

# SmartData Construcción

## Documentación

Proceso de Datos

SmartData Construcción  
Documento producido por PostData  
Agosto del 2019  
CLHM

# Introducción

Este documento tiene por objetivo, realizar una revisión en el tratamiento de los datos a través del entorno de tablas macro excel, jupyter notebook y su posterior alojamiento tanto en local como en la nube, con el fin de alimentar la API de SmartData Construcción.

## Sobre los datos

Los datos corresponden a 93 datasets divididos en 8 tablas generales:

- Macroindicadores
- Sustentabilidad
- Inmobiliarios y Habitacionales
- Casen
- Innovacion
- BIM
- Materiales
- Leed\_CES

Cada una de estas tablas han sido separadas en hojas para la construcción de los datasets finales. Estas tablas de entrada son del tipo 'Excel'.

Para cada una de las tablas antes mencionadas se ha creado una sección dentro de la plataforma una serie de *scripts* en el lenguaje de programación Python, utilizando la herramienta Jupyter Notebook, que facilita una interfaz gráfica amigable de la terminal del computador.

- S1\_Destacados.ipynb
- S2\_Macroindicadores.ipynb
- S3\_Sustentabilidad.ipynb
- S4\_Inmobiliario\_habitacional.ipynb
- S5\_BIM\_Innovacion.ipynb
- S6\_Seguridad.ipynb
- S7\_CChC.ipynb
- LEED\_CES\_GeoCoding.ipynb

Lo que realiza cada uno de estos *notebooks* es tomar la data de ingreso (archivos excel), procesarlos mediante una serie de funciones y generar el formato final con toda la información relacionada para indicador. El formato de salida de los datos, siempre es del tipo Json<sup>1</sup>.

---

<sup>1</sup> Ver <https://en.wikipedia.org/wiki/JSON>

## Section 1: Destacados

```
In [45]: 1 import seaborn as sns
2 import matplotlib.pyplot as plt
3 import json
4 import pandas as pd
5 from itertools import zip_longest as zip
6 from itertools import count
7 import numpy as np
8 import codecs
9
10 ## A. Writing Json file with format
11 def write_json_UTF8(jsonDir, name_file):
12     with open(name_file, 'w') as outfile:
13         json.dump(jsonDir, outfile, ensure_ascii=False)
14     with codecs.open(name_file, 'r', encoding = "iso-8859-1") as file:
15         lines = file.read()
16     with codecs.open(name_file, 'w', encoding = 'utf8') as file:
17         file.write(lines)
18     print("CREATED: ", name_file)
19
```

Imagen: notebook de la sección ‘Destacados’. Aquí se aprecian las librerías y la función que realiza el encoding a UTF-8.

```
MapGraphJs NavVisitjs Destacado_03_IMACON.json x
src > frontend > public > data > Destacado_03_IMACON.json >
1 [{"titulo": "IMACON Variaci3n Mensual", "Fuente": "CCHC", "LinkDatosOriginales": "https://smartdata-demo.sfo2.digitaloceanspaces.com/data/Destacados.xlsx", "comment":
"3ndice Mensual de la Construcci3n. Variaci3n mensual (%). CChC", "linkPlot": "", "glosario": ""}], [{"anio": 1989, "regiones": {"PAIS": {"meses": {"Abr": {"Var Mes": 0,
"tooltip": {"Var Mes": 0, "Fecha": ": "1989/ Abr"}}, "May": {"Var Mes": 1.3, "tooltip": {"Var Mes": 1.3, "Fecha": ": "1989/ May"}}, "Jun": {"Var Mes": 1.7, "tooltip":
{"Var Mes": 1.7, "Fecha": ": "1989/ Jun"}}, "Jul": {"Var Mes": 1.9, "tooltip": {"Var Mes": 1.9, "Fecha": ": "1989/ Jul"}}, "Ago": {"Var Mes": 1.9, "tooltip": {"Var Mes":
1.9, "Fecha": ": "1989/ Ago"}}, "Sep": {"Var Mes": 1.9, "tooltip": {"Var Mes": 1.9, "Fecha": ": "1989/ Sep"}}, "Oct": {"Var Mes": 1.8, "tooltip": {"Var Mes": 1.8, "Fecha":
": "1989/ Oct"}}, "Nov": {"Var Mes": 1.8, "tooltip": {"Var Mes": 1.8, "Fecha": ": "1989/ Nov"}}, "Dic": {"Var Mes": 1.8, "tooltip": {"Var Mes": 1.8, "Fecha": ": "1989/
Dic"}}}}], {"anio": 1990, "regiones": {"PAIS": {"meses": {"Ene": {"Var Mes": 1.8, "tooltip": {"Var Mes": 1.8, "Fecha": ": "1990/ Ene"}}, "Feb": {"Var Mes": 1.4,
"tooltip": {"Var Mes": 1.4, "Fecha": ": "1990/ Feb"}}, "Mar": {"Var Mes": 0.8, "tooltip": {"Var Mes": 0.8, "Fecha": ": "1990/ Mar"}}, "Abr": {"Var Mes": 0, "tooltip":
{"Var Mes": 0, "Fecha": ": "1990/ Abr"}}, "May": {"Var Mes": 0.6, "tooltip": {"Var Mes": 0.6, "Fecha": ": "1990/ May"}}, "Jun": {"Var Mes": 1.2, "tooltip": {"Var Mes": 1.2,
"Fecha": ": "1990/ Jun"}}, "Jul": {"Var Mes": 1.5, "tooltip": {"Var Mes": 1.5, "Fecha": ": "1990/ Jul"}}, "Ago": {"Var Mes": 1.7, "tooltip": {"Var Mes": 1.7, "Fecha": ":
"1990/ Ago"}}, "Sep": {"Var Mes": 1.7, "tooltip": {"Var Mes": 1.7, "Fecha": ": "1990/ Sep"}}, "Oct": {"Var Mes": 1.5, "tooltip": {"Var Mes": 1.5, "Fecha": ": "1990/ Oct"}},
"Nov": {"Var Mes": 1.1, "tooltip": {"Var Mes": 1.1, "Fecha": ": "1990/ Nov"}}, "Dic": {"Var Mes": 0.6, "tooltip": {"Var Mes": 0.6, "Fecha": ": "1990/ Dic"}}}}], {"anio":
1991, "regiones": {"PAIS": {"meses": {"Ene": {"Var Mes": 0.2, "tooltip": {"Var Mes": 0.2, "Fecha": ": "1991/ Ene"}}, "Feb": {"Var Mes": 0.1, "tooltip": {"Var Mes": 0.1,
"Fecha": ": "1991/ Feb"}}, "Mar": {"Var Mes": 0.5, "tooltip": {"Var Mes": 0.5, "Fecha": ": "1991/ Mar"}}, "Abr": {"Var Mes": 0.8, "tooltip": {"Var Mes": 0.8, "Fecha": ":
"1991/ Abr"}}, "May": {"Var Mes": 0.9, "tooltip": {"Var Mes": 0.9, "Fecha": ": "1991/ May"}}, "Jun": {"Var Mes": 1.5, "tooltip": {"Var Mes": 1.5, "Fecha": ": "1991/ Jun"}},
"Jul": {"Var Mes": 0.4, "tooltip": {"Var Mes": 0.4, "Fecha": ": "1991/ Jul"}}, "Ago": {"Var Mes": 3, "tooltip": {"Var Mes": 3, "Fecha": ": "1991/ Ago"}}, "Sep": {"Var
Mes": 0.2, "tooltip": {"Var Mes": 0.2, "Fecha": ": "1991/ Sep"}}, "Oct": {"Var Mes": 2, "tooltip": {"Var Mes": 2, "Fecha": ": "1991/ Oct"}}, "Nov": {"Var Mes": 1.5,
"tooltip": {"Var Mes": 1.5, "Fecha": ": "1991/ Nov"}}, "Dic": {"Var Mes": 1.6, "tooltip": {"Var Mes": 1.6, "Fecha": ": "1991/ Dic"}}}}], {"anio": 1992, "regiones":
{"PAIS": {"meses": {"Ene": {"Var Mes": 1.8, "tooltip": {"Var Mes": 1.8, "Fecha": ": "1992/ Ene"}}, "Feb": {"Var Mes": 1.5, "tooltip": {"Var Mes": 1.5, "Fecha": ": "1992/
Feb"}}, "Mar": {"Var Mes": 1.3, "tooltip": {"Var Mes": 1.3, "Fecha": ": "1992/ Mar"}}, "Abr": {"Var Mes": 1.9, "tooltip": {"Var Mes": 1.9, "Fecha": ": "1992/ Abr"}},
"May": {"Var Mes": 1.5, "tooltip": {"Var Mes": 1.5, "Fecha": ": "1992/ May"}}, "Jun": {"Var Mes": 1, "tooltip": {"Var Mes": 1, "Fecha": ": "1992/ Jun"}}, "Jul": {"Var
Mes": 2.4, "tooltip": {"Var Mes": 2.4, "Fecha": ": "1992/ Jul"}}, "Ago": {"Var Mes": 0.4, "tooltip": {"Var Mes": 0.4, "Fecha": ": "1992/ Ago"}}, "Sep": {"Var Mes": 1.3,
"tooltip": {"Var Mes": 1.3, "Fecha": ": "1992/ Sep"}}, "Oct": {"Var Mes": 0.9, "tooltip": {"Var Mes": 0.9, "Fecha": ": "1992/ Oct"}}, "Nov": {"Var Mes": 0.8, "tooltip":
{"Var Mes": 0.8, "Fecha": ": "1992/ Nov"}}, "Dic": {"Var Mes": 2.5, "tooltip": {"Var Mes": 2.5, "Fecha": ": "1992/ Dic"}}}}], {"anio": 1993, "regiones": {"PAIS":
{"meses": {"Ene": {"Var Mes": 0.2, "tooltip": {"Var Mes": 0.2, "Fecha": ": "1993/ Ene"}}, "Feb": {"Var Mes": 3, "tooltip": {"Var Mes": 3, "Fecha": ": "1993/ Feb"}}, "Mar":
{"Var Mes": 0.1, "tooltip": {"Var Mes": 0.1, "Fecha": ": "1993/ Mar"}}, "Abr": {"Var Mes": 4.3, "tooltip": {"Var Mes": 4.3, "Fecha": ": "1993/ Abr"}}, "May": {"Var Mes":
0.9, "tooltip": {"Var Mes": 0.9, "Fecha": ": "1993/ May"}}, "Jun": {"Var Mes": 0.9, "tooltip": {"Var Mes": 0.9, "Fecha": ": "1993/ Jun"}}, "Jul": {"Var Mes": 0.1,
```

Imagen: estructura de los datos correspondientes a IMACON.

# Instalación de Jupyter Notebook

(Se debe contar con Python3 y Pip instalados<sup>2</sup>)

Las instrucciones para la instalación de Jupyter Notebook se encuentra aquí:

<https://jupyter.org/install>

Una vez instalado, siempre en terminal/command line, se debe ejecutar en la carpeta donde se localizan los notebooks, el comando:

```
$ jupyter notebook
```

Esto abrirá de forma automática el *browser* predeterminado y los notebook para cada sección:

```
$ src > frontend > public
```

## Procesamiento de los datos

Los datos de entrada están ya configurados para cada notebook de cada sección. Por ende, si existe algún nuevo datos que se quiera agregar, basta con respetar el formato de las tablas excel, subirlas a la carpeta antes mencionada (\$ src > frontend > public) y ejecutar los notebooks.

Para hacer correr la totalidad de funciones para los indicadores de cada sección, de debe presionar en cada notebook la opción. 'Run All':

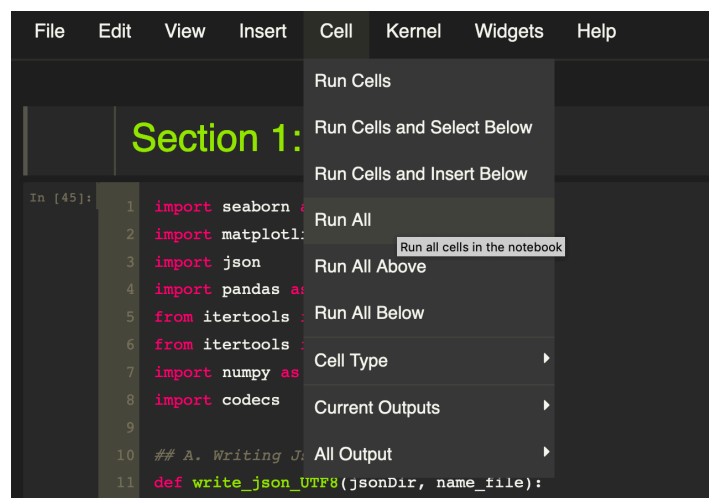


Imagen: comando *run all* en Jupyter Notebook.

---

<sup>2</sup> Ver. <https://tecnonucleous.com/2018/01/28/como-instalar-pip-para-python-en-windows-mac-y-linux/>

Los datos de salida tendrán el formato adecuado que de forma automática letrá cada módulo de visualización. Es decir, las visualizaciones construyen sin mayor mediación las gráficas necesarias.

## Data uploading

Los datos en la aplicación deben tener presencia en dos espacios para poder hacer un render de la visualización. Los datos deben estar presentes tanto en la carpeta 'Data' en la ruta \$ src > frontend > public > data; tanto como en la carpeta 'SmartData-Files/data' en Digital Ocean<sup>3</sup>. El que deban tener presencia en ambos espacios radica en que las visualizaciones, al mismo tiempo, trabajan en local y en la nube.

El proceso de subida de datos a digital ocean es tan sencillo como arrastrar un archivo y configurar los permisos del mismo para que sea visible por el módulo de visualización:

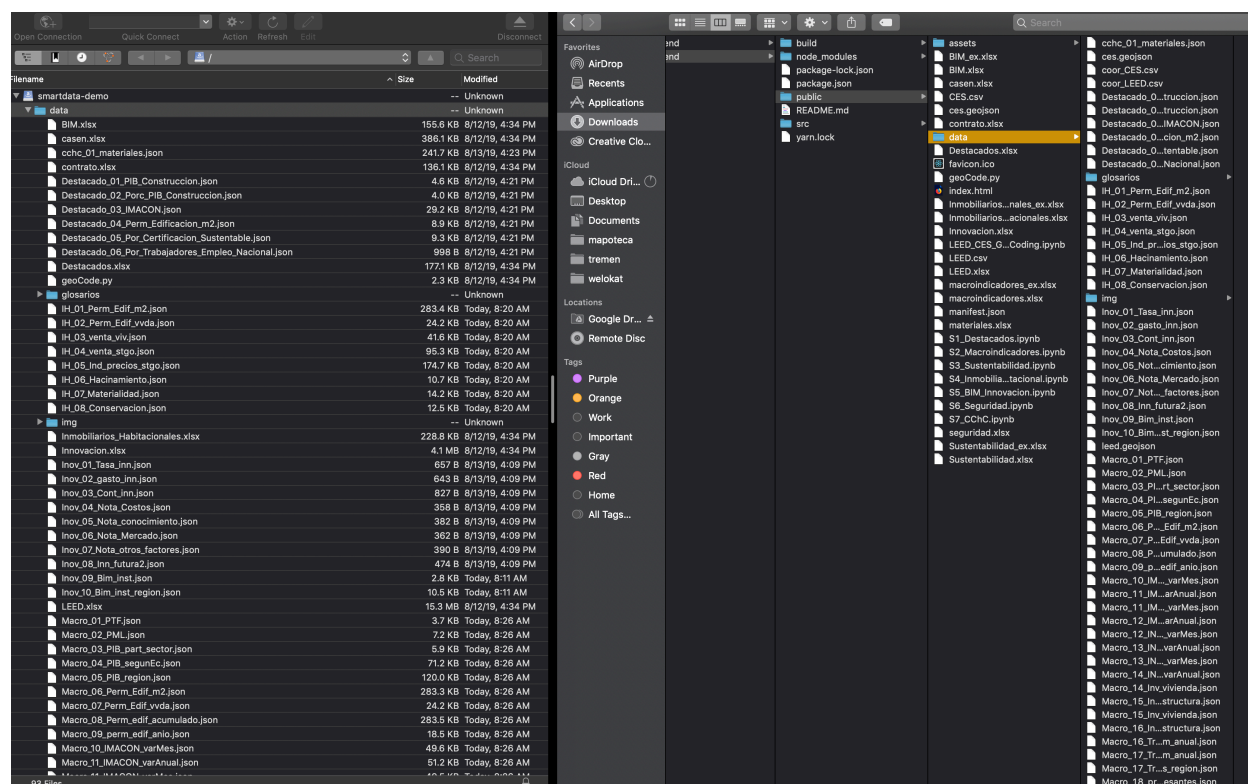
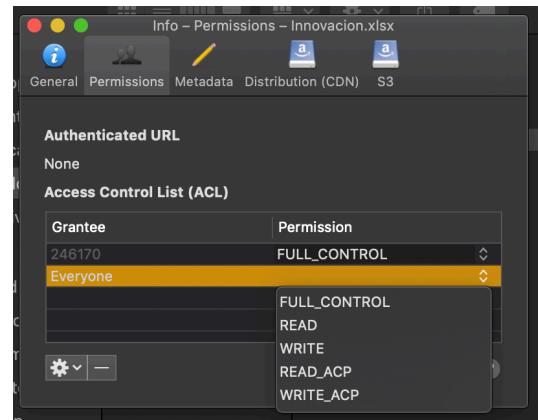
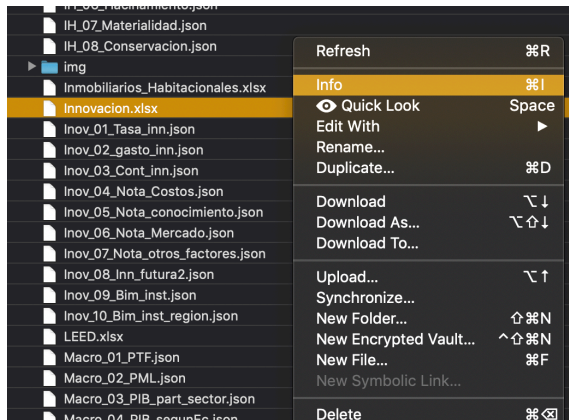


Imagen: a la derecha la carpeta en local src/frontend/public/data con los mismos datasets que encontramos en Digital Ocean (a la izquierda). En este caso, utilizamos CyberDuck para subir datos directamente a nuestro espacio en Digital Ocean.

<sup>3</sup> Ver documentación deployment

# Permisos



Para configurar permisos a cada dataset (o conjunto de datos o carpeta), se debe solamente hacer click derecho en el archivo que ha sido subido a Digital Ocean mediante CyberDuck, luego click en 'info'. Estos nos lleva a la pestaña de permisos en donde presionamos en la rueda (o gear) para añadir un permiso a 'Everyone', el cual configuramos a READ. Estos permite que las visualizaciones tengan todos los permisos necesarios para leer los datos y construir las gráficas.