Question 1

Investigation approach, challenges, and ethics

A. High level investigation plan

1. Define goals and legal constraints

Clarify the specific investigatory questions

Involve HR to confirm whats allowed before accessing employee communications or private files.

2. Collect relevant telemetry

For example network logs, access records, email metadata, supplementary sources

3. Feature engineering and baseline behavior

Build per user baseline profiles

Contextual features e.g. job role, team

4. Anomaly detection

Statistical method eg thresholding for sudden spikes.

Rule-based detection eg many failed admin logins, access to sensitive systems not in role, mass external email attachments.

Unusual sequences of events is like admin access - export - email.

Graph analysis – weird graph patterns.

5.investigation.

If strong evidence exist, perform deeper forensic steps only after approvals.

6. Iterate and document

Keep and audit every access, query, and extract.

Record decisions

B. Challenges distinguishing legitimate vs malicious

Legitimate work can look fishy. e.g. a data scientist legitimately exports

Context missing in logs automated mistake can cause it.

Many false positives .this can cause a lot of wrong alarms.

Insider stealth tactics. Insiders may act slowly and use normal tools to blend in.

Data quality & gaps for example missing logs, inconsistent timestamp.

Mitigations:

Use role specific baselines, combine multiple signals

C. Privacy, ethics, and legal balance

Minimize collection - collect and analyze minimally required metadata before any content review.

Get approvals: involve HR and maintain documented approvals

Least privilege access: limit who can run queries and who can see identifiable employee data

Transparency & policy: ensure internal policies describe monitoring practices and employees have been notified like in terms and conditions

Retention & deletion: only retain investigation items for the time needed

D. Transparency and maintaining ethical standards

Document everything — data sources used, queries run, results approvals and who accessed what thing

Independent review — involve a compliance officer or legal reviewer before taking major actions

Protect employees not involved —limit access to irrelevant PII in outputs.

E. Communicating findings

For technical stakeholders:

Provide thorough items such as timelines and suggested next forensic steps for log done.

Provide repeatable scripts and queries and provide reasoning and confidence metrics for every result.

For non-technical stakeholders:

Provide a high level risk summary, including the information that was accessed, the reason it is sensitive and the degree of malicious intent.

Make use of visuals such as timelines eg risk levels and simplified graphs

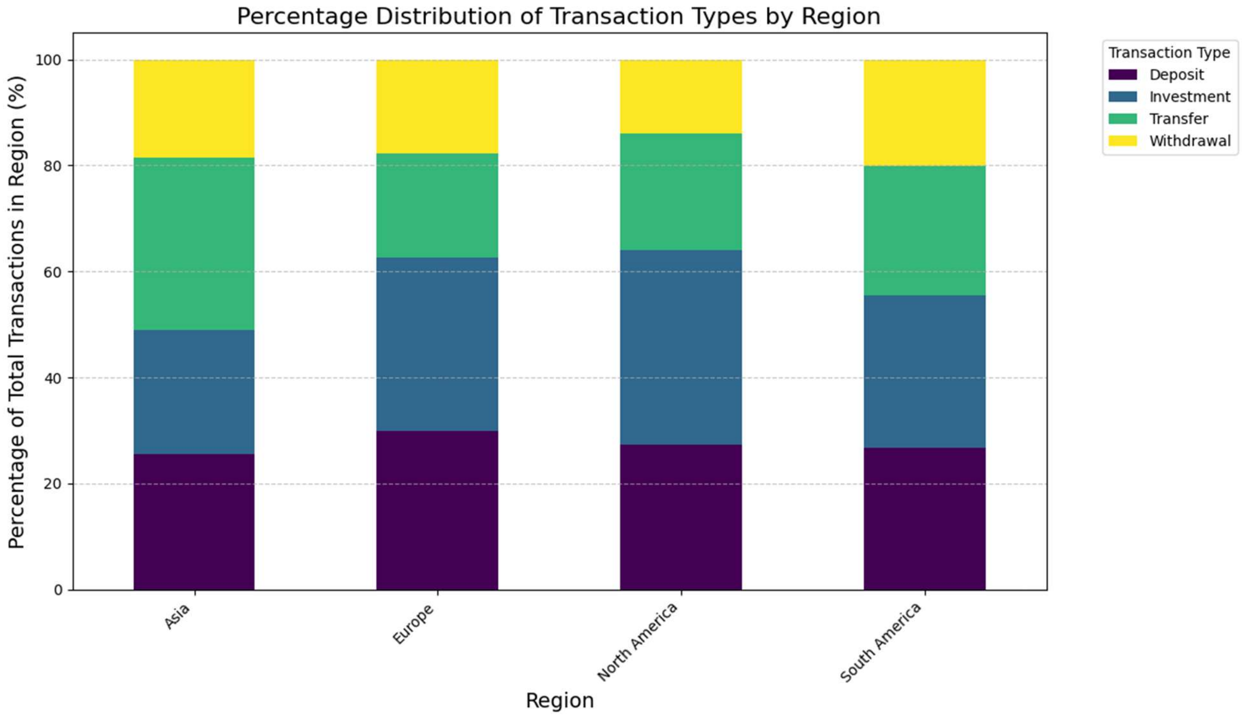
Suggested communication strategy:

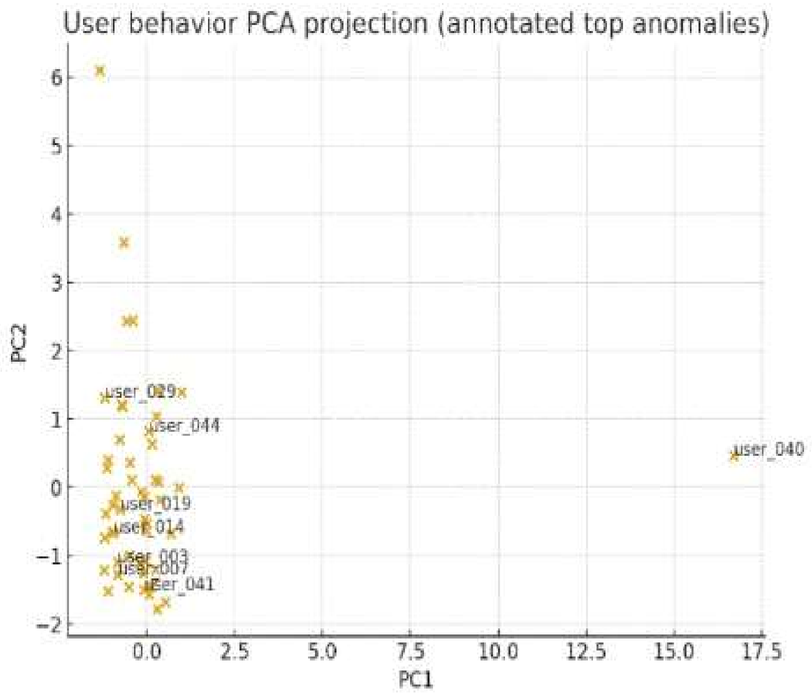
Executive one page summary.

Detailed appendix for technical teams and auditors.

Regular status updates and an outcomes report when the investigation concludes.

Question 2




User behavior PCA projection (annotated top anomalies)
PC2
6
5
4
3
2
1
0
−1
−2
PC1
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
user_029
user_044
user_040
user_019
user_014
user_003
user_007
user_041

[redacted]