



IIC2115 – Programación como Herramienta para la Ingeniería (I/2020)

Laboratorio 6 - Manejo de GIS

Objetivos

- Aplicar los contenidos de manejo de GIS para procesar información geográfica, crear visualizaciones y tomar decisiones en base a esto.

Entrega

- **Lenguaje a utilizar:** Python 3.6
- **Lugar:** repositorio privado en GitHub. Recuerde incluir todo en una carpeta de nombre **L06**.
- **Entrega:** Domingo 5 de julio a las 23:59 hrs.
- **Formato de entrega:**
 - Archivo python notebook (**L06.ipynb**) con su código. Utilice múltiples celdas de texto y código para facilitar la revisión de su laboratorio.
 - Archivo python (**L06.py**) con el mismo código entregado en el notebook.
 - Todos los archivos deben estar ubicados en la carpeta **L06**. No se debe subir ningún otro archivo a la carpeta. Los archivos **.ipynb** y **.py** deben contener la misma solución. No incluya las bases de datos en esta entrega.
- **Descuentos:** El descuento por atraso se realizará de acuerdo a lo definido en el programa del curso. Además de esto, tareas que no cumplan el formato de entrega tendrán un descuento de 0.5 pts.
- **Laboratorios con errores de sintaxis y/o que generen excepciones serán calificadas con nota 1.0.**

- Si su laboratorio es entregado fuera de plazo, tiene hasta el **Lunes 6 de Julio a las 23:59 hrs** para responder el formulario de **entregas fuera de plazo** disponible en el Syllabus.
- Las discusiones en las *issues* del Syllabus en GitHub son parte de este enunciado.
- El uso de librerías externas que sean estructurales en la solución de los problemas no podrán ser utilizadas. Solo se podrán utilizar las que han sido aprobadas en las *issues* de GitHub.

Introducción: En búsqueda de una herramienta urbana parte 3

Debido al éxito de los concursos para el desarrollo de la herramienta urbana definitiva, el Gobierno ha abierto una convocatoria final que busca potenciar aún más el análisis de los datos disponibles e incorporar información geográfica. En particular, se busca el uso de herramientas de GIS (Geographic Information System) para entener y visualizar los datos.

Usted ha visto el nuevo anuncio y ha decidido participar. Para ello, utilizando el conocimiento y los datos correspondientes a las herramientas anteriores, para cumplir una nueva serie de misiones, esta vez relacionadas con el uso del módulo *geopandas* de Python.

Nuevos archivos geográficos

Dado que incluir la información geográfica para todo Chile es una idea ambiciosa, el Gobierno ha disponibilizado información geográfica solo para el sector de Temuco, la que se lista a continuación:

- Shape de mapa base
- Shape de manzanas
- Shape de establecimientos educacionales
- Shape de servicios de salud
- Shape de áreas verdes
- Shape de precipitaciones
- Shape de vialidad

En este mismo link (<https://bit.ly/2T08n0q>) encontrará los datos que hemos estado utilizando y los nuevos datos SHP.

Preparando la información previa

Para poder realizar este laboratorio es importante que prepare la información en DataFrames de la librería pandas. Para ello debe seguir los pasos de la **M1** del **L05**. Es decir, debe traspasar los datos utilizados en el **L04** a DataFrames de *pandas*. Para cumplir con esto, deberá realizar siguientes dos pasos:

- (a) Importe los datos generados en la **M1** del **L04**, directamente desde las tablas en SQLite a DataFrames, generando uno por tabla. Una para Predios y otra para líneas de construcción.
- (b) Importe la información de los archivos **.csv**, incorporándola directamente en los DataFrames.

A contar de este momento, usted debe tener dos DataFrame. Uno de Predios y otro de Líneas de construcción, además, en ellos, ya debió incluirse la información de los diccionarios.

Misiones para la herramienta

Para completar requerimiento solicitado por el Gobierno, usted deberá crear una herramienta que sea capaz de realizar las tareas descritas en cada una de las misiones siguientes:

- M1. Revisión de información geográfica:** Su primera misión consiste en importar todos los recursos geográficos disponibles como GeoDataFrames. Una vez importados, Revise qué datos almacena cada shape. Además, asegúrese de dejar todos los shapes bajo el mismo sistema de referencia (crs). Genere una o dos visualizaciones que le permitan verificar gráficamente que ha dejado todo correctamente referenciado. NOTA: Hay algunos shapes más grandes que otros, sea cuidadoso para visualizarlos. (0.5 puntos)
- M2. Filtrado e incorporación de datos del SII:** Es hora de incorporar los datos que ha preparado en la introducción de este laboratorio.
 - (a) Filtre los datos de Predios y Líneas de construcción para conservar solo los datos correspondientes a las comunas de Temuco y Padre las casa. (0.2 puntos)
 - (b) A partir de la información filtrada, cree un nuevo DataFrame donde cada fila corresponde a una manzana. Para cada manzana, cuente el número de líneas de construcción con código de destino 'H', 'O' y 'C'. Cree una columna para cada registro. (0.3 puntos)
 - (c) De forma análoga a (b), agregue 3 columnas nuevas a su DataFrame de manzanas con el total de m^2 de 'H', 'O' y 'C' de cada manzana. (0.2 puntos)

(d) Mediante el código de la manzana incorpore esta nueva información al shape de manzanas disponible.
(0.3 puntos)

- M3. **Ajuste de datos:** El shape de mapa base será la base de sus mapas. Filtre todos los archivos geográficos (shapes) para que solo contengan geometrías que pertenecen al shape del mapa base. Averigüe el uso de las funciones *overlay* y *sjoin*. Para completar esta misión, genere una o dos visualizaciones que le permitan verificar gráficamente que ha dejado todo con el tamaño adecuado. (1 punto)
- M4. **Evaluando el alcance de la educación:** Investigue la herramienta *buffer*. Construya un *buffer* de 300m al rededor de cada establecimiento educacional. Para cada nuevo *buffer* obtenga el número de hogares que quedan dentro de él. Si el *buffer*, toca una manzana, entonces contabilice el número de hogares de esa manzana. Realice una visualización de lo obtenido. Para ello considere el nuevo shape de manzanas (de la M3) como base de su mapa. Sobre él muestre *buffer* de colegios coloreados de acuerdo al número de hogares contabilizados. (1 punto)
- M5. **Aproximación de accesibilidad a servicios:** Construya un nuevo GeoDataFrame con los centroídes de las manzanas. Basado en estos nuevos puntos, construya 5 *buffers* de 100m, 200m, 300m, 400m y 500m respectivamente. Para cada centroide y en cada buffer, contabilice el número de establecimientos educacionales, servicios de salud, metros totales de vialidad, número de oficinas y m^2 de comercio (Para estas dos últimas, considere que si el *buffer* toca una manzana, entonces deben considerarse los datos de esta manzana). En cada centroíde, calcule la accesibilidad para cada uno de estos servicios sabiendo que la accesibilidad al servicio x , queda determinada como $Acc_x = \sum_{i=1}^5 \frac{1}{100 \cdot i} C_i^x$. Donde i es el *buffer* $i \in \{1, 2, 3, 4, 5\}$ y C_i^x es el conteo del servicio x en el *buffer* i . Finalmente, haga una visualización de las accesibilidades por servicio, mostrando solo el shape de manzanas con una coloración adecuada. (1.5 puntos)
- M6. **Cercanía de áreas verdes:** Para cada centroíde de manzana, determine la distancia mínima al área verde más cercana. Grafique esta nueva información usando el shape de manzanas y una coloración adecuada. (1 punto)

Corrección

Es importante que deje ejecutado todo su trabajo antes de subirlo, de lo contrario se le aplicará un descuento al puntaje total. Para la corrección de este laboratorio, se revisarán los procedimientos desarrollados para responder las diferentes misiones propuestas y la estructura de como utiliza Python en ellos.

Dado lo abierto de las misiones, se espera que las respuestas incluyan análisis exploratorios y visualizaciones que permitan justificar las decisiones tomadas.

Política de Integridad Académica

“Como miembro de la comunidad de la Pontificia Universidad Católica de Chile me comprometo a respetar los principios y normativas que la rigen. Asimismo, prometo actuar con rectitud y honestidad en las relaciones con los demás integrantes de la comunidad y en la realización de todo trabajo, particularmente en aquellas actividades vinculadas a la docencia, el aprendizaje y la creación, difusión y transferencia del conocimiento. Además, velaré por la integridad de las personas y cuidaré los bienes de la Universidad.”

En particular, se espera que mantengan altos estándares de honestidad académica. Cualquier acto deshonesto o fraude académico está prohibido; los alumnos que incurran en este tipo de acciones se exponen a un procedimiento sumario. Ejemplos de actos deshonestos son la copia, el uso de material o equipos no permitidos en las evaluaciones, el plagio, o la falsificación de identidad, entre otros. Específicamente, para los cursos del Departamento de Ciencia de la Computación, rige obligatoriamente la siguiente política de integridad académica en relación a copia y plagio: Todo trabajo presentado por un alumno (grupo) para los efectos de la evaluación de un curso debe ser hecho individualmente por el alumno (grupo), sin apoyo en material de terceros. Si un alumno (grupo) copia un trabajo, se le calificará con nota 1.0 en dicha evaluación y dependiendo de la gravedad de sus acciones podrá tener un 1.0 en todo ese ítem de evaluaciones o un 1.1 en el curso. Además, los antecedentes serán enviados a la Dirección de Docencia de la Escuela de Ingeniería para evaluar posteriores sanciones en conjunto con la Universidad, las que pueden incluir un procedimiento sumario. Por “copia” o “plagio” se entiende incluir en el trabajo presentado como propio, partes desarrolladas por otra persona. Está permitido usar material disponible públicamente, por ejemplo, libros o contenidos tomados de Internet, siempre y cuando se incluya la cita correspondiente.