

# DIPLOMADO EN DATA ANALYTICS

**Universidad del Valle de Guatemala (UVG) es la institución pionera en el campo de business intelligence, analytics y data en la región centroamericana.** Con el lanzamiento de la Maestría en Business Intelligence and Analytics en 2013 y el inicio de la Ingeniería en Ciencia de los Datos en 2020, UVG se ha posicionado como referente de educación superior en el tema a nivel nacional e internacional.

El Programa de Educación Continua (Educon) de UVG ofrece cursos libres dentro del campo de analytics y data science, los cuales han capacitado a cientos de personas durante los últimos 10 años. Uno de ellos es el Diplomado en Data Science, que ofrece a interesados de distintos perfiles la oportunidad de un primer acercamiento a la materia para entender uno de los procesos de transformación digital más importantes actualmente.

El diplomado se conforma de ocho módulos que proporcionan a los participantes conocimientos sólidos sobre la ciencia de datos para luego aplicarlos en el mundo empresarial. A lo largo del programa, se emplean las herramientas computacionales para el procesamiento y análisis de datos que cualquier científico de datos debe dominar.

**EDUCON UVGMASTERS**  
EDUCACIÓN CONTINUA **THINK FORWARD >>**

# MÓDULOS

## 1 R para data science:

R es el lenguaje más utilizado a nivel mundial para el análisis estadístico y la ciencia de datos. Se considera la lengua franca del mundo científico en lo que respecta al análisis, desarrollo de modelos, visualización y comunicación de datos de todo tipo. Este módulo introduce el uso de las principales funciones del ecosistema de R para realizar análisis exploratorios y desarrollar modelos descriptivos y predictivos.

## 2 Python para data science:

Aborda los fundamentos del lenguaje de programación Python en su versión 3.7 o más reciente para utilizar la herramienta en tareas de manipulación y análisis de datos. Trata temas concernientes a la manipulación de datos, la lectura de datos de diferentes fuentes (archivos de texto, bases de datos, hojas de cálculo, entre otros) y su preparación para el proceso analítico, las estructuras de datos fundamentales para el análisis de datos, entre otros.

Indaga en los principales módulos / librerías especializadas para el análisis de datos, desarrollo de modelos y visualización.

## 3 SQL para data science:

Explora los conceptos básicos del lenguaje SQL, la herramienta estándar a nivel mundial para interactuar con Sistemas de Gestión de Bases de Datos Relacionales (RDBMS, por sus siglas en inglés). El curso utiliza la visión de un analista de datos y no un administrador de base de datos (DBA), por lo que su objetivo es la extracción, procesamiento y limpieza de datos almacenados en bases de datos relacionales.

Se cubren las principales y más importantes estructuras para realizar queries y manipulación de datos.

## 4 Análisis exploratorio de datos y modelos descriptivos para data science:

Aborda la exploración de datos tanto en forma analítica como gráfica, explorando modelos univariados para datos cualitativos y cuantitativos, relaciones entre dos o más variables y el análisis multivariado. Se explican los conceptos básicos y las herramientas de análisis descriptivo, además de los tipos de datos, su visualización, entre otros.





## 5 Modelos predictivos, pronósticos y series de tiempo:

Explica el desarrollo y evaluación de modelos predictivos causales y acausales como: regresión logística, regresión multivariada y series de tiempo para la predicción de variables cualitativas y cuantitativas.

## 6 Machine learning / Data mining I:

Aborda los principales modelos de aprendizaje automatizado para data science y data mining utilizando las principales librerías de R y Python. Se utilizan modelos de aprendizaje supervisado para clasificación y predicción, como k-NN o Naive Bayes, así como modelos de aprendizaje no supervisado, incluyendo clusters y modelos jerárquicos.

## 7 Machine learning / Data mining II:

Continuación del curso anterior. Presenta modelos más avanzados de aprendizaje automático para clasificación, reducción de dimensionalidad, entre otros. Se incluye el uso de componentes principales, support vector, machines de ensamblaje, bagging, random forests, entre otros.

En ambos módulos se utiliza R y Python para las aplicaciones prácticas.

## 8 Visualización de datos y diseño de dashboards:

Introducción al uso de herramientas de visualización, así como a la metodología del storytelling para crear visualizaciones de datos que comuniquen efectivamente los principales hallazgos de un proceso analítico.

Se utilizan las principales librerías de gráficos y visualización de datos de R y Python para crear imágenes que pueden ser incorporadas en presentaciones ejecutivas, reportes, publicaciones varias, sitios web, entre otros.



# PERFIL DEL INTERESADO

- Profesionales de distintas áreas del conocimiento, como ingeniería, administración, economía, finanzas, agronomía, química, biología, entre otros.
- Analistas de negocios en organizaciones de cualquier tipo.
- Personas en general que deseen adquirir conocimientos sobre ciencia de datos y analítica para aplicarlos en análisis de datos de cualquier tipo de procesos u operaciones.

El programa no requiere conocimientos previos sobre el tema. Su composición permite al participante llevar un aprendizaje gradual y estructurado, iniciando por los aspectos básicos y fundamentales, hasta llegar a las aplicaciones prácticas del contenido.



■ **MODALIDAD**  
Virtual

■ **DURACIÓN**  
Cada módulo tiene una duración de 16 horas, que suman 128 horas de formación total.

■ **METODOLOGÍA**  
Los cursos son prácticos. Se utilizan sets de datos reales y casos de estudio para ilustrar la aplicación de los conceptos y metodologías impartidas. El software empleado es de acceso libre y gratuito, y puede ser instalado en la computadora del estudiante sin ninguna restricción.

Para obtener el diploma, los participantes deben cursar todos los módulos y desarrollar un proyecto de aplicación que abarque uno o varios de los puntos vistos en el programa.

■ **INVERSIÓN**  
Precio por el programa completo: Q13,500.00  
(con opción a pago con visacuotas).

**Fecha de inicio:** 10 de junio  
Lunes y miércoles de 6:00 a 8:00 p. m.  
**Modalidad virtual**

**INSCRÍBETE EN**  
[bit.ly/DiplomadoData](https://bit.ly/DiplomadoData)

**Departamento de  
Business Intelligence y Data Science**

José Antonio Medrano  
jamedrano@uvg.edu.gt