

R-EVOLUTION: Data Science & Machine Learning Pro

Diseño Curricular (36 horas)

Programa profesional que integra estadística descriptiva e inferencial, ciencia de datos y machine learning aplicado, con un enfoque reproducible, comunicable y listo para producción.

CRONOGRAMA GENERAL

36 horas académicas — 12 sesiones de 3 horas (más Sesión 0 preparatoria)

DESARROLLO POR SESIONES

SESIÓN 0 — Instalación y Administración de Herramientas

Sesión preparatoria (no cuenta en las 36 horas)

	Tema	Entregable
1	Instalación de R, RStudio y gestión de paquetes	Entorno base funcional
2	Configuración de Quarto y Git básico	Proyectos reproducibles desde el inicio
3	Administración de entornos (renv) y verificación	Herramientas listas para el curso

SESIÓN 1 — Fundamentos del Entorno Data Scientist

Horas 1–3

	Tema	Entregable
1	Proyectos reproducibles y organización	Entorno organizado
2	Tipos de datos, vectores, tibbles y coerción segura	Scripts de manipulación básica
3	Introducción a R Markdown / Quarto	Reporte PDF/HTML inicial

SESIÓN 2 — Data Wrangling Profesional con dplyr

Horas 4–6

	Tema	Entregable
1	select(), filter(), arrange(), pipes	Pipeline limpio
2	mutate(), case_when(), across()	Variables derivadas
3	group_by(), summarise(), joins	KPIs consolidados

SESIÓN 3 — Importación y Limpieza de Datos Reales

Horas 7–9

	Tema	Entregable
1	CSV, Excel, APIs REST	Dataset integrado
2	NAs, outliers, validación	Dataset limpio
3	pivot_longer / pivot_wider	Datos en formato analítico

SESIÓN 4 — Visualización con ggplot2

Horas 10–12

	Tema	Entregable
1	Gramática de gráficos	Gráficos exploratorios
2	Facetado y comparaciones	Dashboards estáticos
3	Temas y estilo corporativo	Visual profesional

SESIÓN 5 — Análisis Exploratorio de Datos (EDA)

Horas 13–15

	Tema	Entregable
1	Estadísticos descriptivos avanzados	Reporte EDA
2	Correlaciones y relaciones	Mapas de calor
3	Anomalías y sesgos	Dataset validado

SESIÓN 6 — Inferencia Estadística I

Horas 16–18

	Tema	Entregable
1	Distribuciones y CLT	Simulación
2	Pruebas t y contrastes	Inferencia aplicada
3	Bootstrapping	Intervalos robustos

SESIÓN 7 — Inferencia Estadística II

Horas 19–21

	Tema	Entregable
1	ANOVA y supuestos	Comparaciones
2	Poder estadístico	Tamaño muestral
3	Pruebas no paramétricas	Análisis robusto

SESIÓN 8 — Regresión Lineal

Horas 22–24

	Tema	Entregable
1	Modelos lineales	Ajuste reproducible
2	Diagnóstico del modelo	Validación
3	Predicción e intervalos	Pronósticos

SESIÓN 9 — Machine Learning Supervisado I

Horas 25–27

	Tema	Entregable
1	Train/Test y workflows	Pipeline ML
2	Regresión logística y árboles	Modelos comparados
3	Validación cruzada	Métricas

SESIÓN 10 — Machine Learning Supervisado II

Horas 28–30

	Tema	Entregable
1	Random Forest	Importancia de variables
2	Gradient Boosting / XGBoost	Mejora de performance
3	Tuning de hiperparámetros	Modelo optimizado

SESIÓN 11 — Aprendizaje No Supervisado

Horas 31–33

	Tema	Entregable
1	Clustering k-means	Segmentación
2	PCA	Reducción dimensional
3	Clustering jerárquico	Dendrogramas

SESIÓN 12 — Proyecto Final y Comunicación

Horas 34–36

	Tema	Entregable
1	Presentación del proyecto	Defensa oral
2	Introducción a despliegue	Modelo operativo
3	Cierre y roadmap profesional	Certificación