# Analysis Report for EdTech Startup

**Prepared by: Sandeep Devra**

**Date: 26 December 2024**

## ➢ Table of Contents


1. Introduction
2. Data Cleaning Process
3. Data Visualization Approach
4. Insights and Analysis
    - Class Timings and Time Zones (Question 1)
    - Class Curation (Question 2)
    - Broader Data Analysis
5. Recommendations
6. Conclusion

## ➢ Introduction

This report provides an in-depth analysis of the data collected from an online EdTech startup operating in the US, focusing on kids aged 8-13 years. The data includes enrollment information and feedback from beta testing conducted on a website where parents sign up their kids for live online classes in small groups.

## ➢ Data Cleaning Process

*Initial Data Description:*

- Rows: 6509
- Columns: 38

*Cleaning Steps:*

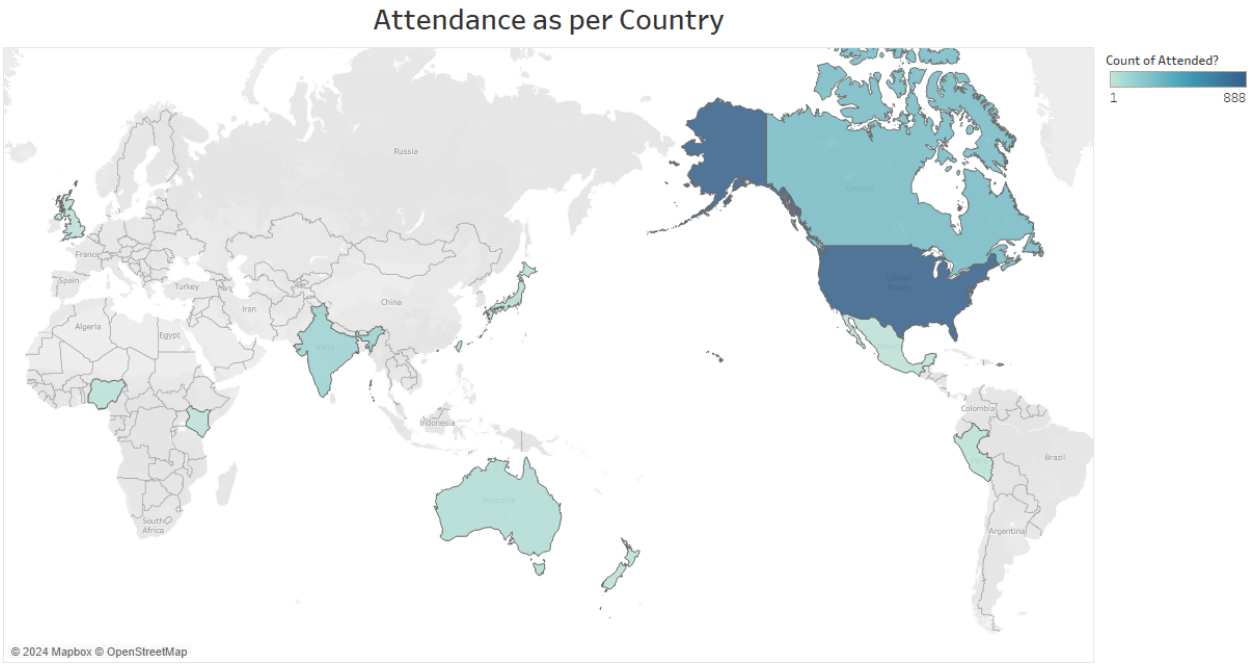
1. **Removal of Unnecessary Columns:**
    - Removed columns that were not relevant to the analysis, reducing the dataset's complexity and ensuring focus on key information.
2. **Handling Missing Values:**
    - Dropped rows with multiple null values to ensure data integrity and reliability.
    - Imputed missing values using statistical methods (mean, mode, median) to maintain dataset consistency and completeness.
3. **Final Dataset:**
    - Rows: 4684
    - Columns: 27
    - Column names: ['Timestamp', 'Parent name', 'Child Name', 'Age', 'Teacher Name', 'Source', 'Class ID', 'Type', 'Class

name', 'Date', 'Day', 'Start time (in PST)', 'End time (in PST)', 'Country', 'Region', 'City', 'Timezone', 'Phase Mapping', 'Topics', 'Attended?', 'Score - Concept', 'Score - Interesting', 'Score - Another class', 'Yes', 'Maybe', 'No', 'Final score', 'Hour', 'DayOfWeek']

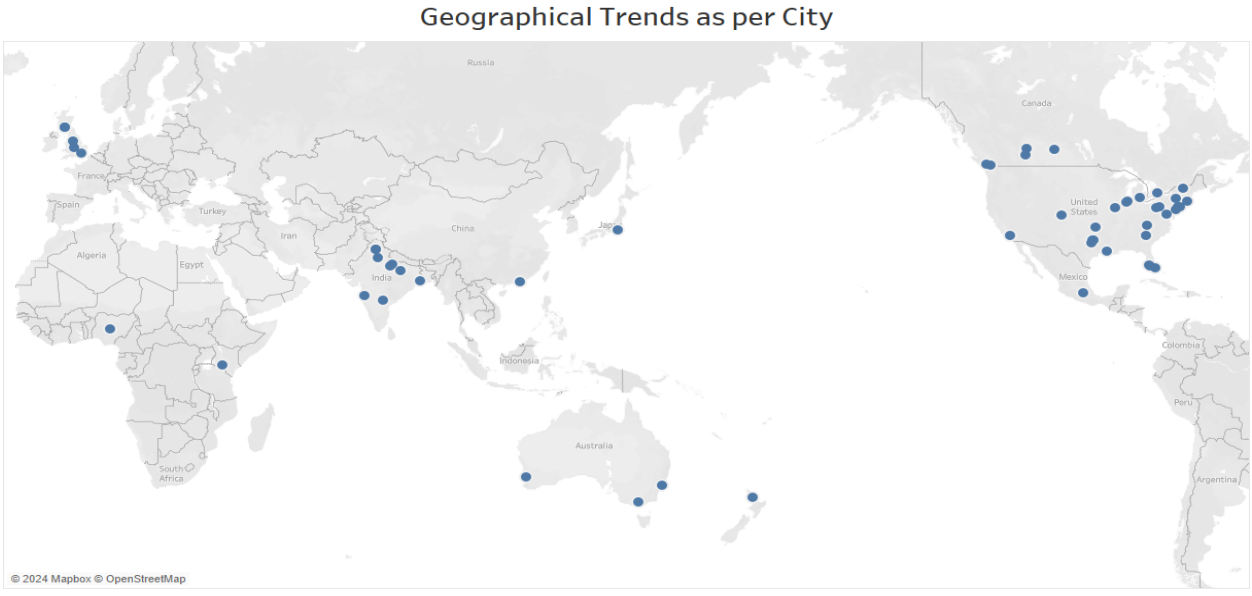
## ➢ Data Visualization Approach

Using Tableau, various visualizations were created to analyze different aspects of the data. The visualizations include:

1. **World Map Visualization:**
    - *Visualization 1*: Shows attendance based on city and country, filtered for attendees marked as "Yes."



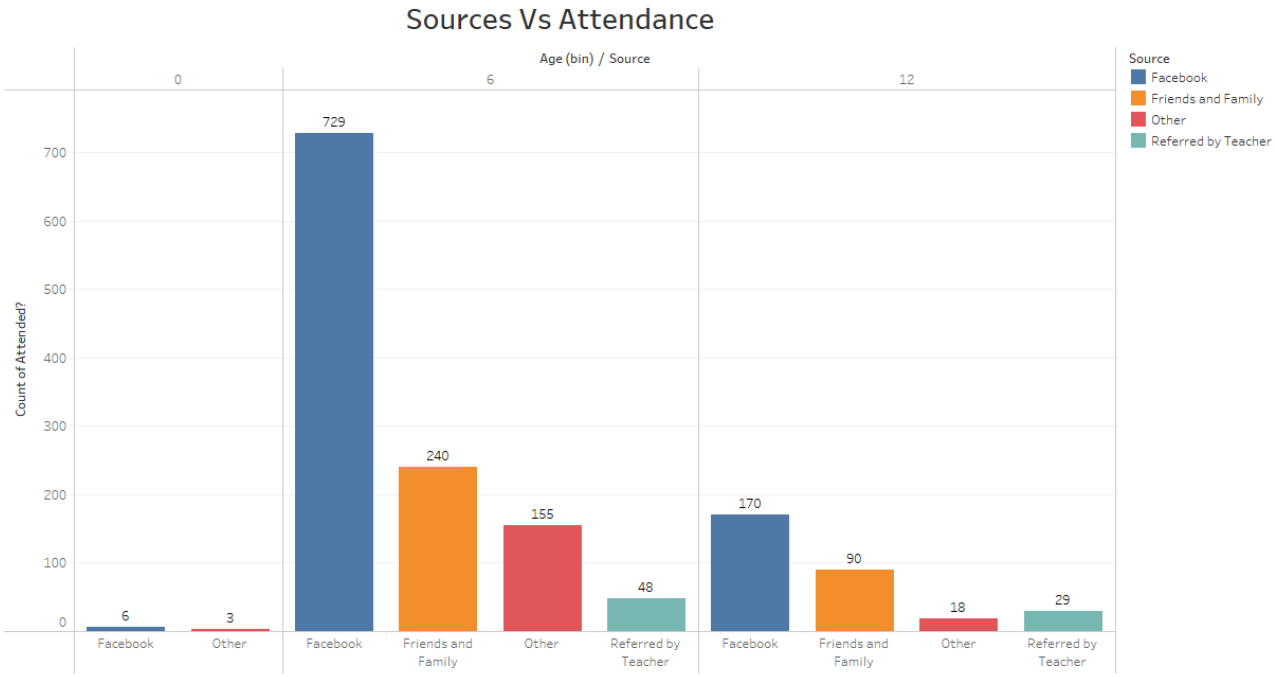

Map based on Longitude (generated) and Latitude (generated). Color shows count of Attended?. Details are shown for Country. The data is filtered on Attended?, which keeps Yes.




Map based on Longitude (generated) and Latitude (generated). Details are shown for Country and City. The data is filtered on Attended?, which keeps Yes.

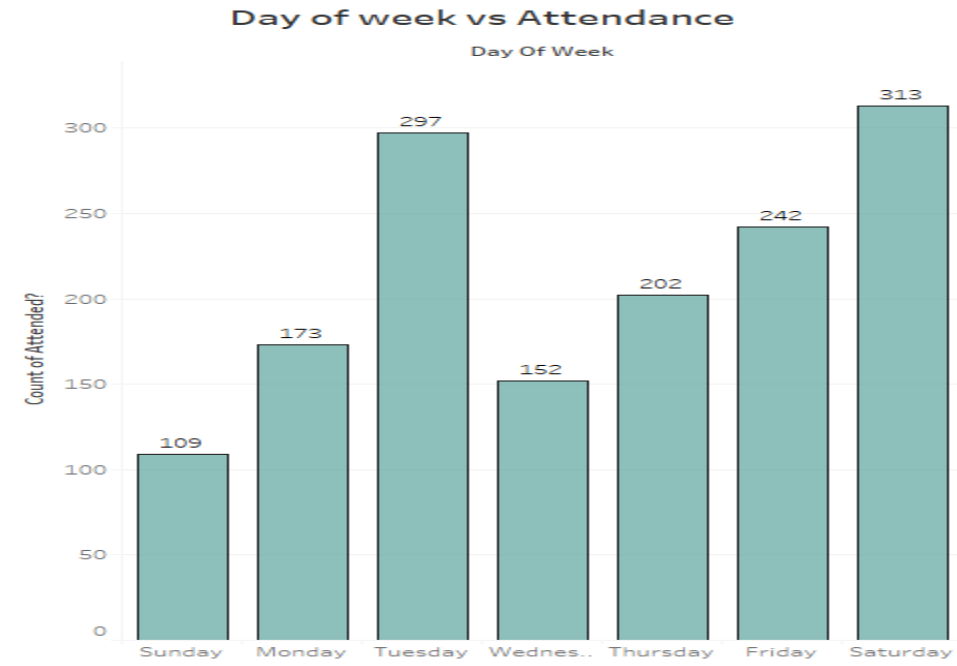
## 2. **Bar Charts:**

- *Visualization 2*: Sources vs. Attendance



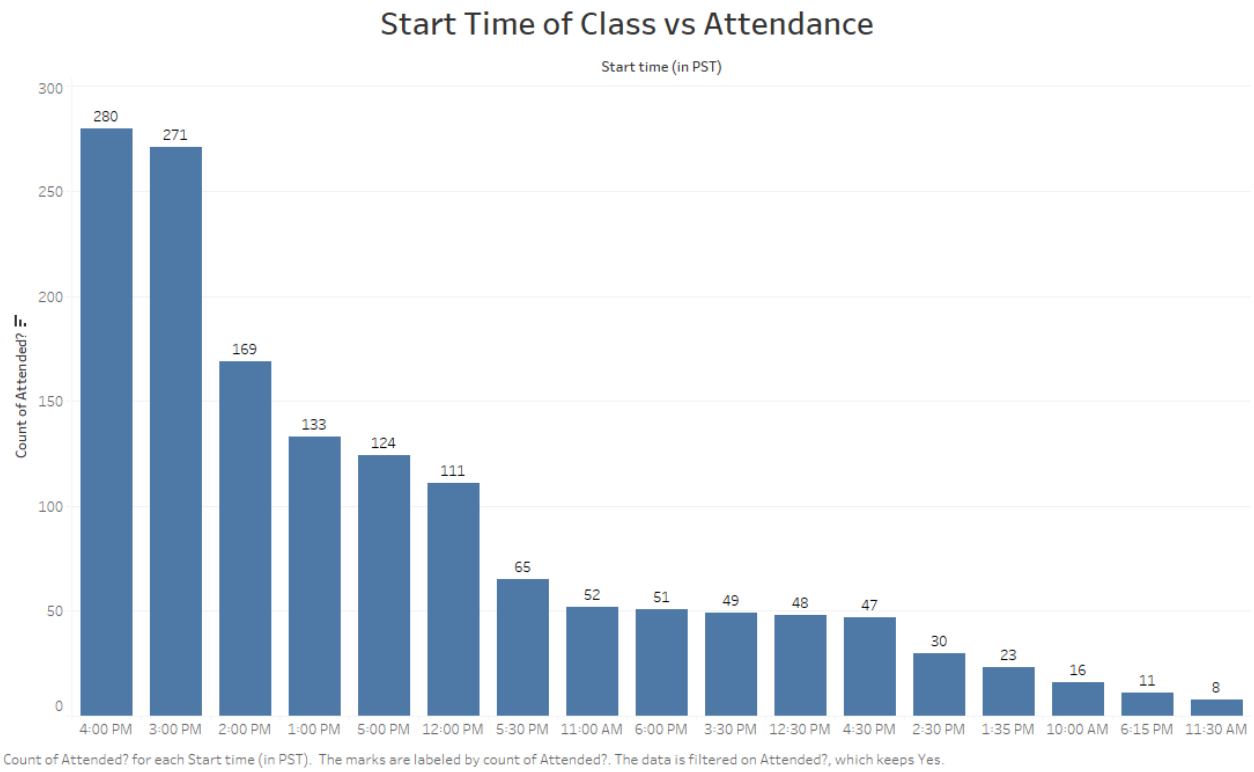

Count of Attended? for each Source broken down by Age (bin). Color shows details about Source. The marks are labeled by count of Attended?. The data is filtered on Attended?, which keeps Yes.

- *Visualization 3*: Attendance trends by day of the week

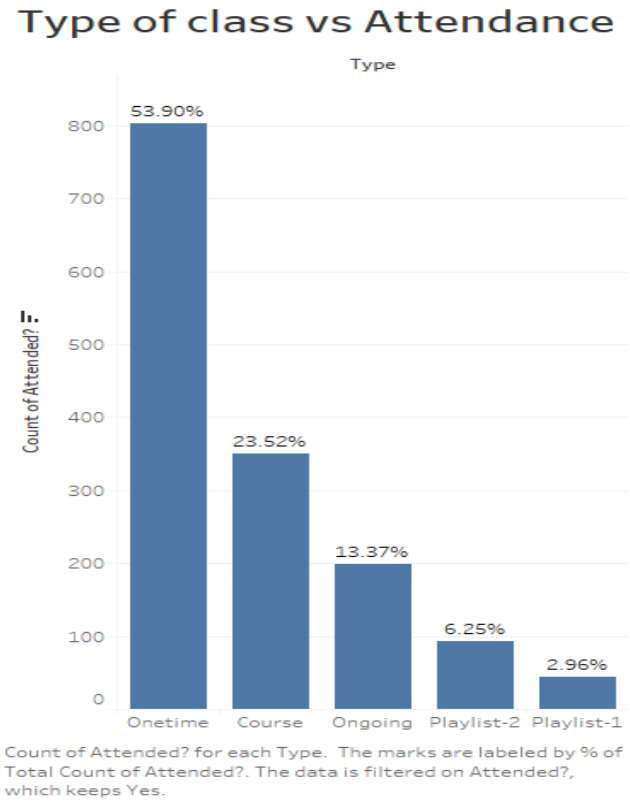



Count of Attended? for each Day Of Week. The marks are labeled by count of Attended?. The data is filtered on Attended?, which keeps Yes.

- *Visualization 4*: Start time of classes in PST vs. Attendance

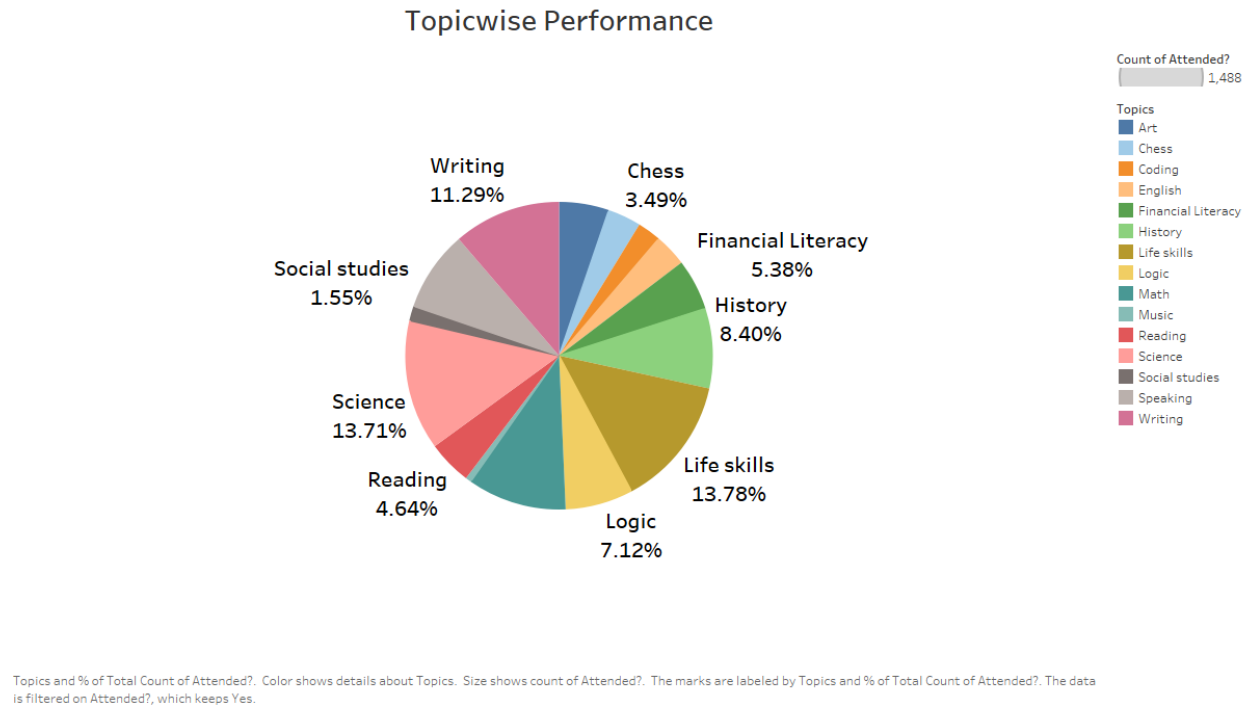



- *Visualization 5*: Type of class vs. Attendance



## 3. **Pie Charts:**

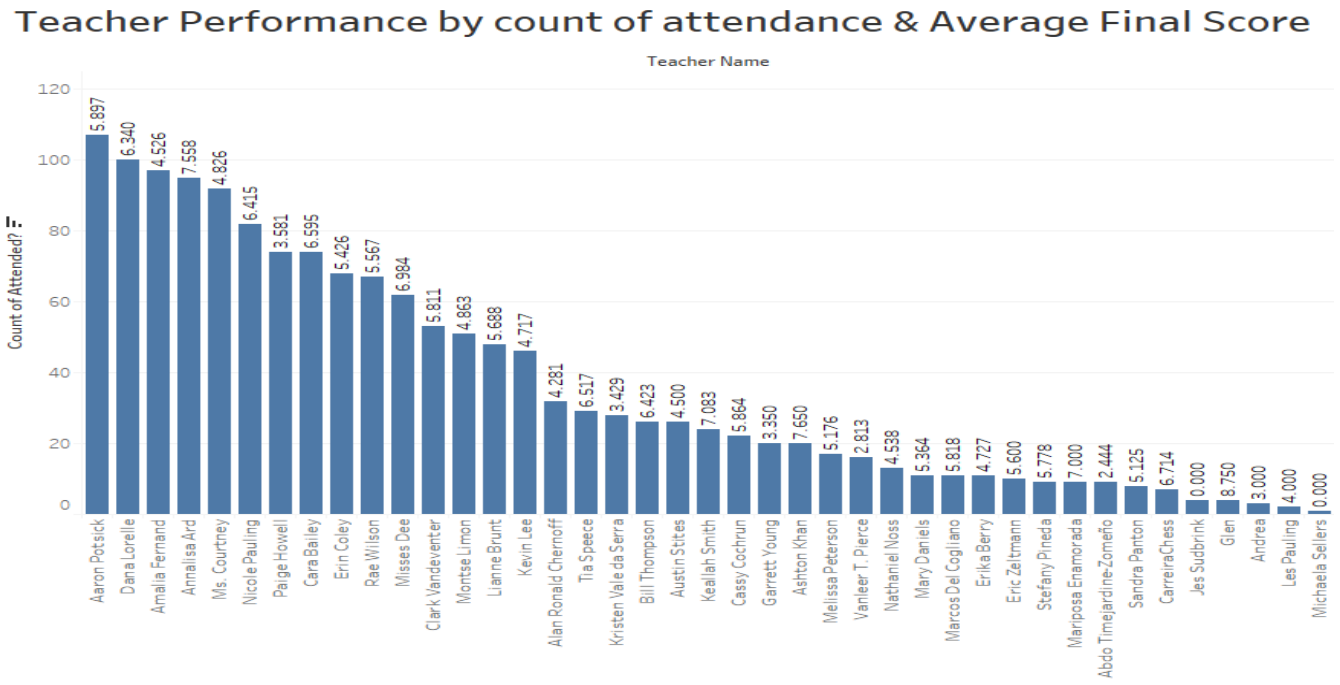
- *Visualization 6*: Topicwise Performance




Topics and % of Total Count of Attended?. Color shows details about Topics. Size shows count of Attended?. The marks are labeled by Topics and % of Total Count of Attended?. The data is filtered on Attended?, which keeps Yes.

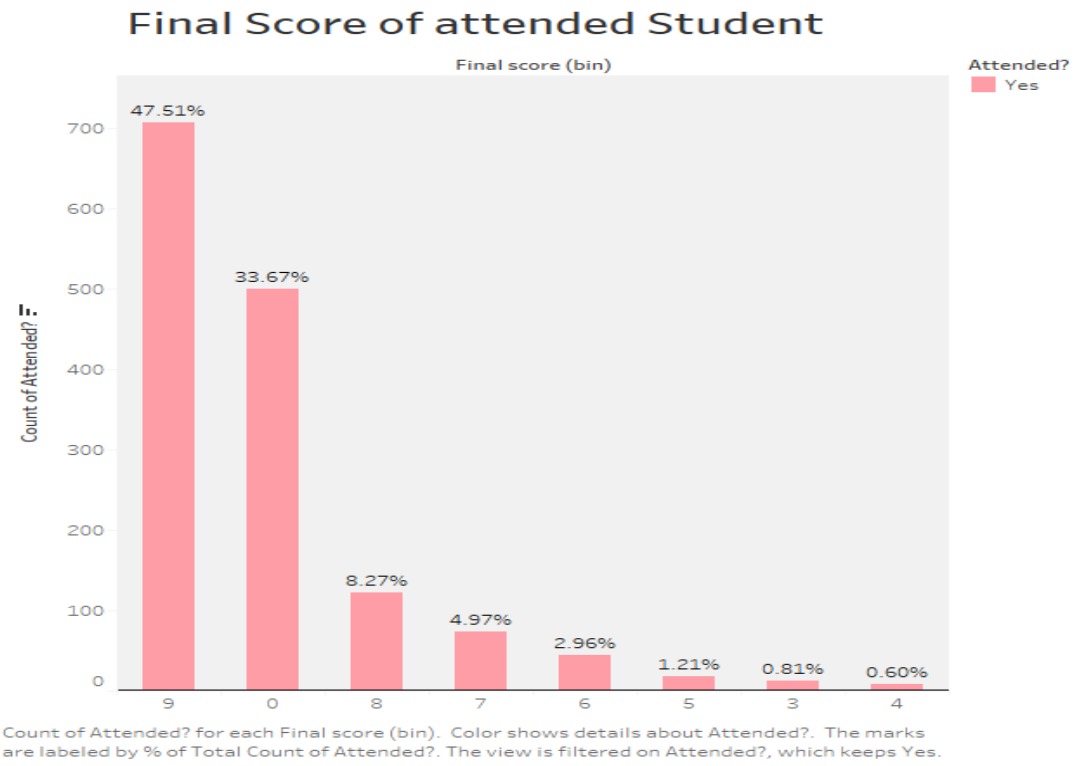
## 4. **Teacher Performance Analysis:**

- *Visualization 7*: Count of Attendance and Average Final Score

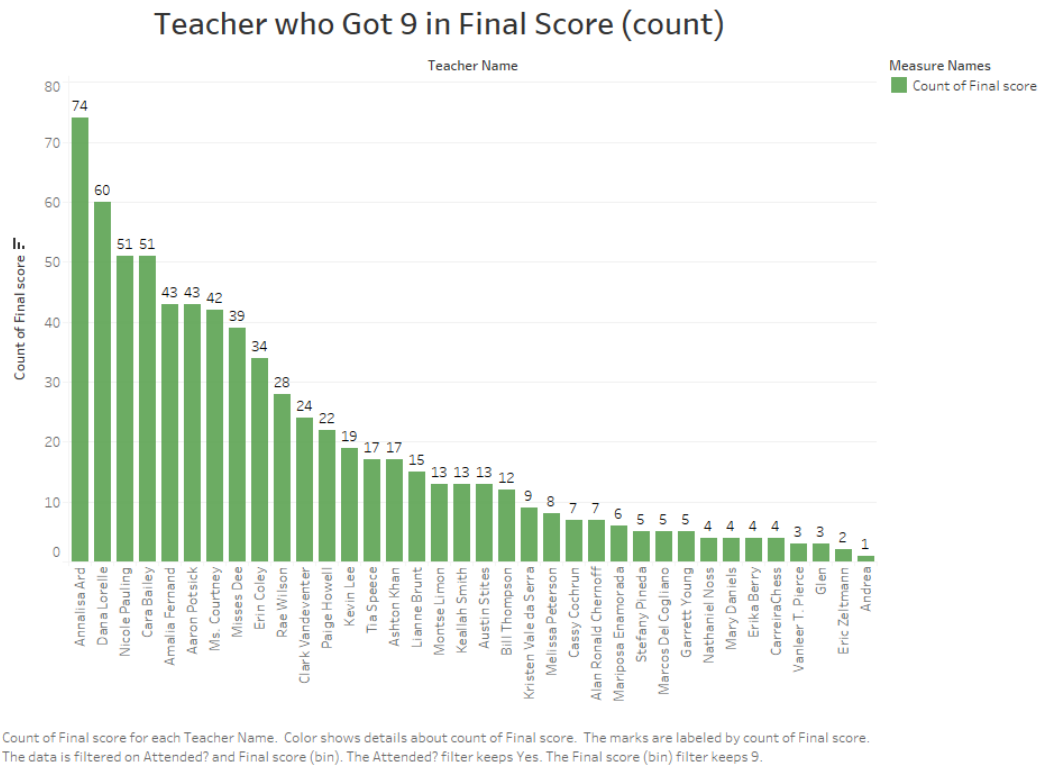



Count of Attended? for each Teacher Name. The marks are labeled by average of Final score. The data is filtered on Attended? and maximum of Final score. The Attended? filter keeps Yes. The maximum of Final score filter ranges from 0 to 9.

- *Visualization 8*: Final Score Distribution for attended students




- *Visualization 9*: Teachers who achieved a final score of 9



- **Insights and Analysis**

**Class Timings and Time Zones (Question 1):**

*Visualization 3: Attendance Trends by Day of the Week*

- This bar chart shows the count of attendance for each day of the week. The highest attendance is observed on Tuesday and Thursday, making them the most popular days for classes.

*Visualization 4: Start Time of Classes in PST vs. Attendance*

- This bar chart indicates the number of attendees for classes starting at different times in PST. The most popular start times are 10 AM, 2 PM, and 4 PM PST.

*Insights:*

- **Recommendation:** Schedule classes at 10 AM, 2 PM, and 4 PM PST on weekdays, particularly on Tuesdays and Thursdays, to maximize attendance.

**Class Curation (Question 2):**

*Visualization 7: Teacher Performance by Count of Attendance and Average Final Score*

- This bar chart compares teachers based on their attendance counts and average final scores. Teachers like Annalisa Ard, Diana Burrell, and Nicole Pauline have high average final scores and significant attendance, indicating strong performance.

*Visualization 8: Final Score Distribution for Attended Students*

- This chart shows the distribution of final scores for students who attended classes. A large proportion of students (47.51%) scored 9, followed by 33.67% who scored 0.

*Visualization 9: Teachers who Achieved a Final Score of 9*

- This chart highlights teachers with the highest counts of students achieving a final score of 9. Notable teachers include Annalisa Ard (74 counts), Diana Burrell (60 counts), and Nicole Pauline (51 counts)

*Insights:*

- **Recommendation:** Assign high-performing teachers such as Annalisa Ard, Diana Burrell, and Nicole Pauline to teach popular topics like Science (13.71%), Life skills (13.78%), and Writing (11.29%) for the launch phase.

➢ **Broader Data Analysis:**

*Visualization 2: Sources vs. Attendance*

- This bar chart breaks down the number of attendees from different sources, categorized by age bins (0, 6, and 12). The highest attendance source is “School Outreach,” followed by “Word of Mouth.”

*Insights:*

- **Actionable Item:** Focus marketing efforts on "School Outreach" and encourage "Word of Mouth" referrals to boost enrollments.

*Visualization 5*: Type of class vs. Attendance

- This bar chart shows the attendance count for different types of classes. The "Onetime" class type has the highest attendance (53.90%), followed by "Course" (23.52%).

*Insights:*

- **Actionable Item:** Emphasize "Onetime" and "Course" classes in your offerings, as they show the highest interest from students.

*Visualization 6*: Topicwise Performance

- This pie chart displays the distribution of attendance across various topics. The most popular topics are Life skills (13.78%), Science (13.71%), and Writing (11.29%).

*Insights:*

- **Actionable Item:** Increase the number of classes in Life skills, Science, and Writing to leverage high student interest.

- **Recommendations**

Based on the analysis, the following recommendations are made:

1. **Optimal Class Timings and Time Zones:**
    - Schedule classes at 10 AM, 2 PM, and 4 PM PST on weekdays, especially on Tuesdays and Thursdays, to maximize attendance.
2. **Class Curation and Teacher Assignments:**
    - Assign high-performing teachers such as Annalisa Ard, Diana Burrell, and Nicole Pauline to teach popular topics like Science, Life skills, and Writing.
3. **Focus on High Engagement Topics:**
    - Offer more classes in Life skills, Science, and Writing to leverage high student interest and engagement.
4. **Marketing Efforts:**
    - Focus marketing efforts on "School Outreach" and encourage "Word of Mouth" referrals to boost enrollments.
5. **Class Offerings:**
    - Emphasize "Onetime" and "Course" classes in your offerings, as they show the highest interest from students.

- **Conclusion**

This report provides a comprehensive analysis of the EdTech startup's dataset, offering insights and recommendations aimed at supporting business growth and improving their offerings. The visualizations and insights provided herein serve as a foundation for informed decision-making in the startup's upcoming launch phase.