



<b>Título del trabajo:</b> Análisis exploratorio de datos y su tratamiento computacional. Aplicación a un Proyecto estadístico
<b>Tutor/a:</b> José Fernando Vera Vera <b>Correo electrónico:</b> jfvera@ugr.es
<b>Cotutor/a:</b>
<b>Departamento responsable:</b> Estadística e Investigación Operativa
<b>Perfil y número de estudiantes al que va dirigido (máximo 2):</b> 2 alumnos <i>(Rellenar sólo en caso de que la propuesta esté realizada a través de estudiante)</i>
<b>Estudiante que propone el trabajo (Nombre, Apellidos, DNI):</b>
<b>Tipo de trabajo</b> <i>(consultar (*))</i> El trabajo está diseñado para abarcar de forma general cualquiera de las tipologías indicadas en la normativa del TFG de Estadística. Por tanto el alumno que realice este TFG complementa de forma completa su formación en el Grado de Estadística.
<b>Competencias</b> <i>(estas son la mínimas; consultar (**) si se considera añadir otras)</i> <i>Competencias generales:</i> G02, G03, G04, G05, G06, G07, G08, G09. <i>Competencias específicas:</i> E01, E02, E03, E04, E05, E08, E09, E10.
<b>Resultados de aprendizaje</b> <i>(estos son los mínimos; añadir otros si se considera)</i> <ul style="list-style-type: none"><li>• Adquirir competencias globales ligadas al desarrollo y aplicación de los conocimientos del Grado.</li><li>• Adquirir competencias ligadas a la búsqueda y organización de información y documentación relevante sobre el tema objeto de estudio.</li><li>• Aplicar el "pensamiento estadístico" y tener capacidad para enfrentarse a las distintas etapas de un estudio estadístico (desde el planteamiento del problema hasta la exposición de resultados).</li><li>• Saber presentar, de forma escrita y oral, la memoria, los resultados y las conclusiones del trabajo realizado.</li></ul>
<b>Antecedentes y resumen del tema propuesto:</b> <p>Desde el punto de vista aplicado, la realización de un Proyecto estadístico constituye sin duda alguna uno de los aspectos básicos en la formación de un Estadístico, tal y como he podido comprobar a lo largo de los más de 15 años durante los cuales que he dirigido proyectos (el equivalente al TFG) en la asignatura de Proyectos Estadísticos de la Diplomatura de Estadística. Uno de los aspectos más importantes en la aplicación real de la estadística lo constituye la fase exploratoria, fundamentalmente realizada mediante la utilización de técnicas exploratorias de Análisis Multivariante, en colaboración con la Estadística Computacional, así como la metodología asociada para el análisis estadístico computacional de los datos frente a un problema real.</p> <p>La temática general de esta propuesta de TFG consiste en la recopilación, descripción de las principales técnicas estadísticas y computacionales para un análisis exploratorio de datos y su metodología de aplicación en un área real como puede ser Ciencias Sociales, Marketing, etc. Esta propuesta de trabajo permite la realización de hasta un máximo de 2 TFG por 2 alumnos en función de tres módulos de actuación: Así, cada trabajo podría centrarse en uno de los tres objetivos del análisis de datos que se describen a continuación: Reducción de la dimensionalidad y representación multidimensional, clasificación y discriminación.</p>



Cada trabajo contemplará una parte teórica asociada a uno de los objetivos descritos anteriormente, una parte computacional que consistirá en la recopilación bibliográfica y descripción general de la metodología para la aplicación práctica del correspondiente objetivo, y su aplicación mediante el análisis de datos con SPSS y R. Finalmente, el TFG contemplará la metodología correspondiente para el análisis de datos en un problema real en relación a un Proyecto Estadístico.

Para cada módulo, esta propuesta de trabajo fin de grado contempla los siguientes objetivos, contenidos y actividades concretas a desarrollar para cada módulo:

- Descripción de la problemática abordada por el módulo correspondiente y del encargo profesional o problemática real abordada.
- Recopilación y breve descripción teórica de las técnicas estadísticas más utilizadas para abordar la problemática concreta del módulo elegido.
- Recopilación y descripción del software estadístico más utilizado para el análisis de los datos relacionados con el módulo.
- Análisis de los datos y comparación de resultados entre las distintas técnicas. Metodología del análisis exploratorio.
- Conclusiones y presentación de resultados.

Cabe la posibilidad de que el análisis real de datos esté relacionado con prácticas en una empresa.

Breve descripción de las actividades presenciales y no presenciales a realizar:

Actividades presenciales (15-30%)	Planteamiento, orientación y supervisión	90 horas
	Exposición del trabajo	1 horas
	Otras:	9
Actividades no presenciales (70-85%)	Preparación del trabajo	250 horas
	Elaboración de la memoria	50 horas
	Otras:	
Total (12 ECTS)		300 horas

Objetivos que se pretenden alcanzar:

- Estudio y descripción de una problemática real según un enfoque exploratorio.
- Recopilación y breve descripción de las técnicas estadísticas relacionadas con el problema y determinación del software adecuado para su aplicación.
- Análisis Exploratorio de los datos siguiendo la metodología general de un Proyecto Estadístico y presentación de resultados.



Bibliografía básica para la puesta en marcha del trabajo:

BIBLIOGRAFÍA BÁSICA

1. BORG, I, GROENEN, P. (2005). Modern Multidimensional Scaling. Theory and Applications. Springer.
2. EVERITT, B., HOTHON, T. (2010). A Handbook of Statistical Analysis using R. Second Edition. CRC Press.
3. GALLARDO, J. A., VERA, J.F. (2004). Técnicas aplicadas de análisis de datos multivariante. Universidad de Granada.
4. HOAGLIN, D., MOSTELLER, F. & TUKEY, J.W.(1983).-Understanding robust and Exploratory Data Analysis. New York. John Wiley & Sons.
5. MARDIA, K.V., KENT, J.T., BIBBY, J.M. Multivariate Analysis (1979). Academic Press.
6. ROBERT, C. P. & CASELLA, G. (2004).-Monte Carlo Statistical Methods. Springer-Verlag. New York.
7. TUKEY, J.W.(1977).-Exploratory Data Analysis. Reading Mass. Addison & Wesley.
8. VERA, J.F. (2004). Técnicas aplicadas de análisis exploratorio de datos. Universidad de Granada.
9. VERA, J.F., RAMOS, E, RUEDA, C, SÁCHEZ, J. (2004). Técnicas para la realización de un Proyecto Estadístico. Universidad de Granada.
10. VERA, J.F., RAYA, R, GALLARDO, J.A. (2004). Estadística Computacional Aplicada.

BIBLIOGRAFÍA COMPLEMENTARIA

1. Abranovic, W. A. (1997): Statistical Thinking and Data Analysis Methods for Managers. Ed. Addison-Wesley. Aguarón J. y otros (1993): Simulación. Colección Textos Docentes, Servicio de Publicaciones de la Universidad de Zaragoza.
2. Carrie, A. (1988): Simulation of Manufacturing Systems. Ed. John Wiley & Sons.
3. Chambers J.M. y otros (1983): Graphical Methods for Data Analysis. Ed. Chapman-Hall.
4. Olson, D.L. (2002): Introduction to Simulation and Risk Analysis. Ed. Prentice-Hall.
5. Hair, J.F.; Anderson, R.E.; Tatham, R.L.; Black, W.C. (1999): Análisis Multivariante, 5a ed. Ed. Prentice Hall
6. Kelton, W.D.; Sadowski, R.P.; Sadowski, D.A. (2002): Simulation with Arena.
7. McGraw-Hill. Law, A.M.; Kelton, W.D. (2000): Simulation Modeling and Analysis, 3rd edition. Ed. McGraw-Hill.
8. Martínez-Arias, R. (1999):El Análisis Multivariante en la Investigación Científica. Colección Cuadernos de Estadística. Ed. La Muralla-Hespérides.
9. Middleton M. R. (1995): Data Analysis Using Excel 5.0. Ed. Duxbury.
10. Naylor T.H. (1982): Experimentos de Simulación en computadoras con modelos de Sistemas Económicos. Ed. Limusa.
11. Pardo L. Y Valdés T. (1987): Simulación. Aplicaciones prácticas en la Empresa.



Ed. Díaz de Santos S.A.

12. Pérez, C. (2001): Técnicas Estadísticas con SPSS. Ed. Pearson Educación.
13. Rial, A.; Varela, J.; Rojas, A.J. (2001): Depuración y Análisis Preliminares de Datos en SPSS. Ed. Ra-Ma.
14. Ríos-Insúa D.; Ríos-Insúa S.; Martín J. (1997): Simulación. Métodos y Aplicaciones. Ed. Rama.
15. Visauta B.(1997): Análisis Estadístico con SPSS para Windows. Ed. McGraw-Hill.
16. Waters, D.(1998): Essential Quantitative Methods. Ed. Addison Wesley Longman.
17. Winston W.(2000): Financial Models using Simulation and Optimization. Ed. Palsade

### Tipo de trabajo (\*):

1. Estudio de profundización en algún tema concreto de Estadística, o como proyecto de aplicación de la misma a estudios o problemas de otros ámbitos científicos o sociales.
2. Realización completa de todas las fases de un proyecto estadístico, bien con auxilio de prácticas en empresas o con prácticas propuestas y dirigidas por el tutor.
3. Estudio de casos, teóricos o prácticos, relacionados con la Estadística.
4. Elaboración de un informe o un proyecto de naturaleza profesional.
5. Elaboración de un plan de empresa.
6. Simulación de encargos profesionales.
7. Trabajos bibliográficos sobre el estado actual de una temática relacionada con la Estadística.
8. Creación y/o empleo de herramientas informáticas para su uso en Estadística.
9. Trabajos de inicio a la investigación.
10. Trabajos cuya finalidad sea la divulgación de la Estadística en diversos contextos.
11. Trabajos sobre Historia de la Estadística.
12. Trabajos relacionados con la docencia de la Estadística.

### Competencias (\*\*)

#### Competencias generales:

**G01.** Poseer los conocimientos básicos de los distintos módulos que, partiendo de la base de la educación secundaria general, y apoyándose en libros de texto avanzados, se desarrollan en la propuesta de título de Grado en Estadística que se presenta.

**G02.** Saber aplicar los conocimientos básicos de cada módulo a su trabajo o vocación de una forma profesional y poseer las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de la Estadística y ámbitos en que esta se aplica directamente.



**G03.** Saber reunir e interpretar datos relevantes para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.

**G04.** Poder transmitir información, ideas, problemas y sus soluciones, de forma escrita u oral, a un público tanto especializado como no especializado.

**G05.** Haber desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía.

**G06.** Saber utilizar herramientas de búsqueda de recursos bibliográficos.

**G07.** Poder comunicarse en otra lengua de relevancia en el ámbito científico.

**G08.** Poseer habilidades y aptitudes que favorezcan el espíritu emprendedor en el ámbito de aplicación y desarrollo de su formación académica.

**G09.** Fomentar y garantizar el respeto a los Derechos Humanos, a los principios de accesibilidad universal, igualdad, y no discriminación; y los valores democráticos, de la cultura de la paz y de igualdad de género.

#### **Competencias específicas:**

**E01.** Conocer los fundamentos básicos del razonamiento estadístico, en el diseño de estudios, en la recogida de información, en el análisis de datos y en la extracción de conclusiones.

**E02.** Conocer, saber seleccionar y saber aplicar, técnicas de adquisición de datos para su tratamiento estadístico.

**E03.** Conocer los fundamentos teóricos y saber aplicar modelos y técnicas estadísticas en estudios y problemas reales en diversos ámbitos científicos y sociales.

**E04.** Saber seleccionar los modelos o técnicas estadísticas para su aplicación en estudios y problemas reales en diversos ámbitos científicos y sociales, así como conocer herramientas de validación de los mismos.

**E05.** Comprender la importancia de la Investigación Operativa como metodología de optimización, toma de decisiones y diseño de modelos particulares para la resolución de problemas en situaciones específicas.

**E06.** Comprender y utilizar básicamente el lenguaje matemático.

**E07.** Conocer los conceptos y herramientas matemáticas necesarias para el estudio de los aspectos teóricos y prácticos de la Probabilidad, la Estadística y la Investigación Operativa.

**E08.** Conocer y saber utilizar aplicaciones informáticas de análisis estadístico, cálculo numérico y simbólico, bases de datos, visualización gráfica y optimización, que sean útiles para la aplicación y desarrollo de las técnicas estadísticas.

**E09.** Conocer los conceptos básicos y habilidades propias de un ámbito científico o social en el que la Estadística o la Investigación operativa sean una herramienta fundamental.

**E10.** Tomar conciencia de la necesidad de asumir las normas de ética profesional y las relativas a la protección de datos y del secreto estadístico, como premisas que deben guiar la actividad profesional como profesionales de la Estadística.