



INSTITUTO FEDERAL DE  
EDUCAÇÃO, CIÊNCIA E TECNOLOGIA  
SÃO PAULO

CAMPUS BARRETOS

## 1- IDENTIFICAÇÃO

**CURSO:** Tecnologia em Análise e Desenvolvimento de Sistemas

**Componente Curricular:** Mineração de Dados Web com *scripts* BASH

**Ministrante:** Vinícius Silva Madureira  
Pereira

**Evento:** FLISOL 2018

**Tipo:** Minicurso

**Total de aulas:** 03

**Total de horas:** 2,5

**Abordagem Metodológica:**

T ( ) P (X) T/P ( )

**Uso de laboratório ou outros ambientes além da sala de aula?**

(x) SIM ( ) NÃO Qual(is)?

Práticas em laboratório para desenvolvimento de sistema de mineração de dados.

## 2 - EMENTA:

O minicurso abrange os conhecimentos sobre *scripts* de Shell, comandos e programas em *command-line*, permissões em arquivos e diretórios em ambientes GNU/Linux, linguagem nativa do BASH, expressões regulares, mineração de dados com processos ETL (Extrair, Transformar e Carregar) e descrição e manipulação de dados em bases de dados relacionais.

## 3 - OBJETIVOS

### GERAL:

Permitir aos alunos que conheçam e desenvolvam com eficiência e eficácia *scripts* de *shell* em BASH, estabelecendo proficiência na codificação de programas em *command-line*. Propiciar elementos fundamentais para a prática profissional em ambientes GNU/Linux, facilitando a formação e capacitando os alunos a compreender linguagens de programação interpretadas. Demonstrar meios de análise e identificação de padrões para a mineração de dados em âmbito Web.

### ESPECÍFICOS:

- Definir conceitos e contexto na utilização de *scripts* de Shell.
- Diferenciar comandos e programas em *command-line*.
- Descrever e praticar a manipulação do sistema permissões de arquivos e diretórios em ambientes GNU/Linux.
- Descrever e praticar a execução dos programas Wget e Curl.
- Conhecer, praticar e aplicar os padrões da linguagem de programação do BASH.
- Conceituar e aplicar a mineração de dados.
- Conhecer formas de averiguar requisições HTTP.
- Apresentar e aplicar Processos ETL (*Extract, Transform and Load*).
- Apresentar, definir e aplicar a descrição e manipulação de dados em bases de dados relacionais com o SQLite3.

## 4 - CONTEÚDO PROGRAMÁTICO:

- Conceitos e contextos de *scripts* de Shell.

- Bourne Again Shell (BASH).
- Comandos e programas no terminal:
  - Comandos nativos BASH;
  - Programas;
  - Parâmetros.
- Programas Wget e Curl.
- Permissões em arquivos ou diretórios:
  - chown;
  - chmod;
  - chgrp;
  - groupmems.
- A Linguagem BASH:
  - Declaração de variáveis e atribuições de valores;
  - Entrada e saída de dados;
  - Operadores aritméticos, relacionais e lógicos;
  - Estruturas de decisão e repetição;
  - Funções;
  - Pipe;
  - Direcionadores;
  - Execução com e sem passagem de parâmetros.
- Mineração de dados:
  - Expressões Regulares (Regular Expression - RegExp);
  - Identificando padrões para extrair dados;
  - Programas: GREP, AWK e SEd;
  - Processos ETL (Extract, Transform and Load):
    - Extração e Transformação com GREP, AWK e SEd;
    - Carregamento no SQLite3:
      - Criação de Base de Dados;
      - Criação de Tabelas;
      - Inserção de registros;
      - Leitura de registros.

#### 5 - BIBLIOGRAFIA BÁSICA:

- [1] JARGAS, A. M. **Shell Script Profissional**. 1 ed. São Paulo: Novatec, 2008.
- [2] NEGUS, C. **Linux a Bíblia. O Mais Abrangente e Definitivo Guia Sobre Linux**. 1 ed. Rio de Janeiro: Alta Books, 2014.
- [3] SILVA, L. C.; FERRAR, D. G; QUERO, P. **Introdução à Mineração de Dados. Conceitos Básicos, Algoritmos e Aplicações**. 1 ed. São Paulo: Saraiva, 2016.

#### 6 - BIBLIOGRAFIA COMPLEMENTAR:

- [1] CEZAR, J. **Programação Shell Linux**. 10 ed. São Paulo: Brasport, 2014.
- [2] MORIMOTO, C. E. **kurumin 7. Guia Prático**. 1 ed. Paraná: Sulina, 2007.
- [3] PASSOS, E. **Datamining. Conceitos, Técnicas, Algoritmos, Orientações e Aplicações**. 2 ed. Rio de Janeiro: Campus, 2015.