

**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

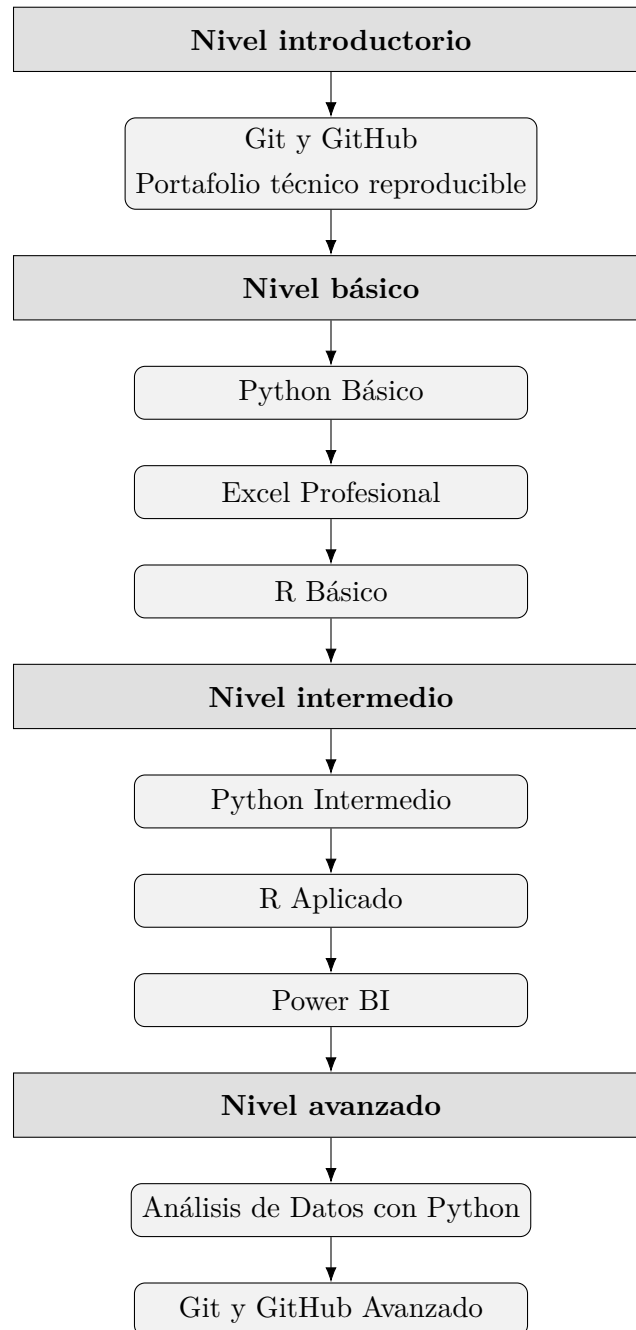
Presentación

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.

3.2. R Aplicado

Descripción general

Este curso desarrolla competencias intermedias en el uso del lenguaje R orientadas al análisis exploratorio de datos mediante herramientas modernas del ecosistema tidyverse.

Se enfatiza la manipulación eficiente de datasets, la transformación de variables, la visualización profesional con `ggplot2` y la generación de flujos de análisis reproducibles mediante scripts estructurados.

El curso constituye la transición entre la estadística descriptiva básica en R y el modelado estadístico aplicado dentro del programa de formación en analítica aplicada.

Datos generales del curso

- Duración: 24 horas
- Modalidad: Taller práctico con proyecto integrador
- Nivel: Intermedio
- Requisito: R Básico
- Entorno de trabajo: R y RStudio
- Recomendado para: Estudiantes interesados en análisis estadístico aplicado y ciencia de datos

Objetivo general

Desarrollar en el estudiante las competencias necesarias para manipular, transformar y visualizar datasets mediante herramientas modernas del ecosistema tidyverse, generando análisis exploratorios reproducibles orientados a la interpretación estadística de información cuantitativa.

Objetivos específicos

Al finalizar el curso, el estudiante será capaz de:

- Importar datasets utilizando librerías especializadas como `read_csv()`
- Manipular datos mediante `dplyr`
- Transformar variables con `mutate()`
- Filtrar observaciones con `filter()`

- Seleccionar variables con `select()`
- Integrar datasets mediante funciones `join()`
- Generar visualizaciones profesionales con `ggplot()`
- Organizar flujos de análisis reproducibles mediante pipelines
- Exportar resultados para documentación técnica

Palabras clave

R, tidyverse, dplyr, ggplot2, manipulación de datos, análisis exploratorio, visualización estadística, transformación de variables, scripts reproducibles, integración de datasets.

Contenido temático

Módulo 1: Introducción al ecosistema tidyverse

- 1.1 Instalación y carga de tidyverse
- 1.2 Estructura de flujo de trabajo con pipelines
- 1.3 Uso del operador pipe `%>%`
- 1.4 Organización de scripts reproducibles

Módulo 2: Manipulación de datos con dplyr

- 2.1 Selección de variables con `select()`
- 2.2 Filtrado de observaciones con `filter()`
- 2.3 Ordenamiento de datos con `arrange()`
- 2.4 Creación de variables con `mutate()`
- 2.5 Resumen de información con `summarise()`

Módulo 3: Agrupamiento y agregación

- 3.1 Uso de `group_by()`
- 3.2 Estadísticas por grupos con `summarise()`
- 3.3 Tablas resumen automatizadas
- 3.4 Interpretación de resultados agrupados

Módulo 4: Integración de datasets

- 4.1 Unión de datasets con `left_join()`
- 4.2 Unión de datasets con `inner_join()`

4.3 Unión de datasets con `right_join()`

4.4 Verificación de consistencia de datos

Módulo 5: Transformación avanzada de variables

5.1 Recodificación de variables

5.2 Transformaciones condicionales con `mutate()`

5.3 Manejo de valores faltantes

5.4 Creación de variables derivadas

Módulo 6: Visualización de datos con `ggplot2`

6.1 Gramática de gráficos con `ggplot()`

6.2 Gráficos de dispersión

6.3 Diagramas de caja

6.4 Histogramas

6.5 Visualizaciones por grupos

6.6 Personalización de gráficos

Módulo 7: Exportación de resultados

7.1 Exportación de tablas procesadas

7.2 Exportación de gráficos

7.3 Generación de reportes básicos

7.4 Organización de resultados reproducibles

Módulo 8: Proyecto integrador

8.1 Importación de dataset real con `read_csv()`

8.2 Limpieza y transformación de datos con `dplyr`

8.3 Integración de datasets mediante `join()`

8.4 Análisis exploratorio estructurado

8.5 Generación de visualizaciones interpretables con `ggplot()`

8.6 Elaboración de script reproducible

Proyecto integrador del curso

El proyecto integrador consistirá en el desarrollo de un análisis exploratorio reproducible de un dataset real utilizando herramientas del ecosistema tidyverse.

El estudiante deberá desarrollar:

- Importación del dataset
- Limpieza y transformación de variables
- Integración de al menos dos fuentes de datos
- Generación de estadísticas descriptivas por grupos
- Elaboración de al menos tres visualizaciones interpretables
- Documentación del análisis mediante script reproducible

Este proyecto constituirá evidencia formal de análisis estadístico aplicado dentro del portafolio técnico del estudiante.

Producto final del curso

Al finalizar el curso, el estudiante contará con:

- Script reproducible de análisis exploratorio en R
- Dataset transformado y estructurado
- Visualizaciones estadísticas interpretables
- Tablas resumen automatizadas
- Evidencia reproducible de análisis aplicado

Criterios de evaluación

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

- Manipulación correcta del dataset (20 %)
- Transformación adecuada de variables (20 %)
- Integración de datasets (20 %)
- Generación de visualizaciones interpretables (20 %)
- Organización y reproducibilidad del análisis (20 %)

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

Importancia dentro del programa formativo

Este curso fortalece las competencias necesarias para el desarrollo posterior de modelado estadístico y aprendizaje automático dentro del programa, especialmente:

- Métodos de Machine Learning con R
- Estadística computacional