

## Trabajo Grupal de Estadística: análisis de datos

---

El proyecto grupal de la asignatura de Estadística para ingeniería consiste en realizar un análisis descriptivo e inferencial en un conjunto de datos dado, utilizando un software para ello (Matlab/R/Excel/Python). Los estudiantes deberán recopilar, almacenar y visualizar la información y generar un informe y una presentación o vídeo donde se expliquen los análisis realizados y sus conclusiones. Se detallan a continuación los contenidos, instrucciones y entregas parciales del proyecto.

### Objetivos, discusión y enfoque

Los objetivos de esta actividad son:

- Promover habilidades de trabajo en equipo y coordinación estudiantil: el trabajo se llevará cabo en grupos de 3 a 4 estudiantes que permanecerán igual a lo largo del curso.
- Desarrollar el aprendizaje autónomo: saber buscar un conjunto de datos apropiado para poder poner en práctica los conocimientos técnicos adquiridos durante la asignatura.
- Analizar y extraer información relevante de un conjunto de datos: es necesario que todos los miembros del grupo discutan sobre los siguientes temas: ¿qué es importante? ¿Qué no lo es? ¿cómo puedo incluirlo de manera clara y objetiva en mi informe?
- Desarrollar el pensamiento y análisis crítico: ¿qué preguntas puedo responder con estos datos? ¿Tienen sentido los resultados? ¿Qué conclusiones se pueden extraer?

### Formación de los grupos (2 de abril)

Los grupos estarán formados por 4 personas.

La fecha límite para la formación de los grupos es el **domingo 2 de abril de 2023 a las 23:59 h.**

Pasada esa fecha se rellenarán, de manera aleatoria, los grupos que tengan aún vacantes.

La entrega del trabajo de manera individual implica una calificación igual a cero.

La manera de recuperar el trabajo grupal en extraordinaria será un examen oral de la asignatura.

*En cada una de las entregas detalladas a continuación, habrá que incluir una tabla con el porcentaje de participación de cada uno de los integrantes del grupo.*

## Instrucciones para elaboración de los documentos

### 1ª Entrega: Búsqueda de datos (9 de abril)

Será necesario buscar y seleccionar un conjunto de datos en internet que cumpla las siguientes condiciones:

- Que la muestra tenga un tamaño mínimo de 30 datos.
- Que haya, al menos, 2 variables numéricas continuas y una variable discreta y categórica.
- Que los datos no representen una serie temporal. Las series temporales son un conjunto de datos particular y tienen un análisis diferente al tratado en esta asignatura.

La propuesta de variables debe ser enviada al profesor para su validación. Se entregará un informe de una página de extensión que contenga una breve explicación de los datos, en especial las variables seleccionadas, especificando la fuente de estos. Además, se planteará una pregunta a la que se pretende dar respuesta mediante los datos expuestos (segunda entrega).

*La fecha límite para la primera entrega con los datos es el domingo 9 de abril de 2023 a las 23:59 h.*

*Calificación máxima de esta entrega: 1 punto.*

*(Cada día de retraso en esta entrega implicará un punto menos en la nota final del trabajo)*

### 2ª Entrega: El análisis descriptivo (28 de abril)

Esta parte del trabajo contempla toda la parte descriptiva de los datos. Se realizarán medidas de tendencia central y de dispersión de las variables seleccionadas, así como visualizaciones de los datos usando distintos tipos de tablas y gráficos: histogramas, diagramas de barra, de caja y bigote, de dispersión, etc. Además, se debe realizar un análisis descriptivo conjunto de las dos variables continuas seleccionadas (regresión lineal y correlación).

Se debe enviar el informe solicitado antes de la fecha establecida de entrega. La longitud del informe debe ser entre 3 y 5 páginas en formato PDF con un tamaño de letra entre 11 y 12 puntos.

#### Requerimientos mínimos

- ✓Tabla de análisis descriptivo de cada variable: media, mediana, moda, rango, desviación típica, varianza.
- ✓Gráficos de las variables: histogramas, diagramas de barra, caja y bigotes, etc.
- ✓Un análisis de regresión para las dos variables continuas, así como su gráfico de dispersión y el coeficiente de correlación.
- ✓Gráficos y tablas que pueden explicar las variables de una manera más completa (diagramas de caja según variable categórica, por ejemplo).
- ✓Discusión crítica de los resultados. En dicha discusión recaerá 1/3 del peso de la evaluación.

*La fecha límite para la 2ª entrega con el análisis descriptivo es el viernes 28 de abril de 2023 a las 23:59 h.*

*Calificación máxima de esta entrega: 4 puntos.*

*(Cada día de retraso en esta entrega implicará un punto menos en la nota final del trabajo)*

### 3ª Entrega: El análisis inferencial (10 de mayo)

En este caso, debéis utilizar vuestros conocimientos adquiridos a través del contenido de la asignatura e intentar analizar las estadísticas de población (media, varianza, etc.). Además, se debe realizar un análisis inferencial para, al menos, dos de las variables seleccionadas al principio del proyecto. La entrega debe incluir métodos de cálculo y una explicación de por qué los resultados obtenidos fueron esperados / inesperados, posibles / imposibles y relevantes / irrelevantes.

Debéis enviar el informe solicitado en la fecha establecida de entrega. La longitud del informe debe ser entre 2 y 3 páginas en formato PDF, con un tamaño de letra entre 11 y 12 puntos.

#### Requerimientos mínimos

- ✓Un intervalo de confianza para cada una de las dos variables continuas.
- ✓Un contraste de hipótesis para cada una de las dos variables continuas.
- ✓Justificación de los contrastes de hipótesis elegidos y explicación de los resultados.

*La fecha límite para la 3ª entrega con el análisis inferencial es el **miércoles 10 de mayo de 2023 a las 23:59 h.***

*Calificación máxima de esta entrega: 2 puntos.*

*(Cada día de retraso en esta entrega implicará un punto menos en la nota final del trabajo)*

### Entrega final: Integración de informes y Presentación final (17 de mayo)

Unificad las 2 entregas anteriores en un solo archivo con una página de título y los nombres de los miembros del equipo claramente representados.

Entre 6 y 9 páginas (excluyendo portada), con letra times o arial y unos 11 puntos.

Se incorporará el feedback recibido por el profesor tras la entrega del análisis descriptivo.

Se elaborará una presentación en PowerPoint (o herramienta similar) a exponer delante de los profesores de la asignatura y el resto de los compañeros y compañeras.

El tiempo de exposición no será superior a 10 minutos. Tras la exposición, los alumnos responderán las preguntas del profesor.

Cada alumno integrante del grupo podrá realizar y presentar una parte del trabajo, pero todos han de entender el trabajo en su totalidad, con el fin de ser capaces de responder a cualquiera de las preguntas del profesor.

*La fecha límite para la entrega final con el informe completo y la presentación es el **miércoles 17 de mayo de 2023 a las 23:59 h.***

*(Cada día de retraso en esta entrega implicará un punto menos en la nota final del trabajo)*

**Presentación en clase (19 de mayo)**

La presentación oral en clase tendrá lugar el **viernes 19 de mayo** en la hora habitual de clase.

El tiempo de exposición no será superior a 10 minutos. Tras la exposición, los alumnos responderán las preguntas del profesor. Cada alumno integrante del grupo podrá realizar y presentar una parte del trabajo, pero todos han de entender el trabajo en su totalidad, con el fin de ser capaces de responder a cualquiera de las preguntas del profesor.

La persona que no realice la exposición oral no podrá optar a los 3 puntos con los que se califica esta parte.

*Calificación máxima de esta presentación: 3 puntos.*