



TE INVITAMOS A NUESTROS CURSOS DE:



SPARK – SPARK SQL

(CON PYSPARK Y AWS)

MODALIDADES: PRESENCIAL y ONLINE (LIVE STREAMING)

Consulta nuestros SERVICIOS en: <https://rhcecam.com>

ANTECEDENTES

En los últimos años la tecnología ha crecido a pasos agigantados y con ella la cantidad de datos que se generan día con día. Redes sociales, IoT, Deep Learning y bases de datos, entre otros, han pasado a ser parte de los activos más importantes de las empresas, a través de los datos se pueden conocer comportamientos, tendencias, áreas de oportunidad, etc. en los diferentes negocios. Esto nos plantea el paradigma de lograr generar el máximo valor posible con estos datos ya que, no todas las herramientas que se usaban tradicionalmente tienen el poder suficiente para manipular la masiva cantidad de información a la que se tiene acceso y con esto aprovechar su máximo potencial.

En la actualidad una de las herramientas con mayor uso en temas de **Big Data** es **Spark** cuya estructura de trabajo en paralelo ha generado una revolución en las áreas de analítica, si a esto sumamos las ventajas de un programa de alto nivel como Python se nos abre un mundo de posibilidades para generar análisis de alto valor.

OBJETIVOS

Este curso está diseñado para:

Aprender a manipular los módulos más importantes de Spark con SQL y Python (Pyspark) para generar análisis y modelado de datos.

- Manipular grandes cantidades de datos con el lenguaje SQL dentro de Spark.
- A través de Spark Context crear RDDs (Resilient Distributed Datasets) que es la estructura base de spark y otras herramientas para la manipulación y análisis de datos.

Te invitamos a formar parte de lo que será una experiencia inolvidable, en donde aprenderás a explotar la volumetría de datos y generar valor a través de la Data. Un curso guiado por un profesional experto que ha trabajado con millones de datos en instituciones financieras y que es parte de nuestro equipo, quien ha impartido en varias ediciones y con mucho éxito los cursos de programación en R, Python y Machine Learning.

Perfil:

El curso está dirigido a

Profesionistas de las licenciaturas de Actuaría, Ingeniería, Computación, Informática, Estadística, Matemáticas, Física, Economía, licenciaturas afines y público en general, que desea incursionar en áreas de Big Data, Ciencia de Datos, BI, o que trabaja en su día a día con muchos datos.

Ponente:

Andrés Octavio: Actuario y coordinador en Bancomer: forma parte de un grupo de Científicos de Datos en el área de ingeniería con tareas encaminadas a desarrollar modelos predictivos para las diferentes Subdirecciones. Trabaja y explota la volumetría de las bases con Spark y modelos en Pyspark, Python y R (Árboles y Bosques de decisión, regresión logística, Perceptrón Multicapa, Cluster y Análisis de componentes principales, entre otros). Cuenta con experiencia docente y forma parte de nuestro equipo RHCECAM, impartiendo con éxito los cursos de R, Python y Machine Learning.

PREREQUISITOS:

PREREQUISITOS:

- Se recomienda tener conocimientos básicos de Python y SQL
- Nociones básicas de conceptos estadísticos como medias, varianza y probabilidad (casos favorables contra casos totales)
- Deseable: Procesador Intel i5 o superior (O su equivalente) y 8GB o más de memoria RAM.

Reserva en este mes y obtén **GRATIS** el curso grabado de **PYTHON BÁSICO**

MÓDULO I: SPARK CON RDDs

Duración 15 horas.

Spark Context es la puerta de entrada al mundo de spark, en este módulo se transmiten los conocimientos para crear y manipular RDDs (Resilient Distributed Datasets) que es la estructura base de spark y de otras herramientas para la manipulación y análisis de datos y la importancia del cómputo en paralelo.

Levantamiento de máquina EC2 en AWS
Instalación de PySpark en una máquina EC2
Introducción a Spark
Creación de RDDs
Carga de datos
Operaciones con RDDs

Map, reduce, filter, union, intersection, group, distinct

KeyFunctions

Persistencia

Variables Broadcast

PRÁCTICA: Construcción de Wordcloud en tiempo real.

Módulo II: SPARK SQL

Duración 15 horas.

En este módulo se transmiten los conocimientos para manipular bases (selección, filtrado, ordenamiento, agrupación y sumariado) con el lenguaje SQL dentro de spark y con las funciones que pyspark incorpora para estas tareas, además de generar cruce de tablas y agrupaciones especiales conocidas como funciones de ventana.

Definición y creación de dataframes en pyspark

Tipos de datos

Datos especiales:

Manejo de fechas y horas en pyspark

Transformación de tipos de datos

Generando consultas SQL con pyspark

Funciones con dataframes

Windows functions

Join's con dataframes

Variables Broadcast

Inversión, Promociones y Fechas

***PROMOCIONES** Vigentes
DURANTE ESTE MES

MODALIDAD	PRECIO NORMAL POR MÓDULO y/o 3 MSI con TDC	DESCUENTO (Reserva con un pago mínimo y asegura tu PROMOCIÓN)
LIVE STREAM	MXN: 5,499 + IVA USD: 360	20% DE DESCUENTO por Módulo y 3 meses sin intereses
PRESENCIAL	MXN: 6,499 + IVA	30% DE DESCUENTO al elegir ambos Módulos o 25% de descuento y 3 MSI.

* Promociones aplicables reservando lugar con un pago mínimo de \$1,000 MXN antes de la fecha mencionada/mes. / Pagos con TDC aplica comisión del 5%

SPARK CON RDDs	SPARK SQL	HORARIOS POR MÓDULO
Inicio: 25 de OCTUBRE	Inicio: 22 de NOVIEMBRE	Domingos 9:00 – 12:45 HRS.

Informes

DESCUENTOS Y PROMOCIONES NO ACUMULABLES.

Precios por persona, precios en MXN agregar IVA del 16%. Para facturación enviar datos fiscales por correo únicamente en el mes de inscripción.

Pagos con cargo a tarjeta de crédito aplica comisión del 5% + IVA.

Cursos ONLINE: Cursos con transmisión en VIVO, en las fechas programadas. (Zona horaria C. de México: A partir de 05-abr UTC-0500, a partir de 25-oct UTC-0600)

Cancelaciones: Si el curso es cancelado por RHCECAM se reembolsa inversión total, en otro caso el participante podrá reprogramar su participación, cambiar de curso o ceder su lugar a un tercero.

Inscripciones

1. **Reservar lugar** mediante transferencia de fondos, depósito bancario, depósito en tiendas de conveniencia y OXXO 's, con tarjeta de crédito a través de PayPal
2. **Realizar pago complementario** con dos días de anticipación previo al inicio del curso y enviar confirmación del mismo vía correo
3. El(los) pago(s) se podrá(n) realizar a través de la(s) siguiente(s) cuenta(s).

Solicita la información de pago AQUÍ o por correo

Enviar confirmación de reservación, pago total o complementario vía email a: informes@rhcecam.com anexando foto de identificación oficial INE/IFE u otra, ficha(s) de depósito(s) bancario y relación de(l) participante(s) [Nombre completo, correo, número de teléfono y licenciatura]

Para más información del curso, nos pueden contactar a:

Facebook: www.facebook.com/Rhcecam

Correo:



informes@rhcecam.com

Teléfonos:



5548663933



5548663933

