



### Laboratorio de Seminario de Sistemas 2

<b>CATEGORÍA:</b>	Obligatorio	<b>SEMESTRE:</b>	Segundo Semestre 2018
<b>AUXILIAR:</b>	Luis Barrios	<b>HORARIO:</b>	19:00 – 20:40
<b>EDIFICIO Y SALÓN:</b>	T-3 314		
<b>DÍAS QUE SE IMPARTE EL LABORATORIO:</b>	Jueves		

#### 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.
- Identificar y comprender los conceptos de Big Data.
- Comprender y explicar los diversos componentes del marco Spark.
- Transformar datos en información.
- Desarrollar una solución de inteligencia empresarial a través de las herramientas de Microsoft
- Diseñar modelos predictivos bajo el concepto de minería de datos.
- Extraer de manera automatizada patrones ocultos, tendencias, anomalías, vinculaciones complejas y significados en grandes volúmenes de datos.
- 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.
- Evolucionar hacia el perfil de Arquitecto Big Data.

#### 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.

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
Practicas (2)	20 pts
Proyectos (1)	35 pts
Examen Final (1)	10 pts
<b>Nota de promoción</b>	<b>100 puntos</b>

## CONTENIDO

### Unidad 1: Cubos Multidimensionales

Definición de un cubo multidimensional  
Indicadores, atributos y jerarquías  
Modelamiento de cubos (Estrella, copo de nieve, constelación)  
Herramientas de análisis(Reportes, Dashboards)  
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

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: Minería de datos: Uso de RapidMiner

Data Access: Conectando con la información  
Exploración de los datos: Descubriendo patrones  
Data Blending: Analisis predictivo

### Unidad 5: 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