Data analysis through visualization

What we did:
In last class we learned about the correlation and the methods to find it.

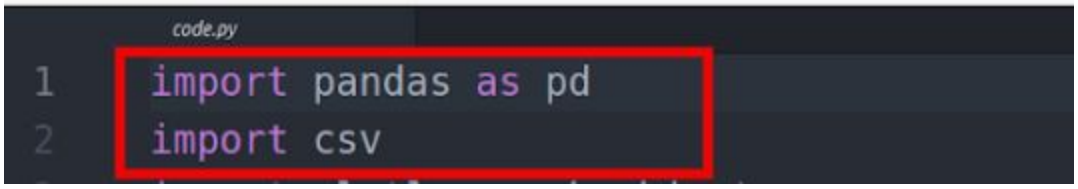
In this class we learned about analysing data through visualization.

How we did it:

1. We saw the data of the 12 students of grade 3 who played lesson 1 on pixel math app.

```
student_id,level,attempt
TRL_xsl,Level 4,1
TRL_xsl,Level 1,1
TRL_xsl,Level 2,1
TRL_xsl,Level 3,1
TRL_xsl,Level 4,1
TRL_xsl,Level 2,1
TRL_xsl,Level 2,1
TRL_xsl,Level 3,1
TRL_xsl,Level 4,1
TRL_xsl,Level 3,1
TRL_xsl,Level 3,1
TRL_xsl,Level 1,1
TRL_xsl,Level 2,1
TRL_xsl,Level 3,1
TRL_xsl,Level 1,1
TRL_xsl,Level 4,1
TRL_xsl,Level 4,1
TRL_xsl,Level 4,0
TRL_xsl,Level 1,1
TRL_xsl,Level 1,1
TRL_xsl,Level 2,1
TRL_xsl,Level 1,1
```

2. We imported pandas and csv to program file.



3. We read the data from the csv file and stored it in df variable.

```
import pandas as pd
import csv
import plotly.graph_objects as go

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

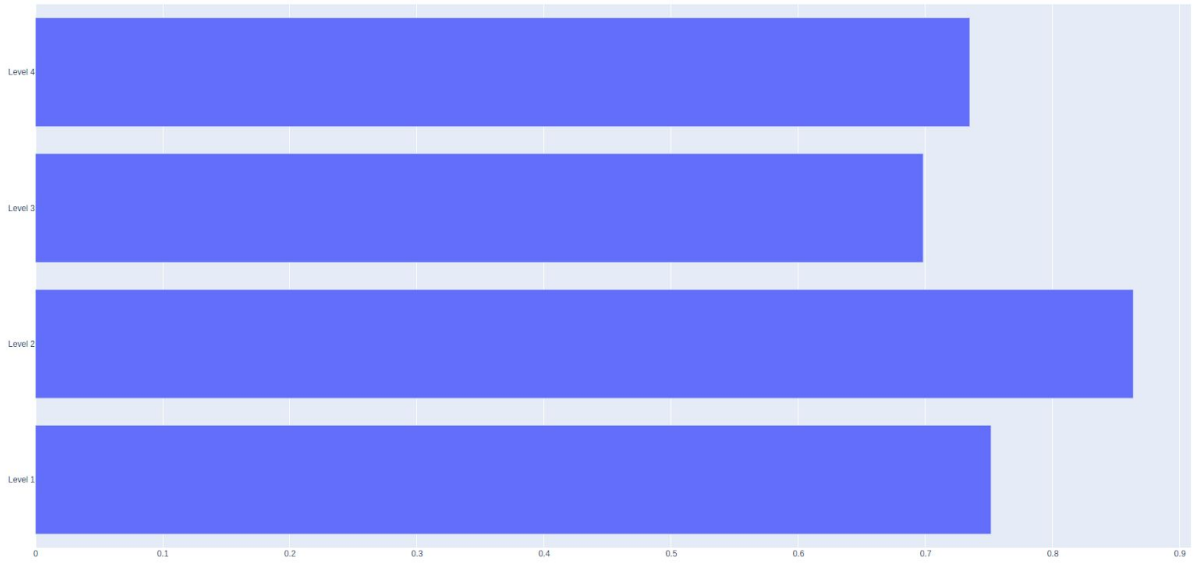
4. We got the mean of the attempts grouping them by the level

```
code.py
1  import pandas as pd
2  import csv
3  import plotly.graph_objects as go
4
5  df = pd.read_csv("data.csv")
6
7  print(df.groupby("level")["attempt"].mean())
8
```

```
[redacted]Pixel Math data$ python3 code.py
level
Level 1    0.751445
Level 2    0.863281
Level 3    0.698113
Level 4    0.734694
Name: attempt, dtype: float64
```

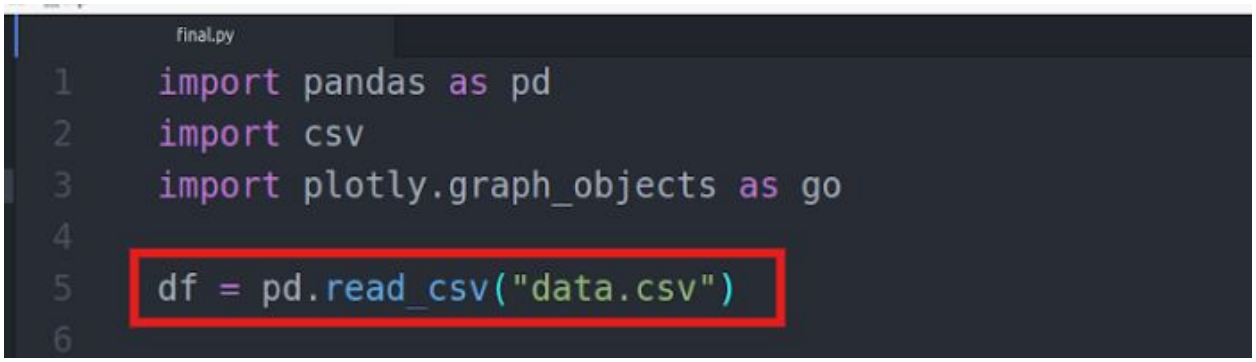
5. Then we plotted the horizontal bar graph based on the data.

```
code.py
1  import pandas as pd
2  import csv
3  import plotly.graph_objects as go
4
5  df = pd.read_csv("data.csv")
6
7  print(df.groupby("level")["attempt"].mean())
8
9  fig = go.Figure(go.Bar(
10             x=df.groupby("level")["attempt"].mean(),
11             y=['Level 1', 'Level 2', 'Level 3', 'Level 4'],
12             orientation='h'))
13
14 fig.show()
15
```



We plotted another bar graph.

1. We created new file ,imported pandas and csv and read the data from csv and stored it in the df variable.



2. We filtered and got the data of the particular student.

```
final.py
1  import pandas as pd
2  import csv
3  import plotly.graph_objects as go
4
5  df = pd.read_csv("data.csv")
6
7  student_df = df.loc[df['student_id'] == "TRL_987"]
8
```

3. We got the mean of the attempts by grouping them by level and using graph_objects we plotted the bar chart for that student.

```
final.py
1  import pandas as pd
2  import csv
3  import plotly.graph_objects as go
4
5  df = pd.read_csv("data.csv")
6
7  student_df = df.loc[df['student_id'] == "TRL_987"]
8
9  print(student_df.groupby("level")["attempt"].mean())
10
11 fig = go.Figure(go.Bar(
12             x=student_df.groupby("level")["attempt"].mean(),
13             y=['Level 1', 'Level 2', 'Level 3', 'Level 4'],
14             orientation='h'))
15
16 fig.show()
17
```

What's next?

In the next class, we will learn about normal distribution.