



masterX

Programa de Especialización:

Estadística y Análisis de Datos en R

Sesiones online en vivo

PLAN DE ESTUDIOS



INFORMACIÓN GENERAL DEL PROGRAMA

Este es el punto de partida de tu carrera en investigación. En este programa, aprenderás habilidades en demanda que te prepararán analizar los datos generados por investigaciones. No se requiere experiencia previa en programación o estadística. Incluye cinco cursos para llevarte de Zero a Hero en el manejo del lenguaje de programación R orientado a investigación científica. Haz clic en los banners para ver la currícula de cada curso.

Inicio: 06 de Junio de 2022

Extensión: 15 semanas (46 sesiones)

Horario: (GMT -5, hora de Lima, Quito, Bogotá)

- Lunes, Miércoles y Viernes de 3:00 a 5:30 pm

Incluye:

- Clases online 100% en vivo (no es obligatoria la asistencia).
- Videos de las grabaciones disponibles posterior a cada sesión de clase.
- Archivos descargables.
- PDFs de diapositivas.
- Certificados (05) con enlace único URL para compartir en redes sociales, como LinkedIn.



Data Science: Estadística y Análisis de Datos con R
Grupo 2022-2

Clases: Lunes, Miércoles y Viernes

Horario: 3 a 5:30 pm (GMT-5, Hora Oficial de Lima, Quito, Bogotá)

Beneficios

- Plataformas masterX y Zoom
- Clases 100% en vivo
- Grabaciones de cada clase disponibles
- Metodología Hands-On
- Profesor especializado en R y docencia

Cursos: 185 horas académicas totales

- C1: Introducción a R para Ciencias (40 horas acad.)
- C2: Gráficos estadísticos avanzados con ggplot2 (35 horas acad.)
- C3: Estadística básica con R (40 horas acad.)
- C4: Análisis multivariados con R (35 horas acad.)
- C5: Modelamiento estadístico avanzado con R (35 horas acad.)

masterX



Profesor: Blgo. Irwing S. Saldaña

Candidato a Magíster (Universidad de San Martín de Porres y Universidad Nacional Mayor de San Marcos). Biólogo con experiencia en análisis estadístico y modelamiento. Experiencia de más de 3 años en enseñanza del lenguaje de programación R a alumnos de diferentes profesiones. Interesado en el desarrollo de la informática ecológica y evolutiva (Eco-Evo Informática), con un enfoque de biología integrativa.

C1 - INTRODUCCIÓN A R PARA CIENCIAS

Módulo 1: Esenciales de R y RStudio

- Instalación y configuración (cubierto con videotutorial previo a la clase inicial)
- Conociendo a RStudio
- Librerías y actualizaciones
- Tipos de archivos en RStudio
- Proyectos en RStudio
- Ayuda en RStudio

Módulo 2: Introducción al lenguaje de programación R

- Primeros pasos con código en R
- Variables u objetos
- Funciones
- Estructuras de datos
- Coerción
- Carga y exportación de datos
- Programación funcional
- Creación de funciones
- Iniciando con loops

Módulo 3: Tidyverse para el manejo de bases de datos

- Datos tidy
- Pipelines y pipe
- tidyR para el manejo de tablas
- Filtrado
- Selección
- Agrupamiento
- Reordenamiento
- Conteo
- Transformación de variables
- Resúmenes estadísticos
- Combinar tablas
- Valores perdidos

C2 - GRÁFICOS ESTADÍSTICOS AVANZADOS CON GGPLO2

Módulo 1: Introducción a ggplot2

- Gramática de gráficos
- Trabajo por capas
- Geométricas
- Stats

- Colores y formas
- Etiquetas
- Facetas
- Modificando los ejes
- Temas
- Exportación de alta calidad

Módulo 2: Gráficos estadísticos manuales

- Visualización de cantidades
- Visualización de histogramas
- Visualización de funciones de densidad
- Visualización de distribuciones en simultáneo
- Visualización de proporciones
- Visualización de asociaciones
- Visualización de series temporales
- Visualización de regresiones

Módulo 3: Gráficos estadísticos semi-automáticos

- Funciones de ggpubr
- Funciones de ggstatsplot
- Funciones de ggcorplot

C3 - ANÁLISIS ESTADÍSTICO BÁSICO CON R

Módulo 1: Introducción a probabilidades

- Probabilidades
- Variables aleatorias
- Distribuciones teóricas de probabilidades
- Distribuciones empíricas de probabilidades
- Función de densidad
- Medidas de tendencia central
- Medidas de dispersión
- Error estándar
- Margen de error
- Intervalos de confianza
- ¿Qué significa el p-valor?
- Nivel de significancia
- Pruebas de hipótesis

Módulo 2: Test estadísticos pareados

- Introducción a pruebas paramétricas y no paramétricas
- Pruebas de normalidad
- Pruebas de T de Student (con variantes)
- Prueba de Wilcoxon para una muestra

- Prueba de Rangos con signo
- Prueba U de MannWhitney
- Supuestos teóricos de las pruebas

Módulo 3: Modelamiento lineal

- Regresión lineal simple
- Verificación de supuestos teóricos
- Valores atípicos en regresiones
- Prueba de ANOVA (y variantes)
- Prueba de Kruskal-Wallis
- Post Hoc (paramétrico y no paramétrico)
- ANCOVA
- Análisis de correlación (Pearson, Spearman, Kendall)

C4 - ANÁLISIS MULTIVARIADOS CON R

Módulo 1: Introducción a estadística multivariada

- Estadística multivariada
- ¿Qué es ordenamiento?
- Espacios N-dimensionales
- Preprocesamiento de datos
- Transformaciones
- Matrices de distancia
- Medidas de distancia

Módulo 2: Métodos multivariados no canónicos

- Agrupamiento jerárquico (clustering)
- Agrupamiento no jerárquico (clustering)
- Análisis de Componentes Principales (PCA)
- Análisis de Correspondencia (CA)
- Análisis de Coordenadas Principales (PCoA)
- Escalamiento Multidimensional No Métrico (NMDS)

Módulo 3: Métodos multivariados canónicos

- Introducción al análisis canónico
- Análisis de Redundancia (RDA, y variantes)
- Gráfico triplot
- Pruebas de permutaciones
- Análisis de Correspondencia Canónica (CCA, y variantes).
- Análisis Discriminante Lineal (LDA)
- Análisis Discriminante Cuadrático (QDA)

C5 - MODELAMIENTO ESTADÍSTICO AVANZADO CON R

Módulo 1: Modelos lineales con efectos mixtos (LMM)

- Breve repaso sobre modelamiento
- Modelos lineales con efectos mixtos
- Definición matemática
- Diferencias entre LMM y ANOVA para muestras repetidas
- Terminología para modelos de efectos mixtos
- Enfoques de análisis Regresión lineal simple y LMM
- Topología de una LMM en R
- Gráficas asociadas con ggplot2
- Selección de modelos

Módulo 2: Modelos generalizados lineales (GLM)

- Distribuciones teóricas de probabilidades en GLM
- Funciones de enlace (link function)
- Topología de un GLM en R
- Regresión logística
- Odds ratio
- Logit function
- Selección de modelos
- Regresión binomial
- Regresión de poisson
- Sobredispersión en modelos
- Regresión binomial negativa
- Trabajando con exceso de ceros
- Regresión de Hurdle, Zero inflados, Zero truncados
- Gráficos asociados con ggplot2

Módulo 3: Modelos generalizados lineales con efectos mixtos (GLMM)

- Introducción a modelos generalizados con efectos mixtos
- Inferencias basadas en máxima verosimilitud
- Integrales con cuadratura de Gauss-Hermit y Laplace
- Regresión logística con efectos mixtos
- Regresión de poisson con efectos mixtos
- Regresión binomial negativa con efectos mixtos
- Gráficos asociados a ggplot2



masterX



www.masterx.org



www.brackinstitute.com



Fan page masterX | Cursos online para científicos



Fan page Instituto de Ciencias Antonio Brack



Chat Instituto de Ciencias



soporte@masterx.org



info@brackinstitute.com