

R para Científicos Una Introducción Práctica

Tarea Evaluación Final

Ranking Mundial de Universidades

Dispones de la base de datos correspondiente al ranking mundial de universidades. En ella se muestran distintas variables relacionadas el ranking de las mejores universidades alrededor del mundo. La intención de este estudio es generar funciones útiles y gráficos que permitan sugerir que variables inciden sobre ese ranking. Los recursos suministrados en este tópico especial (paquetes, uso de funciones, filtrado de vectores, etc) te permitirán resolver las siguientes preguntas.

Los datos se encuentran en el archivo (world-university-rankings/timesData.csv) perteneciente a **Times Higher Education World University Ranking** (<https://www.timeshighereducation.com/world-university-rankings>) Esta institución fue fundada en 2004, proporciona la lista definitiva de las mejores universidades del mundo, evaluadas a través de la enseñanza, la investigación, el panorama internacional, la reputación y mucho más. Los datos de esta institución son de confianza para los gobiernos y las universidades y son un recurso vital para los estudiantes, ayudándoles a elegir dónde estudiar. Sin embargo ha sido criticada por su comercialización y por socavar las instituciones que no instruyen en inglés.

El dataset es un archivo csv con 2603 filas y 14 variables (descritas al final de esta asignación). Con el mismo responderás una serie de preguntas relacionadas con el ranking universitario aplicando las herramientas obtenidas en este.

- 1) Crea una data.frame/tibble denominada **worldUni**. ¿Qué países están representados? ¿Cuál es el país con más universidades en este ranking?
- 2) Crea una función denominada rankUni que te muestre en pantalla el nombre la universidad al suministrarle los argumentos: rango(world_rank) y el año de la evaluación (year). La salida debe ser un vector carácter de longitud 1. Por ejemplo al aplicar rankUni(rank = 8, year=2016) arroja como resultado "Imperial College London". Aplique su función a rangos 1,10 y 100 para los años 2015 y 2016 respectivamente.
- 3) Un estudiante interesado en el puntaje de estas universidades quiere saber de manera inmediata si introduciendo un país aparecerán las universidades y sus rangos. Diseña una función denominada uniCountry que le de respuesta a este estudiante. Esta función usa como argumentos el país(country) y el año(usa por defecto el año 2016). En este caso la salida de la función debe ser una tabla que muestre las universidades de ese país y su ranking mundial. Ten en cuenta que hay países que no están registrados en esta data frame y si lo introduces no se mostrara resultado alguno. Prueba tu función con Argentina, Chile, Colombia, México y Alemania.
- 4) Crea una nueva variable categórica del ranking universitario utilizando tu propio criterio y representa gráficamente para los países americanos su distribución de frecuencia en el año 2016.
- 5) Representa gráficamente la variación anual del ranking universitario para las primeras 10 universidades evaluadas en este estudio.
- 6) Elige un método(generación de data.frame, uso de gráficos) que te permita explorar las universidades con mayor proporción de estudiantes femeninas. ¿Tiene incidencia sobre el ranking evaluado?
- 7) Existe una relación lineal importante entre las variables mostradas en este estudio. Usa gráficos o análisis de regresión que permitan justificar tu respuesta.
- 8) El paquete gapminder trae la información relacionada con continentes, usando las funciones right_join o left_join trata de unir la información de los continentes a tu data frame y realiza una evaluación del ranking universitario por continente. Justificalo gráficamente.

INFORMACIÓN DE CADA VARIABLE

world_rank - Rango mundial para la universidad. Contiene rangos e intervalos (ej. =94 and 201-250).

- **university_name**- nombre de la universidad.
- **country**- país de cada universidad
- **teaching**- puntaje de enseñanza universitaria (ambiente de aprendizaje).
- **international**- Puntaje internacional de la universidad (personal, estudiantes, investigación).
- **research**- Puntaje universitario para investigación (volumen, ingresos y reputación).
- **citations**- Puntaje universitario para citas (influencia de la investigación).
- **income**- Puntaje universitario para ingresos de la industria (transferencia de conocimiento).
- **total_score**- Puntuación total de la universidad, utilizada para determinar el rango.
- **num_students**- Número de estudiantes en la universidad
- **student_staff_ratio**- Número de estudiantes divididos por número de empleados.
- **international_students**- Porcentaje de estudiantes que son internacionales
- **female_male_ratio**- Relación estudiante femenino-estudiante masculino
- **year**- Año del ranking (2011 a 2016 incluido).