Topic	**Data Analysis by visualization**	
Class Description	**Students use class assessment data collected from a mobile app and analyze it by visualizing through drawing graphs. Students identify the trouble spots or concept areas where students need more reinforcement - both at the class level and at student level.**	
Class	**C107**	
Class time	**45 mins**	
Goal	• Understand and study the assessment data collected from the mobile app • Choose a visualization graph to represent the data in a meaningful way • Use tools in pandas to process the data and draw the graph using plotly	

Resources Required	• Teacher Resources
	  ○ Laptop with internet connectivity
	  ○ Earphones with mic
	  ○ Notebook and pen
	• Student Resources
	  ○ Laptop with internet connectivity
	  ○ Earphones with mic
	  ○ Notebook and pen	

Class structure	**Warm Up** **Teacher-led Activity** **Student-led Activity** **Wrap up**	**5 mins** **15 min** **15 min** **5 min**

CONTEXT

- **Give context to the students on the use of data analysis in student assessments.**

Class Steps	Teacher Action	Student Action

Step 1: **Warm Up** **(5 mins)**	Hi <Student Name>! Can you quickly recall what we have learned in the last few classes?	**ESR:** - We learned how to visualize data using bar charts, line charts and scatter plots. - We studied about the central tendency of data- mean, median and mode. - We learned about standard deviation using which we can understand how much data differs from the central tendency. - We also learned about correlation and how to identify if one data depends on the values of others.
	Amazing! Today, we are going to apply some of these concepts to real-life data and gauge the power of analyzing data by visualizing it. Before we do that, can you think of some of the areas where analyzing data might be useful?	**ESR:** The *student mentions some of the areas where they think analyzing data might be useful.* *Some of the areas which students can come up with are banking, budget tracking, etc.*
	What about learning? Can analyzing data be helpful or useful in learning? Allow the student to think for sometime on it.	*The student can come up with varied responses.*
Teacher Initiates Screen Share		

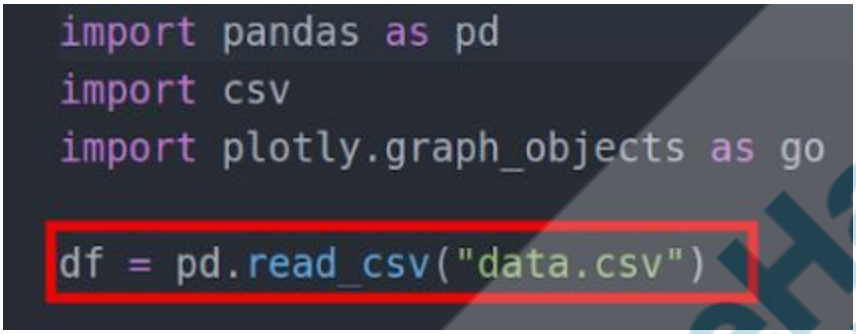
CHALLENGE • **Filter the data using tools in pandas** • **Use Plotly graph objects to represent the data and show the trouble spot for the class of students**		
Step 2: Teacher-led Activity (15 min)	We have some data with us from a live app on PlayStore / AppStore. The app - PixelMath - allows students from different grades to login and solve math questions aligned to their Maths Textbook in a gamified way. Show the csv data to the student. *<Teacher opens the link from Teacher Activity 1>* What do you understand from the data? *<this is a snap of data. see full data by opening the link>*	**ESR:** varied

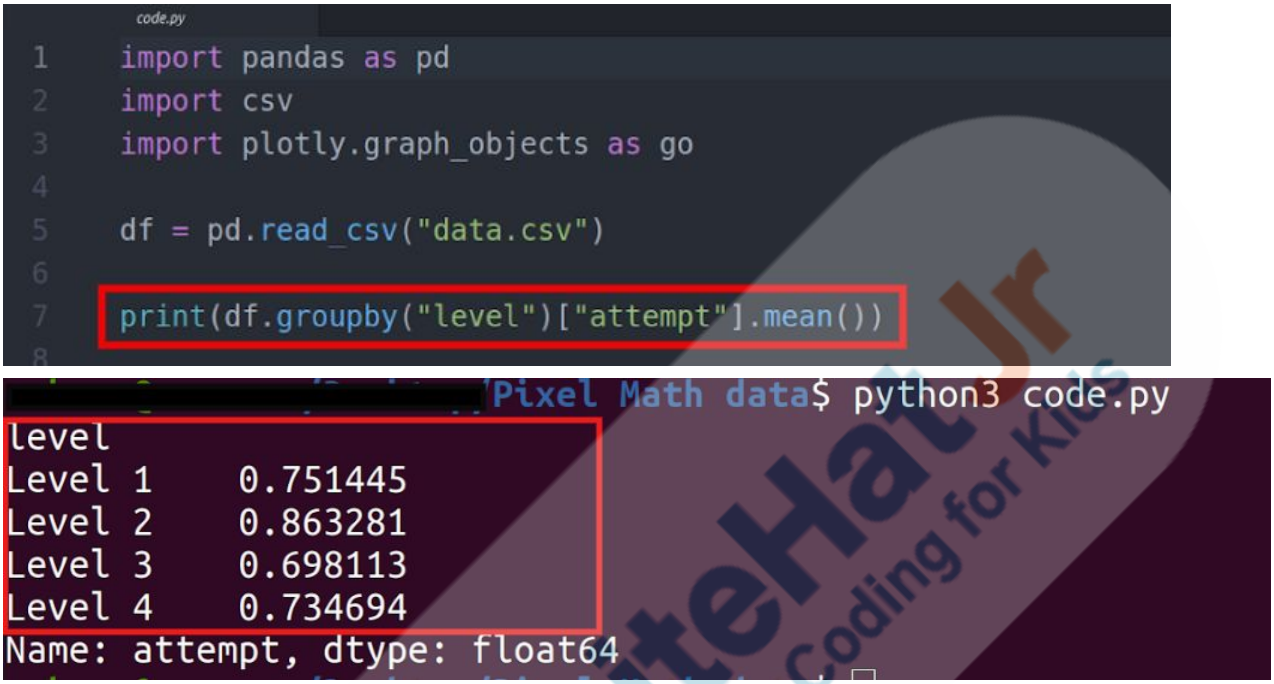
```
student_id,level,attempt
TRL_xsl,Level 4,1
TRL_xsl,Level 1,1
TRL_xsl,Level 2,1
TRL_xsl,Level 3,1
TRL_xsl,Level 4,1
TRL_xsl,Level 2,1
TRL_xsl,Level 2,1
TRL_xsl,Level 3,1
TRL_xsl,Level 4,1
TRL_xsl,Level 3,1
TRL_xsl,Level 3,1
TRL_xsl,Level 1,1
TRL_xsl,Level 2,1
TRL_xsl,Level 3,1
TRL_xsl,Level 1,1
TRL_xsl,Level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

	The data is for 12 students of Grade 3 who played Lesson 1 in the app. The Lesson 1 is divided into four levels aligned to the four different concepts covered in the lesson. Each row in the data represents - unique student id, level which the	

	student was playing and whether the student got the question right (1) or wrong (0). For example- the first row: TRL_xsl,Level 4,1 TRL_xsl - > unique student id Level 4 -> Student is playing level 4 (or concept 4) 1 -> the student has answered the question correctly. The data represents attempts of 12 different students from a class who practiced Lesson 1 using the app. Look at the data again (in the csv file) and see if you can understand the data. Are there any conclusions you can draw by looking at the data?	*The student looks at the raw data and attempts to make some conclusions by looking at the data.*
	What would a teacher - who is teaching these students - like to know from the data? What would each student like to know from the data?	*The teacher would like to know which concept the class is strong in and which concept the class finds challenging.* *The teacher would also like to know the performance of each child in the different concepts.*
	Let us see if we can use the python skills and tools we have to get this information. Let us download the data in our local machine.	*Student observes.*

	Teacher downloads the data on the local machine and places it in a folder titled "PixelMath Data". Let us create a new file in the folder called code.py. We are going to write the code in code.py to analyze the data.	
	We can open code.py on Visual Studio Code or another editor to write the code. What would we like to do first?	**ESR:** We would like to read the data from our data file.
	How do we read the data from a csv file? *Help the student recall how to read data from the csv file.*	**ESR:** We can use dataframes in pandas.
	Let us first import pandas and csv. *Teacher writes code to import pandas and csv.*	*The student helps the teacher in importing pandas and csv.*
		
	pandas has a read_csv function which can read a csv file and store it as a dataframe object. Do you remember what a dataframe object in python is?	**ESR:** Dataframe is a 2D labeled data structure having multiple columns.

	Yes! *Teacher writes code to read the csv file and store it in a dataframe object.* *The teacher can also print the dataframe object and run the code to show what the dataframe object contains.*	*Student observes and guides the teacher to write the code.* *The student sees the output has 1654 different attempts by 12 students of Grade 3 class.*
		
	We want to understand - for each level- what was the performance of all the students. What can we do to get this?	**ESR:** varied
	One way would be to group the data by each level and then get the mean value of all the attempts for each level. Comparing the mean would tell us how the students performed across the different levels.	*Student reflects on this and asks questions.*
	pandas dataframe object has functions which can help us do that. *Teacher writes code to group the data by level and calculate the mean for the attempts in the level.*	*Student tries to understand the code and observes the output.*

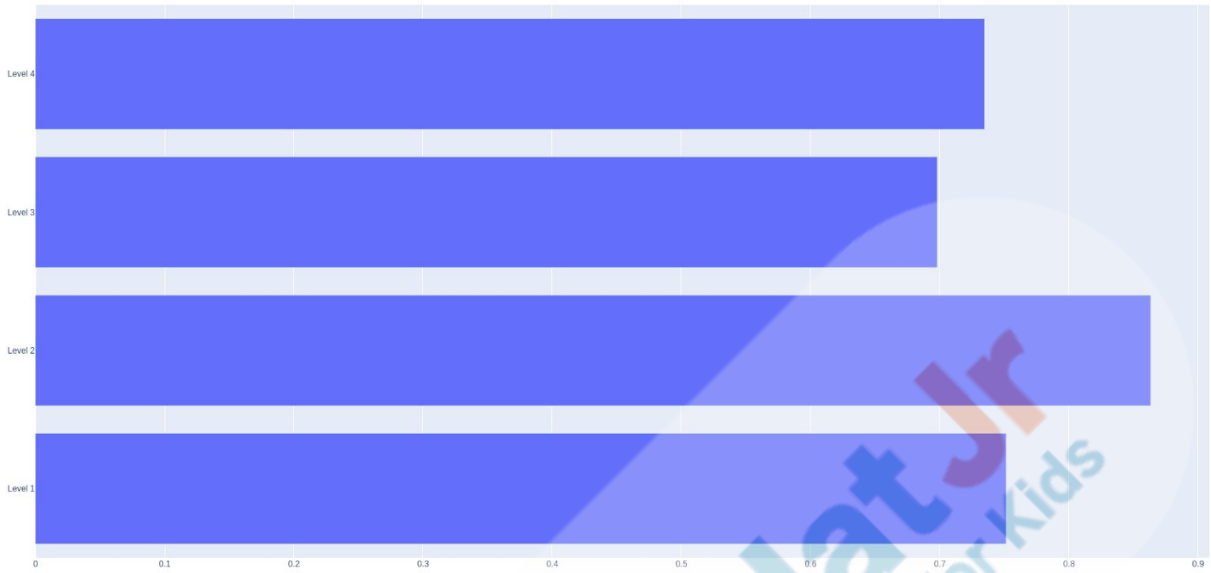
	Teacher runs the code to show the output to the kid.	
	
	What does the output tell you?	**ESR:** It tells us that the students have performed best in Level 2 and least in Level 3. The teacher might have to re-teach or reinforce concept 3 in the class.
	Good analysis. Now, let us try to visually represent it for the teacher. We're going to use the graph_objects in plotly to draw a horizontal bar graph. *Teacher uses plotly graph_objects to draw a horizontal bar graph.*	*Student observes and asks questions.*

	First we'll import plotly graph_objects as go. Using the go.Bar method we'll plot the mean that we found before on X axis and the list of levels on Y axis. Then use fig.show to show the graph.	

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
```

	Teacher runs the code and shows the output. What do you see?	**ESR:** We can clearly see that students have performed the least in level 3 and best in level 2.



	This data gives the teacher inputs on where he/she can spend time with the students.	-
	The data visualization and analysis we did was for all the students in Grade 3. What if we wanted to understand how each student performed across the different levels? For example: Can you do the same analysis for a student with id - TRL_987 and understand how this student performed in the different levels?	**ESR:** Yes!
	Let's try.	

Teacher Stops Screen Share

	Now it's your turn. Please share your screen with me.	

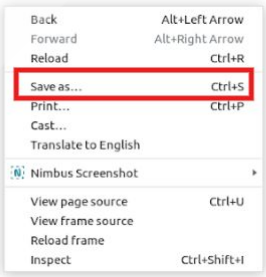
- **Ask Student to press ESC key to come back to panel**
- **Guide Student to start Screen Share**
- **Teacher gets into Fullscreen**

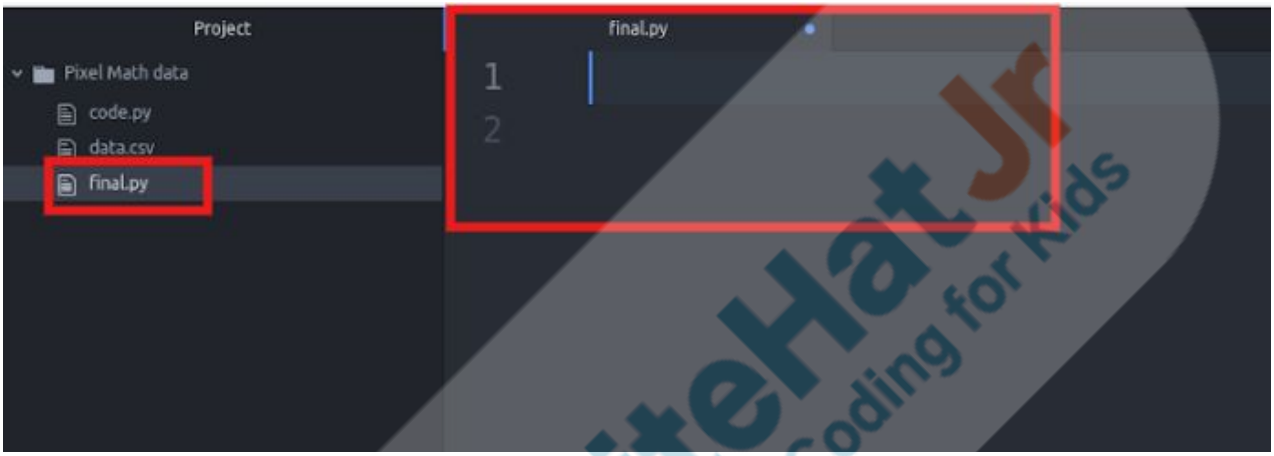
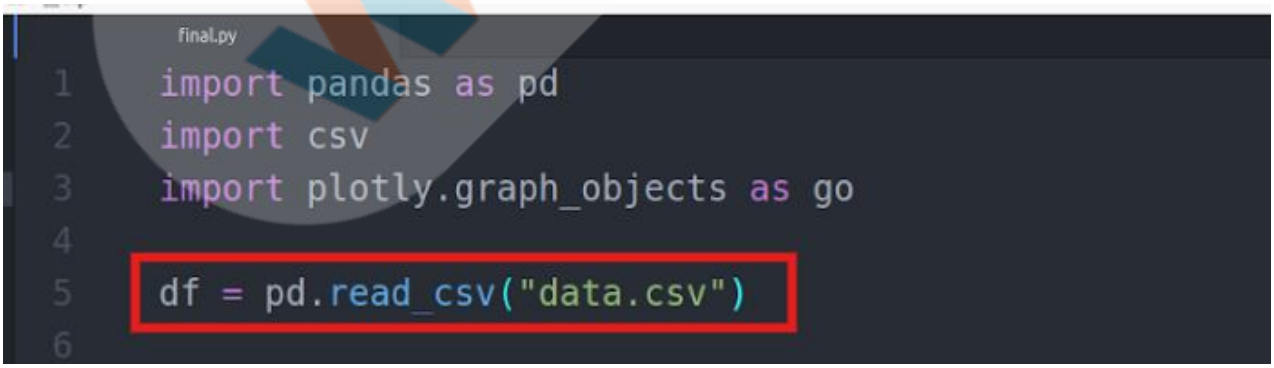
ACTIVITY

- **Filter the data for each student using tools in pandas.**
- **Use Plotly graph objects to represent the data and show the trouble spot for a particular student.**

Step 3: Student-Led Activity (15 min)	*Guide the student to download the data.csv file on their local machine and move it to a folder where they are working.*	*The student downloads the file and moves it to their working directory.* *<Student downloads data from student activity 1>*

```
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
TRL_xsl,Level 2,1
```



	Guide the student to create a new python file where they will write their code in the same working directory.	*The student creates a new file called final.py where they will be writing code for this class.*
		
	Guide the student to read the csv file and store the data in a dataframe object *Ensure that the student has csv and pandas installed.*	*The student writes code to read the data.csv file.*
		

	We want to filter out all data with student id TRL_987. df.loc will help us filter out all the rows with the given student id. We need to create a filter with df['student_id'] == "TRL_987" and pass it to df.loc. We will get a new dataframe object with only the given student id. You can print it. *Guide the student to write code for this.*	*The student writes code to filter out the data with student id TRL_987 and prints it.*
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		
	Now, you can find the mean of the attempts for each level for the student. You can also use the graph_objects in plotly to draw a bar graph for this. *Guide the student to draw a bar graph which visually represents the mean values for each level for the student.*	-

	What does the bar graph show?	*The student looks at the graph and analyzes the performance of the student in different levels.*

	Data visualization can be a powerful tool to give us findings which might go unnoticed otherwise. We'll keep learning more about data visualization and different data analysis in the upcoming classes.	-

Teacher Guides Student to Stop Screen Share

FEEDBACK

- **Appreciate the student for their efforts**
- **Identify 2 strengths and 1 area of progress for the student**

Step 4: **Wrap-Up** **(5 min)**	Let's review what we did in today's class.	**ESR:** - We took assessment data from a math practice app. - We processed it using dataframe objects and made visual representations to analyze the class performance and individual student performance for different concepts.

	I hope you found these simple data visualizations exciting in terms of the communication they make. We'll be doing many other and different kinds of data visualization. Next class, we will study a very useful pattern which all the collected data often follow!	-
Teacher Clicks 		
Additional Activities	*Encourage the student to write reflection notes in their reflection journal using markdown.* Use these as guiding questions: ● What happened today? - Describe what happened - Code I wrote ● How did I feel after the class? ● What have I learned about programming and developing games? ● What aspects of the class helped me? What did I find difficult?	*The student uses the markdown editor to write her/his reflection in a reflection journal.*

Activity	Activity Name	Links
Teacher Activity 1	raw data	https://raw.githubusercontent.com/whitehatjr/Data-Analysis-by-visualisation/master/data.csv

Teacher Activity 2	Solution	https://github.com/whitehatjr/Data-Analysis-by-visualisation
Student Activity 1	Raw data	https://raw.githubusercontent.com/whitehatjr/Data-Analysis-by-visualisation/master/data.csv