**Log Analysis**

**1.0**

# Contents

[1. Contents 1](#_Toc383466729)

[2. Change History 2](#_Toc383466730)

[3. Object 2](#_Toc383466731)

[4. Purpose 2](#_Toc383466732)

[5. Non-Functional Requirements 2](#_Toc383466733)

# Change History

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Date** | **Change Purpose** | **Changed By** | **Version** | **Contact** |
| 24-Mar-14 | Initial Version | Murthy | 1.0 | Ksn.moorthy@gmail.com |
|  |  |  |  |  |

# Object

This document outlines the problem statement of the Log Analysis project.

# Purpose

The following data will be pulled out from the logs.

* Total hits for a particular day.
* Total hits from a particular IP address
* IP address wise total hits.
* Which IP address has made more number of hits?
* Which day has made more number of hits?
* Max number of hits IP address of
* Browser wise total hits.
* Date wise total products sold.

# Non-Functional Requirements

* Flume – Data Ingestion
* Oozie – Work flow
* Map Reduce – Pull the required data and convert into desired format
* Hive – Load the converted data into Hive tables and perform queries
* HBase – Maintain the data that will be available to dashboards
* A Web technology – Display the analyzed data

Many more details yet to be updated.