

# **Inbank Product Data Analyst Internship Tasks – Elchin Huseynov**

## **Task 1: Product Metrics and Critical Events for Inbank's Consumer Loan Online Application Flow**

### **1. Key Product Metrics**

To optimize Inbank's consumer loan application process, the following metrics should be monitored:

#### **Application Funnel Metrics**

- **Application Start Rate** – Percentage of users initiating the loan application.
- **Completion Rate** – Percentage of users who successfully submit the application.
- **Drop-off Rate** – Percentage of users abandoning the process at each stage.
- **Time Spent per Step** – Average duration users spend on each application step.

#### **User Experience Metrics**

- **Form Error Rate** – Frequency of validation errors encountered by users.
- **Field Correction Rate** – Instances where users modify fields before submission.
- **Support Request Rate** – Number of users seeking assistance during the application.

#### **Loan Approval Metrics**

- **Approval Rate** – Percentage of applications approved.
- **Rejection Reasons** – Analysis of common denial reasons.
- **Processing Time** – Average time taken to process applications.

### Financial Metrics

- **Average Loan Amount** – Mean value of loans applied for.
- **Default Rate** – Percentage of approved loans that default.
- **Disbursement Time** – Time from approval to fund disbursement.

### Fraud & Risk Metrics

- **Fraud Detection Rate** – Percentage of applications flagged for potential fraud.
- **Repeat Application Rate** – Users reapplying after rejection.
- **Document Verification Failures** – Rate of unsuccessful document verifications.

### Data Quality Metrics

To ensure reliable decision-making and reporting, Inbank should also monitor data quality KPIs:

- **Completeness** – Percentage of records with all required fields (e.g., transaction date, currency ID).
- **Accuracy** – Data consistency between internal systems and external sources (e.g., exchange rates).
- **Validity** – Percentage of fields matching expected formats or lookup values.
- **Timeliness** – Time taken for data to arrive and be processed (e.g., exchange rates delivered within SLA).

## 2. Critical Events to Track

Tracking the following user interactions and system events is essential for understanding the customer journey and system reliability:

Event Name	Description
Application Initiated	User starts the loan application.
Personal Information Submitted	User provides personal details.
Financial Information Submitted	User submits financial data.
Document Upload	User uploads necessary documents.
Application Submitted	Completion and submission of the application.
Application Approved	Loan application approved.
Application Rejected	Loan application denied.
Funds Disbursed	Loan amount transferred to the user.

### 3. Recommended Tools for Tracking and Analysis

#### Data Orchestration & Processing

- **Apache Airflow** – Used for workflow orchestration, ensuring batch processes (e.g., exchange rate ingestion) run on time.
- **Snowflake** – Scalable cloud data warehouse supporting both storage and in-database processing logic.

#### Data Integration

- **Pentaho** – Used for ETL pipelines, being phased out in favor of Snowflake stored procedures.

#### Data Querying

- **SQL** – Primary language for analyzing data in Snowflake and building reports.

#### Data Visualization & Reporting

- **Microsoft Power BI** – Used for real-time dashboards, funnel analysis, fraud trends, and data quality monitoring.

## 4. Batch Processing and Data Flow Pipelines

Data used in decision-making (e.g., currency rates, blacklist checks, loan statuses) is processed through **scheduled batch pipelines**, often using Airflow. These pipelines:

- Pull from multiple data sources (e.g., banking systems, third-party risk platforms).
- Run ETL or ELT jobs using Snowflake or legacy Pentaho.
- Deliver cleansed, enriched data for reporting and analytics.

Monitoring batch pipeline SLAs and failures is critical for ensuring loan decisions are based on complete and timely data.

## 5. Data Quality Use Case: Currency Exchange Monitoring

Inbank supports loan applications in multiple currencies. Accurate and timely exchange rate data is critical for correct repayment calculations, compliance, and customer trust.

### Challenges

- Rates arrive from different national systems with inconsistent formats and delivery times.
- Data may be missing, delayed, or outdated.
- There is risk of applying incorrect rates if data is not validated.

### Proposed Controls

- Define **critical data elements (CDEs)** like `currency_id`, `exchange_date`, and `exchange_rate_to_eur`.
- Implement **automated checks** for:
  - Missing exchange rates (nulls)
  - Outliers or anomalies in rate fluctuations
  - Timeliness (rates not loaded on time)

## Monitoring Tools

- Use **Snowflake Tasks & Streams** for real-time validation and logging.
- Trigger alerts and pipeline failures through **Airflow** if issues arise.
- Visualize quality issues in **Power BI** for transparency.

## Business Value

- Prevents errors in customer offers and repayments
- Builds trust in international loan offerings
- Reduces manual checks and data disputes
- Enables faster delivery of high-quality financial products

## Task 2: Write a query to return the amounts in euros aggregated by transaction\_date

Please check Elchin\_Inbank\_Product\_Data\_Analyst\_Internship\_Task2.sql file and Readme.md file.