



ANÁLISIS BIG DATA PARA UNA PASTELERÍA DIGITAL

CESAR LEONEL ALIAGA BACARREZA

INTRODUCCIÓN

- “Bake IT!” es un sistema de e-commerce especializado en pastelería artesanal.
- Integra herramientas de Big Data para analizar ventas, clientes e inventarios.
- Busca optimizar procesos y anticipar la demanda.
- Se enmarca en el sector gastronómico digital.

IMPORTANCIA DEL BIG DATA

- Permite gestionar grandes volúmenes de información de ventas y usuarios.
- Facilita el análisis predictivo y la personalización del servicio.
- Reduce pérdidas y mejora la experiencia del cliente.
- Base para la toma de decisiones inteligentes en producción y marketing

JUSTIFICACIÓN

Técnico: arquitectura escalable (Python, PowerBI
PostgreSQL).

Social: mejora atención y personalización.

Económico: reduce desperdicio y aumenta rentabilidad.

5V del Big Data:

- Volumen → +30,000 registros
- Velocidad → datos en tiempo real
- Variedad → estructurados / no estructurados
- Veracidad → control y limpieza
- Valor → decisiones estratégicas

PLANTEAMIENTO DEL PROBLEMA

- Las pastelerías no aprovechan los datos disponibles.
- No existen sistemas que analicen ventas, gustos y producción.
- Se desperdician recursos y se toman decisiones empíricas.
- “Bake IT!” integrará análisis Big Data para convertir datos en conocimiento.

OBJETIVOS

Objetivo General:

Desarrollar un sistema e-commerce inteligente con análisis Big Data para optimizar ventas, inventario y producción.

Objetivos Específicos:

1. Diseñar arquitectura de datos escalable.
2. Aplicar análisis descriptivo y predictivo.
3. Crear dashboards interactivos.
4. Evaluar modelos de IA para predicción.
5. Garantizar calidad y seguridad de datos.

FUENTES DE DATOS

- Datos simulados de usuarios, ventas, productos e inventarios.
- Generados con Python (Faker, Pandas) o scripts SQL.
- Formato: CSV y JSON, con más de 30,000 registros.
- Fases: diseño → generación → validación → análisis.
- Consideraciones éticas: anonimización y uso de datos no reales.

CRONOGRAMA DE DESARROLLO Y ANÁLISIS DEL PROYECTO

Fase	Descripción	Duración Estimada
Fase 1: Diseño del modelo de datos	Definición de entidades, relaciones y estructura de la base de datos simulada.	Semanas 1-2
Fase 2: Generación y simulación de datos	Creación de registros con herramientas como Faker (Python) o scripts SQL.	Semanas 3-4
Fase 3: Integración y limpieza	Validación de coherencia, normalización y eliminación de duplicados.	Semanas 4-5
Fase 4: Análisis y visualización	Aplicación de técnicas descriptivas y predictivas, dashboards y visualizaciones.	Semanas 5-6

**MUCHAS
GRACIAS**

