

UNIVERSIDAD PRIVADA FRANZ TAMAYO
FACULTAD DE INGENIERÍA DE SISTEMAS
CARRERA DE INGENIERÍA DE SISTEMAS



**Primera Entrega del Proyecto Final: Documento de
Fundamentos y Plan de Datos**

Estudiante: Cesar Aliaga Bacarreza - SIS7060055

Docente: Ing. Erique Alejandro Laurel Cossio

Asignatura: Big Data (BDA-911)

LA PAZ – BOLIVIA
2025

1. Introducción

En la actualidad, la digitalización de los procesos comerciales ha transformado de manera profunda la forma en que las empresas del sector gastronómico interactúan con sus clientes. En este contexto, el proyecto “Bake IT!” surge como una plataforma de comercio electrónico especializada en la venta y personalización de productos de pastelería, que busca integrar no solo la gestión de ventas y usuarios, sino también un sistema inteligente de análisis de datos para optimizar la toma de decisiones empresariales.

El sector gastronómico, y particularmente el de la pastelería artesanal, enfrenta retos significativos relacionados con la predicción de la demanda, la personalización de la experiencia del cliente y la optimización del inventario. La variabilidad en los pedidos, los cambios estacionales y la influencia de factores externos como fechas festivas o tendencias en redes sociales generan grandes volúmenes de información que, si se gestionan adecuadamente, pueden convertirse en una ventaja competitiva.

En este sentido, el análisis de datos y el uso de tecnologías Big Data representan una herramienta esencial para comprender el comportamiento del consumidor, identificar patrones de compra y anticipar la demanda de productos. El aprovechamiento de estas capacidades permite a las empresas del sector no solo mejorar su eficiencia operativa, sino también fortalecer la relación con sus clientes mediante recomendaciones personalizadas, promociones dinámicas y un control preciso de la producción.

“Bake IT!” se plantea así como un proyecto integral que combina el desarrollo de un sistema web con una arquitectura orientada al análisis masivo de datos, incorporando módulos de registro de ventas, auditoría de accesos, control de inventarios y análisis de comportamiento de usuarios. La inclusión de técnicas de Big Data en este entorno busca sentar las bases para futuras implementaciones de modelos predictivos que permitan anticipar tendencias de consumo, reducir desperdicios y optimizar los recursos de la empresa.

De esta manera, el presente proyecto se enmarca dentro del campo de la analítica de datos aplicada al comercio electrónico, contribuyendo al desarrollo tecnológico del sector gastronómico y ofreciendo un enfoque innovador que integra la gestión operativa con la inteligencia de negocios.

2. Justificación

El desarrollo del sistema “Bake IT!” responde a la necesidad de modernizar y optimizar los procesos de gestión en el sector de la pastelería, un ámbito que, pese a su dinamismo y creatividad, aún depende en gran medida de decisiones empíricas y procesos manuales. Abordar este desafío desde una perspectiva de Big Data permite transformar la forma en que la información es recolectada, almacenada, procesada y analizada, generando conocimiento útil para mejorar la eficiencia operativa, la calidad del servicio y la rentabilidad del negocio.

La importancia de aplicar técnicas de Big Data en este proyecto radica en la capacidad de manejar grandes volúmenes de datos provenientes de diversas fuentes, tales como registros de ventas, preferencias de clientes, reseñas, patrones de navegación, inventarios y métricas de desempeño de los empleados. Este enfoque posibilita no solo una visión integral del funcionamiento de la empresa, sino también la detección de tendencias de consumo, la predicción de la demanda y la identificación de oportunidades de mejora.

Desde una perspectiva técnica, el proyecto promueve la implementación de arquitecturas escalables y seguras para la gestión de datos masivos, combinando tecnologías modernas como Node.js, PostgreSQL, React y Tailwind CSS, con prácticas de análisis descriptivo y predictivo. En el plano social, impulsa la personalización del servicio y la satisfacción del cliente mediante recomendaciones basadas en datos y experiencias de compra más ágiles. A nivel económico, la capacidad de prever la demanda y optimizar los recursos reduce los costos de producción, evita el desperdicio de insumos y aumenta la rentabilidad.

3. Planteamiento del Problema

En el ámbito de la pastelería moderna, la gestión eficiente de los procesos de producción, inventario y ventas se ha convertido en un desafío constante debido a la alta variabilidad de la demanda y a la diversificación de los productos personalizados. Las pastelerías tradicionales suelen enfrentar dificultades para anticipar qué tipos de productos serán más solicitados en determinadas épocas, qué cantidades producir para evitar pérdidas o cómo ajustar sus estrategias de marketing en función del comportamiento del consumidor. Estas limitaciones derivan, en gran medida, de la ausencia de sistemas de información integrados y del uso insuficiente de técnicas analíticas basadas en datos.

En este contexto, el proyecto “Bake IT!” busca resolver una problemática común en las micro y pequeñas empresas del sector gastronómico: la falta de un sistema centralizado que unifique los datos operativos y permita analizarlos de manera

inteligente. Actualmente, muchos negocios registran sus ventas, pedidos e inventarios de forma manual o en plataformas aisladas, lo que impide generar una visión integral del negocio. Esta desconexión dificulta la toma de decisiones estratégicas, como la planificación de la producción, la gestión del stock o la definición de campañas de marketing personalizadas.

La ausencia de un enfoque analítico impide también comprender el comportamiento de los clientes. Sin herramientas de análisis, resulta complejo identificar patrones de consumo, detectar los productos más rentables o estimar la demanda futura. Esta carencia de conocimiento lleva a decisiones basadas en la intuición, lo que se traduce en ineficiencias operativas, exceso o escasez de inventario y pérdida de oportunidades de venta.

El problema central que se aborda en este proyecto es, por tanto, la necesidad de integrar un sistema de gestión de pastelería con capacidades de análisis Big Data que permita recopilar, procesar y transformar grandes volúmenes de información en conocimiento útil para la toma de decisiones. Se busca que la plataforma “Bake IT!” funcione no solo como un punto de venta en línea, sino también como un sistema inteligente capaz de analizar indicadores clave, generar reportes visuales y aplicar modelos predictivos de demanda y comportamiento de usuarios.

Comprender y aprovechar los datos generados dentro del sistema tiene un valor estratégico significativo. A través del análisis masivo y automatizado, la empresa podrá responder preguntas como:

- ¿Cuáles son los productos más vendidos por temporada?
- ¿Qué factores influyen en el aumento o disminución de la demanda?
- ¿Qué tipo de promociones generan mayor retorno de inversión?
- ¿Qué patrones de comportamiento muestran los clientes recurrentes frente a los nuevos?

Responder estas preguntas permitirá transformar los datos en decisiones, mejorando la planificación de recursos, el control de inventarios y la eficiencia de las operaciones diarias. En consecuencia, el proyecto “Bake IT!” no solo busca desarrollar una plataforma tecnológica, sino también implementar un modelo de gestión basado en el conocimiento y la analítica de datos, fortaleciendo la competitividad y sostenibilidad de las pastelerías digitales.

4. Objetivos de Investigación

4.1 Objetivo General

Desarrollar e implementar un sistema de comercio electrónico inteligente denominado “Bake IT!” orientado al sector de pastelería, que integre herramientas de análisis de datos y técnicas de Big Data para optimizar la toma de decisiones en procesos de venta, producción, inventario y personalización de productos, contribuyendo así a la eficiencia operativa y al fortalecimiento de la competitividad empresarial.

4.2 Objetivos Específicos

1. Diseñar e implementar una arquitectura de datos que permita recopilar, almacenar y procesar grandes volúmenes de información provenientes de las operaciones del sistema (ventas, usuarios, inventarios, accesos, reseñas, etc.).
2. Aplicar técnicas de análisis descriptivo y predictivo sobre los datos recolectados, con el fin de identificar patrones de consumo, tendencias de demanda y oportunidades de mejora en la gestión del negocio.
3. Desarrollar paneles de visualización interactiva (dashboards) que presenten indicadores clave (KPI) sobre ventas, comportamiento de clientes, desempeño de productos y eficiencia del inventario.
4. Evaluar la viabilidad técnica y analítica de incorporar modelos de inteligencia artificial y aprendizaje automático (Machine Learning) para la predicción de la demanda y la recomendación de productos personalizados.
5. Garantizar la calidad, seguridad y ética en el manejo de datos, aplicando políticas de validación, anonimización y auditoría que aseguren la veracidad y confiabilidad de la información analizada.

5. Fuentes de Datos

El proyecto “Bake IT!” Se sustenta en la generación, procesamiento y análisis de datos simulados que representan el funcionamiento real de un sistema de comercio electrónico especializado en pastelería. Dado que el sistema se encuentra en etapa de desarrollo y validación, se optará por simular los conjuntos de datos para emular el comportamiento de los usuarios, las ventas, los productos y la gestión de inventarios, garantizando así un entorno controlado para la aplicación de técnicas de Big Data.

5.1 Descripción de las Bases de Datos Simuladas

Las bases de datos se diseñarán con una estructura coherente con la arquitectura del sistema “Bake IT!”, basada en PostgreSQL y alimentada mediante scripts o generadores automáticos. Los conjuntos principales incluirán:

- **Usuarios:** datos simulados de clientes, emprendedores y administradores, con atributos como edad, ubicación, historial de compras y nivel de satisfacción.
- **Productos:** catálogo de pasteles y postres con variables ficticias (nombre, ingredientes, categoría, precio, tiempo de preparación, nivel de demanda).
- **Ventas y pedidos:** registros de transacciones generadas de manera aleatoria, con fechas, montos, estados de entrega y formas de pago realistas.
- **Inventarios:** movimientos simulados de insumos (harina, azúcar, frutas, coberturas, etc.) que reflejan la dinámica de reposición y consumo.
- **Logs y auditorías:** simulación de registros de accesos, actividad de usuarios y errores del sistema para el análisis de rendimiento.
- **Reseñas y calificaciones:** comentarios generados automáticamente con lenguaje natural básico, para incorporar análisis de sentimientos.

5.2 Volumen y Estructura de los Datos

El conjunto total de datos simulados estará compuesto por al menos 30,000 registros distribuidos entre las tablas principales (usuarios, ventas, productos e inventario).

Los datos serán generados mediante técnicas de simulación controlada usando herramientas como Python (pandas, faker, numpy) o scripts de carga en SQL, con una estructura tabular y semiestructurada (CSV y JSON).

El diseño del dataset buscará reflejar comportamientos realistas, tales como picos de demanda en fechas festivas, productos más vendidos por temporada o usuarios frecuentes con preferencias personalizadas.

5.3 Plan de Generación de Datos Simulados

La adquisición de datos se realizará en tres fases planificadas:

1. **Fase 1 (Semanas 1–2):** definición del modelo de datos y atributos principales de cada entidad.
2. **Fase 2 (Semanas 3–5):** generación de los registros simulados usando *faker* y carga en la base de datos.
3. **Fase 3 (Semanas 6–8):** validación de coherencia, limpieza y preparación de los datasets para análisis exploratorio y visualización.

Además, se contempla ampliar el dataset con datos externos públicos de consumo alimentario y tendencias gastronómicas para complementar el análisis predictivo.

5.4 Consideraciones Éticas y Legales

Al tratarse de datos simulados, no se maneja información personal real ni sensible. No obstante, se mantendrán los principios de ética, seguridad y trazabilidad propios de un sistema Big Data, asegurando que los métodos de generación y análisis sean reproducibles y transparentes. En caso de integrar fuentes públicas, se respetarán los términos de uso y licencias abiertas de los datasets.

5.5 Relación con las 5V del Big Data

- **Volumen:** más de 30,000 registros simulados, escalables en pruebas.
- **Velocidad:** generación automatizada y procesamiento inmediato.
- **Variedad:** datos estructurados (ventas, inventarios) y no estructurados (reseñas, logs).
- **Veracidad:** validación estadística de la coherencia de los datos generados.
- **Valor:** los datos simulados permitirán probar técnicas de análisis predictivo y visualizar escenarios reales de decisión empresarial.