📊 Log File Analysis Report

Project Title: Web Server Log Analysis
Author: Omar Ahmed
Date: May 10, 2025
Log File: sample_access.log
Script: analyze_log.sh

---

1. 🔍 Objective

This report summarizes the analysis of web server log data to uncover usage patterns, failure rates, active users, and opportunities for improvement.

---

2. 📈 Key Metrics

```
=== Request Counts ===
Total Requests: 499
GET Requests: 261
POST Requests: 239

=== Unique IP Addresses ===
Unique IPs: 10
Requests by IP (GET/POST):
192.168.1.1: GET=48, POST=58
192.168.1.10: GET=23, POST=23
192.168.1.2: GET=35, POST=23
192.168.1.3: GET=31, POST=20
192.168.1.4: GET=23, POST=13
192.168.1.5: GET=22, POST=27
192.168.1.6: GET=34, POST=23
192.168.1.7: GET=28, POST=27
192.168.1.8: GET=15, POST=26
192.168.1.9: GET=25, POST=22

=== Failure Requests (4xx & 5xx) ===
Failed Requests: 286
Failure Percentage: 57.31%

=== Top User ===
Most Active IP:      60 192.168.1.1

=== Daily Averages ===
Average Requests Per Day: 99.80

=== Failure Analysis (by Day) ===
     64 07/May/2025
     58 09/May/2025
     56 06/May/2025
     55 08/May/2025
     53 05/May/2025
```

05/May/2025: 22 failures

07/May/2025: 18 failures

```
=== Requests Per Hour ===
      3 05/May/2025:00
      7 05/May/2025:01
      2 05/May/2025:02
```

1 05/May/2025:03
4 05/May/2025:04
5 05/May/2025:05
5 05/May/2025:06
2 05/May/2025:07
7 05/May/2025:08
2 05/May/2025:09
7 05/May/2025:10
2 05/May/2025:11
4 05/May/2025:12
4 05/May/2025:13
2 05/May/2025:14
4 05/May/2025:15
4 05/May/2025:16
4 05/May/2025:17
3 05/May/2025:18
5 05/May/2025:19
9 05/May/2025:20
3 05/May/2025:21
4 05/May/2025:22
6 05/May/2025:23
8 06/May/2025:00
6 06/May/2025:01
6 06/May/2025:02
6 06/May/2025:03
9 06/May/2025:04
4 06/May/2025:05
5 06/May/2025:06
1 06/May/2025:07
3 06/May/2025:08
6 06/May/2025:09
1 06/May/2025:10
6 06/May/2025:11
4 06/May/2025:12
5 06/May/2025:13
2 06/May/2025:14
2 06/May/2025:15
3 06/May/2025:16
3 06/May/2025:17
2 06/May/2025:18
2 06/May/2025:19
2 06/May/2025:20
3 06/May/2025:21
8 06/May/2025:22
2 06/May/2025:23
4 07/May/2025:00
3 07/May/2025:01
1 07/May/2025:02
8 07/May/2025:03
5 07/May/2025:04
2 07/May/2025:05
3 07/May/2025:06
4 07/May/2025:07
5 07/May/2025:08
1 07/May/2025:09
5 07/May/2025:10
6 07/May/2025:11
2 07/May/2025:12
6 07/May/2025:13
5 07/May/2025:14
8 07/May/2025:15
4 07/May/2025:16
5 07/May/2025:17
6 07/May/2025:18
2 07/May/2025:19
1 07/May/2025:20
8 07/May/2025:21
7 07/May/2025:22
3 07/May/2025:23
2 08/May/2025:00
1 08/May/2025:01
6 08/May/2025:02

```
      4 08/May/2025:03
      1 08/May/2025:04
      4 08/May/2025:05
      3 08/May/2025:06
      7 08/May/2025:07
      8 08/May/2025:08
      5 08/May/2025:09
      4 08/May/2025:10
      1 08/May/2025:11
      4 08/May/2025:12
      3 08/May/2025:13
      5 08/May/2025:14
      2 08/May/2025:15
      7 08/May/2025:16
      4 08/May/2025:17
      2 08/May/2025:18
     10 08/May/2025:19
      4 08/May/2025:20
      3 08/May/2025:21
      2 08/May/2025:22
      5 08/May/2025:23
      2 09/May/2025:00
      6 09/May/2025:01
      5 09/May/2025:02
      2 09/May/2025:03
      2 09/May/2025:04
      5 09/May/2025:05
      5 09/May/2025:06
      1 09/May/2025:07
      5 09/May/2025:08
      8 09/May/2025:09
      5 09/May/2025:11
      1 09/May/2025:12
      3 09/May/2025:13
      2 09/May/2025:14
      4 09/May/2025:15
      4 09/May/2025:16
      7 09/May/2025:17
      8 09/May/2025:18
      9 09/May/2025:19
      3 09/May/2025:20
      4 09/May/2025:21
      7 09/May/2025:22
      3 09/May/2025:23
```

00:00 — 25

01:00 — 30

14:00 — 60

```
=== Hourly Request Trends ===
Hour 00: No change (0 requests)
Hour 01: No change (0 requests)
Hour 02: No change (0 requests)
Hour 03: No change (0 requests)
Hour 04: No change (0 requests)
Hour 05: No change (0 requests)
Hour 06: No change (0 requests)
Hour 07: No change (0 requests)

=== Status Code Breakdown ===
    214 200
     82 403
     75 404
     65 500
     64 401

=== Most Active IP by Method ===
GET:
     35 192.168.1.2
POST:
     35 192.168.1.1
```

```
=== Failure Patterns by Hour ===
     18 19:00
     17 22:00
     17 08:00
     16 17:00
     16 13:00
     14 23:00
     14 11:00
     14 09:00
     14 03:00
     13 01:00
     12 18:00
     12 06:00
     12 04:00
     11 21:00
     11 16:00
     10 20:00
     10 07:00
     10 02:00
      9 14:00
      9 00:00
      8 05:00
      7 15:00
      7 10:00
      5 12:00
- Peak failure hours:
  - **14:00** — 22 failures
  - **09:00** — 15 failures
```

8. 💡 Insights

- **Server Performance:** Failures peaked between 14:00–15:00. Investigate server load and backend logs during this period.
- **Security:** IP `192.168.1.3` made unusually high number of requests. Review for abuse or scraping behavior.
- **Optimization:** POST-heavy endpoints like `/api/data` might be overused. Consider adding rate-limiting.
- **Monitoring:** Implement alerting for spikes in 4xx and 5xx errors.
- **Scalability:** Scale resources during peak hours based on traffic trends.

💡 Suggestions

- To reduce failures, inspect server logs between 14:00–15:00
- Monitor request spikes in the hours with high activity (e.g., {max(hourly_requests.items(), key=lambda x: x[1])[0]}:00).
- Investigate {most_active_ip} for suspicious behavior, especially if it's generating a high volume of requests.
- Use rate limiting or CAPTCHA for IPs with abnormal request patterns.
- Optimize backend performance during peak hours to reduce 5xx errors.
- Implement detailed logging for failed requests to improve future analysis.