

Analytics Requirement

Required As Data Frame

- For each month, how many times each individual host has connected to our server? Store data month wise & highest count first.
- For each month, how many times each individual page has been requested from our server? Store data month wise & highest count first.
- For each month, how much data has been downloaded by each individual host that has connected to our server? Store data month wise & by highest count first.
- How much data was sent out as each individual page was downloaded from our server? Store data month wise & by highest count first.

Required As Visualization

- For each data set generated above, prepare suitable visualization (giving reason why the graph is chosen). Limit data to suitable significant number if graph is looking too cluttered.
- Time Series Graph for total hits per day with each month being shown as separate line.
- Time Series Graph for total download size per day with each month being shown as separate line.
- How would you show top “10 most popular pages” per day as Time Series Graph with each month being shown as separate line. Explain how you find “top 10 most popular pages”

Answers Required

Using the above results and also carrying out any other analysis as may be required, provide answers to the following questions.

- Which host has connected the maximum number of times to our server? Give the host name & count of connections from that host.
- Which page that has been requested the maximum number of times from our server? Give the page name & count of the times the page was requested.
- How many unique hosts have connected to our server? Give counts.
- How many unique pages have been requested from our server? Give counts.
- Which host has caused maximum data transfer from our server? Give host name & the data transfer for the host.
- Which page has caused maximum data transfer from our server? Give page name & the data transfer for the page.
- Which page has maximum download size from our server? Give page name & the size for the page.

- What is the download count of the page that has maximum download size from our server?
Give page name & download count
- Which page has minimum download size from our server? Give page name & the size for the page.
- What is the download count of the page that minimum download size from our server? Give page name & the size for the page.