

EP01 (FORMATIVO): ESTADÍSTICA DESCRIPTIVA CON R



Actividad: estadística descriptiva con R

Contexto

Sabemos que uno de los usos de los modelos estadísticos es **describir o caracterizar un conjunto de datos**, lo que aprendimos en la enseñanza secundaria y en cursos anteriores con el nombre de **estadística descriptiva**. Son muchas las herramientas, pero es importante saber usar la que es adecuada para lo que se quiere describir.

Objetivos de aprendizaje

1. Recordar que existen numerosas alternativas para describir datos.
2. Seleccionar una herramienta adecuada considerando los datos que se tienen y las características que se quiere describir.
3. Describir datos usando herramientas disponibles en el entorno R.

Éxito de la actividad

1. El equipo revisa los datos entregados y reconoce la(s) variable(s) involucrada(s) en la pregunta.
2. El equipo elige una herramienta de visualización adecuada para los datos y pregunta(s) planteada(s).
3. El equipo consigue caracterizar los datos entregados, usando el entorno R, y responder la(s) pregunta(s) planteada(s) en consecuencia.

Actividades

1. Obtener el archivo (en la carpeta compartida para este ejercicio) con una selección de datos públicos recopilados por la encuesta Casen 2017 en la Región Metropolitana. La siguiente es una breve descripción de las columnas seleccionadas:
 - **folio**: identificador del hogar
 - **o**: número de orden de la persona dentro del hogar
 - **id.vivienda**: identificador de la vivienda
 - **hogar**: identificación del hogar en la vivienda
 - **region**: region
 - **provincia**: provincia
 - **comuna**: comuna
 - **ing.comuna**: posición en el ranking histórico del ingreso de la comuna (ascendente)
 - **zona**: área geográfica (Urbano, Rural)
 - **sexo**: sexo de la persona registrada
 - **edad**: edad de la persona registrada
 - **ecivil**: estado civil de la persona registrada
 - **ch1**: situación ocupacional de la persona registrada
 - **ytot**: ingreso total
2. Considerar la pregunta asignada a su equipo.
3. Discutir y consensuar qué medidas estadísticas (media, mediana, moda, etc.) y qué forma gráfica ayudaría a responder la pregunta asignada.
4. Construir un script en R que produzca los estadísticos y el gráfico seleccionado.
5. Documentar en el script el objetivo de cada estadístico, una descripción de lo que se ve en el gráfico y la respuesta a la pregunta planteada.
6. Subir la respuesta con el nombre "EP01-respuesta-equipo-*i*", donde *i* es el número de equipo asignado, a la carpeta compartida para este fin antes de las **22:00 hrs. del jueves**.
7. Calcular $j = ((i + 1) \bmod n) + 1$, donde *n* es la cantidad de equipos; descargar la respuesta del equipo *j* y aplicar la rúbrica publicada para este EP.

8. Subir la rúbrica completada, en formato PDF, con el nombre "EP01-resultado-equipo-j" a la carpeta compartida antes de las **22:00 hrs. del viernes.**

Preguntas

Equipo 1:

¿Se encuestaron más o menos la misma cantidad de gente en cada provincia de la RM?

Equipo 2

¿Tienen hombres y mujeres ingresos similares en la RM?

Equipo 3

¿Se distribuye de igual manera la situación ocupacional de los hombres que viven en áreas rurales y quienes viven en áreas urbanas de la RM?

Equipo 4

¿Son similares los ingresos registrados en las diferentes provincias de la RM?

Equipo 5

¿Van los ingresos de las mujeres de la RM incrementándose con la edad?

Equipo 6

¿Cómo diría que es el ingreso de los hombres de la RM? (simétrico/asimétrico, concentrado/disperso, unimodal/multimodal, etc.)

Equipo 7

¿Tiene relación el ingreso de las mujeres de la RM con el riqueza del municipio donde habita?

Equipo 8

¿Cómo diría que es el ingreso de las mujeres de la RM? (simétrico/asimétrico, concentrado/disperso, unimodal/multimodal, etc.)

Equipo 9

El nivel de ingreso de las mujeres de la RM ¿varía con el estado civil?

Equipo 10

Los ingresos de los habitantes de la RM, ¿van incrementándose con la edad de forma similar en hombres y mujeres?

Criterios de evaluación

Pregunta	Categoría	Nivel de logro	Puntos
Obtención y carga de datos	Tabla de datos	Escriben código R que carga el archivo CSV y asegura que los tipos de variables se reflejen en el data frame resultante	3
	Variables	Mencionan explícitamente todas las variables presentes en el archivo CSV que son relevantes para responder la pregunta asignada, reconociendo correctamente sus tipos	3
Estadística descriptiva con R	Medida estadística	Seleccionan una(s) medida(s) estadística adecuada(s) para responder la pregunta asignada y, usando R, la(s) calculan correctamente	4
	Dato agrupado	Construyen un(os) gráfico(s) en R adecuado(s) para resumir las variables de interés y responder la pregunta asignada	4
Cuidado por el detalle	Código fuente	El script R es ordenado, está bien indentado, no contiene sentencias espurias y está bien comentado	2
	Comentarios y respuestas	Las respuestas y comentarios están escritos con buena ortografía (<5 errores) y redacción, usando vocabulario propio de la disciplina y el contexto del problema	2
Total ejercicio práctico			18
Porcentaje de la puntuación máxima para nota 4,0			60.0 %

Sumario de calificaciones

Ocultado a los estudiantes	No
Participantes	25
Fecha de entrega	jueves, 20 de marzo de 2025, 13:00
Tiempo restante	1 día 23 horas

VER TODOS LOS ENVÍOS

CALIFICACIÓN

◀ AVISOS

Ir a...

DATOS, RÚBRICA, RESPUESTAS Y REVISIONES EP01 ▶

Síguenos en:



Dirección de Pregrado - VRA

✉ E-mail: soporte.uvirtual@usach.cl

En caso de presentar problemas con sus datos institucionales, validar datos en <https://mail.usach.cl>, saliendo de su sesión de correo actual. No ocupe datos guardados.