



SISTEMAS INTELIGENTES PARA CIENCIA DE DATOS

Maestría en Ciencia de Datos

Universidad de la Ciudad de Aguascalientes

Anuncios

- La primera actividad ya esta en la plataforma
 - Fecha limite: 23 de febrero
- Es necesario traer laptop a las sesiones
 - Conexión a internet
 - Acceso a su cuenta institucional
- ¿Inglés?
- ¿Dudas?

En esta sesión...

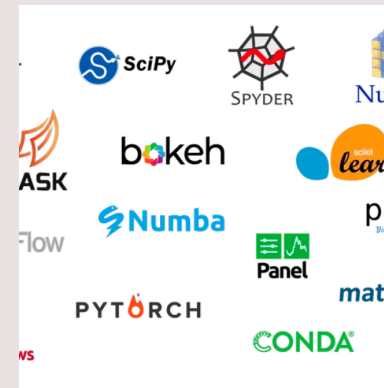
- Preliminares
 - Introducción a
 - Python
 - Jupyter
 - Google Colab
- Recuperación de información
 - HTTP Requests
 - APIs REST

Preliminares

Python

lenguaje de programación

- Características
 - Alto nivel
 - Interpretado
 - Código abierto
 - Multiparadigma
 - Tipos dinámicos
- Ventajas
 - Sintaxis sencilla
 - Gran comunidad
 - Muchas librerías



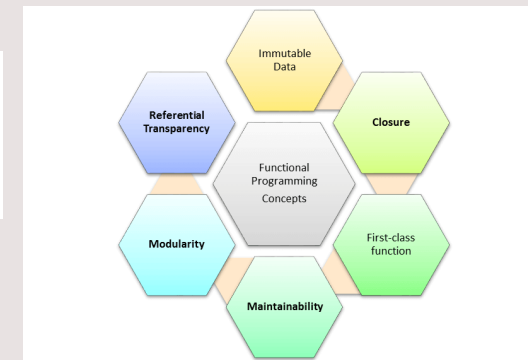
```
foo:
movl $0xFF001122, %eax
addl %ecx, %edx
xorl %esi, %esi
pushl %ebx
movl 4(%esp), %ebx
leal (%eax,%ecx,2), %esi
cmpl %eax, %ebx
jnae foo
retl
```

```
>>> sorted([5, 2, 3, 1, 4])
[1, 2, 3, 4, 5]
```

```
int num1 = 5 + static_cast<int>(12.75);
double num1 = static_cast<double>(5) + 12.75;
char x = static_cast<char>(65);
```



python™



Cerca de 1,360,000,000 resultados (

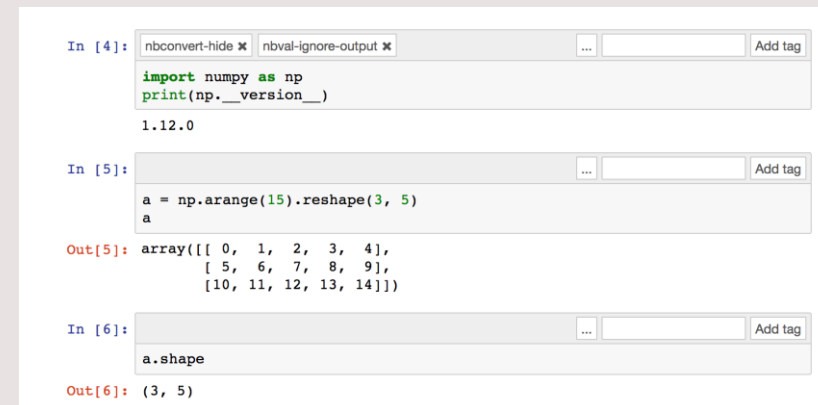
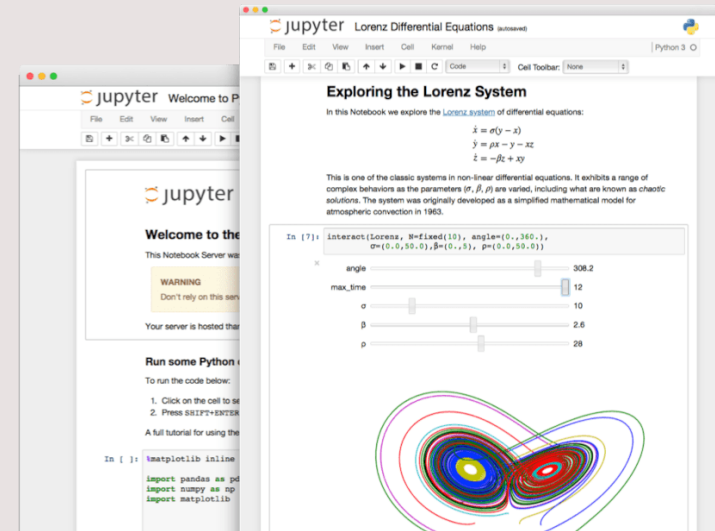
¿Cómo instalar Python?

- Windows
<https://www.python.org/downloads/windows>
- MAC
<https://www.python.org/downloads/macos/>
- Linux
Apt/pacman/dnf
- Docker
https://hub.docker.com/_/python/
- Anaconda
<https://www.anaconda.com/products/distribution>
- Miniconda
<https://docs.conda.io/en/latest/miniconda.html>

Jupyter

JupyterLab

- Basado en cuadernos
 - Combina código, documentación y visualizaciones
 - Permite ejecutar pequeñas secciones de código de manera interactiva
 - Soporta distintos lenguajes
- Fortalezas
 - Excelente para experimentar y hacer prototipos
 - Permite crear documentos para compartir resultados
- Debilidades
 - Complicado en proyectos con muchos archivos
 - No esta diseñado para flujos automáticos



¿Cómo instalar JupyterLab?

- `pip install jupyterlab`
- `conda install -c anaconda Jupyter`
- `mamba install -c conda-forge jupyterlab`
- `docker run -p 10000:8888 jupyter/scipy-notebook:2023-02-13`

- `jupyter-lab`

```
[I 11:31:38.714 NotebookApp] 0 active kernels
[I 11:31:38.714 NotebookApp] The Jupyter Notebook is running at: http://localhost:8888/?token=bf955b36f6957215ee6c774bcfeb23649334fb88f5040087
[I 11:31:38.714 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[W 11:31:38.715 NotebookApp] No web browser found: could not locate runnable browser.
[C 11:31:38.715 NotebookApp]

Copy/paste this URL into your browser when you connect for the first time,
to login with a token:
http://localhost:8888/?token=bf955b36f6957215ee6c774bcfeb23649334fb88f5040087
[I 11:31:58.671 NotebookApp] 302 GET /?token=bf955b36f6957215ee6c774bcfeb23649334fb88f5040087 (:::1) 0.82ms
```



Google Colaboratory

- Versión modificada de Jupyter
- Conectada a tu cuenta de Google
- Acceso a Google Drive
- Gratuito
 - Colab Pro
- Limitantes
 - Hasta 12 horas de ejecución
 - Memoria limitada
 - Cantidad de procesadores
 - GPU mas lentas
 - Código, datos y ejecución están en servidores externos

<https://colab.research.google.com>

The background features a light gray gradient. In the corners, there are several circles of various sizes and colors, including teal, orange, and light green, creating a modern, abstract design.

¿Dudas?

The background features decorative elements in the corners: a cluster of light green, teal, and orange circles in the top-left, and a cluster of light green, teal, and orange circles in the bottom-right.

Recuperación de información

HTTP requests (Petición HTTP)



Ejercicio!!

- Abran una pagina web
- Den clic derecho en cualquier parte
 - Ver código fuente de la pagina



Unfettered Duty To Endanger any Trans People As

ne. Each and every day, reporters and editors
fficult decisions about which issues should
rage will influence not only how people...



BREAKING NEWS

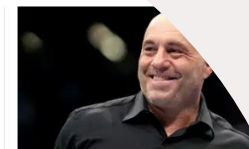
Justice Department Concurs With Matt Gaetz Defense That 17-Year-Olds Pretty Much Ready To Roll

WASHINGTON—Announcing the decision to end its sex-trafficking investigation into the Republican congressman, the Justice Department stated Friday that it concurred with Rep. Matt Gaetz's defense that 17-year-olds were pretty much ready to roll....

LOCAL

'Harry Potter' Fan Always Dreamed Of Receiving Magical Defamation Letter From J.K. Rowling

DOVER, MA—Saying that he waited for this moment since he was 11 years old, local *Harry Potter* fan Graham Wheelan told reporters Friday that he always dreamed of receiving a magical defamation letter from J.K. Rowling. "I'm a defendant! I'm a...



LOCAL

What To Say To Someone Who Is A Fan Of Joe Rogan



LOCAL

Annoyed Man Rates UberEats Driver 3 Stars For Having To Wait 10 Minutes Of The

[illegible]

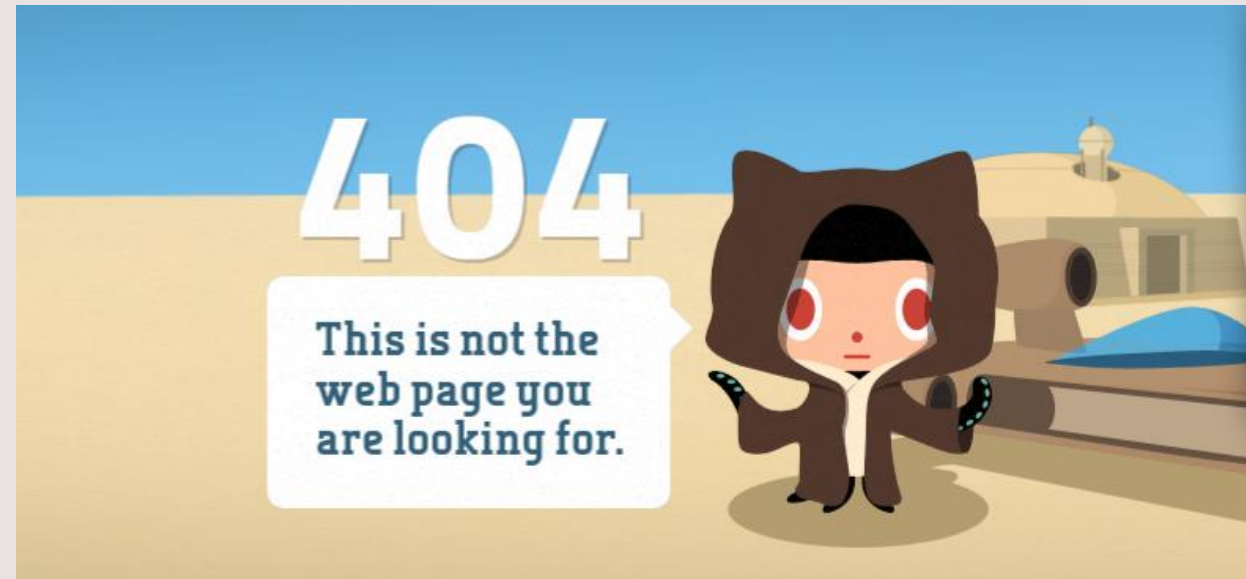
Métodos de petición

- GET
Solicita información
- POST
Envía información
- HEAD
- PUT
- DELETE
- CONNECT
- OPTIONS
- TRACE
- PATCH

Respuestas

Códigos del 100 al 599

- 100s respuestas informales
 - 100 Continue
- 200s respuestas exitosas
 - **200 OK**
 - 201 Creado
 - 202 Aceptado
- 300s respuestas de redirección
 - 301 movido permanentemente
 - 308 redirección temporal
- 400s respuestas de error del cliente
 - **404 No encontrado**
 - 418 Soy una tetera
- 500s respuestas de error del servidor
 - **500 error interno del servidor**
 - 501 no implementado



Python puede hacer requests

- Librería Requests

- `requests.get('https://api.github.com/events')`
- `requests.post('https://httpbin.org/post', data={'key': 'value'})`
- `requests.put('https://httpbin.org/put', data={'key': 'value'})`
- `requests.delete('https://httpbin.org/delete')`
- `requests.head('https://httpbin.org/get')`
- `requests.options('https://httpbin.org/get')`

- Secciones
- cookies
- codificación
- encabezados
- códigos de estatus
- redirecciones



¿Por qué hacer request con Python?

Base de datos			
III Trimestre (ENOE_N)	2022	CSV ↕ 29.1 MB	DBF ↕ 22.8 MB
		DTA ↕ 51.8 MB	SAV ↕ 30.9 MB
II Trimestre (ENOE_N)	2022	CSV ↕ 29.1 MB	DBF ↕ 27.0 MB
		DTA ↕ 51.8 MB	SAV ↕ 30.9 MB
I Trimestre (ENOE_N)	2022	CSV ↕ 27.3 MB	DBF ↕ 25.6 MB
		DTA ↕ 43.5 MB	SAV ↕ 34.0 MB
IV Trimestre (ENOE_N)	2021	CSV ↕ 26.7 MB	DBF ↕ 25.3 MB
		DTA ↕ 41.6 MB	SAV ↕ 32.9 MB
III Trimestre (ENOE_N)	2021	CSV ↕ 33.0 MB	DBF ↕ 31.3 MB
		DTA ↕ 51.8 MB	SAV ↕ 39.6 MB
II Trimestre (ENOE_N)	2021	CSV ↕ 26.3 MB	DBF ↕ 22.5 MB
		DTA ↕ 37.0 MB	SAV ↕ 31.0 MB
I Trimestre (ENOE_N)	2021	CSV ↕ 24.3 MB	DBF ↕ 22.8 MB
		DTA ↕ 39.1 MB	SAV ↕ 30.1 MB
IV Trimestre (ENOE_N)	2020	CSV ↕ 22.8 MB	DBF ↕ 21.6 MB
		DTA ↕ 35.3 MB	SAV ↕ 30.0 MB
III Trimestre (ENOE_N)	2020	CSV ↕ 22.5 MB	DBF ↕ 21.1 MB
		DTA ↕ 38.9 MB	SAV ↕ 27.8 MB

salario mínimo> Horas> Infoenoe> Infolaboral> Informal> Informalidad> Ingresos> pesos> población ocupada> ingresos> establecimientos> Micronegocios> Negocios> PEA> Pequeños establecimientos> percepciones> asalariada> Pruebas> significancia> estadística> Restaurantes> servicios> SCIAN> Subocupación> Tasas> Vínculo laboral> Enoe> ene", "license" : "https://www.inegi.org.mx/inegi/terminos.html", "creator": {"@type": "Organization", "url": "https://www.inegi.org.mx/", "name": "Instituto Nacional de Estadística y Geografía", "contactPoint": {"@type": "ContactPoint", "contactType": "customer service", "email": "atencion.usuarios@inegi.org.mx"}}, "includedInDataCatalog": {"@type": "DataCatalog", "name": "data.gov"}, "distribution": [{"@type": "DataDownload", "encodingFormat": "CSV", "contentUrl": "https://www.inegi.org.mx/contenidos/programas/enoe/15ymas/datosabiertos/2020/conjunto_de_datos_enoe_2020_1t_csv.zip"}], "temporalCoverage": "2020"}</script>

<meta http-equiv="x-ua-compatible" content="IE=edge">

<meta charset="UTF-8">

<meta name="keywords" content="15 años, Actividad, económica, Agricultura, laboral, población, económicamente, activa, asalariados, contrato, Desempleo, Desocupado, trabajar, empleo, jornada, trabajo, edad, empleo, Empresas, salario mínimo, Horas, Infoenoe, Infolaboral, Informal, Informalidad, Ingresos, pesos, población ocupada, ingresos, establecimientos, Micronegocios, Negocios, PEA, Pequeños establecimientos, percepciones, asalariada, Pruebas, significancia, estadística, Restaurantes, servicios, SCIAN, Subocupación, Tasas, Vínculo laboral, Enoe, ene">

<meta name="description" content="La Encuesta Nacional de Ocupación y Empleo (ENOE) es la principal fuente de información sobre el mercado laboral mexicano al ofrecer datos mensuales y trimestrales de la fuerza de trabajo, la ocupación, la informalidad laboral, la subocupación y la desocupación. Constituye también el proyecto estadístico continuo más grande del país al proporcionar cifras nacionales y de cuatro">

<meta name="robots" content="index, follow, odp">

<meta name="COLOR" content="e16c32">

<meta name="URL_IMAGE" content="/img/programas/enchogares/enoe_ch.gif">

<meta name="author" content="Instituto Nacional de Estadística y Geografía (INEGI)">

<link rel="stylesheet" href="/estilos/sitio/con_sitio.min.css">

<link rel="stylesheet" href="/estilos/sitio/plantillas.min.css">

<link href="/programas/css/style.min.css?v=0.0.2" rel="stylesheet">

<script src="/lib/jquery/jquery-1.11.2.min.js"></script>



The image features a light gray background with decorative circles in the corners. The top-left corner has a large light blue circle, a small orange circle, and two small green circles. The top-right corner has a large green circle, a small orange circle, and a medium teal circle. The bottom-right corner has a small yellow circle, a small green circle, a medium light blue circle, and a large yellow circle. The text "¿Dudas?" is centered on the left side of the image.

¿Dudas?

APIs REST

- API=Application Programming Interface → Interfaz de Programación de Aplicaciones
Es un conjunto de reglas de como se deben de comunicar los dispositivos.
- REST = Representational State Transfer → Transferencia de estado representacional
Es un modelo arquitectónico que provee gran flexibilidad
- Principios Básicos
 - Interfaz uniforme: todas las solicitudes al mismo recurso deben resultar iguales
 - Cliente y servidor desacoplados: El cliente y el servidor solo pueden interactuar durante la petiencion y respuesta
 - Sin estado: en la solicitud se debe incluir toda la información necesaria
 - Capacidad de almacenamiento en caché : debe ser posible almacenar los recursos en cache
 - Arquitectura por capas: APIs le pueden preguntar a otras APIs. Eso no debe de importar

¿Por qué usar APIs

- Subconjunto del internet mas ordenado
- Diseñado principalmente para maquinas
- Podemos usar los métodos de requests GET, POST
- Respuestas codificadas (JSON)
- Eficientes
- Fáciles de entender

¿Sabías que...

- INEGI tiene sus propias API de consulta
https://www.inegi.org.mx/servicios/api_indicadores.html
- Es necesario un token
 - Gratuito
 - Registrar correo electrónico
- La pagina cuenta con guías y ejemplos

API del Banco de Indicadores

Versión 2.0



La API de Indicadores del INEGI permite consultar los datos y metadatos de los indicadores disponibles a nivel nacional, por entidad federativa y municipio. Puede utilizar la API para crear aplicaciones que muestren la información directamente de las bases de datos del INEGI en el preciso momento en que se actualiza.

[Lista de feeds](#)

API del DENUÉ

La API del DENUÉ te permite consultar datos de identificación, ubicación, actividad económica y tamaño de más de 5 millones de establecimientos a nivel nacional, por entidad federativa y municipio. Puedes utilizar la API para crear aplicaciones que muestren la información directamente de las bases de datos del INEGI en el preciso momento en que se actualiza.

API de Google Maps

El complemento de Google Maps y su API de JavaScript permiten visualizar la información del INEGI en un globo terráqueo digital 2D, en tus páginas web. Puedes utilizar la API para crear aplicaciones que muestren tu información, la del INEGI y la de otras instituciones, siempre y cuando se encuentren en formato Google Maps o inclusive cargar archivos KML.

API de ruteo

El Instituto Nacional de Estadística y Geografía, a efecto de proveer conocimiento de redes de carreteras y caminos, implementa esta API de Ruteo del tipo sistemas inteligentes de transporte, que da la funcionalidad de ruteo a los sistemas de consulta de mapas del Instituto y aplicaciones externas, para que los usuarios a través de éstas interfaces, puedan planear y analizar la movilidad entre diversos destinos.

▼ Ver más

The background features a light gray gradient. In the corners, there are several circles of various sizes and colors, including teal, orange, and light green, creating a modern, abstract design.

¿Dudas?

Actividad

1. Cuaderno de requests

- descargue datos de otro programa de INEGI
- Descargue un archivo de otra pagina fuera de INEGI

- Cuadernos de Colab
- Fecha limite: 25 febrero
- Todas las celdas deben ejecutar en orden
En caso contrario debe indicarlo el cuaderno

2. Cuaderno de APIs

- Obtenga información de la API de INEGI
- Obtenga información de una API publica fuera de INEGI
- Omitir accesos personales
- Extra kudos si el cuaderno esta correctamente documentado
- Cuidar el tamaño de los cuadernos