



Laboratorio de Seminario de Sistemas 2

<b>CÓDIGO:</b>	<b>798</b>	<b>CRÉDITOS:</b>	<b>3</b>
<b>ESCUELA:</b>	<b>Ciencias y Sistemas</b>	<b>ÁREA A LA QUE PERTENECE:</b>	<b>Desarrollo de Software</b>
<b>CATEGORÍA:</b>	<b>Obligatorio</b>	<b>SEMESTRE:</b>	<b>Segundo Semestre 2024</b>
<b>AUXILIAR:</b>	<b>Sergio Cubur</b>	<b>HORARIO:</b>	<b>14:00 15:50</b>
<b>EDIFICIO Y SALÓN:</b>	<b>Meet</b>		
<b>DÍAS QUE SE IMPARTE</b>	<b>Sabado</b>		

**DESCRIPCIÓN DEL LABORATORIO:**

El laboratorio está diseñado para que al estudiante se le dé una introducción a las herramientas de inteligencia de negocios en sistemas transaccionales y búsqueda de información. Aprenderán a programar soluciones para procesar enormes cantidades de datos y crear modelos a partir de los datos, que permita clasificar la información partiendo de un conjunto de reglas inferidas de los datos.

**OBJETIVO GENERAL:**

Crear proyectos de inteligencia de negocios en las organizaciones utilizando las herramientas de tendencia.

**OBJETIVOS ESPECIFICOS:**

- Reunir, depurar y transformar grandes volúmenes de datos en información estructurada y coherente.
- Identificar y comprender los conceptos de Big Data.
- Comprender conceptos básicos de la ciencia de los datos.
- Ampliar el conocimiento sobre el lenguaje Python, desde la instalación de las herramientas hasta el aprendizaje de nuevas perspectivas de uso.
- Comprender los conceptos sobre Procesamiento masivo paralelo.
- Evolucionar hacia el perfil de Data Scientist.

**METODOLOGÍA:**

Se impartirán clases virtuales, con la participación de los estudiantes para el manejo de aplicaciones orientadas a cubrir los temas del contenido.

Autoaprendizaje y lectura constante de información relacionada al laboratorio.

Exámenes cortos, tareas, prácticas y proyecto, para el control del progreso.

**EVALUACIÓN DEL RENDIMIENTO ACADÉMICO:**

Según el Reglamento General de Evaluación y Promoción del Estudiante de la Universidad de San Carlos de Guatemala, la nota mínima para ganar el laboratorio es de 61 puntos.

#### **PONDERACIÓN:**

<b>Instrumento de evaluación</b>	<b>Ponderación</b>
Proyecto Fase 1	15 pts.
Proyecto Fase 2	20 pts.
Practicas (2)	20 pts.
Tareas (5)	10 pts.
Hojas de Trabajo (5)	10 pts.
Cortos (3)	15 pts.
Examen Final	10 pts.
<b>Total</b>	<b>100 pts.</b>

#### **CONTENIDO:**

##### **Unidad 1: Cubos Multidimensionales**

Análisis multidimensional (OLAP)

Definición y detalles sobre cubos multidimensionales.

Detalles de los tipos de modelos para el modelamiento de cubos.

Proceso de ETL y sus características.

Herramientas para ETL (clasificación)

##### **Unidad 2: BI con herramientas Microsoft**

Introducción a herramientas de Microsoft para solución BI

SSIS de Visual Studio.

SSAS de Visual Studio.

Introducción a la reportería con herramientas de SSRS.

##### **Unidad 3: Análisis de Datos con Python**

Introducción a Python para análisis de datos

Manipulación de datos con Pandas

Visualización de datos con Matplotlib

Datos no estructurados con NLTK

##### **Unidad 4: Introducción a BigQuery**

Qué es BigQuery y sus características principales

Arquitectura de BigQuery

Carga de datos a BigQuery desde archivos CSV

Consultas SQL en BigQuery

# Select, Where, Group By, Order By, Limit

# Funciones de agregación, JOINS, etc

Visualización de resultados de consultas

Integración y concretos de Big Query

# Conexión desde Python

# Ingesta de datos desde Cloud Storage

# Big Query GIS

Ventajas y casos de uso BigQuery vs bases de datos tradicionales