Posgrados CUCEA
El mejor lugar para el talento

MCD | MAESTRÍA EN CIENCIA DE LOS DATOS

Maestría en Ciencia de Datos (MCD)

Programación II

Unidad 1 Ciencia de datos con python

Programación II

Objetivo particular de la unidad: *Conocer acerca de bibliotecas de relacionadas con el análisis de datos en Python para crear software de manera mas rápida y eficiente*

1. Bibliotecas de ciencia de datos para Python

1.1 Ciencia de datos
1.2 Biblioteca NumPy
1.3 Biblioteca Pandas
1.4 Biblioteca rpy2
1.5 Biblioteca Scikit-leam
1.6 Distribución Anaconda
1.7 Visual Studio Code

Ciencia de Datos

Programación II

“La ciencia de datos es el estudio de datos con el fin de extraer información significativa para empresas. Es un enfoque multidisciplinario que combina principios y prácticas del campo de las matemáticas, la estadística, la inteligencia artificial y la ingeniería de computación para analizar grandes cantidades de datos.” *(AWS,2023)*

Ciencia de Datos

Programación II

¿Por qué es importante la ciencia de datos?

La ciencia de datos es importante porque combina herramientas, métodos y tecnología para generar significado a partir de los datos.

(AWS,2023)

Ciencia de Datos

Programación II

Futuro de la ciencia de datos

La inteligencia artificial y las innovaciones del machine learning han hecho que el procesamiento de datos sea más rápido y eficiente. La demanda del sector ha creado un ecosistema de cursos, grados académicos y puestos de trabajo en el campo de la ciencia de datos.

(AWS,2023)

Ciencia de Datos

Programación II

¿Para qué se utiliza la ciencia de datos?

La ciencia de datos se utiliza para estudiar los datos de cuatro maneras principales:

1. Análisis descriptivo

El análisis descriptivo examina los datos para obtener información sobre lo que ha ocurrido u ocurre en el entorno de datos. Se caracteriza por las visualizaciones de datos, como los gráficos circulares, de barras o líneas, las tablas o las narraciones generadas.

(AWS,2023)

Ciencia de Datos

Programación II

2. Análisis de diagnóstico

El análisis de diagnóstico es un examen profundo o detallado de datos para entender por qué ha ocurrido algo. Se caracteriza por técnicas como el análisis detallado, el descubrimiento y la minería de datos o las correlaciones. Se pueden llevar a cabo varias operaciones y transformaciones de datos en un determinado conjunto con el fin de descubrir patrones únicos en cada una de estas técnicas.

(AWS,2023)

Ciencia de Datos

Programación II

3. Análisis predictivo

El análisis predictivo utiliza los datos históricos para hacer previsiones precisas sobre los patrones de datos que pueden producirse en el futuro. Se caracteriza por técnicas como el machine learning, la previsión, la coincidencia de patrones y el modelado predictivo

(AWS,2023)

Ciencia de Datos

Programación II

3. Análisis predictivo

El análisis predictivo utiliza los datos históricos para hacer previsiones precisas sobre los patrones de datos que pueden producirse en el futuro. Se caracteriza por técnicas como el machine learning, la previsión, la coincidencia de patrones y el modelado predictivo

(AWS,2023)

Ciencia de Datos

Programación II

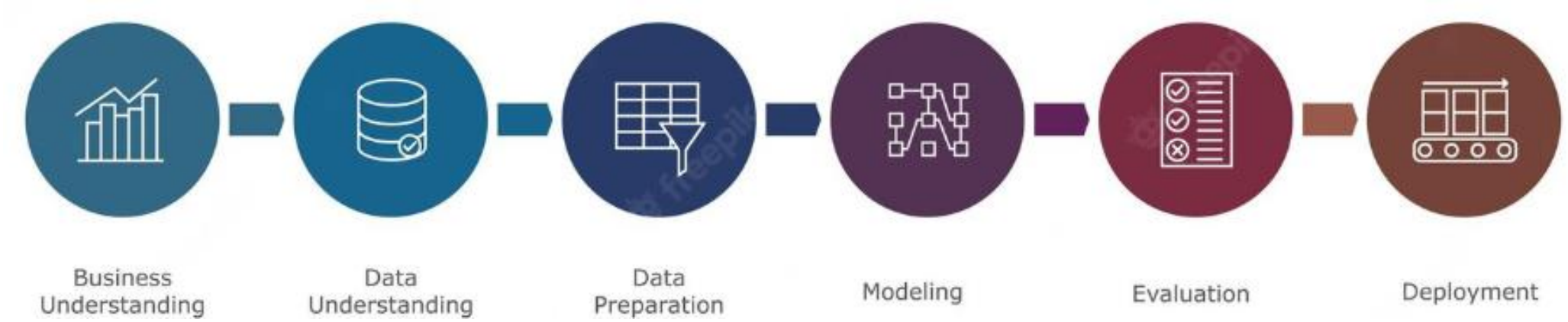
4. Análisis prescriptivo

El análisis prescriptivo lleva los datos predictivos al siguiente nivel. No solo predice lo que es probable que ocurra, sino que sugiere una respuesta óptima para ese resultado. Puede analizar las posibles implicaciones de las diferentes alternativas y recomendar el mejor curso de acción.

(AWS,2023)

Ciencia de Datos

Programación II



Anaconda

Programación II

¿Qué es el Navegador Anaconda?

Anaconda Navigator es una interfaz gráfica de usuario (GUI) de escritorio incluida en la distribución de Anaconda® que le permite iniciar aplicaciones y administrar fácilmente paquetes, entornos y canales de conda sin usar comandos de línea de comandos. Navigator puede buscar paquetes en Anaconda.org o en un repositorio local de Anaconda. Está disponible para Windows, macOS y Linux.

(Anaconda, 2023)

Anaconda

Programación II

¿A qué aplicaciones puedo acceder usando Navigator?

Las siguientes aplicaciones están disponibles de forma predeterminada en Navigator:

- JupyterLab
- Cuaderno Jupyter
- espía
- PyCharm
- Glueviz
- Aplicación naranja 3
- RStudio
- Aviso de Anaconda (solo Windows)
- Anaconda PowerShell (solo Windows)

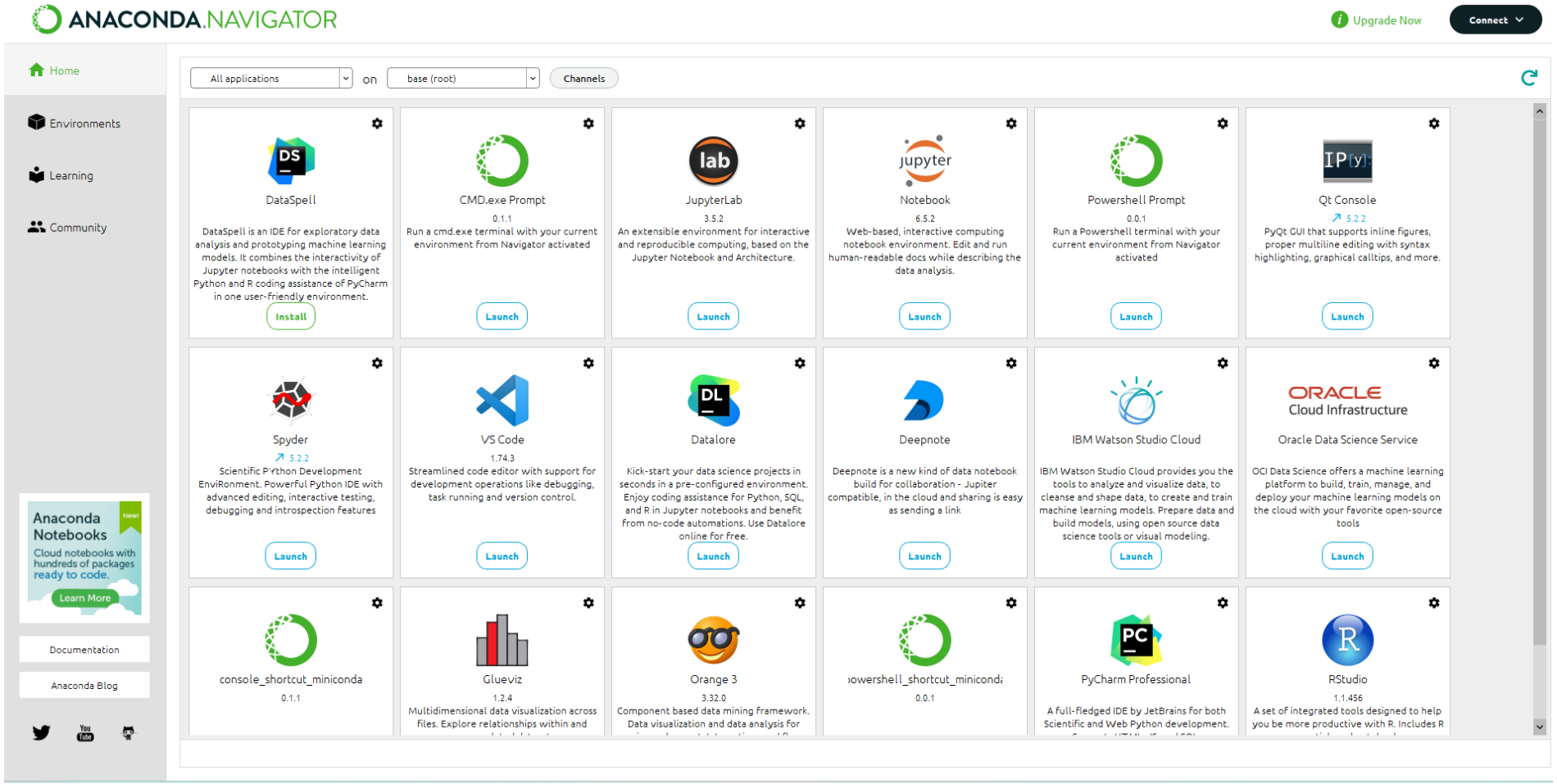
Anaconda

Programación II

https://www.anaconda.com/

Visual Studio Code

Programación II



Visual Studio Code

Programación II

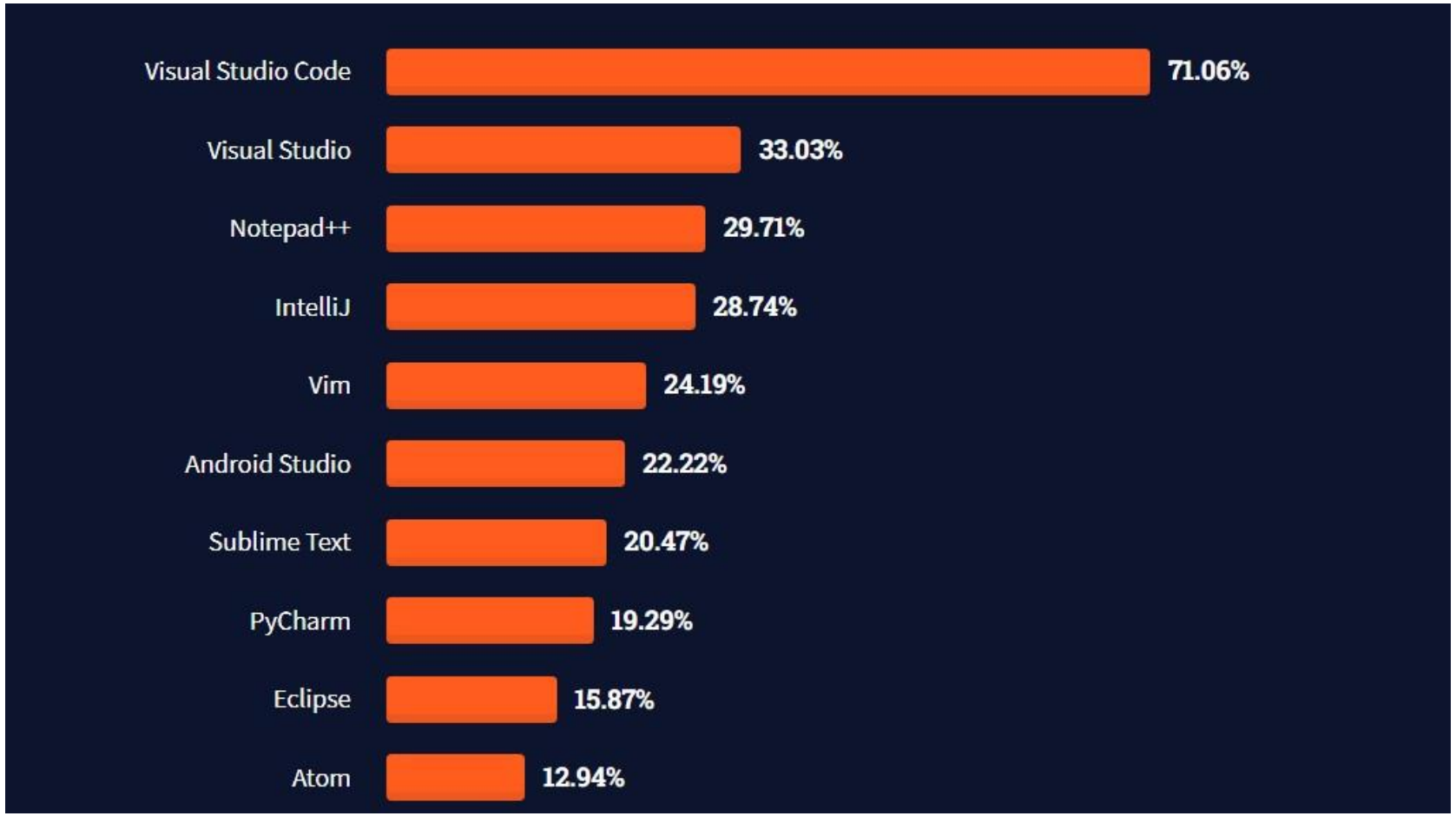
Qué es Visual Studio Code?

Visual Studio Code (VS Code) es un editor de código fuente desarrollado por Microsoft. Es software libre y multiplataforma, está disponible para Windows, GNU/Linux y macOS. VS Code tiene una buena integración con Git, cuenta con soporte para depuración de código, y dispone de un sinnúmero de extensiones, que básicamente te da la posibilidad de escribir y ejecutar código en cualquier lenguaje de programación

(OpenWebinars, 2023)

Visual Studio Code

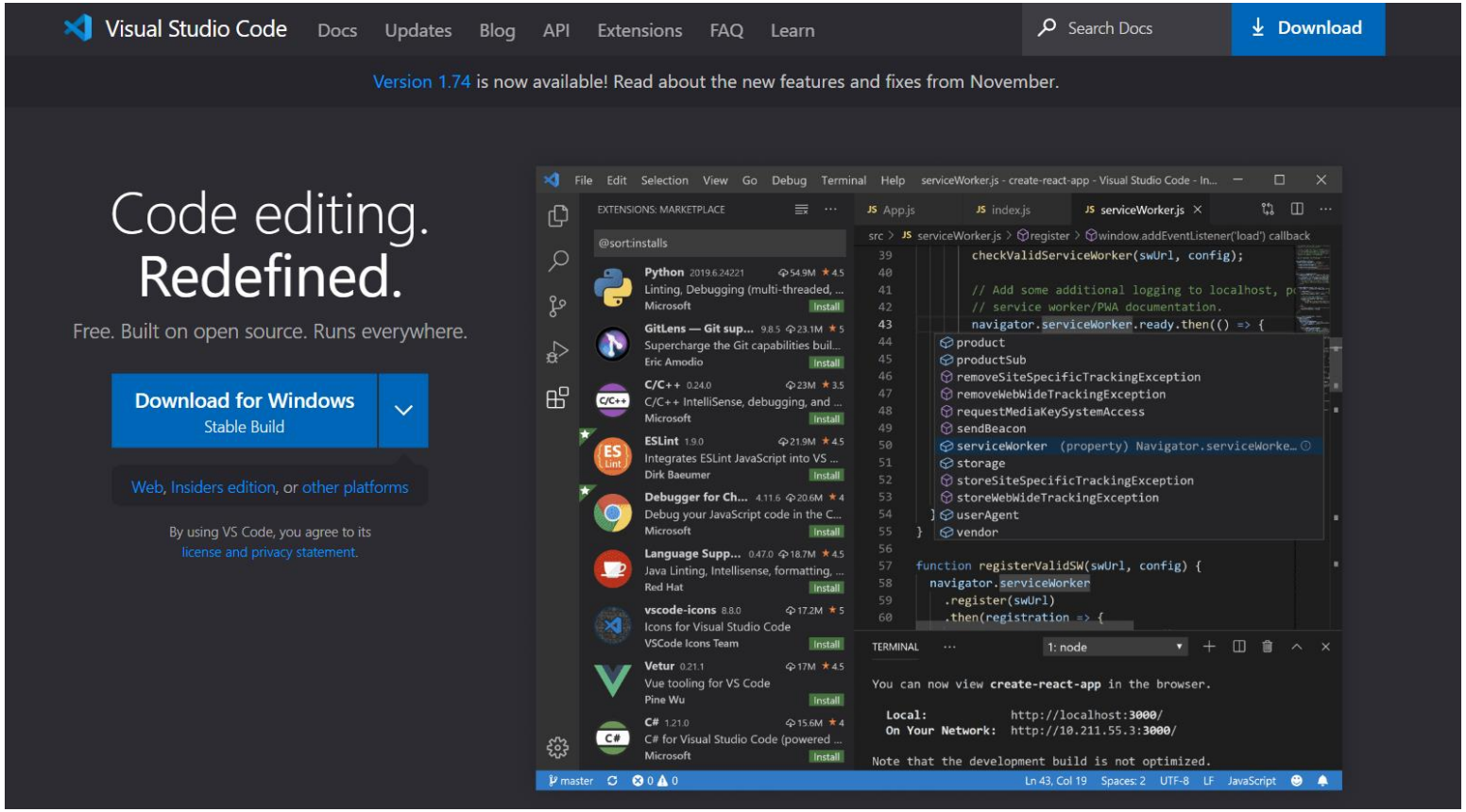
Programación II



(OpenWebinars, 2023)

Visual Studio Code

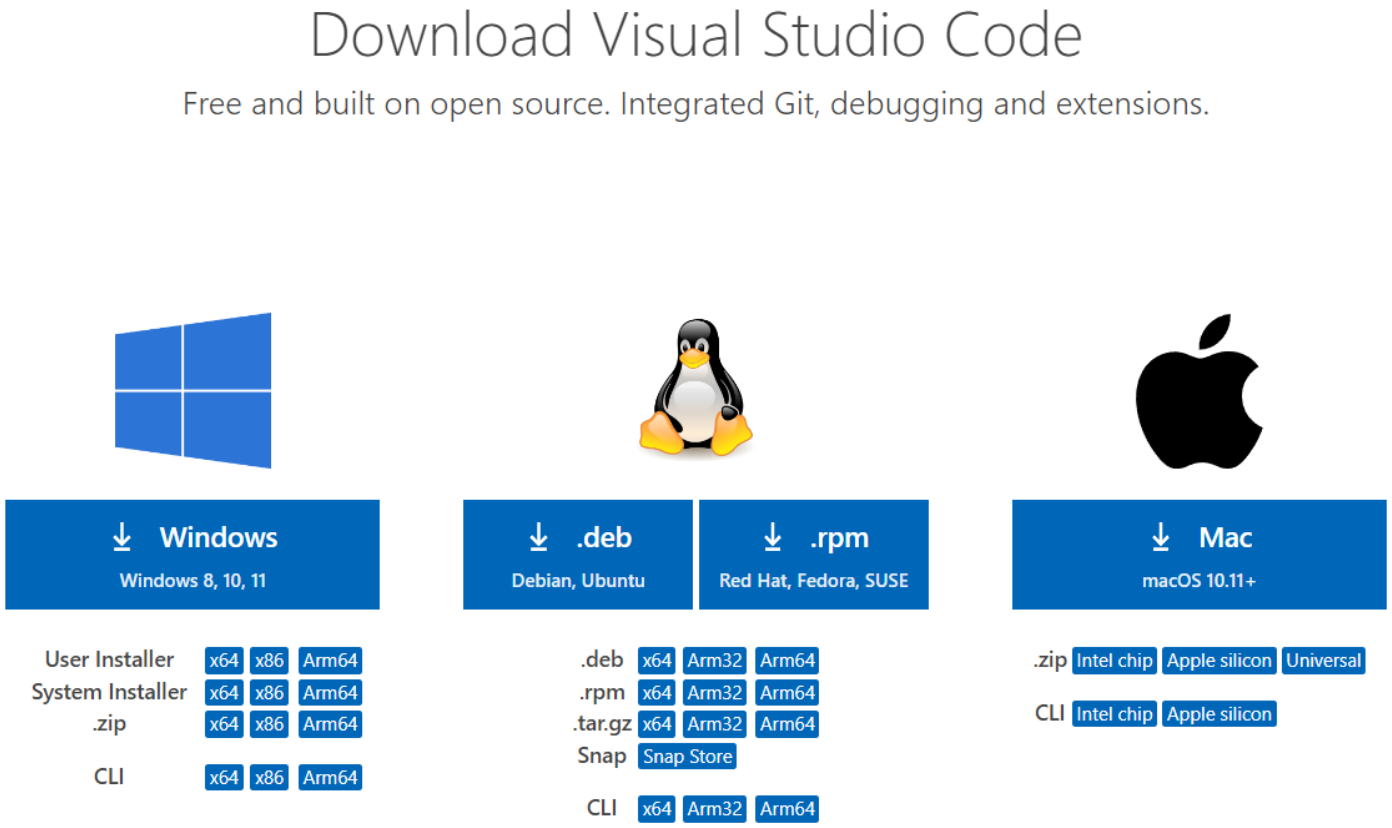
Programación II



https://code.visualstudio.com/

Visual Studio Code

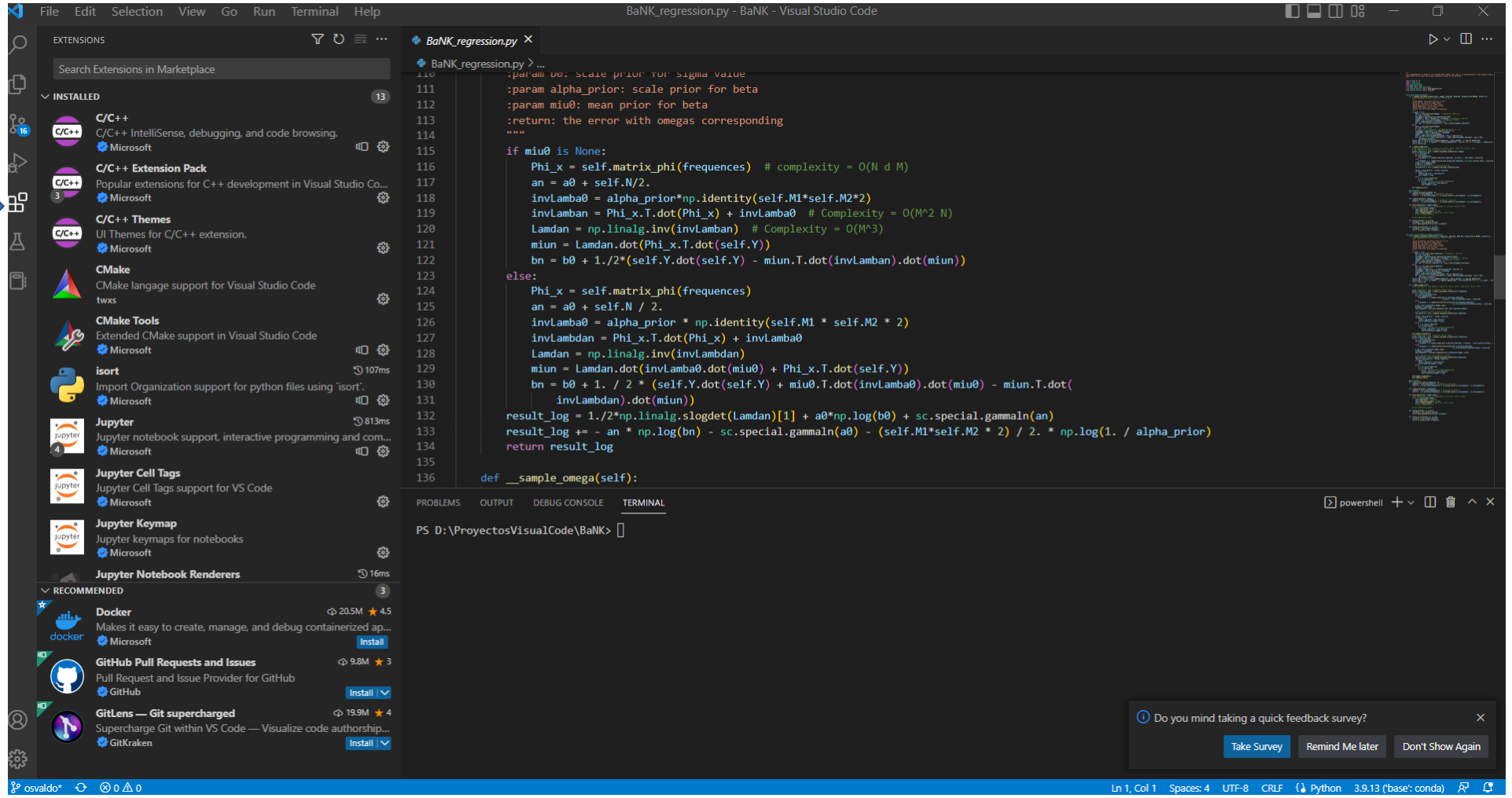
Programación II



https://code.visualstudio.com/

Visual Studio Code

Programación II

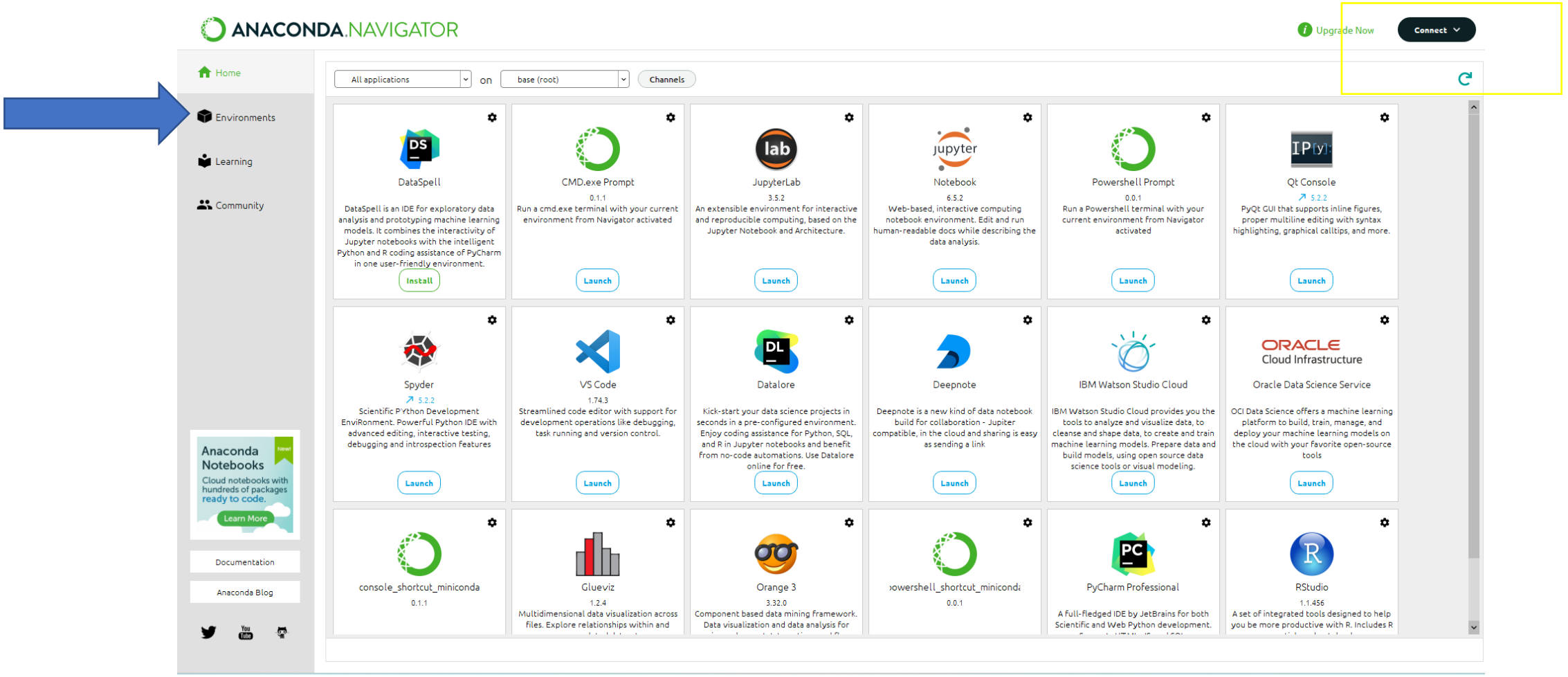


Actividades

- Instalar Python
- Instalar Jupyter
- Crear carpeta
- Abrir carpeta

Environments

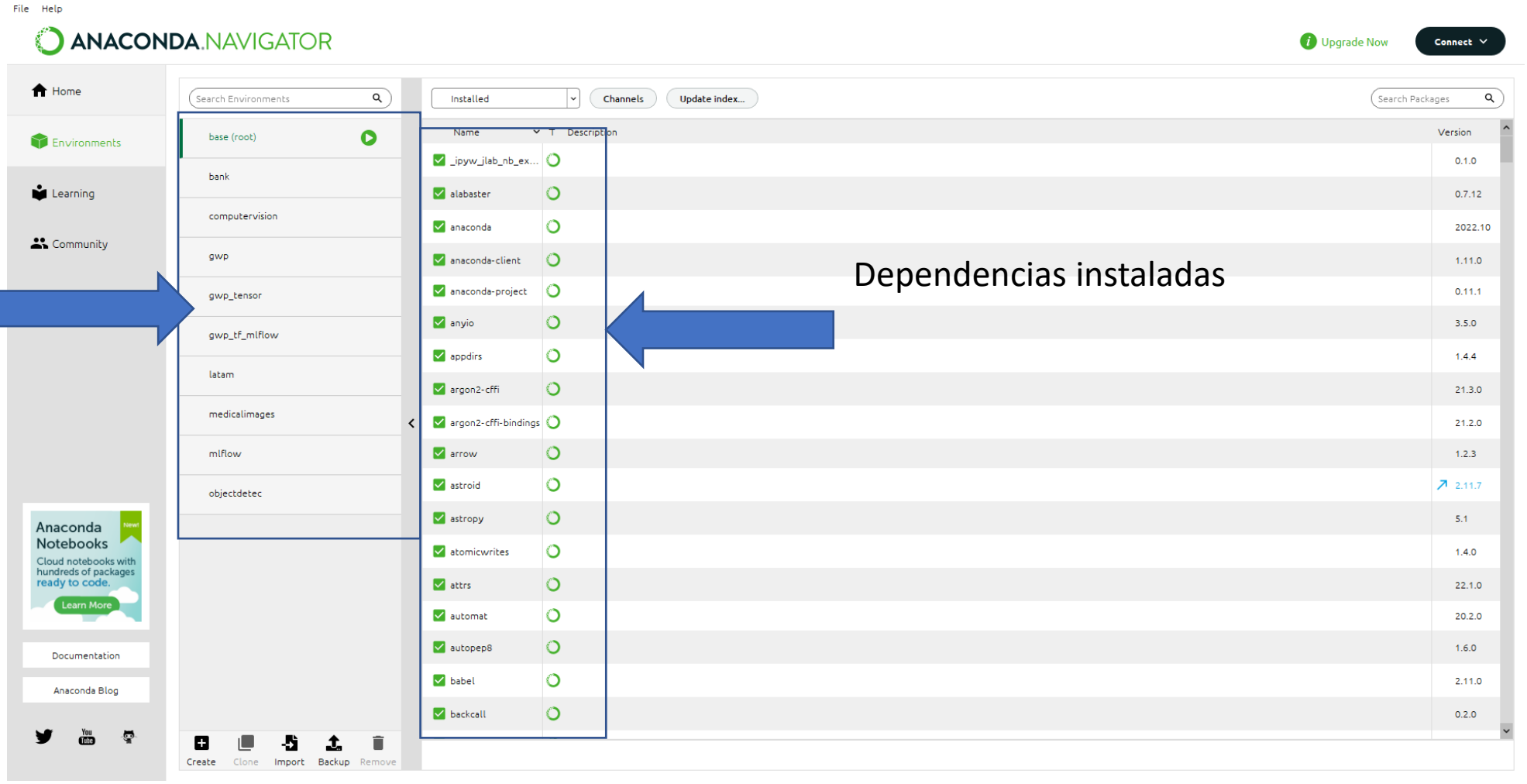
Programación II



Environments

Programación II

Environments



Dependencias instaladas

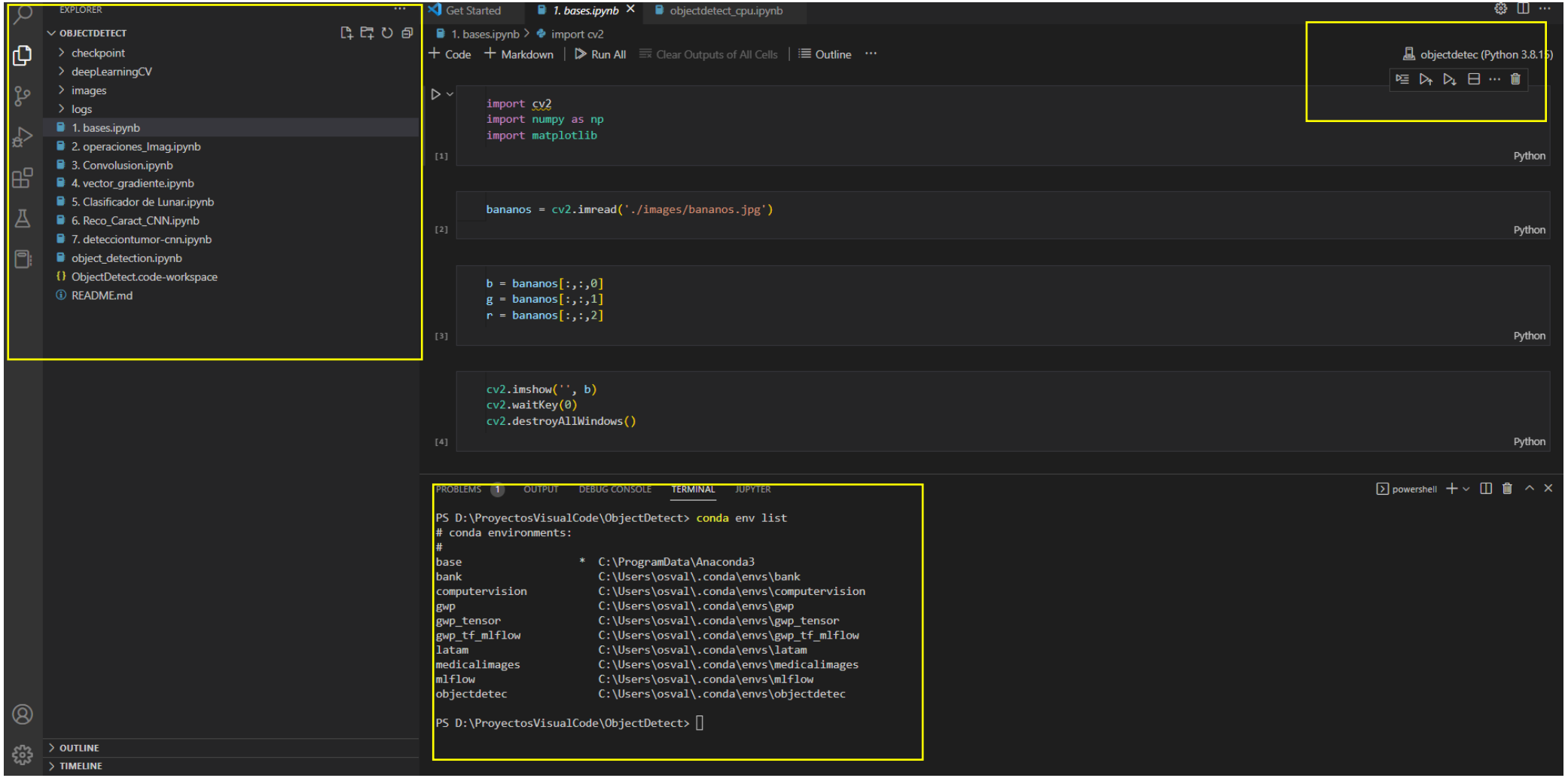
Environments

Programación II

- *conda create -n nombreenv python=3.x*
- *conda remove -n nombreenv –all*
- *conda env list*
- *conda activate nombreenv*
- *conda deactivate*

Environments (Problema común)

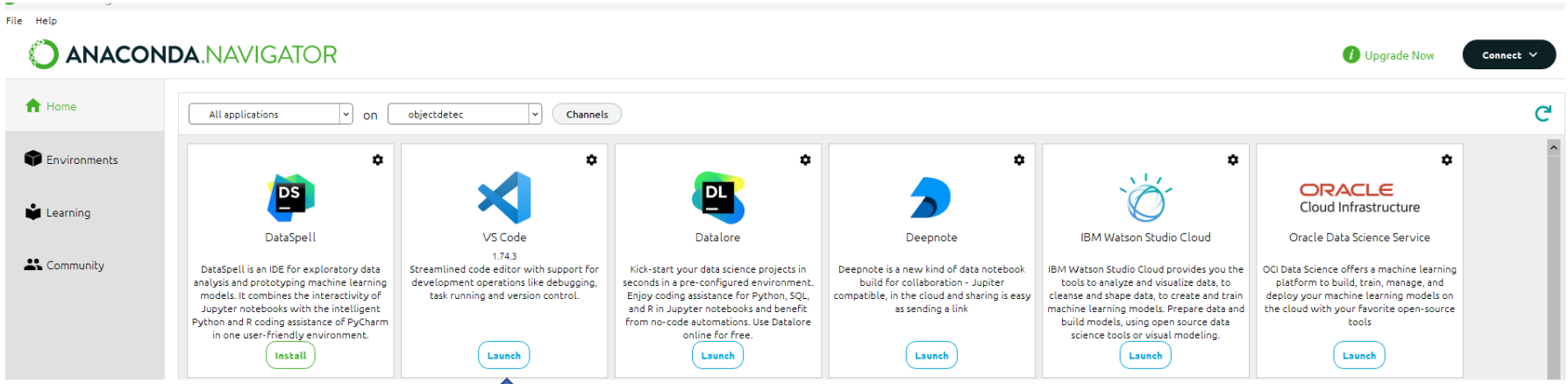
Programación II



Environments (Solución)

Programación II

1



2

3