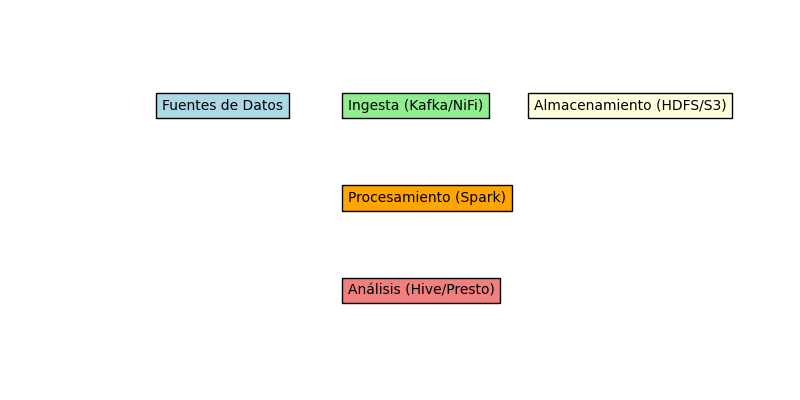
Propuesta de Solución Big Data para MarketTrend S.A.

# 1. Análisis de las 5V’s de Big Data

- Volumen: MarketTrend S.A. recopila grandes cantidades de datos desde múltiples fuentes como redes sociales, sensores, logs de aplicaciones y datos transaccionales.  
- Velocidad: La información llega en tiempo real o casi real, lo que requiere procesamiento inmediato para obtener valor.  
- Variedad: Los datos provienen de fuentes estructuradas (bases de datos), semi-estructuradas (logs) y no estructuradas (texto, imágenes).  
- Veracidad: Es fundamental validar la calidad y confiabilidad de los datos para evitar decisiones erróneas.  
- Valor: El objetivo es extraer conocimiento útil para la toma de decisiones estratégicas.

# 2. Arquitectura Big Data Propuesta

- Adquisición de datos: Se utilizarán herramientas como Apache NiFi o Kafka para la ingesta de datos en tiempo real.  
- Almacenamiento distribuido: Se propone el uso de Hadoop HDFS o Amazon S3 para almacenar grandes volúmenes de datos.  
- Procesamiento distribuido: Apache Spark será la herramienta principal para el procesamiento por su velocidad y escalabilidad.  
- Análisis: Se utilizarán herramientas como Apache Hive, Presto o notebooks con PySpark para análisis exploratorio y modelado.



# 3. Beneficios de la Solución

- Mejora en la toma de decisiones basada en datos en tiempo real.  
- Mayor eficiencia operativa al automatizar la recolección y análisis de datos.  
- Escalabilidad para adaptarse al crecimiento del volumen y variedad de datos.

# 4. Riesgos y Medidas de Seguridad

- Riesgos: Complejidad en la integración de tecnologías, costos de infraestructura, y problemas de calidad de datos.  
- Medidas: Implementar controles de calidad de datos, cifrado en tránsito y en reposo, autenticación y autorización robusta, y monitoreo continuo.