

SOBRE LAS PRÁCTICAS DE LABORATORIO EN LA ESTADÍSTICA DE UN GRADO EN TURISMO

J. CASTELLÓ

castelló@uji.es

Universitat Jaume -I / Departamento de Matemáticas-Castellón

C. GALINDO

galindo@uji.es

Universitat Jaume -I / Departamento de Matemáticas-Castellón

V. MARTÍNEZ

martinez@uji.es

Universitat Jaume -I / Departamento de Matemáticas-Castellón

J. MOMPARDLER

momparle@uji.es

Universitat Jaume -I / Departamento de Matemáticas-Castellón

RESUMEN: El turismo es una actividad económica que representa aproximadamente el 10% del PIB en España. El Grado de Turismo contiene en su primer curso una asignatura de Estadística básica que en la Universitat Jaume I realiza prácticas de laboratorio utilizando el paquete de software estadístico R. En este trabajo, mostramos como pueden realizarse prácticas de laboratorio utilizando, de modo sencillo, un software aparentemente complicado. Para ello realizamos una actividad relacionada con la titulación y de actualidad. Además, mostramos un caso práctico relacionado con el futuro desarrollo profesional del estudiante.

Palabras claves: Estadística, Grado en Turismo, software R.

ABSTRACT: Tourism is an economic activity that represents approximately 10% of GDP in Spain. The Degree of Tourism in its first course contains a course in basic statistics at the Universitat Jaume I laboratory practice using the statistical software package R. In this paper, we show how labs can be performed using, so simple, seemingly complicated software. We carry out an activity related to the degree and topical. Moreover, we show a case study related to the student's future professional development.

Keywords: Statistical. Degree of Tourism. R-Software.

1. Introducción

El turismo es una actividad económica de carácter social que busca el desarrollo de actividades que puedan generar momentos especiales para los seres humanos. En

general, los turistas pretenden visitar un lugar atractivo con el objeto de divertirse, negociar o llevar a cabo cualquier otra actividad relacionada con el ocio o la cultura.

En las últimas décadas, el turismo se ha convertido en una actividad habitual de una buena parte de los habitantes de los países industrializados y como consecuencia es una gran fuente de recursos para los países receptores de turistas. En particular en países como España, esta actividad representa una buena parte del producto interior bruto (PIB) del país. Así en el año 2010 el turismo en España alcanzó el 10,23% de su PIB. Cifras análogas pueden darse dentro de la Comunidad Valenciana (lugar dónde ejercen la docencia, los firmantes de este artículo). La actividad turística se concentra especialmente en las provincias de Alicante y Castellón, donde además del tradicional turismo de sol y playa están cobrando fuerza otras modalidades turísticas que contribuyen a generar empleo estable y a reducir la estacionalidad.

No es pues de extrañar la existencia de títulos universitarios cuyo objetivo sea formar profesionales que puedan incrementar la calidad del servicio. Por lo que respecta a España, según datos de la Estadística Universitaria en España (<https://www.ine.es>) los estudiantes matriculados en la Diplomatura de Turismo (16.364), y los matriculados en el Grado de Turismo según la administración (29.532), permiten estimar que en el curso 2010/11 había 45.896 alumnos inscritos en las titulaciones de Turismo. Además, dentro del ámbito de las Ciencias Sociales y Jurídicas los estudios de Turismo ocupaban el tercer lugar, tras Administración y Dirección de Empresas y Derecho.

Es obvio que un Graduado en Turismo debe tener la capacidad de reunir e interpretar datos relevantes para poder emitir juicios que incluyan una reflexión sobre un tema relevante como éste. A modo de ejemplo, los países receptores de turismo tienen un gran interés en tener información veraz de datos económicos relacionados con el turismo, como pueden ser:

- Los movimientos de los flujos turísticos.
- El gasto promedio que produce.
- Su influencia en la balanza de pagos.
- La inversión que necesita.
- El efecto cambiario de la actividad turística de la que venimos hablando.

A nivel más reducido, cualquier operador turístico necesita también conocer datos similares a los anteriores (con la excepción del último) para que su actividad sea lucrativa y pueda obtener los beneficios esperados mediante la inversión en los lugares óptimos y la oferta de productos que se adecúen a las expectativas de los turistas. Parece pues evidente que éste es un objetivo básico para que el sector crezca y se desarrolle correctamente.

Por consiguiente será obligación de los gobiernos de los estados poner a disposición de los operadores turísticos los datos que les permitan realizar su tarea del mejor modo posible. En este sentido, existen organizaciones nacionales y supranacionales dedicadas a esta tarea. Así a nivel global, la Organización Mundial del Turismo (OMT) es reconocida por las Naciones Unidas como la organización competente para recoger, analizar, publicar, uniformar y mejorar las estadísticas de turismo. La OMT se ha ocupado de liderar iniciativas encaminadas a medir la contribución económica del turismo a las economías nacionales. Destaca entre todas ellas la Cuenta Satélite de Turismo (CST), un marco para reunir estadísticas de turismo y organizar los datos de forma coherente con las cuentas nacionales y, por lo tanto, de manera fiable y comparable en el plano internacional. En particular en 2010 se celebró el Día Mundial de la Estadística, que fue propuesto por la División de Estadística de las Naciones Unidas y proclamado por la Asamblea General de las Naciones Unidas, y tuvo por objeto sensibilizar de los numerosos logros de las estadísticas oficiales. También a nivel español está el Instituto Nacional de Estadística que ofrece, entre otra información, datos referentes a la actividad turística del Reino de España.

Por último, cabe resaltar que durante este año 2013 se está celebrando el Año Internacional de la Estadística, que tiene por objeto el reconocimiento a nivel mundial de la contribución que la Estadística ha realizado a lo largo de los años al progreso de nuestra sociedad. Los principales objetivos de Statistics2013 son:

- Incrementar el conocimiento sobre el impacto que la Estadística tiene en todos los aspectos de la sociedad.
- Promocionar la estadística como profesión, especialmente entre los jóvenes.
- Promover el desarrollo de las Ciencias Estadísticas.

Resulta ya claro a estas alturas que un estudiante de un Grado de Turismo debería poder sacar conclusiones de los datos que se le ofrecen tanto a nivel global como nacional. Esto podrá alcanzarse a través de un conocimiento de los procedimientos y métodos de la Estadística y de la capacidad de entender los grandes números estadísticos que se le pone a su alcance. Así, de poco le servirá entender con razonable profundidad el significado de los diferentes valores y coeficientes numéricos que ofrece la Estadística si no es capaz de tener en sus manos datos fiables para interpretarlos y, recíprocamente, si es capaz de recolectar esos datos e incluso de obtener los valores estadísticos más interesantes pero no es capaz de interpretarlos, tampoco se habrá logrado el objetivo.

En este artículo se pretende explicar cómo se enseña el doble proceso, antes mencionado, a los alumnos de la asignatura *Introducción a la Estadística* (TU0907) del Grado en Turismo de la Universitat Jaume I de Castellón.

2. La asignatura Introducción a la Estadística

Esta es una asignatura de 6 créditos que se imparte en el segundo semestre del primer curso de los futuros Graduados en Turismo. En ella, y según su guía docente, se pretende que el estudiante sepa identificar la necesidad de información estadística que requiere el sector turístico, así como manejar las fuentes estadísticas existentes. Además el alumno también debería ser capaz de analizar, sintetizar y resumir la información estadística mencionada. Aún más, se le pide que sobre ello sea capaz de elaborar y presentar un informe de los resultados obtenidos.

Una lectura detallada de lo anterior, deja bien claro que existe un fuerte desajuste entre los objetivos de la asignatura y el muy escaso tiempo para su desarrollo en el aula. Si acudimos de nuevo a la guía docente, nos encontramos con los siguientes contenidos:

- Descripción de muestras de datos univariantes.
- Descripción de muestras de datos bivariantes.
- Descripción de series de datos temporales.
- Números índices.
- Probabilidad.
- Variable aleatoria.
- Modelos probabilistas de variable aleatoria.
- Inferencia Estadística.

Su desarrollo en cierta profundidad claramente requiere mucho más de 6 créditos. Al final de este artículo aparecen como referencias aquellos textos que se sugiere a los alumnos para seguir la asignatura [1,2,3,4]. No obstante la referencia esencial es [1]. En todos se refleja el objetivo primordial que pretendemos mostrar en este artículo. Se trata de enseñar a los estudiantes a extraer de las fuentes públicas o privadas de que dispongan los datos que necesiten y luego hacer un análisis de ellos. En nuestra opinión es mejor que dominen los cálculos estadísticos sencillos y sepan interpretarlos correctamente. Sabiendo que cuando la magnitud del problema sea muy alta o la interpretación menos clara convendrá acudir a expertos en Estadística que puedan aconsejarle con las muchas y sofisticadas herramientas que se conocen en la actualidad. Con respecto a éstas, al menos el futuro graduado debería saber que existen.

3. La herramienta informática: el programa R

R es un lenguaje y un entorno de programación para análisis de datos y gráficos estadísticos. Es un proyecto GNU y se ha desarrollado en los Laboratorios Bell (antes AT&T, ahora Lucent Technologies) por John Chambers principalmente. R ofrece una gran variedad de datos relacionados con la Estadística. Trabaja con modelos lineales y no lineales, realiza tests estadísticos clásicos, análisis de series temporales, clasificación, clustering y un largo etcétera. También trata con técnicas gráficas. Este lenguaje es muy utilizado en el campo estadístico en general y tiene la ventaja de usar código abierto. Puesto que es de uso común por parte de la comunidad universitaria internacional, R recibe mejoras, e implementaciones de las últimas novedades, de modo constante.

Uno de los puntos fuertes de R es que permite producir informes de calidad en los que se puede incluir símbolos y fórmulas matemáticas. Así, si queremos elaborar un documento dinámico o programable que combine texto y resultados estadísticos, tanto numéricos como gráficos, la mejor opción es utilizar LaTeX y R. Con ello, se obtiene un documento con extensión .tex, que contiene el texto original y los resultados, cálculos y gráficos demandados por el código R insertado. Posteriormente, el archivo .tex se procesará como es habitual en LaTeX para obtener el documento definitivo: .ps, .pdf, etc.

Algo importante para nosotros, por la facilidad que da al alumno en su acceso y uso, es que R está disponible como software libre. Se compila y ejecuta en una amplia variedad de plataformas UNIX y sistemas similares (incluyendo FreeBSD y Linux), Windows y MacOS.

Siendo un poco más explícitos, R es un conjunto integrado de servicios de software para la manipulación de datos, cálculo y representación gráfica. En particular, R incluye:

- Datos: un manejo eficaz y cómodo de datos y almacenamiento.
- Análisis de datos: una amplia colección de herramientas para el análisis de datos.
- Matrices: un conjunto de operadores para los cálculos de estructuras ordenadas, en particular de matrices.
- Gráficos: la posibilidad de obtener gráficos tanto en pantalla como en papel para la interpretación de los datos recibidos.
- Programación: un lenguaje de programación simple y eficaz para desarrollar los cálculos necesarios.

Finalmente, R permite a los usuarios añadir nuevos recursos adicionales mediante la adición de nuevas funciones. Para las tareas computacionalmente intensivas, puede vincularse a C, C++ y Fortran acortando el tiempo de ejecución. Los usuarios

avanzados pueden escribir código C para manipular objetos de investigación directamente. Además R se puede extender a través de paquetes. Hay alrededor de ocho paquetes suministrados con la distribución de R y muchos más están disponibles a través de la familia CRAN de sitios de Internet que cubren una gama muy amplia de la Estadística moderna.

4. La práctica concreta propuesta: Introducción a la Estadística Descriptiva y al R

4.1. La web del INE

Para empezar la práctica, lo primero que se pide al estudiante es que ingrese en Internet y luego a la página web del INE: <http://www.ine.es/>. En ella se accede a algunos contenidos: primero *INEbase*, luego *Servicios* y, finalmente, *Hostelería y Turismo*. En la figura 1 se observa el aspecto de la página.

A continuación observamos que se dispone de los *datos de las Consejerías de Turismo* de cada Comunidad Autónoma. Pasamos pues, por ejemplo, a seleccionar el *año 2008*. Y luego *las plazas estimadas por categoría y meses*. En la figura 2 se pueden observar estos datos.

Operaciones estadísticas relacionadas		Elaboradas por el INE		Elaboradas por otros organismos		Publicaciones	
	Tablas INEbase	Contenido		Tablas INEbase	Contenido	Enlaces de interés	
Encuesta de ocupación hotelera	[icon]	[icon]	Turismo receptor	[icon]	[icon]	Este tema en el IOE	
Encuesta de ocupación en campings	[icon]	[icon]	Turismo nacional	[icon]	[icon]	Legislación estadística de la UE	
Encuesta de ocupación en apartamentos turísticos	[icon]	[icon]	Ingresos y pagos por turismo	[icon]	[icon]	[Image]	
Encuesta de ocupación en alojamientos de turismo rural	[icon]	[icon]					
Índice de precios hoteleros	[icon]	[icon]					
Índice de precios de campings	[icon]	[icon]					
Índice de precios de apartamentos turísticos	[icon]	[icon]					
Índice de precios de alojamientos de turismo rural	[icon]	[icon]					
Indicadores de Rentabilidad del Sector Hotelero	[icon]	[icon]					
Encuesta anual de servicios	[icon]	[icon]					
Estadística de productos en el sector servicios	[icon]	[icon]					
Encuesta sobre la estructura de las empresas hoteleras 1996	[icon]	[icon]					
Encuesta sobre la estructura de empresas de agencias de viajes 1997	[icon]	[icon]					

Operaciones estadísticas elaboradas por el INE		Ambito	Fuente	Frecuencia	Cód. IOE
Encuesta de ocupación hotelera		nacional	INE	mensual	2025

Figura 1. Aspecto de la web del INE con datos de hostelería y turismo

El siguiente paso es guardar los datos en formato Excel y salir de la página. Los usuarios del sistema operativo *MacOS* deben guardar los datos en formato de texto, ya que R con este sistema no puede leer datos desde Excel.

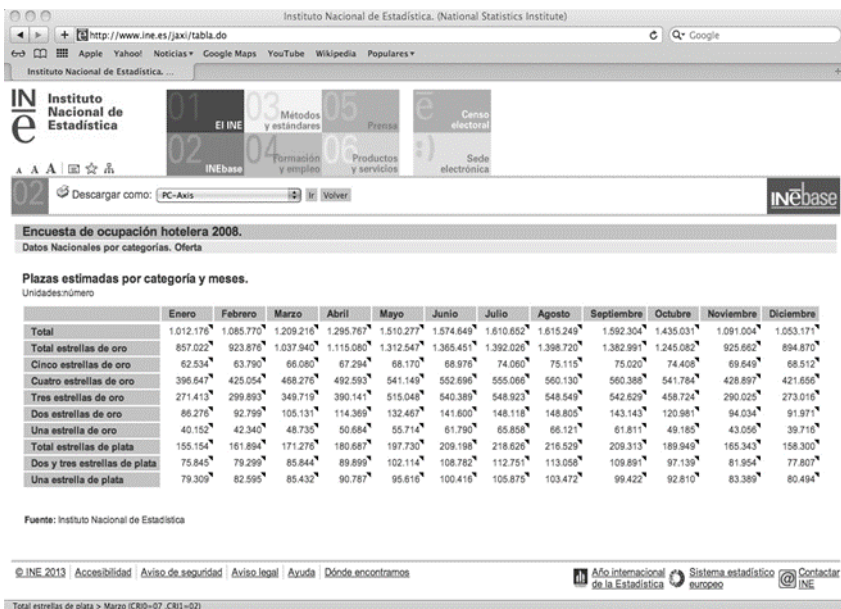


Figura 2. Datos de ocupación hotelera del año 2008 (web INE)

4.2. El paquete R Commander

Nuestra etapa siguiente consistirá en el uso de programa de software R y, en particular, trabajar con el paquete *R Commander* que utiliza ventanas. Consecuentemente cargamos este paquete *Rcmdr*. Mostramos en la figura 3 una imagen de lo que nos debe aparecer.

A partir de ahora, ya podemos trabajar con ventanas. Importamos los datos desde Excel, seleccionando primero *Datos*, para posteriormente acceder a *Importar datos desde Excel* o desde formato de texto si usamos el sistema *MacOS* (en la figura 4 se observa como deben estar cargados los datos para que R los entienda). A partir de aquí, podemos realizar cálculos estadísticos sencillos de la gran cantidad de datos reales que hemos obtenido utilizando, simplemente y de modo sucesivo, las pestañas *Estadísticos*, *Resúmenes* y *Resúmenes numéricos*. La figura 5 muestra lo que debemos ver en la pantalla de nuestro ordenador. En consecuencia, podemos calcular datos estadísticos como: medias, desviaciones, coeficientes de variación, cuartiles, etc. Y lo que es más importante, podemos dedicar nuestro tiempo a interpretar estos datos más que al puro cálculo de los mismos. Para que los estudiantes practiquen por sí mismos, mostramos un ejemplo utilizando los hoteles de 3 estrellas y luego les pedimos a los estudiantes que repitan el proceso de datos.

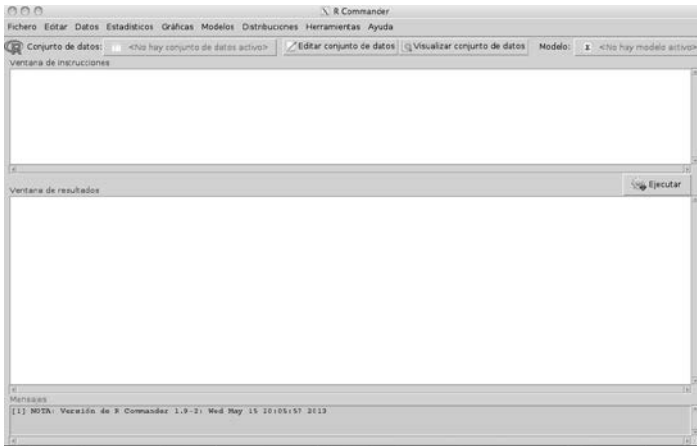


Figura 3. Aspecto del paquete R. Commander

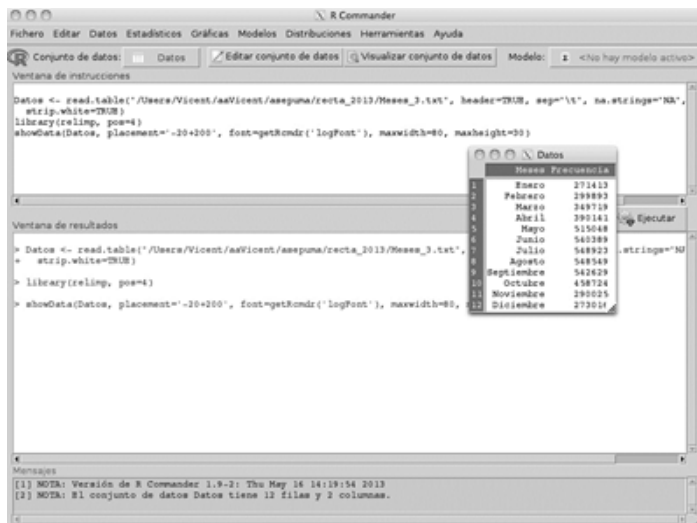
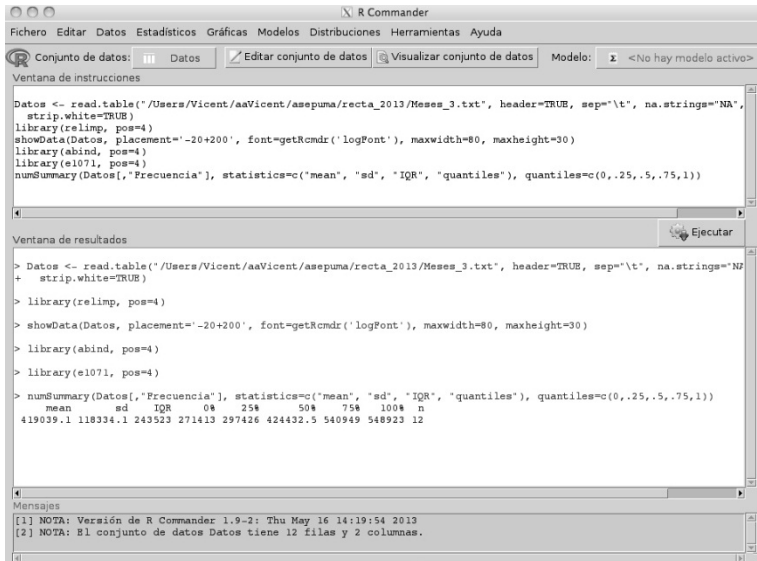


Figura 4. Aspecto de la visualización de datos en R. Commander



The screenshot shows the R Commander interface with the following code in the instruction window and the resulting summary statistics in the results window:

```

Datos <- read.table("/Users/Vicent/aaVicent/asepuma/recta_2013/Meses_3.txt", header=TRUE, sep="\t", na.strings="NA",
strip.white=TRUE)
library(relimp, pos=4)
showData(Datos, placement="-20+200", font=getRcmdr('logFont'), maxwidth=80, maxheight=30)
library(abind, pos=4)
library(e1071, pos=4)
numSummary(Datos[, "Frecuencia"], statistics=c("mean", "sd", "IQR", "quantiles"), quantiles=c(0.25,.5,.75,1))

```

The results window displays the following summary statistics:

	mean	sd	IQR	0%	25%	50%	75%	100%	n
Frecuencia	419639.1	118334.1	243523	271413	297426	424432.5	540949	548923	12

The messages window shows the following notes:

```

[1] NOTA: Versión de R Commander 1.9-2: Thu May 16 14:19:54 2013
[2] NOTA: El conjunto de datos Datos tiene 12 filas y 2 columnas.

```

Figura 5. Estadísticos que muestra R. Commander

4.3. Los gráficos

Una ventaja del programa R es la facilidad que nos ofrece a la hora de obtener gráficos para interpretar los datos de una manera más cómoda. Para ello, tenemos que darle a R los datos para que pueda trabajar con ellos. Ponemos en la ventana de instrucciones de *R Commander* las expresiones siguientes:

```
Meses<-rep(Datos$Meses,Datos$Frecuencia)
```

```
Datos2<-data.frame(Meses).
```

Con esto obtenemos un nuevo conjunto de datos: *Datos2*. Anteriormente el conjunto de datos *Datos* tenía 12 filas y 12 columnas, mientras que ahora el conjunto de datos *Datos2* tiene 5119553 filas y 1 columna. Este formato de datos es el óptimo para conseguir que R los maneje. Ahora R está preparado para obtener gráficos. Para representar un histograma, procederemos de la manera siguiente: seleccionar *Gráficas* y luego *Diagrama de barras* obteniendo así el histograma que se muestra en la figura 6. Observamos que la variable *Meses* se encuentra en orden alfabético, más tarde indicaremos como se puede obtener el orden cronológico.

Análogamente se podría obtener el diagrama de sectores que se muestra en la figura 7, actuando como sigue: seleccionar *Gráficas* y luego *Diagrama de sectores*.

Finalmente, en la figura 8 se representa la ocupación hotelera por meses ordenados cronológicamente. Para ello hemos procedido así: en primer lugar, hemos reorganizado los datos (*Modificar variables del conjunto de datos activo* y *Reordenar niveles de factor*) y definido así una nueva variable *Meses_ord*, que habrá aparecido en el desplegable de *Modificar variables del conjunto de datos activo*; y, en segundo lugar, procediendo como anteriormente (*Gráficas* y luego *Diagrama de barras*) pero con la nueva variable *Meses_ord*.

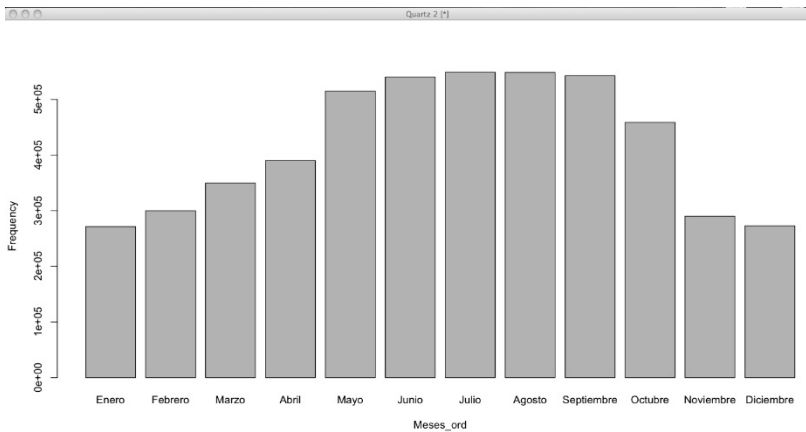


Figura 6. Histograma representando la ocupación hotelera por meses

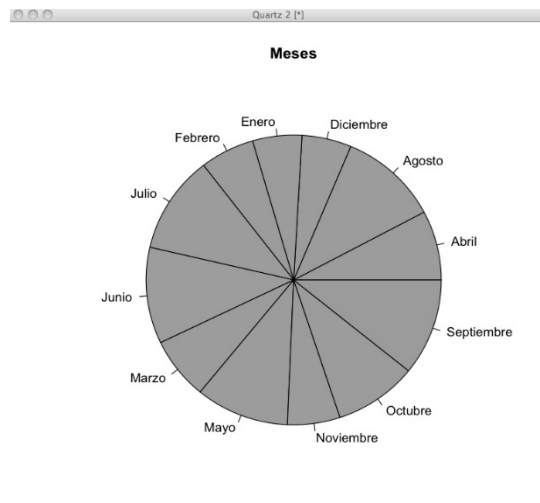


Figura 7. Diagrama de sectores mostrando la ocupación hotelera por meses

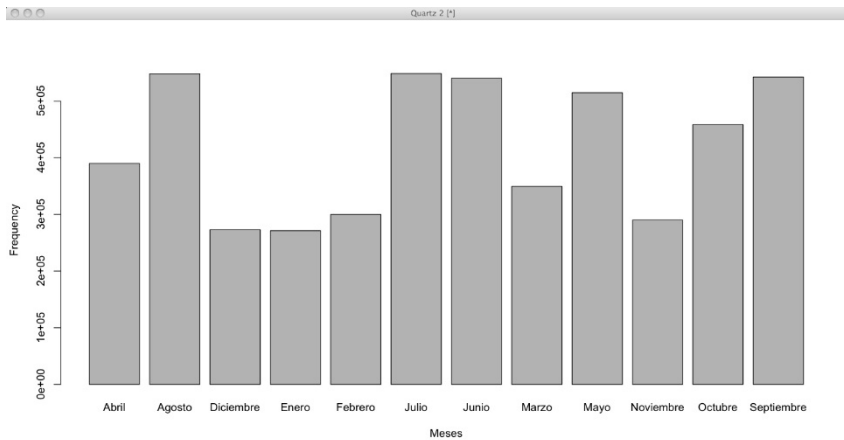


Figura 8. Histograma de ocupación hotelera por meses ordenados cronológicamente

5. Conclusiones

En este trabajo se ha presentado, a modo de ejemplo, una de las prácticas que desarrollamos en la asignatura Estadística del grado de Turismo de la Universitat Jaume I de Castellón. El conjunto de estas prácticas utiliza como software el paquete R cuyas ventajas hemos comentado en secciones anteriores. Los aspectos de mejora que hemos introducido, a partir del curso 2010/11, han sido fundamentalmente los siguientes:

- Obtención y manejo de datos, disponibles en fuentes oficiales, y por tanto fiables y útiles en su ámbito profesional.
- Desarrollo de problemas fuertemente relacionados con la titulación.
- Participación en proyectos reales a los cuales tendrán que enfrentarse en su futura vida profesional.

El resultado de este cambio metodológico ha sido muy positivo. La satisfacción de los alumnos se refleja en la valoración del profesorado, la cual ha aumentado en un 18.4%, según la encuesta de evaluación del profesorado. Además, como se deduce de los datos suministrados en la Tabla 1; el número de alumnos que se presenta a las pruebas ha aumentado en 10 puntos porcentuales y lo mismo ocurre con el rendimiento académico de los mismos, cuyo porcentaje de aprobados aumenta en 11 puntos. Por tanto, el rendimiento y el interés de los alumnos en la asignatura ha aumentado considerablemente con la metodología indicada.

Tabla 1. Rendimiento de los alumnos

Resultados de evaluación en porcentajes					
		Aprobados	Suspensos	No presentados	Totales
Curso	2009/10	50.68	15.07	34.25	100
	2010/11	61.63	15.12	23.25	100
	2011/12	61.98	15.71	22.31	100

Agradecimientos

Este trabajo ha sido parcialmente financiado por Unitat Support Educatiu, Universitat Jaume I, proyecto no.10G136-334.

Referencias Bibliográficas

1. J. Alegre, M. Cladera, C. Juaneda. *Análisi quantitativa de l'activitat turística* (Piramide, 2003).
2. C. Fernández. Instruccions d'estadística descriptiva aplicada al sector turístic (Síntesi, 2004).
3. J. M. Raya. *Estadística aplicada al turismo* (Pearson 2004).
4. A. Ronquillo. *Estadística aplicada al sector turistic*. (Centro de estudios Ramón Areces, 1997).
5. <http://www.ine.es/inebmenu/mnu\ hosteleria.htm>
6. <http://www.iet.tourspain.es/paginas/PubFrontur.aspx\ ?\option=front\&idioma=es-ES>
7. <http://unwto.org/es>
8. <http://www.comunidad-valenciana.org/opencms/opencms/turisme\es/index.jsp>
9. <http://www.educacion.gob.es/educabase>