

# Gerardo Toboso

Especialista en Datos y Analítica

Buenos Aires, Argentina | gerardotoboso1909@gmail.com | 011 15-4045-6207  
linkedin.com/in/gerardo-toboso-512a48290 | github.com/Gerardo1909

## Resumen

---

Especialista en Datos y Analítica con experiencia práctica en el diseño de pipelines de datos automatizados, curación de datasets listos para análisis, y transformación de datos crudos en insights accionables para el negocio. Sólida formación en Python y SQL, con experiencia en soporte de entornos de BI, definición de métricas y construcción de dashboards para la toma de decisiones. Apasionado por la documentación, la calidad de datos y la comunicación clara con stakeholders técnicos y no técnicos. Actualmente cursando la Licenciatura en Ciencia de Datos y buscando crecer en un equipo de analítica colaborativo y orientado a la misión.

## Habilidades

---

**Analítica & BI:** SQL (PostgreSQL), Modelado de Datos (Star Schemas), Definición de Métricas, Análisis Exploratorio, Insights de Negocio, Validación de Datos

**Visualización de Datos:** Power BI (DAX), Tableau, Reportes y Dashboards para Stakeholders

**Ingeniería de Datos:** Python (pandas, numpy), Pipelines ETL, Limpieza de Datos, Parquet, Logging y Monitoreo

**Plataformas y Herramientas:** MongoDB Atlas, Supabase, Git & GitHub, GitHub Actions (CI/CD), Docker, Linux, Jupyter

**Adicional:** Pytest, Pydantic, Scikit-learn, Excel Avanzado (Tablas Dinámicas)

**Idiomas:** Español (Nativo) | Inglés (Fluido - C2)

## Experiencia

---

**Ingeniero de Datos**, NovaMart – ETL de Reportes E-commerce (Proyecto Personal) – [GitHub](#) Nov 2025 – presente

- Diseñé e implementé un pipeline de analítica automatizado para un dataset de e-commerce (10+ tablas relacionadas), reduciendo el tiempo de reportes manuales de 2 horas a menos de 3 minutos.
- Construí un pipeline ETL de extremo a extremo en Python para extraer, limpiar y transformar datos transaccionales en tablas listas para análisis.
- Definí y calculé métricas de negocio clave (mejores clientes, productos más vendidos, tendencias mensuales) para soportar la toma de decisiones basada en datos.
- Identifiqué y resolví problemas de calidad de datos, manejando ~15% de valores faltantes y eliminando ~3% de registros duplicados.
- Optimicé el almacenamiento de datos migrando outputs de CSV a Parquet, reduciendo el tamaño de archivos en 8x.

**Colaborador de Investigación**, MIT LIFT Lab & Proyecto Colaborativo Argentina 2025 – [GitHub](#) Sep 2025 – Dic 2025

- Recolecté y analicé datos de encuestas de pequeños negocios en comunidades locales de Argentina.
- Limpié y procesé datasets usando Python, aplicando técnicas de NLP, segmentación geográfica y análisis estadístico.
- Identifiqué 4 patrones de adopción digital basados en geografía y 2 tendencias estacionales que impactan directamente el rendimiento de ventas.
- Preparé reportes analíticos en inglés para audiencias del MIT, contribuyendo a recomendaciones que destacaron mejoras potenciales en ventas de hasta 18%.

## Educación

---

**Universidad Nacional de San Martín (UNSAM)**, Licenciatura en Ciencia de Datos

Jul 2022 – presente

- 75% completado — Promedio: 9.0 / 10