

# Introducción al Manejo y Visualización de Datos en R

Del 06 de Julio al 20 de Julio del 2018



UNIVERSIDAD PERUANA  
CAYETANO HEREDIA  
FACULTAD DE CIENCIAS Y FILOSOFÍA

FACULTAD DE CIENCIAS Y FILOSOFÍA ALBERTO CAZORLA TALLERI

DEPARTAMENTO ACADÉMICO DE ESTADÍSTICA, DEMOGRAFÍA, HUMANIDADES  
Y CIENCIAS SOCIALES

EDUCACIÓN CONTÍNUA

## I. DATOS GENERALES

1.1	Nombre del curso	<b>INTRODUCCIÓN AL MANEJO Y VISUALIZACIÓN DE DATOS EN R</b>
1.2	Dirigido a	<b>Profesionales y alumnos interesados en el uso del software R.</b>
1.3	Organizado por	<b>Departamento de Estadística, Demografía, Humanidades y Ciencias Sociales de la Facultad de Ciencias y Filosofía de la Universidad Peruana Cayetano Heredia.</b>
1.4	Horas y/o créditos	<b>21 horas / 1 crédito académico.</b>
1.5	Modalidad	<b>Presencial.</b>
1.6	Profesor coordinador	<b>Mg. Gabriel Carrasco Escobar</b>

## II. RESULTADO DE APRENDIZAJE

1. Procesar un archivo de datos estadísticos con R.
2. Aplicar técnicas de análisis exploratorio numérico y gráfico con R.

## III. CONTENIDOS

UNIDAD 1: Introducción a R

1. Introducción a R como lenguaje de programación.
2. Importación y manejo de objetos con R.
3. Tipos y estructuras de datos.

UNIDAD 2. Manejo de datos con R

3. Manejo de datos con comandos base: subset, merge, binding.
4. Medidas de resumen: tendencia central y dispersión.
5. Manejo de datos con dplyr.
6. Introducción a pipelines: tidy, transform.

UNIDAD 3. Análisis gráfico.

7. Visualización de datos con R.
8. Análisis gráfico exploratorio con ggplot2.
9. Introducción a geometría y aesthetics.
10. Extensiones gráficas: mapas y gráficos dinámicos.
11. Introducción a ggmap, plotly y shiny.

#### IV. ESTRATEGIAS DIDÁCTICAS

##### 1. Clases expositivas

Proporcionan la información conceptual y práctica sobre R. Lo realizará el docente en el aula, usando equipos multimedia.

##### 2. Ejercicio de desarrollo

Espacio para que el alumno se ejercite en el uso de las técnicas presentadas. Se utilizará un archivo de datos (proporcionado por el alumno o entregado por el profesor) en el cual se aplicarán progresivamente las técnicas presentadas en el curso; los avances se entregarán según el calendario propuesto por el docente. Cada alumno portará con su laptop para el desarrollo de las actividades prácticas.

##### 3. Proyecto final

Permitirá presentar los logros alcanzados por el alumno. Consiste en la redacción y presentación de un informe de un análisis exploratorio de datos, como resultado del ejercicio de desarrollo.

#### V. EVALUACIÓN

##### 1. Actividades evaluativas:

- a. Avances del ejercicio de desarrollo (AED). Se calculará un promedio sobre 20.
- b. Informe final del ejercicio de desarrollo (IFED). Se calificará sobre 20.

##### 2. Sistema de calificación:

$NF = 0.4 * AED + 0.6 * IFED$   
Se considera aprobado con nota 11.00.

##### 3. Criterios de aprobación:

Para aprobar el curso se considera haber aprobado en las actividades lectivas y haber asistido, cuando menos, al 80 por ciento de clases presenciales.

#### VI. CERTIFICACIÓN

##### 1. Certificado

Cuando el participante haya aprobado satisfactoriamente las actividades según los requisitos establecidos consignándose en el documento la calificación obtenida.

##### 2. Constancia de asistencia

Cuando no obtenga una nota aprobatoria habiendo asistido o cumplido con las actividades programadas según lo establecido.

#### VII. BIBLIOGRAFÍA

1. Brian Caffo. 2016. Statistical inference for data science. Lean Publishing
2. Rafael Irizarry, Michael Love. Data Analysis for the Life Sciences. First printing 2016. Chapman and Hall/CRC
3. Robert Kabacoff. 2015. *R in Action: Data Analysis and Graphics with R*. Manning Publications Co., Greenwich, CT, USA.
4. Brian Caffo. 2016. Developing Data Products in R. Lean Publishing

Enlaces de internet.

1. The R Manuals: <https://cran.r-project.org/manuals.html>
2. R Contributed Documentation: <https://cran.r-project.org/other-docs.html>
3. Stackoverflow: <https://stackoverflow.com/questions/tagged/r>
4. Quick-R: <https://www.statmethods.net/>

## ANEXO 1

### PROFESORES PARTICIPANTES

<b>Grado ó Título</b>	<b>Nombre</b>	<b>Apellidos</b>	<b>Departamento Académico</b>
Mg.	Gabriel	Carrasco Escobar	Departamento de Estadística, Demografía, Humanidades y Ciencias Sociales
BSc.	Edgar	Manrique Valverde	Departamento de Estadística, Demografía, Humanidades y Ciencias Sociales
BSc.	Jorge	Ruiz Cabrejos	Departamento de Estadística, Demografía, Humanidades y Ciencias Sociales

## ANEXO 2

### PROGRAMACIÓN DE ACTIVIDADES

Sesión	Día y fecha	Hora	Contenido	Docente responsable
1	Viernes 06/07/2018	10:00 - 13:00	Presentación del curso. Introducción a R como lenguaje de programación: Importación y manejo de objetos Tipos y estructura de datos. Fuentes de información para trabajo Final.	G. Carrasco J. Ruiz
2	Lunes 09/07/2018	10:00 - 13:00	Manejo de datos con comandos base: subset, merge, binding. Medidas de resumen: medidas de tendencia central, medidas de dispersión	G. Carrasco J. Ruiz
3	Miércoles 11/07/2018	10:00 - 13:00	Visualización de datos con comandos base: histogramas, boxplots, frecuencias y mosaicos	G. Carrasco J. Ruiz
4	Viernes 13/07/2018	10:00 - 13:00	Manejo de datos con dplyr. Introducción a pipelines: tidy, transform.	G. Carrasco E. Manrique
5	Lunes 16/07/2018	10:00 - 13:00	Análisis gráfico exploratorio con ggplot2. Introducción a geometría y aesthetics	G. Carrasco E. Manrique
6	Miércoles 18/07/2018	10:00 - 13:00	Extensiones gráficas: Mapas y gráficos dinámicos. Introducción a ggmap y plotly.	G. Carrasco E. Manrique
7	Viernes 20/07/2018	10:00 - 13:00	Presentación de Trabajo Final	G. Carrasco E. Manrique