

Entrenamiento Especializado

Ciencia de los Datos aplicando técnicas de análisis multivariante.

Descripción

Un elemento central en la Ciencia de Datos es la aplicación de las técnicas de Análisis Multivariante, las cuales permiten analizar de manera simultánea el comportamiento de múltiples características o variables relacionadas con un conjunto de objetos (personas, empresas, entidades, entre otras) facilitando establecer relaciones y generar grupos. En este entrenamiento especializado el participante avanzará en el aprendizaje de los procedimientos asociados con la Ciencia de los Datos (Data Science) con un enfoque práctico, partiendo de conceptos intermedios de estadística y análisis de datos, hasta llegar al abordaje de técnicas multivariantes que proporcionan mayor tecnicidad y complejidad analítica, permitiendo la apropiación de los conceptos y procedimientos asociados así como su aplicación práctica a cualquier tipo de industria o área de investigación.

Dirigido a

Todas las personas con un conocimiento intermedio de estadística interesadas en aprender y aplicar las principales técnicas de análisis multivariado para el análisis científico de datos.

Objetivo

- Entender los conceptos clave de las principales técnicas de Análisis Multivariante y poder diferenciarlas entre sí.
- Identificar las aplicaciones y utilidades de las distintas técnicas para el análisis de datos.
- Utilizar software especializado para el desarrollo de los algoritmos así como para la validación matemática y estadística de cada uno de ellos.

Temario

Sesión 1:

Introducción al Análisis Multivariante

Términos y conceptos relevantes en la aplicación del análisis multivariante.

Escalas de Medida.

Principales Técnicas del Análisis Multivariante.

Introducción al Manejo de STATA

Ayuda y recursos.

Descripción de datos.

Edición de datos.

Creación y cambio de datos.

Administración de variables.

Resumen de estadísticas.

Distribución de Frecuencias, Tabulación Cruzada y Pruebas de Hipótesis

Distribuciones de Frecuencia

Medidas de Ubicación

Medidas de Variabilidad

Medidas de Forma.

Sesión 2:

Inferencia-Intervalos de Confianza y Pruebas de Hipótesis

Diferencia de Medias.

Pruebas de Asociación.

Pruebas Paramétricas.

Pruebas no Paramétricas.

Análisis de Varianza de una Vía (ANOVA)

Sesión 3:

Regresión Lineal Múltiple

Modelo lineal general.

Estimación del modelo de regresión.

Pruebas de significancia de los parámetros.

Evaluación del modelo.

Multicolinealidad.

Criterios para adicionar o eliminar variables, Procedimientos para la selección de variables, Regresión escalonada (Stepwise):

Selección hacia adelante (Forward) y Eliminación hacia atrás (Backward)

Validación de supuestos.

Sesión 4:

Regresión Logística y Análisis Discriminante

Regresión Logística.

Modelo de Análisis Discriminante.

Estadísticos Asociado.

Supuestos y Limitaciones.

Análisis de resultados.

Sesión 5:

Clasificación por Grupos: Análisis de Conglomerados

Visión General.

Estadísticos Asociados al Análisis de Conglomerados.

Supuestos y Limitaciones.

Análisis de resultados.

Escalamiento Multidimensional (EM)

Visión General.
Estadísticos Asociados.
Supuestos y Limitaciones.
Análisis de resultados.

Sesión 6:

Análisis Factorial y de Componentes Principales (PCA)

Visión General.
Componentes principales y análisis factorial.
Estadísticos asociados.
Supuestos y Limitaciones.
Análisis de resultados.

Análisis de Correspondencias

Visión General.
Estadísticos Asociados.
Supuestos y Limitaciones.
Análisis de resultados.

Taller de aplicación

- Identificación de necesidades según área de trabajo de los participantes
- Resolución de inquietudes y vías de aplicación
- Ejemplos de aplicación de la Ciencia de Datos en gobierno
- Ejemplos de aplicación en finanzas
- Ejemplos de aplicación en Investigación de Mercados
- Otras aplicaciones

Instructores

Miller Janny Ariza

Doctorando en Análisis de Datos-Data Science Universidad Complutense de Madrid, Magíster en Economía de la Pontificia Universidad Javeriana, Lic. en Matemáticas de la Universidad Distrital, y Estadístico de la Universidad Nacional de Colombia. Se ha desempeñado en cargos de dirección en áreas de Riesgo y Modelaje en el Banco Caja Social, Consultor estadístico para entidades como UNODC-Oficina de las Naciones Unidas contra la Droga y el Delito, ITS soluciones estratégicas en proyectos conjuntos con el DNP (Departamento Nacional de Planeación-regalías), Investigador en estudios desarrollados por CIET (Centro de Investigación para la Epidemia del Tabaquismo) y Consultor estadístico regional en Latinoamérica y el Caribe para la Fundación Capital en proyectos de Inclusión Financiera.

Información General

Duración

30

Fecha de Inicio

Jue. 31 de Dic de 2020

Horarios

Fecha inicio:

Julio 23 de 2019

Programación de las sesiones:

Lunes, martes y miércoles.

Tres horas por sesión.

Hora de inicio por país

3:30 p.m. San José de Costa Rica

4:30 p.m. Bogotá - Lima - Quito - CDMX

5:30 p.m. Caracas - La Paz - Santiago de Chile

6:30 p.m. Buenos Aires

Herramientas de apoyo



Políticas

Para conocer en detalle las condiciones y políticas de participación en los eventos, ingrese a nuestro link

<https://www.software-shop.com/formacion/politicas>

Inscríbete ahora