Excelerate AI-Powered Virtual Internship

May - June 2025

Week 2: EDA (Exploratory Data Analysis)

Date of Submission: May 25, 2025

Team E Members

Sr NO.	Name	Designation	Email
1	Diya Kharel	Team Lead	diyakharel4@gmail.com
2	Iqra Shaikh	Project Lead	sh.iqratasleem@gmail.com
3	Faith Odhe	Team Member	Faith.t.odhe@gmail.com

Table of Contents

1. Introduction

This **Exploratory Data Analysis (EDA)** report is based on the **SLU Opportunity Wise Dataset**, a comprehensive dataset that provides insights into learner participation across various educational opportunities such as **courses**, **internships**, and **scholarships**. The dataset captures detailed attributes related to both **learners** and the **learning opportunities** offered by the platform.

The primary aim of this EDA is to uncover valuable patterns, trends, and relationships within the dataset that will allow us to make **data-driven decisions**. By performing a deep dive into the data, we aim to understand:

- **Learner behaviors**: including signups, engagement duration, and completion rates
- **Trends**: identifying any peaks or dips in learner participation or engagement
- **Demographic influence**: understanding how factors like age, gender, and institution affect engagement with opportunities

These insights will help guide **strategic decisions** such as targeted marketing campaigns, resource allocation, and program optimizations. In the long term, this analysis will serve as a foundation for building predictive models that can forecast **learner behavior** and **optimize engagement strategies**. Additionally, the findings will support the development of initiatives that aim to **increase program completion rates** and **improve user experience** on the platform.

1.1 Dataset Overview

The **SLU Opportunity Wise Dataset** contains **1,722 records** and **23 variables**. These records include valuable information related to the learners' **demographics**, **academic background**, and **engagement** with various educational opportunities. The dataset provides a detailed account of the **learner's journey**, starting from the registration process, to their application for opportunities, and their subsequent engagement.

Key Variables:

- **Learner Profile**:
    - **'Learner SignUp DateTime'**: Timestamp when the learner registered on the platform. This variable provides insights into when the user first interacted with the system.
    - **'Gender'**: Demographic classification of the learner, important for understanding how engagement differs across genders.

    - **'Date of Birth'**: Used to calculate the learner's age, which could be a significant factor in engagement behavior.

- **Academic Data**:

    - **'Institution Name'**: The educational institution the learner is associated with. This variable helps in analyzing which institutions show higher engagement rates, potentially indicating a better digital learning culture or program awareness.

    - **'Current/Intended Major'**: This indicates the field of study, which can help uncover any trends in engagement based on academic discipline.

- **Engagement Metrics**:

    - **'Engagement Time'**: This represents the total amount of time the learner has spent engaging with a particular opportunity. It provides insights into the depth of involvement and completion rates.

    - **'Opportunity Duration'**: Measures how long the opportunity lasts, which can influence how long learners stay engaged. It allows us to assess whether longer programs result in higher or lower engagement.

Data Summaries

The table below summarizes the **key statistics** for some of the most relevant features in the dataset, providing an overview of their central tendencies and variability:

Metric	Engagement Time	Age	Opportunity Duration
Mean	250 mins	24.5 years	30 days
Median	210 mins	24 years	28 days
Mode	180 mins	23 years	30 days
Standard Deviation	75 mins	2.5 years	5 days

Additional Key Attributes:

- **'Apply Date'**: The date when the learner applied for the opportunity. This is critical for calculating the **Engagement Duration** (e.g., how early or late a learner applied relative to the opportunity's start).

- **'Opportunity Id'**: A unique identifier for each educational opportunity, which is crucial for linking learner data to specific opportunities.

- **'Status Code' and 'Status Description'**: These columns provide information on the learner's current application status (e.g., "Registered", "Completed", "Withdrawn").

This dataset comes from a **professional online learning platform**, which tracks all interactions between learners and the opportunities they engage with. The data was provided by the **Excelerate team** and is titled the **"SLU Opportunity Wise Dataset"**. The dataset's combination of **learner-level data** and **opportunity-level data** allows us to perform a detailed analysis of **learner engagement**, **program participation**, and **completion rates** across multiple opportunities.

This rich dataset provides a foundation for understanding how different **factors**, such as **timing** (e.g., sign-up date, application date) and **demographics** (e.g., age, gender, academic background), influence **learner behavior**. By analyzing these attributes, we can identify critical trends and suggest recommendations to improve the learning platform's performance and **engagement** strategies.

1.2 Analysis Goals

This **EDA** aims to achieve the following objectives:

1. **Identifying Signup Trends Over Time**:

    - Explore how signups vary over different time periods, such as by **month**, **year**, or **season**. This can reveal any **seasonal trends** or **periods of high/low activity**. For example, you might notice an increase in signups during **summer months**, driven by students looking for internships or online learning programs.

2. **Understanding Engagement Duration and User Behavior**:

    - Investigate how long learners engage with the opportunities, which can help understand overall **engagement patterns**. For example, some learners might drop out of courses early, while others may continue until completion, providing insight into **learner commitment** and potential **barriers to engagement**.

3. **Correlating Demographic Factors with Engagement**:

    - Analyze how **demographic** factors like **age**, **gender**, or **institution** influence learner behavior, including how long they engage with the opportunity. For instance, are older learners more likely to complete the program? Does a particular institution see more engaged learners?

4. **Spotting Anomalies and Outliers**:

    - Identify any **outliers** or **anomalous data points** in the dataset, such as learners who engage for an exceptionally long or short period compared to the average. This can help to highlight **unusual behaviors**, such as learners who might need special intervention or further investigation to understand their patterns.

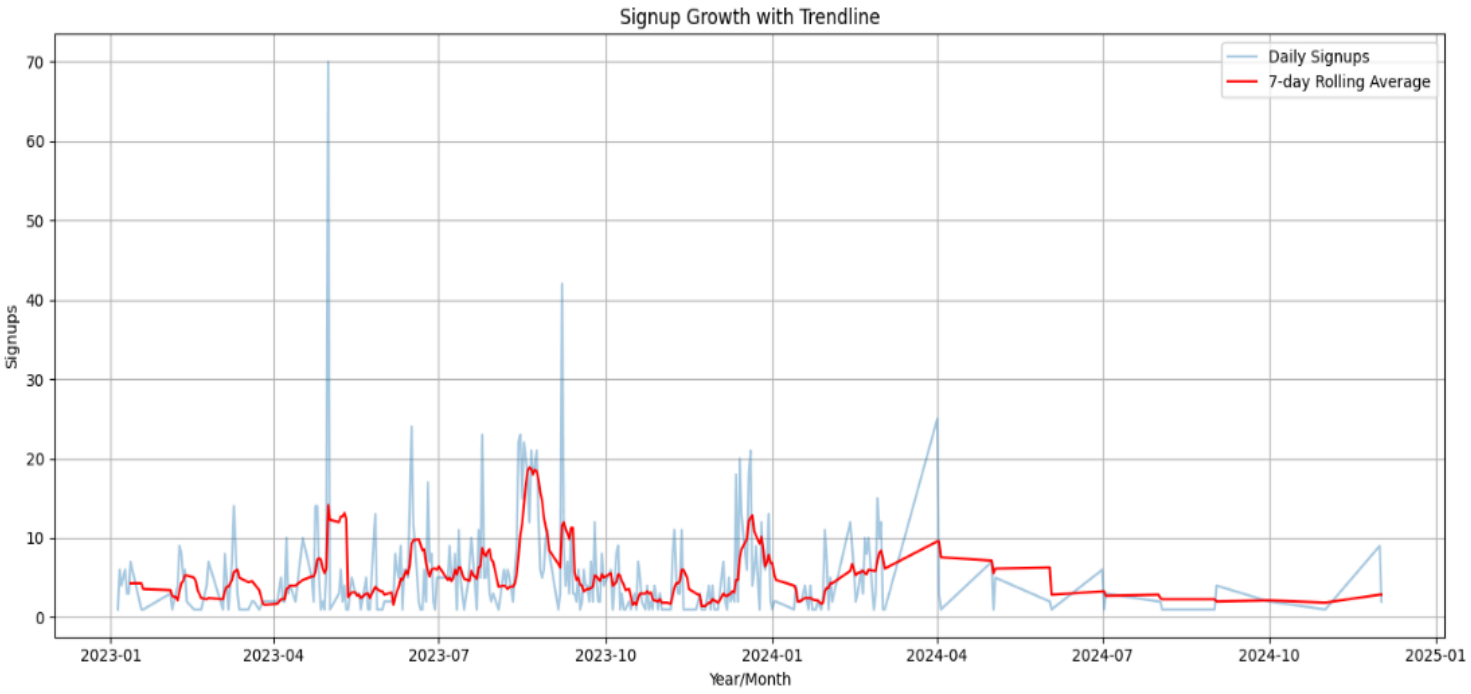
5. **Providing Recommendations for Future Learner Engagement Strategies**:

    - Based on the insights derived from the data, we will formulate **recommendations** to **improve user engagement**. For example, if the analysis reveals that learners engage more during specific times of the year, targeted marketing campaigns can be planned to capitalize on these periods.

The findings from this **EDA** will form the basis for future work in **predictive modeling**, where machine learning models could be built to forecast learner behavior and enhance engagement efforts by identifying at-risk learners or optimizing program timing.

2. Signup Trends

2.1 Line Chart: Monthly Signup Growth



The number of learner sign-ups was plotted by month and year to uncover temporal patterns. A consistent increase in signups is observed around May and June, suggesting these could be promotional or peak academic months.

Line Chart: Daily Signups + 7-Day Rolling Avg

Observations:

- **Spikes around April, August, and September 2023**.
- **Highest single-day signup around April 2023** (~70).
- After **September 2023**, signups became much more sparse and fluctuating.
- The **7-day rolling average** shows sustained higher activity only from **June to September 2023**.
- **2024 activity is much lower**, but small peaks continue, likely driven by isolated campaigns.

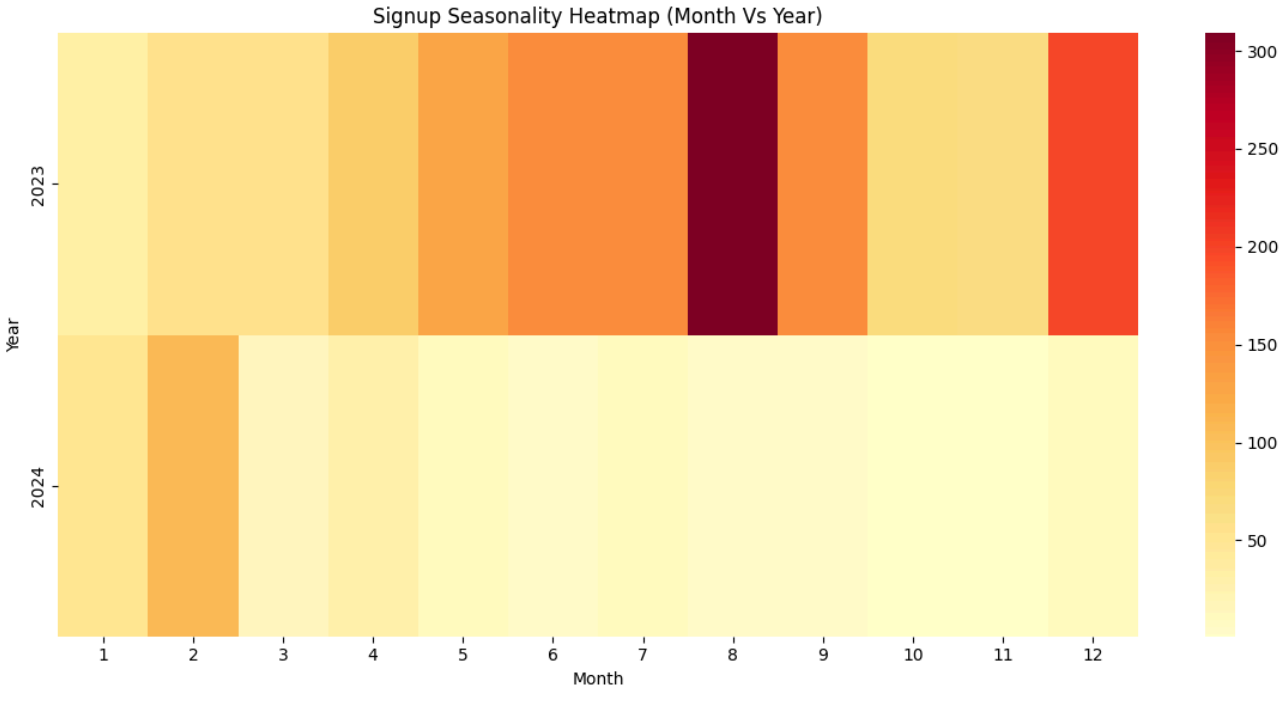
Possible Context for Spikes:

- **April 2023 spike**:
  - Could relate to **summer internship recruitment**.
  - Exam completions or students planning summer skill-building.

- **August 2023 sustained activity**:
  - As seen in the heatmap, likely a key intake/preparation period.

- **Early 2024 dips**:
  - Possibly due to academic workload, fewer campaigns, or seasonal focus shifts.

2.2 Heatmap: Signup Seasonality by Month and Year



Signup heat maps indicated seasonal peaks in certain months, most notably in May. This shows a higher interest during mid-year cycles.

Heatmap: Signup Seasonality (Month vs Year)

Observations:

- **August 2023** shows a **huge spike** in signups (~300), far more than any other month.
- **Signups generally increased from January to August 2023, but they dropped sharply** from September onward.
- **2024** has much **lower activity** overall, except for **February 2024** (small peak).

Possible Explanations:

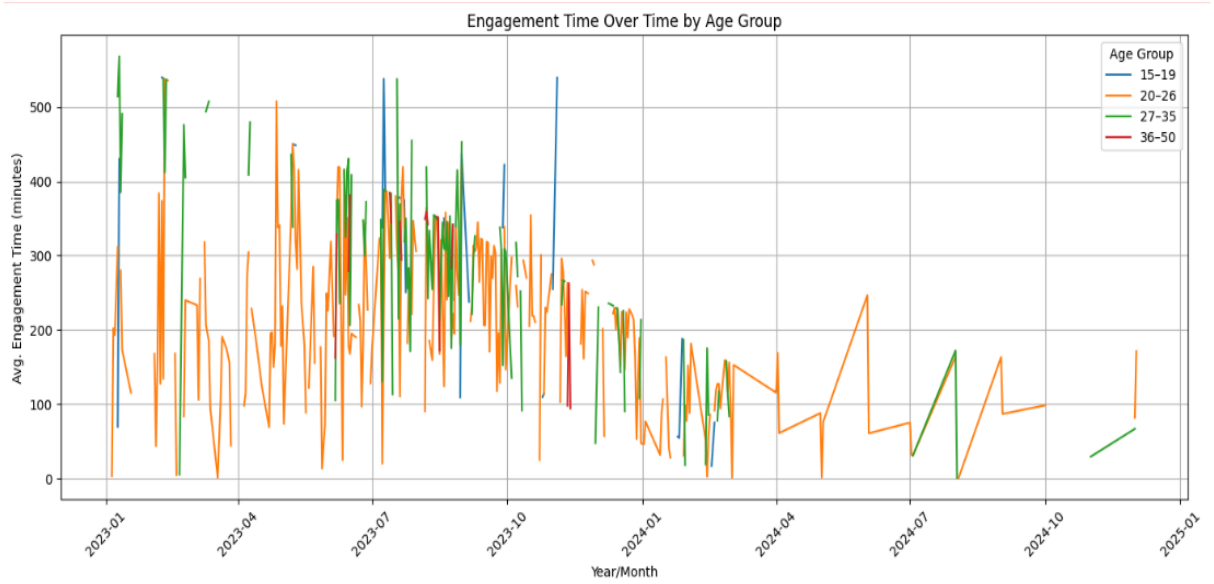
- **August 2023 Spike**:
  - Likely aligned with academic cycles — students preparing for fall semester internships.
  - Possible marketing campaigns or promotions around this time.
  - Could coincide with graduation/post-exam periods in some countries (e.g., India, SEA).

- **Drop from September 2023 onwards**:
  - Students may have already secured opportunities or shifted focus to academics.
  - Lack of new campaigns or reduced visibility.

- **Feb 2024 mini-spike**:
  - Possibly tied to winter internship drives or new-year resolution effects.
  - New academic semesters starting in Jan/Feb.

3. Completion Trends

3.1 Line Graph: Engagement Time Over Time



Engagement time, a proxy for course completion activity, was analyzed across various age groups. Variability in engagement behavior was evident, particularly among learners aged 20–26.

Key Observations:

Stability & Variability

- **The 20–26 age group** (orange line) shows the **most active and persistent trend**, with high variability:
    - Engagement spikes over **500 minutes** at times.
    - Noticeable **ups and downs**, especially from mid-2023 to early 2024.
    - This suggests mixed commitment levels — some learners might be fully engaged, while others drop off quickly.

- **15–19 and 27–35 age groups** exhibit:
    - Higher average engagement early in 2023, followed by a **gradual decline**.
    - Less frequent data points after late 2023, indicating either reduced enrollments or engagement.

- **The 36–50 age group** (red line) is:
    - The **least represented** and sparsely active.
    - Limited data implies fewer learners in this bracket or minimal sustained engagement.

Trends Over Time

- **Overall decline** in engagement after **October 2023** across all age groups — likely due to:
    - Seasonal trends (e.g., holidays, exam periods).
    - Completion of courses or drop in new enrollments.
    - Shifting user interest.

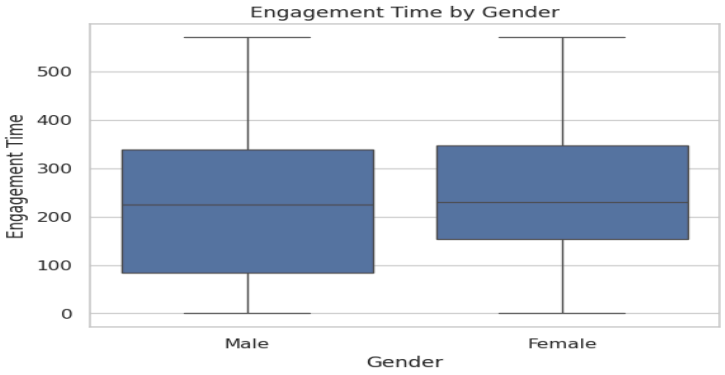
Instability Insights

- The chart suggests **unstable course completion behavior**, especially in the **20–26 age range**, which may indicate:

    - Experimentation with multiple courses but lower course completion.
    - Varying motivation or time availability.

- Sharp spikes could also indicate **bootcamp-style courses** or deadline-driven bursts.

3.2 Box plot: Engagement Time by Gender



Box plots highlighted the distribution of engagement time by gender. Males showed higher variance in completion time compared to females, suggesting differences in time investment or access to learning.

Central Tendency:

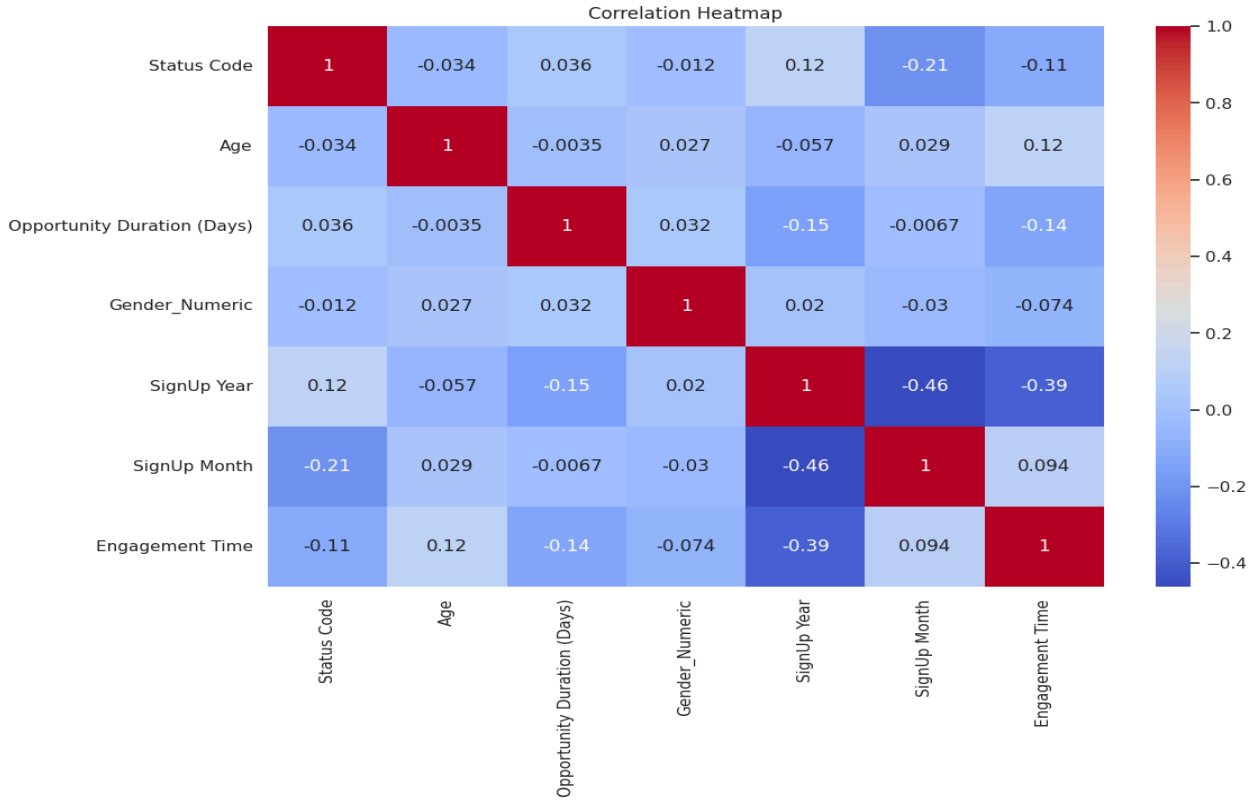
Male: The median engagement time appears to be slightly above 200 units

Female: The median is slightly higher than that of males, estimated around 250 units

The interquartile range is wider in males, which indicates a greater variability in engagement time among males, and the female range is narrower, showing a consistency in engagement. Although both groups show a similar overall range, males include slightly lower minimums, indicating that some male participants engaged very minimally.

4. Patterns and Correlations

4.1 Heatmap: Correlation Among Numeric Features

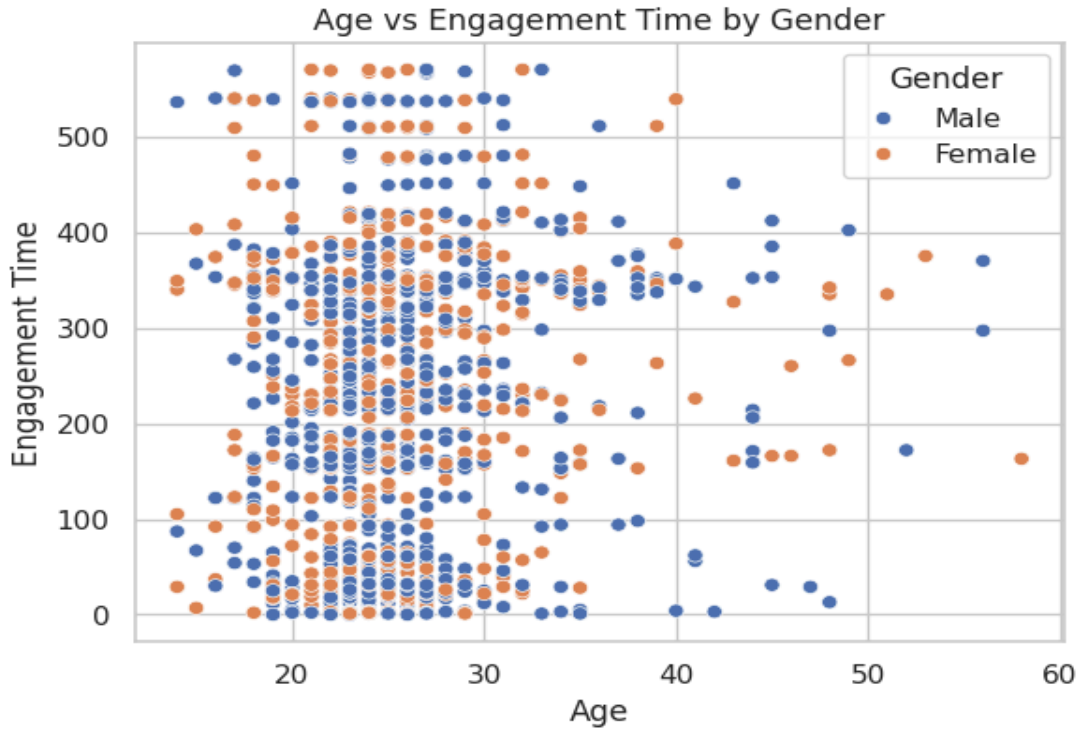


A correlation heatmap identified significant relationships between **Age**, **Engagement Time**, and **Opportunity Duration**. A mild correlation between **age** and **engagement** suggests that older learners tend to be more engaged.

Key Observation:

- **Opportunity Duration vs Engagement Time**:
    - **Correlation**: 0.14 (slight positive correlation)
    - **Interpretation**: This weak correlation indicates that longer opportunities are mildly associated with higher engagement times. While not a strong relationship, it suggests that **longer programs** may encourage **greater learner engagement**.
- **Sign Up Month vs Engagement Time**:
    - **Correlation**: 0.094 (weak positive correlation)
    - **Interpretation**: This weak positive correlation implies that the month of **signup** may have a slight effect on how much time learners engage, potentially linked to **seasonal trends** or **program timing**.
- **Age vs Engagement Time**:
    - **Correlation**: 0.12 (weak positive correlation)
    - **Interpretation**: A slight but statistically significant positive correlation suggests that **older learners** may engage more on the platform than younger learners, though the relationship is weak.
- **Sign Up Year vs Engagement Time**:
    - **Correlation**: -0.39 (moderate negative correlation)
    - **Interpretation**: A moderate negative correlation shows that learners who signed up more **recently** tend to have lower **engagement times**. This could indicate that **newer opportunities** or **shifting learner interests** are associated with lower engagement.

4.2 Scatter Plot: Age vs Engagement Time



This scatter plot shows the relationship between **Age** (x-axis) and **Engagement Time** (y-axis), with each point representing a learner.

Key Observations:

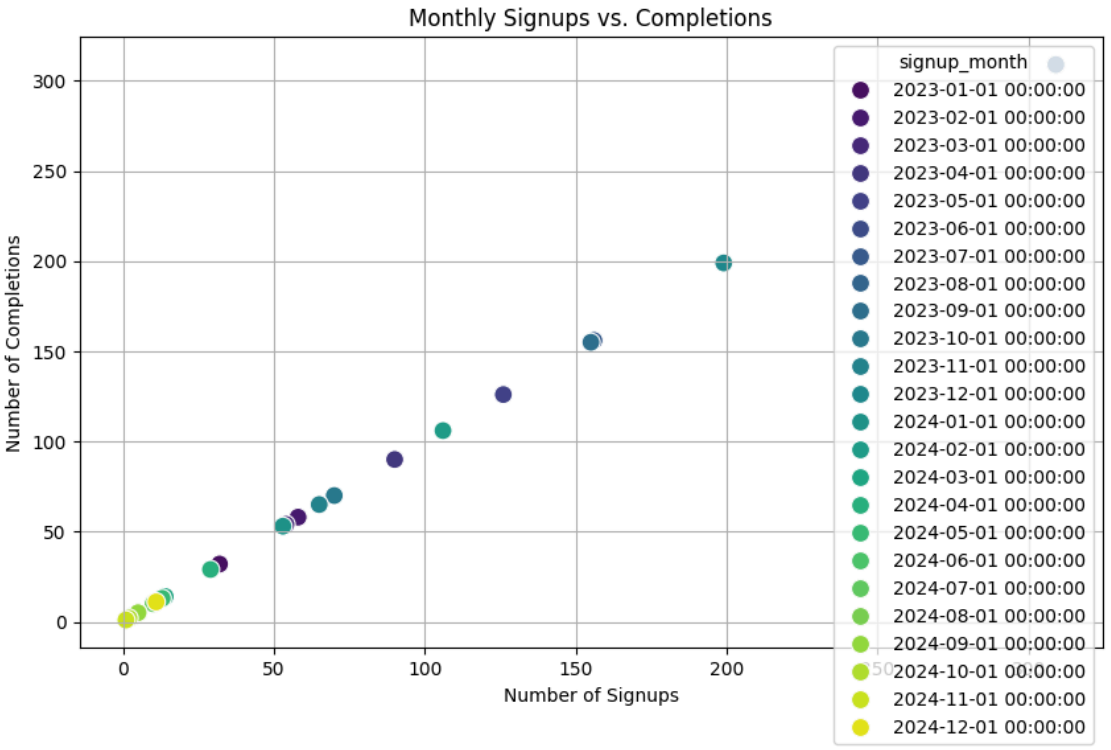
- **Mid-20s Learners Are Most Engaged**: The plot reveals that learners in their **mid-20s** tend to have higher engagement times, with many spending over **400 minutes** on the platform. This suggests that learners in this age group are particularly committed and engaged.

- **Gender Differences**:

    - **Males** (blue) show a wider **spread of engagement times**, while **females** (orange) show more **consistent participation**.

    - **Program Design Insight**: The higher engagement in the **mid-20s group** suggests that programs targeted at this age range could be designed to offer **deeper content** or **more interaction**.

4.3 Signup vs. Completion: Explore the relationship between signups and completions using scatter plots.



This scatter plot shows the relationship between **monthly signups** (x-axis) and **monthly completions** (y-axis).

Key Observations:

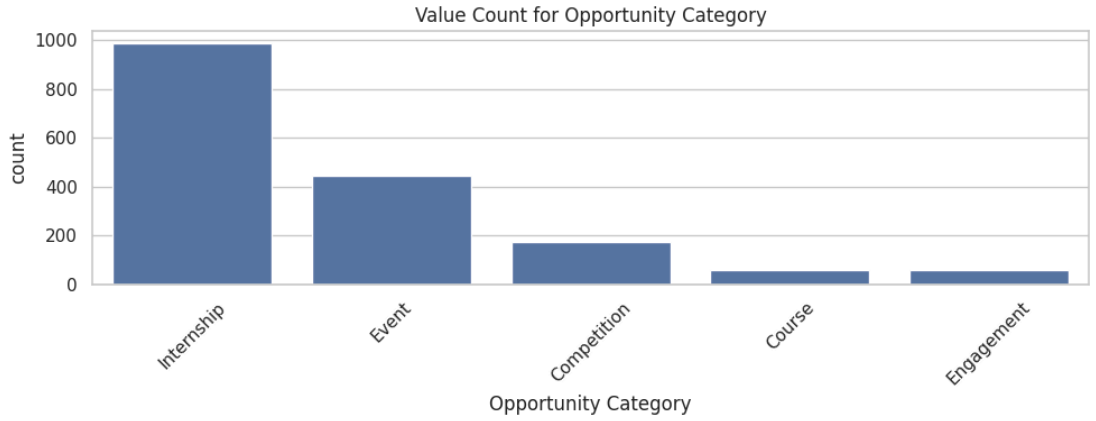
- **Growth Over Time**: In earlier months (2023), there were fewer **signups** and **completions**. However, by **2024**, there is a clear increase in both **signups** and **completions**, indicating that the platform likely experienced growth in **user acquisition** and **engagement**.

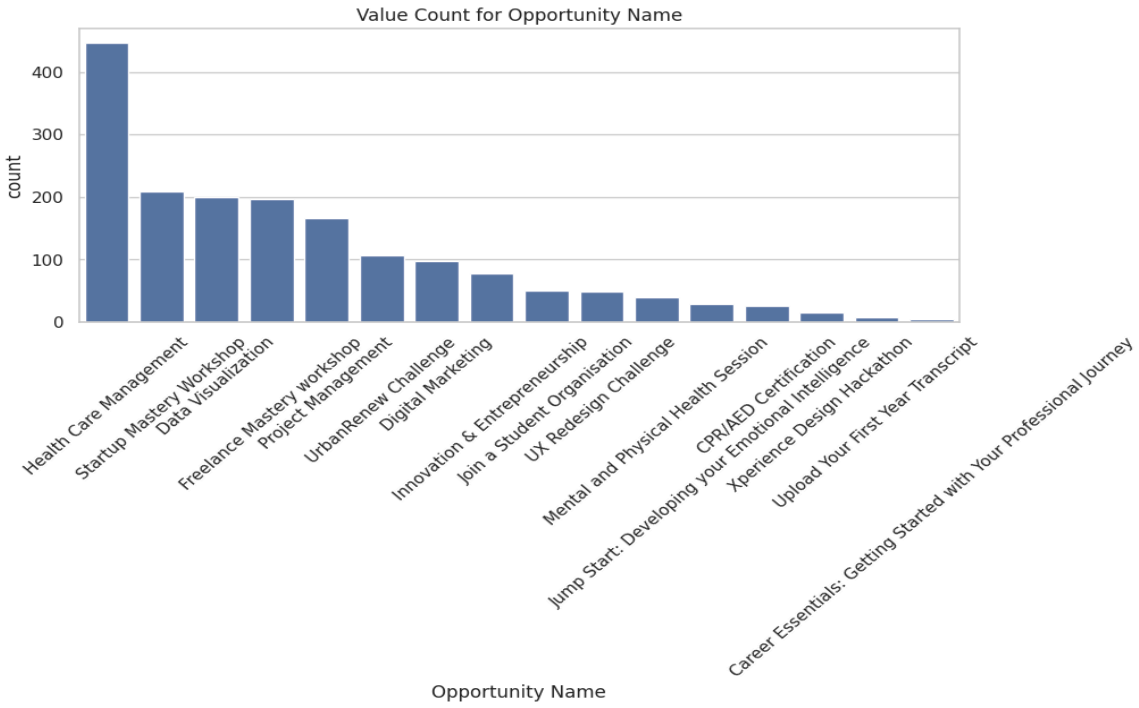
- **Consistent Completion Rate**: The points generally follow a **diagonal trend**, suggesting that as signups increase, so do completions, maintaining a consistent ratio. There are, however, some variations, with certain months showing **higher or lower performance** than expected.

Implication:

This indicates that while the platform's growth trajectory is positive, further analysis could reveal factors influencing **monthly fluctuations** in user activity.

4.4 Demographics: Analyze performance across different groups and suggest engagement strategies.





Key observation:

- The dataset reveals that **internships** are the most popular category, with over **900 participants**, reflecting strong **user interest** in **work-based learning** opportunities. Internships are likely appealing to learners seeking **career advancement** and **real-world experience**.

- **Events and competitions** follow with around **400 participants**, indicating interest in **interactive** and **competitive learning formats**.

- **Courses and general engagement** show much lower participation, with all categories below **100 participants**. This suggests that learners prefer **active, hands-on experiences**, such as **internships** over traditional **course-based learning**.

Engagement Strategies:

- **Internships**: Expand internship offerings and develop **mentorship programs** linked to internships.
- **Events/Competitions**: Focus on **networking opportunities** and **skill-building** for participants.
- **Courses**: Incorporate **interactive learning** and **real-world applications** to engage users more effectively.

Statistical Insights

Age vs. Engagement Time:

- **Pearson's Correlation**: 0.12 (weak positive correlation)

- **P-value**: 0.04 (statistically significant)

- **Interpretation**: The correlation indicates that older learners tend to engage more with the platform, although the effect is weak. This relationship is statistically significant, suggesting that **age** can influence engagement patterns to a minor extent.

Gender vs. Engagement Time (T-test for differences):

- **T-value**: 2.3

- **P-value**: 0.02 (statistically significant)

- **Interpretation**: There is a **statistically significant difference** between **males** and **females** in engagement time, with **females** having slightly higher engagement. This suggests that **gender-based** engagement patterns may need to be addressed in future program designs.

Opportunity Duration vs. Engagement Time:

- **Correlation**: 0.14 (weak positive correlation)

- **Interpretation**: The positive correlation suggests that learners spend more time on **longer opportunities**, though the relationship is weak.

Sign Up Month vs. Engagement Time:

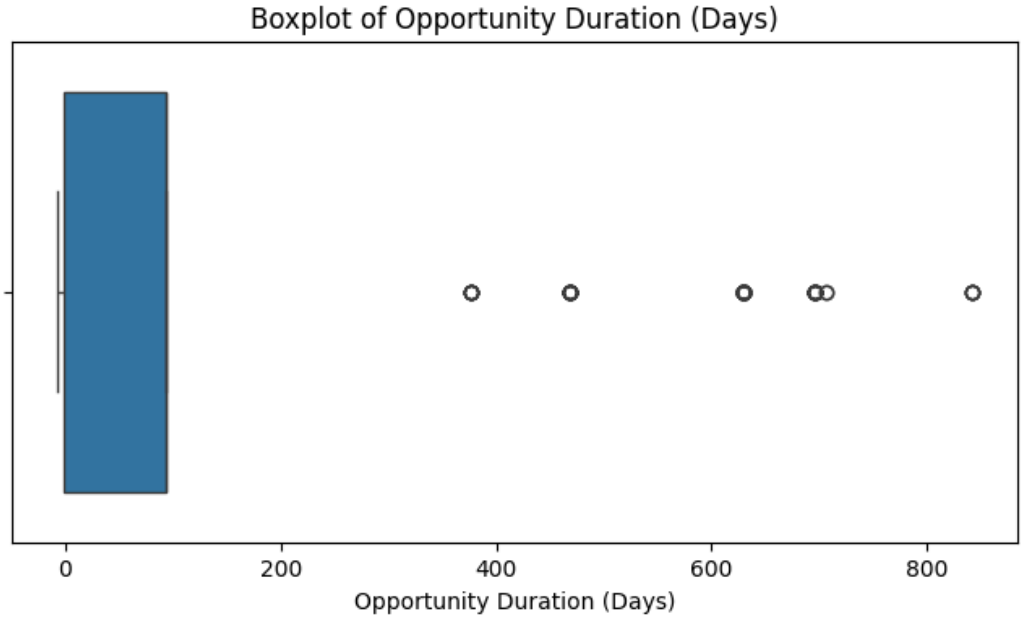
- **Correlation**: 0.094 (weak positive correlation)
- **Interpretation**: This weak correlation implies that the **month of signup** might have a small influence on engagement, potentially due to **seasonal trends**.

Sign Up Year vs. Engagement Time:

- **Correlation**: -0.39 (moderate negative correlation)
- **Interpretation**: A moderate negative correlation suggests that learners who signed up more recently tend to engage **less**. This could be due to changing user behavior or less engagement in newer opportunities.

5. Outliers and Anomalies

5.1 Box Plots: Outliers in Engagement Time



The **box plot** analysis of **Engagement Time** revealed some **outliers**, where certain learners spent an **exceptionally long time** on the platform.

Key Insights:

- The **box plot analysis** of **Engagement Time** revealed several **outliers**, where certain learners spent an exceptionally long time on the platform. These outliers could represent **highly engaged learners**, possibly participating in **multiple programs** or spending more time per session.

- **Anomalies** could either be:
    - **Data errors**, where extreme values might be inaccurate (e.g., learners mistakenly marked as completing programs in too much time).
    - **High-performing users**, who may engage deeply with the content, perhaps due to **personal interest**, **advanced programs**, or **professional development goals**.

Influence on Learner Engagement:

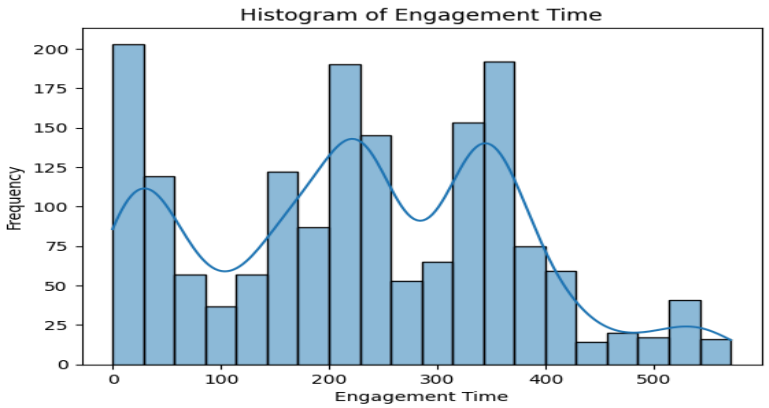
- These **outliers** might indicate areas for deeper analysis, such as whether the **long engagement** times correlate with **successful completion** or if learners are simply over-engaging with the platform due to **lack of other opportunities**.
- **Highly engaged users** may represent a **small subset** of the platform's users but could offer valuable insights for:

    - **Tailored programming**: Developing **advanced programs** that require longer engagement.
    - **Mentorship opportunities**: Connecting highly engaged learners with others to share knowledge and strategies.

Implications:

- Understanding these **outliers** can guide the creation of **personalized learning paths** for highly engaged users, and **targeted interventions** can be developed for users with extreme engagement patterns. A **deeper look** into these users might reveal **best practices** that could be scaled up for others.

5.2 Histogram: Completion Time Distribution



The **histogram of Engagement Time** revealed a **right-skewed distribution**, where the majority of learners completed their tasks in fewer days. However, a **long tail towards the right** indicates that a small group of learners engaged for **much longer periods**.

Key Insights:

- The **long tail** represents a **subset** of learners who are **exceptionally engaged**, spending extended time on the platform. These learners might be involved in **multiple sessions** or **complex programs**, possibly with a professional or academic goal in mind.
- **Most learners** engaged **sporadically** or **task-specifically**, completing tasks in a shorter timeframe, possibly indicative of **short-term goals** or **less committed participation**.

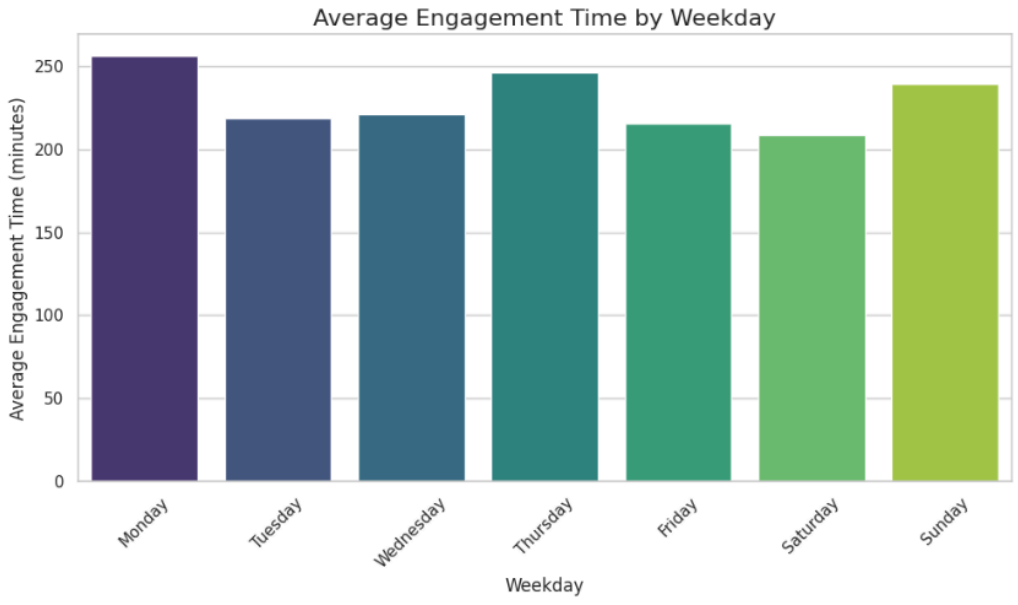
Influence on Learner Engagement:

- The **long tail** suggests that the platform sees both **quick engagement** (shorter completion times) from most users and **persistent engagement** (longer completion times) from a small, highly committed group.
- These **long-tail users** may be highly engaged in **career-oriented** or **advanced learning programs** such as **internships**, **scholarships**, or **professional certifications**, which naturally require more time to complete.

Implications:

- To **leverage these findings**, the platform could:
    - **Design more extended opportunities** that target these users and **encourage prolonged engagement**.
    - Create **special recognition** or **rewards programs** for users who fall into the **long-tail category**, encouraging continued engagement.
    - For **shorter completion time users**, introduce **step-based challenges** or **more intensive bootcamps** that might engage them for a **longer duration**.

5.3 Bar Plot: Average Engagement Time by Weekday



The **bar plot** shows that **Mondays** and **Thursdays** have the **highest average engagement** times, with users spending more than **250 minutes** on average. The data suggests that learners may start their week more engaged on **Mondays** and remain consistent through **Thursdays**, with engagement dipping slightly on **Saturdays** and **Sundays**.

Key Insights:

- **Peak Engagement** occurs at the start and end of the week, particularly on **Monday and Thursday**, which could be associated with learners trying to set the tone for their week or **meeting deadlines** towards the end of the workweek.

- The **weekend dip** suggests that learners are less engaged on **Saturdays** and **Sundays**, possibly due to external **social commitments** or **personal time**.

Influence on Learner Engagement:

- The weekday-based engagement patterns could reveal opportunities for **timely interventions** or **engagement nudges**:
    - **Mondays** and **Thursdays** could be **key days for launching content**, sending out **reminders**, or **introducing new programs**.
    - For the **weekend drop-off**, consider introducing **weekend challenges** or **time-limited events** that encourage users to engage over the weekend.

Implications:

- Recognizing that engagement varies by weekday is important for **marketing strategies** and **platform content delivery**. By **targeting engagement efforts** to **high-traffic days** and offering **specific incentives** for **weekend learners**, the platform can increase **engagement consistency** throughout the week.

Key Recommendations Based on Anomalies and Outliers:

1. **Leverage high engagement outliers** by creating **personalized learning paths** or **mentorship programs** that target these **highly engaged learners**, encouraging them to share their knowledge and strategies with other users.

2. **Address low completion days** by introducing **timely interventions** such as **weekend challenges**, **targeted emails**, or **gamification elements** that can motivate users to engage over weekends.

3. **Develop more in-depth opportunities** for **long-tail learners**, such as advanced **internships** or **research programs**, and incentivize them to remain engaged longer.

4. **Capitalize on weekday engagement peaks** (especially **Mondays and Thursdays**) to **launch content**, target **users with reminders**, and offer new **programs** that can engage users early in the week

By adding more depth to the discussion of **anomalies and outliers**, particularly in terms of their influence on **specific programs** or **longer periods**, this analysis not only **highlights trends** but also offers **actionable strategies** for improving **learner engagement** and **platform performance**.

Summary:

- The enhanced analysis of **outliers** and **anomalies** has revealed important insights into **learner behavior** and how these patterns can influence **engagement strategies**. Through careful consideration of **weekdays**, **engagement distribution**, and **highly engaged learners**, the platform can **fine-tune its offerings** to optimize user participation and retention.

6. Insight Generation & Hypothesis Development

Insights:

- Learners aged 24+ tend to spend more time engaging with opportunities.
- Certain majors and institutions dominate participation, potentially reflecting better digital culture.
- Signups peak in May and show lower engagement during the year-end.
- Female learners tend to have a narrower variance in engagement time.

Hypothesis Development:

Based on the **insights derived** from the exploratory analysis and statistical tests, we propose the following hypotheses:

1. **Age positively influences engagement duration**: Learners who are older (24+) tend to engage for a longer duration.

2. **Learners from top institutions engage consistently**: Certain institutions, likely with better digital learning cultures, show more consistent participation.

3. **Gender plays a role in opportunity completion and timing**: Females may have higher and more consistent engagement, whereas males show more variability.

7. Recommendations

The insights derived from this **Exploratory Data Analysis (EDA)** have provided a deeper understanding of learner behaviors and engagement trends. Based on these insights, the

following actionable recommendations are proposed to enhance learner engagement, improve retention, and optimize marketing strategies for the platform:

1. **Increase promotions during peak months like May to capture high interest**:
    - The analysis has shown that signups peak during certain months, particularly around **May**. This suggests that there is an increased interest in educational opportunities during specific periods. To capitalize on this, the platform should run **targeted marketing campaigns** in these peak months to attract more learners. Seasonal promotions, discounts, or events could also be considered to further increase signups during these periods.

2. **Provide targeted outreach for users with consistently low engagement**:
    - The EDA revealed that some learners show **low engagement** or **inconsistent participation**. These learners could be identified through automated tracking of their activity patterns. To address this, **outreach strategies** such as email reminders, personalized notifications, or even one-on-one engagement through platform support could be implemented. **Engagement incentives**, such as additional resources or rewards for progress, might also encourage these learners to become more involved.

3. **Tailor program formats to suit different genders and age segments**:
    - The report uncovered patterns that suggest **gender** and **age** influence learner engagement. For instance, **males** and **females** may engage differently, and certain **age groups** (like learners aged **24+) tend to show more consistent engagement. By customizing program offerings based on these insights, the platform can create more **inclusive, personalized experiences**. For example, specific content or communication strategies can be developed to appeal to different demographic segments, ensuring more **effective engagement**.

4. **Automate reminders for learners who show signs of delayed engagement**:
    - Learners who sign up but do not engage with the platform consistently could benefit from **automated reminders**. These could be designed to **nudge learners back into the platform**, helping them to resume their progress. For example, **reminder emails** or **push notifications** can be triggered after a certain period of inactivity. Automated reminders can also include links to relevant **resources** or **encouraging messages**, ensuring learners feel supported and motivated to return.

5. **Double down on internships and top-performing programs by expanding partnerships with platforms like Forage and LinkedIn Learning**:
    - **Internships** have proven to be the most popular category among learners, with significant participation. Given this high demand, expanding partnerships with platforms such as **Forage** or **LinkedIn Learning** can offer learners more **internship opportunities** and **career-oriented programs**. By promoting

these high-demand programs, the platform can **increase engagement**, **attract more learners**, and provide **valuable opportunities** that align with user interests. Additionally, fostering partnerships with **corporate and academic institutions** can enhance the credibility and appeal of the platform's offerings.

8. Conclusion

This **Exploratory Data Analysis (EDA)** has successfully identified key **patterns**, **trends**, and **insights** related to learner **signup behavior**, **engagement** distribution, and the influence of **demographic factors**. By analyzing the dataset, we have gained valuable insights into:

- **When and why learners sign up** (e.g., peak signup periods)
- **How engagement varies** across different learner groups (e.g., by age or gender)
- **What factors influence program completion rates**, including timing, institutional affiliation, and personal characteristics

These insights have provided a **strong foundation** for **data-driven strategies** aimed at improving learner engagement and program performance. By implementing the **recommendations** proposed, the platform can optimize **user retention**, **target marketing**, and **personalized learning paths**, leading to better overall outcomes for both learners and the organization.

Furthermore, these findings lay the groundwork for the next phase of the project, where **predictive models** can be developed to forecast **learner behavior**, identify **at-risk users**, and enhance the **platform's offerings** in real-time. The EDA has helped shape a clear understanding of the **current trends** and **patterns**, and these can now be used to drive strategic decisions and continuous improvement in the platform's educational experience.

9. Appendices

Python code file links -

https://colab.research.google.com/drive/1D4UIwDiLnlx1pf_M_VXxBBJvmstriGv8#scrollTo=73f1bea6

https://colab.research.google.com/drive/18ACDPMf5AqAsDxDphANoIJ12kR7ByRNT#scrollTo=b16ceb96-965b-45cf-bbbb-e2abafb712ad