

**INSTITUTO DE EVALUACIÓN DE TECNOLOGÍAS EN SALUD E
INVESTIGACIÓN, IETSI-ESSALUD
DIRECCIÓN DE INVESTIGACIÓN EN SALUD**

PROGRAMA DE FORMACIÓN CIENTÍFICA: ANÁLISIS DE DATOS 2023

I. DATOS GENERALES		
1.1	Programa de Formación Científica	Análisis de Datos 2023
1.2	Número Total de Horas:	22 horas de actividad sincrónica
1.3	Créditos académicos	1 crédito
1.4	Duración	(8 días) 24, 26, 31 de mayo de 2023 2, 9, 14, 16, 21 de junio de 2023
1.5	Horario	6:00 p.m. a 9:00 p.m.
1.6	Público objetivo	Profesionales de la salud de EsSalud
1.7	Número de participantes	30
1.8	Requisito	Ninguno
1.9	Lugar	Virtual - Plataforma Zoom
1.10	Coordinación	Dirección de Investigación en Salud (DIS), IETSI
1.11	Docentes	Investigadores RENACYT

II. ASPECTOS GENERALES

El Instituto de Evaluación de Tecnologías en Salud e Investigación (IETSI), tiene dentro de sus funciones el promover y desarrollar investigación en EsSalud, para el beneficio de sus asegurados. Así mismo, la directiva que regula el desarrollo de la investigación en salud, establece que la dirección de investigación en salud, debe de promover actividades de investigación en salud, entre las cuales se enmarca el entrenamiento en aspectos metodológicos.

Este entrenamiento está dirigido a personal de salud que se encuentre laborando en EsSalud, que cuente con entrenamiento básico en estadística aplicada a ciencias de la salud, estadística

o ciencia de datos. Su propósito es capacitar y brindar herramientas en análisis bioestadístico de datos para preguntas de investigación explicativas. Se revisarán conceptos relacionados a la redacción científica y los componentes de un artículo de investigación. El entrenamiento se realizará de manera virtual empleando la plataforma Zoom. Al finalizar el entrenamiento, el alumno estará en la capacidad de manejar y preprocesar datos, analizarlos descriptivamente y evaluar algunas relaciones de interés desde una perspectiva de modelado descriptivo.

III. COMPETENCIAS DEL PROGRAMA

- Usar R y RStudio para el manejo y limpieza de base de datos de estudios de investigación en salud.
- Realizar un análisis descriptivo e inferencial con un paquete estadístico de datos para la investigación en salud.
- Responder preguntas de modelado descriptivo mediante regresión lineal y logística
- Generar tablas y reporte reproducibles de estos mediante el uso de markdown a través de Quarto en R/RStudio.

IV. RESULTADOS DE APRENDIZAJE

Al finalizar las unidades, los profesionales serán capaces de:

- **Módulo 1:** Manejar y pre-procesar datos con R.
- **Módulo 2:** Realizar un análisis descriptivo de los datos
- **Módulo 3:** Realizar procedimientos de inferencia estadística a través de pruebas de hipótesis y estimación de intervalos de confianza

V. CONTENIDOS

- **Módulo 1:** Manejo de datos con R
 - Importar y exportar datos.
 - Seleccionar columnas y filas.
 - Transformar columnas y filas.
 - Insertar y modificar metadatos.
- **Módulo 2:** Análisis descriptivo de datos con R
 - Inspección inicial de datos.
 - Descripción de datos numéricos
 - Descripción de datos categóricos
 - Descripción de la covariación de datos
- **Módulo 3:** Inferencia estadística con R

- Pruebas de hipótesis paramétricas y no paramétricas.
- Estimación de intervalo de confianza

VI. ESTRATEGIAS DIDÁCTICAS

- **Clases teóricas:** Los alumnos recibirán las clases teóricas de manera virtual. Inicialmente, los docentes dictarán las sesiones académicas, y los participantes brindarán comentarios respecto al tema. Posteriormente, se entregarán las diapositivas de las sesiones académicas por correo electrónico que serán leídas por los participantes. Las ponencias serán grabadas y se almacenarán en un repositorio académico virtual que será compartido a los participantes. Los participantes recibirán material de lectura complementario que ayudará a consolidar las sesiones teóricas.
- **Clases prácticas (Talleres):** Los alumnos recibirán las clases prácticas de manera virtual. Previamente, se hará entrega de conjunto de datos de estudios reales publicados en repositorios abiertos o realizados por IETSI. Durante las clases prácticas, guiados por el docente a cargo, se realizará el análisis de los datos aplicando los métodos enseñados durante la sesión. Los talleres serán grabados y se almacenarán en un repositorio académico virtual que será compartido a los participantes. Los conjuntos de datos y el código de análisis de datos estarán disponible en un repositorio abierto en Github. Adicionalmente, se dispondrá de 15 minutos adicionales por cada clase práctica para poder hacer preguntas directas al docente a cargo, para solucionar cualquier duda o consulta, o si es que algún alumno lo solicita, hacer una revisión específica de algún aspecto del análisis de datos.

Conjunto de datos para analizar en casos aplicados: Siempre que sea posible, trabajaremos con conjunto de datos reales u obtenidos mediante simulación de estudios reales. También haremos uso del conjunto de datos del paquete “medicaldata” de R.

EXAMEN FINAL

Al final del curso habrá un conjunto de datos problema que corresponderán a un conjunto de datos con una pregunta de investigación y objetivo de modelado pre-establecido. Los alumnos tendrán una semana para realizar los análisis solicitados y presentarlos como tarea calificada. Asimismo, se compartirá un solucionario al final del cierre de cada entrega.

VII. EVALUACIÓN

Para efectos de evaluar la asistencia, los participantes deberán haber asistido a mínimo 13/16 sesiones académicas. La nota mínima aprobatoria para que el participante apruebe el entrenamiento es de 13.00 (trece) en el sistema vigesimal.

Para aprobar el entrenamiento, los participantes deberán cumplir con dos requisitos:

- Cumplir con el requisito de asistencia del entrenamiento (13/16 sesiones académicas).
- Aprobar el examen final del entrenamiento (calificación mayor o igual a 13.0 en el examen final).

Se otorgará constancia de participación a quienes asistan mínimo al 80% de clases (equivalente a 13 de las 16 ponencias), y un certificado válido por 1 crédito académico a quienes adicionalmente aprueben la evaluación final.

VIII. BIBLIOGRAFÍA

- Measurement scales and their summary statistics BMJ Evidence-Based Medicine 2004;9:164-166. Enlace: <https://ebm.bmj.com/content/9/6/164.2.citation-tools>
- Baillie M, le Cessie S, Schmidt CO, Lusa L, Huebner M, for the Topic Group “Initial Data Analysis” of the STRATOS Initiative (2022) Ten simple rules for initial data analysis. PLoS Comput Biol 18(2): e1009819. <https://doi.org/10.1371/journal.pcbi.1009819>
- Cummings P, Rivara FP. Reporting Statistical Information in Medical Journal Articles. Arch Pediatr Adolesc Med. 2003;157(4):321–324. doi:[10.1001/archpedi.157.4.321](https://doi.org/10.1001/archpedi.157.4.321)

Sesión 2:

- Batra, Neale, et al. The Epidemiologist R Handbook. 2021. DOI 10.5281/zenodo.4752646. Capítulos 17. Enlace: <https://epirhandbook.com/en/index.html>
- Greenland, S., Senn, S.J., Rothman, K.J. et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol 31, 337–350 (2016). <https://doi.org/10.1007/s10654-016-0149-3>
- Altman, N., Krzywinski, M. P values and the search for significance. Nat Methods 14, 3–4 (2017). <https://doi.org/10.1038/nmeth.4120>
- Are confidence intervals better termed “uncertainty intervals”? BMJ 2019; 366 doi: <https://doi.org/10.1136/bmj.15381> (Published 10 September 2019)
- Valentin Amrhein, David Trafimow & Sander Greenland (2019) Inferential Statistics as Descriptive Statistics: There Is No Replication Crisis if We Don’t Expect Replication, The American Statistician, 73:sup1, 262-270, DOI: 10.1080/00031305.2018.1543137
- Hawkins AT, Samuels LR. Use of Confidence Intervals in Interpreting Nonstatistically Significant Results. JAMA. 2021;326(20):2068–2069. doi:10.1001/jama.2021.16172
- Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. Nature. marzo de 2019;567(7748):305-7. <https://www.nature.com/articles/d41586-019-00857-9>

ANEXO 1. PROGRAMACIÓN DE ACTIVIDADES

No de sesión	Día	Horario	Actividades de Aprendizaje	Docente Responsable
MÓDULO 1: MANEJO DE DATOS CON R				
0	24/05/2023	17:45– 18:00	Palabras de bienvenida y presentación del programa	Director del IETSI Dr. Santillana - Gerente de la Dirección de Investigación en Salud
1	24/05/2023	18:00 – 19:15	Clase teórica: Verificación de instalación de R, Rstudio y Quarto. Primeros pasos con R. Importar datos con R.	Dr. Percy Soto Becerra
	24/05/2023	19:15 – 19:30	Descanso	
2	24/05/2023	19:30 – 21:00	Clase práctica: Taller de trabajo en equipo: Primeros pasos con R. Importar datos con R.	Dr. Percy Soto Becerra Dr. Virgilio Failoc Rojas Dr. Diego Urrunaga Pastor Dr. Jorge Osada Liy MSc Stephanie Montero Trujillo Dr. Víctor Juan Vera Ponce Mg.Oliver Elorreaga Reyes
MÓDULO 1: MANEJO DE DATOS CON R				
3	26/05/2023	18:00 – 19:15	Clase teórica: Introducción a la visualización de datos con ggplot2.	Dr. Percy Soto Becerra
	26/05/2023	19:15 – 19:30	Descanso	
4	26/05/2023	19:30 – 21:00	Clase práctica: Introducción a la visualización de datos con ggplot2.	Dr. Percy Soto Becerra Dr. Virgilio Failoc Rojas Dr. Diego Urrunaga Pastor Dr. Jorge Osada Liy MSc Stephanie Montero Trujillo Dr. Víctor Juan Vera Ponce Mg.Oliver Elorreaga Reyes
MÓDULO 1: MANEJO DE DATOS CON R				
5	31/05/2023	18:00 – 19:15	Clase teórica: Manejo de datos con DPLYR 1: Seleccionar y ordenar filas y columnas. Crear y modificar columnas. Etiquetas y metadatos. Helpers.	Dr. Percy Soto Becerra
	31/05/2023	19:15 – 19:30	Descanso	

6	31/05/2023	19:30 – 21:00	Clase práctica: Manejo de datos con DPLYR 1: Seleccionar y ordenar filas y columnas. Crear y modificar columnas. Etiquetas y metadatos. Helpers.	Dr. Percy Soto Becerra Dr. Virgilio Failoc Rojas Dr. Diego Urrunaga Pastor Dr. Jorge Osada Liy MSc Stephanie Montero Trujillo Dr. Víctor Juan Vera Ponce Mg.Oliver Elorreaga Reyes MSc Stephanie Montero
MÓDULO 2: ANÁLISIS DESCRIPTIVO CON R				
7	02/06/2023	18:00 – 19:15	Clase teórica: Inspección inicial de datos. Análisis descriptivo de datos 1.	Dr. Percy Soto Becerra
	02/06/2023	19:15 – 19:30	Descanso	
8	02/06/2023	19:30 – 21:00	Clase práctica: Inspección inicial de datos. Análisis descriptivo de datos 1.	Dr. Percy Soto Becerra Dr. Virgilio Failoc Rojas Dr. Diego Urrunaga Pastor Dr. Jorge Osada Liy MSc Stephanie Montero Trujillo Dr. Víctor Juan Vera Ponce Mg.Oliver Elorreaga Reyes
MÓDULO 2: ANÁLISIS DESCRIPTIVO CON R				
9	09/06/2023	18:00 – 19:15	Clase teórica: Análisis descriptivo de datos 2.	Dr. Percy Soto Becerra
	09/06/2023	19:15 – 19:30	Descanso	
10	09/06/2023	19:30 – 21:00	Clase práctica: Análisis descriptivo de datos 2.	Dr. Percy Soto Becerra Dr. Virgilio Failoc Rojas Dr. Diego Urrunaga Pastor Dr. Jorge Osada Liy MSc Stephanie Montero Trujillo Dr. Víctor Juan Vera Ponce Mg.Oliver Elorreaga Reyes
MÓDULO 3: INFERENCIA ESTADÍSTICA CON R				
11	14/06/2023	18:00 – 19:15	Clase teórica: Prueba de hipótesis y Estimación 1.	Dr. Percy Soto Becerra
	14/06/2023	19:15 – 19:30	Descanso	
12	14/06/2023	19:30 – 21:00	Clase práctica: Prueba de hipótesis y Estimación 1.	Dr. Percy Soto Becerra Dr. Virgilio Failoc Rojas Dr. Diego Urrunaga Pastor Dr. Jorge Osada Liy MSc Stephanie Montero Trujillo Dr. Víctor Juan Vera Ponce Mg.Oliver Elorreaga Reyes
MÓDULO 3: INFERENCIA ESTADÍSTICA CON R				
13	16/06/2023	18:00 – 19:15	Clase teórica: Prueba de hipótesis y Estimación 2.	Dr. Percy Soto Becerra
	16/06/2023	19:15 – 19:30	Descanso	

14	16/06/2023	19:30 – 21:00	Clase práctica: Prueba de hipótesis y Estimación 2.	Dr. Percy Soto Becerra Dr. Virgilio Failoc Rojas Dr. Diego Urrunaga Pastor Dr. Jorge Osada Liy MSc Stephanie Montero Trujillo Dr. Víctor Juan Vera Ponce Mg.Oliver Elorreaga Reyes
MÓDULO 3: INFERENCIA ESTADÍSTICA CON R				
15	21/06/2023	18:00 – 19:15	Clase teórica: Tablas y reporte reproducibles.	Dr. Percy Soto Becerra
	21/06/2023	19:15 – 19:30	Descanso	
16	21/06/2023	19:30 – 21:00	Clase práctica: Tablas y reporte reproducibles.	Dr. Percy Soto Becerra Dr. Virgilio Failoc Rojas Dr. Diego Urrunaga Pastor Dr. Jorge Osada Liy MSc Stephanie Montero Trujillo Dr. Víctor Juan Vera Ponce Mg.Oliver Elorreaga Reyes
Hasta el 23 de junio de 2023			Conjunto de datos problema (examen final)	