



Laboratorio de Seminario de Sistemas 2

CÓDIGO:	798	CRÉDITOS:	3
ESCUELA:	Ciencias y Sistemas	ÁREA A LA QUE PERTENECE:	Desarrollo de Software
CATEGORÍA:	Obligatorio	SEMESTRE:	Segundo Semestre 2022
AUXILIAR:	Sergio González	HORARIO:	10:40 - 12:20
EDIFICIO Y SALÓN:	Meet		
DÍAS QUE SE IMPARTE EL LABORATORIO:	Viernes		

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:

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

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: Herramientas para científicos de datos

Tipos de herramientas que debemos conocer como científicos de datos

Herramientas más utilizadas

Introducción a R

Introducción a Python (orientado al análisis de datos)

Unidad 4: Procesamiento masivo paralelo, Big Data

Introducción a Hadoop.

Funcionamiento de HDFS

Funcionamiento de MapReduce

Hadoop y su vinculación con Big Data.

Ejemplos sobre los diferentes sistemas de almacenamiento sobre Hadoop.

Introducción a Spark

Arquitectura Spark y sus componentes

Ejemplos de utilización de Spark