DIPLOMADO
DE ACTUALIZACIÓN PROFESIONAL

DISEÑO DE INVESTIGACIÓN Y ANÁLISIS DE DATOS CON SOFTWARE R

COCID
Colegio Científico de Datos

REGISTRO OFICIAL SEEM 17DAP57319

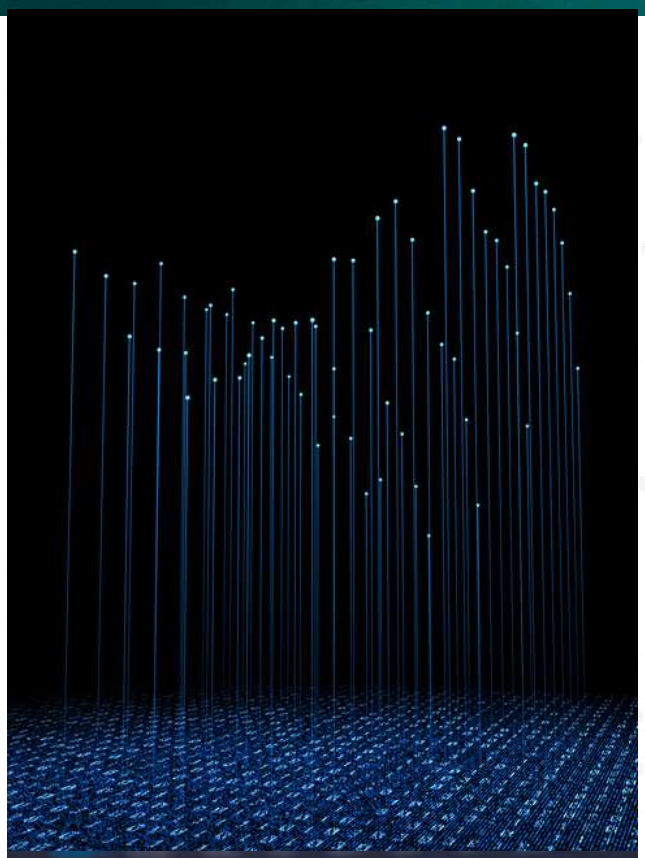
Inicia desde ahora

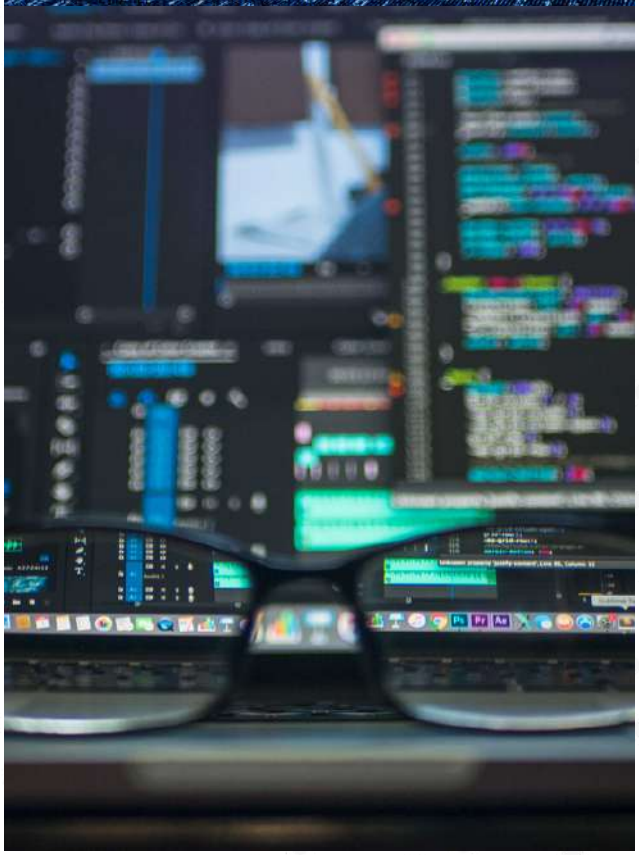
100% VIRTUAL

www.cocid.mx

Información

COCID
Colegio Científico de Datos





Objetivo:
Convertirse en un científico de datos profesional con la herramienta del lenguae R, aprendiendo su potencialidad para el procesamiento de datos, interpretando de forma correcta los resultados obtenidos y así innovar en sus proyectos de investigación.

A quién va dirigido:
- A cualquier persona que quiera adentrarse en el mundo de la investigación
- Mínimo nivel de licenciatura
- Conocimientos básicos de estadística requeridos
- No se requieren conocimientos de programación en lenguaje R

Comienza desde este momento

200 hrs. de estudio, (promedio de estudio estimado de 6 meses)

Modalidad: Virtual

Costo mediante plataforma PayPal $15,000.00
Costo mediante transferencia electrónica $13,500 incluye un 10% de descuento al elegir el pago en esta modalidad.

Transferencia

Vía transferencia electrónica o depósito bancario

Banco BBVA
PROMOTORA DE FORMACIÓN ASML, S.C
Cuenta: 0113456544

Clave interbancaria: 012542001134565445

PayPal

Vía Paypal (con tarjeta de crédito)

1.- Ingresar en el siguiente link
https://www.paypal.me/COCID2017
2.- Seleccionamos la opción "Enviar"
3.- Colocamos el monto a pagar $15,000.00
4.- Elegimos los meses deseados a pagar de 1 a 6 meses

Imangen de Ejemplo

Si requiere factura, favor de enviar los datos vía correo electrónico a administracion@cocid.mx o a través del formulario de Inscripción

Procedimiento de inscripción

1 **Llenar el formulario de inscripción**

Escanee el código para llenar el formulario

2 **Realizar el pago del diplomado y enviar por correo el comprobante de pago:** cocid.direccion@gmail.com o administracion@cocid.mx

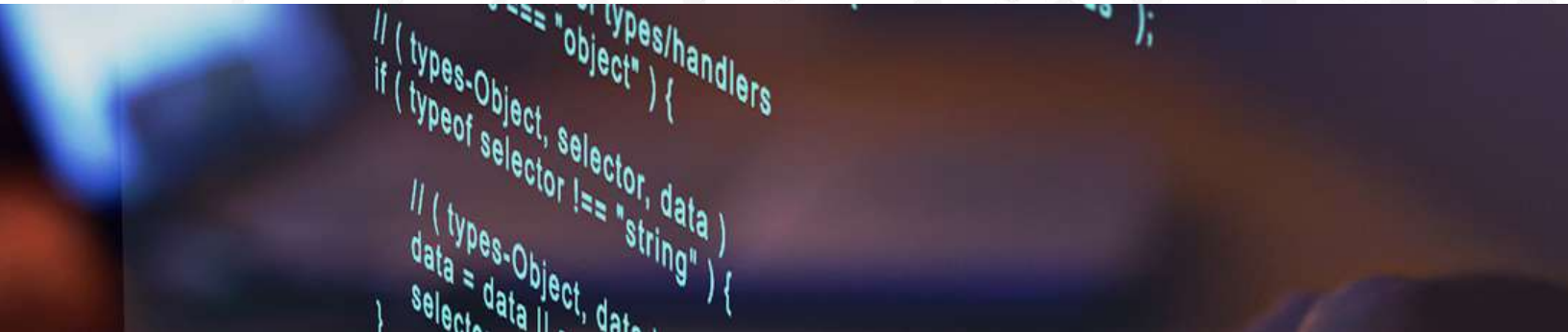
3 **Tendrá acceso al diplomado en menos de 24 hrs** desde el envío del comprobante de pago

Información

COCID
Colegio Científico de Datos

¿Por qué estudiar el Diplomado de Actualización Profesional: Diseño de Investigación y Análisis de Datos con software R, con registro oficial 17DAP57319 -SEEM-?

- 1) Comenzará de cero en programación.
- 2) Solo se le recomienda tener nociones básicas de estadística aplicadas a las ciencias experimentales.
- 3) Comprenderá y sabrá aplicar las nociones básicas para poder llevar a cabo una investigación, desde el póster de concepto, cálculos muestrales hasta los distintos tipos de diseños de investigación, entre otros.
- 4) Aprenderá las técnicas estadísticas más utilizadas en investigación.
- 5) Aplicará las técnicas clásicas paramétricas y no paramétricas, predictivas y de análisis multivariado.
- 6) Conocerá y sabrá en qué situaciones aplicar las técnicas robustas en las diferentes técnicas para el tratamiento de falta de supuestos y presencia de datos anómalos.
- 7) Aprenderá técnicas de machine learning.
- 8) Trabajará estrategias ETL y bases de datos (big data) con cubos OLAP.
- 9) Aprenderá a trabajar bases de datos con entorno tidyverse.
- 10) Al finalizar obtendrá un diploma oficial con valor curricular por valor de 200 horas y/o 12.5 créditos de la Secretaría de Educación del Estado de Morelos con registro 17DAP57319.



Información

COCID
Colegio Científico de Datos

¿Qué integra la enseñanza virtual?

moodle

- Plataforma Moodle
- Forma de trabajo intuitiva
- 24/7
- Toda la información integrada en módulos y apartados
- Con todo el temario completo, 58 clases grabadas, material complementario y actividades
- Explicación pormenorizada en cada módulo
- Evaluación automatizada
- Acceso a biblioteca virtual
- Tutor especializado en comunicación constante
- Soporte y asistencia técnica inmediata

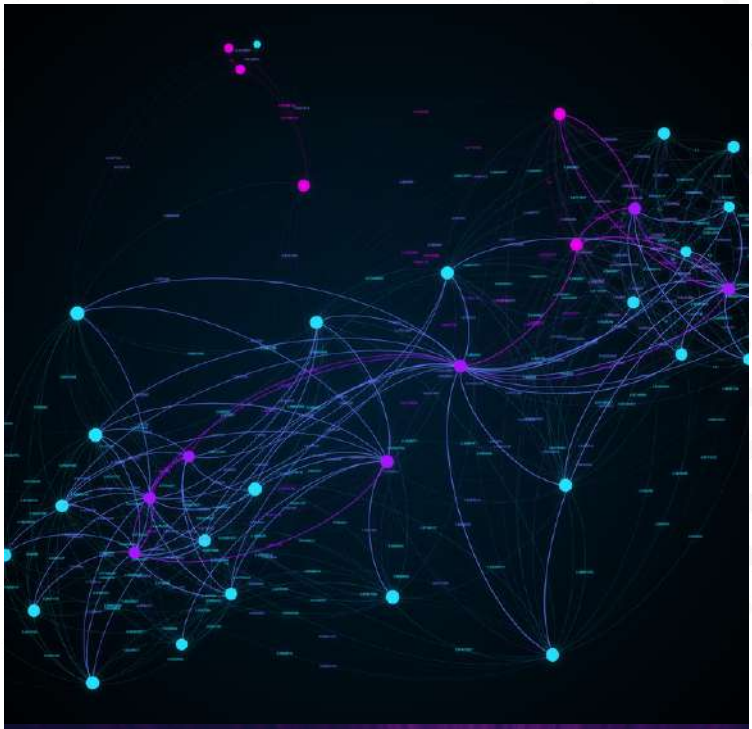
Más de 500 investigadores de México e Hispano América formados con nuestros diplomados

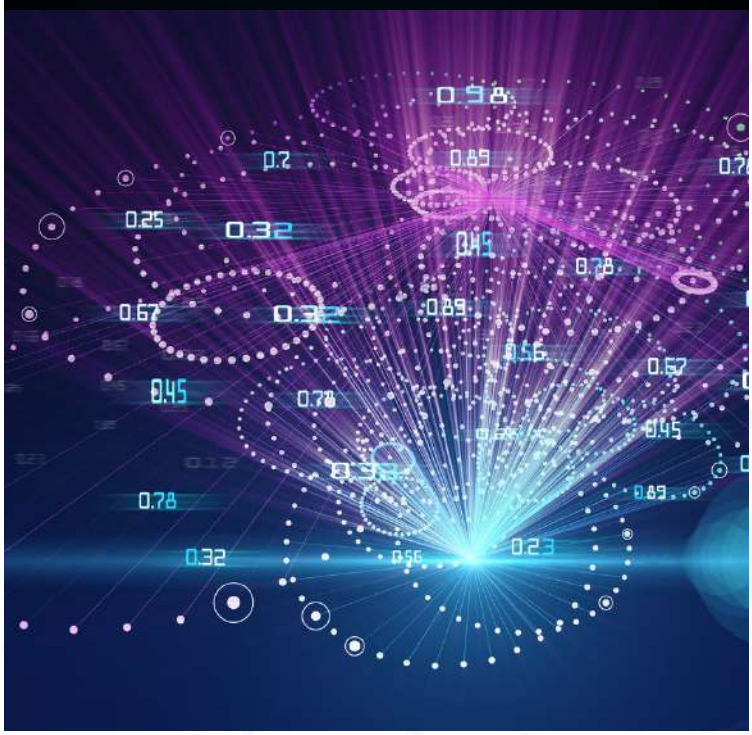
Contenido

COCID
Colegio Científico de Datos

1 Módulo I. DISEÑO DE INVESTIGACIÓN (40 hrs)

TEMA 1. FUNDAMENTOS Y BASES EMPÍRICAS.
TEMA 2. ELECCIÓN DEL TEMA, OBJETIVOS E HIPÓTESIS.
TEMA 3. MUESTREOS, TIPOS, DATOS Y CONTROLES.
TEMA 4. TIPOS DE DISEÑOS DE INVESTIGACIÓN.





2 MÓDULO II. TÉCNICAS CLÁSICAS (40 hrs)

TEMA 1. R SOFTWARE.
1.1.- Descripción.
1.2.- Historia.
1.3.- Características
1.4.- Instalación
1.5.- Formatos.
1.6.- Transponer datos de hoja de cálculo.
1.7.- Cargado de archivos.
1.8.- Trabajo de bases de datos.

TEMA 2. ESTADÍSTICA DESCRIPTIVA I.
2.1.- Gráficos de barras.
2.2.- Gráficos de sectores.
2.3.- Histograma.
2.4.- Nubes de puntos.
2.5.- Gráficos de cajas.
2.6.- Gráficos para tablas de doble entrada.
2.7.- Conclusión.

TEMA 3. ESTADÍSTICA DESCRIPTIVA II.
3.1.- Medidas de posición y dispersión.
3.1.1.- Media.
3.1.2.- Mediana.
3.1.3.- Cuasivarianza.
3.1.4.- Cuasidesviación típica.
3.1.5.- Cuantiles.
3.1.6.- Resumen.
3.2.- La distribución normal.
3.3.- Medidas para investigación.
3.4.- Conclusión.

TEMA 4. PRUEBAS NO PARAMÉTRICAS.
4.1.- Test de Wilcoxon.
4.2.- Test de Wilcoxon – Mann-Whitney.
4.3.- Test de Kolmogorov – Smirnov.
4.4.- Test de Kruskal – Wallis.
4.5.- Conclusión.

TEMA 5. CHI-CUADRADO Y T DE STUDENT.
5.1.- Chi- cuadrado de Pearson.
5.1.1.- Introducción.
5.1.2.- Formulación.
5.1.3.- Resolución con R.
5.1.4.- Conclusión.
5.2.- T-Student.
5.2.1.- Distribución t-Student.
5.2.3.- conclusión.

TEMA 6. ANÁLISIS DE VARIANZA. DISEÑO COMPLETAMENTE ALEATORIZADO.
6.1.- Un factor clásico. HSD Tukey.
6.2.- Un factor robusto. HSD Tukey.
6.3.- Dos factores clásico.
6.4.- Dos factores robusto.

TEMA 7. ANÁLISIS DE VARIANZA. MEDIDAS REPETIDAS.
7.1.- Introducción.
7.2.- Formulación.
7.3.- Análisis de Varianza. Medidas Repetidas con R.
7.4.- Conclusión.

TEMA 8. REGRESIÓN Y CORRELACIÓN LINEAL SIMPLES.
8.1.- Regresión lineal simple.
8.1.1.- Introducción.
8.1.2.- Modelo.
8.1.3.- Contraste.
8.1.4.- Tabla de análisis de varianza.
8.1.5.- Regresión con R.
8.1.6.- Conclusión.
8.2.- Correlación simple.
8.2.1.- Introducción.
8.2.2.- Coeficiente de correlación de Pearson.
8.2.3.- Correlación lineal con R.
8.2.4.- Conclusión.

TEMA 9. REGRESIÓN MÚLTIPLE.
9.1.- Modelo.
9.2.- Contraste de la regresión lineal múltiple.
9.3.- Tabla de análisis de la varianza para la regresión lineal múltiple.
9.4.- Estimación de la varianza común.
9.5.- Contraste de hipótesis sobre los coeficientes de regresión.
9.6.- Regresión múltiple con R.
9.7.- Selección secuencial de variables por pasos.

TEMA 10. CORRELACIÓN MÚLTIPLE.
10.1.- Introducción.
10.2.- Cálculo con R.
10.3.- Representaciones gráficas.
10.4.- Conclusión.

735 339 27 95
www.cocid.mx
PROMOTORA DE FORMACIÓN ASML SC
PayPal VISA

Contenido

COCID
Colegio Científico de Datos

3 MÓDULO III. TÉCNICAS AVANZADAS (40 hrs)

TEMA 1. TÉCNICAS ACTUALES EN REGRESIÓN. TRATAMIENTO DE DATOS ANÓMALOS.
1.1.- Introducción.
1.2.- Bibliotecas.
1.3.- Formulación.
1.4.- Ejemplo completo:
1.4.1.- Regresión lineal por mínimos cuadrados.
1.4.2.- Técnicas actuales en Regresión (tratamiento de datos anómalos).
1.4.2.1.- Recta de Huber.
1.4.2.2.- LTS-Least Trimmed Squares.
1.4.2.3.- LMS-Least Median of Squares.
1.4.2.4.- Recta MM.
1.5.- Conclusión.

TEMA 2. REGRESIÓN POISSON.
2.1.- Introducción.
2.2.- Cálculo con R.
2.3.- Conclusión.

TEMA 3. MODELOS DE REGRESIÓN LOGÍSTICA. LOGIT Y PROBIT.
3.1.- Introducción.
3.2.- Modelos logit y probit.
3.3.- Modelos de regresión logit con R.
3.4.- Modelos de regresión probit con R.
3.5.- Conclusión.

TEMA 4. REGRESIÓN SUAVIZADA.
4.1.- Introducción.
4.2.- Formulación.
4.3.- Regresión suavizada con R.
4.4.- Conclusión.

TEMA 5. REGRESIÓN MÚLTIPLE ROBUSTA.
5.1.- Introducción.
5.2.- Regresión de Huber.
5.3.- Regresión múltiple robusta de Huber.
5.4.- MM-estimadores.
5.5.- Regresión múltiple M-estimadores.
5.6.- Conclusión.

TEMA 6. CORRELACIÓN SIMPLE Y MÚLTIPLE ROBUSTAS.
6.1.- Introducción.
6.2.- Correlación robusta para dos variables.
6.2.1.- Correlación de porcentaje ajustado.
6.2.2.- Correlación winsorizada.
6.2.3.- M-estimadores de Goldberg e Iglewicz.
6.3.- Correlación robusta para p variables.
6.3.1.- Correlación de porcentaje ajustado entre p variables.
6.3.2.- Correlación winsorizada entre p variables.
6.4.- Conclusión.

TEMA 7. MODELOS LINEALES GENERALIZADOS UNIVARIANTES.
7.1.- Desarrollo.
7.2.- Conclusión.

TEMA 8. MODELOS LINEALES MIXTOS GENERALIZADOS. DISEÑO POR BLOQUES ALEATORIOS.
8.1.- Introducción.
8.2.- Formulación.
8.3.- Modelos lineales generalizados y mixtos generalizados con R. Diseño por bloques aleatorios.
8.4.- Conclusión.

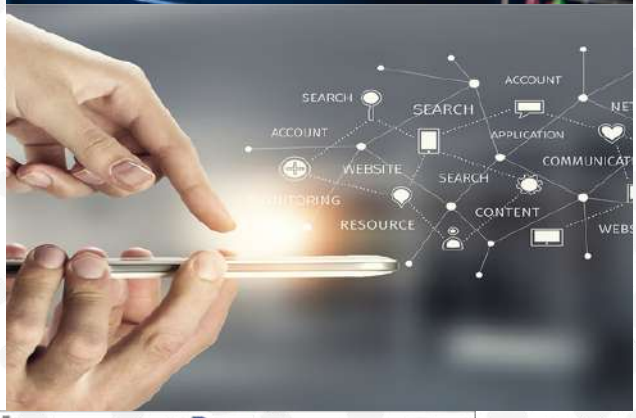
TEMA 9. ÁRBOLES DE REGRESIÓN Y CLASIFICACIÓN (CARTS).
9.1.- Introducción.
9.2.- Árboles de regresión con R.
9.3.- Prunned o podado del Árbol.
9.4.- Árboles de clasificación con R.
9.5.- Conclusión.

TEMA 10. MODELOS ADITIVOS GENERALIZADOS GAM.
10.1.- Introducción.
10.2.- Formulación.
10.3.- Modelos GAM con R.
10.4.- Conclusión.

TEMA 11. ANÁLISIS DE SUPERVIVENCIA.
11.1.- Introducción.
11.2.- Cálculo completo en R. Uno y dos grupos.

TEMA 12. MODELOS SARIMA DE SERIES TEMPORALES.
12.1.- Introducción.
12.2.- Objetivos y usos en las diferentes ciencias.
12.3.- Paquetes a instalar en R.
12.4.- Bibliotecas a abrir en R.
12.5.- Cómo introducir los datos en R.
12.6.- Ejemplo completo:
12.6.1.- Cálculo con R.
12.6.2.- Filtrado lineal.
12.6.3.- Modelos SARIMA:
12.6.3.1.- Identificación del modelo.
12.6.3.2.- Estimación de los parámetros.
12.6.3.3.- Diagnosis.
12.6.3.4.- Predicción.
12.6.3.5.- Test de serie estacionaria.
12.6.4.- Cointegración de Series.
12.7.- Conclusión.

TEMA 13. PARÁMETROS GRÁFICOS.
13.1.- Gráfico plot. Función legend.
13.2.- Recta de regresión.
13.3.- Histograma.
13.4.- Gráfico de cajas.
13.5.- Gráficos de barras con barras de error.



735 339 27 95
www.cocid.mx
PROMOTORA DE FORMACIÓN ASML SC
PayPal
VISA

Contenido

COCID
Colegio Científico de Datos

4

Módulo 4.TÉCNICAS DE ANÁLISIS MULTIVARIANTE (40 hrs)

TEMA 1. ANÁLISIS DE CORRESPONDENCIAS.
1.1.- Introducción.
1.2.- Análisis de correspondencias bidimensional con R.
1.3.- Análisis de correspondencias múltiple con R.
1.4.- Conclusión.

TEMA 2. ANÁLISIS FACTORIAL.
2.1.- Introducción.
2.2.- Formulación.
2.3.- Análisis factorial con R.
2.4.- Conclusión.

TEMA 3. ANÁLISIS DE COMPONENTES PRINCIPALES.
3.1.- Introducción.
3.2.- Formulación.
3.3.- Análisis de componentes principales con R.
3.4.- Representaciones gráficas
3.5.- Conclusión.

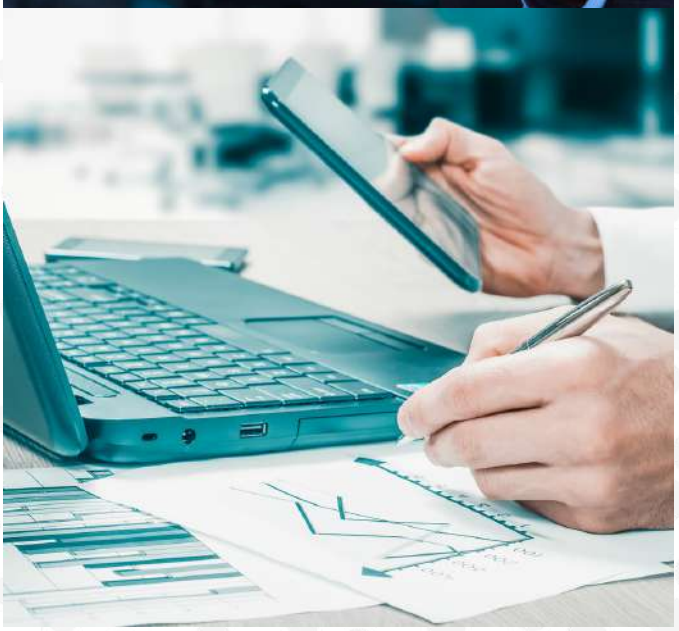
TEMA 4. ANÁLISIS DISCRIMINANTE.
4.1.- Introducción.
4.2.- Formulación.
4.3.- Análisis discriminante con R.
4.4.- Validación cruzada.
4.5.- Método de los k vecinos más próximos.
4.6.- Método de los k vecinos más próximos. Validación cruzada.
4.7.- Análisis discriminante. K grupos y clasificación desconocida.
4.8.- Conclusión.

TEMA 5. ESCALADO MULTIDIMENSIONAL.
5.1.- Introducción.
5.2.- Formulación.
5.3.- Escalado multidimensional con R. Métrico y no métrico.
5.4.- Cálculo de matrices.
5.5.- Conclusión.

TEMA 6. ANÁLISIS CLUSTERS.
6.1.- Introducción.
6.2.- Formulación.
6.3.- Análisis clusters con R.
6.4.- Elección del número de clusters.
6.5.- Conclusión.

5

Módulo 5. . BASES DE DATOS. BIG DATA Y ETL (40 hrs)
TEMA 1. TRABAJO DE BIG DATA Y CUBOS OLAP.
1.1.- OLAP (On-Line Analytical Processing – Procesamiento analítico en línea)
1.1.1.- Los Cubos OLAP
1.1.2.- Gestores de bases de datos con OLAP
1.1.3.- Hechos y dimensiones
1.1.4.- Operaciones OLAP
1.1.4.1.- «Rebanada»
1.1.4.2.- «Dados»
1.1.4.3.- Enrollar
1.1.4.4.- Profundizar
1.1.4.5.- Pivote



735 339 27 95
www.cocid.mx
PROMOTORA DE FORMACIÓN ASML SC
PayPal VISA

Contáctanos

administracion@cocid.mx
cocid.direccion@gmail.com
735 339 27 95