**PSG COLLEGE OF TECHNOLOGY**

**DEPARTMENT OF APPLIED MATHEMATICS AND COMPUTATIONAL SCIENCES**

**M.Sc (DS) – WEB ANALYTICS LAB**

**PROBLEM SHEET – I**

1. Analyze the Web Server Log File and print the following
2. **List the Frequent Hosts**
3. **List the number of Unique Hosts**

#### **List the visit length of the user in a particular web page**

#### **List the demographic and system statistics**

#### **List the visitor path**

#### **List the top entry page, top exit page and most popular pages**

#### **Exploring 404 Response Codes**

1. **List 404 Response Code Records**
2. **Count 404 Response Codes**
3. **List the Top Twenty 404 Response Code Endpoints**
4. **List 404 Response Codes per Day**
5. **Write a program to record the user behavior in a web site**
6. **Write a program to collect data from a web site using page tagging.**
7. **Write a program to gather metrics using a log file.**