

Pauta de evaluación Proyectos: Certamen I Taller Data Science

En este certamen se evaluará el proyecto de dos maneras:

1. A través de una presentación de 10 minutos a realizar en la sesión del taller, que resuma toda la propuesta presentada en el reporte.
2. Un reporte del proyecto de 5 páginas como máximo. Debe contener las secciones indicadas en el siguiente apartado, incluyendo la sección de auto-reflexión.

Ejemplo de Secciones para el Reporte del Proyecto

La estructura del reporte que esperamos debe contener, como mínimo, las secciones que aquí se indican. Si tienes dudas sobre lo que se debe indicar en cada sección, preguntanos en Piazza.

- Título: debe ser explícito con la pregunta de investigación y el contexto de esta.
 - El peor título posible: “Proyecto de Taller de Data Science”
 - Mal título: “Entendiendo la Movilidad en Santiago”
 - Buen título: “Caracterización de Elección de Modo de Transporte”
 - The best título: “¿Qué variables consideran las personas de Santiago al elegir modo de transporte?”
- Resumen: debe describir el contenido del informe. Una o dos frases por cada una de las secciones siguientes.
- Introducción:
 - Indicar motivación (por ej., mejorar la movilidad de una ciudad), y la relevancia del tema escogido (por ej., si mejora la movilidad, mejora nuestra calidad de vida). Especular sobre posibles aplicaciones (por ej, realizar intervenciones en la ciudad para incentivar el uso de transporte público, disminuir congestión, etc.).
- Plan de investigación:
 - Indicar preguntas de investigación, posible plan de investigación (indicar pasos y los datos necesarios para concretarlo).
 - Ejemplo preguntas de investigación: ¿qué variables socio-económicas de infraestructuras están relacionadas con la elección de modo de transporte? ¿son las mismas variables igual de importantes para personas de distintas comunas?
 - Ejemplo de plan de investigación:
 - Elegiremos diez variables individuales para las personas de la encuesta: edad, sexo, comuna de origen, ingreso del hogar, nivel educacional, ...
 - A cada persona le asignaremos otras variables de acuerdo a su contexto geográfico: distancia al paradero más cercano, distancia a la estación de metro más cercana, disponibilidad de servicios de salud, de educación, etc.

- Análisis exploratorio: agregaremos para cada comuna las distribuciones de las variables, considerando también el modo de viaje.
 - En caso de encontrar diferencias visuales interesantes, utilizaremos alguna métrica o técnica que nos permita asociar o correlacionar las variables con los modos de viaje. Los resultados de esta métrica nos permitirán establecer los niveles de asociación, que evaluaremos después en una regresión.
 - En caso de no encontrar diferencias en general, proponemos evaluar las variables en comunas específicas que tengan realidades distintas. Por ej., Las Condes y Pedro Aguirre Cerda.
 - A partir de los resultados elaboraremos un set de intervenciones posibles que permitan mejorar el transporte en esas comunas.
 - Los resultados serán presentados utilizando mapas de coropletas y visualizaciones estadísticas.
- Fuentes de Datos:
 - Deben indicar cuáles fuentes de datos utilizarán para responder la pregunta de investigación. Cada fuente incluye:
 - Título, descripción (1 párrafo), fuente, link. Ejemplo: “La Encuesta Origen-Destino 2012 de Santiago contiene viajes realizados en la ciudad, con variables demográficas y económicas. Cada viaje incluye variables que describen el viaje en sí (distancia, tiempo, propósito, puntos de origen y destino, modo de viaje elegido). Fue publicado por la DTPM y está disponible en <http://udd.cl/eod2012>”.
 - Descripción de la fuente de datos. Cantidad de observaciones (ej., 20 mil hogares), cantidad de tablas o archivos incluidos (ej., una tabla de viajes, otra de hogares, otra de personas).
 - Un proyecto debiese utilizar más de una fuente de datos. Si es una sola fuente, debiese ser muy rica en variables (decenas o centenas), observaciones (cien mil o más), y contexto (ej., toda la movilidad de una ciudad). En el plan de investigación de ejemplo, se aumentó la información de la EOD con datos geoespaciales de OpenStreetMap.
- Discusiones y trabajo futuro:
 - Indicar las fortalezas, debilidades y limitaciones del proyecto. Sé crítica/o.
 - Lecciones aprendidas (¿qué aprendiste haciendo el reporte que no sabías antes de comenzarlo?)
 - Trabajo futuro (¿qué hubieses agregado al proyecto si tuvieses más tiempo disponible?)
- Referencias:
 - Indica los trabajos en que te has basado para el desarrollo de tu proyecto. También indica el origen de las fuentes de datos.
 - Asegúrate de usar referencias reales de artículos científicos, y no simplemente una dirección web con un pdf. Indica autor(es), journal/conferencia y año.

El reporte debe incluir una sección de auto-reflexión. Cada integrante del equipo debe reflexionar sobre el trabajo realizado, indicando su nombre y la respuesta a cada una de las preguntas.

1. Preguntas de conocimiento:
 - ¿Qué conceptos aprendí haciendo este trabajo? ¿Con qué otros conceptos los relaciono?
 - ¿Qué conozco del tema? ¿Qué conceptos me falta por aprender?
 - ¿Qué conclusiones puedo sacar de este trabajo?
2. Preguntas sobre el proceso:
 - ¿Qué habilidades trabajé haciendo este proyecto?
 - ¿Qué pasos debo seguir para comenzar un proyecto de ciencia de datos?
 - ¿Qué pasos requirieron más tiempo? ¿Cuánto comprendí las instrucciones?
3. Preguntas sobre las actitudes:
 - ¿Cuán sistemático he sido? ¿Cuánto interés tengo en trabajar en el proyecto?
 - ¿Le he dedicado suficiente atención y concentración a la resolución del proyecto?
 - ¿Colaboré con mis compañeros? ¿Qué rol asumo cuando trabajo en equipo?

