

**UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE FILOSOFIA Y LETRAS**

DEPARTAMENTO:DEPARTAMENTO DE ANTROPOLOGÍA

SEMINARIO: Herramientas cuantitativas para ciencias sociales con R

MODALIDAD DE DICTADO: VIRTUAL (según Res. (D) N° 732/20 y normativa específica dispuesta a los efectos de organizar el dictado a distancia)

PROFESOR/A: Cardillo Marcelo

CUATRIMESTRE: 1° bimestre

AÑO: 2024

CÓDIGO N°:

UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE FILOSOFIA Y LETRAS
DEPARTAMENTO DE CIENCIAS ANTROPOLÓGICAS
SEMINARIO: Herramientas cuantitativas para ciencias sociales con R
CUATRIMESTRE Y AÑO: 1º Bimestre 2024
CODIGO N°:

PROFESOR/A: Dr. Marcelo Cardillo

EQUIPO DOCENTE COLABORADOR: Lic. Eugenia Carranza

a. Fundamentación y descripción

Un aspecto fundamental en cualquier etapa de la investigación en ciencias sociales como la Antropología y la Arqueología es el la selección, manejo, curado y procesamiento de la información cualitativa y cuantitativa. De este hecho, surge la necesidad de comprender las herramientas fundamentales del manejo numérico de los datos. La enseñanza de estos métodos tiene tanto una dimensión teórica como otra práctica, la que puede realizarse a partir programas (software) específicos. Sin embargo, la práctica en el manejo de estas herramientas computacionales queda actualmente fuera de la currícula de las carreras dictadas en la facultad, debido al tiempo y recursos que esto demanda y por ende, la dificultad de darle un espacio más sustancial dentro del dictado de las materias afines; como en el caso Métodos Cuantitativos en Antropología. Creemos que esta falencia puede ser reducida a partir del dictado de un seminario de grado focalizado en la práctica y aplicación de las herramientas estadísticas cuya fundamentación teórica y práctica se dicta en la mencionada cátedra durante el primer cuatrimestre. De esta manera se podrá integrar ambas dimensiones (conceptual y práctica) a través del aprendizaje de un programa de lenguaje libre, abierto como R.

b. Objetivos del seminario:

El objetivo central de este seminario es brindar a los estudiantes las herramientas metodológicas necesarias para realizar la selección, curado y aplicación de las técnicas estadísticas básicas, las cuales están comprendidas dentro del dictado de la materia Métodos Cuantitativos en Antropología.

El objetivo secundario es acercar a los estudiantes una herramienta estadística poderosa, de lenguaje abierto y acceso libre, como es R. Este programa es uno de los más empleados en la investigación en todos los campos disciplinares en la actualidad, incluyendo el manejo de todo tipo de datos y permite su integración con inteligencia artificial abierta (Open IA) que asiste en la escritura, análisis e interpretación de código en R.

Por último y no de menor importancia, es lograr, a través de la fusión de las herramientas teóricas con la práctica, un aprendizaje integral que estimule el desarrollo del pensamiento cuantitativo en los estudiantes de la carrera.

c. **Contenidos organizados en unidades temáticas:**

1. **R como lenguaje.**

Características del lenguaje abierto y replicación del estudio científico. Concepto de dato y variable en R. Instalación del programa, paquetes, repositorios y bases de datos. Selección del directorio de trabajo. Introducción a R studio. Actualización de los paquetes. El empleo de inteligencia artificial Open AI en R alcances y limitaciones.

2. **Concepto de población, muestra, dato y variable. Importar y exportar datos a R. Curado y manejo inicial de los datos.**

Qué es una población y una muestra. Tipos de datos y variables en R. Objetos. Vectores, matrices. Construcción de matrices. Importar datos en R de distintos formatos. Configuración de R studio. Ayuda en R e interpretación de código a través de Open IA.

3. **Operaciones cuantitativas básicas con vectores y matrices.**

Tipos de datos y su manejo: Arrays, data.frames, listas y matrices en R. Como adicionar, sustraer o renombrar elementos. Operadores lógicos, numéricos y relacionales. Curado inicial de los datos, manejo de datos ausentes. Exportar datos de R. El concepto y utilidad de los script o comandos. Uso de Open IA como asistente virtual en el curado de datos

4. **Estadística Univariada. Análisis descriptivo de datos cualitativos y cuantitativos.**

Funciones y paquetes básicos para la estadística descriptiva. Aplicación de las funciones summary () y str (). Cálculo de medidas de tendencia central, localización y dispersión más comunes. Aplicación de operadores para el cálculo manual de estadísticos. Operadores básicos para la generación de gráficos en R. Gráficos de barras, histogramas y gráfico de cajas.

5. **Estadística Univariada y bivariada. Análisis descriptivo de datos cuantitativos.**

Transformación de datos. Comparación entre dos variables cuantitativas, la correlación r de Pearson y de Spearman. Gráficos de dispersión. Herramientas básicas para generar, modificar y salvar gráficos en R en distintos formatos.

6. **Estadística inferencial, aspectos fundamentales.**

El lenguaje y los tipos de regresión en R. Generación y control de los supuestos del modelo de regresión. Representación gráfica. Otros tipos de regresión.

7. **Test de hipótesis para datos cualitativos y cuantitativos.**

Estimación puntual y por intervalos. Test de hipótesis. Teorema central y del límite con ejemplos y aplicaciones en R. Test paramétricos y no paramétricos: sus características básicas y el control de los supuestos. Transformaciones útiles. Tablas de contingencia y test de Chi2, Test de la t . y Prueba U de Mann Whitney.

8. **Error estadístico y potencia del test, selección del tamaño de la muestra.**

Error de la prueba de hipótesis (errores de tipo I y II) el rol del tamaño de la muestra. Efecto tamaño y potencia del test.

d. **Bibliografía obligatoria:**

1.

Cardillo Marcelo. 2018. Introducción a la estadística para arqueólogos con R. Primera parte: métodos descriptivos e inferenciales uni y bivariados. 163 p. Editado por IMHICIHU-CONICET.

Grolemund Garrett y Hadley Wickham. R for Data Science. <https://es.r4ds.hadley.nz/> (en castellano). Capítulo 1 y 2.

2.

Cardillo Marcelo. 2018. Introducción a la estadística para arqueólogos con R. Primera parte: métodos descriptivos e inferenciales uni y bivariados. 163 p. Editado por IMHICIHU-CONICET, pp. 17-48.

Grolemund Garrett y Hadley Wickham. R for Data Science. <https://es.r4ds.hadley.nz/> (en castellano). Capítulo 3.

Freddy Hernández y Olga Usuga. 2020. Manual de R. Capítulos 3, 4 y 8
<https://fhernanb.github.io/Manual-de-R/>

3.

Cardillo Marcelo. 2018. Introducción a la estadística para arqueólogos con R. Primera parte: métodos descriptivos e inferenciales uni y bivariados. 163 p. Editado por IMHICIHU-CONICET, pp. 17-48.

Grolemund Garrett y Hadley Wickham. R for Data Science. <https://es.r4ds.hadley.nz/> (en castellano). Capítulo 4 y 5.

4.

Cardillo Marcelo. 2018. Introducción a la estadística para arqueólogos con R. Primera parte: métodos descriptivos e inferenciales uni y bivariados. 163 p. Editado por IMHICIHU-CONICET, pp. 49-63.

Grolemund Garrett y Hadley Wickham. R for Data Science. <https://es.r4ds.hadley.nz/> (en castellano). Capítulo 3 y 7.

Freddy Hernández y Olga Usuga. 2020. Manual de R. Capítulo 9.
<https://fhernanb.github.io/Manual-de-R/>

5.

Cardillo Marcelo. 2018. Introducción a la estadística para arqueólogos con R. Primera parte: métodos descriptivos e inferenciales uni y bivariados. 163 p. Editado por IMHICIHU-CONICET, pp. 67-77.

Freddy Hernández y Olga Usuga. 2020. Manual de R. Capítulo 20.
<https://fhernanb.github.io/Manual-de-R/>. Capítulo 13.

6. Cardillo Marcelo. 2018. Introducción a la estadística para arqueólogos con R. Primera parte: métodos descriptivos e inferenciales uni y bivariados. 163 p. Editado por IMHICIHU-CONICET, pp. 78-89.

Grolemund Garrett y Hadley Wickham. R for Data Science. <https://es.r4ds.hadley.nz/> (en castellano). Capítulo 22 y 23.

Amat Rodrigo J. 2016. Correlación lineal y Regresión lineal simple.
https://www.cienciadedatos.net/documentos/24_correlacion_y_regresion_lineal

7.

Cardillo Marcelo. 2018. Introducción a la estadística para arqueólogos con R. Primera parte: métodos descriptivos e inferenciales uni y bivariados. 163 p. Editado por IMHICIHU-CONICET, pp. 90-140

Freddy Hernández y Olga Usuga. 2020. Manual de R. Capítulo 20.
<https://fhernanb.github.io/Manual-de-R/>

Amat Rodrigo J. 2017. T-Test. https://www.cienciadedatos.net/documentos/12_t-test

Amat Rodrigo J. 2017. Test de Wilcoxon-Mann-Whitney como alternativa al t-test.
https://www.cienciadedatos.net/documentos/17_mann%E2%80%93whitney_u_test

Amat Rodrigo J. 2017. Test estadísticos para variables cualitativas: test exacto de Fisher, chi-cuadrado de Pearson, McNemar y Q-Cochran.
https://www.cienciadedatos.net/documentos/22.2_test_exacto_de_fisher_chi-cuadrado_de_pearson_mcnemar_qcochran

8.

Cardillo Marcelo. 2018. Introducción a la estadística para arqueólogos con R. Primera parte: métodos descriptivos e inferenciales uni y bivariados. 163 p. Editado por IMHICIHU-CONICET, pp. 155-153.

Bibliografía complementaria para cada una de las unidades temáticas:

Kabacoff Robert. 2013. R in Action: Data Analysis and Graphics with R.

Mike Baxter y Hilary Cool (2016). Basic Statistical Graphics for Archaeology with R: Life Beyond Excel.
https://www.researchgate.net/publication/320592558_Basic_Statistical_Graphics_for_Archaeology_with_R_Life_Beyond_Excel

Mike Baxter (2015). Notes on Quantitative Archaeology and R.
https://www.researchgate.net/publication/277931925_Notes_on_Quantitative_Archaeology_and_R

Fuentes:

La ejercitación para las distintas unidades se realizará a través de bases de datos disponibles en R como `archdata()` de bases de datos arqueológicas de distinto tipo, `titanic()` datos de composición, género, clase social y supervivencia diferencial de los pasajeros del Titanic, así como bases de datos generadas a partir de la votación del senado en relación a la legalización del aborto en el 2018. Se realizarán ejercitaciones de desarrollo individual y asistidas mediante Open IA para la detección de errores y la escritura eficiente de código abierto en R.

Software:

R Development Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>. Open IA (ChatGPT 3.5) <https://openai.com/>. R Studio. <https://posit.co/products/open-source/rstudio/>

e. Organización del dictado de seminario: carga horaria y distribución entre modalidades de clase (teóricos, teórico-prácticos, trabajos prácticos, talleres u otras), carácter de las actividades para las distintas modalidades de clase planificadas:

Seminario: Bimestre de verano

Total de horas semanales: 8 hs. (modalidad sincrónica: 4 horas teórico-prácticas y modalidad diacrónica 4 horas prácticas).

Total de horas cuatrimestrales: 64 hs.

Modalidad de trabajo

El seminario se dictará en modalidad virtual de 8 horas por semana, dividido en dos módulos de 2 (dos) clases semanales de 4 horas cada una. El primer módulo será de 2 (dos) horas y de carácter sincrónico. En esta modalidad se prevé realizar una primera parte de la clase con una presentación de cada unidad temática en base a la bibliografía obligatoria, ejemplificando con casos de estudio y datos acordes. Los casos de estudio incluyen tanto datos arqueológicos como antropológicos o de interés social. Al finalizar esta clase se entregará la consigna práctica.

La segunda parte de la clase será práctica de dos (2) horas de duración y de carácter diacrónico. Durante esta modalidad se resolverán preguntas específicas de la realización de las consignas prácticas a través del Campus Virtual de la FFyL y de distintos medios virtuales y de ser necesario, se programará un evento sincrónico.

La organización de los temas planteados, tal como se mencionó en la fundamentación está en base a los que están incorporados dentro de la currícula de la carrera, en particular los comprendidos por la materia Métodos Cuantitativos en Antropología. Un requisito para la

realización de este seminario es el empleo de una computadora personal y/o de escritorio donde se descargarán los programas y tutoriales. Al respecto, es importante destacar que R funciona a través de scripts o líneas de comando. Este factor facilita la ejecución del programa y la resolución de un problema de manera no ambigua, ya que la línea de comando puede ser almacenada como texto y compartida entre usuarios. Asimismo, esto facilitará la práctica y el manejo del software requerido, así como la aplicación de las distintas herramientas metodológicas.

El programa utilizado (R) es de acceso libre. Asimismo, la bibliografía obligatoria está digitalizada bajo la misma modalidad de acceso.

Para los encuentros sincrónicos se realizarán a través de la plataforma Microsoft Teams que provee la FFyL, para cumplir con el módulo práctico así como con otros medios de comunicación durante la cursada de utilizará el Campus Virtual de la Facultad así como las Plataformas Google Classroom y Teams.

f. Organización de la evaluación: régimen de promoción y formas y criterios de evaluación a utilizar.

La evaluación del seminario comprenderá tomar en cuenta dos instancias de calificación. La primer nota se genera por el cumplimiento de la ejercitación programada para cada unidad temática, que será entregada con regularidad por los alumnos durante la cursada. La segunda calificación proviene de la realización del trabajo final. Este trabajo de carácter monográfico será el análisis integral de un caso de estudio antropológico o arqueológico que incorpore las distintas unidades temáticas tanto en aspectos conceptuales como prácticos. La calificación final resultará del promedio de la nota de cursada y del trabajo final integrador.

Es condición para alcanzar la REGULARIDAD del seminario:

- i. entregar la totalidad de los trabajos prácticos resueltos;
- ii. obtener una nota promedio con un mínimo de 4 (cuatro) para la cursada (la nota emerge del promedio de todos los trabajos prácticos).

Los/as estudiantes que cumplan con los requisitos mencionados podrán presentar el trabajo final integrador que será calificado con otra nota. La calificación final resultará del promedio de la nota de cursada y del trabajo final integrador.

Si el trabajo final integrador fuera rechazado, los/as interesados/as tendrán la opción de presentarlo nuevamente antes de la finalización del plazo de vigencia de la regularidad. El/la estudiante que no presente su trabajo dentro del plazo fijado, no podrá ser considerado/a para la aprobación del seminario.

VIGENCIA DE LA REGULARIDAD: El plazo de presentación del trabajo final de los seminarios es de 4 (cuatro) años posteriores a su finalización.

g. Recomendaciones

Haber realizado previamente -o estar cursando en ese mismo cuatrimestre- la materia Métodos Cuantitativos en Antropología, ya que allí se dictan los conceptos teóricos necesarios para una comprensión integral de los métodos que serán aplicados.

Firma Firma

Aclaración:Dr Marcelo Cardillo
Cargo: Profesor

Aclaración:Lic. Eugenia Carranza
Cargo: Docente colaborador