



SANTANDER PRODUCT RECOMMENDATION

Team 7
Arpit Rawat
Vaishali Lambe
Nishant Gandhi



Use cases

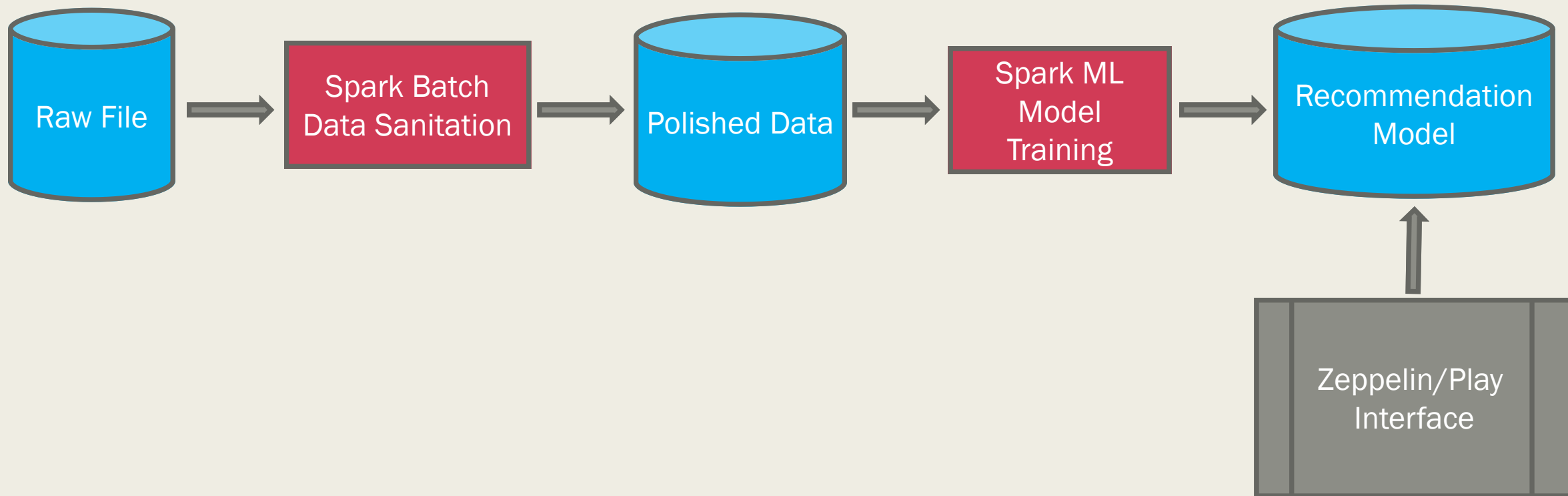
1. Actor – Customer

When customer is a new customer [customer for less than 6 months], system will recommend specific banking products to the new customer

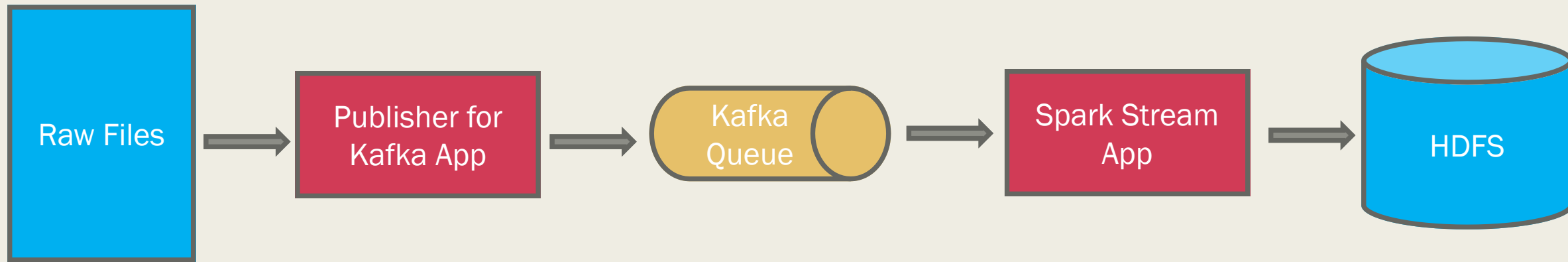
2. Actor - Customer

When customer is an existing customer [customer for more than 6 months], system will recommend the relevant products on the basis of previous months product consumption of a customer

Execution Pipeline



ETL Pipeline(Additional)



Data Source: <https://www.kaggle.com/c/santander-product-recommendation/data>

Data Size: ~2.3GB [~930K rows]

Milestones/Sprints:

Sprint#	Timeline	Tasks
Sprint 1	03/19/2018 - 03/25/2018	Environment Set Up, Data Cleaning, Data Visualization, Unit Test Cases, Testing
Sprint 2	03/26/2018 - 04/01/2018	ML Spark, Data Modelling, Integration with pipeline, Testing
Sprint 3	04/02/2018 - 04/08/2018	Model fitting and cross validation, Testing, Accuracy, Optimization, UI framework integration
Sprint 4	04/09/2018 - 04/15/2018	Re-Testing and finishing, Preparation of Final Presentation

What you program in Scala:

- Data Extraction, Data Cleaning [IntelliJ IDEA]
- Data Visualization[Zeppelin]
- Data Modelling [Spark]
- User Interface [Play Framework]

Code Repository Location:

https://github.com/vaishalilambe/Team7_Santander_Product_Recommendation

https://gitlab.com/nishantgandhi99/Team_7_Santander_Product_Recommendation

Acceptance Criteria:

- Predictive model accuracy will be approx. 60-70%
- UI efficiency to get any recommendation of a banking product for pre-defined customer will be approx. 5-10sec [It means almost 90% customers will receive UI response on recommended products approx. in 5-10 seconds time]

Goals of the Project:

- Learn and enhance skills on Scala programming, Apache Spark, Data Analysis and Apache Kafka [Kafka is additional]
- Achieve the stated acceptance criteria [refer prior slide# 7]
- Meet deadlines of each sprint and activities associated with it
- Collaborative learning and knowledge sharing, to have good team work
- Sharpen presentation skills in a group