

SISTEMA NACIONAL DE CAPACITACION DISEÑO DE LA ACTIVIDAD

Nombre

INTRODUCCIÓN A CIENCIA DE DATOS (NIVEL 1)

Código INAP IN41037/24 **Estado** Activo

Programa)Actividades Transversales **Área** Sistemas, procesos y tecnologías

Fundamentación

Subtema: Gestión administrativa, de la información y los datos.

La ciencia de datos es la rama de la computación que se encarga de entender, procesar y extraer valor a partir de los datos. Hoy en día existen innumerables fuentes de datos y herramientas para recolectar datos, desde el log de un website hasta la estructura de una Red Social pasando por bases de datos en todo tipo de formatos y pertenecientes a los más diversos dominios. Particularmente el sector público recolecta grandes cantidades de datos de las personas u organismos del país. Los censos, pagos, impuestos, declaraciones juradas, formularios, entre otros, son elementos que contienen una gran cantidad de información. Ahora la pregunta es: ¿Qué se hace o qué se puede hacer con esos datos? Este curso da una introducción a la ciencia de datos basándose en los fundamentos de cómo los datos se pueden entender, procesar y extraer valor orientado en los entes públicos. Además, enseña a cómo utilizar las herramientas para que cada participante pueda introducirse en el mundo de las ciencia de datos.

Contribución esperada

- Potenciar las capacidades y las competencias de los agentes públicos en relación con la Ciencia de Datos.
- Promover el uso de herramientas útiles para el manejo de datos y la toma de decisiones.

Perfil del participante

Trabajadores de la administración pública nacional.

Objetivos

Se espera que los participantes logren:

- Identificar los temas teóricos y prácticos que engloba la Ciencia de Datos.
- Comprender cómo encarar un proyecto de Ciencia de Datos.
- Identificar herramientas útiles para el manejo de datos.
- Comprender el panorama actual de los algoritmos de predicción y recomendación.

Contenido

Unidad 1: Fundamentos de la Ciencia de Datos

Introducción a la Ciencia de Datos
Fundamentos de la Ciencia de Datos
Datos en los organismos públicos y privados
Casos de éxito y fracaso de la Ciencia de Datos en el mundo.

Unidad 2: Proyectos de Ciencia de Datos
Cómo abordar un proyecto de Ciencia de Datos
Regresión, Clasificación, Sistemas de Recomendación e Identificación de patrones
Tópicos de la Ciencia de Datos en la Administración Pública
Obtener y depurar los datos
Detectar y eliminar anomalías
Formateo de los datos

Unidad 3: Visualización
Elementos de la Visualización
Objetivos
Principios básicos
Principios de Tufte
Gráficos discretos vs continuos.
Gráficos 3D
Manipulación malintencionada de los gráficos en la política

Unidad 4: Proyecto de Ciencia de Datos
Paso a paso cómo abordar un proyecto de ciencia de datos
Introducción a Machine Learning
Manipulación de datos
Práctica de formateo de datos
Visualizaciones para el análisis exploratorio.
Visualizaciones para la comunicación de resultados
Entrenamiento de varios algoritmos de Machine Learning.

Estrategias metodológicas y recursos didácticos

Al ingresar en el aula virtual los participantes encontrarán una bienvenida y la presentación general del curso, que permitirá navegar por el programa con los objetivos, los contenidos, la secuencia de recursos y actividades y los requisitos de aprobación. También se podrá acceder a la Guía del Participante, que brindará orientaciones para que los participantes puedan recorrer el entorno virtual sin dificultades.

El curso se estructurará a lo largo de cuatro unidades. Cada clase abordará una unidad de contenido temático y se desarrollará a partir de la presentación de un audiovisual grabado por el equipo docente.

Para la presentación y el desarrollo de los temas, se adoptarán estrategias que posibiliten un diseño teórico - práctico amigable, inclusivo, dinámico y dialogado para captar la atención e interpelar en todo momento, a los participantes, respecto de sus conocimientos y experiencias previas vinculadas con la temática del curso, a la vez que les permita tener un rol activo a lo largo del curso AG.

En la primera parte de cada presentación se expondrán aspectos conceptuales que orienten y promuevan la comprensión de cada uno de los temas previstos. A continuación, se presentarán algunos ejemplos prácticos o configuraciones en entornos reales para una mejor comprensión de los temas expuestos.

Además de los videos de las clases se facilitarán otros materiales didácticos y recursos digitales complementarios en diversos formatos: infografías, presentaciones en Power Point, recursos disponibles en Internet, realizando el correspondiente anclaje con la teoría propuesta.

Dependiendo del tema de cada unidad didáctica, se utilizarán algunos de los siguientes recursos:

Diagramas de arquitectura

Diagramas de despliegue

Diagramas de secuencia

Ejemplos de código fuente en diferentes lenguajes de programación

Ejemplos de utilización de programas, tales como IDEs (entornos integrados de desarrollo), navegadores, etc.

En cada una de las unidades estará disponible el material teórico para abordar los temas desarrollados durante las clases. Sobre la base de los recursos disponibles en el aula virtual y del material de lectura obligatoria provisto por el equipo docente, los participantes deberán resolver un cuestionario autoadministrado orientado a monitorear los avances en su propio proceso de aprendizaje.

Al finalizar el curso y como actividad final integradora, los participantes deberán resolver un cuestionario de carácter obligatorio.

Descripción de la modalidad

Virtual autogestionado.

Bibliografía

Unidad 1

Bibliografía Obligatoria

Loiseau, Matías (2022). Apunte de clase: Introducción a la Ciencia de Datos.

Bibliografía Opcional

Wikipedia (2022). Ciencia de datos. Disponible en:

https://es.wikipedia.org/wiki/Ciencia_de_datos Verificado: 10/08/2022

Oracle (2022). ¿Qué es la Ciencia de Datos?. Disponible en:

<https://www.oracle.com/ar/what-is-data-science/> Verificado: 10/08/2022

Unidad 2

Bibliografía Obligatoria

Loiseau, Matías (2022). Apunte de clase: Introducción a la Ciencia de Datos.

Unidad 3

Bibliografía Obligatoria

Loiseau, Matías (2022). Apunte de clase: Introducción a la Ciencia de Datos.

Unidad 4

Bibliografía Obligatoria

Loiseau, Matías (2022). Apunte de clase: Introducción a la Ciencia de Datos.

Evaluación de los aprendizajes

De proceso: Realización de una actividad de evaluación por unidad de carácter obligatoria, con preguntas de opciones múltiples en diversos formatos, para recuperar los conceptos/prácticas centrales de cada clase y orientado a monitorear los avances en el propio proceso de aprendizaje.

De producto: Resolución de un cuestionario final integrador, de carácter obligatorio, con preguntas de opciones múltiples en diversos formatos y distintas a las planteadas en la evaluación de proceso.

Instrumentos para la evaluación

a) Evaluación de los aprendizajes:

Tanto para las evaluaciones de proceso como de producto, se utilizarán:

Cuestionarios de preguntas cerradas de opción múltiple en diversos formatos.

b) Evaluación de la actividad:

Encuesta de satisfacción de los participantes mediante un cuestionario semiestructurado.

Requisitos de Asistencia y aprobación

Aquellos participantes que deseen un certificado de aprobación del curso, deberán:

- Realizar y aprobar en tiempo y forma las 4 actividades de evaluación propuestas (una por unidad) con, al menos, el 60% de las respuestas correctas.
- Realizar y aprobar en tiempo y forma la actividad de evaluación integradora con, al menos, el 60% de las respuestas

correctas. El número máximo de intentos para aprobar el cuestionario será de 3.

Duración (Hs.)

12

Detalle sobre la duración

12 horas distribuidas a lo largo de tres semanas.

3 horas, correspondientes a las 4 clases virtuales, disponibles en material audiovisual (30 minutos de duración, aproximadamente, de cada audiovisual).

6 horas de lectura de bibliografía obligatoria y realización de actividades complementarias. (Una hora y media, aproximadamente, para cada unidad).

3 horas para la resolución de 5 evaluaciones autoadministradas. (1 cuestionario por unidad de 30 minutos y 1 cuestionario final integrador de una hora, aproximadamente).

Lugar

Campus INAP - <https://campus.inap.gob.ar>

Perfil Instructor

Profesional o estudiante avanzado en sistemas, informática o tecnología con experiencia en docencia y nivel senior en desarrollo de software full stack.

Información del docente:

Ing. Matías Loisau.

CV Resumido: Ingeniero en Informática. Actualmente trabajando como Científico de Datos para la Dirección de Ciencia de Datos de la subdirección de sistemas de la Administración Federal de Ingresos Brutos (AFIP) de la República Argentina.

Docente e investigador universitario en programación de Sistemas Embebidos.

Origen de la demanda

INAP

Prestadores Docentes

CUIT/CUIL	APELLIDO Y NOMBRE
	ACTIVIDAD,AUTOADMINISTRADA