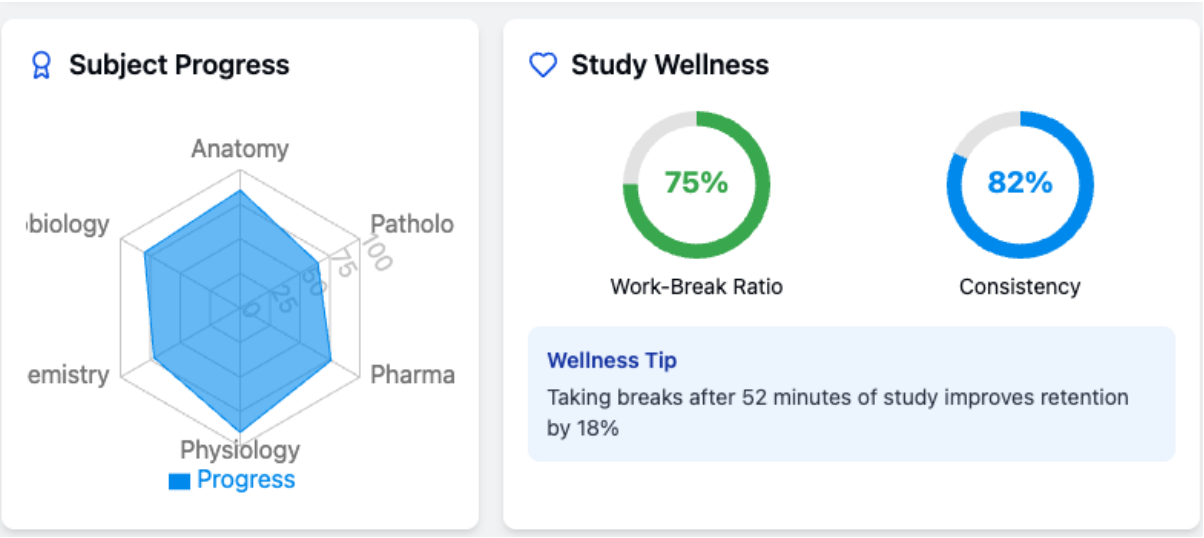


Subject Progress Calculation (Subject Based)

Subject Progress Score = (Completed Pages / Total Pages) × 60% +

(Flashcard Success Rate) × 20% +

(Time Spent / Recommended Time) × 20%

Example for Anatomy:

- Completed Pages: 180/200 = 0.9

- Flashcard Success Rate: 75%

- Time Spent: 45 hours / 50 recommended hours = 0.9

- Subject Progress Score = (0.9 × 0.6) + (0.75 × 0.2) + (0.9 × 0.2) = 0.54 + 0.15 + 0.18 = 0.87 (87%)

Study Wellness Calculation

Work-Break Ratio = (Total Study Time - Break Time) / Break Time

Optimal ratio is 3:1 to 5:1

Example:

- Total Study Time: 40 hours/week

- Break Time: 10 hours/week

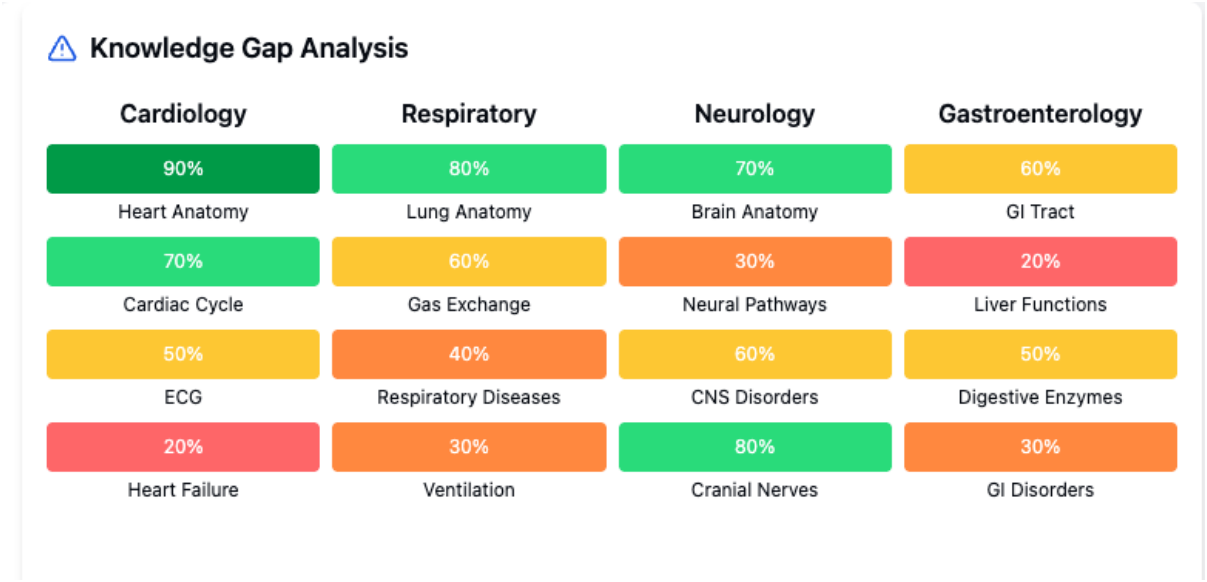
- Work-Break Ratio = 30/10 = 3 (optimal)

Consistency Score = (Days with study activity / Total days in period) × 100%

Example:

- Study activity on 24 days out of 30

- Consistency Score = (24/30) × 100% = 80%



maturity

Knowledge Gap Heatmap Calculation (Topic Based)

Topic Mastery Score = (Correct Flashcard Answers / Total Attempts) × 40% +

(Completion Rate) × 30% +

(Time Spent / Recommended Time) × 30%

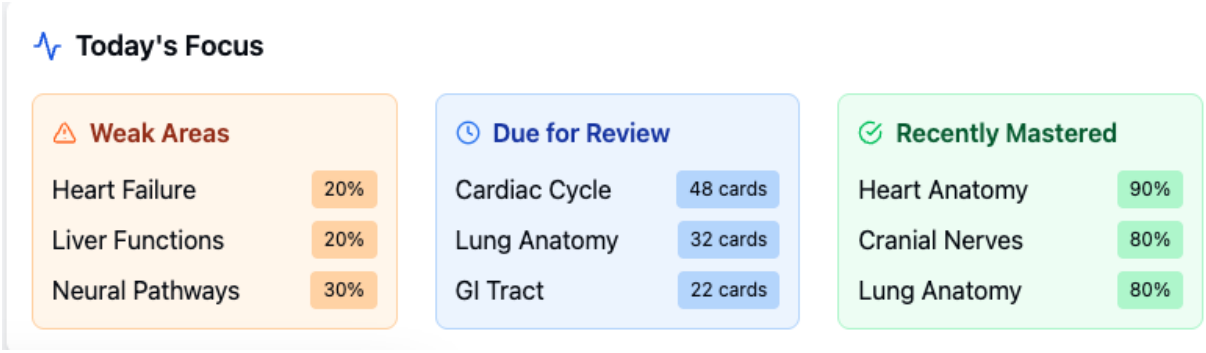
Example for "Heart Failure":

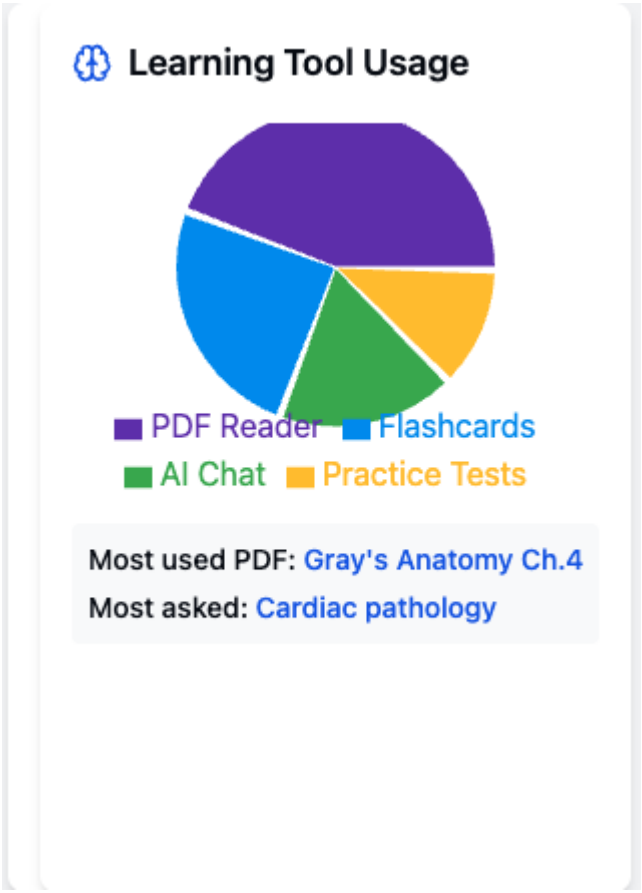
- Flashcard Performance: 15 correct out of 25 attempts = 60%

- Completion Rate: 8 pages read out of 12 = 67%

- Time Spent: 3 hours / 5 recommended hours = 60%

- Mastery Score = (0.6 × 0.4) + (0.67 × 0.3) + (0.6 × 0.3) = 0.24 + 0.201 + 0.18 = 0.621 (62%)





Today's Focus:(Topic based)

Use LLM to identify weak area suggestions the knowledge gap data with lowest score will be used

Learning Tool Usage: (Feature based)

Highest used tool in a pie chart calculation

Interactive Concept Network (Topic based)

Concept Connection Strength

Connection Strength = (Co-occurrence in Study Material) × 40% +

(Sequential Learning Pattern) × 30% +

(Related Quiz Performance) × 30%

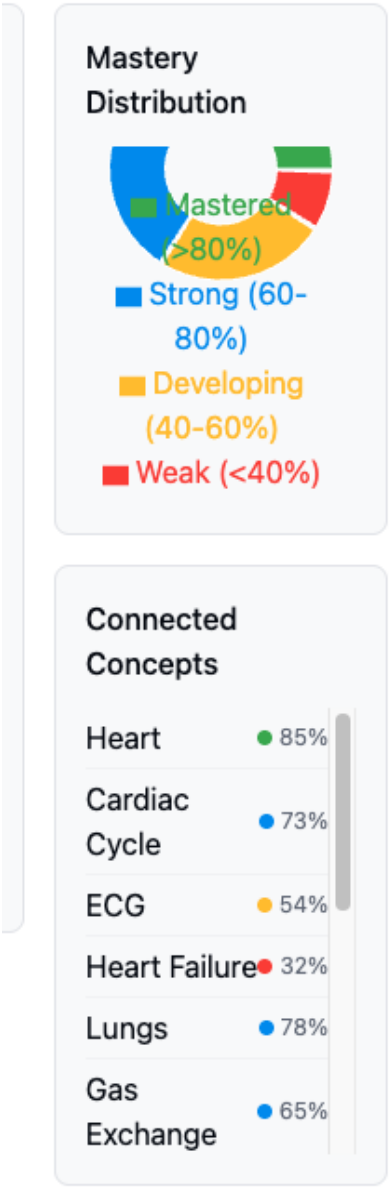
Example for "Heart → Cardiac Cycle":

- Co-occurrence: Found together on 8 out of 10 pages = 80%

- Sequential Learning: User studied these topics in sequence 7 out of 10 times = 70%

- Related Quiz Performance: 75% success when questions involve both concepts

- Connection Strength = (0.8 × 0.4) + (0.7 × 0.3) + (0.75 × 0.3) = 0.32 + 0.21 + 0.225 = 0.755 (76%)



Concept Mastery Distribution(Topic based)

For each mastery level (Mastered, Strong, Developing, Weak):

Count of concepts at each level / Total concepts × 100%

Example:

- 24 Mastered concepts (>80% mastery)

- 35 Strong concepts (60-80% mastery)

- 28 Developing concepts (40-60% mastery)

- 13 Weak concepts (<40% mastery)

- Total: 100 concepts

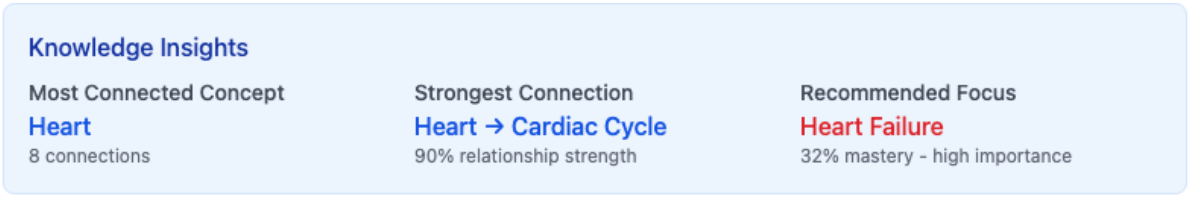
Distribution:

- Mastered: 24%

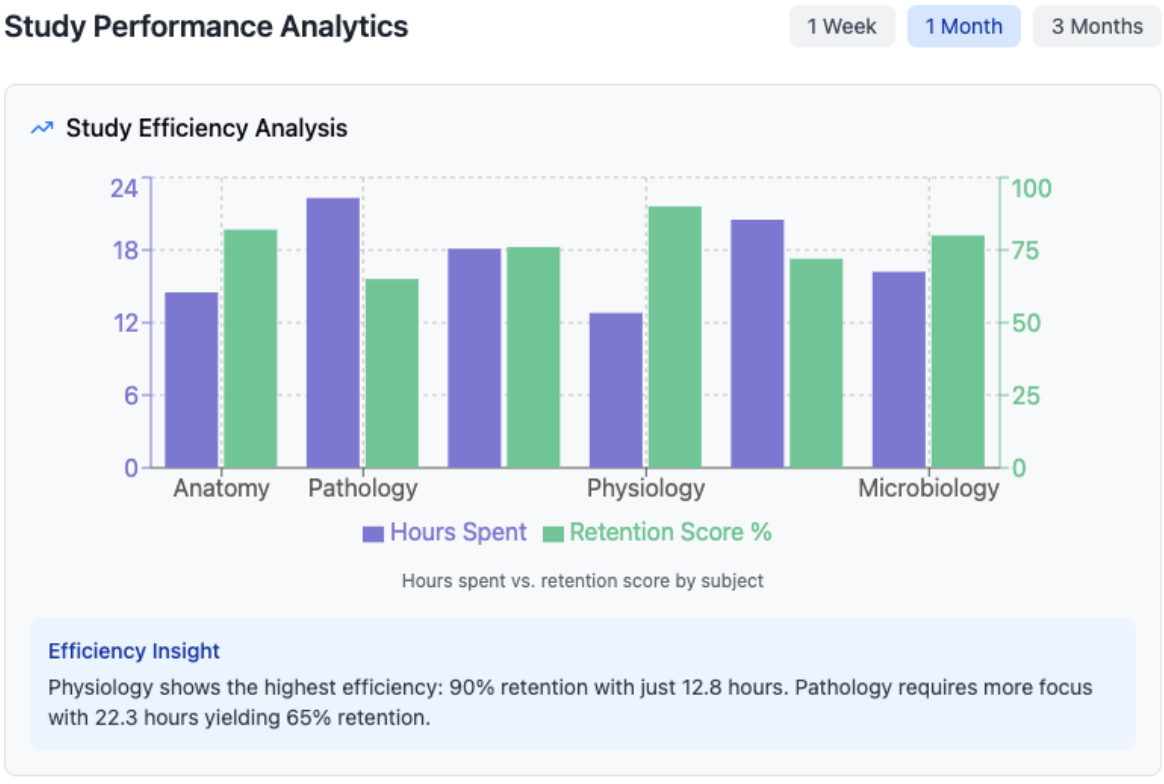
- Strong: 35%

- Developing: 28%

- Weak: 13%



Knowledge Insights(AI based)



Study Performance Analytics - Sample Calculations

Study Efficiency Analysis

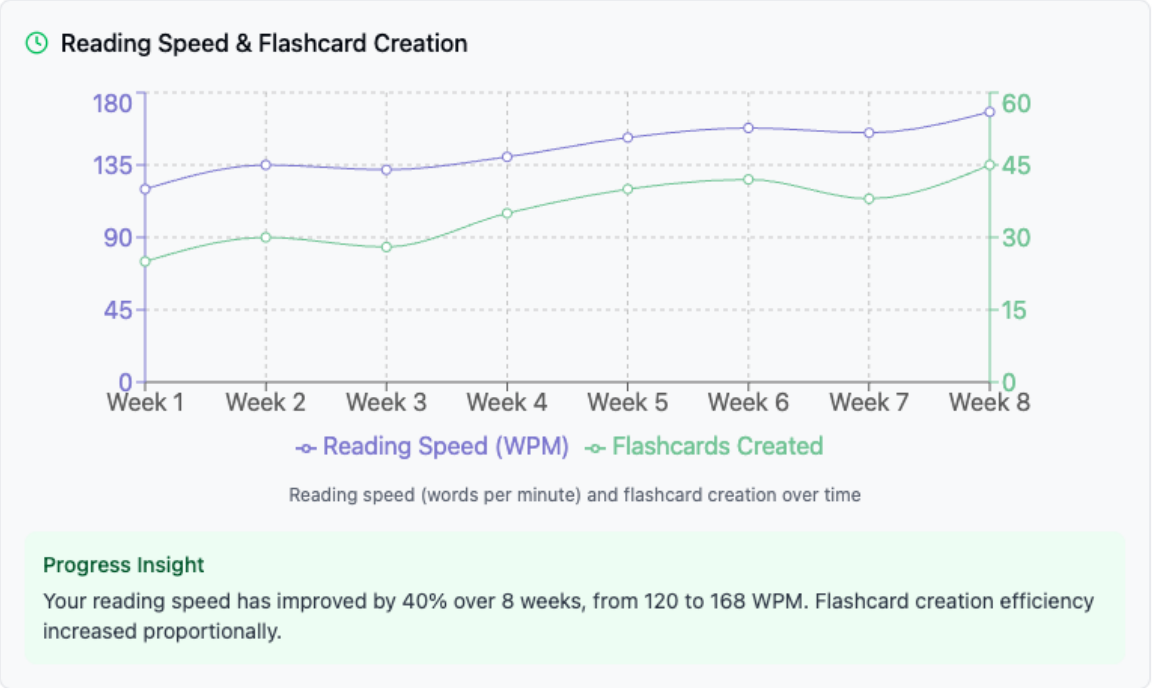
Efficiency Score = (Retention Score) / (Hours Spent) × Normalization Factor

Example for Physiology:

- Retention Score: 90%
- Hours Spent: 12.8
- Normalization Factor (to scale to 0-100): 1.2
- Efficiency Score = (90/12.8) × 1.2 = 8.4 (very efficient)

Compare to Pathology:

- Retention Score: 65% (Calculations can be found on page 12)
- Hours Spent: 22.3
- Efficiency Score = (65/22.3) × 1.2 = 3.5 (less efficient)



Reading Speed Calculation

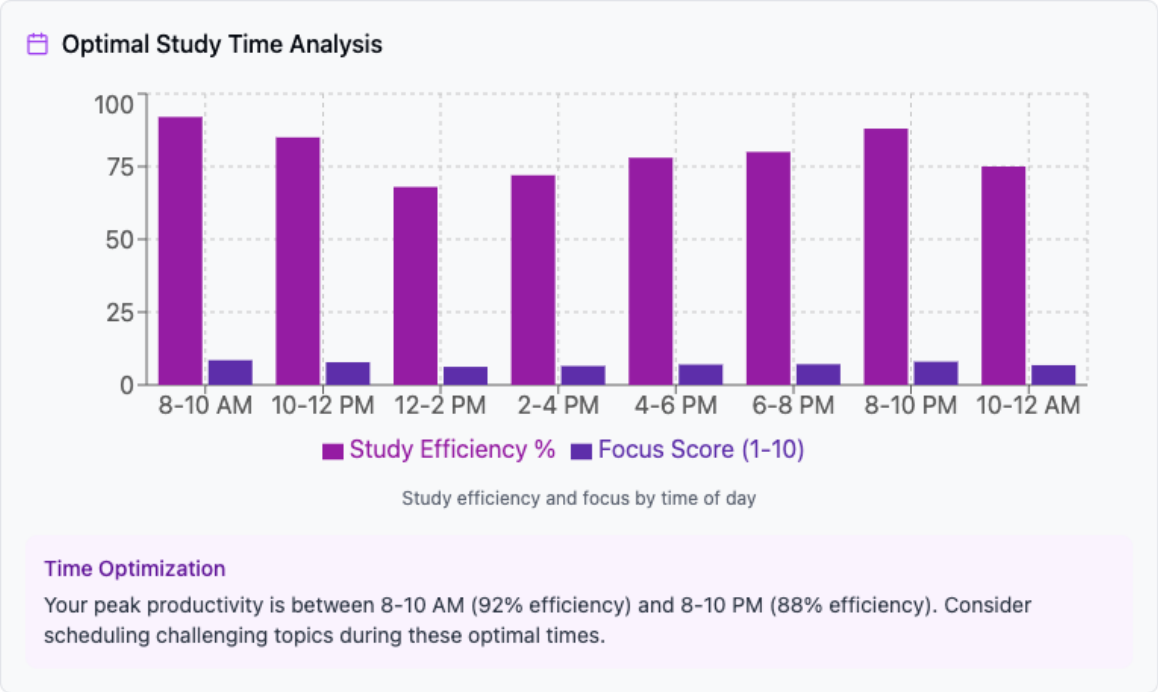
Reading Speed (WPM) = Words Read / Minutes Spent Reading

Example:

- Week 1: 6,000 words / 50 minutes = 120 WPM

- Week 8: 8,400 words / 50 minutes = 168 WPM

- Improvement: 40% increase over 8 weeks



Optimal Study Time Analysis

Study Efficiency by Time = (Pages Completed × Comprehension Score) / Time Spent

Example for 8-10 AM:

- Pages Completed: 15

- Comprehension Score (from quiz results): 92%

- Time Spent: 2 hours

- Efficiency = (15 × 0.92) / 2 = 6.9 (highest efficiency)

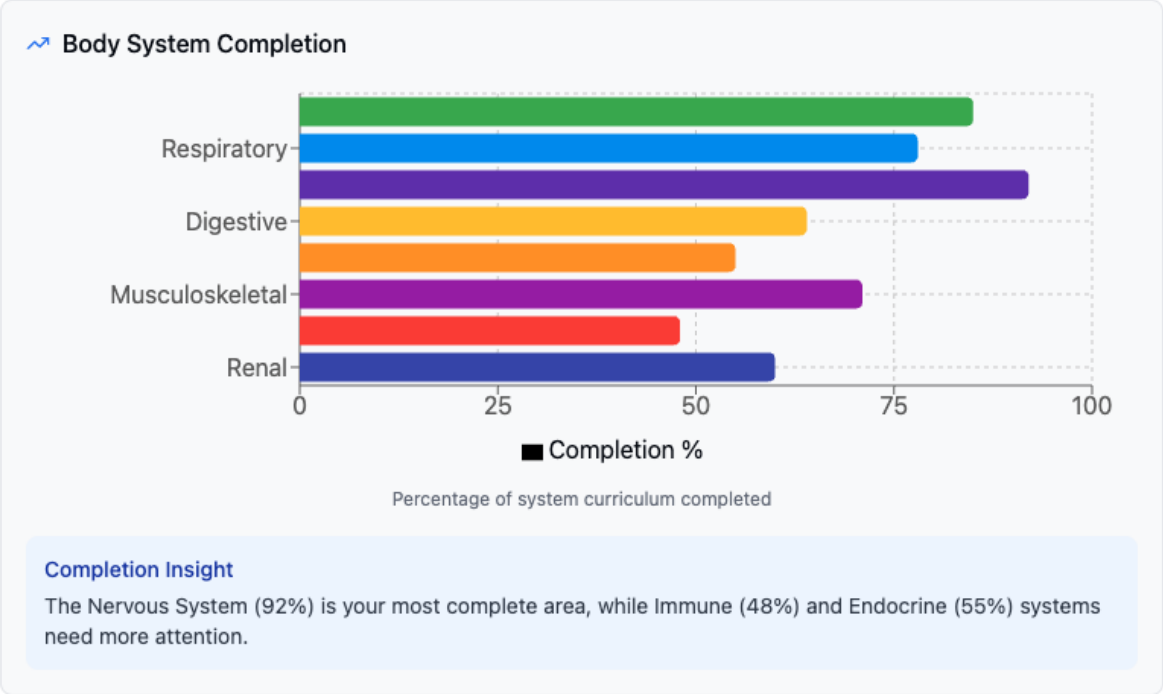
Compare to 12-2 PM:

- Pages Completed: 12

- Comprehension Score: 68%

- Time Spent: 2 hours

- Efficiency = (12 × 0.68) / 2 = 4.08 (lower efficiency)



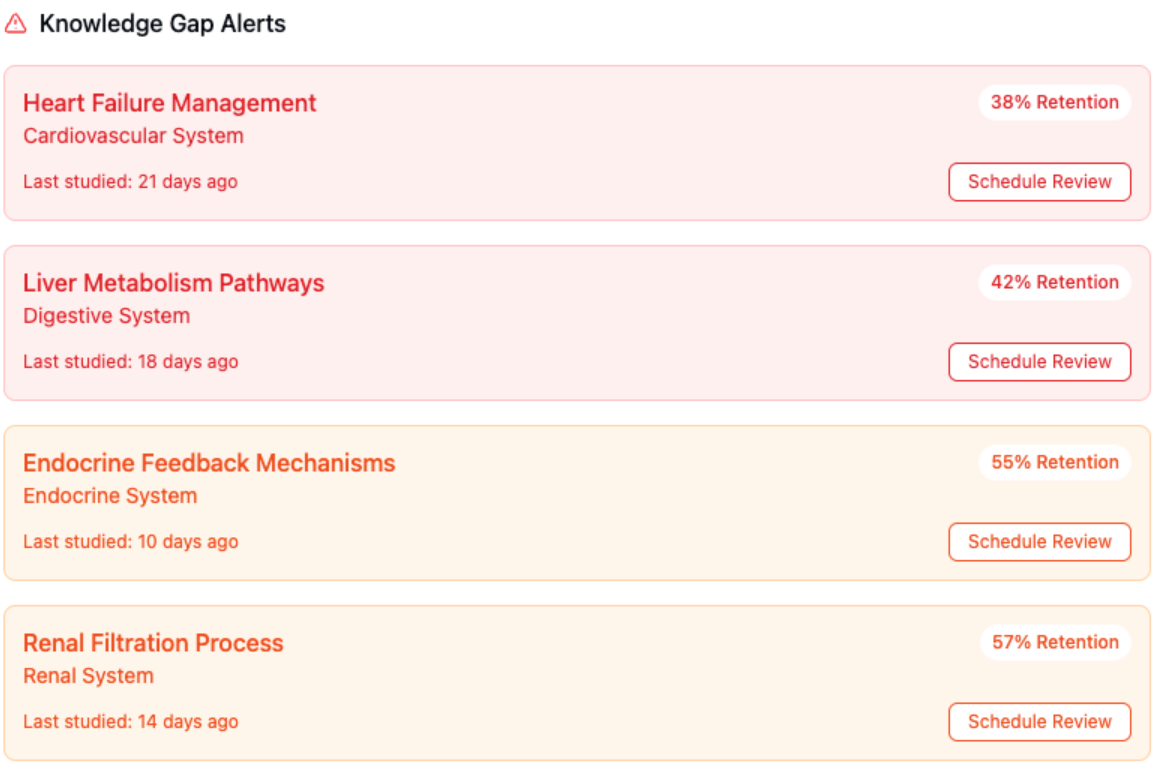
Study Efficiency by Time = (Pages Completed × Comprehension Score) / Time Spent

Example for 8-10 AM:

- Pages Completed: 15
- Comprehension Score (from quiz results): 92%
- Time Spent: 2 hours
- Efficiency = (15 × 0.92) / 2 = 6.9 (highest efficiency)

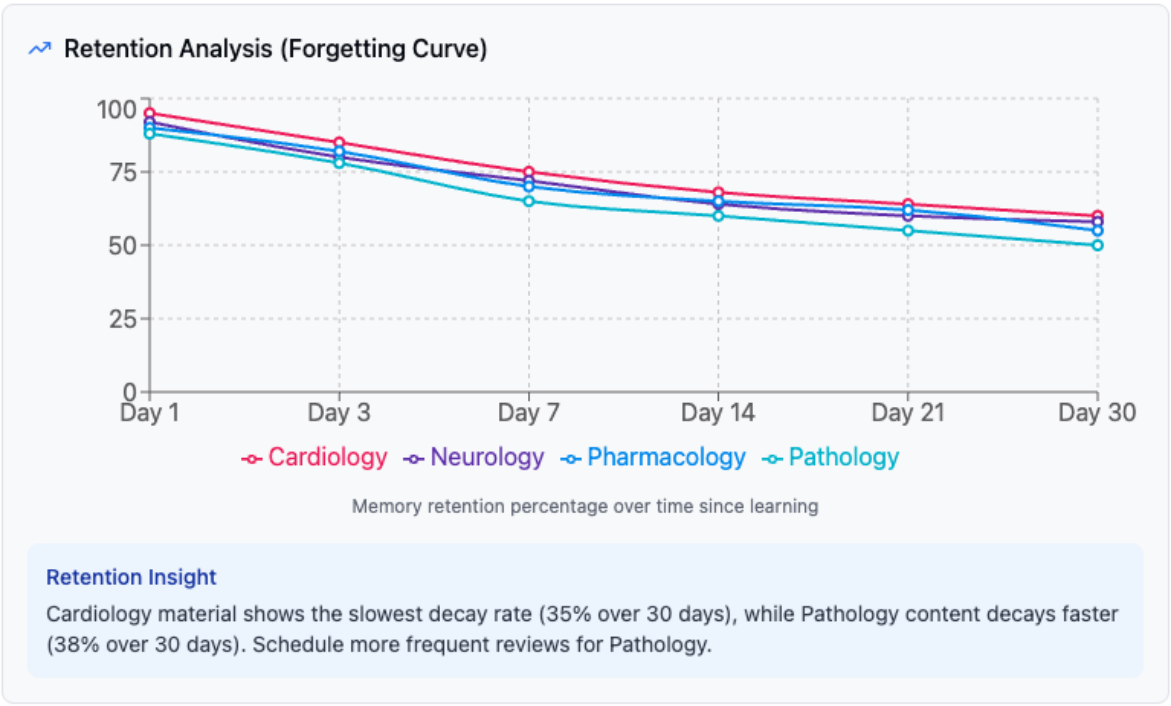
Compare to 12-2 PM:

- Pages Completed: 12
- Comprehension Score: 68%
- Time Spent: 2 hours
- Efficiency = (12 × 0.68) / 2 = 4.08 (lower efficiency)



Knowledge Gap Alerts

Personalized Recommendations



Retention Rate Calculation

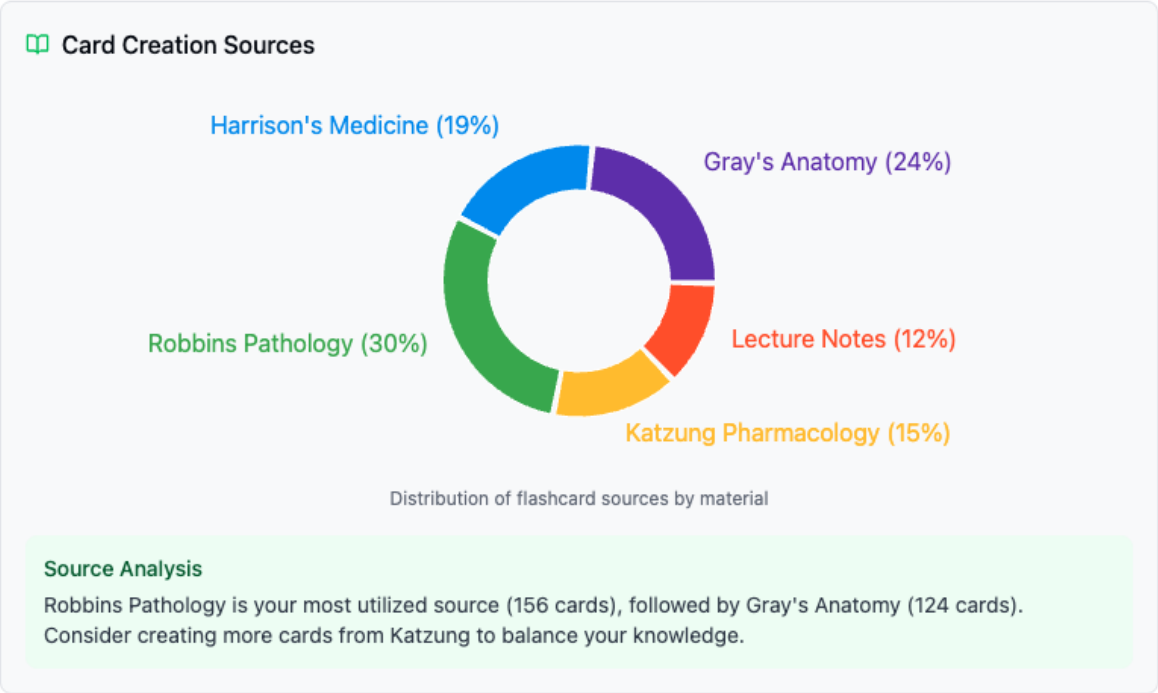
Retention Rate = (Number of Correct Answers) / (Total Attempts) × 100%

Example for Cardiology cards at different intervals:

- Day 1: 95 correct out of 100 = 95%
- Day 3: 85 correct out of 100 = 85%
- Day 7: 75 correct out of 100 = 75%
- Day 30: 60 correct out of 100 = 60%

Forgetting Curve Decay Rate = (Initial Retention - Final Retention) / Days

= (95% - 60%) / 30 days = 1.17% per day

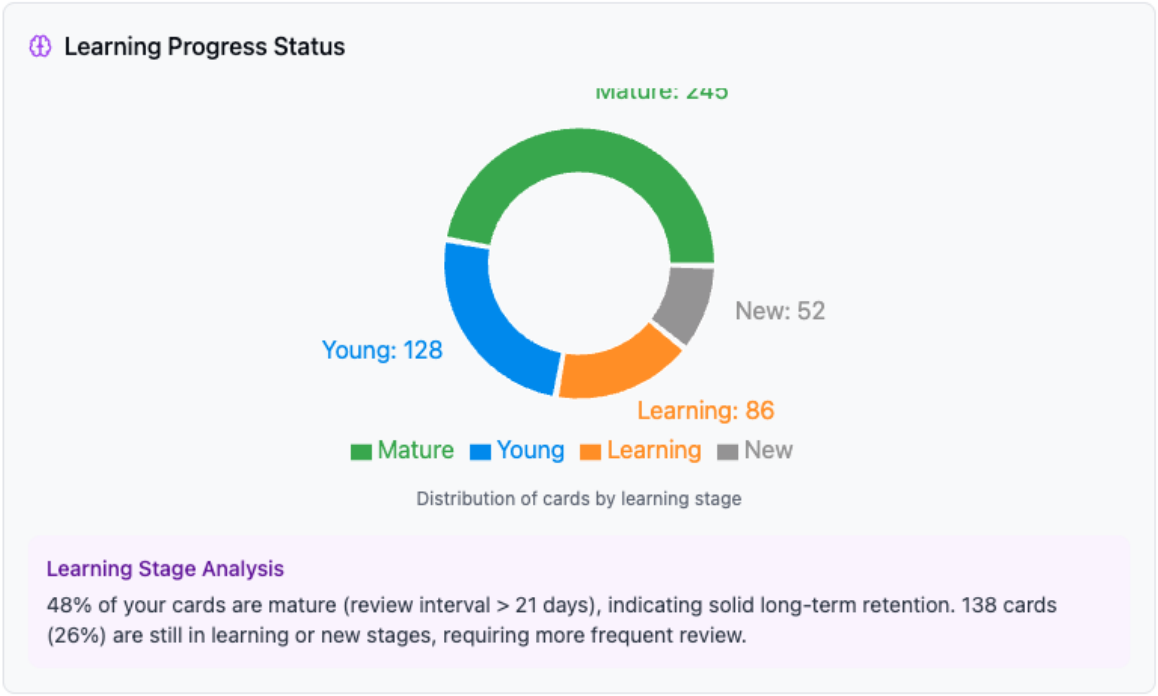


Source Analysis Calculation

Source Distribution = (Cards from Source) / (Total Cards) × 100%

Example:

- Robbins Pathology: 156 cards out of 521 total = 30%
- Gray's Anatomy: 124 cards out of 521 total = 24%
- Harrison's Medicine: 98 cards out of 521 total = 19%
- Katzung Pharmacology: 78 cards out of 521 total = 15%
- Lecture Notes: 65 cards out of 521 total = 12%



Learning Progress Status

Difficulty Score = (Average Number of Attempts) × 50% +
(Percentage of "Again" Responses) × 30% +
(Time Spent per Card) × 20%

Example for Heart Failure cards:
- Average Attempts: 3.8 out of max 5 = 76%
- "Again" Responses: 65%
- Average Time: 12 seconds out of typical 8 seconds = 150%
- Difficulty Score = (0.76 × 0.5) + (0.65 × 0.3) + (1.5 × 0.2) = 0.38 + 0.195 + 0.3 = 0.85 (high difficulty)

Cards Due on Day N = Σ (Cards with Interval = N)

Example:

- Today: 48 cards due (sum of all cards with interval = 0)

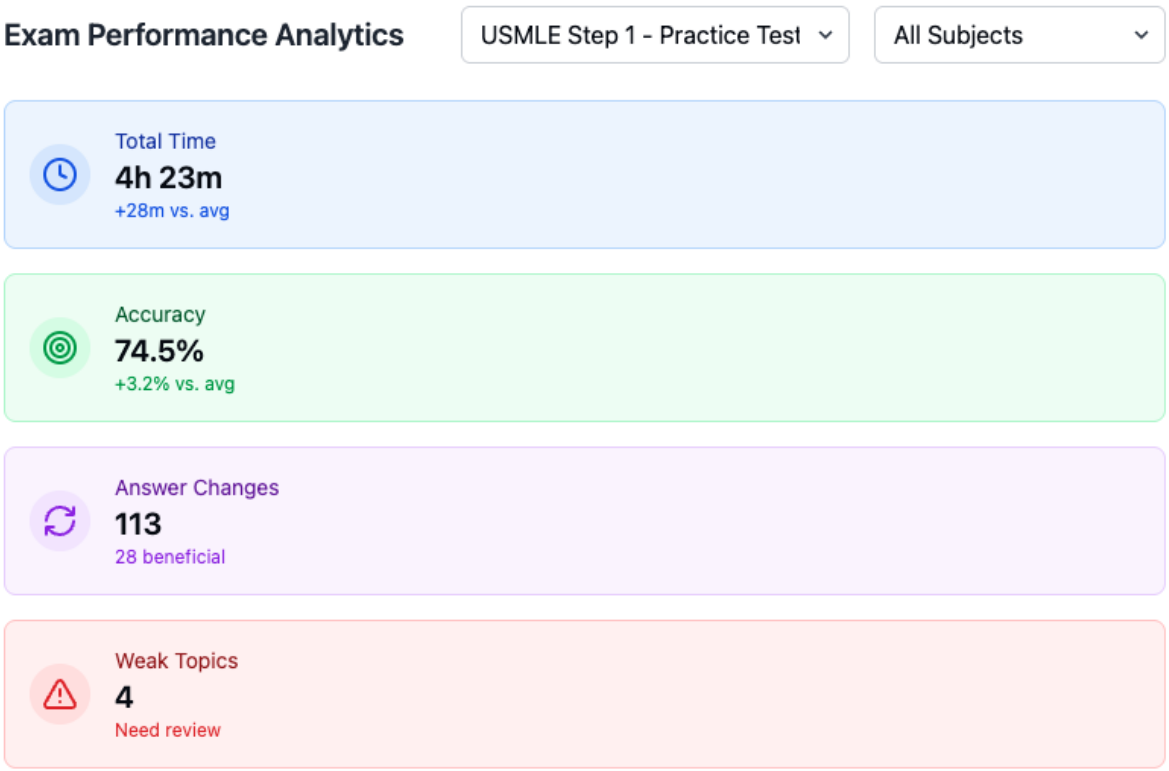
- Tomorrow: 65 cards (all cards with interval = 1)

- In 6 days: 78 cards (all cards with interval = 6)

Workload Prediction = Σ (Cards Due on each day in forecast period)

= 48 + 65 + 42 + 38 + 25 + 52 + 78 = 348 cards in next week

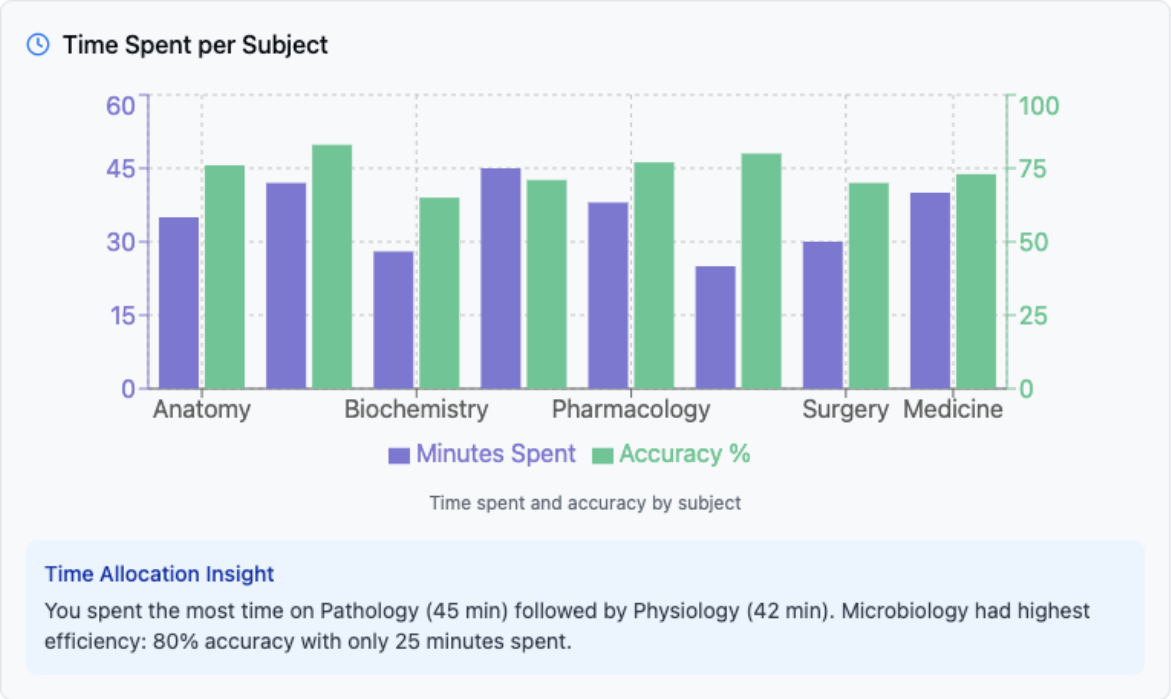
Practice Module:



Exam Performance Analytics

The exam module analytics leverage key data points captured during practice tests:

1. **Time tracking metrics** (time per question, subject, topic)
2. **Confidence levels** (low/medium/high self-ratings)
3. **Answer change patterns** (initial vs. final answers)
4. **Navigation behavior** (skips, revisits, answering order)
5. **Topic and subject performance** (accuracy, speed, confidence)



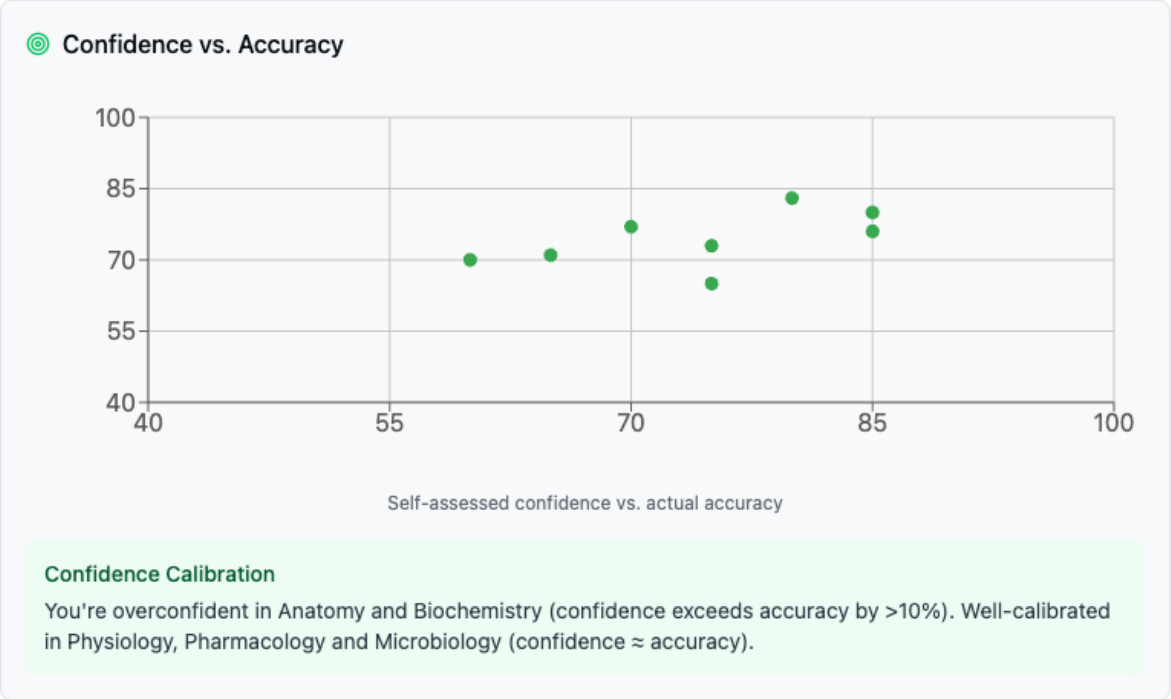
1. Time Spent per Subject vs. Accuracy (continued)

Sample Calculation:

```
For Pathology:
- Total time spent: 45 minutes
- Number of questions: 35
- Correct answers: 25
- Accuracy = (25/35) × 100% = 71%

For Microbiology:
- Total time spent: 25 minutes
- Number of questions: 20
- Correct answers: 16
- Accuracy = (16/20) × 100% = 80%

Time Efficiency = Accuracy / Time Spent
Pathology: 71% / 45 min = 1.58% per minute
Microbiology: 80% / 25 min = 3.2% per minute (more efficient)
```

Insights: The visualization reveals which subjects yield the best return on time investment. In this example, Microbiology shows higher efficiency with 80% accuracy

despite less time spent, suggesting better knowledge retention or question-answering strategy in this area.



2. Confidence vs. Accuracy Calibration

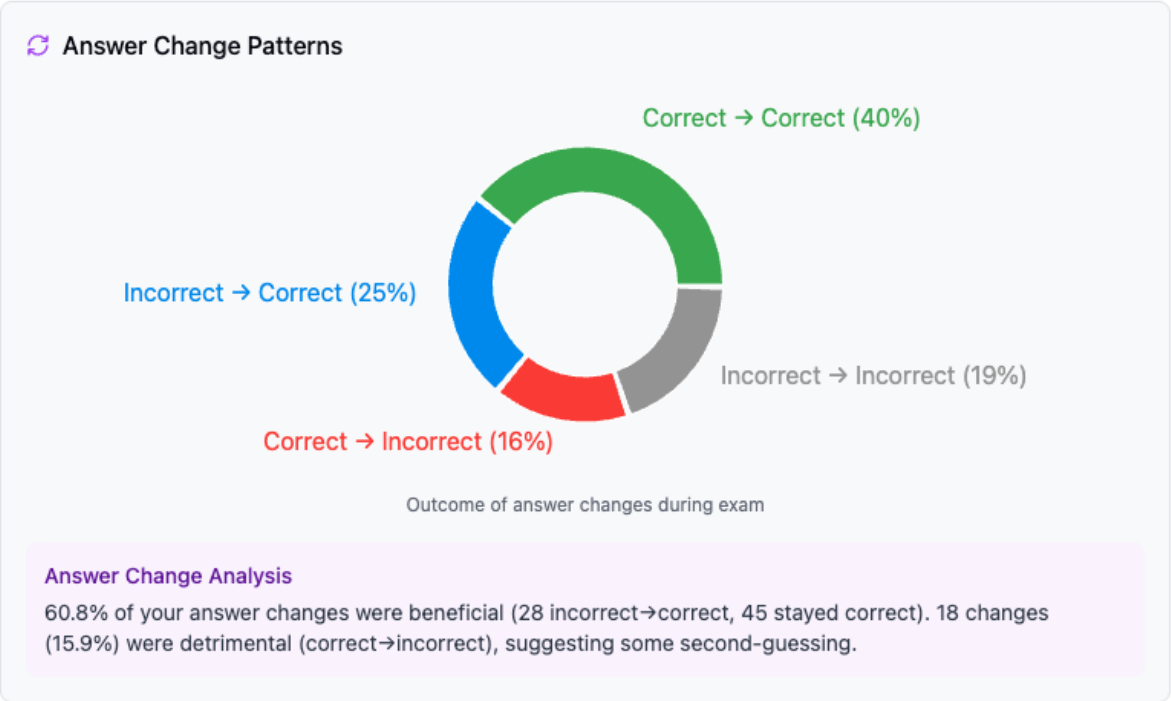
This scatterplot visualization compares your self-assessed confidence with your actual performance across subjects.

Sample Calculation:

```
For Anatomy:
- Average confidence rating: 85% (based on confidence markers)
- Actual accuracy: 76%
- Calibration gap = 85% - 76% = +9% (slight overconfidence)

For Physiology:
- Average confidence rating: 80%
- Actual accuracy: 83%
- Calibration gap = 80% - 83% = -3% (slight underconfidence)
```

Insights: Points above the diagonal line indicate overconfidence (confidence > accuracy), while points below indicate underconfidence. Well-calibrated subjects fall close to the diagonal. The visualization helps identify subjects where you might be

overestimating your knowledge (risking mistakes) or underestimating it (potentially second-guessing correct answers).



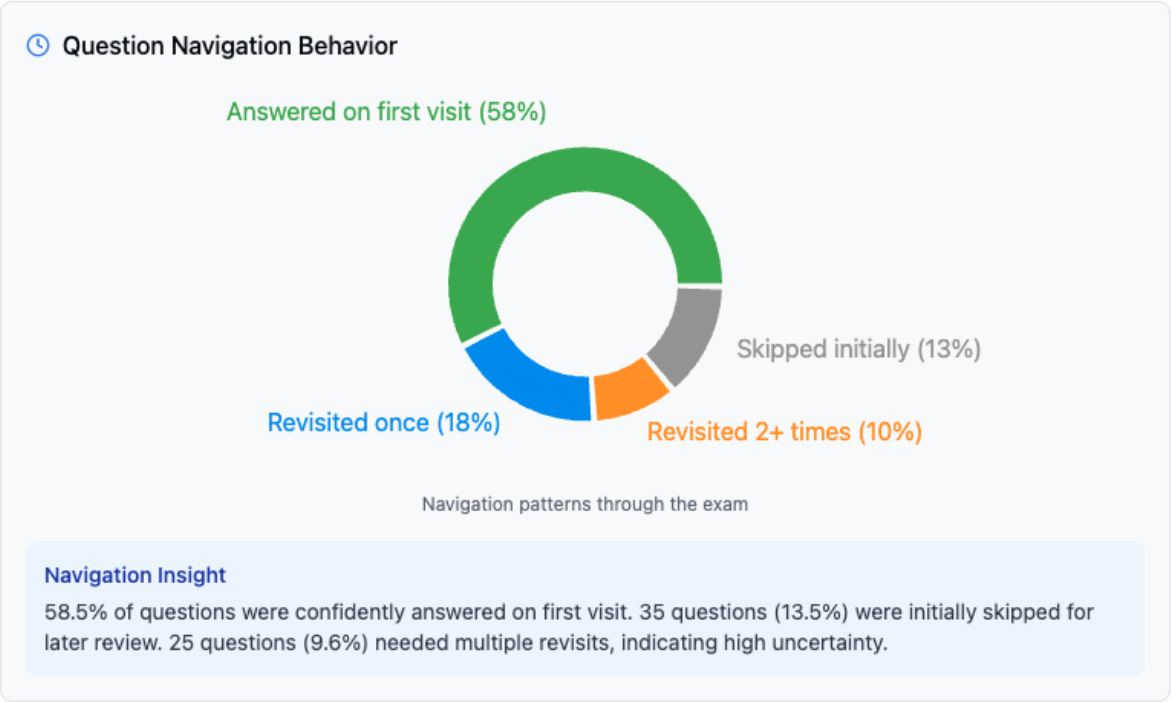
3. Answer Change Patterns

This pie chart shows the outcomes of changed answers during the exam.

Sample Calculation:

```
Total answer changes: 113
- Correct → Correct: 45 (39.8%)
- Incorrect → Correct: 28 (24.8%) [Beneficial]
- Correct → Incorrect: 18 (15.9%) [Detrimental]
- Incorrect → Incorrect: 22 (19.5%)

Net benefit = Beneficial changes - Detrimental changes
            = 28 - 18 = +10 correct answers
```

Insights: This analysis challenges the conventional wisdom that you should "always stick with your first answer." In this example, changing answers resulted in a net gain of 10 correct answers, suggesting that thoughtful reconsideration can be beneficial.



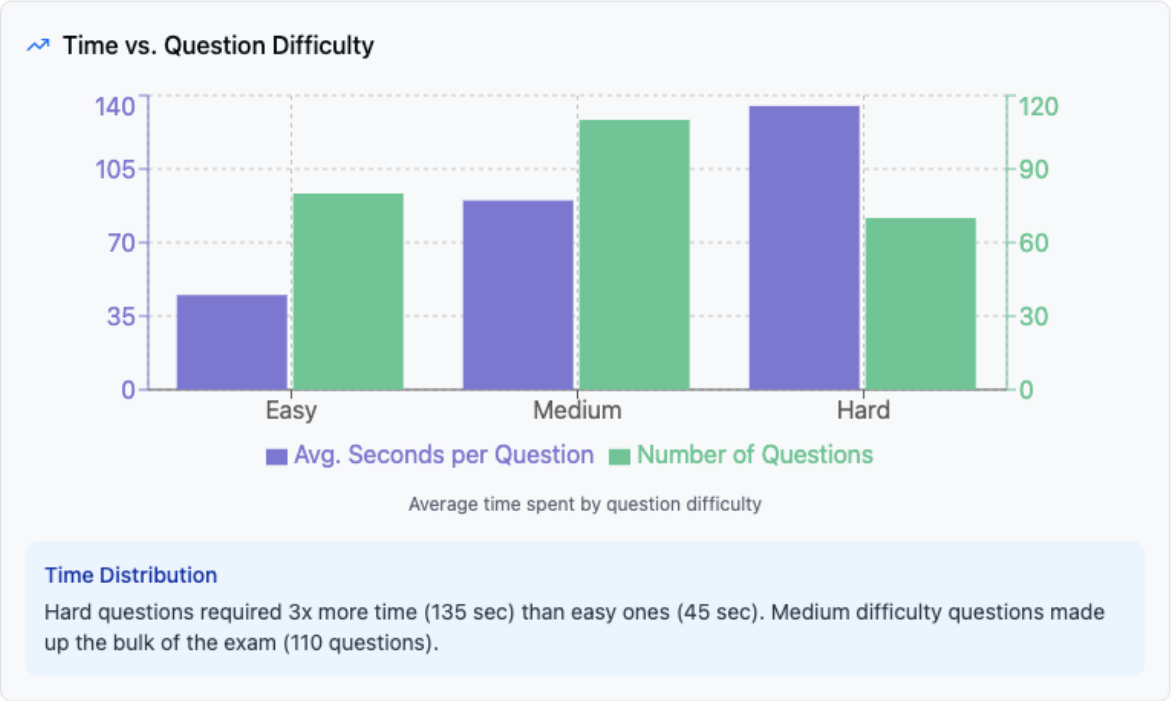
4. Question Navigation Behavior

This visualization analyzes how you move through the exam questions.

Sample Calculation:

```
Total questions: 260
- Answered on first visit: 152 (58.5%)
- Revisited once: 48 (18.5%)
- Revisited 2+ times: 25 (9.6%)
- Skipped initially: 35 (13.4%)

Revisit ratio = (Questions revisited / Total questions) × 100%
              = (48 + 25) / 260 × 100% = 28.1%
```

Insights: High revisit rates may indicate uncertainty in specific topics. The 9.6% of questions requiring multiple revisits are particularly worth examining as they likely represent significant knowledge gaps or conceptual confusion.



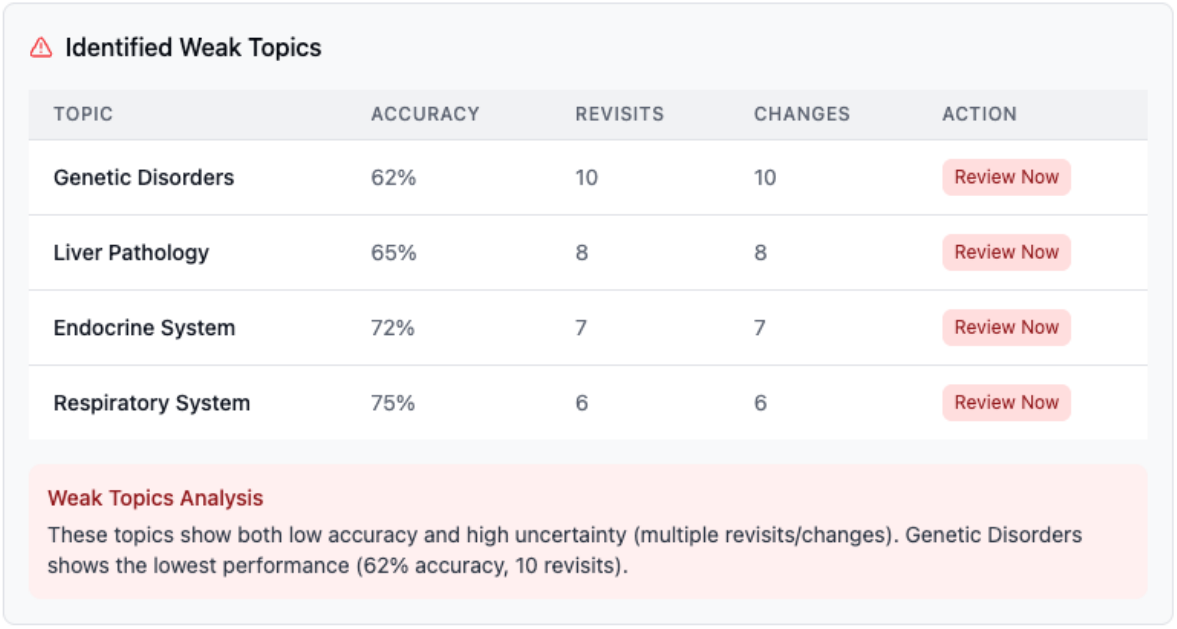
5. Time vs. 1328

This chart shows the relationship between question difficulty and time spent.

Sample Calculation:

```
Average time spent:
- Easy questions (80 questions): 45 seconds
- Medium questions (110 questions): 90 seconds
- Hard questions (70 questions): 135 seconds

Time ratio (Hard:Easy) = 135/45 = 3:1

Speed analysis by difficulty:
- Easy: 80 questions × 45 sec = 3,600 sec (60 min)
- Medium: 110 questions × 90 sec = 9,900 sec (165 min)
- Hard: 70 questions × 135 sec = 9,450 sec (157.5 min)
```

Insights: Hard questions require 3 times more time than easy ones, with medium questions taking exactly the middle value. Despite being fewer in number, hard questions consume a disproportionate amount of total exam time (157.5 minutes vs. 60 minutes for easy questions).



6. Weak Topics Identification

This table identifies topics requiring urgent review based on multiple metrics.

Sample Calculation:

```
For Genetic Disorders:
- Accuracy: 62%
- Revisits: 10
- Answer changes: 10

Weakness score = (100% - Accuracy) × (1 + Revisits/10) × (1 + Changes/10)
               = 38% × 2 × 2 = 152

For Cardiac Physiology:
- Accuracy: 92%
- Revisits: 2
- Answer changes: 2

Weakness score = (100% - 92%) × (1 + 2/10) × (1 + 2/10)
               = 8% × 1.2 × 1.2 = 11.52
```

Insights: The weakness score formula weighs not just low accuracy but also uncertainty indicators (revisits and changes). Genetic Disorders has a weakness score

of 152 compared to Cardiac Physiology's 11.52, making it a much higher priority for review.

Topic Performance Breakdown

TOPIC	ACCURACY	CONFIDENCE	AVG TIME (SEC)	ANSWER CHANGES	CALIBRATION
Cardiac Physiology	92%	85%	75	2	Well-calibrated
CNS Anatomy	85%	78%	90	5	Well-calibrated
Liver Pathology	65%	70%	110	8	Well-calibrated
Antibiotics	78%	75%	85	4	Well-calibrated
Endocrine System	72%	68%	95	7	Well-calibrated
Renal Physiology	88%	80%	80	3	Well-calibrated
Genetic Disorders	62%	55%	115	10	Well-calibrated
Respiratory System	75%	72%	90	6	Well-calibrated

7. Topic Performance Breakdown

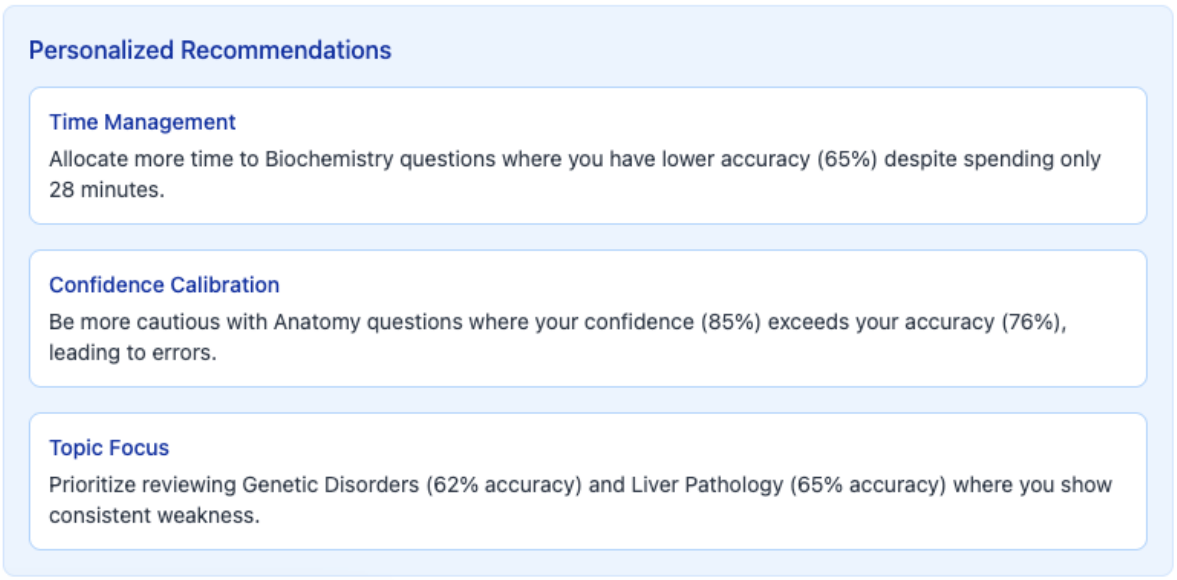
This detailed table provides a comprehensive view of performance across specific medical topics.

Sample Calculation:

```
For CNS Anatomy:
- Accuracy: 85%
- Confidence: 78%
- Average time: 90 seconds
- Answer changes: 5

Calibration = Confidence - Accuracy = 78% - 85% = -7%
```

```
(Underconfident)

For Liver Pathology:
- Accuracy: 65%
- Confidence: 70%
- Average time: 110 seconds
- Answer changes: 8

Calibration = Confidence - Accuracy = 70% - 65% = +5% (Slight
overconfidence)
```

Insights: This granular breakdown helps identify specific topics (rather than broad subjects) that need attention. It also reveals behavioral patterns - for example, the data suggests underconfidence in CNS Anatomy despite strong performance, while showing slight overconfidence in Liver Pathology where performance is weaker.



Educational Value of These Analytics

These analytics provide medical students with actionable insights to improve their exam performance:

1. **Time allocation optimization** - Identifying which subjects deserve more study time based on performance gaps
2. **Confidence calibration** - Learning when to trust or question their instincts

3. **Answer change strategy** - Understanding when reconsidering answers helps or hurts
4. **Weak area prioritization** - Focusing review on topics with lowest performance and highest uncertainty
5. **Exam-taking behavior awareness** - Recognizing patterns in navigation and time management

System Architecture

Main Functionalities:

- Flashcard
- Exams
- PDF Analytics

Schema

- Flashcard Tables:
    - Flashcard Analytics (to store the last seen of a subject)
    - Flashcard-Review-History (Stores each review made by user)
    - Flashcard-User (Each Flashcard created by user)
- Pdf Tables:
    - Pdf-Topic-Table (stores each topic extracted from the each page of pdf)
    - Pdf-Session-Table (stores users activity on a particular page)
    - Pdf-Completion-History (Stores pageNo when a page is completed reading)
- Application Analytics:
    - Application-User-Analytics (stores each user's application data)
- Generic:
    - Subject-Table (stores all the subject data of the application)
    - Subject-Analytics (stores each user's data regarding the subject)
    - Topic-Analytics-dev (stores each user's analytics regarding each subject)

Exam tables:

- Exam Tables (stores exam details)
- Exam-Question-Junction (maps question and question types to junction)

- Exam-Attempt (stores attempt made by each user)
- Exam-answer (stores each answer details of a user)
- Question Type Tables:
    - Fill_blanks
    - case_based
    - mcq
    - script
    - matching

Services:

- Flashcard-Service
- Exam-Service
- Pdf-Analytics
- Graph

Routes

- Pdf-Service
    - Create topic
    - Create Session and update analytics

When users session information is given the flow goes as follows
sessionKey is created using pdfId, userId and pageNo
if existing session then increment by the currentSessionTime
retrieve thresholdTime from PDF_TOPIC_TABLE and compare with current time
  if increased then store it in PAGE_COMPLETION_HISTORY_TABLE
retrieve data for the pdf from PDF_ANALYTICS_TABLE -> increase the totalTimeSpent
retrieve data for APPLICATION_USER_ANALYTICS_TABLE -> increase time for pdf reading
retrieve data for subject from SUBJECT_ANALYTICS_TABLE -> increase time for subject
retrive data for topic from TOPIC_ANALYTICS_TABLE -> increase time for topic

    - Create Multiple Session
        - Create session but for multiple pages for a given pdf
    - Check Topic for pdf
        - Returns all topic for a pdf
- Flashcard Service
    - Space Repetition Algorithm
        - lapsesCount
        - ease factor

        - interval
        - review count
    - Create Flashcard
        - lapsesCount and interval set to 0
        - ease factor's default value set to 2.5
    - get flashcard by id
    - bookmark flashcard
    - get bookmarked flashcard
    - Submit Review
        - Based on difficulty set the quality from 0 to 5
        - if quality less than 3
            - interval = 1
            - review count = 0
            - lapses count++
        - else
            - review count++
            - interval = interval * ease factor
        - ease factor recalculated based on the quality
        - next due date is set based on new interval
    - Get due flashcards
        - All flashcards before today's date are returned
    - Again clicks
        - When user clicks again button, the quality value is set to 1
        - Interval is set to next day
        - Spaced repetition algorithm is applied
    - Get all flashcard by subjects
    - Get dashboard data
        - Get due flashcards based on the dates before today's date
        - Get bookmarked flashcard
        - Get recent subjects of the user from the FLASHCARD_ANALYTICS TABLE
        - Get the no of flashcards for each book
        - Counts no of flashcard by subjcts
        - Counts no of flashcard by book
        - Calculates maturity of each flashcard and divides it into 4 categories: New, Learning, Young, Mature
        - Returns all of these datas for the dashboard

Graphs Service Routes:

- getAllSubjectProgress

    - Scan all subjects from subjects table
    - For each subject, calculates Flashcard success rate, Reading progress, Time efficiency
    - DB accessed: SUBJECT_TABLE, SUBJECT_USER_ANALYTICS, FLASHCARD_TABLE
- getKnowledgeGapHeatmap
    - Get all subjects and group topics by subject from SUBJECT_TABLE and TOPIC_ANALYTICS_TABLE
    - For each topic, calculate mastery score using flashcard performance(Retrieved from flashcard table), completion rate (retrieved from TOPIC_ANALYTICS_TABLE) and time efficiency (from TOPIC_ANALYTICS_TABLE
- getLearningToolUsage
    - Extract time values from analytics: flashCardTime, pdfTime, chatbotTime, practiceTestTime
    - APPLICATION_USER_ANALYTICS table is used
- getConceptMasteryDistribution
    - Get all topics for the user
    - calculates mastery score (the same used in knowledge gap heatmap)
    - Groups into 4 categories
    - Same db
- getStudyPerformanceAnalytics
    - Get all subjects from subjects table
    - Calculate retention score: (correctAttempts / totalAttempts) × 100 (retrieved from flashcard table)
    - Get hours spent from subject analytics (fetch from subject_analytics table)
- getWeeklyReviewPattern
    - FLASHCARD_TABLE and FLASHCARD_REVIEW_HISTORY tables used
    - Fetches for the newly created flashcard from FLASHCARD_TABLE and reviewed flashcards from FLASHCARD_REVIEW_HISTORY. from the previous 7 days
- getBodySystemCompletion
    - Group flashcards by bodySystem field
        - Count total flashcards
        - Count completed flashcards (repetitions > 0 && lapsesCount <= repetitions)
        - Calculate completion percentage
- getRetentionAnalysis
    - SUBJECT_TABLE and FLASHCARD_REVIEW_HISTORY_TABLE

    - Define time intervals: [1, 3, 7, 14, 21, 30] days
    - For each subject and time interval:
        - Filter reviews within the time cutoff
        - Calculate retention rate: (correctAnswers / totalAttempts) × 100%
- timeSpentPerSubject
    - Get time spent per subject with accuracy data
    - Get subject time data from subject analytics
    - Get all flashcard review history from FLASHCARD_REVIEW_HISTORY_TABLE for accuracy calculation
    - Map flashcard IDs to subjects using flashcard table
    - Calculate accuracy by subject: correctReviews / totalReviews
- getUserOverallAnswerPatterns
    - DB used
        - EXAM_ATTEMPT
        - EXAM_ANSWER
        - EXAM_QUESTIONS_JUNCTION
        - Question Type Tables
        - EXAM_TABLE
    - Get all completed exam attempts for user
    - Analyze answer change patterns (Incorrect→Correct, Correct→Incorrect)
    - navigation patterns (revisits, skips, first-visit answers)
- getExamPerformanceAnalytics
    - EXAM_ATTEMPT, EXAM_ANSWER, EXAM_QUESTIONS_JUNCTION, Question Type Tables
    - Get all completed attempts from EXAM_ATTEMPT table
    - Process each answer to calculate:
        - Total time spent and average per question
        - Overall accuracy percentage
        - Answer changes (beneficial vs detrimental)
        - Topic and difficulty performance
- incrementBreakTime
    - Increment user's break time
    - APPLICATION_USER_ANALYTICS table used
    - Add new break time to existing break time
- getWorkBreakRatio
    - Calculate work-break ratio for study wellness
    - APPLICATION_USER_ANALYTICS table used
    - Calculate total study time: chatbotTime + flashCardTime + pdfTime + practiceTestTime
    - Get total break time

    - Calculate work-break ratio: (totalStudyTime - totalBreakTime) / totalBreakTime
    - Calculate consistency: activeDays / daysSinceCreation
- getFlashcardAnalyticsDetailed
    - FLASHCARD_TABLE used

Calculate multiple distributions:

    - Maturity levels (using calculateMaturity function)
    - Material source distribution
- calculateExamAttemptStats:
    - EXAM_ATTEMPTS and EXAM_ANSWERS table used
    - for each attempt sum up and average the confidence and accuracy by going through each answer retrieved from EXAM_ANSWERS
- getUserWeakTopics:
    - EXAM_ATTEMPT, EXAM_ANSWER, EXAM_QUESTIONS_JUNCTION and Question Type Tables
    - Get all completed exam attempts and their answers
    - Map questions to topics using question type tables
    - For each topic, calculate:
        - Accuracy percentage
        - Number of incorrect answers
        - Average revisits and changes per question
    - Identify weak topics (accuracy < 70%)
- Update Learning tool usage
    - APLICATION_USER_ANALYTICS used
    - according to the incremented time, updated the fields

- Exam-Service:
- Analyse the graphs