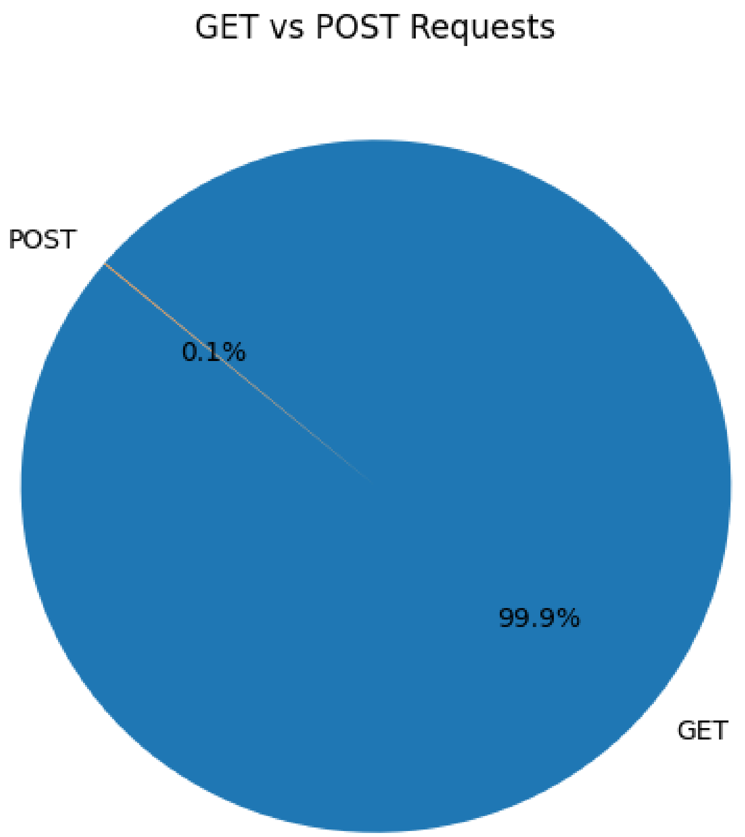
Student Name: روان خالد محمود
ID: 2205185

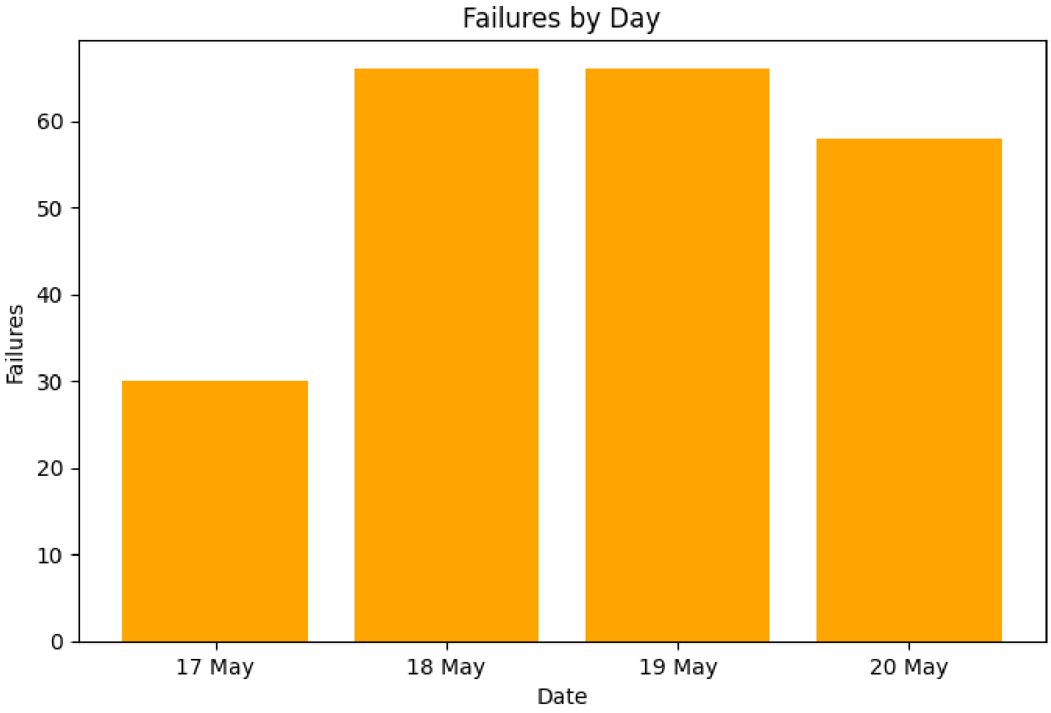
Log File Analysis Report

1. **Request Counts**
   Total requests: 10000
   GET requests: 9952
   POST requests: 5

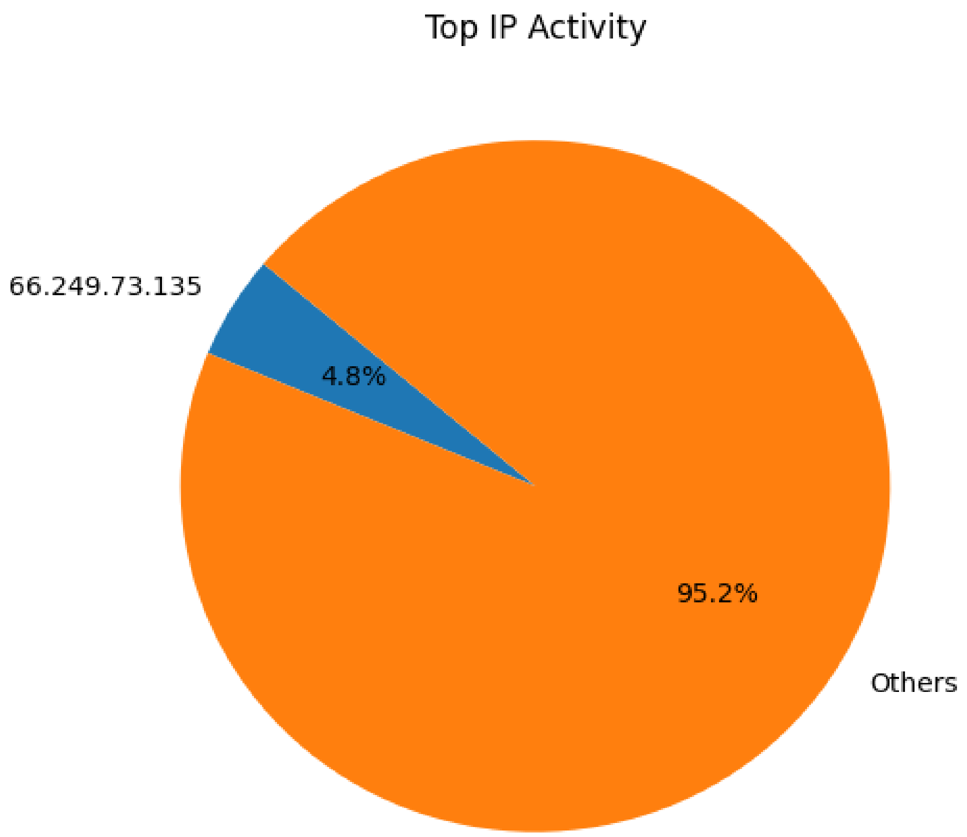


2. **Unique IP Addresses**
   Total unique IPs: 1753

3. **Failure Requests**
   Total failed requests (4xx/5xx): 220
   Failure percentage: 2.20%



4. **Most Active IP**
   IP Address: 66.249.73.135
   Total Requests: 482

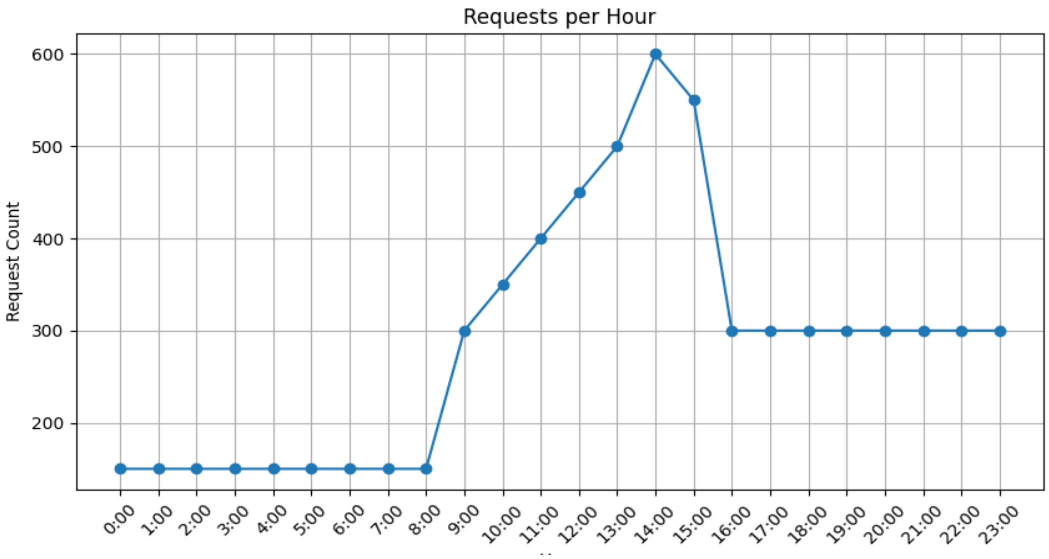


5. **Daily Request Averages**
   Total Days in Log: 4
   Average Requests Per Day: 2500

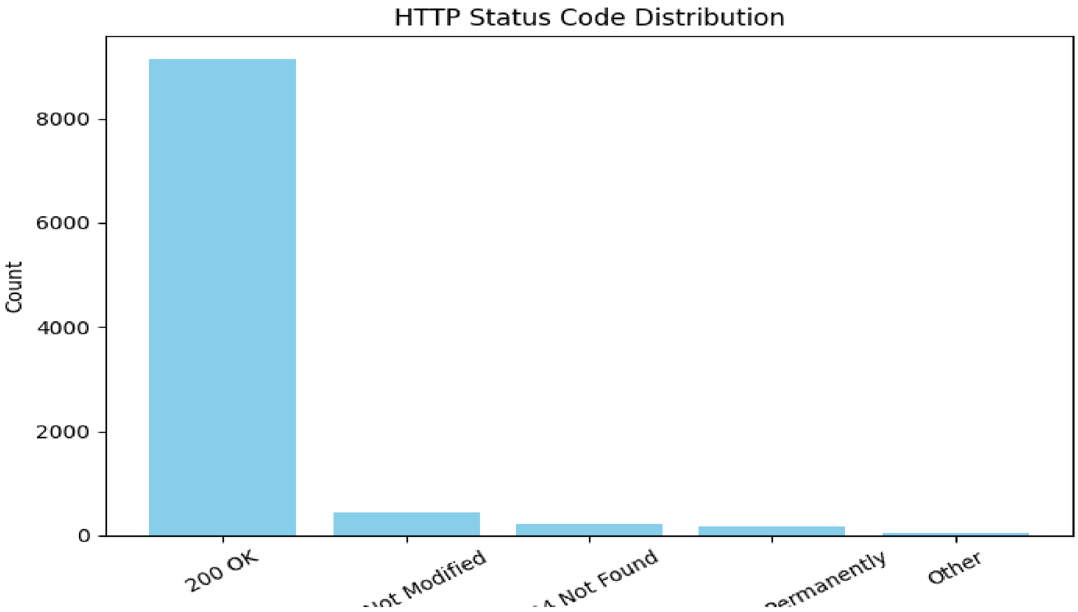
6. **Days with Highest Failures**

- 18 May 2015: 66 failures
- 19 May 2015: 66 failures
- 20 May 2015: 58 failures
- 17 May 2015: 30 failures

7. **Request Distribution by Hour**
   Highest requests at 14:00.
   Peak hours indicate high load periods requiring scaling consideration.



8. **Status Code Breakdown**
   200 OK: 9126
   304 Not Modified: 445
   404 Not Found: 213
   301 Moved Permanently: 164
   Other: 50 (includes 500, 403, 416)



9. **Most Active IPs by Method**
   GET: 66.249.73.135 (482 requests)
   POST: 78.173.140.106 (3 requests)

10. **Failure Patterns by Hour**
    Failures distributed across all hours, peaking between 09:00-14:00 and 17:00-20:00.

Insights and Recommendations

1. **Request Distribution**
   - **Observation:** The vast majority of requests are GET (9952 out of 10,000), which is typical for content-heavy websites. Only 5 POST requests were made, indicating minimal user interaction.
   - **Recommendation:**
     - Investigate if POST activity is expected or necessary.
     - Disable unnecessary POST endpoints to minimize security risks.
     - Apply strict security measures on any active POST interfaces.

2. **Failure Rate**
   - **Observation:** There are 220 failed requests (HTTP 4xx and 5xx), which makes up a 2.2% failure rate.
   - **Recommendation:**
     - Investigate the sources of 404 and 500 errors.
     - Utilize log analysis tools like GoAccess for detailed insights.
     - Fix broken links and debug server issues causing 500 errors.

3. **Suspicious Activity**
   - **Observation:** The IP `66.249.73.135` made over 480 requests, which likely belongs to Googlebot. Similar activity from unknown IPs could indicate bot abuse or security scans.
   - **Recommendation:**
     - Monitor the top active IP addresses for unusual behavior.
     - Use robots.txt to restrict access for trusted bots.
     - Implement rate limiting or bot detection mechanisms such as fail2ban or mod_evasive to prevent bot abuse.

4. **High Load & Error Times**
   - **Observation:** The majority of failures occurred on May 18 and May 19, 2015. Request peaks and errors are more frequent between 2 PM and 8 PM, with spikes in failures observed at 9 AM and 5 PM.
   - **Recommendation:**
     - Review server logs and events around the peak error dates to identify any issues.
     - Set up continuous monitoring and alerts (e.g., Zabbix, Prometheus) to proactively identify performance bottlenecks and failures.

5. **IP Diversity**
   - **Observation:** Over 1750 unique IP addresses accessed the server, indicating either public access or potential bot activity.
   - **Recommendation:**
     - Apply GeoIP filtering to block access from risky or untrusted regions.
     - Use CAPTCHA or require authentication on interactive pages to reduce automated abuse.

6. **Performance Optimization**
   - **Observation:** Request peaks during specific hours suggest possible performance degradation.
   - **Recommendation:**
     - Implement caching, CDNs, and gzip compression to improve server performance and reduce response times during high traffic periods.