Optimizing Customer Satisfaction: A Deep Dive into Spar Nord Bank ATM Transactions Data Analysis Project Plan

Version 3.0

Revision History

Date	Version	Description	Author
02/19/2023	1.0	Initial Project Plan, Abstract, Deliverables	Rashmi, Bhavana, Jay, Sawan
03/26/2023	2.0	Technology stack, tools to be used, Data cleaning and pipeline designing, AWS pipeline creation	Rashmi, Bhavana, Jay, Sawan
05/04/2023	3.0	Documentation, report and visualizations	Rashmi, Bhavana, Jay, Sawan

Table of Contents

1. Introduction

1.1 Purpose of this document

The purpose of this document is to provide a detailed project description of the application called Optimizing Customer Satisfaction: A Deep Dive into Spar Nord Bank ATM Transactions Data Analysis, which is designed to provide a more efficient, secure, and customer-oriented banking experience. This document includes details about organization, roles, deliverables, project risks, time plans and financial plans.

1.2 Intended Audience

This document shall be used in all phases of the project as a guideline. Intended audiences of this project are all project stakeholders:

- Project Supervisor
- Project Leader
- Team Members

1.3 Scope

This document defines the project plan of Spar Nord Bank ATM Transactions Data Analysis application. The overview includes objectives of the project, organization of the project team, development process that is going to be used during the project, assessment of possible risks, communication used between project stakeholders and project plan that includes time schedule and activity plan.

1.4 Definitions and acronyms

1.4.1 Definitions

Keyword	Definitions
The name of the project	Optimizing Customer Satisfaction: A Deep Dive into Spar Nord Bank ATM Transactions Data Analysis
Project Supervisor	Prof. Andrew Bond
Project Leader	Bhavana Prasad Kote
Team Members	Jay Dattoo Dale Rashmi Shree Veeraiah Sawan Shivanand Beli
Milestone	May 2, 2023

Git	https://github.com/rasho330/DATA-228-Big-Data-Tech-App
Scrum	An iterative and incremental agile software development method for managing software projects and product or application development
Kunagi	Web-based tool for integrated agile project management and collaboration based on Scrum
Scrum sprint	The basic unit of development in Scrum
Scrum master	Ensures the smooth working of the Scrum team and enforces Scrum practices
Product owner	Responsible for product management and its quality

1.4.2 Acronyms and abbreviations

Acronym or abbreviation	Definitions
SNB	Spar Nord Bank
ATD	ATM Transaction Data

1.5 References

1. http://www.scrum.org/
2. http://kunagi.org/

2. Background and Objectives

Abstract

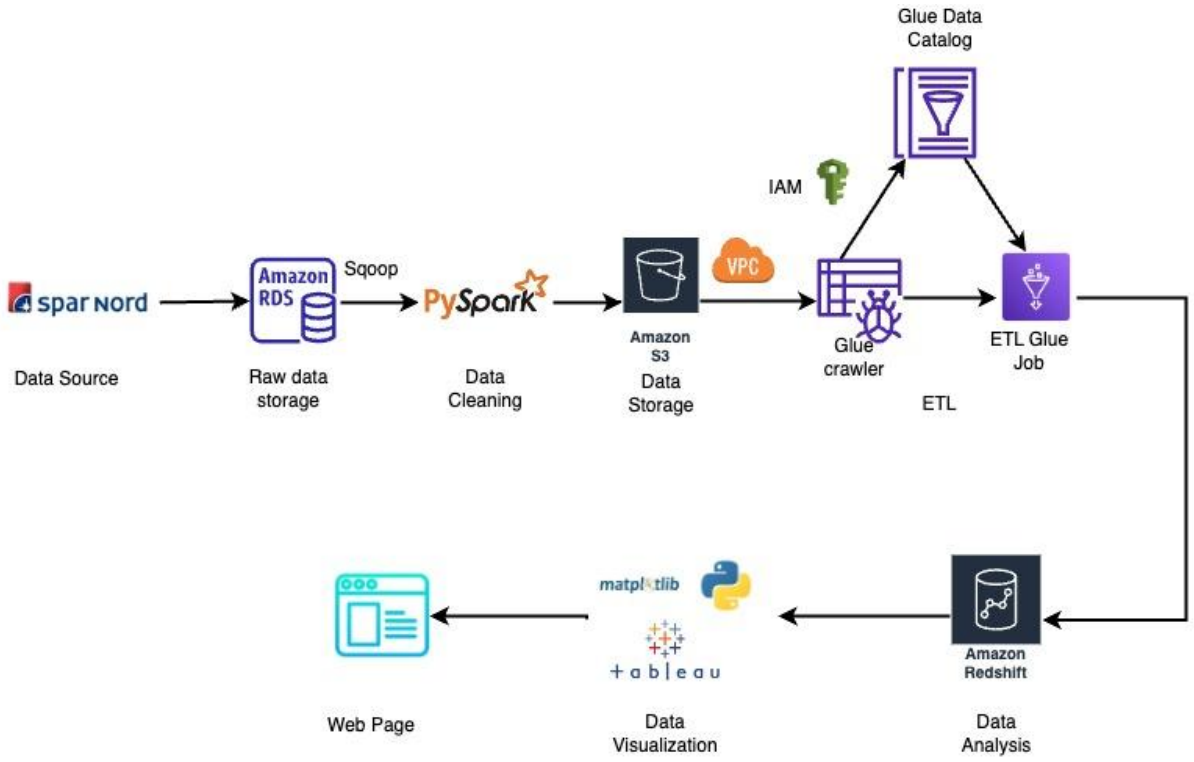
Spar Nord Bank is a financial institution that aims to provide a unique blend of personal advice and service supported by active involvement and cutting-edge digital solutions. The bank's operations are based on a franchise-inspired model where locally based ownership is the primary driver of customer relations and business volume, backed by strong central support. In 2017, Spar Nord Bank made its ATM transaction data publicly available and encouraged analysis of the information to gain insights into ATM usage patterns. The goal is to comprehend the behavior of customer withdrawals and perform an analysis utilizing various AWS services like S3, construct a batch ETL pipeline that will extract transactional

data from an RDS, transform it, and load it into target dimensions and facts on the Redshift Data Mart. The analysis of ATM usage patterns enables banks to enhance their operations, including ATM placement, network design, and cash management, which leads to decreased costs, heightened security, and an optimized ATM network. Furthermore, examining ATM transaction data assists in detecting fraud and minimizing financial losses. Ultimately, this analysis results in a more efficient, secure, and customer-oriented banking experience.

Objectives

The project analyzes Spar Nord Bank ATM transaction data. The analysis will discover patterns, trends, and customer behaviors to optimize the bank's ATM network, decrease expenses, improve security, and improve customer service. We will create a batch ETL pipeline that extracts transactional data from an RDS, transforms it, and loads it into target dimensions and facts on Redshift Data Mart. ETL will be fast, precise, and scalable for big data. S3 will store and manage transactional data, Redshift will store target data, and other AWS services will be used to efficiently execute the ETL process and analyze data. By identifying high-traffic regions, peak usage hours, and cash management demands, Spar Nord Bank can optimize its ATM network. This optimization should reduce bank costs and improve user satisfaction. ATM transaction data helps detect fraud and other security concerns. The project should discover problematic patterns and make security suggestions.

3. Architecture & High Level Design



4. Organization

The Pie Sparkers

4.1 Project group

Name	Initials	Responsibility (roles)
Bhavana Prasad Kote	BPK	Analysis, ETL, Visualization
Jay Dattoo Dale	JDD	Preprocessing, Analysis, Visualization, Deployment
Rashmi Shree Veeraiah	RSV	Analysis, ETL, Visualization
Sawan Shivanand Beli	SSB	Preprocessing, Redshift, Visualization, Deployment

4.2 Customer

The target customers are listed below:

- Spar Nord Bank
- Other Financial Institutions
- Project Supervisor

5. Development process

The project will use below mentioned tools throughout the development process

Tool Name	**Purpose**
AWS S3	S3 bucket to store raw transaction data from Spar Nord Bank ATMs.
AWS Glue	Transform and clean the raw transaction data and write it to a new S3 bucket.
PySpark	Use Spark to preprocess data and perform machine learning techniques to optimize customer satisfaction.
Python	Write Python scripts to preprocess and clean the data.
AWS Redshift	Create a Redshift cluster and use Glue to load the cleaned data from the S3 bucket into Redshift. Perform analysis using queries.
Tableau	Visualizations and interactive dashboards.

6. Deliverables

Phase	**Output**	**Planned week**	**Promised week**	**Late +/-**	**Delivered week**
Abstract	Data selection and Abstract submission	Feb 2nd week	Feb 2nd week		Feb 12, 2023
Design	Technology stack, tools to be used, Data	March 2nd week	March 4th week	+1	March 25, 2023

	cleaning and pipeline designing				
Coding	AWS pipeline creation	April 1st week	April 1st week		April 5, 2023
Documentation	Data documentation, report and visualization	April 3rd Week	April 4th Week	+1	April 24, 2023
Project Presentation	Project will be presented to intended audience and stakeholders	May 2nd, 2023	May 2nd, 2023		May 9, 2 2023

7. Project risks

N/A

8. Communication

- Google Meet
- Scrum meetings

8.1 Collaboration

Pair Programming

8.2 Git

All source code and finished documentation will be uploaded to Github repository.

Repository URL: https://github.com/rasho330/DATA-228-Big-Data-Tech-App

- Bhavana: https://github.com/Bprasad07
- Jay: https://github.com/Jaydale3221
- Rashmi: https://github.com/rasho330
- Sawan: https://github.com/Saavnbeli

9. Project plan

9.1 Time schedule

Id	Milestone Description	Responsible Dept./Initials	Finished week Plan
C1	Abstract	Team	Feb 2nd week
C2	Design	Team	March 2nd week
C3	Code	SSB	April 1st week
C4	ETL- AWS Glue	BPK	April 2nd week
C5	Analysis- PySpark	RSV	April 2nd week
C6	Analysis- Redshift	JDD	April 2nd week
C7	Code Testing	JDD, SVB	April 3rd week
C8	Visualization	Team	April 3rd week
C9	Website	BPK, RSV	April 4th week

C10	Documentation	Team	May 1st week

9.1.1 Remarks

N/A

9.2 Test plan

N/A

9.2.1 Testing Remarks

N/A

10. References

Data Source:
https://github.com/rasho330/DATA-228-Big-Data-Tech-App/blob/main/atm_data_part1.csv.zip
https://github.com/rasho330/DATA-228-Big-Data-Tech-App/blob/main/atm_data_part2.csv.zip
https://www.sparnord.com/
2.5M Danish ATM Transactions from 2017. (2019, January 16).
Kaggle.https://www.kaggle.com/datasets/sparnord/danish-atm-transactions