# **Description on Log Parsing Project**

For my coding project, I wrote a program that parses logs and generates a related report. The focus was on analysing traffic logs from an Apache web server and identifying any suspicious actions within. This came about through careful log analysis and the setup of alerts based on specific policies to identify the potential questionable actions. I handled the parsing of those logs myself. That process generated two separate output files. After that, I reviewed the findings to identify patterns that might indicate harmful actions.  
  
Firstly, I coded a Python parser that extracted necessary fields from the logs, such as timestamps, IP addresses and request paths. The logs were combined into a summary.csv file to display statistics on high-traffic IPs, the distribution of HTTP response codes, and the most accessed endpoints. To identify any unusual activity, I set up alert policies that detect IPs with high request volumes, repeated login attempts, or frequent access to sensitive paths like /admin or /.env. This is to make it easier to see what type of activity occurred and where suspicious patterns might exist.  
  
After parsing was complete, the results were recorded in different CSV files. "summary.CSV" displayed the overall statistics of the logs as mentioned earlier. "alerts.CSV" showed suspicious or unusual activities detected in the logs while they were parsed. The files helped me identify malicious intent within the web server. The results of the audit process were then added to the overall report, which summarised the activities identified and the steps taken to address them, using log analysis.  
  
After writing up the report, I wrote a conclusion summarising the activity detected and why it might be considered suspicious. I outlined the strengths and gaps I had identified and offered clear, practical recommendations to improve monitoring and strengthen security practices.

# **Reflection on Log Parsing Project**

My overall aim was to analyse and evaluate whether the activity within the logs showed any signs of suspicious behaviour. I achieved this by comparing the behaviour observed in the logs with typical patterns of malicious activity, explaining which rules were met, and providing recommendations for actions to address them. I also clearly demonstrated the steps I took by outlining the dataset, writing the parsing script, and showing how the alerts were generated and documented within the overall report.

To improve this project, I would implement the use of Splunk software to display log statistics rather than code and CSV files. Splunk allows you to display data through a variety of graphs, which are easier to communicate to non-technical teams and clearly show trends in the gathered data. I did initially try to use Splunk as part of the project, but due to the time constraints and the additional learning curve involved within the short time remaining, I didn't successfully integrate it into the final outcome.