



**Laboratorio de Seminario de Sistemas 2**

<b>CATEGORÍA:</b>	Obligatorio	<b>SEMESTRE:</b>	Segundo Semestre 2020
<b>AUXILIAR:</b>	Breyner Cortez	<b>HORARIO:</b>	12:20 – 14:00
<b>EDIFICIO Y SALÓN:</b>	Meet		
<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.
- Transformar una gran cantidad de datos en información.
- Comprender conceptos básicos de la ciencia de los datos.
- Hacer uso del lenguaje Python, desde la instalación de las herramientas hasta el aprendizaje de generar modelos de machine learning.
- 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
Prácticas (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.  
Definición de conceptos sobre DataWarehouse.  
Proceso de ETL y sus características.  
Definición y conceptos de Datamart.  
Detalles de los tipos de modelos para el modelamiento de cubos.

### Unidad 2: Solución de BI con herramientas Microsoft

Definición de Business Intelligences y aplicaciones en el mundo real.  
Introducción a herramientas de Microsoft para solución BI  
Usos de SSIS de Visual Studio.  
Usos de SSAS de Visual Studio.  
Introducción a la reportería con herramientas de SSRS.

### Unidad 3: Procesando Big Data con Apache Spark

Definición y conceptos de Big Data.  
¿Porqué usar Spark?  
Arquitectura Spark y sus componentes  
Introducción a Spark y a RDD  
Definición y conceptos de Data Science.

### Unidad 4: Procesamiento masivo paralelo y Hadoop

Definición sobre procesamiento masivo paralelo.  
Introducción a Hadoop.  
Hadoop y su vinculación con Big Data.  
Ejemplos sobre los diferentes sistemas de almacenamiento sobre Hadoop.  
¿Para qué y cómo se usa Hadoop?