

Nombre de la actividad formativa

Análisis de datos con JUPYTER NOTEBOOK

Modalidad	Estrategia metodológica	Tipo de capacitación	Duración
Por participación	Capacitación no presencial-virtual *	Curso	16.0 horas efectivas **

* Sesiones telepresenciales y simultáneas a través de la Plataforma Zoom.

** Cuatro sesiones semanales y simultaneas de 3 horas y 1 hora semanal de trabajo a distancia.

Descripción: Jupyter Notebook es un entorno de trabajo interactivo que permite desarrollar código, principalmente en Python. Su capacidad para combinar código, texto elementos multimedia en un mismo documento lo convierten en la herramienta ideal para llevar la práctica del análisis de datos al siguiente nivel. Este taller introduce a los participantes en las aplicaciones de Jupyter para el análisis profesional de datos. Comenzando desde cómo reemplazar eficazmente operaciones de manejo de datos inconvenientemente llevadas a cabo a través de Excel u herramientas similares, hasta la implementación de algoritmos de Minería de Datos.

Dirigido a: Usuarios regulares de información cuantitativa, con un dominio alto de Excel o con conocimientos básicos de programación, que deseen dar el salto a técnicas y metodologías de análisis de datos más sofisticadas.

Requisito: Los participantes deberán trabajar con su computadora personal con sistemas operativos Windows o Mac. Visitar el siguiente enlace: <https://www.anaconda.com/distribution/#download-section> y descargar Python 3.7 según el sistema operativo correspondiente (el archivo es pesado por lo que la descarga puede demorarse inclusive hasta una hora dependiendo de la velocidad de la conexión). Solamente es necesario descargar el archivo, las indicaciones de instalación se discutirán durante la primera sesión del taller.

Objetivo general

Gestionar apropiadamente todas las fases que comprenden el análisis de un conjunto de datos a través de las funcionalidades ofrecidas por un Jupyter Notebook.

Objetivos específicos	Contenido temático	Actividades a realizar (tarea semanal)
<p>Semana 1</p> <p>Explorar el entorno de programación de los JUPYTER NOTEBOOKS</p>	<ol style="list-style-type: none"> 1. ¿Qué son los Jupyter Notebooks? 2. Ventajas y desventajas con respecto a otras herramientas 3. Instalación y configuración 4. Adquisición de datos 5. Manipulación de datos 	<p>En clase (3 horas)</p> <ol style="list-style-type: none"> a. Presentación de contexto y conceptos b. Instalar y configurar un Jupyter Notebook c. Ejercicio de exploración de funcionalidades básicas de un Jupyter Notebook

**ESTE DOCUMENTO Y SU CONTENIDO ES PROPIEDAD INTELECTUAL DEL
COLEGIO DE CIENCIAS ECONÓMICAS DE COSTA RICA**

Objetivos específicos	Contenido temático	Actividades a realizar (tarea semanal)
		Tarea (1 hora) Búsqueda y construcción de un conjunto propio
Semana 2 Diseñar flujos completos de análisis de datos en JUPYTER NOTEBOOK (foco en tabulación y visualización)	<ol style="list-style-type: none"> 1. Lectura del conjunto de datos 2. Operaciones de limpieza Operaciones de tabulación y visualización 	En clase (3 horas) a. Ejercicio de exploración de datos: estadísticas policiales Tarea (1 hora) Análisis exploratorio del conjunto propio
Semana 3 Diseñar flujos completos de análisis de datos en JUPYTER NOTEBOOK (foco en minería de datos)	<ol style="list-style-type: none"> 1. Lectura del conjunto de datos 2. Operaciones de limpieza 3. Operaciones de tabulación y visualización 4. Operaciones de análisis (Minería de datos) 	En clase (3 horas) a. Ejercicio de exploración de datos: contrataciones públicas Tarea (1 hora) Tabulación y visualización del conjunto propio
Semana 4 Diseñar flujos completos de análisis de datos en JUPYTER NOTEBOOK (foco en reporte y comunicación de resultados)	<ol style="list-style-type: none"> 1. Lectura del conjunto de datos 2. Operaciones de limpieza 3. Operaciones de tabulación y visualización 4. Operaciones de análisis (Minería de datos) 5. Operaciones de reporte y comunicación de resultados 	En clase (3 horas) a. Ejercicio de exploración de datos: exportaciones Tarea (1 hora) Interpretación de resultados del conjunto propio

Facilitador

Dagoberto Herrera Murillo

Científico de datos, docente universitario y consultor en proyectos de inteligencia competitiva.

Cuenta con una maestría en Big Data Management & Analytics con mención Cum Laude y una especialidad en minería de procesos. Las cuales fueron impartidas conjuntamente por la Universidad Libre de Bruselas, la Universidad Politécnica de Cataluña y la Universidad Tecnológica de Eindhoven. Graduado con honores como Ingeniero en Negocios y Tecnologías de la Información por el Tecnológico de Monterrey en México.