

FACULDADE DE TECNOLOGIA SENAC DE GOIÁS
CURSO GESTÃO DA TECNOLOGIA DA INFORMAÇÃO
PROGRAMAÇÃO ORIENTADA A OBJETOS

JOÃO NETTO NATAL PINHEIRO
JOÃO PAULO NASCIMENTO OLIVEIRA
PAULO ROBERTO VIEIRA
RUBEM DE OLIVEIRA VIEIRA

Analizador de Logs do Servidor Web

Goiânia
2018

JOÃO NETTO NATAL PINHEIRO
JOÃO PAULO NASCIMENTO OLIVEIRA
PAULO ROBERTO VIEIRA
RUBEM DE OLIVEIRA VIEIRA

Analizador de Logs do Servidor Web

Trabalho apresentado como requisito parcial de nota da disciplina Programação Orientada a Objetos do curso Gestão da Tecnologia da Informação da Faculdade de Tecnologia SENAC.

Orientação: Alice Mota Faleiro

A proposta solicitada no módulo Programação Orientada a Objetos se relaciona diretamente com dois outros módulos (Estatística Aplicada e Laboratório de Redes de Computadores), pede-se que faça implementação do analisador de logs do servidor web e cálculos estatísticos, em outras palavras o software desenvolvido deve ler um arquivo de log, extrair os dados e fazer os seguintes cálculos: Média, Moda, Mediana, Desvio Padrão e Frequência, aplicados por hora do acesso, endereço IP e browser.

- Recuperar o conteúdo do arquivo de log de acesso do servidor httpd Apache;
- Definir o caractere separador de campo (ex.: - ou | ou ,);
- Criar uma classe que represente os campos recuperados do arquivo de log;
- Desenvolva a lógica para fazer a separação dos campos necessários;
- Desenvolver a lógica para fazer os cálculos estatísticos solicitados;
- Desenvolva a classe Arquivo, na qual deve existir dois métodos estáticos um responsável pela leitura dos dados do arquivo e que retorna o ArrayList preenchido e outro que recebe o ArrayList preenchido e grava no arquivo.
- Desenvolver a classe java Executável, utilizando Interface Gráfica, que manipule a execução do programa que com o ArrayList preenchido efetue os cálculos estatísticos, descritos na unidade e Estatística.

A seguir apresentamos os prints de tela da execução do software. Será disponibilizado no site do projeto integrador de cada integrante do grupo o arquivo .jar para download e também o arquivo log já tratado, caso queira verificar o software em execução.

A primeira tela esboça as funções do software, exceto os gráficos. No canto superior esquerdo está a opção de **Escolher o Arquivo** log, após selecionado o arquivo, deve-se clicar em **Analisar**, enquanto a análise ocorre, é aberto uma janela que apresenta a porcentagem de conclusão, com a análise finalizada o resultado do processamento do arquivo é preenchido (Quantidade de linhas analisadas e o tempo de análise).

Análise de Log

Resultado do processamento do arquivo

Arquivo: Escolher Arquivo
 Analisar

Quantidade de linhas analisadas: -
 Tempo de análise: - s

Tabela de dados brutos Gerar Estatística

Info Gráficos

Acessos por IP

Média: -
 Moda: -
 Mediana: -
 Desvio Padrão: -
 Variância: -

Tabela Cálculos

Acessos por Browser

Browser mais utilizado: -----

Tabela Frequência

Acesso por Hora

0 -- 1	1 -- 2	2 -- 3	3 -- 4	4 -- 5	5 -- 6	6 -- 7	7 -- 8	8 -- 9	9 -- 10	10 -- 11
0	0	0	0	0	0	0	0	0	0	0

Média: -
 Classe Modal: -
 Classe Mediana: -
 Desvio Padrão: -
 Variância: -

Tabela ROL Tabela Cálculo

No segundo print poderá ser observado os campos acima citados preenchidos, o próximo passo é preencher as tabelas, isso se dará clicando em **Tabela de dados brutos**. Esta ação irá gerar uma nova janela com os dados (IP, Data/Hora, Browser) preenchidos, são os dados extraídos do arquivo de log.

Análise de Log

Resultado do processamento do arquivo

Arquivo: C:\Users\UsoaPaulo\Documents\output.log Escolher Arquivo
 Analisar

Quantidade de linhas analisadas: 788718
 Tempo de análise: 14.009 s

Tabela de dados brutos Gerar Estatística

Info Gráficos

Acessos por IP

Média: -
 Moda: -
 Mediana: -
 Desvio Padrão: -
 Variância: -

Tabela Cálculos

Acessos por Browser

Browser mais utilizado: -----

Tabela Frequência

Acesso por Hora

0 -- 1	1 -- 2	2 -- 3	3 -- 4	4 -- 5	5 -- 6	6 -- 7	7 -- 8	8 -- 9	9 -- 10	10 -- 11
0	0	0	0	0	0	0	0	0	0	0

Média: -
 Classe Modal: -
 Classe Mediana: -
 Desvio Padrão: -
 Variância: -

Tabela ROL Tabela Cálculo

Tabela de dados brutos

IP	Data	Browser
108.62.115.226	Thu Feb 22 11:36:49 B.	Bot
77.139.51.223	Thu Feb 22 11:37:02 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:03 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:07 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:08 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:08 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:10 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:11 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:12 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:39 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:41 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:41 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:41 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:41 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:42 B.	Chrome
77.139.51.223	Thu Feb 22 11:37:45 B.	Chrome
178.255.155.2	Thu Feb 22 11:41:49 B.	Bot
82.103.139.165	Thu Feb 22 11:46:49 B.	Bot
54.36.148.173	Thu Feb 22 11:48:51 B.	Bot
85.17.156.76	Thu Feb 22 11:51:49 B.	Bot
141.8.183.17	Thu Feb 22 11:52:44 B.	Bot

IP	Data	Browser
108.62.115.226	Thu Feb 22 11:36:49 B...	Bot
77.139.51.223	Thu Feb 22 11:37:02 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:03 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:07 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:08 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:08 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:08 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:10 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:11 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:12 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:39 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:41 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:41 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:41 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:41 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:41 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:41 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:41 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:42 B...	Chrome
77.139.51.223	Thu Feb 22 11:37:45 B...	Chrome
178.255.155.2	Thu Feb 22 11:41:49 B...	Bot
82.103.139.165	Thu Feb 22 11:46:49 B...	Bot
54.36.148.173	Thu Feb 22 11:48:51 B...	Bot
85.17.156.76	Thu Feb 22 11:51:49 B...	Bot
141.8.183.17	Thu Feb 22 11:52:44 B...	Bot

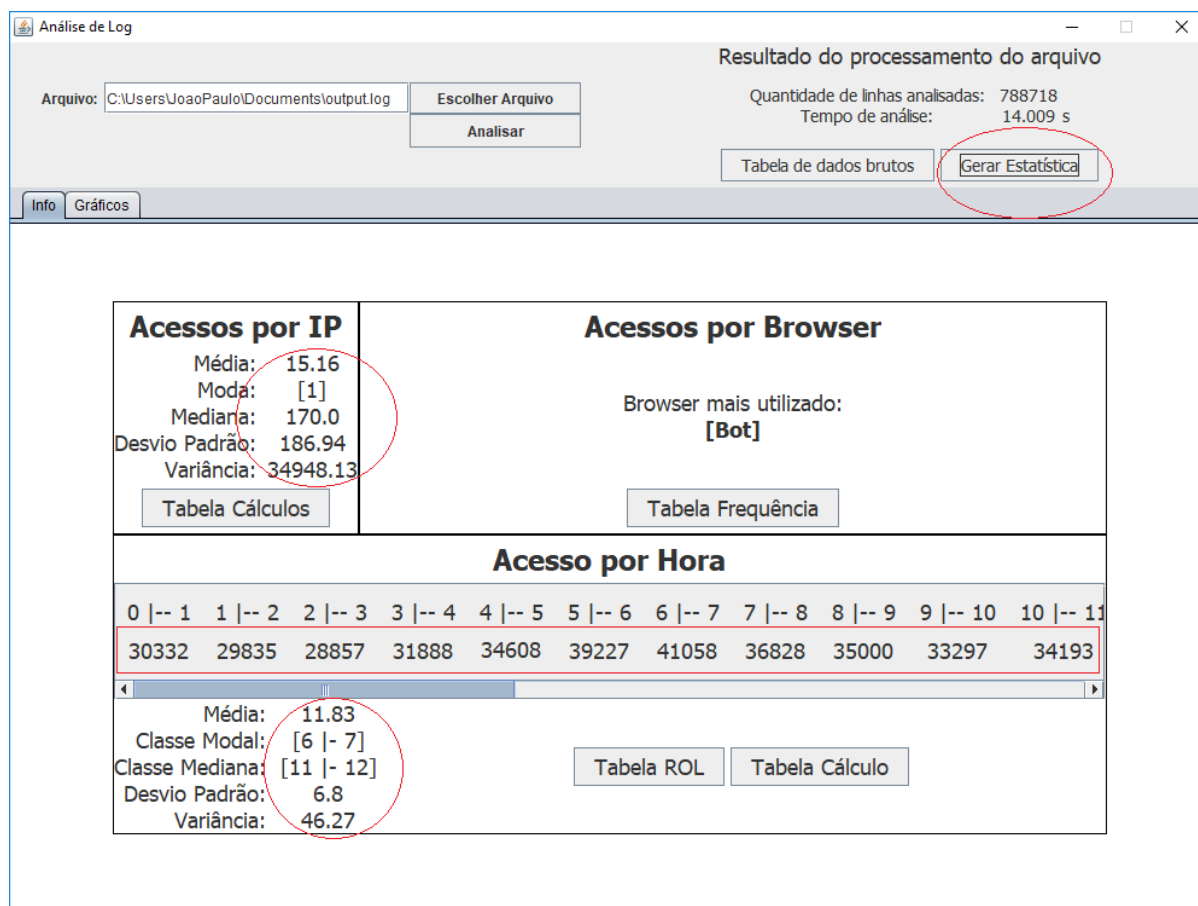
O passo seguinte é fazer os cálculos estatísticos, para isso basta clicar no botão **Gerar Estatística**, todos os cálculos relacionados aos dados extraído do arquivo de log e que estão na tabela de dados brutos serão executados e apresentados em cada painel que estão divididos pelo seu respectivo título.

No primeiro painel é apresentado os dados de **Acessos por IP**, a média de acesso, moda, mediana, desvio padrão e a variância, para uma melhor visualização, foi desenvolvido a opção de clicar em **Tabela de Cálculos**, esta ação irá gerar uma tabela dos dados extraídos do arquivo analisado.

No painel de **Acessos por Browser**, será apresentado o tipo de navegador que mais tem registro de acesso, caso clique em **Tabela de Frequência**, será apresentado a frequência de uso de cada Browser.

O painel de **Acesso por Hora**, é o retângulo que está na parte inferior da tela, apresenta os dados de acesso subdivididos por classes de hora de acesso, é calculado quantos acessos ocorreram por exemplo entre 6 e 7 da manhã (6 |-- 7), mas também mostra os cálculos de média de acesso, moda, mediana, desvio

padrão e a variância. Neste painel tem o botão da **Tabela Rol** que exibe de forma crescente a frequência de acesso por hora, o outro botão **Tabela de Cálculo** assim como as anteriores apresenta uma tabela com os dados de acesso por hora, separados em classes (hora/hora).



A outra aba que existe na tela do software e a de **Gráficos**, após selecionar e analisar o arquivo, gerar estatística, os gráficos são preenchidos automaticamente, esta aba também possui 3 painéis (Acessos por IP, Browser, Acessos por Hora), porém os cálculos são apresentados em gráficos, e não somente no formato alfanumérico, é uma análise visual bastante utilizada, ressaltamos que ao para o ponteiro do mouse sobre um dado ele apresenta a informação relativa, por exemplo: Se parar o ponteiro do mouse em alguma barra do gráfico de barras, será exibido qual a faixa de hora que se refere e também a quantidade de acesso que ocorreram no intervalo em questão, o scroll do mouse também interfere em expandir ou diminuir a imagem.

