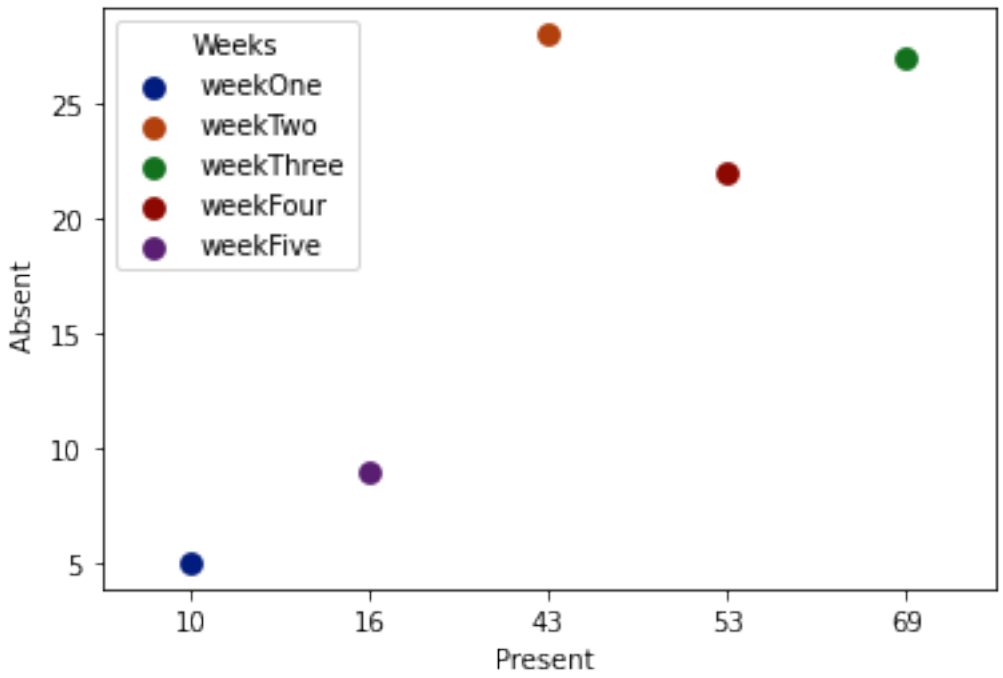
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5 entries, 0 to 4
Data columns (total 7 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   Present          5 non-null      int64
 1   Absent           5 non-null      int64
 2    E               5 non-null      int64
 3   Present (total)  1 non-null      float64
 4   Absent (total)   1 non-null      float64
 5   E(total)         1 non-null      float64
 6   Weeks            5 non-null      object
```

```
dtypes: float64(3), int64(3), object(1)
memory usage: 408.0+ bytes
```

a) let us get to know how many times people were present throughout the five weeks

```
sns.pointplot(data=datas, x='Present ', y='Absent ', hue='Weeks',
palette='dark')

<AxesSubplot:xlabel='Present ', ylabel='Absent '>
```



from the above we can see that the number of people that were present at the programme increased with time during the week which indicates that there is interest in the course studied.

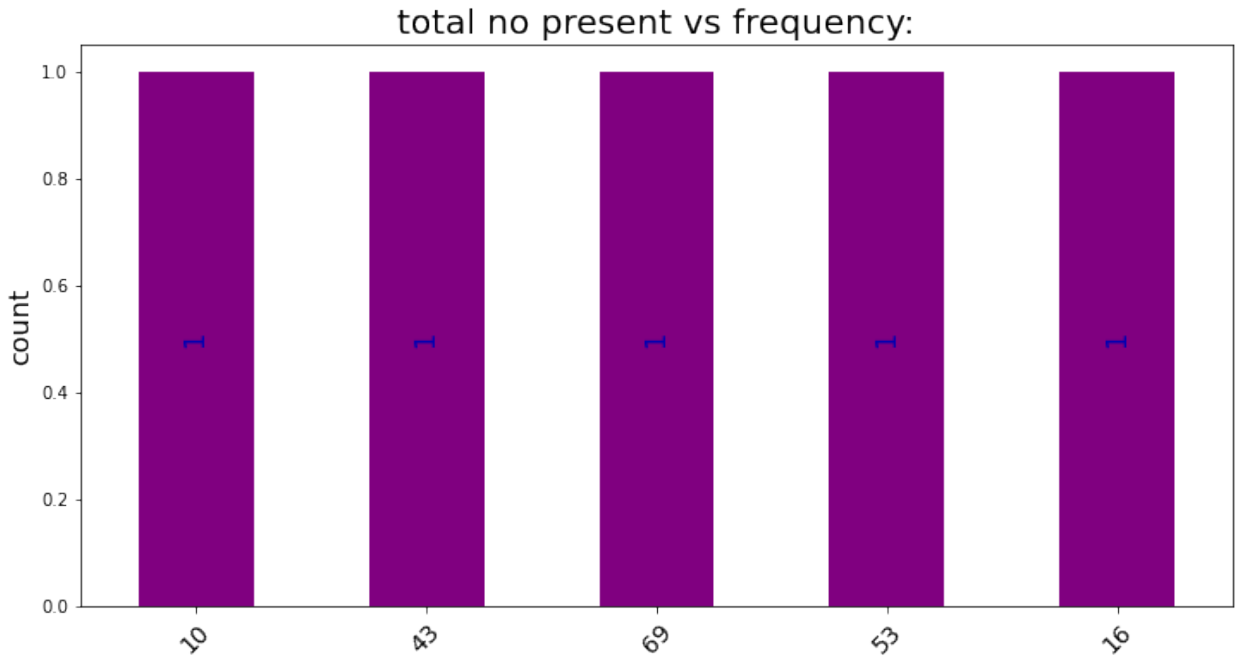
this includes that the instructor does a very good job at communicating with his students also they find it condusive.

In week 3 there was high level of attendance.

b)lets know the number of people that were consistent.

```
plt.figure(figsize=(12, 6))
ax = datas['Present '].value_counts().head(10).plot(kind = 'bar',color
= 'purple')
plt.title('total no present vs frequency:', fontsize = 20)
plt.ylabel('count', fontsize = 16)
plt.xticks(fontsize = 14, rotation = 45)
```

```
plt.bar_label(ax.containers[0], label_type = 'center',fontsize = 16,
rotation = 90, color = '#0000b3')
plt.show()
```



from here we can see that theb number of people who were present were consistent through out the weeks.

```
datas['Absent '].value_counts()
```

```
5     1
28    1
27    1
22    1
9     1
Name: Absent , dtype: int64
```

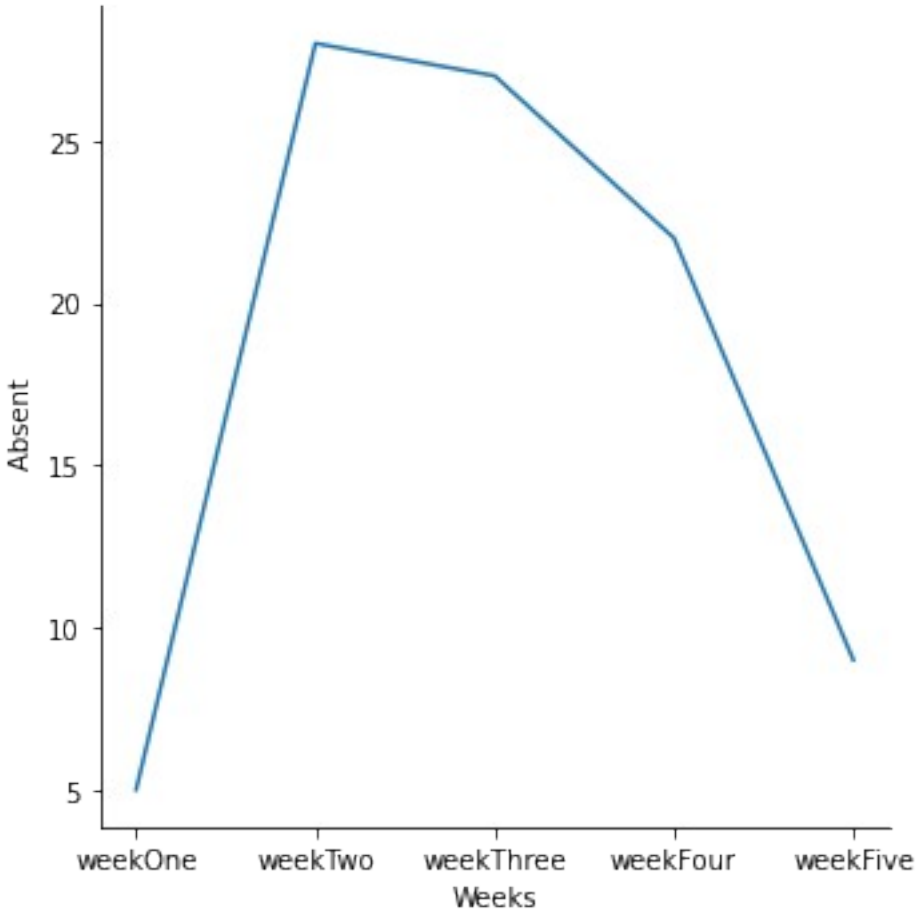
```
datas['Present '].value_counts()
```

```
10    1
43    1
69    1
53    1
16    1
Name: Present , dtype: int64
```

c) let us look at the number of people who were absent .

```
sns.relplot(data=datas, x='Weeks', y='Absent ',kind='line')
```

```
<seaborn.axisgrid.FacetGrid at 0x7fe6003199a0>
```



from the above plot we can see that the least number of times people were absent is in week one, in week two we recorded a higher frequency in absentees .

so week one has it for me

D)lets relate the names with their attendances

```
import numpy as np
import pandas as pd
import seaborn as sns
import plotly.express as px
import matplotlib.pyplot as plt
%matplotlib inline

from urllib.request import urlretrieve

data= pd.read_csv('talent for startup batch a .csv')
data.head()
```

```
                   Names  No of Presents No of Absents No of E  No of
days
0      Folashade Ayodele              13          none    none
```

```
13
1    Ayomide Osundahunsi               13          none    none
13
2     Bright Ndoma Obaji               12          none     one
13
3      Ugbodaga Emmanuel               13          none    none
13
4  Eseyin Rowland Mayowa                9           two     two
13
```

```
data.isnull().sum()
```

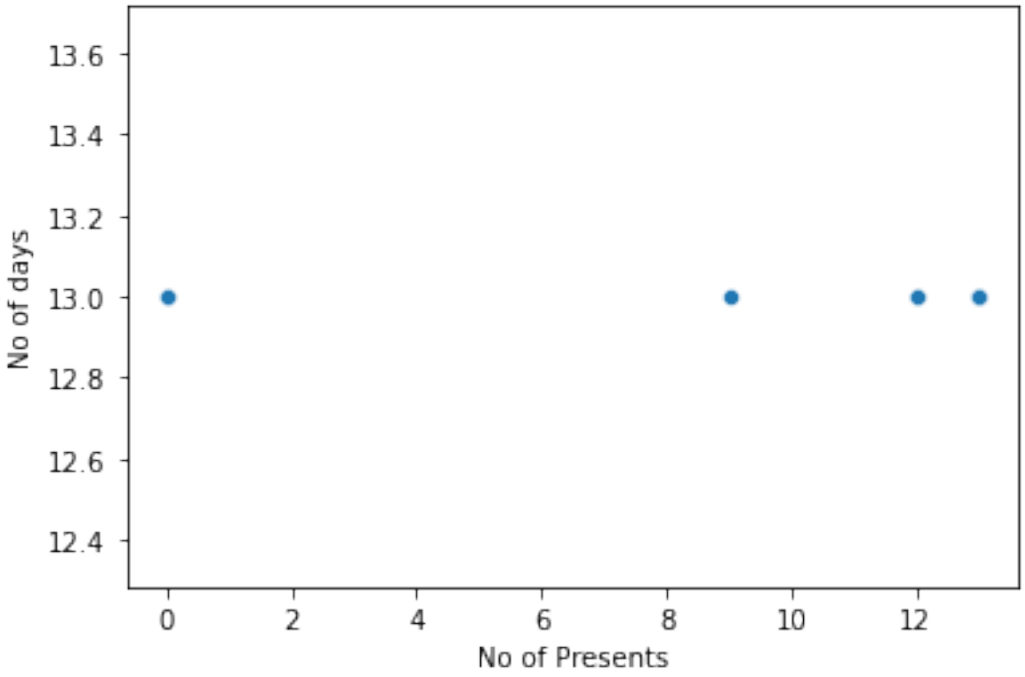
```
Names              0
No of Presents     0
No of Absents      0
No of E            0
No of days         0
dtype: int64
```

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9 entries, 0 to 8
Data columns (total 5 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   Names           9 non-null      object
 1   No of Presents  9 non-null      int64 
 2   No of Absents   9 non-null      object
 3   No of E         9 non-null      object
 4   No of days      9 non-null      int64 
dtypes: int64(2), object(3)
memory usage: 488.0+ bytes
```

```
sns.scatterplot(data=data, x='No of Presents', y='No of days')
```

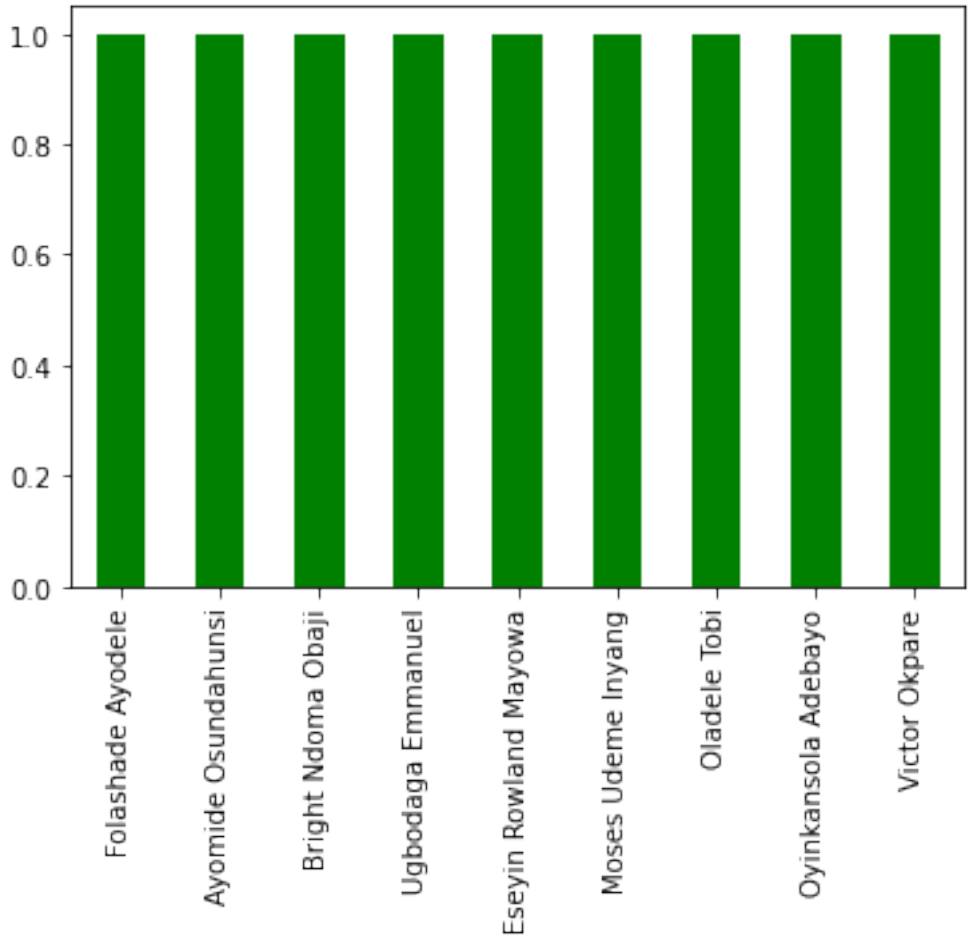
```
<AxesSubplot:xlabel='No of Presents', ylabel='No of days'>
```



d) from the plot we can see that the number of days spent during the 5 weeks program is 13 days and from the graph we can see that from the top ten names included 5 people attended all the classes from the first top ten (batch a)

```
data['Names'].value_counts().plot(kind='bar',color = 'green')
```

```
<AxesSubplot:>
```



e)from this we can see that there is no one signing in twice, also we can get the mean ,median and mode from this plot

from number C we can see clearly that attendance of absent people was at its peak at week 2. How can we improve attendance?

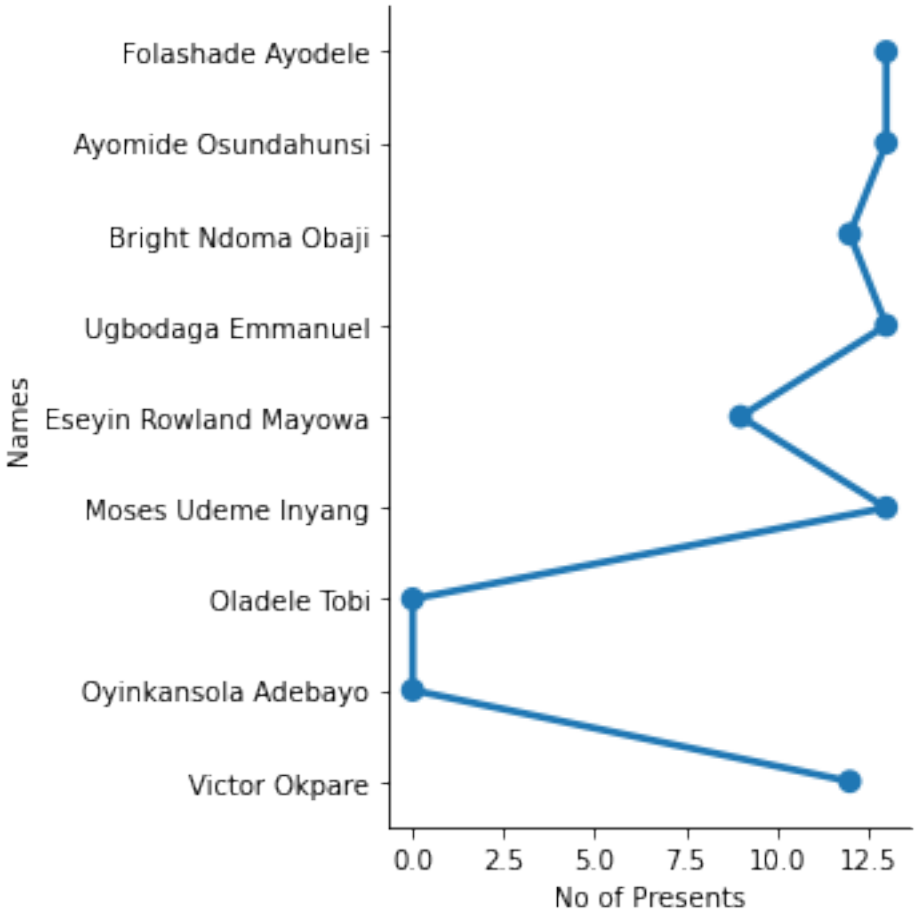
i) Transportation: how close is the school to their homes? enabling communication by gfiving group works with aid communication hence making them reach out to eachother.

ii) Also encouraging them to know their locations so that whoever has a car can assist in mobility while coming.

conclusion:

There has been a positive outcome in coming to class which indicates interest in the course.

```
sns.catplot(x = 'No of Presents',y ='Names',data= data,kind='point')
plt.show()
```



f) From this plot we can see the first ten people and how many times they came around within the 13 days of learning