



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

Laboratorio 2

Aspectos generales

- **Objetivo:** evaluar individualmente el aprendizaje sobre análisis de datos tabulares con Pandas, en un problema práctico.
- **Entrega:** Parte 1 lunes 01/09 a las 17:30, Parte 2 domingo 07/09 a las 23:59 hrs. Archivos Python Notebook L2_1.ipynb y L2_2.ipynb con las soluciones de las partes 1 y 2 del laboratorio, respectivamente. Los archivos deben estar ubicados en la carpeta L2 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

En este laboratorio utilizará un conjunto de datos que recopila información sobre hoteles internacionales, usuarios que realizaron reservas y reseñas que describen su experiencia durante la estadía. Estos datos están divididos en tres archivos:

- **hotels.csv**: cada fila describe un hotel, incluyendo atributos como ubicación geográfica, categoría, y puntajes de referencia o “base” en distintas dimensiones (ej. limpieza, comodidad, ubicación, etc.).
- **users.csv**: cada fila corresponde a un usuario de la plataforma, con información demográfica y de contexto relevante para caracterizar su comportamiento de reseña.
- **reviews.csv**: cada fila representa una reseña realizada por un usuario sobre un hotel en una fecha determinada.

En base a los campos recién descritos y utilizando las librerías presentadas en clases, deberá responder una serie de preguntas relacionadas con el análisis de datos tabulares en Python.

Parte 1

- a) Analice y describa la relación entre el puntaje global entregado por los usuarios y los sub-puntajes de las distintas dimensiones de la reseña, es decir, cómo se calcula. Evalúe en qué medida ambos son consistentes y comente posibles discrepancias.
- b) Compare el desempeño observado de los hoteles con los puntajes de referencia que estos poseen en cada dimensión. Genere un índice que capture la diferencia entre ambas mediciones y determine qué hoteles aparecen sistemáticamente mejor o peor evaluados que su base.
- c) Evalúe si existen diferencias notorias en la forma en que reseñan los usuarios según la fecha en que se unieron a la plataforma, comentando patrones que considere relevantes.

Parte 2

- a) Analice si los usuarios tienden a evaluar de manera distinta cuando el país de residencia coincide con el del hotel. Comente posibles sesgos y patrones que encuentre.
- b) Estudie la evolución temporal de los puntajes de reseñas a lo largo de los años y meses disponibles. Identifique periodos atípicos y discuta en qué medida afectan la interpretación general de la calidad hotelera.

- c) Construya un indicador que caracterice a los usuarios en términos de la severidad o benevolencia de sus reseñas. Analice cómo este indicador influye en la evaluación general de los hoteles.
- d) Considere la ubicación geográfica de los hoteles y los puntajes asignados a la dimensión de ubicación. Evalúe si existe coherencia entre la localización real y la percepción reportada en las reseñas, identificando casos que se aparten de lo esperado.
- e) Analice el puntaje de valor por dinero otorgado a los hoteles, considerando simultáneamente otras dimensiones de calidad y atributos del hotel. Determine qué hoteles aparecen como “sobrevalorados” o “subvalorados” en este sentido.
- f) Estudie cómo varía la evaluación de los hoteles dependiendo del tipo de viajero que los visita. Proponga un escenario alternativo de composición de viajeros y estime cómo cambiaría el ranking de hoteles bajo esa hipótesis.
- g) Investigue la presencia de patrones sospechosos en las reseñas. Discuta en qué medida la existencia de estos patrones altera la reputación de los hoteles.

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/>).