DOCUMENTATION
DATA ANALYSIS
OF KIDS ENROLLMENT DATASET

BY
SAURAV KUMAR SINGH
U21CH116
SVNIT,SURAT

TABLE OF CONTENT

❖ Problem Statement

A dataset sheet of a log of all the classes that have hosted on Zoom, with student attendance details and feedback scores given by students for each class.

- We'd like you to review the data, analyse it, and provide insights, observations, and actionable items based on your findings.

- Looking at the dataset, please suggest which timings and timezone should we schedule classes during launch? (Aim - our classes should sell out)

- Looking at the dataset, please suggest what should our class curation be. (which teachers should teach during launch, and which classes should they teach)

Overview:
*We are an online edtech startup operating in the US for kids aged **8-13 years**. Currently, we have a beta testing website where parents can sign up their kids for live online classes conducted in small groups of students.*

In the dataset, you'll find information collected after parents submit an enrollment for their child. We also collect feedback from kids after each class.

Here's what a few columns in the dataset indicate:
- ***Attended:*** *Did the child attend the class?*
- ***Concepts:*** *Were the concepts or topics taught in the class clear to the child?*
- ***Interested:*** *Was the topic interesting and enjoyable for the child?*
- ***Another Class?:*** *Is the child interested in taking another class with the same teacher?*

❖ Approach to Solve Problem

• Data Cleaning in Excel.

1. Delete columns like timestamp, ID for sequential class, class part number

2. Apply filter to every column

3. Empty cell from child name columns, parent columns

4. Remove outlier and empty cells from age columns like age can't be 118 or more

5. Remove wrong input from users like communication preference cannot be numeric and add Not Given as a defeat value to all empty values of the communication preference column.

6. Deleted column source details because it was not useful in giving insides and added 'Other' as the default value for the source column.

7. Deleted all empty values and wrong input values of a time zone

8. Deleted all empty values of Phase Mapping

9. Deleted all empty values of the topic column where the Attended Column has empty values. Then delete the remaining empty values of the Attended Column.

10. Remove all duplicates from the table.

11. Modify want another slot columns empty value into NO and means blanks into NO.

12. Delete N/A and empty values from final Score, yes, No, score- concepts, score interesting, and score class columns.

13. Parent name and child name columns values into Sentence case.

14. Columns like Concepts, Interesting, and Another class now contain "Unknown" where data was missing.

15. At last, Converted the format table by 'clicking format as Table'

• Pandas Profiling report of the Enrollment dataset for a better understanding of columns

⇒ Pandas Profiling report in HTML format. It contains a detailed exploratory data analysis (EDA) summary of the dataset. (*Python code file is being shared*)

- **Key components**:

$\Rightarrow$ **Overview Statistics**:
- o Number of variables (columns) and observations (rows).
- o Count of missing values.
- o Types of variables (numerical, categorical, datetime, etc.).

$\Rightarrow$ **Columns with Issues After data cleaning**:
- o Variables with high missing values or constant values.
- o Variables with high cardinality (too many unique categories).

$\Rightarrow$ **Correlations**:
- o Significant correlations between numeric variables.

$\Rightarrow$ **Key Highlights**:
- o Variables with the highest influence (e.g., strong correlations, highest variability).
- o Any specific warnings or anomalies.

$\Rightarrow$ [Link to Pandas Profiling Report](#)

- **Creating Pivot Table and Chart In Excel**

  - ➤ **Class Timing Analysis**: A table showing the start and end times of classes, along with the total number of children attending from all time zones, to determine the best time to hold classes.

  - ➤ **Age Group Analysis**: A table showing the ages of enrolled children and the number of children in each age group, along with a chart to identify which age group enrolls the most.

  - ➤ **Enrollment Sources**: A table showing the sources through which children are enrolling and the total number for each source, with a chart to visualize this data.

  - ➤ **Country-wise Enrollment**: A table showing the number of children enrolling from each country, along with a chart to identify which country has the highest enrollment.

  - ➤ **Interest in Additional Slots**: A table showing how many children want additional class slots versus those who do not.

  - ➤ **Popular Class Types**: A table listing the types of classes offered, the number of children enrolled in each type, and a chart to highlight the most popular class types.

  - ➤ **Attendance by Day**: A table showing the number of children attending classes on each day, along with a chart to visualize attendance trends.

- ➢ **Best Time Zone**: A table showing which time zone is most suitable for conducting classes so that all children can attend, along with a chart.

- ➢ **Communication Preferences**: A table showing the communication preferences of children, along with a chart to display this information.

- ➢ **Teacher Class Ratings**: Two tables showing which teacher should teach which class, along with ratings for concept clarity, class engagement, and overall performance.

- ➢ **Age-wise Teacher Preference**: A table showing which ages of children are attending classes with each teacher and how many are in each group.

- ➢ **Monthly Enrollment Trends**: A table showing the number of children enrolling each month, along with a chart to display the trends visually.

❖ Finding and Observation from Tables and Charts

- ➢ **Total Enrollments**: After cleaning the data, there are 4,047 valid enrollments.

- ➢ **Unique Students**: The total number of unique students is 430.

- ➢ **Unique Teachers**: There are 62 unique teachers.

- ➢ **Popular Class Timing**: The most children (249) attend classes between 4 PM and 5 PM across all time zones.

- ➢ **Age Group**: Children aged 8 to 13 enroll the most.

- ➢ **Enrollment Source**: Facebook is the most common source of enrollments.

- ➢ **Country-wise Enrollment**: The majority of enrollments come from the United States and Canada.

- ➢ **Class Slot Preference**: Most children prefer not to have an additional slot.

- ➢ **Popular Class Types**: "Onetime" and "Course" classes have the highest enrollments.

- ➢ **Best Time Zones**: The time zones with the most attendance are EST, MST, and CST.

- ➢ **Popular Class Days**: Most children attend classes from Monday to Friday.
- ➢ **Preferred Communication**: Gmail is the most commonly used communication method by children.

- ➢ **Top-rated Teacher**: Annalisa Ard is the highest-rated teacher for concept explanation and making topics interesting. She is especially popular for "Creative Writing 101," where students request additional classes, and her class receives the highest feedback scores.

- ➢ **Most Popular Teacher**: Aaron Potsick is the teacher with the most enrollments across all age groups and has the second-best rating after Annalisa Ard.

- ➢ **Enrollment Months**: The highest enrollments occur in May and February.

❖ Verifying 7 Most important findings using SQL.

> - *(SQL query file is being shared in the given drive)*
> - **Verification Process**: We are selecting a random sample of 3,000 enrollments (75% of the cleaned dataset with 4,047 enrollments) for verification.
> - **Purpose of Sampling**: The idea is to confirm that the findings are reliable. If the insights are consistent for both the full dataset and the sample, they can be trusted.
> - **Inconsistent Findings**: However, if the insights from the full dataset and the sample do not match, the findings cannot be considered completely accurate.

1. Through Facebook maximum number of enrolments had happened.

Source	No_of_children
Facebook	1711
Friends and Family	628
Other	368

2. From the United States maximum number of enrolments had happened.

Country	No_of_children
US	1378
IN	381
Other	315
CA	307
GB	109
AU	82
TW	81
PR	36
MX	28
HK	18
IE	18
JP	15

3. Onetime and Course are types of classes that have maximum enrolment.

Type	No_of_children
Onetime	1328
Course	750
Ongoing	412
Playlist-2	210
Playlist-1	88

4. Maximum number of children enrolling from Monday to Friday Classes.

Day	No_of_children
Monday	554
Thursday	539
Friday	484
Wednesday	459
Tuesday	401
Saturday	230
Sunday	121

5. The most common Preference of communication by the children is Gmail.

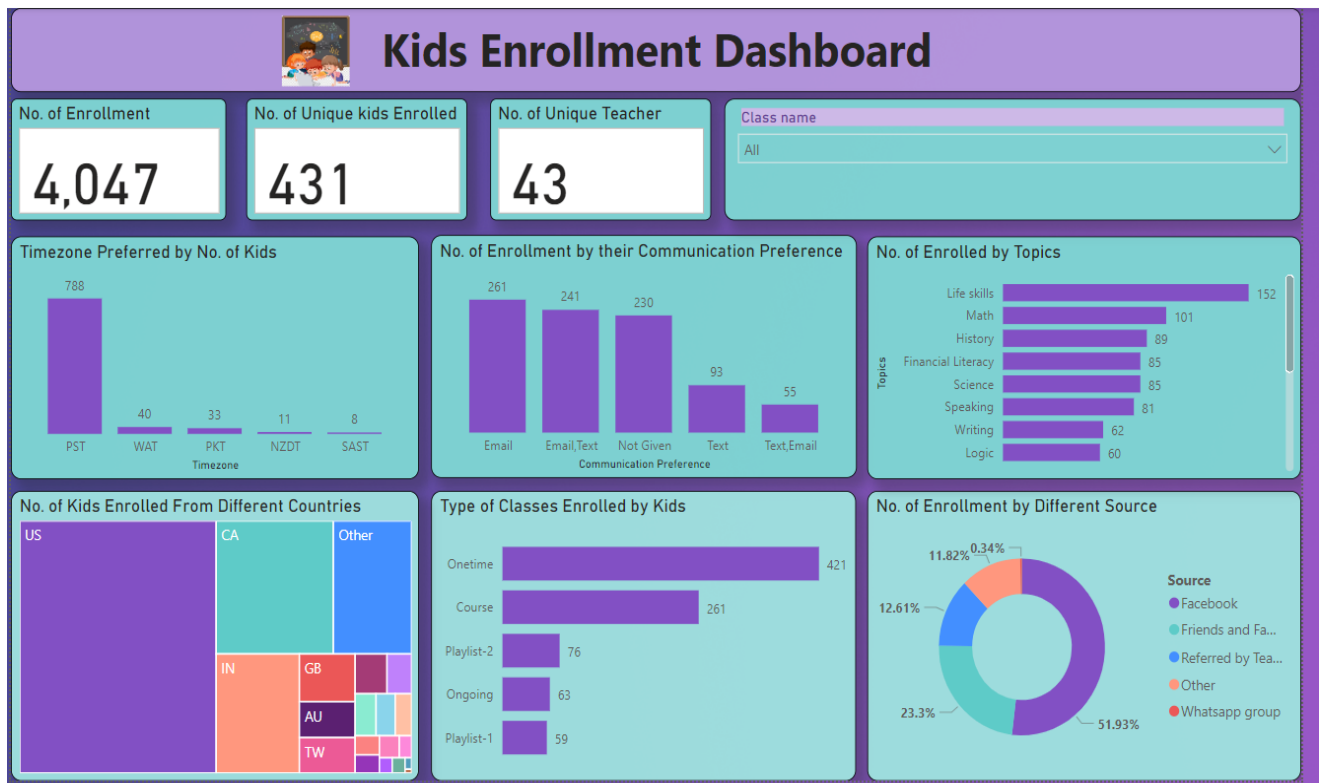| Communication_Preference	No_of_children
Email	1125
Not Given	875
Text	397
Email,Text	298
Text,Email	91
WhatsApp,Email	1
WhatsApp	1

6. Maximum number of children attending classes between 4 to 5 pm from all time zones.

Start_time	End_time	Timezone	No_of_student
16:00	17:00	EST	64
15:00	16:00	EST	52
16:00	17:00	CST	51
16:00	17:00	MST	47
15:00	16:00	CST	45
15:00	16:00	MST	39
17:00	18:00	EST	38
13:00	14:00	EST	32
14:00	15:00	EST	32
16:00	17:00	PST	30
12:00	13:00	EST	29
13:00	14:00	MST	25

7. Annalisa Ard is the best-rated teacher in concept explanation and making topics interesting children also demanded another class (specifically for Creative Writing 101) and the final score of the class was highest during feedback.

Teacher_Name	Class_name	No_of_children
Annalisa Ard	Creative Writing 101	23
Dana Lorelle	Introduction to Creative Writing	21
Rae Wilson	Introduction to Coding with Python	19
Amalia Fernand	Science Explorer Club	16
Ms. Courtney	Unveiling The Secrets of Anatomy	15
Dana Lorelle	Vocabulary Workshop	15
Amalia Fernand	Crazy About Crystals: Mineral Testing	14
Misses Dee	Mind Maze Adventures	14
Austin Stites	Introduction to Fractions with Pokemon Battles ...	13
Cara Bailey	Draw It: Halloween Cat	12
Tia Speece	The Week In News: A Current Events Discussio...	12
Aaron Potsick	Tuesdays with Aaron	12

❖ Power BI Dashboard for Representing my Findings

Kids Enrollment Dashboard

No. of Enrollment	No. of Unique kids Enrolled	No. of Unique Teacher	Class name
4,047	431	43	All

Timezone Preferred by No. of Kids
- PST: 788
- WAT: 40
- PKT: 33
- NZDT: 11
- SAST: 8

No. of Enrollment by their Communication Preference
- Email: 261
- Email,Text: 241
- Not Given: 230
- Text: 93
- Text,Email: 55

No. of Enrolled by Topics
- Life skills: 152
- Math: 101
- History: 89
- Financial Literacy: 85
- Science: 85
- Speaking: 81
- Writing: 62
- Logic: 60

No. of Kids Enrolled From Different Countries
US, CA, Other, IN, GB, AU, TW

Type of Classes Enrolled by Kids
- Onetime: 421
- Course: 261
- Playlist-2: 76
- Ongoing: 63
- Playlist-1: 59

No. of Enrollment by Different Source
- 11.82%, 0.34%
- 12.61%
- 23.3%
- 51.93%

Source
- Facebook
- Friends and Fa...
- Referred by Tea...
- Other
- Whatsapp group

Top classes Enrolled by Kids
- An introduction to Chess!: 122
- Long Division: 103
- Science Explorer Club: 93

Teacher Name
All

Months wise Enrollment of Kids
- January: 113
- February: 45
- March: 68
- April: 92
- May: 264
- June: 68
- July: 45
- August: 82
- September: 33
- October: ...
- November: 40
- December: ...

No. of Enrollment by Attendance
- No: 477
- Yes: 244
- Unenrolled: 134
- Class cancelled: 25

Top Teacher Preferred by No. of Kids
- Aaron Potsick: 308
- Annalisa Ard: 286
- Amalia Fernand: 255
- Misses Dee: 241
- Dana Lorelle: 227

Days wise Enrollment of Kids
- Monday: 170
- Thursday: 144
- Friday: 141
- Tuesday: 140
- Wednesday: 119
- Saturday: 92
- Sunday: 74

How much the classes were Interesting rated by kids
- 3.35%
- 10.27%
- 86.38%

Score - Interesting
- 3
- 2
- 1

How much the classes were able to clear concept rated by kids
- 2.17%
- 10.83%
- 87%

Score - Concept
- 3
- 2
- 1

❖ Suggestions and Actionable

➢ **Suggested Class Timings**: Based on the analysis, classes should be scheduled between 3 PM and 5 PM (PST). For time zones, focus on EST, MST, and CST during the launch to ensure high enrollment and attendance.

➢ **Class Curation Plan**: Recommendations for which teachers should teach and the specific classes they should handle during the launch are provided below.

Teacher	Class Name
Annalisa Ard	Creative Writing 101
Dana Lorelle	Introduction to Creative Writing
Rae Wilson	Introduction to Coding with Python
Amalia Fernand	Science Explorer Club
Ms. Courtney	Unveiling The Secrets of Anatomy
Dana Lorelle	Vocabulary Workshop
Amalia Fernand	Crazy About Crystals: Mineral Testing
Misses Dee	Mind Maze Adventures
Austin Stites	Introduction to Fractions with Pokémon Battles - Multiplication & Division
Cara Bailey	Draw It: Halloween Cat
Tia Speece	The Week In News: A Current Events Discussion Class
Aaron Potsick	Tuesdays with Aaron
Erin Coley	Introduction to Acting!
Kevin Lee	Chess Fundamentals Unveiled: A Beginner's Journey in Mastery
Ashton Khan	Wealthy Mindset: Dream Big
Paige Howell	Reading Sleuths - Find the Clues to Better Reading Comprehension and Fluency
Cara Bailey	Draw It: Dolphin
Dana Lorelle	Introduction to Poetry
Misses Dee	Creative Writing: Write a Fanfiction Short Story
Nicole Pauling	The Red Class: Scary History, Legends, Myths, and Cryptids - Winchester Mystery House

Annalisa Ard	Creative Writing with Miss Ard
Aaron Potsick	Algebra with Aaron
Ms. Courtney	The Statue of Liberty & Independence
Aaron Potsick	Logic Detective: How to Spot Bad Arguments and Errors in Logic
Erin Coley	Impromptu Speaking Competition Club
Ms. Courtney	Our Eyes and Sight
Cara Bailey	Draw It: Thanksgiving Turkey
Ms. Courtney	What are Rainbows and how are they made?
Misses Dee	4-Day Camp for Note-Taking Practice: Take Notes With the Teacher
Annalisa Ard	Explore World Mythologies Through a Cat's Eyes: Cats in Mythology
Aaron Potsick	Unlock Your Social Skills: Confidence, Communication and Friendship
Nicole Pauling	History Mystery: Nellie Bly
Nicole Pauling	Forgotten History: Mysteries and Extraordinary Stories - The Disappearance of Malaysian Flight MH370
Clark Vandeventer	Secrets Unveiled Series
Amalia Fernand	Scales and Slime: What are Reptiles and Amphibians?
Nicole Pauling	Forgotten Histories and Mysteries: The Dancing Plague
Annalisa Ard	Hogwarts Lessons in Creative Writing: Spellbinding Techniques
Keallah Smith	Banks and Money Management
Clark Vandeventer	Travel the World with Mr. Clark: South Korea and North Korea
Kevin Lee	Financial Literacy- Understand the Value of Money
Annalisa Ard	Timeless Tales by Annalisa
Lianne Brunt	Long Division

Marcos Del Cogliano	Beginner Spanish for Young Learners
Garrett Young	Acting & Improv Club for Beginners
Melissa Peterson	The Week In News: A Current Events Discussion Class
Aaron Potsick	The Art of Argument
Nicole Pauling	Forgotten History: Mysteries and Extraordinary Stories - Egyptian mysteries and riddles
Montse Limon	Word Problems
Annalisa Ard	Virtual Field Trip: Travel Across the World - Thailand
Nicole Pauling	Forgotten Histories and Mysteries: Wendigo
Amalia Fernand	Fun with Fossils: Discover Palaeontology!
Kristen Vale da Serra	Weather Explorer Club
Cara Bailey	Secrets Unveiled Series
Nicole Pauling	Escape from Alcatraz
Nathaniel Noss	An introduction to chess!
Paige Howell	Become a Reading Detective
Dana Lorelle	Introducing William Shakespeare!
Paige Howell	Reading Sleuths: Find the Clues to Better Reading Comprehension
Annalisa Ard	History of Magic: Part 1
Annalisa Ard	History of Magic: Part 2
Dana Lorelle	Creative Writing: Hyperbole
Cara Bailey	Draw It: Sea Turtles
Erin Coley	Memorize Anything Quickly!
Cara Bailey	Getting Started With Procreate on Your iPad
Rae Wilson	Introduction to Essay Writing
Cassy Cochrun	Public Speaking and Debate
Ashton Khan	Learning to Overcome Fears and Be Confident in Yourself
Mary Daniels	Long Division
Nicole Pauling	Forgotten History: Mysteries and Extraordinary Stories - The Molasses Flood
Erin Coley	Fast Thinking Word Games

Ms. Courtney	Let's Explore the World! Virtual Field Trip: Great Barrier Reef
Lianne Brunt	Financial Literacy 101 (Grade 7 - Grade 9)
Erika Berry	3rd Grade Math Summer Camp Readiness
Dana Lorelle	Creative writing Paradox
Stefany Pineda	Fun with Fractions
Ashton Khan	Effective Study Habits: Improving Our Grades
Amalia Fernand	Junior Geologists: How do Rocks Form?
Erin Coley	Convince Anyone of Anything: Persuasive Speaking
Bill Thompson	The Week In News: A Current Events Discussion Class
Lianne Brunt	Financial Literacy Saving and Budgeting
Mariposa Enamorada	Financial Literacy Want and Needs and Hard Choices
Kristen Vale da Serra	The Dragon Who Hated Outdoors - Read Along and Comprehension
Glen	Long Division
Carreira Chess	Chess Club 500
Clark Vandeventer	Travel the World with Mr. Clark : England
Clark Vandeventer	Travel the World with Mr. Clark : India
Aaron Potsick	The Baxter Creek Murder: Murder Mystery Game
Nicole Pauling	Forgotten History: Mysteries and Extraordinary Stories - DB Cooper Escape
Montse Limon	Angles
Melissa Peterson	Spelling Bee
Montse Limon	Game theory
Stefany Pineda	Let's Travel to Peru and Learn Spanish Along the Way! Beginner Spanish
Vanleer T. Pierce	Algebra 1 Summer Camp
Annalisa Ard	Birds in Mythology

Paige Howell	Reading Comprehension Skills for Middle School Students
Annalisa Ard	Enchanting Prose: A Magical Writing Workshop with Harry Potter - Part 1
Keallah Smith	Letâ€™s Be Wise: How Do Credit Cards Work?
Cara Bailey	Draw It: Red Panda
Clark Vandeventer	Meet the President: Joe Biden
Montse Limon	Introduction to Geometry
Bill Thompson	Public Speaking & Debate - Level 1
Cara Bailey	Draw It: Dreaming Dragon
Kevin Lee	Master the Art of Communication
Paige Howell	Learn World Geography through Animals Around the World
Clark Vandeventer	Mind & Matter: Managing Time, Stress & Philosophical Thought
Alan Ronald Chernoff	Introduction and History of Money
Cara Bailey	Letter/ Number Drawing Class
Misses Dee	How to Write the Perfect Paragraph Step-By-Step and Guided Practice - Part 1
Misses Dee	How to Write the Perfect Paragraph Step-By-Step and Guided Practice - Part 2
Lianne Brunt	Creative Writing Series- Animals
Mariposa Enamorada	Science: Weathering and Erosion
Cassy Cochrun	Let's Tell a Story! Practice Public Speaking, Grammar & Sentence Structure
Rae Wilson	Middle School Essay Writing Booster
Keallah Smith	Time Value of Money
Cara Bailey	Draw It: Magical Unicorn
Carreira Chess	Chess Club
Bill Thompson	Guided Impromptu Speaking

Misses Dee	Creative Short Story Writing- Roll the Dice to Write!
Cara Bailey	Draw It: Polar Bear
Misses Dee	Journal Writing in the Diary of a Wimpy Kid Style!
Alan Ronald Chernoff	Financial Bubbles: History€™s Biggest Blunders
Vanleer T. Pierce	Solving Percentage Word Problems
Tia Speece	Master the Art of Communication
Andrea	Home Alone Math & Multiplication Escape Room
Bill Thompson	Public Speaking and Debate Part 1
Cassy Cochrun	Spelling Bee
Rae Wilson	2nd & 3rd Grade Reading Comprehension & Writing Paragraphs
Ms. Courtney	Unveiling Secrets of Skeletal System
Erin Coley	Impromptu Speaking Club: Thanksgiving
Keallah Smith	What is the Stock Market?
Eric Zeltmann	Street Fighting Math
Eric Zeltmann	How to stop a robot uprising - Mathematical Paradoxes
Clark Vandeventer	Historic Clashes: 5 Game Changing Battles
Annalisa Ard	Enchanting Prose: A Magical Writing Workshop with Harry Potter - Part 2
Nicole Pauling	Forgotten History: Mysteries and Extraordinary Stories - The Great Emu War
Cassy Cochrun	Disney Debate Club â€"Intro to Public Speaking & Debate Class

> **Focus on Facebook Marketing**: We should prioritize Facebook ads and posts to maximize class reach and enrollment.

> **Promote Onetime and Course Classes**: Encourage teachers to create more "Onetime" and "Course" classes to boost sales.

- ➢ **Standardize Gmail for Communication**: Use Gmail as the primary communication method for contacting students.

- ➢ **Schedule Classes Monday to Friday**: Conduct all classes from Monday to Friday for better attendance and engagement.

- ➢ **Expand Reach in Key Countries**: Focus on growing our reach in the United States, Canada, and India to achieve maximum business potential.

- ➢ **Highlight Top-Rated Teachers**: Reward and promote content from the highest-rated teachers, such as Annalisa Ard and Aaron Potsickis, on our platform.