

# Formulario de Programa de Estudios Estándar

1. Se necesita editar lo que está en [rojo]. Algunos de los [rojos] son ejemplos. Algunos de los [rojos] son opcionales.
2. La hora de crédito del curso no debe cambiarse.
3. Se deben incluir los contenidos requeridos del curso.
4. El detalle de los contenidos del curso debe incluir todos los contenidos agregados.

# Introducción a Big Data

## Información Básica

### Información del curso

1. Título: Introducción a Big Data
2. Código: [Código del curso]
3. Año lectivo: [2024]
4. Semestre: [Primero, Segundo]
5. Tipo de curso: [Optativo, Requerido]
6. Departamento: Centro de Innovación TIC
7. Año/Nivel: [Prerrequisito, **Básico**, Intermedio, Avanzado]
8. Horas de crédito:
  - Créditos Totales: -
  - Créditos de Clases: -
  - Créditos Experimento/Práctico: -
  - Créditos de Diseño: -
  - Otros: -
9. Formato de clase: **Charlas y sesiones prácticas.**

### Hora y lugar

1. Días: [Martes y viernes]
2. Hora: [11:30 a.m. a 12:45 p.m.]
3. Ubicación: [Lugar]

### Información del instructor

1. Nombre: [Nombre del instructor]
2. Oficina: [Ubicación de la oficina del instructor]
3. Contacto: [Correo electrónico o número de teléfono del instructor]

## Horas de oficina

[El instructor anunciará el horario de oficina para apoyo adicional y aclaraciones fuera del horario de clases programado. Estas sesiones son una oportunidad para que los estudiantes discutan los materiales del curso, busquen ayuda con las tareas y obtengan comentarios sobre su progreso.]

## Notas adicionales

1. [Se anima a los estudiantes a comunicarse con el instructor a través de los datos de contacto proporcionados para cualquier consulta o inquietud relacionada con el curso.]
2. [Es importante asistir tanto a las sesiones teóricas como a las prácticas para beneficiarse plenamente de la estructura del curso.]
3. [El nombre del instructor, la ubicación de la oficina y la información de contacto serán proporcionados por el departamento respectivo o a través del portal en línea del curso. Los estudiantes deben asegurarse de tener estos detalles para una comunicación efectiva durante toda la duración del curso.]

# Prerrequisito

[SIT22001 Programación I]

# Descripción del Curso

Introducción a Big Data es un curso fundamental diseñado para presentar a los estudiantes el vasto y en evolución campo del análisis de big data. En una era en la que la generación de datos es ubicua y exponencial, la capacidad de gestionar, procesar y extraer valor de forma eficaz a partir de grandes conjuntos de datos es invaluable en todos los sectores. Este curso ofrece una inmersión profunda en los conceptos fundamentales de big data, incluida su definición, características y los desafíos que presenta. A través de una combinación de conocimientos teóricos y aplicaciones prácticas, los estudiantes explorarán cómo los big data están transformando las industrias, impulsando la innovación y dando forma a nuestra comprensión del mundo.

El plan de estudios está estructurado para proporcionar una descripción general completa de las tecnologías y metodologías de big data. Comenzando con los conceptos básicos de almacenamiento y gestión de datos, los estudiantes aprenderán sobre marcos informáticos distribuidos como Hadoop y Spark, que permiten el procesamiento eficiente de grandes conjuntos de datos. El curso cubrirá aspectos clave del análisis de datos, incluida la minería de datos, algoritmos de aprendizaje automático y métodos estadísticos diseñados para entornos de big data. Se pone énfasis en las habilidades prácticas, y los estudiantes participan en ejercicios y proyectos prácticos que simulan escenarios de análisis de datos del mundo real.

Una característica importante de este curso es su enfoque en las herramientas y plataformas que son fundamentales en el ecosistema de big data. Los estudiantes adquirirán competencia en el uso de bases de datos NoSQL para almacenamiento escalable, Apache Spark para procesamiento de big data y herramientas de visualización de datos para comunicar conocimientos. Estas habilidades son fundamentales para afrontar los desafíos técnicos de big data y aprovechar su potencial para obtener conocimientos prácticos. El curso brindará a los estudiantes acceso a estas tecnologías, ofreciendo un entorno de pruebas para la exploración y el aprendizaje.

Más allá de las habilidades técnicas, el curso aborda las consideraciones éticas y de privacidad inherentes al análisis de big data. Los estudiantes participarán en discusiones y estudios de casos que resaltan las implicaciones sociales de la recopilación y el análisis de datos, incluidas cuestiones de seguridad de datos, derechos de privacidad y uso ético de los datos. Este enfoque holístico garantiza que los graduados no sólo sean técnicamente competentes sino también éticamente conscientes, preparados para navegar las complejidades de la era digital con integridad.

"Introducción al Big Data" está dirigido a estudiantes y profesionales deseosos de profundizar en el campo de la ciencia de datos, así como a cualquier persona interesada en comprender el impacto del big data en la sociedad y los negocios. Al final del curso, los participantes tendrán una base sólida en los principios y técnicas de big data, estarán equipados con las habilidades para analizar grandes conjuntos de datos y contribuir a los procesos de toma de decisiones basados en datos. Este curso allana el camino para estudios adicionales en ciencia y análisis de datos, ofreciendo un trampolín hacia temas avanzados y áreas especializadas dentro del campo.

### **Características clave:**

- **Plan de estudios integral:** el plan de estudios del curso, que abarca una amplia gama de temas, desde los conceptos básicos de la gestión de datos hasta técnicas de análisis avanzadas, está cuidadosamente estructurado para proporcionar una ruta de aprendizaje paso a paso. Comienza con una descripción general de los conceptos de big data y avanza hacia temas más complejos, como algoritmos de aprendizaje automático diseñados para big data y la implementación de soluciones de big data utilizando tecnologías de última generación.
- **Experiencia de aprendizaje práctico:** a través de una serie de ejercicios prácticos, estudios de casos y trabajos de proyectos, los estudiantes aplicarán conceptos teóricos a escenarios del mundo real. Este enfoque práctico garantiza que los participantes no sólo comprendan los fundamentos teóricos de las tecnologías de big data, sino que también adquieran la confianza para aplicar estas herramientas y técnicas en la resolución de problemas complejos.
- **Herramientas y tecnologías de vanguardia:** el curso integra las últimas herramientas de procesamiento de big data y plataformas de análisis, como bases de datos Hadoop, Spark y NoSQL. Los estudiantes obtendrán experiencia de primera mano con estas tecnologías, preparándolos para las demandas de la industria y la investigación.

- **Habilidades relevantes para la industria:** dado que los datos continúan impulsando la toma de decisiones en casi todos los sectores, las habilidades que se enseñan en este curso son muy relevantes en una amplia gama de industrias. Los graduados de este curso estarán bien posicionados para desempeñar funciones que requieran análisis de datos, gestión de datos y toma de decisiones basada en datos.
- **Implicaciones éticas y sociales:** un aspecto único del curso es su enfoque en las consideraciones éticas y de privacidad de big data. Los estudiantes explorarán el impacto de la recopilación y el análisis de datos en la privacidad, la seguridad y la sociedad, preparándolos para abordar estos importantes temas en sus vidas profesionales.
- **Entorno de aprendizaje colaborativo:** el curso promueve un entorno de aprendizaje colaborativo a través de proyectos grupales, sesiones de revisión por pares y foros de discusión. Esto fomenta el intercambio de ideas y perspectivas, enriqueciendo la experiencia de aprendizaje.
- **Instrucción y soporte de expertos:** impartido por instructores con amplia experiencia en tecnologías y aplicaciones de big data, el curso ofrece orientación y soporte expertos. Se proporcionan horas de oficina y recursos adicionales para apoyar el aprendizaje y el trabajo de proyectos de los estudiantes.

Este curso es ideal para estudiantes, profesionales y cualquier persona interesada en comprender los fundamentos de big data y desarrollar las habilidades para analizar grandes conjuntos de datos. Ya sea que su objetivo sea avanzar en su carrera en ciencia de datos, mejorar sus capacidades analíticas o simplemente obtener una comprensión más profunda de cómo los grandes datos están transformando nuestro mundo, este curso ofrece conocimientos valiosos y habilidades prácticas.

## Objetivo del Curso

Al completar con éxito este curso, los estudiantes podrán:

1. Comprender los conceptos clave y los desafíos del Big Data.
2. Identificar y aplicar tecnologías big data apropiadas para el análisis y procesamiento de datos.
3. Utilizar herramientas de análisis de big data para analizar grandes conjuntos de datos.
4. Desarrollar y ejecutar proyectos de big data desde la recopilación de datos hasta el conocimiento.
5. Evaluar las consideraciones éticas y de privacidad en el análisis de big data.

**1. Comprender los conceptos clave y los desafíos de Big Data:** los estudiantes obtendrán una comprensión profunda de lo que constituye Big Data, incluidas sus características definitorias, como volumen, velocidad, variedad y veracidad. Explorarán las

complejidades y los desafíos asociados con la gestión de big data, como los problemas de almacenamiento, el procesamiento de datos y la gestión de la calidad de los datos. Esta base permitirá a los estudiantes apreciar el alcance de los big data y sus implicaciones en diversos dominios.

## **2. Identificar y aplicar tecnologías de Big Data apropiadas para el análisis y**

**procesamiento de datos:** este objetivo se centra en equipar a los estudiantes con el conocimiento y las habilidades para seleccionar e implementar las tecnologías adecuadas para tareas específicas de Big Data. Los estudiantes aprenderán sobre marcos informáticos distribuidos (por ejemplo, Hadoop y Spark), soluciones de almacenamiento de big data (por ejemplo, bases de datos NoSQL) y modelos de procesamiento de datos. Comprenderán las fortalezas y limitaciones de cada tecnología, lo que les permitirá tomar decisiones informadas basadas en los requisitos de sus proyectos.

## **3. Utilizar herramientas de análisis de Big Data para analizar grandes conjuntos de**

**datos:** los estudiantes recibirán capacitación para utilizar una variedad de herramientas y plataformas para el análisis de datos, incluidas herramientas de visualización de datos, software estadístico y bibliotecas de aprendizaje automático. Este objetivo enfatiza la experiencia práctica, y los estudiantes realizan ejercicios prácticos para aplicar estas herramientas en el análisis de conjuntos de datos del mundo real. El objetivo es desarrollar competencias para extraer información útil a partir de conjuntos de datos complejos, una habilidad muy valorada en diversas industrias.

## **4. Desarrollar y ejecutar proyectos de Big Data desde la recopilación de datos hasta**

**las ideas:** este objetivo integral tiene como objetivo guiar a los estudiantes a través del ciclo de vida completo de un proyecto de Big Data. Desde la recopilación inicial de datos y el preprocesamiento hasta el análisis y la obtención de conocimientos, los estudiantes aprenderán a gestionar proyectos de forma eficaz. Esto incluye formular preguntas de investigación, seleccionar metodologías apropiadas, realizar análisis e interpretar los resultados. El trabajo del proyecto fomentará las habilidades de resolución de problemas, la creatividad y la capacidad de comunicar información compleja con claridad.

## **5. Evaluar las consideraciones éticas y de privacidad en el análisis de Big Data:**

un componente esencial del curso, este objetivo aborda el panorama ético de Big Data. Los estudiantes explorarán los dilemas éticos que plantea la recopilación y el análisis de datos a gran escala, incluidas las preocupaciones sobre la privacidad, la seguridad de los datos y el potencial de sesgo en las decisiones basadas en datos. Las discusiones y los estudios de casos fomentarán el pensamiento crítico sobre las responsabilidades de los científicos de datos y el impacto de su trabajo en los individuos y la sociedad.

Al lograr estos objetivos, los estudiantes saldrán del curso con una comprensión completa de big data, equipados con las habilidades técnicas para manejar herramientas y tecnologías de big data, las habilidades de gestión de proyectos para supervisar iniciativas de big data y el marco ético para navegar las complejas cuestiones que rodean la privacidad y la seguridad de los datos. Esta fundación los preparará para estudios adicionales en ciencia de datos o carreras en industrias donde el big data juega un papel fundamental.

# Política de Calificación

[Calificación absoluta:

La calificación del curso se estructura de la siguiente manera para evaluar la comprensión, el compromiso y la aplicación práctica del material del curso por parte de los estudiantes:

1. Asistencia: 20% de la nota final. La asistencia regular es crucial ya que refleja el compromiso y la participación del estudiante en el proceso de aprendizaje.
2. Tareas y Pruebas: 50% de la nota final. Este componente integral evalúa la comprensión de los estudiantes del material del curso a través de tareas y pruebas regulares. Estas tareas están diseñadas para reforzar el aprendizaje y garantizar un compromiso continuo con el contenido del curso.
3. Examen Parcial: 15% de la nota final. El examen de mitad de período evalúa la comprensión de los estudiantes de los conceptos y principios clave cubiertos en la primera mitad del curso.
4. Examen Final: 15% de la nota final. El examen final se centra en todo el contenido del curso, poniendo a prueba la comprensión general de los estudiantes y su capacidad para integrar diferentes temas aprendidos a lo largo del curso.
5. Crédito adicional: Se pueden otorgar puntos adicionales a la calificación final según las contribuciones excepcionales del estudiante al curso. Esto incluye participación activa, trabajo en proyectos innovadores o cualquier esfuerzo adicional que mejore la experiencia de aprendizaje para ellos y sus compañeros.

Esta política de calificación está diseñada para evaluar de manera justa el desempeño de los estudiantes en diferentes aspectos del curso, fomentando el esfuerzo constante, la participación activa y una comprensión profunda de la materia.]

## Libros de Texto y Otros Materiales

[No se especifica ningún libro de texto principal.]

[Se proporcionarán materiales del curso]

[Libro de referencia]

Bibliografía: Título, Autor(es), Año de publicación, Editorial.

La imagen de la portada del libro se puede insertar aquí.

# Tareas y Exámenes

Incorporar tareas, exámenes y proyectos en el curso "Introducción al Big Data" es fundamental para evaluar la comprensión de los estudiantes, reforzar el aprendizaje y aplicar los conocimientos teóricos a situaciones prácticas. Así es como puedes estructurar estos componentes:

## Tareas

Objetivo: Las tareas están diseñadas para reforzar los conceptos discutidos en conferencias y lecturas, brindando a los estudiantes experiencia práctica en análisis de datos, uso de herramientas de big data y habilidades de resolución de problemas.

Estructura:

- Conjuntos de problemas semanales: cubriendo aspectos teóricos y prácticos de los temas semanales. Estos pueden incluir tareas de manipulación de datos, ensayos breves sobre los desafíos de big data y ejercicios con herramientas de big data como Hadoop o Spark.
- Tareas de análisis de datos: los estudiantes recibirán conjuntos de datos para limpiar, analizar y obtener información, utilizando herramientas como la biblioteca pandas de Python, Apache Spark o consultas SQL en bases de datos NoSQL.
- Evaluación: Las tareas se calificarán en función de la precisión, la creatividad en la resolución de problemas y la capacidad de aplicar conocimientos teóricos a tareas prácticas. Los comentarios se centrarán en mejorar la comprensión y la aplicación de las tecnologías de big data.

## Exámenes parciales y finales

Objetivo: Los exámenes evalúan la comprensión del material del curso por parte de los estudiantes, incluida la comprensión de conceptos clave, la capacidad para analizar problemas de big data y el conocimiento de las tecnologías y herramientas apropiadas para abordar estos problemas.

Estructura:

- Examen Parcial: una combinación de preguntas de opción múltiple, respuesta corta y resolución de problemas que cubren la primera mitad del curso. Esto puede incluir preguntas sobre los fundamentos de big data, análisis de datos básicos y uso introductorio de herramientas de big data.
- Examen Final: es comprensivo, abarca todo el curso. Incluye una combinación de tipos de preguntas, con un enfoque en la integración del conocimiento en diferentes temas, como la elección de tecnologías apropiadas para escenarios de big data determinados y consideraciones éticas en el análisis de datos.

- Evaluación: los exámenes se calificarán según la profundidad de comprensión, la precisión de las respuestas y la capacidad de integrar y aplicar los conceptos del curso a situaciones nuevas.

## Proyecto final

Objetivo: El proyecto final permite a los estudiantes aplicar lo que han aprendido en un escenario integral del mundo real. Los proyectos implicarán la recopilación, el procesamiento, el análisis y la presentación de resultados de datos, haciendo hincapié en el desarrollo de una solución completa de big data.

Estructura:

- Propuesta: los estudiantes envían una propuesta de proyecto que describe el problema elegido, las fuentes de datos, el plan de análisis y los resultados esperados.
- Informes de Progreso: actualizaciones periódicas para realizar un seguimiento del progreso, los desafíos encontrados y los resultados preliminares.
- Presentación: una presentación formal del proyecto a la clase, destacando el problema, la metodología, el análisis y los conocimientos derivados.
- Informe Final: un informe detallado que incluye antecedentes de investigación, metodología, análisis, resultados y conclusiones. El informe también debe reflejar el proceso del proyecto y cualquier consideración ética abordada.
- Evaluación: los proyectos se evaluarán en función de su originalidad, complejidad, aplicación de técnicas de big data, profundidad del análisis, calidad de los conocimientos y presentación general. Se pondrá especial énfasis en la aplicación práctica de herramientas y tecnologías, así como en la capacidad de superar desafíos típicos de proyectos de big data.

La incorporación de estos componentes en el curso garantiza una evaluación equilibrada de los conocimientos teóricos y las habilidades prácticas de los estudiantes, preparándolos para los desafíos del mundo real en el campo de big data.

# Actividad del Curso

Incorporar actividades diversas y atractivas en un curso de "Introducción a Big Data" es esencial para mejorar el aprendizaje, fomentar la participación y brindar experiencia práctica. A continuación se presentan descripciones detalladas de diversas actividades del curso diseñadas para alcanzar estos objetivos:

## 1. Clases Interactivas

Objetivo: Introducir y explicar conceptos, tecnologías y desafíos centrales en big data. Las conferencias están diseñadas para ser interactivas, fomentando preguntas, debates y comentarios en tiempo real.

Actividades:

- Oradores invitados: invitar a expertos o investigadores de la industria a compartir conocimientos sobre tendencias actuales, tecnologías y aplicaciones de big data del mundo real.
- Encuestas y cuestionarios en vivo: utilice herramientas como Kahoot o Poll Everywhere durante las conferencias para evaluar la comprensión, fomentar la participación y brindar comentarios instantáneos sobre conceptos clave.

## 2. Laboratorios Prácticos

Objetivo: Proporcionar experiencia práctica con herramientas y técnicas de big data. Los laboratorios están estructurados para complementar los temas de las conferencias, lo que permite a los estudiantes aplicar lo que han aprendido en un entorno controlado.

Actividades:

- Ejercicios de procesamiento de datos: utilizar conjuntos de datos para practicar la limpieza, manipulación y procesamiento de datos utilizando bases de datos Hadoop, Spark o NoSQL.
- Talleres de visualización: enseñar a los estudiantes a usar herramientas de visualización (por ejemplo, bibliotecas Tableau, PowerBI o Python como Matplotlib y Seaborn) para crear visualizaciones de datos interesantes.

## 3. Proyectos Grupales

Objetivo: Desarrollar el trabajo en equipo, la gestión de proyectos y las habilidades técnicas trabajando en un proyecto integral de big data desde su concepción hasta su presentación.

Actividades:

- Reuniones de inicio del proyecto: los grupos discuten ideas, objetivos y planificación del proyecto con el instructor para garantizar la viabilidad y relevancia.
- Registros semanales: reuniones breves con el instructor para discutir el progreso, los desafíos y recibir orientación.
- Revisiones de pares: los grupos presentan su progreso a la clase para recibir comentarios, fomentando un entorno de aprendizaje colaborativo.

## [4. Estudios de casos y sesiones de discusión

Objetivo: Analizar los desafíos y soluciones de big data del mundo real, fomentando el pensamiento crítico, la resolución de problemas y las consideraciones éticas.

Actividades:

- Análisis de estudios de casos: grupos o individuos analizan estudios de casos de aplicaciones de big data, centrándose en la tecnología utilizada, los desafíos enfrentados y los resultados obtenidos.
- Mesas redondas: organice debates sobre temas como la privacidad de los datos, las implicaciones éticas de la recopilación de datos y el futuro del big data, fomentando perspectivas diversas y el pensamiento crítico.]

## **[5. Talleres sobre herramientas y tecnologías de big data**

Objetivo: Proporcionar capacitación en profundidad sobre tecnologías específicas de big data, mejorando la competencia técnica y la familiaridad con las herramientas estándar de la industria.

Actividades:

- Talleres específicos de software: sesiones centradas en Hadoop, Spark o bases de datos NoSQL específicas, dirigidas por usuarios experimentados o representantes de los proveedores de software.
- Cuadernos de ciencia de datos: utilizar cuadernos de Jupyter o Google Colab para ejercicios de programación en Python, integrando la teoría con la práctica y promoviendo habilidades de codificación en análisis de datos y aprendizaje automático.]

## **[6. Foros en línea y aprendizaje colaborativo**

Objetivo: Facilitar el debate continuo, el apoyo y el intercambio de recursos fuera del horario de clase, aprovechando la tecnología para mejorar el aprendizaje y la comunidad.

Actividades:

- Foros de discusión: utilizar plataformas como Canvas o Moodle para que los estudiantes publiquen preguntas, compartan recursos y discutan temas o tareas de clases.
- Toma de notas colaborativa: fomentar el uso de documentos compartidos (por ejemplo, Google Docs) para la toma de notas colaborativas y guías de estudio, promoviendo un sentido de comunidad y aprendizaje compartido.]

Al integrar estas actividades en el plan de estudios, el curso "Introducción a Big Data" se convierte en una experiencia de aprendizaje dinámica y atractiva. Los estudiantes no solo obtienen una comprensión profunda de los principios y tecnologías de big data, sino que también desarrollan habilidades prácticas, pensamiento crítico y una mentalidad colaborativa esenciales para el éxito en el campo.

# Cronograma del Curso

[Contenidos obligatorios del curso]

[SE DEBEN agregar más contenidos que los contenidos obligatorios del curso]

1. Tecnologías de Big Data
2. Análisis y Minería de Datos
3. Almacenamiento de Big Data
4. Aprendizaje Automático con Big Data
5. Herramientas de Análisis de Big Data
6. Consideraciones Éticas y Tendencias Futuras

Semana	Clase
1	
2	
3	
4	Examen Parcial / Proyecto
5	
6	
7	
8	Examen Final / Proyecto

# Contenidos del Curso

**[DEBE actualizarse con otros contenidos]**