**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**

**FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS**

**PROGRAMA (MAESTRÍA EN ESTADISTICA APLICADA)**

**MANEJO DE DATOS**

Profesor Principal: Msig. Dalton Noboa Macias

Correo: [dgnoboa@espol.edu.ec](mailto:dgnoboa@espol.edu.ec)

**Trayectoria Académica:**

* Magister en Sistemas de Información Gerencial
* Diplomado Superior en Auditoría Informática
* Ingeniero en Estadística Informática
* Bachiller en Ciencias Computacionales

**Trayectoria en ESPOL:**

* Decano de Grado de la ESPOL (2021-Nov 2022)
* Director de Admisiones de la ESPOL (2016 – 2021)
* Coordinador Académico (2011- 2016)
* Profesor Titular de la ESPOL (2011 hasta la actualidad)
* Profesor Ocasional de la ESPOL (2005-2011)
* Profesor Ocasional USM (2006-2010)

**Docencia en Grado:**

* Estadística
* Software Estadístico
* Matemática
* Fundamentos de Programación
* Bases de Datos
* Sistemas de Información
* Auditoría Informática

**Docencia en Postgrado:**

* Estructura de Datos y algoritmos
* Programación y Simulación
* Estadística
* Manejo de Datos

**Curso de Manejo de Datos**

**Objetivo General:**

Presentar esquemas de recuperación, procesamiento, almacenamiento, depuración y limpieza de datos, a través de diferentes procedimientos y técnicas aplicadas en software estadístico que permitan un posterior análisis adecuado de los datos

**Descripción del Curso:**

Este curso corresponde a la unidad de formación disciplinar avanzada de la maestría en Estadística Aplicada y está diseñado para impartir los conceptos introductorios al SQL, los datos estructurados y los datos no estructurados, y su aplicación con software estadístico. Se abarca las técnicas de extracción de diferentes fuentes, así como la integración, la depuración, la transformación y el almacenamiento de datos.

**Contenido:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Unidad** | **Nombre de las Unidades y Subunidades** | **Horas de componentes** | | |
| **Contacto con el profesor** | **Práctico-Experimental** | **Aprendizaje Autónomo** |
|  | 1. Introducción a SQL    1. Conceptos sobre bases de datos    2. Creación de una base de datos con sus respectivos componentes en SQL    3. Creación de procedimientos almacenados    4. Mantenimiento de la información de una base de datos con SQL    5. Triggers | 8 | - | 16 |
|  | 1. Datos estructurados y no estructurados    1. Proveniencia y esquema de tratamiento de los datos estructurados    2. Proveniencia y esquema de tratamiento de los datos no estructurados    3. Modelo de Datos    4. Modelo Relacional    5. Modelo analítico | 8 | - | 16 |
|  | 1. Recopilación, integración y manipulación y almacenamiento de datos    1. Conexión con bases de datos desde software estadístico    2. Selección y recuperación de datos (Queries SQL) de una sola tabla, por medio de comandos del software estadístico.    3. Selección y recuperación de datos (Queries SQL) de tablas relacionadas, por medio de comandos del software estadístico.    4. Funciones de datos agregados (agrupados)    5. Exportación e importación de los datos.    6. Almacenamiento de datos en dataframe y archivos. | 8 | - | 16 |
|  | 1. Selección, limpieza y transformación de datos    1. Definición del proceso ETL    2. Aplicaciones informáticas para el ETL    3. Separa, aplica y combina datos utilizando funciones de software estadístico (dplyr, stringr)    4. Filtra datos, ordena y añade campos computados, aplicando funciones de software estadístico (filter, arrange, mutate)    5. Limpieza de datos utilizando funciones de software estadístico (tidyr, gather, spread, separate, unite, lubridate). | 8 | - | 16 |
|  | **TOTAL DE HORAS** | 32 | 0 | 64 |

**Actividades Planificadas:**

* Exposición del profesor y uso de software estadístico (viernes 30 de junio, sábado 1 de julio, jueves de 6 de julio)
* 3 talleres grupales
* 2 quiz individuales
  + Desde las 14h00 del 1 de julio hasta las 23h59 del 2 de julio
  + Desde las 08h00 del 5 de julio hasta las 18h00 del 6 de julio
* 1 Examen 8 de julio

**Políticas de calificaciones:**

* Quiz 10%
* Talleres 40%
* Examen 50%