



UNIVERSIDAD PERUANA
CAYETANO HEREDIA

INSTITUTO DE MEDICINA TROPICAL ALEXANDER VON HUMBOLDT
EDUCACIÓN CONTINUA

I. DATOS GENERALES

1.1	Nombre del curso	Introducción a R y Tidyverse
1.2	Dirigido a	Público general
1.3	Organizado por	Laboratorio de Innovación en Salud
1.4	Horas y/o créditos	16 horas
1.5	Modalidad	Online
1.6	Profesor coordinador	

II. RESULTADOS DE APRENDIZAJE

Al finalizar el curso de manera satisfactoria, el participante podrá:

- Importar conjuntos de datos de los principales formatos que se trabajan en proyectos de investigación.
- Dar formato inicial al conjunto de datos mediante reconocimiento de variables y modificaciones en las mismas.
- Ejecutar modificaciones en variables pre-existentes y nuevas variables que se introduzcan.
- Trabajar con múltiples bases de datos y unificarlas de acuerdo al objetivo de investigación planteado.

CONTENIDOS

<ul style="list-style-type: none">- Introducción a ciencia de datos- Manejo y reconocimiento inicial de Rstudio- Creación y manejo de estructura de proyectos en rstudio- Creación de vectores	Sesión 1
<ul style="list-style-type: none">- Uso de funciones y argumentos- Instalación y manejo de paquetes- Introducción a tidyverse- Importación de datos de principales formatos	Sesión 2

<ul style="list-style-type: none"> - Reconocimiento de variables (función glimpse) - Anidación de funciones con pipe (%>%) - Cálculo de variables (introducción a mutate) - Contabilización y ordenamiento en base a variables (uso de count y arrange). - Explicación de tidy data - Modificación a la estructura de los datos mediante pivot longer y wider 	Sesión 3
	Taller 1
<ul style="list-style-type: none"> - Selección de variables de interés - Uso de condicionales para filtrar casos - Aspectos adicionales sobre la creación de variables - Solicitar resumen de datos (summarise) - Considerar variables para subdividir análisis (group_by) 	Sesión 4
<ul style="list-style-type: none"> - Manejo de múltiples conjuntos de datos mediante funciones join. - Cálculo de variables por filas (rowwise) - Apilamiento de datos por fila (bind_rows) y columna (bind_cols) 	Sesión 5
<ul style="list-style-type: none"> - Introducción a Rmarkdown para reporte de análisis - Elaboración de gráficos con el paquete ggcharts. Gráfico de barras e histograma 	Sesión 6
	Taller 2

III. ESTRATEGIAS DIDÁCTICAS
<ul style="list-style-type: none"> ● Exposición declarativa virtual ● Material bibliográfico ● Ejercicios en clase ● Actividades de reforzamiento

IV. EVALUACIÓN
<p>Para concluir de manera satisfactoria todo el contenido del curso, se llevarán a cabo dos evaluaciones:</p> <ol style="list-style-type: none"> 1. La primera evaluación contará con 10 preguntas relacionadas al primer y segundo módulo. 2. Finalmente se evaluará un pequeño documento reproducible escrito en Rmarkdown y publicado en GitHub.

V. CERTIFICACIÓN

Se otorgará:

1. **Certificado**

- Al participante que lo solicite y que haya aprobado satisfactoriamente las actividades y evaluaciones en las fechas programadas, según los requisitos establecidos consignándose en el documento la calificación obtenida.

2. **Constancia de participación o asistencia**

- Al participante que lo solicite y que no haya aprobado pero que haya realizado como mínimo el 80% de las actividades programadas.

VI. BIBLIOGRAFÍA

1. Brendan R. E. Ansell. Introduction to R - tidyverse.
2. Bret Larger. Statistics 240 Course Notes.
3. Felix Lennert. An(other) introduction to R.
4. Garrett Golemund. Hands On Programming with R.
5. Hadley Wickham. R for Data Science.
6. Nana Kim. Introduction to R (Part 2)
7. Steph Locke. Data manipulation with R
8. Roger D. Peng. R Programming for Data Science.