**Data Analyst Intern Assignment: Analyse User Engagement on an Online Learning Platform**

**Background Scenario:**

Zylentrix runs an online training platform where students enrol in different learning tracks (e.g., Digital Marketing, Python, UI/UX, etc.). The platform records student behaviour, such as logins, course completions, time spent on lessons, and feedback.

You’ve been given 3 CSV files:

1. students.csv – Contains student info (ID, Name, Age, Gender, Location, Enrolment Date)
2. course\_activity.csv – Tracks user activity (Student ID, Course ID, Date, Time Spent [mins], Completion %)
3. feedback.csv – Records feedback (Student ID, Course ID, Rating [1-5], Feedback Text)

**Tasks & Deliverables**

**1. Data Cleaning & Preparation**

* Identify and handle missing, duplicate, or inconsistent data
* Convert data types where necessary (e.g., dates, numeric fields)
* Prepare a clean dataset ready for analysis

**2. Exploratory Data Analysis (EDA)**

Answer the following using Python (preferably with pandas/matplotlib/seaborn/plotly) or any other analytics tool of your choice:

* What is the overall average completion rate across courses?
* Which course has the highest and lowest average engagement time?
* How does engagement differ by age group or location?
* What is the average feedback rating per course?
* Is there a correlation between completion rate and feedback rating?
* Identify top 3 student segments based on engagement and satisfaction.

**3. Visualisations**

Create at least 3–5 visualisations (bar charts, heatmaps, scatter plots, etc.) to:

* Show engagement patterns by course and demographic
* Compare feedback ratings
* Highlight trends over time

**4. Insights & Recommendations**

* Summarise your top 5 insights from the analysis
* Suggest 3 data-driven recommendations to improve student engagement or course performance
* Format your findings in a short slide deck (PDF or Google Slides, max 5 slides) or include it in a structured report

**Submission Format**

* Submit either:
  + A Jupyter Notebook with markdown explanations, OR
  + A Google Colab link, OR
  + A well-documented Python script
* Include your visualisations and recommendations in a separate PDF/Google Slides
* Upload your files or links via the [**Google Form Submission**](https://docs.google.com/forms/d/e/1FAIpQLSeB_tJQH4QmyNC3b9233QzX5fww1FwC-0aXZ8Bql3elKgIwrA/viewform?usp=header)by the date mentioned in the mail**.**

**Notes**

* Use any preferred tool: Python, R, Power BI, Tableau, or Excel (Python preferred).
* Focus on producing meaningful, business-oriented analysis rather than just coding.