

**Programas de Estudio del
Centro de Formación en Analítica Aplicada y
Reconversión Profesional**

CFAARP

Ciudad de México, México

Carlos Ernesto Martínez Rodríguez

7 de mayo de 2026

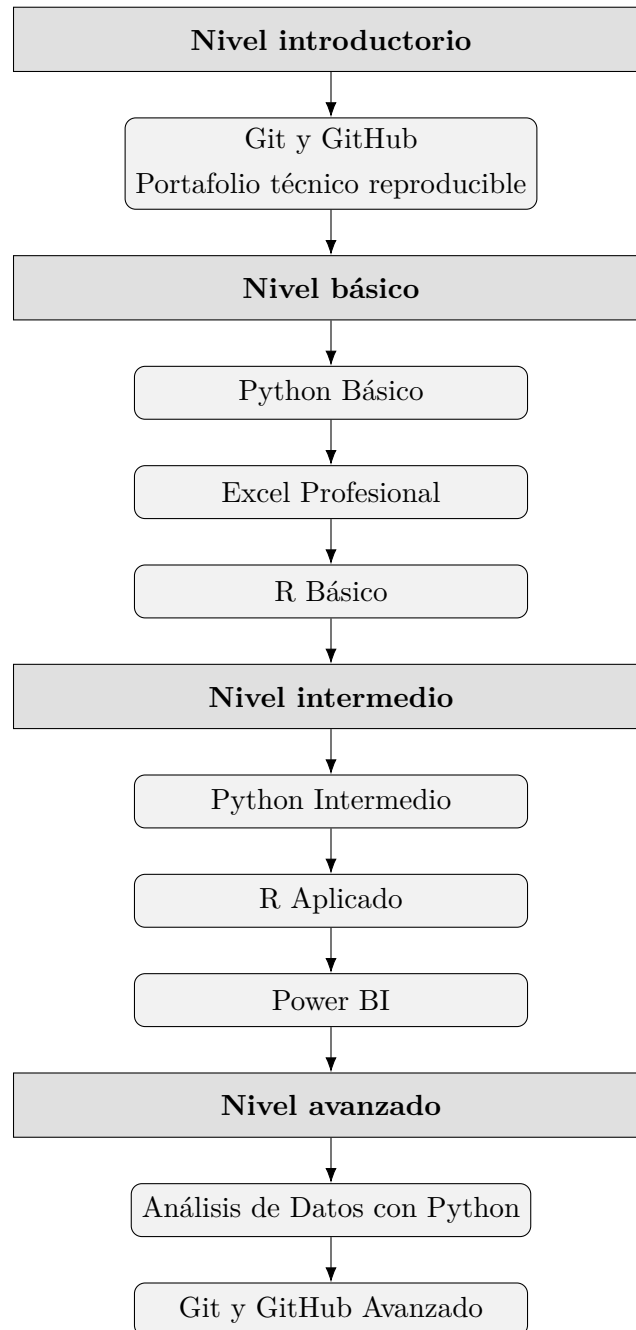
Presentación

El presente documento reúne los programas de estudio correspondientes a la oferta académica inicial del **Centro de Formación en Analítica Aplicada y Reconversión Profesional**, diseñada para facilitar procesos de transición laboral hacia áreas de programación y análisis de datos mediante un modelo educativo práctico, modular y progresivo.

Los cursos aquí descritos se estructuran con base en competencias técnicas aplicadas, desarrollo de portafolio profesional reproducible y uso de herramientas alineadas con el entorno laboral actual.

Este conjunto de programas constituye la base curricular del modelo académico institucional en su fase piloto y se organiza como una trayectoria progresiva de formación en analítica aplicada estructurada en niveles introductorio, básico e intermedio, orientados al desarrollo de competencias técnicas transferibles a contextos académicos y profesionales contemporáneos.

Mapa curricular del programa



El mapa curricular presenta la secuencia formativa sugerida del programa, organizada por niveles progresivos. El curso introductorio de Git y GitHub establece la base metodológica para la construcción de portafolio técnico reproducible, mientras que los niveles básico, intermedio y avanzado desarrollan competencias acumulativas en programación, análisis estadístico, visualización de datos y documentación profesional.

4. Nivel avanzado

4.1. Análisis de Datos con Python

Descripción general

Este curso introduce al estudiante en el uso de Python como herramienta profesional para la manipulación, análisis y visualización de datos mediante librerías especializadas del ecosistema científico.

Se desarrollan competencias en limpieza de datasets, transformación de variables, análisis exploratorio y generación de visualizaciones interpretables utilizando bibliotecas como `pandas`, `numpy`, `matplotlib` y `seaborn`.

El curso constituye la transición entre la programación general en Python y su aplicación directa en ciencia de datos.

Datos generales del curso

- Duración: 24 horas
- Modalidad: Taller práctico con proyecto integrador
- Nivel: Intermedio
- Requisito: Python Intermedio
- Entorno de trabajo: Python + VS Code o Jupyter Notebook
- Recomendado para: Estudiantes interesados en ciencia de datos y analítica computacional

Objetivo general

Desarrollar en el estudiante la capacidad de manipular, transformar y analizar datasets mediante herramientas del ecosistema científico de Python para generar análisis exploratorios reproducibles orientados a la interpretación de información cuantitativa.

Objetivos específicos

Al finalizar el curso, el estudiante será capaz de:

- Importar datasets utilizando `pandas`
- Manipular estructuras tipo `DataFrame`

- Limpiar datasets con valores faltantes
- Filtrar y seleccionar subconjuntos de datos
- Crear variables derivadas
- Integrar datasets múltiples
- Generar estadísticas descriptivas básicas
- Construir visualizaciones interpretables
- Documentar análisis mediante notebooks reproducibles

Palabras clave

Python, pandas, numpy, matplotlib, seaborn, análisis exploratorio, limpieza de datos, visualización estadística, DataFrame, notebooks reproducibles.

Contenido temático

Módulo 1: Introducción al análisis de datos con Python

- 1.1 Ecosistema científico de Python
- 1.2 Introducción a `numpy`
- 1.3 Introducción a `pandas`
- 1.4 Flujo de trabajo reproducible con notebooks

Módulo 2: Importación de datos

- 2.1 Lectura de archivos CSV con `read_csv()`
- 2.2 Lectura de archivos Excel con `read_excel()`
- 2.3 Exploración inicial del dataset
- 2.4 Tipos de variables en un DataFrame

Módulo 3: Limpieza de datos

- 3.1 Identificación de valores faltantes
- 3.2 Eliminación de registros incompletos
- 3.3 Imputación básica de datos
- 3.4 Conversión de tipos de variables

Módulo 4: Manipulación de DataFrames

- 4.1 Selección de columnas

- 4.2 Filtrado de registros
- 4.3 Ordenamiento de datos
- 4.4 Creación de variables derivadas
- 4.5 Uso de `groupby()`

Módulo 5: Integración de datasets

- 5.1 Concatenación de datasets
- 5.2 Unión mediante `merge()`
- 5.3 Integración por claves comunes
- 5.4 Validación de consistencia

Módulo 6: Estadística descriptiva

- 6.1 Medidas de tendencia central
- 6.2 Medidas de dispersión
- 6.3 Tablas resumen
- 6.4 Análisis por grupos

Módulo 7: Visualización de datos

- 7.1 Introducción a `matplotlib`
- 7.2 Introducción a `seaborn`
- 7.3 Histogramas
- 7.4 Diagramas de dispersión
- 7.5 Diagramas de caja
- 7.6 Visualizaciones comparativas por grupos

Módulo 8: Proyecto integrador

- 8.1 Importación de dataset real
- 8.2 Limpieza de datos
- 8.3 Transformación de variables
- 8.4 Integración de múltiples fuentes
- 8.5 Análisis exploratorio estructurado
- 8.6 Generación de visualizaciones interpretables

Proyecto integrador del curso

El estudiante desarrollará un análisis exploratorio reproducible utilizando un dataset real que incluya:

- Importación del dataset
- Limpieza y transformación de variables
- Integración de múltiples fuentes de datos
- Generación de estadísticas descriptivas
- Elaboración de visualizaciones interpretables
- Documentación mediante notebook reproducible

Este proyecto constituirá evidencia formal de competencias en análisis de datos con Python dentro del portafolio técnico del estudiante.

Producto final del curso

Al finalizar el curso, el estudiante contará con:

- Notebook reproducible de análisis exploratorio
- Dataset transformado
- Visualizaciones estadísticas interpretables
- Tablas resumen automatizadas
- Evidencia técnica para portafolio profesional

Criterios de evaluación

La evaluación del curso se realizará mediante el proyecto final conforme a la siguiente ponderación:

- Limpieza correcta del dataset (20 %)
- Transformación de variables (20 %)
- Integración de datasets (20 %)
- Visualizaciones interpretables (20 %)
- Organización reproducible del notebook (20 %)

La acreditación del curso requiere cumplir satisfactoriamente al menos el 80 % de los criterios establecidos.

Importancia dentro del programa formativo

Este curso constituye la base para cursos posteriores de futura oferta tales como:

- Métodos de Machine Learning con Python
- Estadística computacional