

SEMINARIO DE SISTEMAS 2 - SECCIÓN P

CÓDIGO:	798	PUNTEO NETO LABORATORIO:	25
ESCUELA DE INGENIERÍA EN:	CIENCIAS Y SISTEMAS	ÁREA A LA QUE PERTENECE:	DESARROLLO DE SOFTWARE
PRE REQUISITO:	285 Sistemas Operativos 2 775 Sistemas de Bases de Datos 2 797 Seminario de Sistemas 1	POST REQUISITO:	7999 Seminario de Investigación EPS Sistemas 799 Seminario de Investigación
CATEGORÍA:	OBLIGATORIO	VIGENCIA:	PRIMER SEMESTRE 2026

Descripción del Laboratorio

El laboratorio del curso está diseñado para que él y la estudiante puedan introducirse y utilizar las herramientas de inteligencia de negocios en sistemas transaccionales y búsqueda de información. Agregado se aprenderá a crear y diseñar soluciones para procesar enormes cantidades de datos y crear modelos adecuados a partir de estos, dichos modelos permitirán clasificar la información partiendo de un conjunto de reglas inferidas de los datos.

Resumen de Ponderaciones y Tiempo de Auto-aprendizaje

TIPO	PONDERACIÓN	HORAS DE AUTO-APRENDIZAJE
Actividades en Clase	15	0
Proyectos	50	42
Prácticas	20	34
Tareas	5	14
Examen Final	10	0
TOTAL	100	90

Equipo Académico

Coordinador del Área

Nombre: Ing. Cesar Fernandez	Correo electrónico: cesarafc1967@gmail.com
--	--

Docente

Nombre del Docente Luis Alberto Vettorazzi España	Correo electrónico del Docente luisvettorazzi@gmail.com
---	---

	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado
Día	X					
Horario	9:00 – 12:20					
Lugar	MEET					

Tutor(es)

Nombre del Tutor	Javier Andrés Valdez González	
Correo electrónico institucional	2991917610101@ingenieria.usac.edu.gt	

Tipo		Lunes	Martes	Miércoles	Jueves	Viernes	Sábado
Clase	Día			x			
	Horario			17:20 – 19:00			
	Lugar			MEET			
Atención al Estudiante	Día						
	Horario						
	Lugar						

Índice

Descripción del Laboratorio	1
Resumen de Ponderaciones y Tiempo de Auto-aprendizaje	1
Equipo Académico	2
Coordinador del Área	2
Docente	2
Tutor(es)	2
Competencias Vinculadas al Perfil del Egresado	5
Competencias Específicas	5
Competencias Generales	6
Competencias del Laboratorio	6
Competencia(s) Específica(s)	6
Competencia(s) General(es)	6
Diseño Didáctico	7
Sesión de Diagnóstico	7
Evaluación de conocimientos previos	7
Presentación del tutor	7
Presentación de los estudiantes	7
Presentación del programa del curso	7
Evaluación de conocimientos del laboratorio actual	7
Sesión No. 1, Unidad No. 1 Fundamentos de Business Intelligence y Análisis	9
Multidimensional	9
Área Actitudinal (Saber ser)	9
Área de Conocimiento (Saber)	9
Área de Habilidades (Saber Hacer)	10
Sesión No. 2, Unidad 1 Fundamentos de Business Intelligence y Análisis	10
Multidimensional - Unidad 2 Procesos ETL y Modelado de Datos para Análisis	10
OLAP	10
Área Actitudinal (Saber ser)	10
Área de Conocimiento (Saber)	10
Área de Habilidades (Saber Hacer)	11
Sesión No. 3, Unidad No. 2 Procesos ETL y Modelado de Datos para Análisis	11
OLAP	11
Área Actitudinal (Saber ser)	11
Área de Conocimiento (Saber)	11
Área de Habilidades (Saber Hacer)	12
Sesión No. 4, Unidad No. 2 Procesos ETL y Modelado de Datos para Análisis	13
OLAP	13
Área Actitudinal (Saber ser)	13

Área de Conocimiento (Saber)	13
Área de Habilidades (Saber Hacer)	13
Sesión No. 5, Unidad No. 3 - Arquitectura de Datos Modernos y Herramientas BI Empresariales	14
Área Actitudinal (Saber ser)	14
Área de Conocimiento (Saber)	14
Área de Habilidades (Saber Hacer)	14
Sesión No. 6, Unidad No. 3- Arquitectura de Datos Modernos y Herramientas BI Empresariales	15
Área Actitudinal (Saber ser)	15
Área de Conocimiento (Saber)	15
Área de Habilidades (Saber Hacer)	15
Sesión No. 7, Unidad No. 4 - Visualización Analítica, KPI y Scorecards en Plataformas BI	16
Área Actitudinal (Saber ser)	16
Área de Conocimiento (Saber)	16
Área de Habilidades (Saber Hacer)	17
Sesión No. 8, Unidad No. 5 - Análisis de Datos Masivos: Big Data, Hadoop y Procesamiento Distribuido	17
Área Actitudinal (Saber ser)	17
Área de Conocimiento (Saber)	17
Área de Habilidades (Saber Hacer)	18
Sesión No. 9, Unidad No. 5 - Análisis de Datos Masivos: Big Data, Hadoop y Procesamiento Distribuido	18
Área Actitudinal (Saber ser)	18
Área de Conocimiento (Saber)	18
Área de Habilidades (Saber Hacer)	19
Sesión No. 10, Unidad No. 6 - Ciencia de Datos: Modelado Predictivo y Análisis de Texto con Python	19
Área Actitudinal (Saber ser)	19
Área de Conocimiento (Saber)	19
Área de Habilidades (Saber Hacer)	20
Sesión No. 11, Unidad No. 7 -Gobierno de Datos y Machine Learning en Entornos Cloud con BigQuery	20
Área Actitudinal (Saber ser)	20
Área de Conocimiento (Saber)	21
Área de Habilidades (Saber Hacer)	21
Rúbrica de Evaluación	22
Normativa Académica y Ética del Curso	22
Bibliografía	23
E-Grafía	23

Competencias Vinculadas al Perfil del Egresado

Competencias Específicas

No.	Competencia
1	Demuestra destreza y habilidad en la selección, uso y adaptación de herramientas metodológicas, tecnológicas, equipos especializados y en la lectura e interpretación de datos pertinentes al contexto de su ejercicio profesional.
2	Toma decisiones profesionales con base en fundamentos teóricos, datos e información pertinente, válida y confiable.
3	Aplica los conocimientos de su disciplina en la elaboración, fundamentación y defensa de argumentos para prevenir y resolver problemas complejos en su campo profesional, identificando y aplicando innovaciones.
4	Identifica oportunidades y riesgos para la innovación y adaptación de conocimientos y tecnologías para resolver problemas.
5	Demuestra pensamiento crítico, actitud investigativa y rigor analítico en el planteamiento y la resolución de problemas complejos.

Competencias Generales

No.	Competencia
1	Maneja e Interpreta adecuadamente datos masivos, sean estos estructurados o no estructurados, facilitando su visualización e interpretación de forma eficaz en apoyo a la toma de decisiones.
2	Aplica principios básicos de ingeniería, ciencias de computación y sistemas de información y comunicación, en la formulación y resolución adecuada de problemas complejos.
3	Aplica estándares de calidad, eficiencia y seguridad en la implementación adecuada de soluciones de software, hardware y TIC en general.
4	Actualiza permanentemente sus conocimientos relacionados con TIC en general, apoyándose en las estrategias de aprendizaje apropiadas
5	Construye soluciones integrales trabajando en forma colaborativa y propositiva en equipos interdisciplinarios, en forma presencial o utilizando plataformas virtuales.

Competencias del Laboratorio

Competencia(s) Específica(s)

No.	Competencia	Nivel de Aprendizaje
1	Diseña soluciones de inteligencia de negocios utilizando herramientas BI como SSIS, SSAS, SSRS, Python y BigQuery con el fin de transformar datos en información útil para apoyar la toma de decisiones en entornos organizacionales	Crear
2	Diseña modelos de análisis multidimensional y cubos OLAP utilizando herramientas como Power BI y Excel para representar datos y facilitar la toma de decisiones en contextos empresariales	Aplicar
3	Implementa procesos ETL eficientes empleando herramientas como SSIS, Pentaho o Talend asegurando la transformación y carga efectiva de datos para su posterior análisis	Analizar
4	Desarrolla tableros e indicadores estratégicos mediante SSAS, SSRS y Power BI para reflejar proyecciones, KPIs y análisis presupuestario conforme a un modelo de Balanced Scorecard	Evaluuar

Competencia(s) General(es)

No.	Competencia	Nivel de Aprendizaje
1	Utiliza entornos de procesamiento distribuido con bibliotecas como PySpark y herramientas Big Data para procesar grandes volúmenes de datos estructurados y no estructurados	Aplicar
2	Administra proyectos de análisis de datos en la nube con herramientas como BigQuery y Google Cloud Platform aplicando principios de gobierno de datos, seguridad y control de acceso	Evaluuar

Diseño Didáctico

Sesión de Diagnóstico

Evaluación de conocimientos previos

Se aplicará una actividad diagnóstica con el objetivo de identificar el nivel de conocimientos y habilidades que los estudiantes poseen al inicio del curso. No influye en la nota final, pero es obligatoria para todos los estudiantes.

Tipo de Actividad	Descripción
Cuestionario	Cuestionario en Quizziz para validar el conocimiento previo de los estudiantes y mapear su conocimiento con temas de Seminario de Sistemas 1.

Presentación del tutor

El tutor se presenta formalmente al grupo, compartiendo su formación académica, experiencia profesional y educativa, así como sus expectativas sobre el curso. También se abordan aspectos como normas de convivencia, canales de comunicación, disponibilidad para consultas y métodos de acompañamiento.

Presentación de los estudiantes

Se escogen un grupo de estudiantes al azar. En su presentación, se les pedirá que comparten información básica como su nombre, intereses personales o profesionales, experiencias previas relacionadas con el curso y sus expectativas. Esta actividad busca promover la interacción, el reconocimiento entre pares y la construcción de un entorno participativo y respetuoso.

Presentación del programa del curso

Se presenta el contenido del programa del curso, se aclaran dudas y se fomenta el compromiso del estudiante con su aprendizaje.

Evaluación de conocimientos del laboratorio actual

Se realiza una evaluación o práctica que permite conocer el grado de familiaridad de los estudiantes con las herramientas, entornos o competencias técnicas necesarias para el laboratorio actual.

Tipo de Actividad	Descripción
Cuestionario	Dentro del mismo cuestionario se validará la afinidad de los estudiantes con herramientas y conocimientos del tema del laboratorio

Sesión No. 1

Unidad No. 1

Fundamentos de Business Intelligence y Análisis Multidimensional

Área Actitudinal (Saber ser)

Nombre del valor:
Pensamiento crítico Se fomenta al discutir qué es BI, motivando al estudiante a reflexionar sobre cómo se toman decisiones con datos y si dichas decisiones son objetivas o manipuladas.
Responsabilidad Se desarrolla al introducir el rol de los datos en la toma de decisiones organizacionales, promoviendo el uso ético y preciso de la información.
Rigor analítico Se aplica al comparar los niveles de madurez analítica y tipos de información, lo cual exige una interpretación cuidadosa y estructurada del contenido.
Innovación Se impulsa al presentar OLAP como enfoque innovador para modelar y analizar datos, ampliando el pensamiento del estudiante sobre herramientas más allá de Excel o reportes simples

Área de Conocimiento (Saber)

Competencia(s)	
Diseña modelos de análisis multidimensional y cubos OLAP utilizando herramientas como Power BI y Excel para representar datos y facilitar la toma de decisiones en contextos empresariales	
Tema	Subtema
Fundamentos de BI y análisis OLAP	¿Qué es Business Intelligence?
Fundamentos de BI y análisis OLAP	Datos para la toma de decisiones
Fundamentos de BI y análisis OLAP	Niveles de madurez analítica
Fundamentos de BI y análisis OLAP	Tipos de información
Fundamentos de BI y análisis OLAP	Introducción al análisis OLAP
Fundamentos de BI y análisis OLAP	Ventajas del análisis multidimensional
Fundamentos de BI y análisis OLAP	Comparativa entre OLAP y herramientas de visualización

Área de Habilidades (Saber Hacer)

Competencia	
Diseña modelos de análisis multidimensional y cubos OLAP utilizando herramientas como Power BI y Excel para representar datos y facilitar la toma de decisiones en contextos empresariales	
Tipo de Actividad	Ponderación
Actividad	1

Sesión No. 2,
Unidad 2
Procesos ETL y Modelado de Datos para Análisis OLAP

Área Actitudinal (Saber ser)

Nombre del valor:
Organización
Se fortalece al construir modelos jerárquicos con dimensiones y medidas, donde cada elemento debe ocupar un lugar específico en la estructura.
Responsabilidad
Se aplica al enseñar que un mal diseño de cubos impacta directamente en la calidad de las decisiones empresariales basadas en esos datos.
Disciplina
Se promueve al seguir paso a paso el modelado multidimensional (estrella, copo de nieve, constelación), exigiendo precisión y constancia.
Pensamiento lógico
Se entrena al construir jerarquías y relaciones entre dimensiones, lo que implica visualizar estructuras complejas de datos de forma racional.

Área de Conocimiento (Saber)

Competencia(s)	
Diseña modelos de análisis multidimensional y cubos OLAP utilizando herramientas como Power BI y Excel para representar datos y facilitar la toma de decisiones en contextos empresariales	
Tema	Subtema
Diseño y optimización de cubos multidimensionales	Definición de cubo multidimensional
Diseño y optimización de cubos multidimensionales	Dimensiones y jerarquías
Diseño y optimización de cubos multidimensionales	Medidas y agregaciones
Diseño y optimización de cubos multidimensionales	Modelos Estrella, Copo de Nieve, Constelación
Diseño y optimización de cubos multidimensionales	Optimización de cubos multidimensionales

Área de Habilidades (Saber Hacer)

Competencia	
Diseña modelos de análisis multidimensional y cubos OLAP utilizando herramientas como Power BI y Excel para representar datos y facilitar la toma de decisiones en contextos empresariales	
Tipo de Actividad	Ponderación
Actividad	1

Sesión No. 3

Unidad No. 2

Procesos ETL y Modelado de Datos para Análisis OLAP

Área Actitudinal (Saber ser)

Nombre del valor:
Adaptabilidad
Se promueve al comparar herramientas ETL comerciales, open source y cloud, enseñando que el estudiante debe adaptarse según el contexto de uso.
Responsabilidad
Se fomenta al mostrar que errores en la transformación de datos pueden causar pérdidas de calidad e incluso decisiones erróneas a gran escala.
Análisis crítico
Se desarrolla al evaluar las ventajas y limitaciones de cada herramienta ETL y su aplicabilidad en escenarios reales.
Disciplina tecnológica
Se refuerza al aplicar flujos ordenados de ETL, donde cada fase debe seguir una secuencia lógica y técnica correcta.

Área de Conocimiento (Saber)

Competencia(s)	
Implementa procesos ETL eficientes empleando herramientas como SSIS, Pentaho o Talend asegurando la transformación y carga efectiva de datos para su posterior análisis	
Tema	Subtema
Fundamentos del proceso ETL y su aplicación en OLAP	¿Qué es ETL?
Fundamentos del proceso ETL y su aplicación en OLAP	Fases del proceso ETL
Fundamentos del proceso ETL y su aplicación en OLAP	Extracción, Transformación y Carga
Fundamentos del proceso ETL y su aplicación en OLAP	Características del ETL
Fundamentos del proceso ETL y su aplicación en OLAP	Rol del ETL en OLAP
Fundamentos del proceso ETL y su aplicación en OLAP	Desafíos del proceso ETL

Área de Habilidades (Saber Hacer)

Competencia	
Implementa procesos ETL eficientes empleando herramientas como SSIS, Pentaho o Talend asegurando la transformación y carga efectiva de datos para su posterior análisis	
Tipo de Actividad	Ponderación
Actividad	1

Sesión No. 4

Unidad No. 2

Procesos ETL y Modelado de Datos para Análisis OLAP

Área Actitudinal (Saber ser)

Nombre del valor:
Innovación Se activa al conocer arquitecturas modernas de BI como bases en memoria y MPP, alentando al estudiante a explorar nuevas soluciones.
Pensamiento estratégico Se desarrolla al analizar qué arquitectura conviene según el tipo de organización, volumen de datos y necesidades.
Adaptabilidad Se refuerza al integrar conceptos tradicionales de BI con nuevos servicios como IAAS/PAAS y herramientas como SSAS/SSRS.
Responsabilidad tecnológica Se aplica al planificar soluciones BI que deben ser seguras, escalables y sostenibles en ambientes reales.

Área de Conocimiento (Saber)

Competencia(s)	
implementa procesos ETL eficientes empleando herramientas como SSIS, Pentaho o Talend asegurando la transformación y carga efectiva de datos para su posterior análisis	
Tema	Subtema
Herramientas ETL y tendencias modernas	Tipos de herramientas ETL
Herramientas ETL y tendencias modernas	Ejemplos de herramientas comerciales y Open Source
Herramientas ETL y tendencias modernas	Herramientas ETL en la nube
Herramientas ETL y tendencias modernas	Criterios de selección
Herramientas ETL y tendencias modernas	Tendencias actuales
Herramientas ETL y tendencias modernas	Casos de uso reales

Área de Habilidades (Saber Hacer)

Competencia	
implementa procesos ETL eficientes empleando herramientas como SSIS, Pentaho o Talend asegurando la transformación y carga efectiva de datos para su posterior análisis	
Tipo de Actividad	Ponderación
Actividad	1

Sesión No. 5**Unidad No. 3****Arquitectura de Datos Modernos y Herramientas BI Empresariales****Área Actitudinal (Saber ser)**

Nombre del valor:
Seguridad y ética profesional Se refuerza al discutir buenas prácticas para proteger datos en la nube y configurar servicios BI de manera segura.
Responsabilidad Se trabaja al implementar virtualización, control de acceso y monitoreo en arquitecturas BI reales.
Pensamiento analítico Se aplica al evaluar las diferencias y relaciones entre SSIS, SSAS y SSRS como componentes de una solución integrada.
Innovación tecnológica Se promueve al usar herramientas modernas y modelos de nube híbrida para ampliar el alcance de las soluciones empresariales.

Área de Conocimiento (Saber)

Competencia(s)	
Integra soluciones BI con bases de datos modernas y servicios en la nube usando tecnologías como SSMS, SSAS y servicios IAAS/PAAS considerando aspectos de seguridad, virtualización y arquitectura de datos distribuida	
Tema	Subtema
Bases modernas y arquitectura Microsoft BI	Bases en memoria, columnares y MPP
Bases modernas y arquitectura Microsoft BI	Appliances analíticos
Bases modernas y arquitectura Microsoft BI	Arquitectura de BI con Microsoft
Bases modernas y arquitectura Microsoft BI	Herramientas Microsoft BI (SSMS, SSIS, SSAS, SSRS)
Bases modernas y arquitectura Microsoft BI	Diferencias entre SSIS, SSAS y SSRS

Área de Habilidades (Saber Hacer)

Competencia	
Integra soluciones BI con bases de datos modernas y servicios en la nube usando tecnologías como SSMS, SSAS y servicios IAAS/PAAS considerando aspectos de seguridad, virtualización y arquitectura de datos distribuida	
Tipo de Actividad	Ponderación
Actividad	1

Sesión No. 6

Unidad No. 3

Arquitectura de Datos Modernos y Herramientas BI Empresariales

Área Actitudinal (Saber ser)

Nombre del valor:
Pensamiento estratégico
Se desarrolla al definir KPIs y estructuras de Scorecard que alinean los datos con objetivos de negocio.
Rigor analítico
Se refuerza al crear reportes con DAX y MDX, donde cada fórmula debe ser lógica y precisa.
Responsabilidad
Se promueve al trabajar con información crítica como presupuestos, proyecciones y desempeño organizacional.
Innovación aplicada
Se fomenta al automatizar reportes y visualizaciones, mostrando cómo optimizar el flujo de información ejecutiva

Área de Conocimiento (Saber)

Competencia(s)	
Integra soluciones BI con bases de datos modernas y servicios en la nube usando tecnologías como SSMS, SSAS y servicios IAAS/PAAS considerando aspectos de seguridad, virtualización y arquitectura de datos distribuida	
Tema	Subtema
Microsoft BI + Cloud y Seguridad	Introducción a SSIS
Microsoft BI + Cloud y Seguridad	Virtualización y federación de datos
Microsoft BI + Cloud y Seguridad	IAAS, PAAS, SAAS en BI
Microsoft BI + Cloud y Seguridad	Seguridad en la nube
Microsoft BI + Cloud y Seguridad	Nubes híbridas y monitoreo

Área de Habilidades (Saber Hacer)

Competencia	
Integra soluciones BI con bases de datos modernas y servicios en la nube usando tecnologías como SSMS, SSAS y servicios IAAS/PAAS considerando aspectos de seguridad, virtualización y arquitectura de datos distribuida	
Tipo de Actividad	Ponderación
Actividad	1

Sesión No. 7

Unidad No. 4

Visualización Analítica, KPI y Scorecards en Plataformas BI

Área Actitudinal (Saber ser)

Nombre del valor:
Pensamiento estratégico Se desarrolla al definir KPIs y estructuras de Scorecard que alinean los datos con objetivos de negocio.
Rigor analítico Se refuerza al crear reportes con DAX y MDX, donde cada fórmula debe ser lógica y precisa.
Responsabilidad Se promueve al trabajar con información crítica como presupuestos, proyecciones y desempeño organizacional.
Innovación aplicada Se fomenta al automatizar reportes y visualizaciones, mostrando cómo optimizar el flujo de información ejecutiva.

Área de Conocimiento (Saber)

Competencia(s)	
Desarrolla tableros e indicadores estratégicos mediante SSAS, SSRS y Power BI para reflejar proyecciones, KPIs y análisis presupuestario conforme a un modelo de Balanced Scorecard	
Tema	Subtema
Modelado multidimensional y KPIs con SSAS	Diseño de cubos OLAP con KPIs
Modelado multidimensional y KPIs con SSAS	Modelos tabulares
Modelado multidimensional y KPIs con SSAS	Consultas con MDX y DAX
Modelado multidimensional y KPIs con SSAS	Rolling Forecast y análisis presupuestario
Visualización con SSRS y Scorecards	Reportes con parámetros y filtros
Visualización con SSRS y Scorecards	Semáforos y visualizaciones condicionales
Visualización con SSRS y Scorecards	Balance Scorecard con SSRS o Power BI

Área de Habilidades (Saber Hacer)

Competencia	
Desarrolla tableros e indicadores estratégicos mediante SSAS, SSRS y Power BI para reflejar proyecciones, KPIs y análisis presupuestario conforme a un modelo de Balanced Scorecard	
Tipo de Actividad	Ponderación
Actividad	1

Sesión No. 8**Unidad No. 5****Análisis de Datos Masivos: Big Data, Hadoop y Procesamiento Distribuido****Área Actitudinal (Saber ser)**

Nombre del valor:
Rigor técnico
Se exige al conectar con bases NoSQL, crear visualizaciones estadísticas y trabajar con subplots complejos.
Pensamiento crítico
Se fomenta al interpretar datos no estructurados y representar hallazgos de forma visual y comprensible.
Responsabilidad en el análisis
Se refuerza al presentar dashboards que pueden influir en decisiones estratégicas dentro de una organización.

Área de Conocimiento (Saber)

Competencia(s)	
Utiliza entornos de procesamiento distribuido con bibliotecas como PySpark y herramientas Big Data para procesar grandes volúmenes de datos estructurados y no estructurados	
Tema	Subtema
Ánalysis y ecosistemas Big Data	¿Qué es Big Data? (4 V)
Ánalysis y ecosistemas Big Data	Arquitectura de Hadoop
Ánalysis y ecosistemas Big Data	Fundamentos de Python para análisis
Ánalysis y ecosistemas Big Data	Exploración de datos en Pandas
Ánalysis y ecosistemas Big Data	Carga desde data lakes

Área de Habilidades (Saber Hacer)

Competencia	
Utiliza entornos de procesamiento distribuido con bibliotecas como PySpark y herramientas Big Data para procesar grandes volúmenes de datos estructurados y no estructurados	
Tipo de Actividad	Ponderación
Actividad	1

Sesión No. 9

Unidad No. 5 -

Análisis de Datos Masivos: Big Data, Hadoop y Procesamiento Distribuido

Área Actitudinal (Saber ser)

Nombre del valor:
Pensamiento crítico
Se trabaja al construir modelos y analizar sus predicciones, evaluando si tienen sentido según el problema planteado.
Ética en el manejo de datos
Se aplica al trabajar con datos textuales de personas (sentimientos), cuidando no generar sesgos o interpretaciones erróneas.
Disciplina analítica
Se desarrolla al aplicar CRISP-DM paso a paso, desde recolección hasta validación de resultados.
Persistencia investigativa
Se refuerza al probar distintos modelos y métricas hasta encontrar la solución más adecuada.

Área de Conocimiento (Saber)

Competencia(s)	
Utiliza entornos de procesamiento distribuido con bibliotecas como PySpark y herramientas Big Data para procesar grandes volúmenes de datos estructurados y no estructurados	
Tema	Subtema
Procesamiento distribuido con PySpark	Introducción a Spark y PySpark
Procesamiento distribuido con PySpark	Comparativa Pandas vs PySpark
Procesamiento distribuido con PySpark	Conexión a bases NoSQL
Procesamiento distribuido con PySpark	Análisis de datos no estructurados
Procesamiento distribuido con PySpark	Procesamiento de volúmenes grandes

Área de Habilidades (Saber Hacer)

Competencia	
Utiliza entornos de procesamiento distribuido con bibliotecas como PySpark y herramientas Big Data para procesar grandes volúmenes de datos estructurados y no estructurados	
Tipo de Actividad	Ponderación
Actividad	1

Sesión No. 10

Unidad No. 6

Ciencia de Datos: Modelado Predictivo y Análisis de Texto con Python

Área Actitudinal (Saber ser)

Nombre del valor:
Ética de la información
Se aplica al enseñar el valor de proteger datos sensibles, definir permisos y evitar exposiciones indebidas.
Responsabilidad
Se promueve al administrar roles, recursos y costos dentro de entornos cloud, impactando directamente el control de los datos.
Pensamiento estratégico
Se fomenta al diseñar arquitecturas y políticas de datos escalables, seguras y optimizadas para distintas áreas del negocio.
Rigor técnico
Se requiere para configurar correctamente permisos, tipos de almacenamiento y precios en Google Cloud.

Área de Conocimiento (Saber)

Competencia(s)	
Construye modelos predictivos y análisis de sentimientos usando bibliotecas Python como Scikit-learn y NLTK siguiendo la metodología CRISP-DM para resolver problemas reales a partir de datos estructurados/textuales	
Tema	Subtema
Visualización avanzada para ciencia de datos	Matplotlib y Seaborn
Visualización avanzada para ciencia de datos	Dashboards con subplots
Visualización avanzada para ciencia de datos	Gráficos estadísticos
Visualización avanzada para ciencia de datos	Rol del científico e ingeniero de datos
Visualización avanzada para ciencia de datos	CRISP-DM y Scikit-learn
Visualización avanzada para ciencia de datos	Modelo de regresión
Visualización avanzada para ciencia de datos	Clasificación y Bag of Words

Área de Habilidades (Saber Hacer)

Competencia	
Construye modelos predictivos y análisis de sentimientos usando bibliotecas Python como Scikit-learn y NLTK siguiendo la metodología CRISP-DM para resolver problemas reales a partir de datos estructurados/textuales	
Tipo de Actividad	Ponderación
Actividad	1

Sesión No. 11

Unidad No. 7

Gobierno de Datos y Machine Learning en Entornos Cloud con BigQuery

Área Actitudinal (Saber ser)

Nombre del valor:
Innovación aplicada
Se trabaja al utilizar BigQuery ML para crear modelos predictivos directamente desde SQL en la nube.
Pensamiento crítico
Se aplica al evaluar los resultados de los modelos y reflexionar sobre su aplicabilidad y limitaciones.
Rigor técnico
Se exige para optimizar consultas, crear vistas materializadas y configurar correctamente el pipeline de datos.
Responsabilidad
Se fortalece al entregar un modelo y análisis final que puede influir en decisiones reales o simuladas dentro del proyecto.

Área de Conocimiento (Saber)

Competencia(s)	
Administra proyectos de análisis de datos en la nube con herramientas como BigQuery y Google Cloud Platform aplicando principios de gobierno de datos, seguridad y control de acceso	
Tema	Subtema
Gobierno de datos con BigQuery	¿Qué es BigQuery?
Gobierno de datos con BigQuery	Arquitectura, precios y permisos
Gobierno de datos con BigQuery	Aplicación del gobierno de datos
Modelado y ML con BigQuery	Optimización de consultas
Modelado y ML con BigQuery	Tablas materializadas y vistas
Modelado y ML con BigQuery	BigQuery ML
Modelado y ML con BigQuery	Exportación a herramientas externas

Área de Habilidades (Saber Hacer)

Competencia	
Administra proyectos de análisis de datos en la nube con herramientas como BigQuery y Google Cloud Platform aplicando principios de gobierno de datos, seguridad y control de acceso	
Tipo de Actividad	Ponderación
Cuestionario	5

Rúbrica de Evaluación

Cada una de las actividades del laboratorio (proyectos, prácticas, tareas y otras) cuenta con una rúbrica de evaluación específica, la cual está detallada en el documento que se entrega al estudiante al momento de asignar la actividad. Estas rúbricas describen los criterios de evaluación, niveles de desempeño esperados y la ponderación correspondiente de cada aspecto evaluado.

Es **responsabilidad del estudiante** leer detenidamente la rúbrica asignada antes de iniciar el desarrollo de la actividad. Comprender los criterios de evaluación no solo permite orientar adecuadamente el trabajo, sino también mejorar el desempeño académico y fomentar la autorregulación del aprendizaje.

En caso de no recibir la rúbrica al momento de la asignación, el estudiante **debe solicitarla directamente al tutor académico**, ya que constituye una herramienta esencial para el cumplimiento de los objetivos de aprendizaje y la evaluación transparente.

Normativa Académica y Ética del Curso

En concordancia con el perfil del estudiante de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, se espera un alto nivel de compromiso con la excelencia académica y la ética profesional. Por ello, que se establece los siguientes lineamientos de carácter obligatorio que regulan el comportamiento académico del estudiante:

Plagio y copias

- Todo proyecto será sometido a verificación para confirmar su autoría y originalidad, con la finalidad de evitar cualquier plagio, copia o que la actividad no haya sido realizada por el estudiante.
- Cualquier evidencia de lo antes descrito en las distintas actividades será sancionada con una calificación de 0 (cero) y el caso será reportado al Docente quien a su vez informará a la Escuela de Ciencias y Sistemas para su seguimiento institucional.

Prórrogas y reposiciones

- No se otorgarán prórrogas para entregas de actividades.
- No se permitirá la reposición de proyectos bajo ninguna circunstancia.

Requisitos para evaluación final del curso

- Es obligatorio aprobar el laboratorio para tener derecho a la evaluación final del curso.
- La calificación de prácticas, proyectos y otras actividades que se indique será asignada de forma presencial, en la fecha y hora establecidas por el tutor académico.

Asistencia

- Para obtener la nota del laboratorio, se requiere un mínimo del 80% de asistencia a las sesiones de laboratorio.
- En caso de inasistencia, sólo se aceptarán justificaciones válidas respaldadas por constancia oficial.

Entregas

- No se aceptarán entregas tardías de tareas, prácticas, exámenes cortos, exámenes finales o proyectos sin justificación.

Medio oficial de entrega

- La plataforma UEDI de la Facultad será el único medio oficial para la entrega de actividades del curso.

E-Grafía

- Apache Software Foundation. (2025, 20 de agosto). Apache Hadoop 3.4.2 documentation. <https://hadoop.apache.org/docs/stable/>
- Apache Software Foundation. (s. f.). Overview — Apache Spark documentation. Recuperado el 20 de diciembre de 2025, de <https://spark.apache.org/docs/latest/>
- Apache Software Foundation. (s. f.). PySpark documentation (Python API). Recuperado el 20 de diciembre de 2025, de <https://spark.apache.org/docs/latest/api/python/index.html>
- Balanced Scorecard Institute. (s. f.). BSC basics. Recuperado el 20 de diciembre de 2025, de <https://balancedscorecard.org/bsc-basics/>
- Bird, S., Klein, E., & Loper, E. (s. f.). Natural Language Processing with Python (NLTK book). Recuperado el 20 de diciembre de 2025, de <https://www.nltk.org/book/>
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide [PDF]. Recuperado el 20 de diciembre de 2025, de <https://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/14.2/es/CRISP-DM.pdf>
- Google. (s. f.). BigQuery documentation. Recuperado el 20 de diciembre de 2025, de <https://docs.cloud.google.com/bigquery/docs>
- Google. (s. f.). BigQuery IAM roles and permissions. Recuperado el 20 de diciembre de 2025, de <https://docs.cloud.google.com/bigquery/docs/access-control>
- Google. (s. f.). Identity and Access Management (IAM) documentation. Recuperado el 20 de diciembre de 2025, de <https://docs.cloud.google.com/iam/docs>
- Google. (s. f.). Introduction to ML in BigQuery. Recuperado el 20 de diciembre de 2025, de <https://docs.cloud.google.com/bigquery/docs/bqml-introduction>
- Kimball Group. (s. f.). Star schemas and OLAP cubes. Recuperado el 20 de diciembre de 2025, de <https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/star-schema-olap-cube/>
- Microsoft. (2024, 5 de febrero). Cubes in Multidimensional Models. Microsoft Learn. <https://learn.microsoft.com/en-us/analysis-services/multidimensional-models/cubes-in-multidimensional-models?view=sql-analysis-services-2025>
- Microsoft. (2024, 26 de septiembre). SQL Server Integration Services (SSIS). Microsoft Learn. <https://learn.microsoft.com/en-us/sql/integration-services/sql-server-integration-services?view=sql-server-ver17>
- Microsoft. (2025, 6 de junio). Introducción a SQL Server Analysis Services (SSAS). Microsoft Learn. <https://learn.microsoft.com/es-es/analysis-services/ssas-overview?view=sql-analysis-services-2025>
- Microsoft. (2024, 31 de enero). MDX function reference (MDX). Microsoft Learn. <https://learn.microsoft.com/en-us/sql/mdx/mdx-function-reference-mdx?view=sql-server-ver17>
- Microsoft. (2025, 17 de marzo). DAX function reference. Microsoft Learn. <https://learn.microsoft.com/en-us/dax/dax-function-reference>
- Microsoft. (2024, 30 de diciembre). Understand star schema and the importance for Power BI. Microsoft Learn. <https://learn.microsoft.com/en-us/power-bi/guidance/star-schema>
- Microsoft. (2025, 16 de junio). ¿Qué es SQL Server Reporting Services (SSRS)? Microsoft Learn. <https://learn.microsoft.com/es-es/sql/reporting-services/create-deploy-and-manage-mobile-and-paginated-reports?view=sql-server-ver17>
- Microsoft. (s. f.). Power Pivot: Overview and learning. Microsoft Support. Recuperado el 20 de diciembre de 2025, de <https://support.microsoft.com/en-us/office/power-pivot-overview-and-learning-f9001958-7901-4caa-ad80-028a6d2432ed>
- MongoDB, Inc. (s. f.). MongoDB documentation. Recuperado el 20 de diciembre de 2025, de <https://www.mongodb.com/docs>
- Pentaho. (2025, 23 de julio). Pentaho Data Integration 10.2. Pentaho Documentation. <https://docs.pentaho.com/pdia-data-integration>
- Qlik. (s. f.). Functional architecture of Talend Data Integration. Qlik Help. Recuperado el 20 de diciembre de 2025, de <https://help.qlik.com/talend/en-US/studio-getting-started-guide-data-integration/8.0/functional-architecture>
- scikit-learn developers. (s. f.). User guide. Recuperado el 20 de diciembre de 2025, de <https://scikit-learn.org/stable/>

https://scikit-learn.org/stable/user_guide.html
Reis, J., & Housley, M. (2022). Fundamentals of Data Engineering: Plan and Build Robust Data Systems. O'Reilly Media.