

PAUTAS PARA LOS PROYECTOS DE ANÁLISIS DE DATOS

INTRODUCCIÓN

La propuesta de desarrollar un Proyecto de Análisis de Datos (Proyecto de Análisis de Datos) encuentra su primera justificación, como señalan Anderson y Loynes (1987), en que la Estadística es inseparable de sus aplicaciones, mientras que su justificación final es la utilidad de la disciplina para resolver problemas externos a la misma.

El proyecto puede consistir en la resolución de un problema, la verificación de una hipótesis, la propuesta de una alternativa que mejore el estado actual del arte, una combinación de éstas, u otra que surja de la situación particular de cada sistema de datos en el que el proyecto se desarrolle. Así, los proyectos deben concebirse como verdaderas investigaciones asequibles al nivel de los alumnos, procurando integrar la Estadística dentro del proceso más general de la investigación, ubicándolos en la posición de tener que pensar en preguntas tales como (Graham, 1987): ¿Cuál es mi problema? ¿Necesito datos? ¿Cuáles? ¿Cómo puedo obtenerlos? ¿Qué significa este resultado en la práctica?

El inicio consiste en el planteo de un problema práctico que pueda resolverse utilizando la herramienta estadística. Batanero y Díaz (2011) aportan las siguientes ideas:

Puntos a tener en cuenta para plantear las preguntas de investigación:

- ¿Qué se quiere probar?
- ¿Qué datos se necesitan?
- ¿Qué se tiene que medir / observar / preguntar?
- ¿Cómo encontrar los datos?
- ¿Qué hacer con ellos?
- ¿Es posible hacerlo?
- ¿Se cree que aparecerán dificultades durante el proceso? ¿Cuáles?
- ¿Se logrará responder las preguntas de investigación?
- ¿Para qué servirán los resultados?

Una buena evaluación debe asegurar que el estudiante aprende y no sólo que aprueba. Algunos puntos que podrían tenerse en cuenta en la evaluación de un proyecto son los siguientes:

- **Pregunta de interés.** Si es la pregunta de investigación es relevante, está claramente enfocada y expuesta. Si es una pregunta que se puede abordar con los conocimientos del estudiante. En este apartado podría tenerse también en cuenta la definición de las variables, la descripción de cómo se pueden medir, la exposición correcta de los objetivos y, en el caso de que sea pertinente, la exposición de las hipótesis.
- **Diseño de la investigación.** Un mismo problema se puede abordar de muchas formas diferentes. Para evaluar el diseño se debe tener en cuenta si se especificó la forma en que el estudiante aborda el problema, incluyendo la descripción de población y muestra y el modelo en que los estudiantes recopilaron datos. Se tendrá en cuenta si los datos permiten resolver la cuestión investigada.
- **Análisis de datos.** Se debe valorar si el análisis de datos es adecuado al tipo de variables y a la pregunta de investigación, si se respetan los supuestos de aplicación de los diferentes métodos y si los métodos están correctamente aplicados.
- **Conclusiones.** Las conclusiones han de ser consistentes con el análisis; los datos deben apoyar las conclusiones obtenidas. Además han de relacionarse con las preguntas de investigación, objetivos e hipótesis.
- **Reflexión sobre el proceso.** Es interesante incluir una reflexión sobre las limitaciones del estudio y sugerencias de cómo mejorar el diseño o el análisis.
- **Presentación de resultados.** La presentación, incluyendo claridad y corrección de los gráficos, la organización adecuada en secciones y apartados y si la expresión oral y escrita es correcta y pertinente.

- Creatividad y originalidad. El último punto a valorar es la originalidad del trabajo y creatividad del alumno.

OBJETIVOS DEL PROYECTO DE ANÁLISIS DE DATOS

- Recopilar, organizar, depurar, almacenar, representar y analizar sistemas de datos sencillos.
- Trabajar con sistemas de datos reales.
- Comprender la importancia de la información fiable.
- Introducir una filosofía exploratoria en el análisis de datos.
- Exponer el trabajo.
- Producir un informe.

GRUPOS DE TRABAJO

Cada grupo de trabajo debe estar integrado por no más de tres estudiantes.

EXIGENCIAS MÍNIMAS DEL ANÁLISIS ESTADÍSTICO

- Un análisis de datos (exploratorio y descriptivo).
- Un análisis inferencial para una muestra y otro para la comparación de dos muestras (estimación de parámetros, pruebas de hipótesis, prueba de bondad de ajuste, prueba de independencia).

ETAPAS Y ACTIVIDADES PARA DESARROLLAR EL INFORME

Etapas I

La Etapa I del proyecto incluye las actividades 1 a 5 que se detallan a continuación. Debe presentarse para su revisión en Informe de Avance N° 1, de acuerdo al cronograma previsto para el proyecto.

- 1. Definición del título del proyecto.**
- 2. Redacción de los objetivos del trabajo.**

Establecer qué se pretende con la investigación; a tal fin se deben establecer los objetivos del proyecto. Es conveniente tener en cuenta que durante la investigación pueden surgir objetivos adicionales, modificarse los objetivos iniciales, e incluso ser sustituidos por nuevos objetivos, según el rumbo que tome la investigación.
- 3. Descripción del sistema de datos.**

Redactar una breve descripción global del sistema de datos, que tenga en cuenta los siguientes aspectos:

 - a) Procedencia:** Se refiere a si los datos son de elaboración propia, anuarios estadísticos, experimentos realizados en laboratorio, trabajo de campo, instituciones públicas o privadas, internet, prensa, simulación, etc. Mencionar la fuente, fecha, lugar de origen, responsables y toda información pertinente.
 - b) Técnicas de recolección de los datos:** Indicar si los datos fueron obtenidos mediante observación, encuesta, medición u otras técnicas. Si la información no es de elaboración propia, al momento de solicitarla y obtenerla, preguntar cómo fue recopilada para después saber responder a este ítem.
 - c) Definir la unidad de análisis.**
- 4. Definición de variables.**

Este estudio comprende: la propia definición, clasificación, asignación de la escala de medición y codificación de las mismas.
- 5. Presentación de los datos.**

El informe debe incluir los datos estudiados sin procesar, razón por la cual se elegirá la forma más conveniente para su presentación. Una forma, por ejemplo, podría ser mediante tablas. Si la base de datos es muy extensa, al momento del informe de avance, deberá presentarse al menos una parte de la misma.

Etapla II

Esta etapa incluye la actividad 6, según se detalla a continuación. Debe presentarse en Informe de Avance N° 2, de acuerdo al cronograma previsto para el proyecto:

6. **Explorar y analizar los datos** (Análisis con variables estadísticas cualitativas y cuantitativas)
- Organización y presentación de los datos: construcción de distribuciones de frecuencias absolutas y relativas, simples y acumuladas.
 - Trabajo con datos sin agrupar y datos agrupados en intervalos de clases.
 - Construcción de gráficas: de sectores, de barras, polígono de frecuencias, histogramas, ojiva, dispersograma, gráfica de caja.
 - Medidas de tendencia central: media, mediana y moda.
 - Medidas de posición no centrada: deciles, cuartiles, percentiles, fractiles.
 - Medida de posición relativa: puntuación Z o valor Z
 - Asimetría.
 - Apuntamiento.
 - Medidas de dispersión: rango, rango intercuartil, varianza, desviación estándar y coeficiente de variación.
 - Datos apartados: atípicos y anómalos.
 - Cálculo de probabilidades (Si es pertinente).
 - Tablas de contingencia: frecuencias conjuntas, marginales y condicionales
 - Estimación de parámetros.
 - Pruebas de hipótesis
 - Pruebas de independencia.
 - Bondad del ajuste.
 - Análisis de varianza.

Cada análisis debe estar acompañado de su correspondiente interpretación y, de ser oportuno, efectuar las conclusiones parciales correspondientes.

Etapla III

La etapa final incluye la actividad 7 y corresponde al Informe Final del proyecto.

7. **Redacción del Informe final.**

La redacción debe realizarse teniendo en cuenta las siguientes pautas:

Contenidos básicos del informe:

a) Portada: Debe incluir el título del proyecto, el nombre de los autores, el nombre de la cátedra, el nombre de la institución académica y la fecha de presentación del informe (mes y año).

b) Tabla de contenidos: Debe indicarse apartados, subapartados y número de página.

c) Resumen: Constituye el contenido esencial del informe, por lo que puede concebirse como una versión en miniatura del trabajo. Usualmente se incluye el planteamiento del problema, los métodos utilizados, los resultados más importantes y las principales conclusiones, todo resumido en no más de doscientas cincuenta palabras.

Debe escribirse en pretérito, porque se refiere a un trabajo ya realizado, no debe presentar ninguna información ni conclusión que no figure en el informe, no debe citar referencias bibliográficas (salvo raras excepciones). El criterio es tal que, si alguien interesado en la temática lee el resumen, tenga elementos suficientes para tomar la decisión de leer el informe completo o abandonar la lectura del mismo.

d) Introducción: Es la primera sección del informe propiamente dicho y tiene por finalidad suministrar suficientes antecedentes para que el lector pueda comprender y evaluar los resultados del estudio sin necesidad de consultar publicaciones anteriores sobre el tema. Incluye el planteamiento de los objetivos del proyecto, el contexto general de la investigación (cómo y

dónde se realizó), la descripción de la base de datos y las variables estudiadas, definición de términos, problemas y limitaciones.

e) Metodología y resultados: Procedimientos y técnicas empleadas, análisis de los resultados.

f) Conclusiones y recomendaciones: Es un apartado importante del trabajo. En cuanto a las recomendaciones, se deben redactar pensando que futuros estudiantes podrían estar interesados en estudiar el mismo tema. Si así ocurriera, podría suceder que el avance del trabajo constituyese la base de una nueva investigación. Por lo tanto, las recomendaciones podrían referirse a la necesidad de incorporar nuevas variables, profundizar algunos estudios, dificultades encontradas que pudieron evitarse, propuesta de nuevos proyectos a partir de los resultados obtenidos, etc.

g) Bibliografía

h) Anexos. Resultan útiles para describir con mayor profundidad ciertos materiales sin distraer la lectura del texto principal o cuerpo del informe, o evitar que dichos materiales rompan con el formato del informe. A modo de ejemplo, en el anexo podría incluirse la base de datos sin procesar (datos en bruto), el cuestionario utilizado en una encuesta, las fórmulas utilizadas en los cálculos, etcétera.

8. **Presentación:**

Se debe enviar a la cuenta (monica.guitart@um.edu.ar), en la fecha prevista en el cronograma, una copia digital del:

- Informe Final en Microsoft Word o cualquier formato de texto equivalente

¿QUÉ SE TENDRÁ EN CUENTA A LA HORA DE EVALUAR EL PROYECTO DE ANÁLISIS DE DATOS?

- Entrega de los informes de avance en término
- Características del sistema de datos y variables seleccionadas
- Campo de aplicación
- Conceptos y propiedades estadísticas
- Notaciones y representaciones
- Técnicas y procedimientos estadísticos utilizados
- Presentación general del informe
- Actitudes destacables:
 - Valoración de la utilidad de la herramienta estadística para analizar datos obtenidos mediante la experimentación, observación, encuesta o medida.
 - Valoración de la utilidad de la estadística para identificar relaciones de asociación entre variables.
 - Valoración de la utilidad y complejidad de las estadísticas oficiales y la importancia de colaborar en encuestas y censos para obtener datos fiables.
 - Valoración de la estadística como herramienta del científico y profesional de la ingeniería.
 - Espíritu crítico con relación a la información estadística publicada.
 - Estética y la claridad en la construcción de tablas y gráficos estadísticos.
 - Concienciación sobre la posibilidad de transmitir información sesgada en una gráfica mal construida.
 - Reflexión sobre las dificultades de la codificación y de cómo ésta introduce siempre una simplificación de la realidad.
 - Reflexión sobre las propias intuiciones incorrectas en relación a los experimentos aleatorios.
 - Reflexión sobre las tendencias y dispersiones en los datos; sobre el excesivo énfasis en los prototipos y el hecho de que éstos con frecuencia son modelos que no se dan en la realidad.
 - Espíritu crítico en el uso de paquetes estadísticos y sus opciones por defecto.