

Institución Universitaria Politécnico Grancolombiano
Facultad de Ingeniería, Diseño e Innovación
Escuela de Ciencias Básicas
Lineamientos mínimos para el Proyecto de Investigación Formativa
Cursos de Estadística

Introducción

El trabajo de investigación formativa lo define cada grupo de trabajo con base en un problema disciplinar relevante. En principio, esa motivación debe estar en coherencia con la disponibilidad de datos. Algunas fuentes recomendadas, no únicas, para conseguir bases son:

- Kaggle <https://www.kaggle.com/datasets>
- Datos abiertos Colombia <https://www.datos.gov.co/>
- ICES: <https://www.ices.gov.co/web/guest/investigadores-y-estudiantes-posgrado/acceso-a-bases-de-datos>
- Repositorios de investigación: <https://www.nature.com/sdata/policies/repositories>

Es importante que la base de datos tenga pocos problemas de inconsistencia y valores perdidos. Además, se debe garantizar que contenga variables cualitativas y cuantitativas.

Este trabajo se irá construyendo de manera progresiva en cada corte del curso. Los avances deben aplicar, como mínimo, algunos de los contenidos vistos en clase. La metodología de trabajo debe seguir los lineamientos de data storytelling (contar historias con datos). Es recomendable adquirir el siguiente texto para una aproximación conceptual: https://www.amazon.com/DataViz-Choose-Right-Chart-Bite-Size-ebook/dp/B09B47Z2K9/ref=sr_1_4?dchild=1&keywords=dataviz&qid=1630338364&sr=8-4

Sin embargo, el grupo es libre de consultar otra literatura acerca de data storytelling, y referenciarla en su entrega.

La entrega debe seguir la estructura esencial de una historia, como mínimo debe contemplar lo siguiente:

Inicio (primer corte)

- Definición del problema

- Justificación: ¿Por qué es importante ese problema? Enuncie las Consecuencias: ¿Qué pasaría si el problema se resuelve? ¿Qué se gana? ¿Qué sufrimiento se libera?
- Visualización del problema y su justificación con estadísticas descriptivas.
- Personajes involucrados: organizaciones, personas, cosas.
- Hipótesis: ¿Qué hallazgos son esperables de acuerdo con literatura previa o conocimiento de expertos?

Nudo (segundo corte)

Describa con precisión la metodología estadística por medio de la cual resolverá el problema. Explique la naturaleza de la herramienta en términos rigurosos, más no especializados. Asegúrese de que está escribiendo para un público amplio.

En esta sección considere las limitaciones de la solución, asegúrese de acotar lo que se espera obtener, con qué restricciones y lo que no se logrará resolver.

Desenlace (Tercer corte)

Presente los resultados de su trabajo. Asegúrese de relacionar los resultados con el problema y la hipótesis.

Después de los resultados defina unas conclusiones sintéticas.

Haga al menos una recomendación respecto a las acciones recomendables para los actores o interesados derivadas de los resultados y las conclusiones.

Defina lo que no se resolvió por completo. Lo que no se pudo hacer y/o podría seguirse trabajando en otras investigaciones.

Criterios de entrega

Simpleza: no significa falta de rigor. Procure utilizar un lenguaje preciso, limpio y comprensible para un público amplio.

Síntesis: Privilegie **gráficos de contraste** (ojalá máximo entre dos posibilidades: ejemplo: mayor que, menor que). **Gráficos de relacionamiento:** asociación entre variables, entidades. Prefiera las estadísticas descriptivas multivariadas. En la interpretación de los gráficos, evite narrar, repetir con palabras lo que es evidente en las gráficas y tablas. Todo texto y toda palabra en la exposición oral, debería estar enfocada en aportar una interpretación, una lectura no evidente.

Evite el análisis, es decir, una descripción descriptiva exhaustiva de cada variable. Sin embargo, esto puede incluirlo en un anexo.

Ortografía y redacción: asegúrese de que el texto no tenga errores ortográficos. Procure una redacción corta, puntual y comprensible.

Aplicabilidad: cada elemento gráfico, numérico, tabla, debe aportar a la trama (solución del problema).

Estructura de las entregas

Es deseable, aunque no estrictamente necesario, que las entregas se hagan en una herramienta profesional de visualización como Google datastudio, Power BI o Tableau. Si no se tiene experiencia en el uso de estas herramientas, la opción de Google es la más recomendable.

En caso de no utilizar una herramienta de visualización, la entrega se debe hacer en formato PDF exclusivamente.

El análisis de datos debe hacer en un archivo documentado (explicado) y ordenado de R o Excel. Otras herramientas de análisis se podrían considerar, pero deben ser conciliadas con el docente. En todo caso, este documento es anexo.