



IIC2115 – Programación como Herramienta para la Ingeniería (II/2025)

Laboratorio 4

Aspectos generales

- **Objetivo:** evaluar individualmente el aprendizaje sobre análisis de datos geoespaciales, en un problema práctico con datos de la ciudad de Santiago.
- **Entrega:** Parte 1 lunes 13/10 a las 17:30, Parte 2 domingo 19/10 a las 23:59 hrs. Archivos Python Notebook L4_1.ipynb y L4_2.ipynb con las soluciones de las partes 1 y 2 del laboratorio, respectivamente. Los archivos deben estar ubicados en la carpeta L4 del repositorio privado. Entregas que no cumplan con esta especificación no serán corregidas.
- **Formato de entrega:** Utilice múltiples celdas de texto y código para estructurar los archivos. Todas las celdas utilizadas deben estar ejecutadas al momento de entregar el ejercicio, de modo que las salidas generadas sean visibles. **Adjunte además el historial completo de su interacción con asistentes basados en IA, en caso de existir.** Entregas que no cumplan el formato tendrán un descuento de 0,5 puntos.
- **Entregas atrasadas:** el descuento por atraso para la Parte 1 es de 1 punto cada 10 minutos o fracción. El descuento por atraso para la Parte 2 es de 1 punto por cada hora o fracción.
- **Issues:** Las discusiones en las *issues* del Syllabus que sean relevantes para el desarrollo de la evaluación, serán destacadas. Así mismo, el uso de librerías externas que solucionen aspectos fundamentales del problema no podrán ser utilizadas. Solo se podrán utilizar las que han sido aprobadas en las *issues*, previa consulta de los estudiantes.
- **Entregas con errores de sintaxis y/o que generen excepciones en todas las ejecuciones serán calificados con nota 1.0.**

Descripción del problema

Con el fin de evaluar los contenidos de análisis de datos geoespaciales, en este laboratorio deberá realizar una serie de procesamiento y visualizaciones de datos de la ciudad de Santiago. No existe *a priori* un resultado correcto para cada ejercicio. Considere para el desarrollo los datos geoespaciales disponibles en el sitio del curso, que contienen diversos elementos de interés asociados a la ciudad de Santiago. Se recomienda explorar inicialmente el contenido de los archivos asociados y familiarizarse con el formato en que está almacenada la información. Identifique las columnas que contiene cada base de datos.

Descripción de los datos:

- Datos base: Datos ya utilizados en la materia del curso con áreas urbanas, distritos censales, comunas, líneas y estaciones de metro.
- Accidentes: Cada registro tiene un identificador único, junto con las coordenadas de latitud y longitud. Los accidentes están clasificados por año, región, y comuna, y se especifica la zona (urbana o rural) donde ocurrieron, además de las calles involucradas y la dirección completa. También se incluyen los números de fallecidos, heridos graves, menos graves y leves. Finalmente, se detalla la geometría del punto geográfico que representa la ubicación del accidente.
- Delitos: Este conjunto de datos contiene registros de distintos tipos de delitos ocurridos en la Región Metropolitana de Santiago, Chile, entre enero de 2018 y diciembre de 2020. La información contiene datos primarios del CEAD-SPD, compilados y publicados por el IMFD para HDX. Estos abarcan una amplia variedad de incidentes, que van desde infracciones menores como ruidos molestos o comercio ambulante, hasta delitos de mayor gravedad como robos en lugares habitados y robos con violencia. Los datos están organizados a nivel comunal e incluyen información mensual que permite analizar la evolución y distribución espacial de la actividad delictual en la ciudad.
- Seguridad: Contiene indicadores perceptuales de seguridad para 121.351 ubicaciones de la ciudad de Santiago. A modo resumido, estos indicadores cuantifican a partir de una variable continua, la percepción visual de seguridad del entorno urbano de los ciudadanos.

Más detalles sobre la metodología utilizada para construir esta base de datos pueden leerse [aquí](#).

Parte 1

- a) Utilizando los datos de accidentes, construya una visualización que permita identificar zonas urbanas de alta y baja intensidad de tránsito o siniestralidad. Comente qué tan bien cree que los datos de accidentes representan el flujo vehicular real y qué otras fuentes podrían complementar el análisis.
- b) Identifique zonas urbanas que podrían considerarse “críticas” por la coincidencia de factores negativos. Discuta los criterios utilizados para definir y detectar estas zonas, así como las limitaciones de su análisis.
- c) Elija un fenómeno descrito por los datos y analícelo a distintas escalas espaciales. Discuta cómo cambia la interpretación de los resultados según la escala urbana elegida y qué nivel considera más útil para la toma de decisiones.
- d) Explore si existen fronteras geográficas fuertemente marcadas y determinadas por alguna variable. Justifique el esquema de agregación espacial utilizado para la variable elegida. Proponga una forma de visualizarlas y comente qué tipo de procesos urbanos podrían explicar estas diferencias.

Parte 2

- a) Entrene un modelo predictivo que permita estimar la percepción visual de seguridad de un punto de la ciudad, utilizando como insumos información geoespacial disponible en los datos (por ejemplo, distancia a estaciones de metro, número de delitos cercanos, comuna, densidad de accidentes, etc.). Discuta las decisiones tomadas respecto al tipo de modelo, las *features* y la agregación espacial. Evalúe el desempeño obtenido y analice en qué zonas de la ciudad el modelo parece fallar más.
- b) Construya un nuevo conjunto de datos geoespaciales que identifique zonas homogéneas dentro de la ciudad, a partir de información existente. Las zonas pueden definirse según criterios estadísticos, espaciales o visuales (por ejemplo, clusters o polígonos derivados de cuadrículas). Justifique las variables utilizadas y los umbrales o criterios de segmentación aplicados. Discuta si las nuevas fronteras obtenidas coinciden o no con los límites administrativos (comunas o distritos censales, por ejemplo).
- c) Evalúe si existe una relación entre la cercanía a ciertas infraestructuras urbanas (por ejemplo, estaciones de metro o ejes viales principales) y la frecuencia de incidentes de seguridad o accidentes. Proponga una forma cuantitativa de medir esta relación y discuta sus resultados.
- d) Genere un nuevo conjunto de puntos geoespaciales que representen ubicaciones hipotéticas para instalar cámaras de seguridad o dispositivos de control vial. Fundamente su elección utilizando patrones

observados en los datos. Discuta los criterios técnicos o sociales que podrían influir en la decisión final.

- e) Simule un escenario hipotético de intervención urbana, por ejemplo, la apertura de una nueva estación de metro o la implementación de medidas de seguridad en una zona específica. Analice cómo este cambio podría alterar las condiciones observadas en otras variables y proponga una forma de cuantificar o visualizar el impacto esperado.

Consideraciones sobre la metodología de trabajo

En este laboratorio no solo interesa el resultado final, sino también el proceso seguido para construir la solución. Por ello, se espera que el desarrollo de su entrega se organice a partir de una **división progresiva de cada problema en subproblemas más simples**, siguiendo un enfoque de *divide y vencerás*. La subdivisión debe reflejarse tanto en el diseño de las soluciones, como en la organización del notebook: cada celda debe ser temáticamente coherente, evitando reunir todo el código en una única celda. Todos los supuestos y simplificaciones que se utilicen deben quedar claramente explicitados en el notebook. Además, siempre que sea posible, se recomienda complementar con diagramas.

Otro componente clave es el uso de **prompts estructurados** para interactuar con herramientas de IA. Cada vez que se utilice un asistente para generar código, se debe registrar en una celda de texto el prompt utilizado, incluyendo al menos:

- Propósito y requisitos,
- Entradas y salidas,
- Restricciones y supuestos,
- Hitos intermedios,
- Convenciones.

Este nivel de detalle permite mantener claridad sobre lo que se espera del código y facilita evaluar cómo la IA fue utilizada en el proceso de desarrollo.

El único caso donde está permitido usar código generado por IA sin necesidad de dar estructura al prompt, es cuando este código es luego progresivamente modificado y/o mejorado por el estudiante mismo.

IMPORTANTE: los prompts estructurados no debe haber sido escritos por un asistente basado en IA, deben ser completamente escritos por el estudiante que entrega el laboratorio.

Finalmente, todo el código debe estar acompañado de **tests explícitos**. Un test, en este contexto, es simplemente un fragmento de código que demuestra que otro fragmento funciona como se espera, considerando casos no triviales. No se exige un test unitario por cada método, pero sí que se diseñen pruebas convincentes que permitan validar comportamientos relevantes. Para cada caso probado, debe quedar visible el resultado esperado y el resultado obtenido, de modo que cualquier lector pueda verificar si el comportamiento es correcto.

Es importante enfatizar que **adjuntar el historial de interacción con un asistente de IA no constituye por sí mismo una justificación válida ni suficiente**. Si bien dicho historial debe incluirse como material adicional, no sustituye en ningún caso los requisitos de formato, estructuración y testeo descritos anteriormente, los cuales son obligatorios y forman parte de la evaluación.

Corrección

Para la corrección se revisarán los procedimientos desarrollados para responder las diferentes preguntas y cómo estos cumplen con la materia del capítulo y la metodología de trabajo antes descrita. Dado lo abierto de las preguntas, se espera que todas las respuestas incluyan análisis y visualizaciones que permitan justificar las decisiones tomadas.

Política de Integridad Académica

Los/as estudiantes de la Escuela de Ingeniería de la Pontificia Universidad Católica de Chile deben mantener un comportamiento acorde a la Declaración de Principios de la Universidad. En particular, se espera que mantengan altos estándares de honestidad académica. Cualquier acto deshonesto o fraude académico está prohibido; los/as estudiantes que incurran en este tipo de acciones se exponen a un Procedimiento Sumario. Es responsabilidad de cada estudiante conocer y respetar el documento sobre Integridad Académica publicado por la Dirección de Docencia de la Escuela de Ingeniería.

Específicamente, para los cursos del Departamento de Ciencia de la Computación, rige obligatoriamente la siguiente política de integridad académica. Todo trabajo presentado por un/a estudiante para los efectos de la evaluación de un curso debe ser hecho **individualmente** por el/la estudiante, **sin apoyo en material de terceros**. Por “trabajo” se entiende en general las interrogaciones escritas, las tareas de programación u otras, los trabajos de laboratorio, los proyectos, el examen, entre otros.

En particular, si un/a estudiante copia un trabajo, o si a un/a estudiante se le prueba que compró o intentó

comprar un trabajo, **obtendrá nota final 1.1 en el curso** y se solicitará a la Dirección de Pregrado de la Escuela de Ingeniería que no le permita retirar el curso de la carga académica semestral.

Por “copia” se entiende incluir en el trabajo presentado como propio, partes hechas por otra persona. En caso que corresponda a “copia” a otros estudiantes, la sanción anterior se aplicará a todos los involucrados. En todos los casos, se informará a la Dirección de Pregrado de la Escuela de Ingeniería para que tome sanciones adicionales si lo estima conveniente.

También se entiende por copia extraer contenido sin modificarlo sustancialmente desde fuentes digitales como Wikipedia o mediante el uso de asistentes inteligentes como ChatGPT, Gemini o Copilot. Se entiende que una modificación sustancial involucra el análisis crítico de la información extraída y en consecuencia todas las modificaciones y mejoras que de este análisis se desprendan. Cualquiera sea el caso, el uso de fuentes bibliográficas, digitales o asistentes debe declararse de forma explícita, y debe indicarse cómo el/la estudiante mejoró la información extraída para cumplir con los objetivos de la actividad evaluativa.

Obviamente, está permitido usar material disponible públicamente, por ejemplo, libros o contenidos tomados de Internet, **siempre y cuando se incluya la referencia correspondiente**.

Lo anterior se entiende como complemento al Reglamento del Estudiante de la Pontificia Universidad Católica de Chile (<https://registrosacademicos.uc.cl/reglamentos/estudiantiles/>). Por ello, es posible pedir a la Universidad la aplicación de sanciones adicionales especificadas en dicho reglamento.

Compromiso del Código de Honor

Este curso suscribe el Código de Honor establecido por la Universidad, el que es vinculante. Todo trabajo evaluado en este curso debe ser propio. En caso que exista colaboración permitida con otros/as estudiantes, el trabajo deberá referenciar y atribuir correctamente dicha contribución a quien corresponda. Como estudiante es un deber conocer el Código de Honor (<https://www.uc.cl/codigo-de-honor/>).