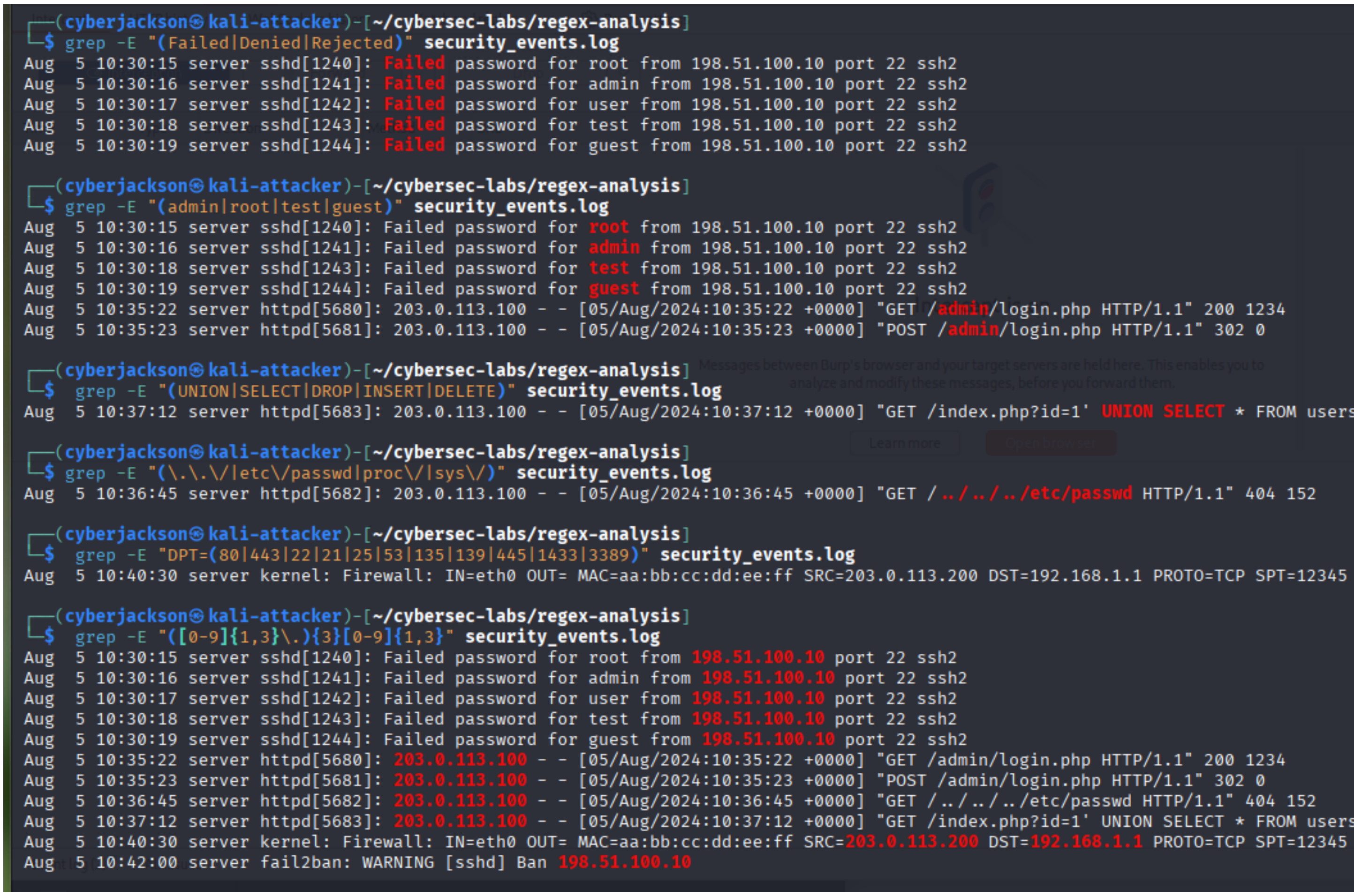
Introduction:

In this lab I will continue my dive into Linux Basics by applying regular exression (regex) to log analysis. Regex is a powerful pattern-matching tool that is essential in the cybersecurity profession. Use cases such as detection of intrusion attempts, analyzing malware behavior, and extracting indicators of compromise (IoCs) from log files are routinely tackled by security analysts with regex. This lab will focus on useful regex applications in real-world log analysis scenarios.

To wrap my head around regex in this lab I will construct regex for pattern matching, combine regex with grep for advanced log analysis, identify security-relevant patterns in log files, then extract those patterns, and finally combine this all to apply regex to real-world incident response situations.

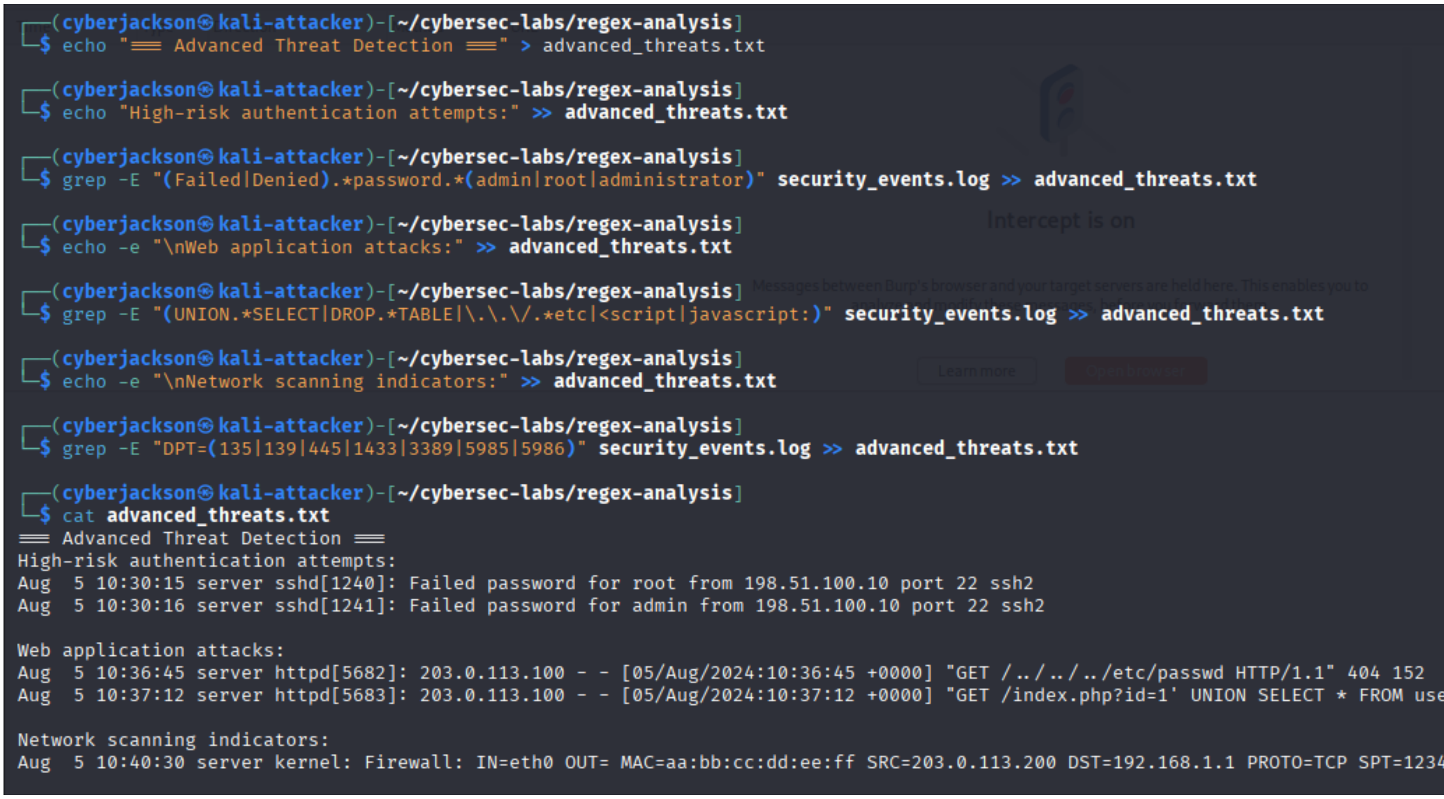
Exercise: Extended Regular Expressions

OR operator, multiple attack pattern detection, port range matching, and complex IP pattern matching



```
┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -E "(Failed|Denied|Rejected)" security_events.log
Aug  5 10:30:15 server sshd[1240]: Failed password for root from 198.51.100.10 port 22 ssh2
Aug  5 10:30:16 server sshd[1241]: Failed password for admin from 198.51.100.10 port 22 ssh2
Aug  5 10:30:17 server sshd[1242]: Failed password for user from 198.51.100.10 port 22 ssh2
Aug  5 10:30:18 server sshd[1243]: Failed password for test from 198.51.100.10 port 22 ssh2
Aug  5 10:30:19 server sshd[1244]: Failed password for guest from 198.51.100.10 port 22 ssh2

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -E "(admin|root|test|guest)" security_events.log
Aug  5 10:30:15 server sshd[1240]: Failed password for root from 198.51.100.10 port 22 ssh2
Aug  5 10:30:16 server sshd[1241]: Failed password for admin from 198.51.100.10 port 22 ssh2
Aug  5 10:30:18 server sshd[1243]: Failed password for test from 198.51.100.10 port 22 ssh2
Aug  5 10:30:19 server sshd[1244]: Failed password for guest from 198.51.100.10 port 22 ssh2
Aug  5 10:35:22 server httpd[5680]: 203.0.113.100 - - [05/Aug/2024:10:35:22 +0000] "GET /admin/login.php HTTP/1.1" 200 1234
Aug  5 10:35:23 server httpd[5681]: 203.0.113.100 - - [05/Aug/2024:10:35:23 +0000] "POST /admin/login.php HTTP/1.1" 302 0

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$  grep -E "(UNION|SELECT|DROP|INSERT|DELETE)" security_events.log
Aug  5 10:37:12 server httpd[5683]: 203.0.113.100 - - [05/Aug/2024:10:37:12 +0000] "GET /index.php?id=1' UNION SELECT * FROM users

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -E "(\.\.\/|etc\/passwd|proc\/|sys\/)" security_events.log
Aug  5 10:36:45 server httpd[5682]: 203.0.113.100 - - [05/Aug/2024:10:36:45 +0000] "GET /../../../etc/passwd HTTP/1.1" 404 152

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$  grep -E "DPT=(80|443|22|21|25|53|135|139|445|1433|3389)" security_events.log
Aug  5 10:40:30 server kernel: Firewall: IN=eth0 OUT= MAC=aa:bb:cc:dd:ee:ff SRC=203.0.113.200 DST=192.168.1.1 PROTO=TCP SPT=12345

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$  grep -E "([0-9]{1,3}\.){3}[0-9]{1,3}" security_events.log
Aug  5 10:30:15 server sshd[1240]: Failed password for root from 198.51.100.10 port 22 ssh2
Aug  5 10:30:16 server sshd[1241]: Failed password for admin from 198.51.100.10 port 22 ssh2
Aug  5 10:30:17 server sshd[1242]: Failed password for user from 198.51.100.10 port 22 ssh2
Aug  5 10:30:18 server sshd[1243]: Failed password for test from 198.51.100.10 port 22 ssh2
Aug  5 10:30:19 server sshd[1244]: Failed password for guest from 198.51.100.10 port 22 ssh2
Aug  5 10:35:22 server httpd[5680]: 203.0.113.100 - - [05/Aug/2024:10:35:22 +0000] "GET /admin/login.php HTTP/1.1" 200 1234
Aug  5 10:35:23 server httpd[5681]: 203.0.113.100 - - [05/Aug/2024:10:35:23 +0000] "POST /admin/login.php HTTP/1.1" 302 0
Aug  5 10:36:45 server httpd[5682]: 203.0.113.100 - - [05/Aug/2024:10:36:45 +0000] "GET /../../../etc/passwd HTTP/1.1" 404 152
Aug  5 10:37:12 server httpd[5683]: 203.0.113.100 - - [05/Aug/2024:10:37:12 +0000] "GET /index.php?id=1' UNION SELECT * FROM users
Aug  5 10:40:30 server kernel: Firewall: IN=eth0 OUT= MAC=aa:bb:cc:dd:ee:ff SRC=203.0.113.200 DST=192.168.1.1 PROTO=TCP SPT=12345
Aug  5 10:42:00 server fail2ban: WARNING [sshd] Ban 198.51.100.10
```

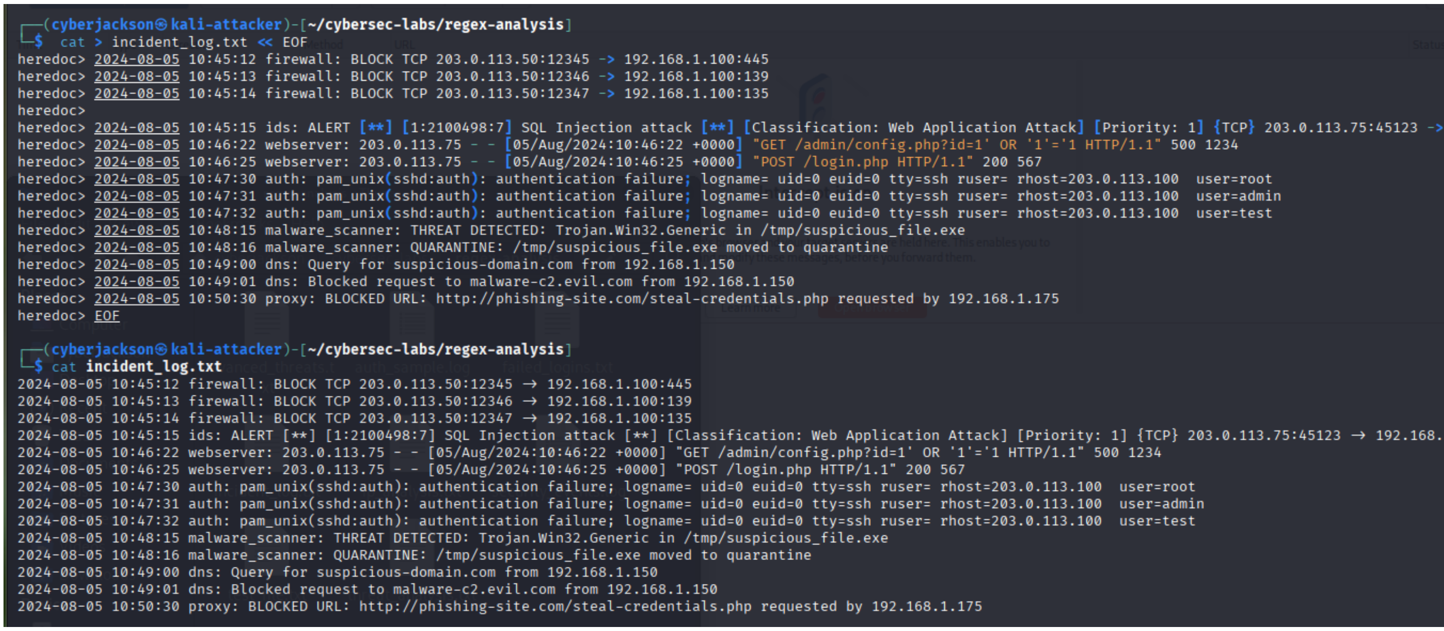
Advanced log parsing



```
┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ echo "=== Advanced Threat Detection ===" > advanced_threats.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ echo "High-risk authentication attempts:" >> advanced_threats.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -E "(Failed|Denied).*password.*(admin|root|administrator)" security_events.log >> advanced_threats.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ echo -e "\nWeb application attacks:" >> advanced_threats.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -E "(UNION.*SELECT|DROP.*TABLE|\.\.\/.*etc|<script|javascript:)" security_events.log >> advanced_threats.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ echo -e "\nNetwork scanning indicators:" >> advanced_threats.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -E "DPT=(135|139|445|1433|3389|5985|5986)" security_events.log >> advanced_threats.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ cat advanced_threats.txt
=== Advanced Threat Detection ===
High-risk authentication attempts:
Aug  5 10:30:15 server sshd[1240]: Failed password for root from 198.51.100.10 port 22 ssh2
Aug  5 10:30:16 server sshd[1241]: Failed password for admin from 198.51.100.10 port 22 ssh2

Web application attacks:
Aug  5 10:36:45 server httpd[5682]: 203.0.113.100 - - [05/Aug/2024:10:36:45 +0000] "GET /../../../etc/passwd HTTP/1.1" 404 152
Aug  5 10:37:12 server httpd[5683]: 203.0.113.100 - - [05/Aug/2024:10:37:12 +0000] "GET /index.php?id=1' UNION SELECT * FROM use

Network scanning indicators:
Aug  5 10:40:30 server kernel: Firewall: IN=eth0 OUT= MAC=aa:bb:cc:dd:ee:ff SRC=203.0.113.200 DST=192.168.1.1 PROTO=TCP SPT=1234
```

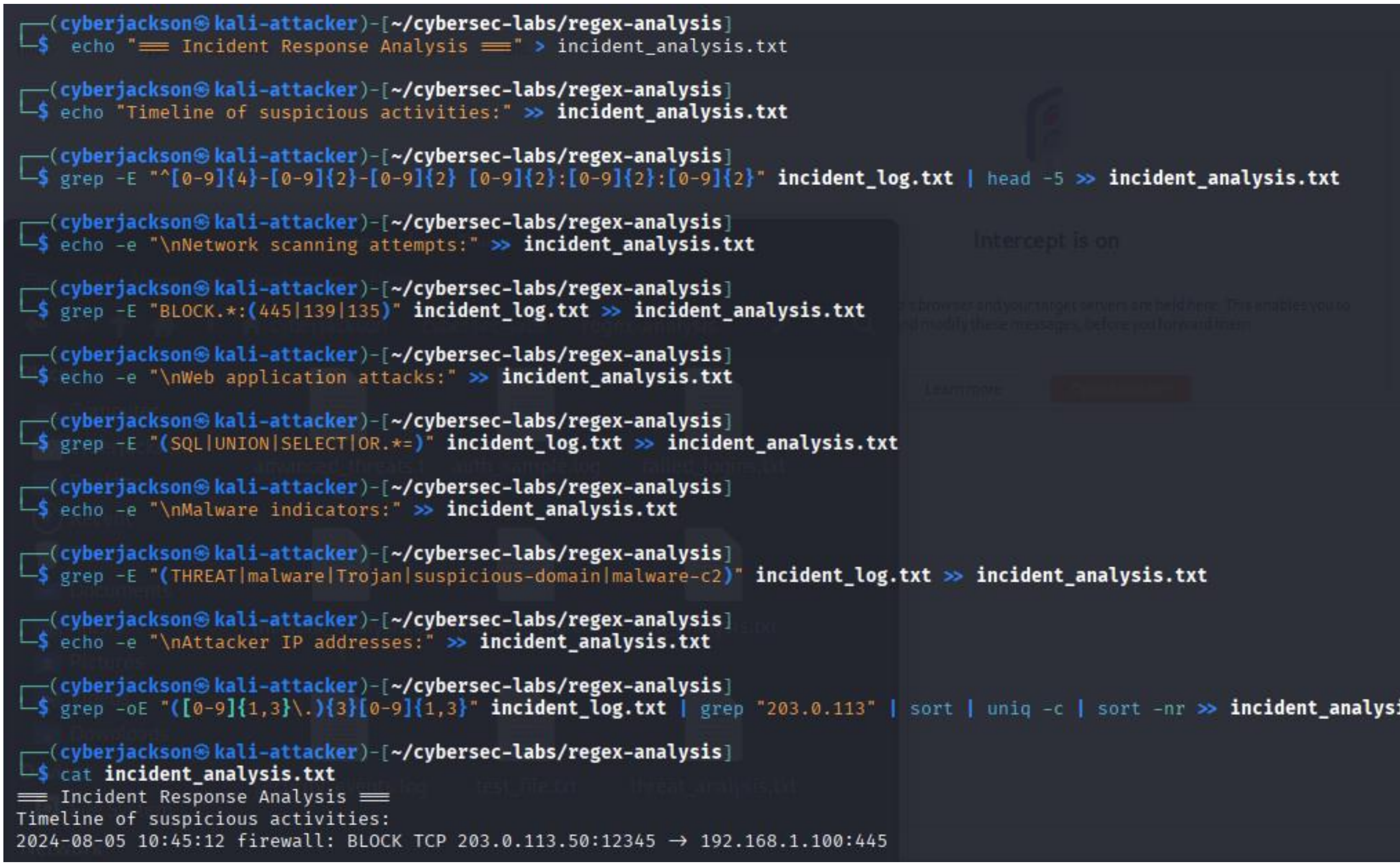
Exercise: Real-World Log Analysis Scenarios

Create incident response log



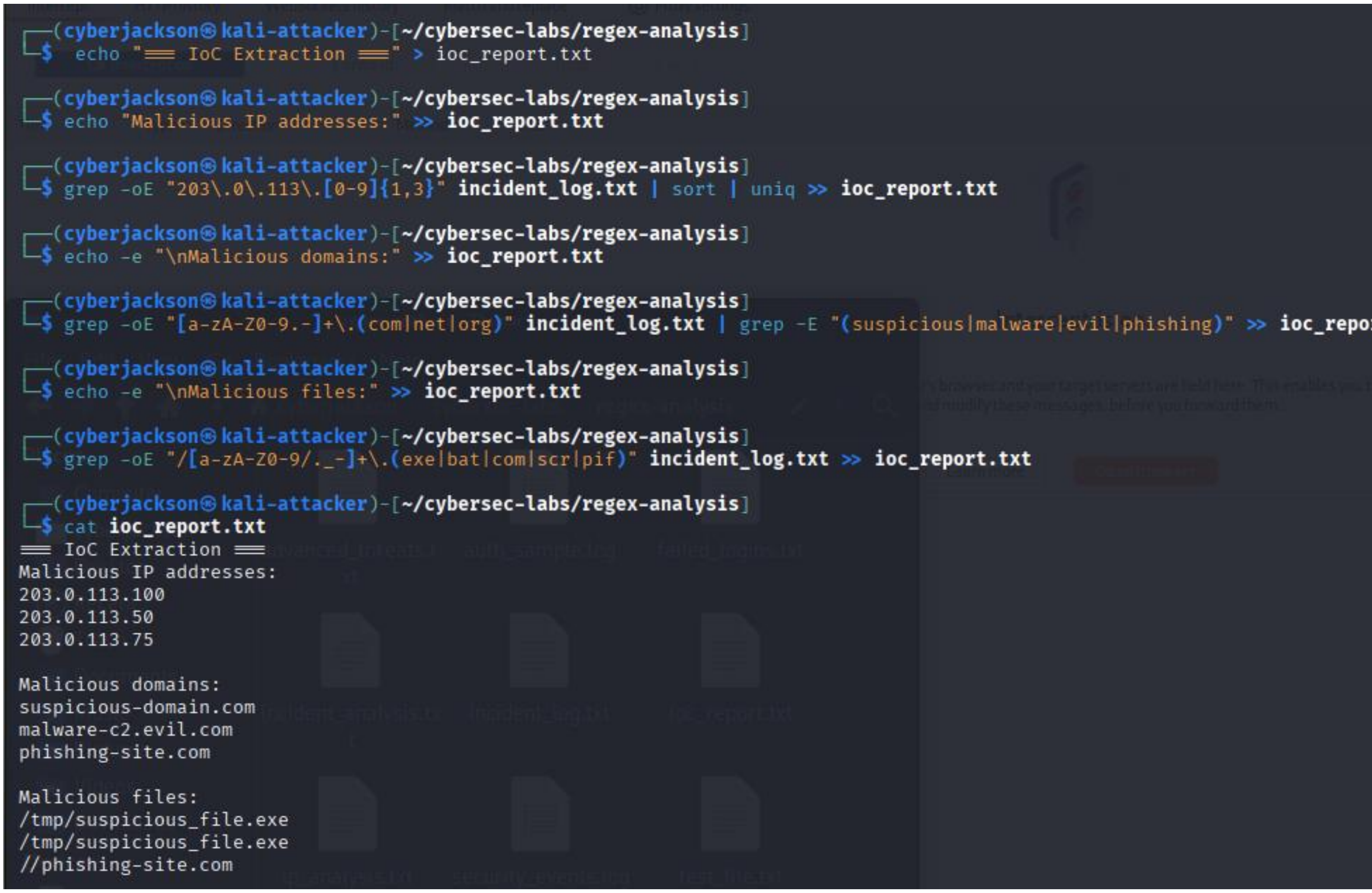
```
┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ cat > incident_log.txt << EOF
heredoc> 2024-08-05 10:45:12 firewall: BLOCK TCP 203.0.113.50:12345 -> 192.168.1.100:445
heredoc> 2024-08-05 10:45:13 firewall: BLOCK TCP 203.0.113.50:12346 -> 192.168.1.100:139
heredoc> 2024-08-05 10:45:14 firewall: BLOCK TCP 203.0.113.50:12347 -> 192.168.1.100:135
heredoc>
heredoc> 2024-08-05 10:45:15 ids: ALERT [**] [1:2100498:7] SQL Injection attack [**] [Classification: Web Application Attack] [Priority: 1] {TCP} 203.0.113.75:45123 ->
heredoc> 2024-08-05 10:46:22 webserver: 203.0.113.75 - - [05/Aug/2024:10:46:22 +0000] "GET /admin/config.php?id=1' OR '1'='1 HTTP/1.1" 500 1234
heredoc> 2024-08-05 10:46:25 webserver: 203.0.113.75 - - [05/Aug/2024:10:46:25 +0000] "POST /login.php HTTP/1.1" 200 567
heredoc> 2024-08-05 10:47:30 auth: pam_unix(sshd:auth): authentication failure; logname= uid=0 euid=0 tty=ssh ruser= rhost=203.0.113.100  user=root
heredoc> 2024-08-05 10:47:31 auth: pam_unix(sshd:auth): authentication failure; logname= uid=0 euid=0 tty=ssh ruser= rhost=203.0.113.100  user=admin
heredoc> 2024-08-05 10:47:32 auth: pam_unix(sshd:auth): authentication failure; logname= uid=0 euid=0 tty=ssh ruser= rhost=203.0.113.100  user=test
heredoc> 2024-08-05 10:48:15 malware_scanner: THREAT DETECTED: Trojan.Win32.Generic in /tmp/suspicious_file.exe
heredoc> 2024-08-05 10:48:16 malware_scanner: QUARANTINE: /tmp/suspicious_file.exe moved to quarantine
heredoc> 2024-08-05 10:49:00 dns: Query for suspicious-domain.com from 192.168.1.150
heredoc> 2024-08-05 10:49:01 dns: Blocked request to malware-c2.evil.com from 192.168.1.150
heredoc> 2024-08-05 10:50:30 proxy: BLOCKED URL: http://phishing-site.com/steal-credentials.php requested by 192.168.1.175
heredoc> EOF

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ cat incident_log.txt
2024-08-05 10:45:12 firewall: BLOCK TCP 203.0.113.50:12345 → 192.168.1.100:445
2024-08-05 10:45:13 firewall: BLOCK TCP 203.0.113.50:12346 → 192.168.1.100:139
2024-08-05 10:45:14 firewall: BLOCK TCP 203.0.113.50:12347 → 192.168.1.100:135
2024-08-05 10:45:15 ids: ALERT [**] [1:2100498:7] SQL Injection attack [**] [Classification: Web Application Attack] [Priority: 1] {TCP} 203.0.113.75:45123 → 192.168.
2024-08-05 10:46:22 webserver: 203.0.113.75 - - [05/Aug/2024:10:46:22 +0000] "GET /admin/config.php?id=1' OR '1'='1 HTTP/1.1" 500 1234
2024-08-05 10:46:25 webserver: 203.0.113.75 - - [05/Aug/2024:10:46:25 +0000] "POST /login.php HTTP/1.1" 200 567
2024-08-05 10:47:30 auth: pam_unix(sshd:auth): authentication failure; logname= uid=0 euid=0 tty=ssh ruser= rhost=203.0.113.100  user=root
2024-08-05 10:47:31 auth: pam_unix(sshd:auth): authentication failure; logname= uid=0 euid=0 tty=ssh ruser= rhost=203.0.113.100  user=admin
2024-08-05 10:47:32 auth: pam_unix(sshd:auth): authentication failure; logname= uid=0 euid=0 tty=ssh ruser= rhost=203.0.113.100  user=test
2024-08-05 10:48:15 malware_scanner: THREAT DETECTED: Trojan.Win32.Generic in /tmp/suspicious_file.exe
2024-08-05 10:48:16 malware_scanner: QUARANTINE: /tmp/suspicious_file.exe moved to quarantine
2024-08-05 10:49:00 dns: Query for suspicious-domain.com from 192.168.1.150
2024-08-05 10:49:01 dns: Blocked request to malware-c2.evil.com from 192.168.1.150
2024-08-05 10:50:30 proxy: BLOCKED URL: http://phishing-site.com/steal-credentials.php requested by 192.168.1.175
```

Complete Incident Analysis



```
┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$  echo "=== Incident Response Analysis ===" > incident_analysis.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ echo "Timeline of suspicious activities:" >> incident_analysis.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -E "^[0-9]{4}-[0-9]{2}-[0-9]{2} [0-9]{2}:[0-9]{2}:[0-9]{2}" incident_log.txt | head -5 >> incident_analysis.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ echo -e "\nNetwork scanning attempts:" >> incident_analysis.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -E "BLOCK.*:(445|139|135)" incident_log.txt >> incident_analysis.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ echo -e "\nWeb application attacks:" >> incident_analysis.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -E "(SQL|UNION|SELECT|OR.*=)" incident_log.txt >> incident_analysis.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ echo -e "\nMalware indicators:" >> incident_analysis.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -E "(THREAT|malware|Trojan|suspicious-domain|malware-c2)" incident_log.txt >> incident_analysis.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ echo -e "\nAttacker IP addresses:" >> incident_analysis.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -oE "([0-9]{1,3}\.){3}[0-9]{1,3}" incident_log.txt | grep "203.0.113" | sort | uniq -c | sort -nr >> incident_analysi

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ cat incident_analysis.txt
=== Incident Response Analysis ===
Timeline of suspicious activities:
2024-08-05 10:45:12 firewall: BLOCK TCP 203.0.113.50:12345 → 192.168.1.100:445
```

Create IoC (Indicators of Compromise) extraction



```
┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$  echo "=== IoC Extraction ===" > ioc_report.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ echo "Malicious IP addresses:" >> ioc_report.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -oE "203\.0\.113\.[0-9]{1,3}" incident_log.txt | sort | uniq >> ioc_report.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ echo -e "\nMalicious domains:" >> ioc_report.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -oE "[a-zA-Z0-9.-]+\.(com|net|org)" incident_log.txt | grep -E "(suspicious|malware|evil|phishing)" >> ioc_repo

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ echo -e "\nMalicious files:" >> ioc_report.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ grep -oE "/[a-zA-Z0-9/._-]+\.(exe|bat|com|scr|pif)" incident_log.txt >> ioc_report.txt

┌──(cyberjackson㉿kali-attacker)-[~/cybersec-labs/regex-analysis]
└─$ cat ioc_report.txt
=== IoC Extraction ===
Malicious IP addresses:
203.0.113.100
203.0.113.50
203.0.113.75

Malicious domains:
suspicious-domain.com
malware-c2.evil.com
phishing-site.com

Malicious files:
/tmp/suspicious_file.exe
/tmp/suspicious_file.exe
//phishing-site.com
```

Conclusion: It is clear that regex is essential in cybersecurity and very powerful for analyzing security logs, forensic data, and other large datasets. The most interesting aspect of this lab and regex for me is how extended regular expressions allow us to do even more by combining expressions and using OR logic. These extended expressions which are built from the basic syntax of regex allow for much more specific yet complex parsing of logs.

The most difficult part of regex for me both before and after this lab is grasping the wide range of flags and delimiters and their accompanying syntax. More practice and work with regex will suffice in helping me here. Also, websites like regex101 allow me to play with and learn more about regex.