

CARLOS MARIO HOYOS RÍOS

Ingeniero de Datos | Data Engineer

Cali, Colombia | carloshoyosrios@gmail.com | Portafolio: ingcarlosmariohoyos.github.io

linkedin.com/in/carlosmario-hoyosrios | github.com/ingcarlosmariohoyos

Disponibilidad: presencial, híbrido o remoto

PERFIL PROFESIONAL

Ingeniero de Datos e Ingeniero de Sistemas, Especialista en Visual Analytics y Big Data, con más de 7 años de experiencia en el sector bancario y financiero. Construyo pipelines ETL/ELT sobre Data Lake en AWS (S3, Athena) que generan información regulatoria y de negocio confiable para áreas como Riesgos y Captaciones. Experiencia con SQL, Python (Pandas, Polars), dbt, Dagster y Power BI (DAX), modelado dimensional (star schema) y arquitectura medallion. Formación en Apache Spark (PySpark) y conocimiento práctico de Microsoft Fabric. Enfocado en la calidad e integridad de los datos, la automatización y las soluciones escalables.

HABILIDADES TÉCNICAS

- **Lenguajes:** SQL (avanzado), PL/SQL, Python (Pandas, Polars, NumPy, SQLAlchemy), DAX
- **Ingeniería de datos:** ETL/ELT, Data Lake, Data Warehouse, modelado dimensional (star schema), arquitectura medallion (Bronze/Silver/Gold), calidad de datos
- **Cloud:** AWS (S3, Athena); Microsoft Fabric (pipelines, OneLake, KQL/Eventhouse, Real-Time Intelligence) – conocimiento práctico
- **Transformación y orquestación:** dbt, Dagster (Dagster Cloud), Prefect, Pentaho Data Integration
- **Procesamiento distribuido:** Apache Spark, PySpark, Spark SQL (nivel básico)
- **Bases de datos:** Oracle, SQL Server, PostgreSQL (Supabase), AWS Athena, IBM AS/400
- **BI y visualización:** Power BI (modelado de datos, DAX, dashboards), Excel avanzado
- **Herramientas:** Git, GitHub, Docker, formatos Parquet y CSV

EXPERIENCIA LABORAL

Desarrollador BI (Ingeniería de Datos) | Banco de Occidente Ene. 2026 – Actualidad

Cali, Colombia (remoto)

Proyecto RTILB (Riesgo de Tasa de Interés del Libro Bancario): informes regulatorios que se transmiten a la Superintendencia Financiera de Colombia, generados con modelos en Python sobre un Data Lake en AWS.

- Diseño y mantengo pipelines ETL/ELT entre AWS S3 y las capas de reporte para las áreas de Riesgos y Captaciones.
- Normalizo archivos Parquet: ajuste de esquemas inconsistentes, homologación de campos y eliminación de lógica hardcoded.
- Valido la consistencia e integridad de los datos procesados con consultas optimizadas en AWS Athena.
- Mantengo dashboards de Power BI alimentados por flujos de datos en AWS con información de rentabilidad.
- Atiendo requerimientos de analítica con cumplimiento de ANS.

Logros

- Migré procesos ETL de Pandas a Polars y reduje en un 20 % el tiempo de ejecución.
- Migré un tablero de Power BI de ambiente local a la nube, reconectando su origen de datos a la arquitectura en AWS.
- Estandaricé datasets heterogéneos para que los archivos Parquet cumplan los estándares de los modelos de analítica.

Ingeniero de Incidentes, Eventos y Problemas | Banco Unión Jun. 2025 – Nov. 2025

Cali, Colombia (híbrido)

- Analicé datos de incidentes con SQL para identificar patrones, tendencias y causas raíz (RCA).
- Diseñé e implementé pipelines ETL con Pentaho Data Integration para consolidar datos operativos de fuentes heterogéneas.
- Construí reportes y métricas de seguimiento de SLA con SQL avanzado, con 100 % de cumplimiento de los acuerdos de servicio.

Ingeniero Operador de Servicios TI | Banco Unión Ene. 2023 – Jun. 2025

Cali, Colombia

- Extraje y actualicé información financiera con SQL (Oracle) para el cierre bancario diario.

- Validé y concilié datos transaccionales en IBM AS/400, asegurando la integridad de los cierres operativos.
- Monitoreé servidores (Zabbix) y ejecuté procesos automatizados con GoAnywhere.

Logros

- Reduje en un 15 % el tiempo del cierre operativo diario rediseñando el plan de ejecución del proceso.
- Mejoré en un 20 % la eficiencia de consultas SQL críticas para la entrega de reportes.
- Aceleré en un 90 % el reprocesamiento de transacciones pendientes mediante automatización de consultas.

Técnico Mesa de Ayuda | Banco Unión

Ene. 2018 – Dic. 2022

Cali, Colombia

- Soporte a aplicativos bancarios y coordinación del equipo de soporte; creé una base de conocimiento que aumentó en un 30 % la eficiencia en requerimientos recurrentes.

Técnico Mesa de Ayuda | Banco W

Ene. 2015 – Dic. 2017

Cali, Colombia

- Configuración de equipos y conectividad de puntos de venta; diseñé un checklist que redujo en un 60 % los reprocesos.

PROYECTOS

Marketplace Data Warehouse – arquitectura medallion

github.com/ingcarlosmariohoyos/marketplace_dw_project | Python, dbt, Dagster Cloud, PostgreSQL (Supabase), Docker

- Pipeline ELT de punta a punta (Bronze → Silver → Gold) para datos de ventas, productos y clientes de un e-commerce.
- Modelado en dbt: staging y esquema en estrella (1 tabla de hechos y 5 dimensiones) con integridad referencial (PK/FK).
- Orquestación basada en assets con Dagster Cloud y schedule diario; despliegue 100 % en la nube y documentación técnica.

Análisis estadístico con Python y Power BI (dataset Pokémon Gen 1–6)

github.com/ingcarlosmariohoyos/Pokemon_analytics | Python (Pandas, NumPy, Matplotlib, Seaborn), Power BI, DAX

- Limpieza, EDA y segmentación de más de 750 registros, con tablero interactivo en Power BI y documentación del proceso.

EDUCACIÓN

- Especialización en Visual Analytics y Big Data – Fundación Universitaria Internacional de La Rioja (UNIR), 2026
- Ingeniería de Sistemas – Institución Universitaria Antonio José Camacho, 2022
- Tecnología en Sistemas – Centro Colombiano de Estudios Profesionales (CECEP), 2016

CERTIFICACIONES

- Big Data y Spark: Ingeniería de datos con Python y PySpark – Udemy, 2025
- AWS Cloud Quest: Cloud Practitioner (Training Badge) – Amazon Web Services, 2025
- Curso de Programación en Python – UNIR, 2025
- AI Fluency: Framework and Foundations – Anthropic, 2026
- Power BI TOTAL: Analista de Datos Avanzado – Udemy, 2024
- Ultimate GIT: gestiona tus proyectos y trabaja en equipo – Udemy, 2024
- Universidad Excel: Básico, Intermedio y Avanzado – Udemy, 2024
- Microsoft Fabric – laboratorios prácticos de Microsoft Learn (pipelines, Eventhouse, KQL, gobierno de datos)

IDIOMAS

- Español: nativo
- Inglés: B1, en formación activa (Centro Colombo Americano)