



Laboratorio de Seminario de Sistemas 2

CATEGORÍA:	Obligatorio	SEMESTRE:	Segundo Semestre 2019
AUXILIAR:	Glen Abraham Calel Robledo	HORARIO:	9:00 – 10:40
EDIFICIO Y SALÓN:	T3 – Laboratorio 013		
DÍAS QUE SE IMPARTE EL LABORATORIO:	Lunes		

DESCRIPCIÓN DEL LABORATORIO:

A través de este curso los alumnos 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 Específicos:

- Implementar proyectos de Inteligencia de negocios.
- Transformar datos en información.
- Desarrollar una solución de inteligencia empresarial a través de las herramientas de Microsoft (Integration Service, Analisis Services y Reporting Services).
- Comprender y explicar los diversos componentes del marco Apache Spark.
- Comprender conceptos básicos de la ciencia de los datos.
- Hacer uso del lenguaje R, desde la instalación del programa hasta la realización de los fundamentos de la Ciencia de los Datos.
- Evolucionar hacia el perfil de Data Scientist.

METODOLOGÍA:

Se impartirán clases magistrales, 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.

Aprendizaje y práctica del lenguaje R por medio de la plataforma DataCamp.

Exámenes cortos, tareas, prácticas y proyectos, 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**.

PROCEDIMIENTO

Instrumento de evaluación	Ponderación
Tareas (4)	10 pts
Cortos (3)	15 pts
Hojas de trabajo (4)	10 pts
Prácticas (2)	20 pts
Proyecto Fase 1	15 pts
Proyecto Fase 2	20 pts
Examen Final (1)	10 pts
Nota de promoción	100 puntos

CONTENIDO

Unidad 1: Cubos Multidimensionales

Herramientas de análisis (Reportes, Dashboards, Cubos)

Definición de un cubo multidimensional

Indicadores, atributos y jerarquías

Modelamiento de cubos (Estrella, copo de nieve, constelación)

Operaciones sobre un cubo OLAP(Slice, Dice, Drill up & down, Drill & Roll across, Pivot)

Unidad 2: Solución de BI con herramientas Microsoft

Introducción y uso de SQL Server Integration Services

Creación de cubos en SQL Analysis Services

Reportería en SQL Reporting Services

Unidad 3: Procesando Big Data con Apache Spark (Implementado con Hadoop)

Que es Apache Spark

Arquitectura Spark y sus componentes

Modos de desarrollo en Spark

Framework Spark

Preparación del entorno

Introducción a Spark y a RDD

Transformaciones sobre RDD

Unidad 4: Programando con R

¿Qué es R?

Conociendo a R

Configuración

Aritmética básica y objetos

Operadores lógicos

Scripts en R

Importar datos - read.csv

Funciones básicas

Tipos de objetos

Seleccionando datos

Bucles y programación con R

Gráficos