

Pedro Durán Matador

Data Analyst/Scientist



Acerca de mí

Tengo experiencia como Data Analyst: ETL, reporting, algoritmos de Machine Learning, y desarrollo web (FrontEnd)

Entusiasta de la tecnología y la programación.

Competencias Digitales

Lenguajes de Programación

Python, R, MySQL, Java, Kotlin.

Data Management

SQLite, MongoDB, MariaDB.

Business Intelligence

Tableau, Power BI, Trello, Qlik Sense.

Machine Learning

TensorFlow, scikit, pandas, openCV, Data Wrangler.

Google Cloud

Google Analytics , Big Query, Looker, Google Colab.

Habilidades

- Capacidad de análisis
- Habilidades comunicativas
- Trabajo en equipo
- Resolución de problemas
- Creatividad

Idiomas

Inglés	B1
Francés	A2

Formación Académica

Certificación Profesional (BIG DATA) The Valley Digital School Business

253 horas dic 2022 –feb 2023 Badajoz

Financial Markets Course

50 horas nov 2019 – dic 2019 Universidad de Yale (Online)

Grado | Comunicación Audiovisual

Sept 2016 – Sept 2023

Experiencia Laboral

Data Analyst/Scientist

CL GRUPO INDUSTRIAL

Julio 2023 / Diciembre 2023

Preprocesamiento y Modelado de Datos: EDA, ETL, Reporting, calidad del dato, Data Wrangler.

Business Analysis: Dashboard (KPIs), Qlik Sense, Power BI, Trello, Jira.

Administración de BBDD: MySQL, MongoDB, MariaDB.

Scraping: Selenium y BeautifulSoup, API.

Machine Learning: Proyectos en Python, R.
Visión computacional (OpenCV y TensorFlow),
OCR y Prophet.

Desarrollo de software/herramientas: Tkinter y Qt Editor.

Desarrollo de apps mobile: Android Studio, Kotlin y Java, Adalo (NoCode), Thunkable.

IA Generativa en Python: creación de prompts personalizados de análisis, comparativo, diseño web y creador de ejecutables.

GitHub: Documentación, control de versiones, bugs.

CONTACTO



618511172

pedrod985@gmail.com

Data Analysis:

- **EDA:** Creación y seguimiento de las diferentes fases del EDA, detección de outliers y storytelling.
- **Modelado de Datos:** usando Python + SQL. Querys en Java y Python.
- **ETL:** Limpieza de datos mediante librerías en Python: pandas, numpy, plotly, matplotlib, seaborn y herramientas data wrangler.
- **Data Visualization:** Reportes y creación de dashboard usando KPIs.
- **SQL:** Administración de Bases de Datos con SQL, querys para reporting usando el software Power BI (desktop y online) y la suite de Qlik (mashup & extension).
- **Microsoft Suite + Azure:** herramientas de Microsoft, generación de macros en Excel, creación de conexiones a través de Python y Java (JDBC) para administrar BBDD con SQL.
- **Servidor en Flask, Docker y Django:** Desarrollo de dashboard personalizable para web service y establecer conexiones con Azure o AWS.
- Experiencia con frameworks de frontend (React/Angular ,Bootstrap) y Javascript.
- **Ingeniería prompt:** uso de diversos modelos de IA para generar contenido Front y gestión backend en Python.

Lista de Proyectos: (Python, Java JDBC, SQL, Azure y ML)

1. Mejora de rendimiento y eficiencia de las consultas de SQL mediante Python + ML:

Modelo de Machine Learning entrenado para reducir la carga en la base de datos, optimizar el tiempo de ejecución y mejorar la recuperación de datos.
Limpieza de datos y Análisis de datos exploratorio para suprimir la carga inicial.

2. Comparar que las estructuras de las tablas en Salesforce y Azure SQL estén sync.

Este método es fundamental para garantizar la integridad y consistencia de datos entre ambos sistemas. Comparar dos listas y devolver las diferencias encontradas para solucionar el problema.

3. Crear clases en Java que permitan establecer conexiones con BBDD de SQL:

Clases en Java para conectar y desconectar automáticamente bases de datos, útil si se trabaja con SQL Server.

Usando la biblioteca HikariCP, gestionar conexiones a diferentes bases de datos Informix.

4. Automatización de flujos de trabajo en Airflow:

Ejecución de scripts de análisis de datos, generación de informes y visualización de datos de forma automatizada. Envío posterior vía email o cargarlos al servidor.

5. Construir Pipeline de datos (ETL):

Un flujo de trabajo de Airflow que extrae datos de diferentes fuentes, los transforma según las necesidades del proyecto y los carga en una base de datos

6. Integración con herramientas de visualización de datos:

Integración con herramientas de visualización de datos como Power BI, Tableau, etc.
Para crear informes interactivos y cuadro de mandos.
Alojamiento en web services y aplicaciones.

Data Science (Machine Learning y Visión Computacional):

- **Object Tracking (OpenCV):** Desarrollo de sistemas de seguimiento de objetos en vídeo utilizando la biblioteca OpenCV en Python.
- **Blurring PDF:** Aplicación en Python para desenfocar partes sensibles de documentos PDF, utilizando técnicas de procesamiento de imágenes.
- **Forecast de Prophet:** Implementación del modelo Prophet para pronósticos de series temporales, utilizando técnicas de regresión lineal en Python.
- **OCR (Tesseract):** Extracción de texto de imágenes utilizando una herramienta de reconocimiento óptico de caracteres.
- **Expresiones Regulares:** Análisis de documentos mediante expresiones regulares para identificar patrones específicos de texto.
- **Automatización de Procesos:** Desarrollo de sistemas para automatizar tareas repetitivas (UIPath), utilizando técnicas de scripting y programación. Usabilidad de APIs.

Desarrollo de Software y Aplicaciones Móviles:

- **Tkinter, Flask y Qt Editor:** Creación de interfaces gráficas de usuario utilizando las librerías en Python.
- **Diseño de Aplicaciones Móviles:** Experiencia en el desarrollo de aplicaciones móviles utilizando Android Studio, Java, Kotlin y Adalo.
- **Lenguaje Python:** Creación y gestión de entornos virtuales, documentación y optimización de código en Python.
- **Data Wrangler:** Limpieza y preparación de datos en archivos .xlsx y .csv utilizando Python y R para garantizar la calidad de los datos.
- **Diseño UI/UX, despliegue web:** HTML+CSS3+JS para diseño técnico y visual.

IA Generativa y Creación de Prompts:

- **Generación de imágenes y texto:** Desarrollo de sistemas de generación de imágenes y texto para satisfacer las necesidades corporativas utilizando modelos de IA generativa.
- **Creación de Prompts:** Elaboración y análisis de códigos y plantillas de emails utilizando técnicas de NLP.
- **Webs Generadas y comparativa de modelos:** Creación de páginas web generadas automáticamente utilizando modelos de IA, y evaluación comparativa de diferentes modelos para seleccionar el más adecuado.
- **Finalización Automática de Tareas:** Utilización de prompts para finalizar tareas en plataformas como Trello y GitHub de manera automática mediante sistemas IA.



Google Analytics

Power BI

