IEUK 2025: Engineering Task

Name: Tanisha Srivastava

Key Findings

1. **Suspicious Requests and Attack Patterns**
   Multiple path traversal attempts (e.g. **/api/v5/database/../../../**) from IP **45.133.1.1** indicate likely reconnaissance. Brute-force login attempts with weak credentials (e.g. **guest, Test123**) from IPs like **185.220.100.77** point to credential stuffing bots.

2. **Slow Endpoints (≥500ms)**
   Several search-related endpoints (e.g. **/search?duration=...**) show high average response times due to inefficient queries or lack of backend indexing.

3. **High-Volume IPs**
   IPs like **45.133.1.1** and **35.185.0.156** made over 5,000 requests, suggesting bot scraping, or even early-stage DDoS behaviour.

4. **Traffic by Country**
   While most traffic is from legitimate regions (UK, US, DE), high volumes from risk-prone countries (e.g. RU, IR, CN) could indicate elevated threat exposure.

5. **Status Codes**
   - 200 OK: Normal (354k+)
   - 404 Not Found: High (38k+) — may indicate bots probing or broken internal links
   - 401/403: ~6,500 — failed logins
   - 429: Rate limiting triggered (1,821) — positive
   - 5xx errors (~8,000): Points to backend instability.

6. **Popular Endpoints**
   Pages like **/contact, /about, /podcasts/...** show genuine user interest and are good candidates for caching.

Recommendations

- Implement IP blocking, rate limiting, and WAF rules against path traversal and brute-force attempts.
- Optimise slow endpoints via query tuning and caching (e.g. ElasticSearch).
- Address 404 errors by cleaning broken internal links.

- Monitor unusual spikes by country or IP, and alert on repeated failed logins.

Assumptions

- Access logs are complete and user-agents are correctly parsed.
- Log timestamps and IPs are reliable.
- Internal servers and APIs are identifiable from endpoint structure.

Cost Considerations

- WAF setup and IP threat reputation tools (e.g. AbuseIPDB): ~£30–£100/month.
- Developer time for query optimisation: ~£500–£1000.
- Monitoring tools (e.g. Grafana, Prometheus): open-source, but hosting may cost £10–£50/month.

How to Run the Code

Link: https://github.com/Tanisha-gitcodes/ieuk-task-2025-main

Prerequisites:

- Python 3.10 or later installed
- **pandas** library installed

Run this command if needed:

```
pip install pandas
```

Steps:

1. Place the log file (**sample-log.log**) in the same directory as **log_analyser.py.**
2. Open a terminal or PowerShell in that directory.

Run the script using:

```
python log_analyser.py
```

3. Or, if you're using a specific version:

```
python3 log_analyser.py
```