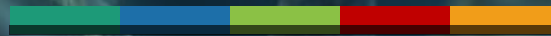


Métodos Estadísticos - Python

César A. Serna

Enero 2018



Herramientas y Métodos



Contenido



1

Introducción al Software

Uso del software de lenguaje y programación estadística.

2

Análisis Exploratorio de Datos.

Resumen descriptivo de variables, según su naturaleza.

3

Muestreo e Inferencia

Muestreo Probabilísticos, Pruebas de hipótesis paramétricas y no paramétricas. Análisis de varianza de una vía, de dos vías, verificación de supuestos, pruebas de comparación de medias.

4

Ajuste de un Modelo de Regresión.

Resumen de ajustes del modelo de regresión. Generando un diagrama de diagnóstico. Ajustar un modelo de regresión polinomial y Regresión robusta.

5

Clasificación:

Preparación de los conjuntos de datos de entrenamiento y prueba
Construyendo un modelo de clasificación con árboles de particiones recursivas
Visualizando un árbol de particionamiento recursivo



Python

Software de lenguaje y programación - estadística

Sobre "Python"

Es un lenguaje de programación y un entorno para el cálculo estadístico y elaboración de gráficas básicas avanzadas; Python es un lenguaje de programación poderoso y fácil de aprender. Cuenta con estructuras de datos eficientes y de alto nivel y un enfoque simple pero efectivo a la programación orientada a objetos. La elegante sintaxis de Python y su tipado dinámico, junto con su naturaleza interpretada, hacen de éste un lenguaje ideal para scripting y desarrollo rápido de aplicaciones en diversas áreas y sobre la mayoría de las plataformas.

Es uno de los programas estadísticos más conocidos por su capacidad para trabajar con grandes bases de datos y por poseer la mayoría de los análisis de última generación. Python es un programa amplio y flexible de análisis estadístico y gestión de información capaz de trabajar con datos procedentes de distintos formatos proporcionando, desde sencillos gráficos de distribuciones y estadísticos descriptivos, hasta análisis estadísticos complejos que permiten descubrir relaciones de dependencia e interdependencia, establecer clasificaciones de sujetos y variables, predecir comportamientos, etc.



Introducción

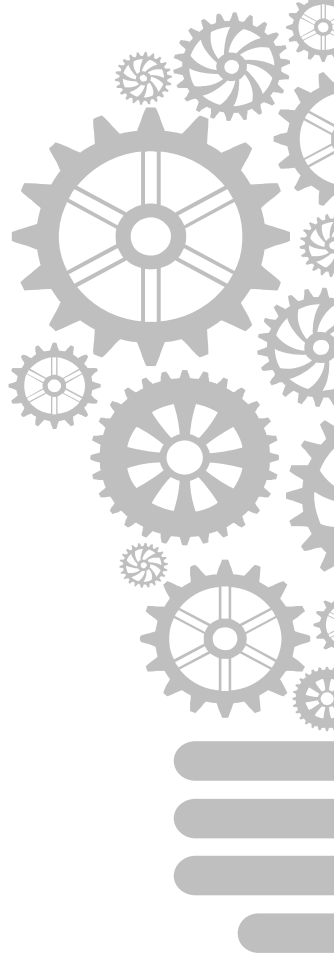
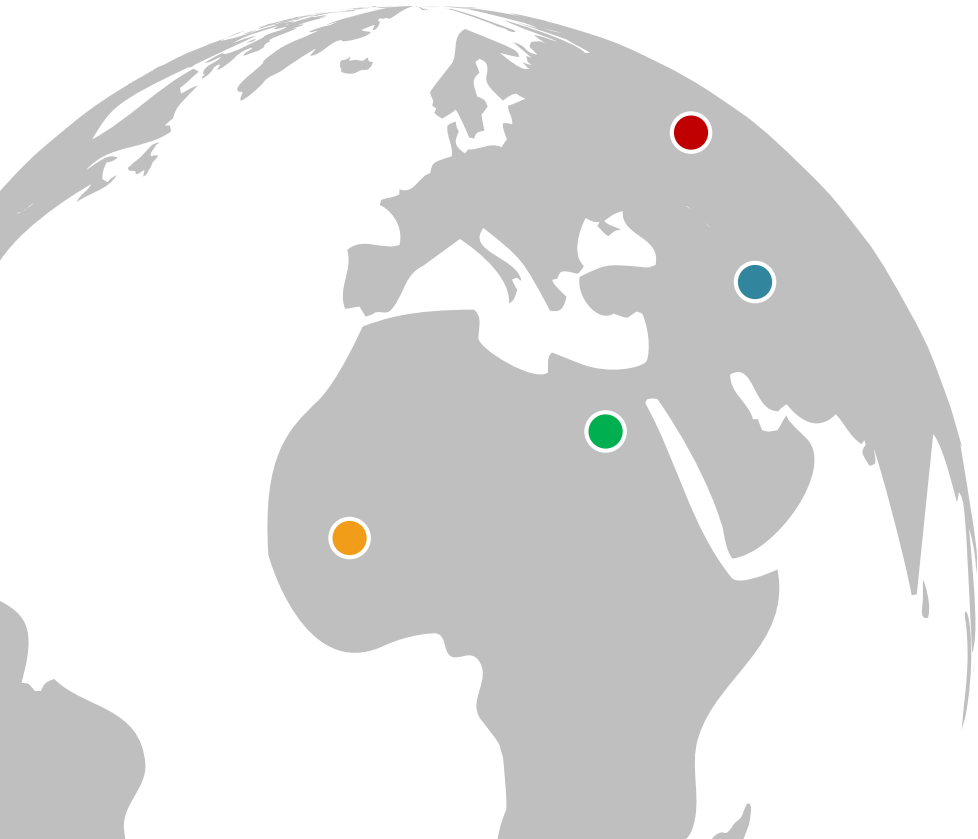
NOCIONES BÁSICAS DE

Análisis Estadístico con Python

Ventajas: *Gratuito, Gran cantidad de paquetes disponibles y muy flexible, Una variedad de bases de datos incluidas, Amplio soporte en internet.*

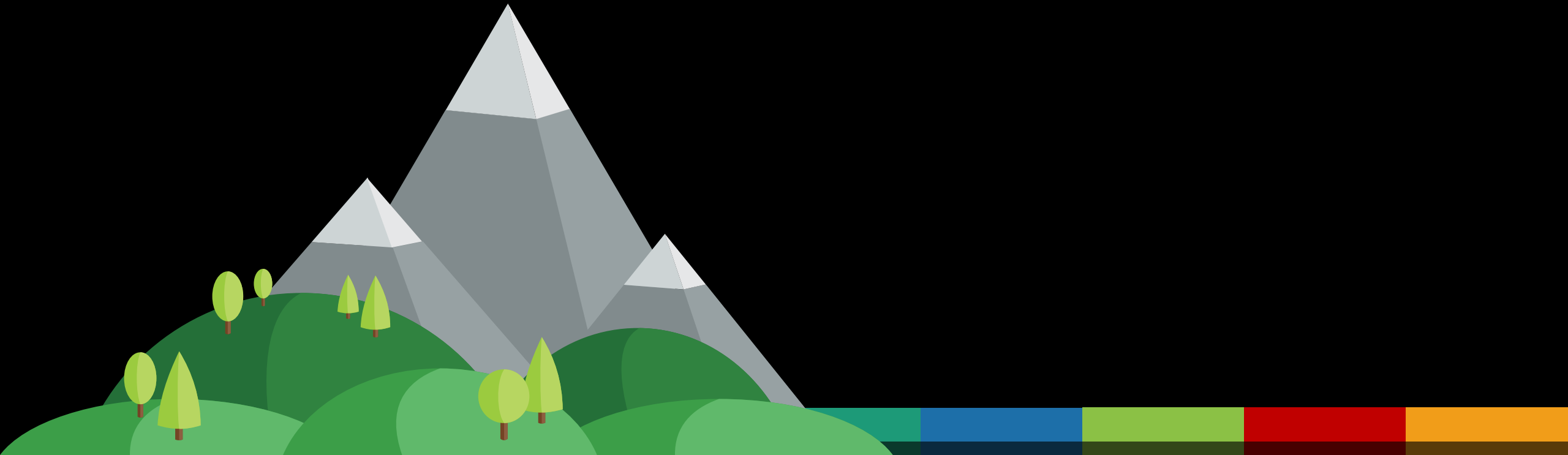
Desventajas:

Está mucho más orientado hacia programación, por tanto la curva inicial de aprendizaje muy inclinada.



Sesión 1:

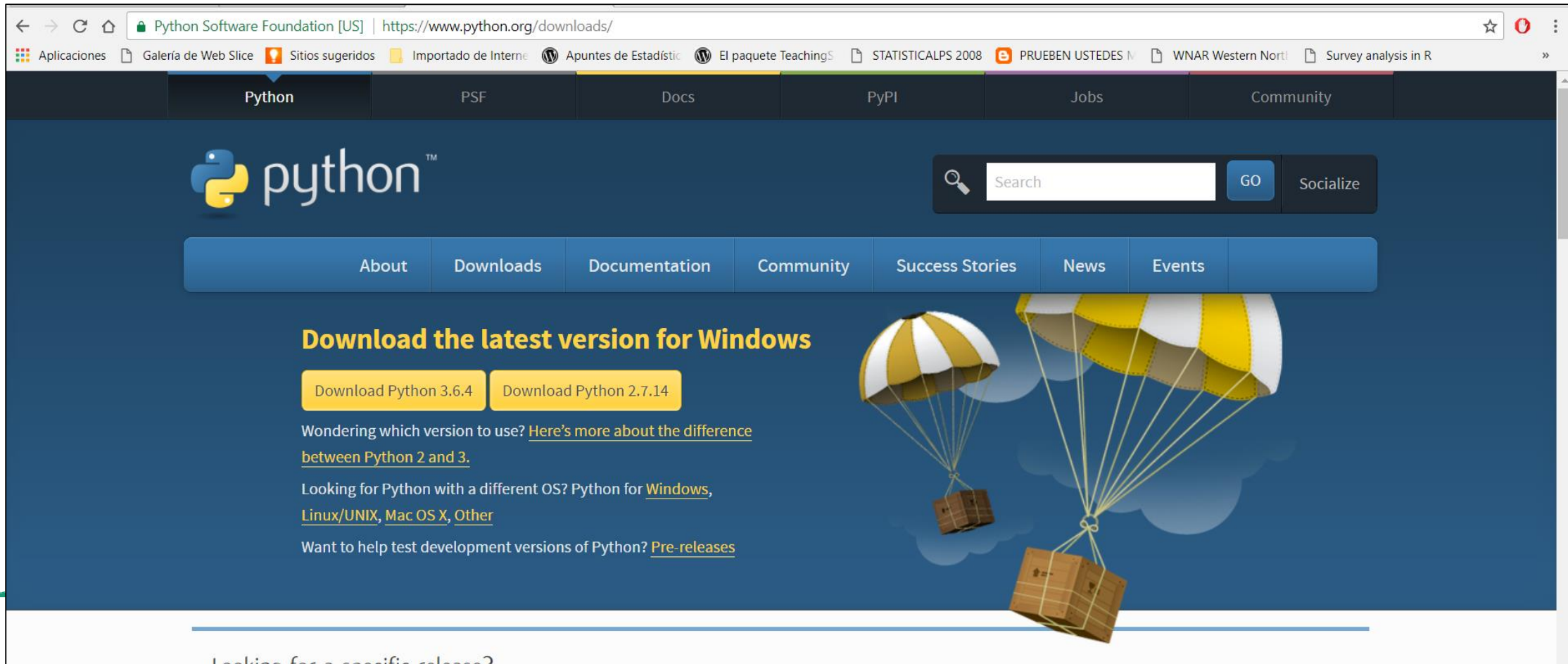
Exploración de la Herramienta Analítica: Python



Un Paseo Rápido por la Herramienta

DESCARGA BÁSICA DE SOFTWARE: <https://www.python.org/downloads/>

Anaconda:



The screenshot shows the Python Software Foundation website's download page. The browser's address bar displays the URL <https://www.python.org/downloads/>. The page features a dark blue header with the Python logo and a search bar. Below the header is a navigation menu with links for About, Downloads, Documentation, Community, Success Stories, News, and Events. The main content area is titled "Download the latest version for Windows" and includes two yellow buttons: "Download Python 3.6.4" and "Download Python 2.7.14". Below these buttons, there are links for "Here's more about the difference between Python 2 and 3.", "Python for Windows, Linux/UNIX, Mac OS X, Other", and "Pre-releases". The background of the page features an illustration of two parachutes with cargo boxes hanging from them, set against a blue sky with clouds.

Python Software Foundation [US] | <https://www.python.org/downloads/>

Python PSF Docs PyPI Jobs Community

python™ Search GO Socialize

About Downloads Documentation Community Success Stories News Events

Download the latest version for Windows

Download Python 3.6.4 Download Python 2.7.14

Wondering which version to use? [Here's more about the difference between Python 2 and 3.](#)

Looking for Python with a different OS? Python for [Windows](#), [Linux/UNIX](#), [Mac OS X](#), [Other](#)

Want to help test development versions of Python? [Pre-releases](#)

Looking for a specific release?

Un Paseo Rápido por la Herramienta

DESCARGA BÁSICA DE SOFTWARE: *Anaconda*: <https://anaconda.org/anaconda/python>

The image shows two overlapping browser windows. The background window displays the Anaconda Cloud profile page for the organization 'Anaconda', created on Sep 18, 2015, located in Austin, Texas. It lists various packages such as nodejs, sqlite, spyder, elasticsearch-dsl, ant, conda, lxml, libxslt, libxml2, and cssselect. The foreground window shows the package page for 'anaconda / packages / python 3.6.4', described as a 'General purpose programming language'. It features tabs for 'Conda', 'Files', 'Labels', and 'Badges'. The 'Conda' tab is active, showing the license (PSF), home page (http://www.python.org/), development documentation (https://docs.python.org/devguide/), and version documentation (https://www.python.org/doc/versions/). It also indicates 202499 total downloads. Below this, the 'Installers' section lists various operating system and architecture combinations for the 'conda install' command, including linux-ppc64le, osx-32, linux-64, win-32, osx-64, linux-32, and win-64, all for version v3.6.2 to v3.6.4.

Es seguro | <https://anaconda.org/anaconda>

Aplicaciones | Galeria de Web Slice | Sitios sugeridos | Importado de Interne | Apuntes de Estadístic | El paquete TeachingS | STATISTICALPS 2008 | PRUEBEN USTEDES M

ANACONDA CLOUD Search Anaconda Cloud Gallery About Pricing Anaconda Help Download

Profile

Anaconda

Organization created on Sep 18, 2015

Anaconda, Inc.
Austin, Texas
<https://www.anaconda.com/what-is-...>

2 Packages View all (1241)

- nodejs 12 hours and 7 minutes ago
- sqlite 14 hours and 7 minutes ago
- spyder 1 day and 19 hours ago
- elasticsearch-dsl 1 day and 19 hours ago
- ant 1 day and 19 hours ago
- conda 3 days and 3 hours ago
- lxml 3 days and 12 hours ago
- libxslt 3 days and 12 hours ago
- libxml2 3 days and 12 hours ago
- cssselect 3 days and 13 hours ago

Es seguro | <https://anaconda.org/anaconda/python>

Aplicaciones | Galeria de Web Slice | Sitios sugeridos | Importado de Interne | Apuntes de Estadístic | El paquete TeachingS | STATISTICALPS 2008 | PRUEBEN USTEDES M | WNAR Western

ANACONDA CLOUD Search Anaconda Cloud Gallery About Pricing Anaconda Help Download Anaconda Sig

anaconda / packages / python 3.6.4

General purpose programming language

Conda Files Labels Badges

License: PSF

Home: <http://www.python.org/>

Development: <https://docs.python.org/devguide/>

Documentation: <https://www.python.org/doc/versions/>

202499 total downloads

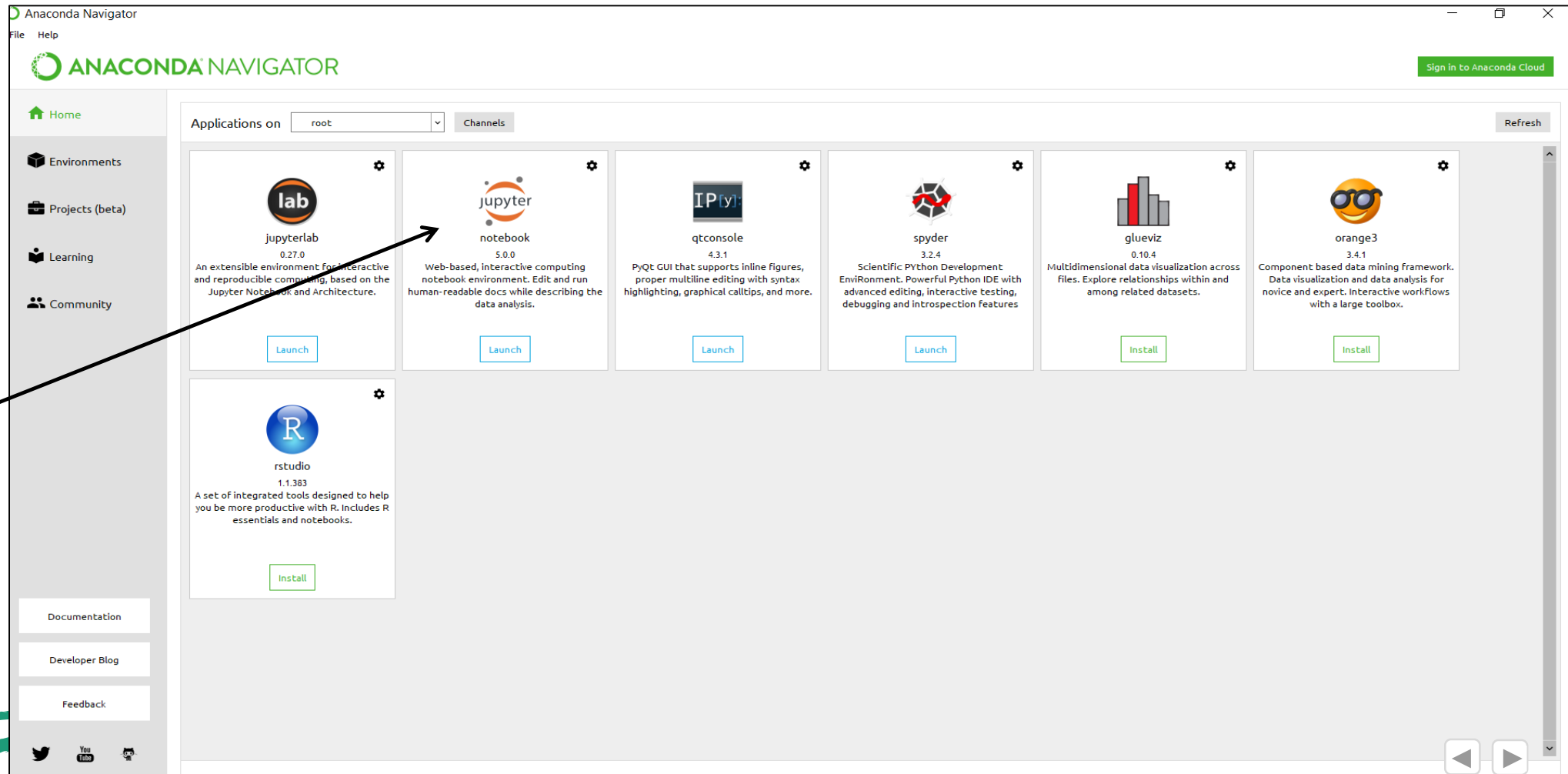
Installers

conda install ?

- linux-ppc64le v3.6.2
- osx-32 v3.4.3
- linux-64 v3.6.4
- win-32 v3.6.4
- osx-64 v3.6.4
- linux-32 v3.6.4
- win-64 v3.6.4

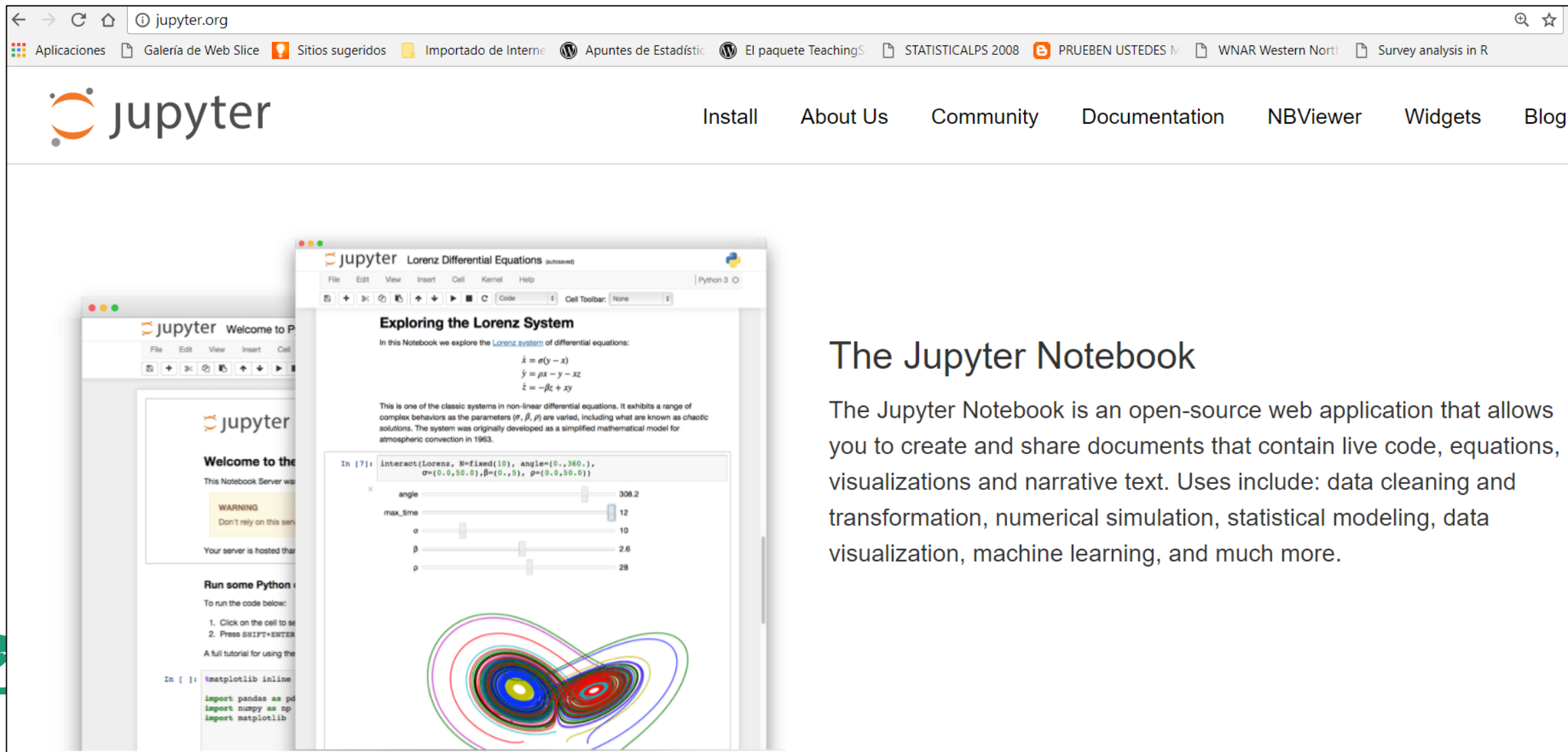
Un Paseo Rápido por la Herramienta

Windows:
Lanzamiento de
Anaconda



Un Paseo Rápido por la Herramienta

SISTEMA DE CÓDIGOS EN EL NOTEBOOK - Jupyter



The Jupyter Notebook

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

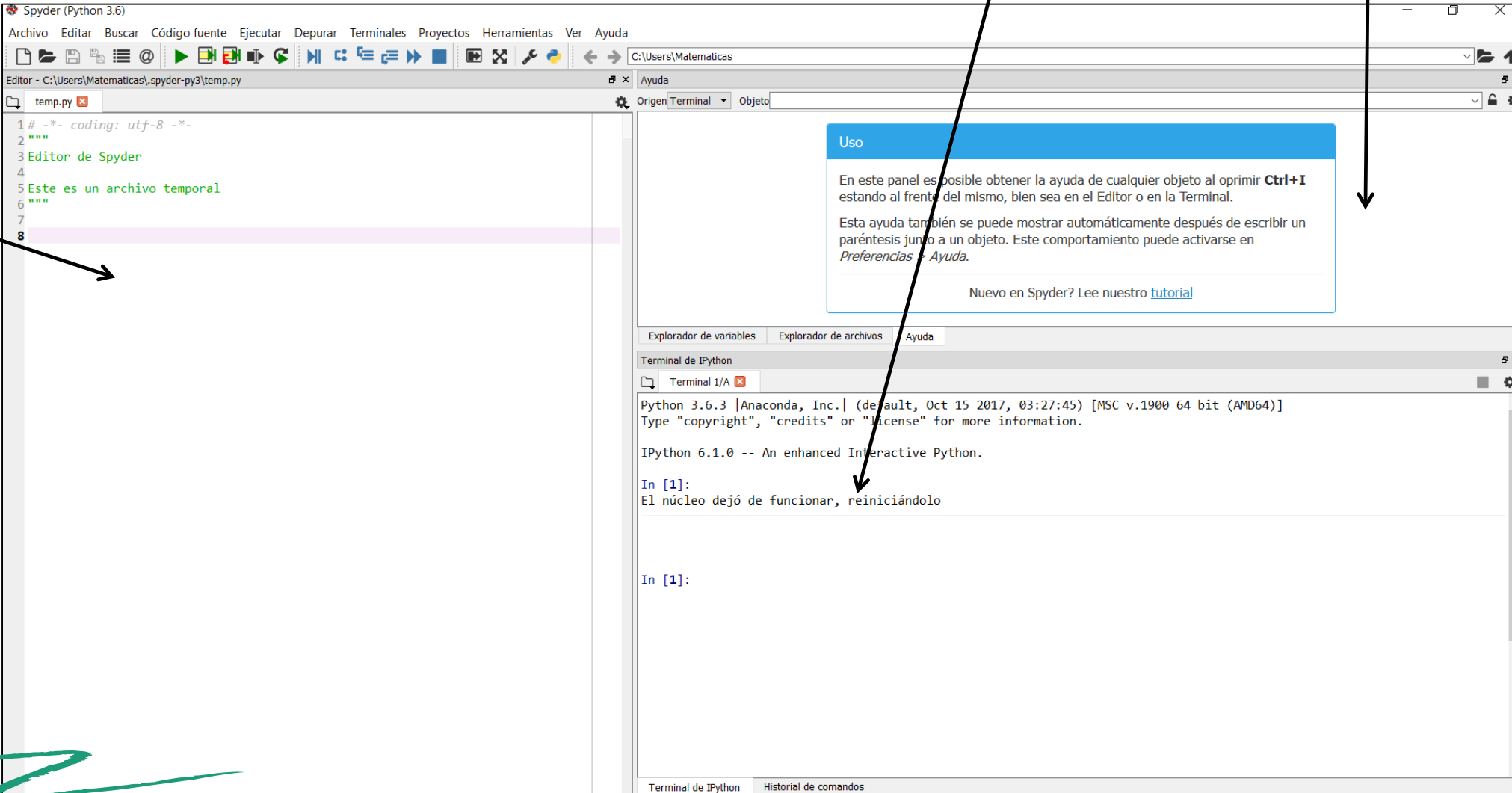
Un Paseo Rápido por la Herramienta

ASISTENTE DE PROCESAMIENTO - Spayder

Consola

Ventana de gráficas

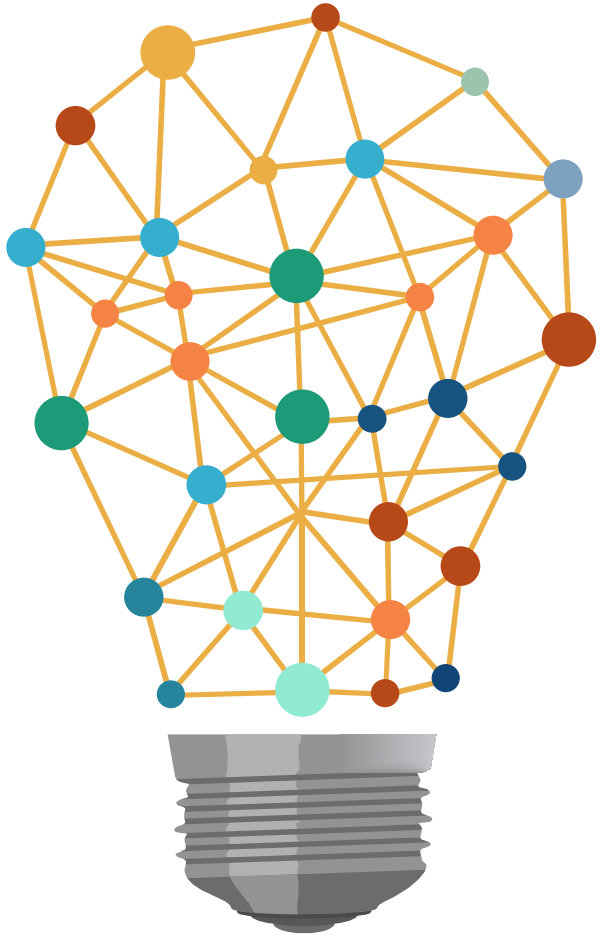
script





Un Paseo Rápido por la Herramienta

Demos Rápidos en Jupyter



Un Paseo Rápido por la Herramienta

PROCEDIMIENTO PARA INTRODUCIR DATOS EN Jupyter:

Uno de los elementos básicos para dar inicio al procesamiento de bases de datos es tener claro el tipo de variable que se desea procesar.



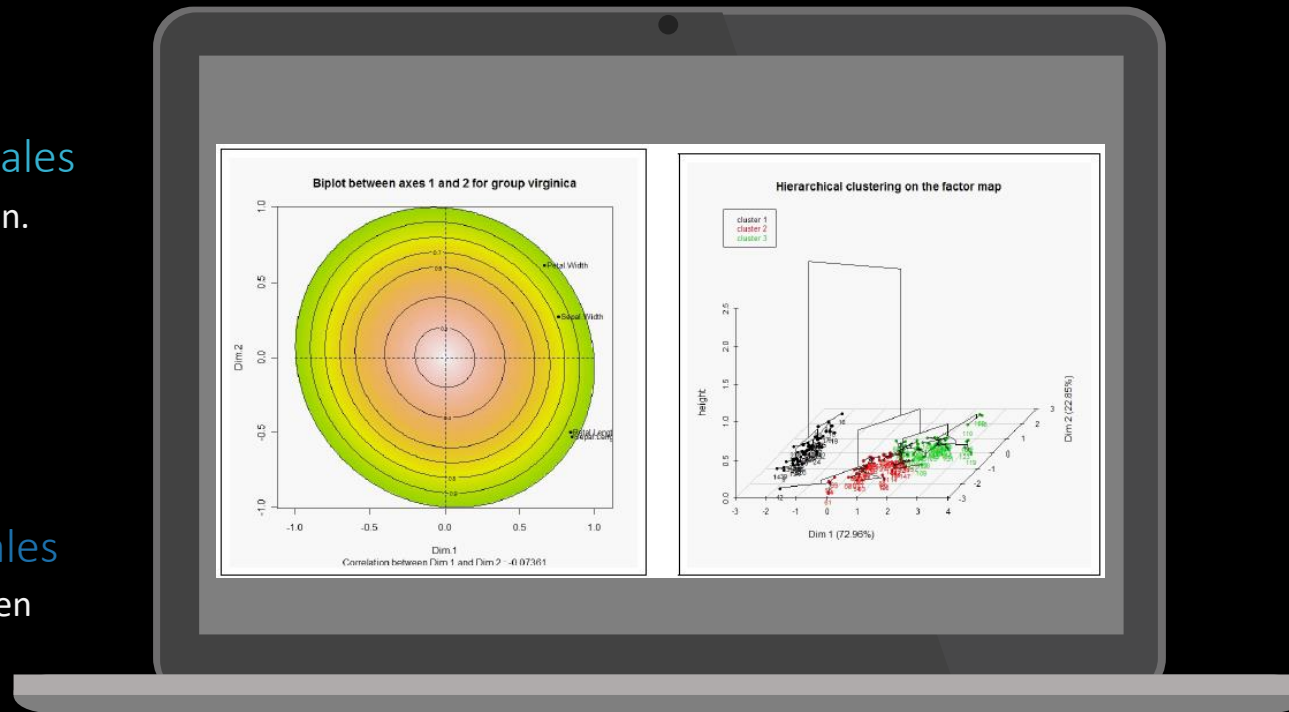
Variables Nominales

Variables cualitativas sin orden.



Variables Ordinales

Variables cualitativas con orden



Variables de Intervalo

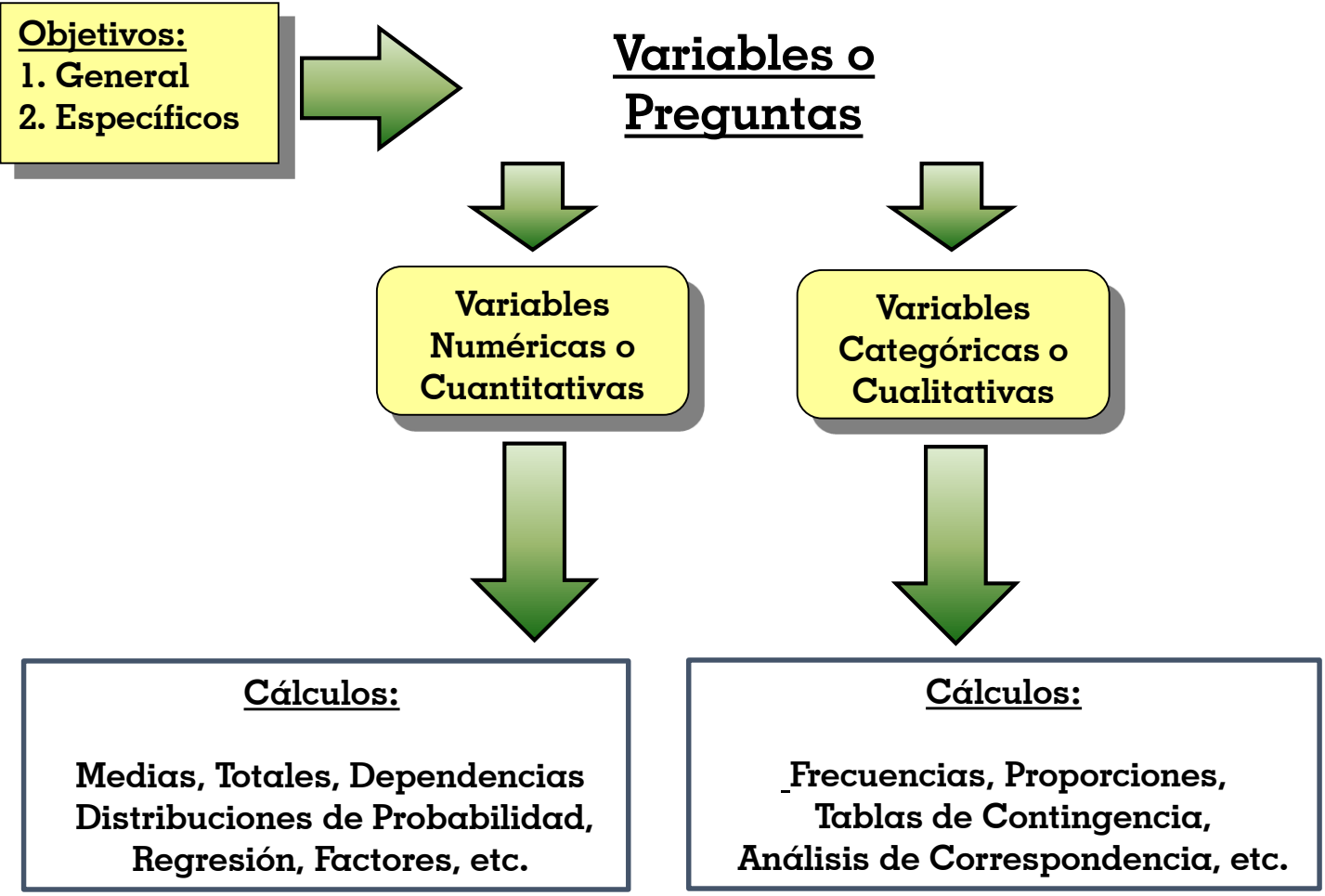
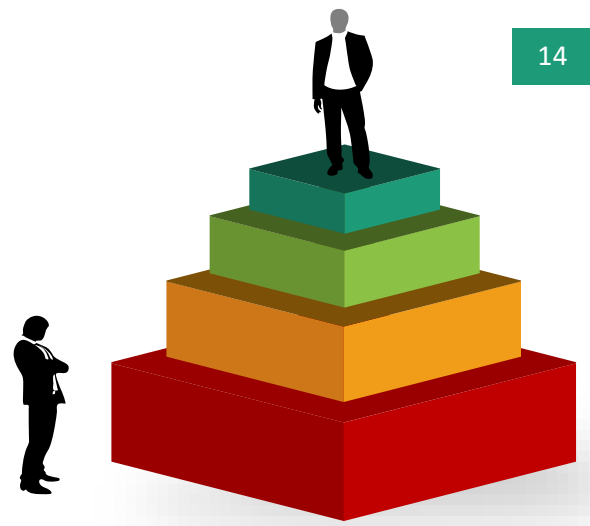
El cero es un número que no representa ausencia del atributo.



Variables de Razón

El cero SI representa ausencia del atributo.

Pasos Previos al Procesamiento



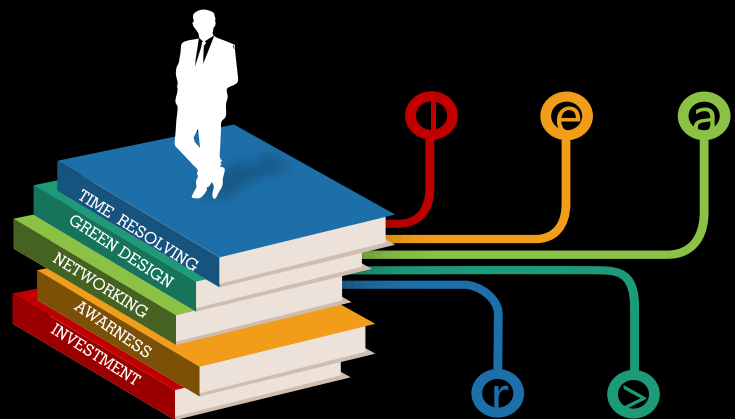
Pasos:

Antes de desarrollar cualquier análisis de información es importante tener claros el Objetivo General y los Objetivos Específicos de la investigación, pues de ellos depende el tipo de tratamiento estadístico que se aplicará.

- 1 **Objetivo General:**
- 2 **Objetivos Específicos.**
- 3 **Identificación del Tipo de Variable**
- 4 **Análisis Estadístico**



Bibliografía



LEVIN / RUBÍN /BALDERAS / DEL VALLE / GÓMEZ. Estadística para Administración y Economía. Pearson.



MENDENHALL /BEAVER/BEAVER. Introducción a la Probabilidad y Estadística.



LIND DOUGLAS / MARCHAL WILLIAMS / WATHEN SAMUEL. Estadística Aplicada a los negocios y la economía. Edición 13a. Editorial Mc. Graw Hill. 2008.



WEIER RONALD. Introducción a la Estadística para los Negocios.



ANDERSON / SWEENEY / WILLIAMS. Estadística para Administración y Economía. Cengage Learning.



JAY L. DEBORE. Probabilidad y Estadística para Ingeniería y Ciencias.



César Augusto Serna

Magister en Ciencias Estadística

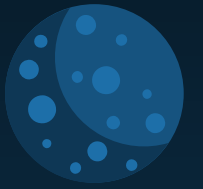
Universidad Nacional de Colombia

Consultor Estadístico

Bogotá - Colombia

📞 300 2914052

✉️ cesarserna29@gmail.com



Thanks
for Coming

