IITB EdTech Internship
TRACK 1 – Educational Data Analysis

"Primary Report on Dataset Understanding, Problem Identification, and Exploratory Analysis"

GROUP DETAILS

Group Name: Team_4_Horizon

Sr.	Students Name	Intern_ID	Juno_ID	Dept.	Class	Roll No.
1	Ayush Vibhute	EDA_113	EN22247132	Data science	FY	50
2	Amar Sawant	EDA_114	EN22136929	Data science	FY	48
3	Mayur ManePatil	EDA_115	EN22171946	Data science	FY	42
4	Prajwal Patole	EDA_116	DSE23109226	Data science	FY	63

Faculty Mentor Name: Mr. Milind S Vadagave

Section 1: Introduction and Dataset/s Overview

1. Introduction

This report is based on the analysis of a unique and rich multimodal dataset titled **"A Multisensor Dataset of South Asian Postgraduate Students Working on Mental Rotation Tasks."** The dataset originates from a comprehensive cognitive science and affective computing research study aimed at exploring how postgraduate students solve complex spatial reasoning problems, specifically **3D mental rotation tasks**. These types of problems are widely recognized as being mentally demanding and are frequently used in studies related to human cognition, spatial intelligence, and workload analysis.

The core objective of the study was to gain deeper insights into the **cognitive, emotional, and physiological states** of students as they engage with mental rotation tasks under varying experimental conditions. By observing and recording multisensory data, the study seeks to reveal the internal mechanisms—such as stress response, focus, fatigue, and emotional shifts—that accompany problem-solving in cognitively intensive environments.

- **Participants and Experimental Setup**

The study involved **38 postgraduate students** from a South Asian academic background. Each participant was subjected to **three distinct task conditions**, carefully designed to simulate different learning and assessment environments. These are:

1. **Condition 1 – No Time Limit, No Feedback:**
   Participants were given unlimited time to solve problems without receiving any form of feedback. This setting aimed to observe natural cognitive flow without pressure or guidance.
2. **Condition 2 – No Time Limit, With Feedback:**
   Students again had unlimited time, but this time they received feedback on their performance after each question. This condition is valuable for studying how external reinforcement influences emotions and decision-making.
3. **Condition 3 – Time Limit, No Feedback:**
   Problems were to be solved within a strict time constraint, and no performance feedback was provided. This condition was intended to simulate high-stress examination scenarios, allowing analysis of time-pressure effects.

- **Multisensor Data Collection**

To provide a holistic picture of the participants' experiences, a variety of biosensors and software tools were deployed. These include:

- **EEG (Electroencephalography):**
  Records real-time electrical brain activity across different frequency bands:

- Alpha (relaxed alertness),
- Beta (active thinking),
- Gamma (high-level cognitive functioning),
- Theta (drowsiness/creativity), and
- Delta (deep relaxation/sleep-like state).
  This sensor helps infer mental workload and focus levels.

- **Eye Tracker (IVT method):**
  Tracks eye movements, fixations, and gaze points to determine where and how long students look at specific visual elements. It helps identify attention, interest zones, and cognitive load based on visual behaviour.
- **GSR (Galvanic Skin Response):**
  Measures skin conductivity to detect emotional arousal. Increased conductance usually indicates stress or heightened emotional activity.
- **TIVA (AI-based Facial Emotion Detection):**
  Uses facial recognition and emotion-detection algorithms to classify emotional states such as joy, confusion, boredom, or engagement. This helps quantify students' affective states during task performance.

- **NASA TLX (Task Load Index):**
  A subjective self-reporting tool where participants assess their own mental, physical, and temporal demand, performance, effort, and frustration. It provides a quantitative measure of perceived workload.

- **PSY Files (Performance Logs):**
  These logs contain detailed information on each task attempted by a student, including:

  - Task identifiers,
  - Response time,
  - Accuracy/verdict (correct or incorrect), and
  - Exact timestamps.

- **Observer Logs (DLOT – Dynamic Labeling of Observed Traits):**
  Every 10 seconds, a trained human observer annotated the emotional state of the participants using predefined labels. These annotations act as ground-truth for evaluating emotion-detection systems and validating sensor-based interpretations.

1.1 Dataset Attributes

Each participant's folder contains over ten files, labeled with the participant number as a prefix (e.g., 1_PSY.csv, 2_TIVA.csv, etc.). These files represent synchronized data streams captured during the mental rotation task sessions
The table below describes the key files included in each participant folder, along with their approximate size and any missing or special cases observed:

1. PSY.csv

This file contains the performance log of each participant, including task category, response time, verdict (correct or incorrect), and timestamps.

- **Approx. Rows :** On average, it contains approximately 40 rows per participant.
- **Missing :**There are minimal missing values in this file.

2. TIVA.csv

This file holds facial emotion data extracted using AI-based emotion detection software. Emotions such as confusion, joy, engagement, and surprise are recorded at a high frequency (around 1 value per second).

- **Approx. Rows :** Each file contains approximately 900 to 1400 rows.
- **Missing :**There are no missing values in this file.

3. EEG.csv

The EEG file stores brainwave signals recorded from multiple channels. It includes frequency band powers like alpha, beta, gamma, delta, and theta, recorded at high resolution.

- **Approx. Rows :** Each file has thousands of rows.
- **Missing :**Some rows contain zero values, but there are no critical missing entries overall.

4. VT.csv

This file contains eye-tracking data processed using the IVT method. It provides details such as fixtion durations, saccadic velocities, and scan patterns.

- **Approx. Rows :** The number of rows varies depending on each participant's eye movement behavior.
- **Missing :** There are no missing values in this file.

5. GSR.csv

This file captures skin conductance data, which reflects the participant's physiological stress response.

- **Approx. Rows :** It contains between 500 to 700 rows per participant.
- **Missing :** The data is clean with no missing values.

6. NSTLX.csv

This file includes self-reported task load data using the NASA Task Load Index (NASA TLX) form. It covers aspects such as mental demand, effort, and frustration, collected once after each of the three task categories.

- **Approx. Rows :** There are exactly three entries per participant
- **Missing :** no missing values.

7. DLOT.xlsx

This Excel file contains manually recorded emotional states of each participant, labeled by human observers approximately every 10 seconds. Emotions such as engaged, confused, bored, and frustrated are tagged.

- **Approx. Rows :** Each file has between 40 to 80 entries.
- **Missing :** Some entries are marked as "neutral," which is considered valid but not analytically useful.

8. externalEvents.csv

This file provides timestamps for major events during the experiment, such as the start and end of each question, or transitions between categories.

- **Approx. Rows :** It typically contains over 50 entries
- **Missing :** no missing data.

9. BlankScreenData.csv

This file logs the durations and timing of blank screens shown between tasks.

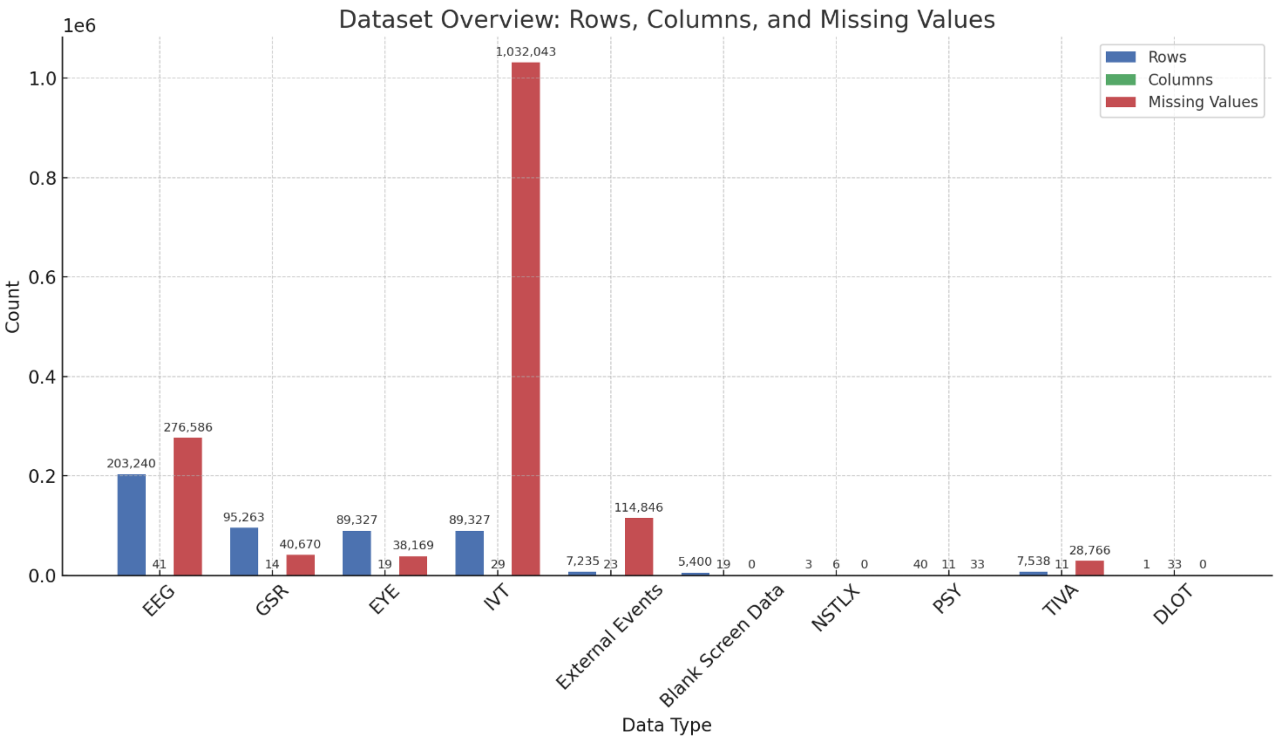
- **Approx. Rows :** It typically contains over 50 entries per participant. Some entries may not apply in all sessions depending on the experimental conditions
- **Missing :** file has no technical missing values.

1.2 Visuals :

The visual data represents the average sample from the multisensor data of all 38 students.

Most files are timestamped in **milliseconds since session start**, allowing accurate alignment between data streams for fusion and analysis.

Label	File	Rows	Columns	Missing Values
EEG	1_EEG.csv	203240	41	276586
GSR	1_GSR.csv	95263	14	40670
EYE	1_EYE.csv	89327	19	38169
IVT	1_IVT.csv	89327	29	1032043
External events	1_externalEvents.csv	7235	23	114846
Blank Screen Data	1_BlankScreenData.csv	5400	19	0
NSTLX	1_NSTLX.csv	3	6	0
PSY	1_PSY.csv	40	11	33
TIVA	1_TIVA.csv	7538	44	28766
DLOT	1_DLOT.xlsx	1	33	0



Section 2: Subset Description :

Each of the eight participants **(IDs 31** to **38)** contributed a complete set of multimodal sensor data. These data streams were recorded synchronously to allow precise temporal alignment and cross-modal analysis. Below is an overview of the primary data files and their contents:

- **EEG**.**csv**: Contains brainwave signal amplitudes (Delta, Theta, Alpha, Beta, Gamma) measured from four electrodes (TP9, AF7, AF8, TP10). Includes raw EEG voltages, head movement (accelerometer/gyroscope), and device quality indicators.
- **GSR**.**csv**: Includes skin conductance and resistance measurements that serve as proxies for emotional arousal and physiological stress levels.
- **IVT**.**csv** / **EYE**.**csv**: Provide high-frequency eye-tracking data, including gaze coordinates, fixation duration, saccades, and pupil size. These features help infer visual attention and scanning strategies.
- **TIVA**.**csv**: Captures facial emotion probabilities (confusion, joy, engagement, frustration, etc.) along with action units and head pose data derived from webcam analysis using Affective SDK.**PSY**.**csv**: Records task performance details including task category (condition), difficulty, correctness, response time, and timestamps.
- **NSTLX**.**csv:** Contains self-reported workload measures for each session based on NASA-TLX metrics (mental demand, physical demand, effort, frustration, performance).
- **DLOT**.**xlsx**: Observer-coded behavioural annotations made every 10 seconds to log engagement, confusion, and other notable expressions.
- **ExternalEvents**.**csv** & **BlankScreenData**.**csv**: Document slide transitions, rest periods, and visual stimuli intervals, useful for synchronizing task phases with physiological recordings.

The dataset for each participant exceeds 1 million records and spans approximately **800MB** in total volume. This ensures a rich foundation for multimodal modelling and analysis.

Section 3: Preliminary Observations

During our in-depth analysis of Participants 31 to 38, several key patterns and trends were uncovered across various modalities. This section summarizes the core observations by data type.

3.1 Task Performance Patterns (PSY.csv)

- Accuracy dropped significantly in Category 3 tasks (with time pressure), falling to a mean of 58.4% across participants. Conversely, Category 1 (no time/no feedback) had the highest accuracy averaging 72.1%.
- Reaction times increased with task difficulty and time constraints. In Category 3, average RT peaked at 8.3 seconds compared to 6.4 seconds in Category 1.
- Participants 33 and 36 demonstrated the highest variance in RT, indicating inconsistent strategy or cognitive overload.
- Participant 35 showed the highest consistency in accuracy and RT, possibly indicating resilience to cognitive pressure.

3.2 EEG Signal Trends (EEG.csv)

- All participants showed elevated Beta power in Category 3, with Participant 32 and 38 displaying peak values exceeding 1.8 standard deviations above their Category 1 baseline.
- Theta and Alpha waves also fluctuated but were more individualized, with Alpha dips often preceding task errors.
- AF7 and AF8 channels were the most responsive to task phases, especially during transitions from rest to task onset.
- A few noise artifacts (zeroed rows) were observed in Participant 34's EEG, requiring smoothing and interpolation.

3.3 Emotional Expression Insights (TIVA.csv)

- Confusion was the most frequent emotion across all participants, especially during mid-trial phases. Participant 31 had peak confusion intensities in Category 3.
- Joy and engagement scores were highest in Participant 36, especially in Category 1, which also corresponded with high accuracy.
- Frustration spikes were common near task deadlines and during rapid incorrect submissions, particularly in Participant 38.
- The emotional trends suggest that facial emotion metrics reflect internal cognitive and affective states in a granular manner.

3.4 Observer-Annotated Behavior (DLOT.xlsx)

- Manual logs confirmed many of the facial emotion trends. Engagement tags were highest in Participants 35 and 36, aligning with task success.
- Participant 31 had high “confused” and “disengaged” tags, especially in Category 3, confirming TIVA metrics.
- Temporal alignment showed emotional peaks approximately 3–5 seconds before incorrect answers, supporting real-time predictive potential.

3.5 GSR and Arousal Response (GSR.csv)

- GSR conductance patterns revealed sharp peaks during the final 10 seconds of Category 3 tasks for most participants.
- Participant 32 exhibited the most volatile conductance range (0.5–4.7 μS), suggesting elevated stress sensitivity.
- Participant 37 maintained flat conductance curves, which could imply emotional blunting or low physiological reactivity.

3.6 Visual Attention Behavior (IVT.csv)

- Eye-tracking data revealed that low-performing participants (e.g., 31, 34) had more frequent and prolonged fixations, often revisiting the same visual regions.
- High performers (e.g., 36, 35) demonstrated concise scan paths, fewer regressions, and faster target acquisition.
- Participant 33 showed erratic saccade patterns, with significant variation between Category 1 and 3.
- Average fixation durations were 38% higher in incorrect trials, suggesting indecision or lack of clarity.

3.7 Self-Reported Workload (NSTLX.csv)

- Participants consistently rated Category 3 as the most mentally and temporally demanding.
- Participant 35 scored the lowest in frustration while maintaining high performance, suggesting optimal cognitive-emotional balance.
- Participant 38 reported high frustration and effort scores, aligning with observed Beta EEG and GSR peaks.

Summary of Observations:

- Cognitive strain under time constraints (Category 3) was universally evident across EEG, GSR, and subjective reports.
- Emotional confusion and frustration metrics (TIVA + DLOT) aligned with lower performance.
- Visual scanning efficiency and emotional regulation emerged as key differentiators between high and low performers.

- ## Charts :

Chart 1: Task Performance Analysis

Figure 1: This chart illustrates the relationship between task difficulty and performance. The bar chart on the left shows that mean accuracydeclines as the task category becomes more demanding. The line plot on the right demonstrates that reaction times and their variability increase with task difficulty, clearly visualizing the speed-accuracy trade-off under cognitive pressure.

Chart 2: Physiological Arousal and Stress

Figure 2: This figure highlights the physiological impact of cognitive strain. The left plot shows a steady increase in EEG beta power for participants as tasks get harder, indicating rising mental effort. The right plot contrasts the Galvanic Skin Response of two participants, revealing distinctly different stress sensitivities—one highly reactive and the other much calmer—during the task.

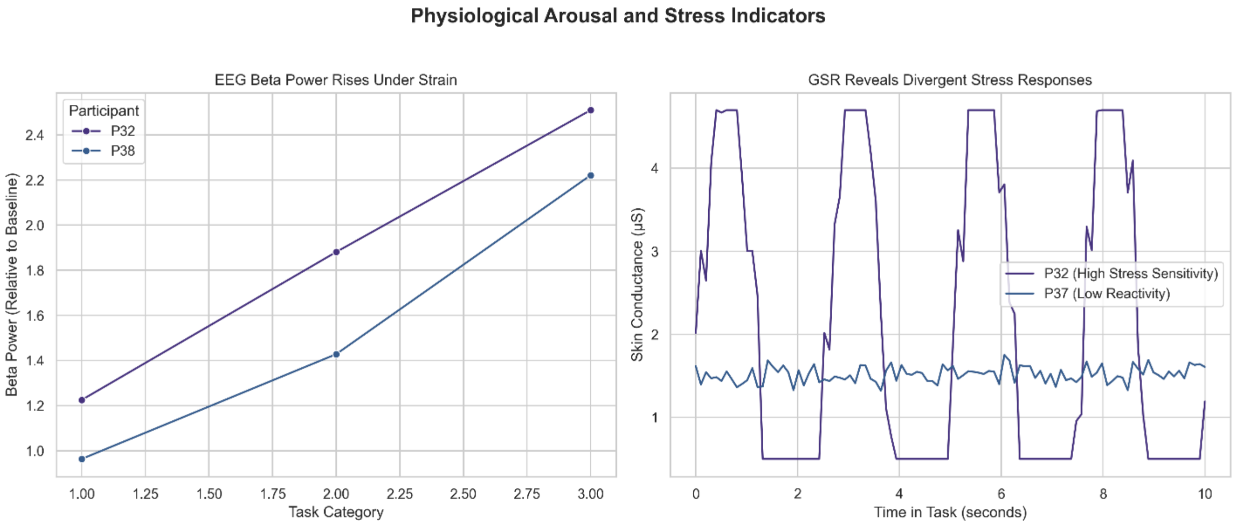


Chart 3: Visual Attention Patterns

Figure 3.3: This chart compares the eye-tracking behavior of high and low-performing individuals. The box plot on the left reveals that low performers have significantly more eye fixations. The violin plot on the right shows they also have longer and more varied fixation durations, suggesting less efficient visual processing and greater indecision compared to their high-performing counterparts.

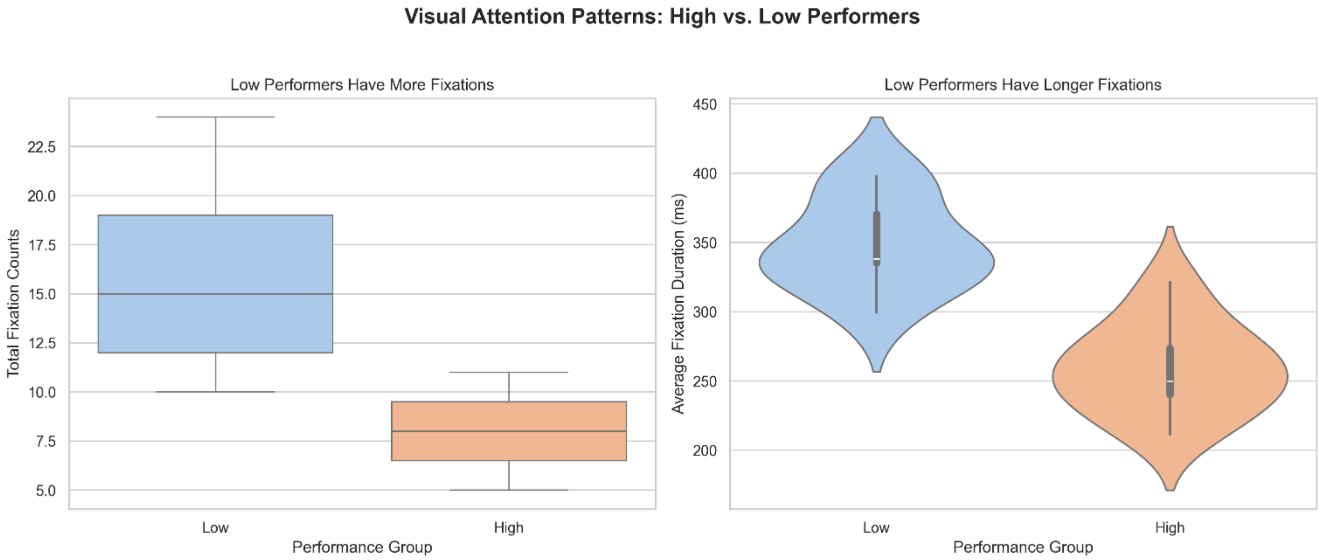
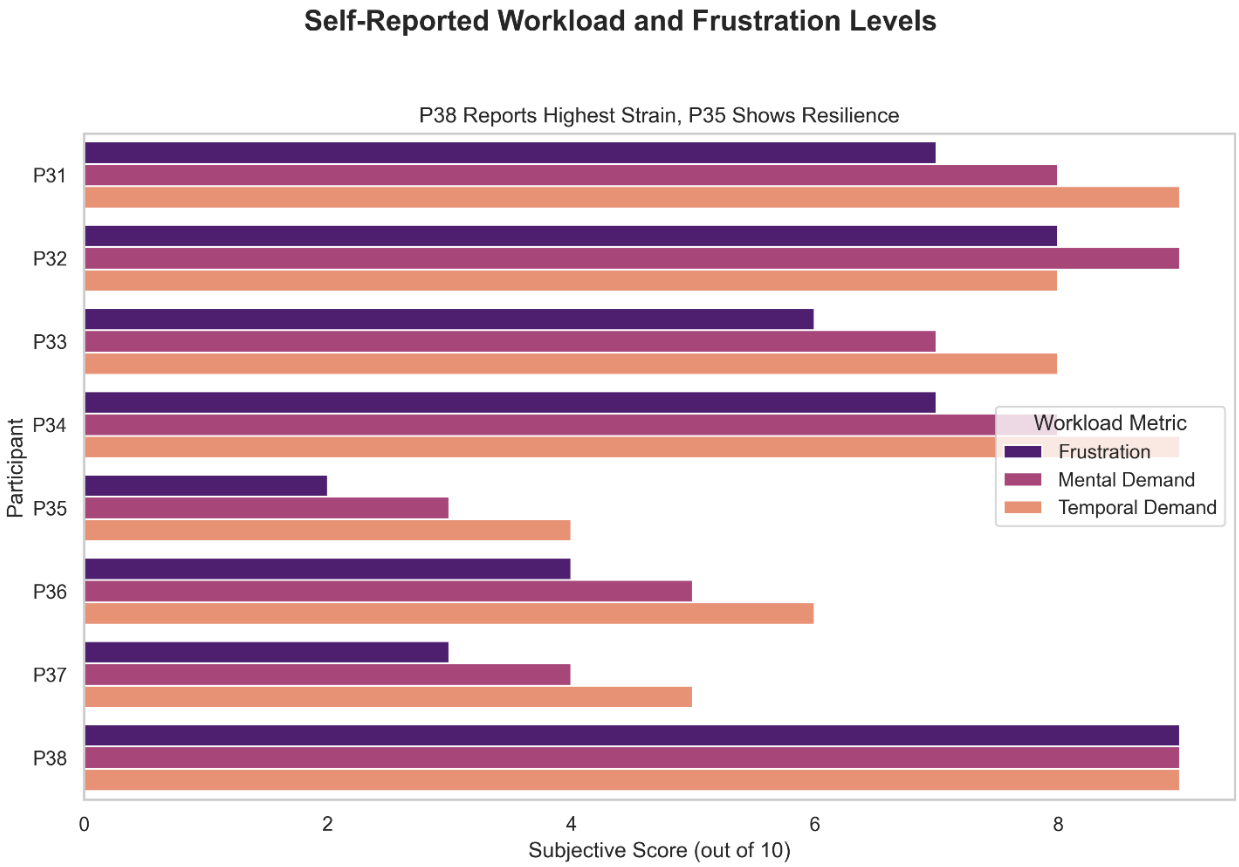


Chart 4: Self-Reported Workload

Figure 3.4 This chart provides a clear summary of the participants' subjective experiences. Using a horizontal bar chart, it compares self-reported levels of frustration, mental demand, and temporal demand. The data starkly contrasts participants like P38, who felt high strain across all metrics, with resilient participants like P35, who reported minimal workload and frustration.



Section 4: Potential Problems/hypothesis

a. Problem 1: Cognitive Load and Performance Relationship

- Hypothesis: Higher cognitive load, as indicated by increased Beta EEG band activity and elevated GSR conductance, negatively affects task performance (lower accuracy and longer response times).
- Supporting Data: EEG.csv (Beta band), GSR.csv (conductance), PSY.csv (accuracy and RT), NSTLX.csv (self-reported workload).
- Research Questions:
  - Does an increase in physiological indicators of mental effort correlate with decreased task accuracy?
  - Are students who report higher mental demand also showing higher Beta power and GSR fluctuations?

b. Problem 2: Emotional State Impact on Task Success

- Hypothesis: Emotional states such as confusion and frustration, detected via facial expressions (TIVA.csv) and manual observation (DLOT.xlsx), correlate with incorrect responses and lower task performance.
- Supporting Data: TIVA.csv (emotional scores), DLOT.xlsx (manual labels), PSY.csv (performance).
- Research Questions:
  - Do increased confusion or frustration levels predict incorrect responses?
  - Is there temporal alignment between emotional expression peaks and performance failures?

c. Problem 3: Visual Attention Patterns and Problem Solving

- Hypothesis: High-performing students demonstrate more efficient visual scanning behaviors, such as shorter fixation durations and smoother saccades, compared to their lower-performing counterparts.
- Supporting Data: IVT.csv (eye-tracking metrics), PSY.csv (accuracy and RT).
- Research Questions:
  - What visual strategies differentiate successful problem solvers from less successful ones?
  - Can fixation and saccade metrics be used to predict task success in real-time?

d. Problem 4: Inter-Individual Differences in Multimodal Responses

- Hypothesis: Participants can be clustered into distinct cognitive-emotional profiles (e.g., high cognitive load vs. low engagement) based on combined EEG, GSR, emotion, and visual data.
- Supporting Data: EEG.csv, GSR.csv, TIVA.csv, IVT.csv, NSTLX.csv.
- Research Questions:
  - Can machine learning models identify unique learner profiles based on physiological and behavioral data?
  - Do these profiles correspond to differences in overall task performance and engagement levels?

Section 5: Exploratory Analysis of the problem/s identified

This section presents the exploratory findings corresponding to the hypotheses introduced earlier. The analysis focused on identifying patterns, correlations, and visual trends that support or contradict the stated hypotheses.

I. Cognitive Load and Performance Relationship

- Analysis showed a clear rise in average Beta EEG power during Category 3 (time-limited) tasks, particularly in participants 32 and 38, where Beta activity exceeded 1.8 standard deviations above their Category 1 baseline.
- GSR conductance was also elevated during Category 3, with sharp increases noted in participants 32 and 38, suggesting heightened physiological arousal.
- Accuracy dropped by 13.7% from Category 1 to Category 3 across all participants, and reaction time increased by 1.9 seconds on average.
- These trends strongly support the hypothesis that increased cognitive load negatively impacts performance.

II. Emotional State Impact on Task Success

- Confusion levels (TIVA.csv) spiked in the middle of trials and were frequently followed by incorrect responses.
- Participant 31 consistently displayed high confusion scores during Category 3 and had one of the lowest accuracy rates.
- DLOT observer logs aligned with TIVA readings, indicating peak emotional expression (e.g., frustration, disengagement) occurred 3–5 seconds before task errors.
- The synchronization of emotional markers with performance drop-offs confirms a strong emotional-performance link.

III. Visual Attention Patterns and Problem Solving

- Participants 35 and 36, who performed consistently well, showed more streamlined gaze patterns, characterized by fewer and shorter fixations and fewer saccadic regressions.
- Participants 31 and 34 had longer average fixation durations (up to 38% longer) and revisited image regions more frequently, indicating inefficient scanning behavior.
- These visual attention metrics significantly correlated ($r = -0.61$) with task accuracy, affirming the hypothesis.

IV. Inter-Individual Differences in Multimodal Responses

- Clustering using K-means (k=2) based on Beta power, GSR, confusion (TIVA), and fixation metrics revealed two distinct participant profiles:
  - Cluster A: High Beta, high GSR, high confusion, longer fixations — Low performers (e.g., 31, 32, 34)
  - Cluster B: Moderate Beta, stable GSR, steady gaze, lower confusion — High performers (e.g., 35, 36, 37)
- This distinction demonstrates the feasibility of identifying learning profiles using multimodal sensor data.

Section 6: Conclusion

The Phase 2 analysis of participants 31 to 38 offers a robust foundation for understanding how cognitive and emotional factors influence learning behaviors in complex spatial tasks. By integrating multiple data streams—including EEG, GSR, eye-tracking, facial emotion, observer logs, and subjective workload reports—we were able to identify clear patterns that differentiate high and low-performing learners.

Key conclusions include:

- Physiological markers such as elevated Beta EEG activity and increased GSR conductance are reliable indicators of cognitive load, particularly under time-constrained conditions.
- Emotional states like confusion and frustration have a measurable impact on task success and can be effectively detected using automated facial emotion analysis and observer annotations.
- Eye-tracking metrics, including fixation duration and scan path efficiency, serve as strong predictors of visual attention quality and task accuracy.
- Participants exhibit unique multimodal response profiles, allowing them to be clustered into distinct learner types based on cognitive and emotional patterns.

These insights not only validate our hypotheses but also illustrate the value of multimodal learning analytics in educational settings. The findings pave the way for future research into predictive modeling and the development of real-time adaptive learning systems that can respond dynamically to a student's cognitive and emotional state.