

# Proyecto 1

---

## Extracción de datos usando Python (RE y Pandas)

### Requisitos previos:

- Instalar Python 2 en el equipo a utilizar (Se recomienda el uso de Jupyter Notebook).

### Condiciones generales:

- 1) La salida debe estar en formato .csv
- 2) Debe contener el mismo número de líneas para cada campo y los valores deben coincidir con los del archivo original.
- 3) Cada campo debe extraerse con una función y las expresiones regulares deben estar asignadas a una variable.
- 4) En deadline para este proyecto es el lunes a las 6:00 pm.

### Instrucciones:

Se le proporciona un archivo llamado city.txt, de este debe extraer la siguiente información:

**(Como aparece -> Como se debe llamar en el csv)**

- Permit Issue Date -> permit\_date (*formato: yyyy-mm-dd*)
- Job Address -> house\_street\_address
- APN -> house\_parcel\_number
- Lot # -> lot
- Permit Type -> permit\_type
- Permit Description -> permit\_description
- Valuation -> cost (*formato: NO puede incluir \$, parte decimal, ni "."*)
- Sq. Ft. -> square\_footage
- Contractor -> company\_name
- Property Owner -> house\_owner\_name

**Nota:** Ninguno de los campos puede contener "," o ";".

- Utilice el código de ayuda suministrado para guiarse.
- Comente sus dudas públicamente en el slack.
- Evite compartir su código, el plagio no es permitido.
- Cumpla con las fecha de entrega.
- Se anexa un archivo llamado city.pdf para verificar la información que se extrae.