Capstone Project:
Mobile Companies Presence in YouTube

Eric Solis Montufar

Business Case

- In this project, we identify the mobile companies that are reachig more audience in YouTube videos.

- Tech influencers review cellphones and tech products, however, is not easy to assess the reach that a particular company has through this media content.

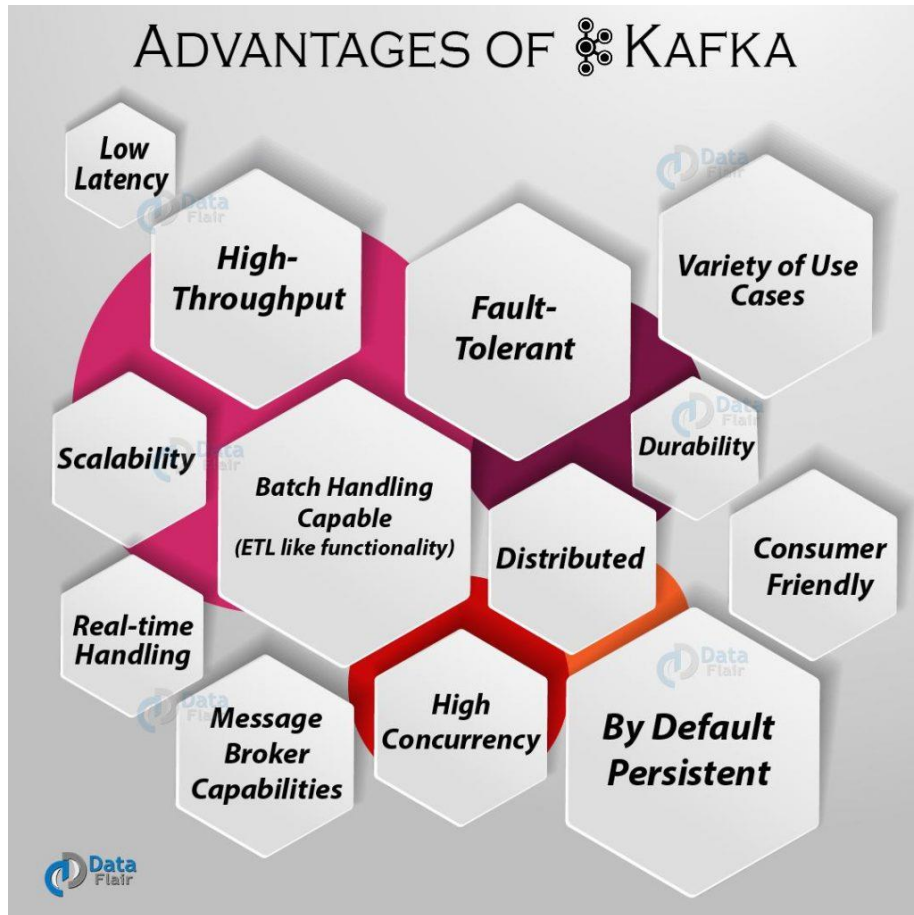# Tools

- YouTube API
- Apache Kafka
- Apache Spark
- Python
- MongoDb
- MySQL
- Plotly Falcon

Integrate → Prepare/Curate → Explore/Refine → Analyze → Visualize

Why kafka?

ADVANTAGES OF Kafka

Low Latency

High-Throughput

Fault-Tolerant

Variety of Use Cases

Scalability

Batch Handling Capable (ETL like functionality)

Durability

Distributed

Consumer Friendly

Real-time Handling

Message Broker Capabilities

High Concurrency

By Default Persistent

Why Spark?

APACHE Spark™

Pros
- Speed
- Ease of Use
- Advanced Analytics
- Dynamic in Nature
- Multilingual
- Apache Spark is powerful
- Increased access to Big data
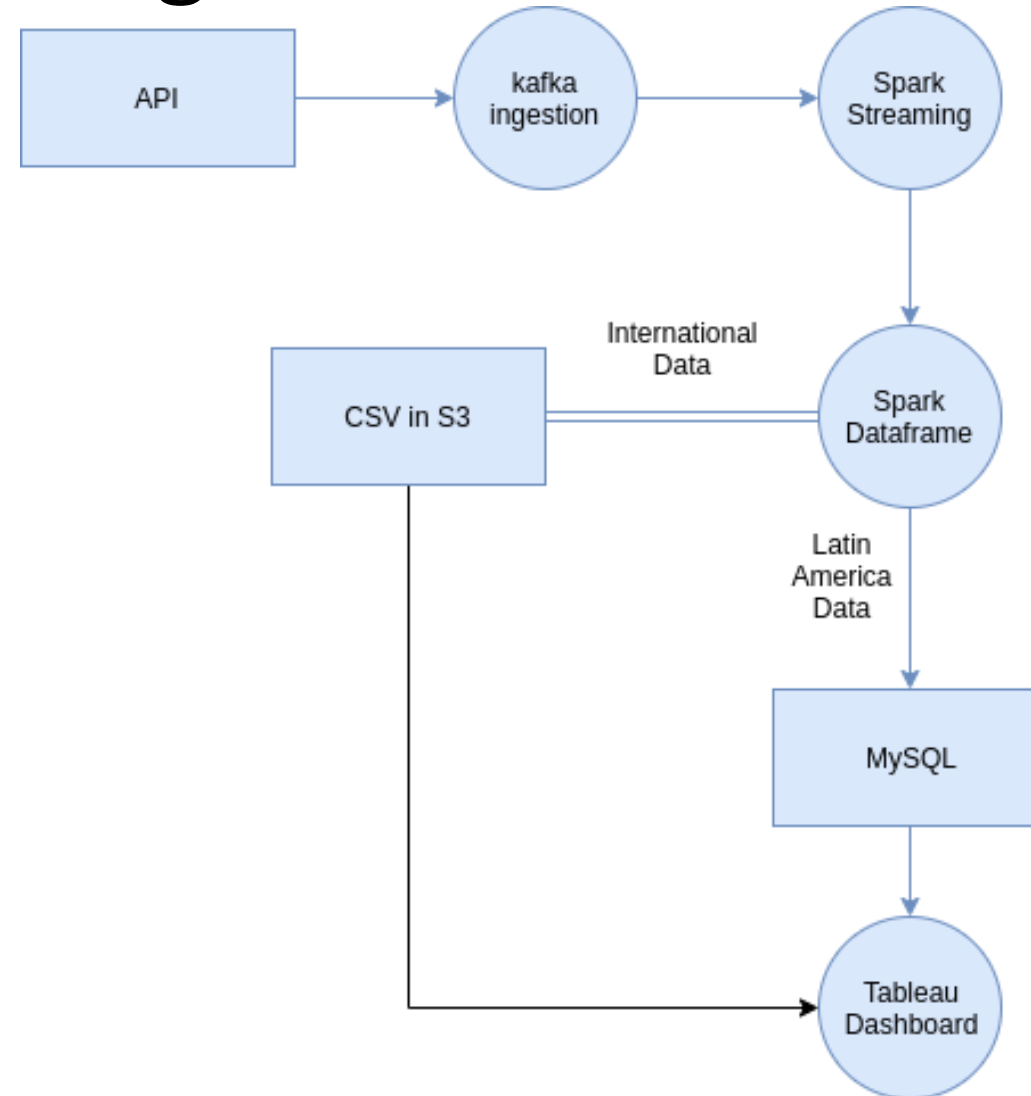- Demand for Spark Developers

Why MySQL?

- The application requires many transactions

- Easy to query

- Easy to connect with BI Tools

Data Flow Diagram

{title: Oneplus,   views: 47k   likes: 1.5k   comments: 287}