

Programas Asíncronos

SÍLABO

I. DATOS GENERALES

Curso	:	Scala Fundamentals
Área	:	Big Data
Docente	:	Marcos Santa Cruz V.
Duración	:	4 horas académicas

II. SUMILLA

El curso se desarrollará mediante la metodología asíncrona, que comprende la revisión de los videos grabados, las lecturas y la resolución de la evaluación planteada. Le permitirá al participante conocer los fundamentos de la programación empleando el lenguaje Scala, los diversos componentes disponibles para el manejo de Dataset, así como la forma de integrarlo con Spark para el procesamiento de grandes volúmenes de datos. El curso comprende los siguientes temas generales: a) Motivación y setup de entorno, b) Programación Básica en Scala I y c) Programación Básica en Scala II y d) Introducción a Spark RDD y Spark SQL.

III. REQUISITOS

- Conocimiento de programación básica.

IV. DIRIGIDO A

- Profesionales, egresados, estudiantes e interesados en conocer los fundamentos de la programación empleando el lenguaje Scala, los diversos componentes disponibles para el manejo de conjuntos de datos, así como la forma de integrarlo con Spark para el procesamiento de grandes volúmenes de datos.

V. OBJETIVOS – LOGROS

El curso le permitirá al participante:

- Conocer los fundamentos de la programación en el lenguaje Scala y configurar un entorno de ejecución amigable.

- Escribir código orientado a objetos con Scala y aprender a compilarlos.
- Entender cómo funciona el ecosistema Spark.
- Utilizar Spark para el procesamiento en paralelo de grandes volúmenes de datos.

VI. METODOLOGÍA

Los contenidos se desarrollarán siguiendo la secuencia teoría – práctica con evaluaciones modulares, utilizando videos grabados y lecturas como recursos didácticos.

VII. CONTENIDOS

Total horas	4 horas
Sesión 1	
• Tema 1: Motivación para el uso de Scala - Spark.	10 min.
• Tema 2: Instalación de JDK.	5 min.
• Tema 3: Instalación del Scala y Sbt	10 min.
• Tema 4: Instalación de IDE: IntelliJ	10 min.
• Tema 5: Mi primer programa en Scala Repl	5 min.
• Tema 6: Tipos de datos y operadores	20 min.
• Tema 7: Estructuras condicionales	10 min.
• Tema 8: Estructuras repetitivas	10 min.
• Tema 9: Programación Funcional	10 min.
Sesión 2	
• Tema 10: Estructuras de datos (Tupla, List, Array y Map)	30 min.
• Tema 11: Introducción a clases y objetos	10 min.
• Tema 12: Constructores Primarios y Secundarios	10 min.
• Tema 13: Objetos Singleton	10 min.
• Tema 14: Herencia y Objetos compañeros	10 min.
• Tema 15: Compilación de archivos scala	5 min.
• Tema 16: Entendimiento del ecosistema Spark	20 min.
• Tema 17: Resilient Distributed Datasets	30 min.
• Tema 18: Spark SQL (Dataframes y Datasets)	15 min.

VIII. SISTEMA DE EVALUACIÓN

La nota final del curso se obtendrá a partir de los siguientes rubros de evaluación y pesos asociados:

Rubro	Peso
Evaluación	100%

IX. Docente**Marcos Santa Cruz Vásquez**

Resumen

- Ingeniero de Sistemas de la UNI con un MBA especializado en Big Data
- Estudios complementarios de Ciencia de Datos y Estadística en el MIT y Universidad de San Diego.
- Más de 8 años de experiencia en temas de Data & Analytics trabajando modelos de machine learning con Big Data, liderando equipos multidisciplinarios.
- Experiencia liderando proyectos de procesamiento de datos digitales y transaccionales de gran escala, en diferentes tecnologías cloud.