Rakamin Academy x syariah btpn

# ETL for Bank BTPN Syariah Credit Card Customer

## BTPN Syariah - Data Engineer

Presented by
Gilang Wiradhyaksa

Microsoft SQL Server Integration Services

Apache Airflow

tableau SOFTWARE

Rakamin Academy

# Gilang Wiradhyaksa

## Data Engineer

I am a Fullstack Software Developer transitioning to Data Science. Experienced in building a Web Application, API Web Service and Background Service.

Possess an understanding of statistical analysis, machine learning, and data visualization techniques, combined with strong programming skills in Python and proficiency in SQL.

My experience includes data preprocessing, feature engineering and model development.


Surabaya, Indonesia

gilang.wirad@gmail.com

in/gilangwiradhyaksa/

# Courses and Certification

**Hacktiv8 / Data Science** | Certificate — November, 2023

**Inixindo / Android Studio - Basic** | Certificate — March, 2018

# About Company

Since its inception as Sharia Business Unit of PT Bank Tabungan Pensiunan Nasional Tbk in 2010, BTPN Syariah has included and reached the segment that had not been touched by banking sector, that is the inclusive communities segment.

With the mandate to deliver empowerment activities and financial literacy for the women in the country, BTPN Syariah provides access and banking products and services in Sharia principles for them to affirm the intention to realize the aspiration for a better life.

On 14 July 2014, BTPN Syariah was officially registered as the 12th Sharia Commercial Bank in Indonesia by the spin off PT Bank Tabungan Pensiunan Nasional Tbk's Sharia Business Unit and the conversion of PT Bank Sahabat Purna Danarta ("BSPD").

# Project Portfolio

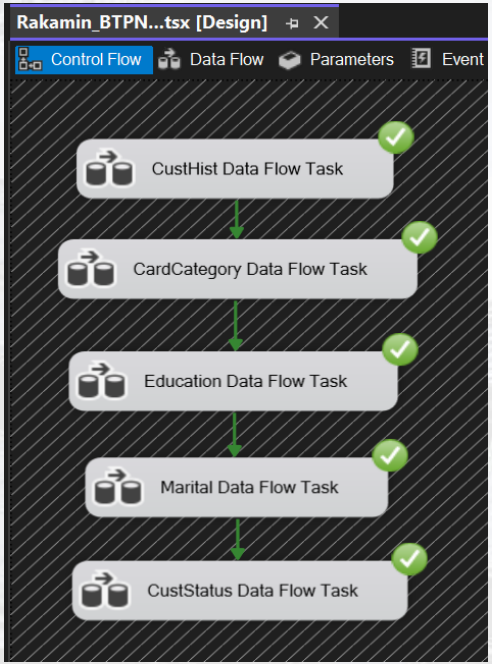
This project is an **ETL** (Extract Transform Load) project. The data contains **Bank BTPN Syariah Credit Card customer**. The project is designed to **gather** data from various sources, **transform** it into a consistent format suitable for analysis and then **load** it into a data warehouse or database for querying and reporting.

This ETL project perform a batch processing with scheduler for automation using Airflow. Starting from **SSIS**, extract data from CSV file then transform it and load it into database and another CSV file. Then run this SSIS package (dtsx file) using **Apache Airflow** scheduler through a bash command **everyday at 5 AM**.

Those process generate a **clean data** in CSV format, then this clean data is used to create **visualization in Tableau**. This visualization can help Data Analyst to get some **insight** and help management to **reduce customer churn**.

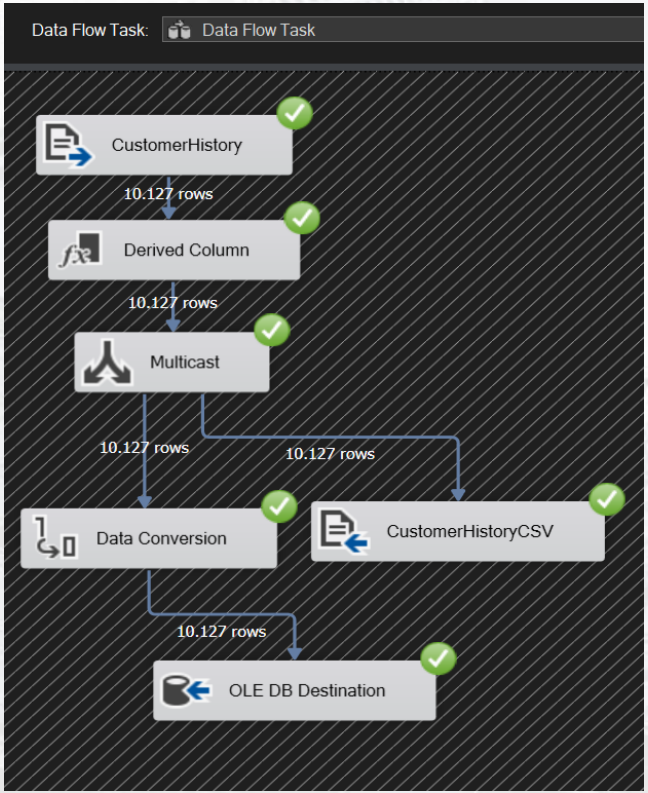
**Project explanation video here!**

# 1. SQL Server Integration Services (SSIS)



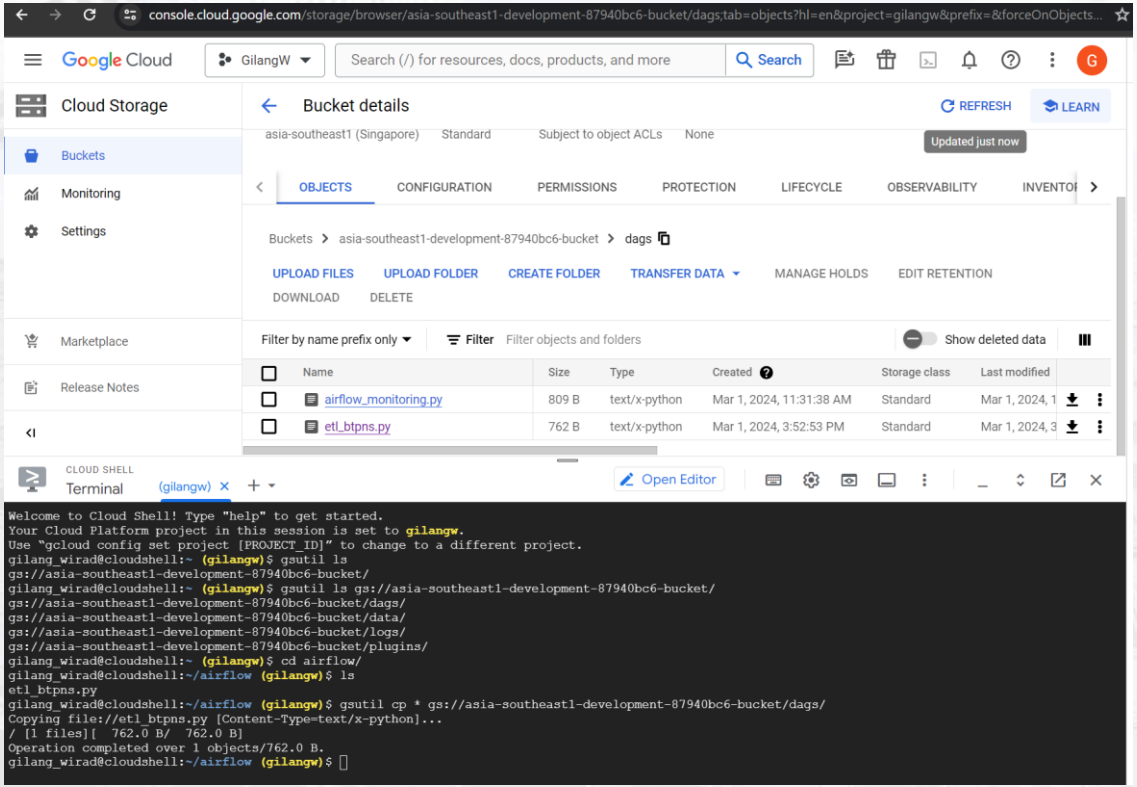

- Start from creating a new SSIS package using Microsoft Visual Studio.
- Create a new Flat File Connection and choose Bank BTPN Syariah CSV file as sources.
- Create five dataflow tasks with the provided data which is Customer History, Card Category, Education, Marital and Customer Status.

## SSIS Package (.dtsx) here!



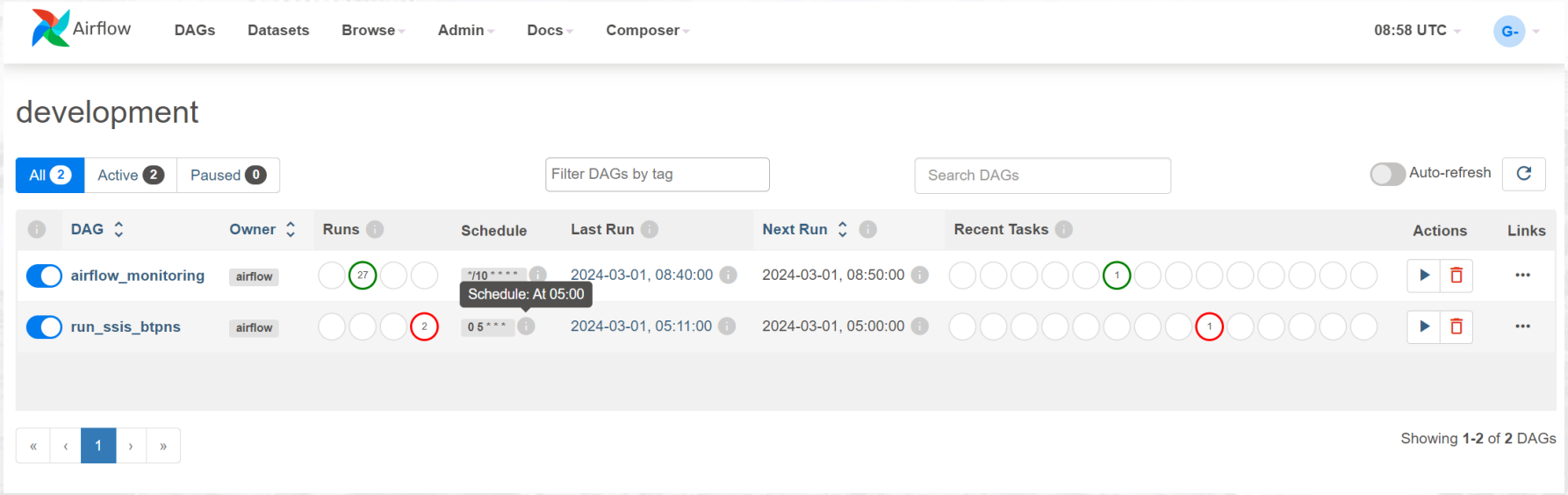

- Inside the data flow task, create Flat File Sources to Extract Data. Choose data source from connection manager which has been created previously.
- Begin the process of transforming data such as Derived Column, Data Type Conversion etc.
- Split the process to two or more output. So it can be treated individually. In this case data conversion is only done for data that will be stored in the database.
- Load the transformation result to CSV and Database simultaneously.

# 2. Creation of workflows on Apache Airflow




First create Airflow DAG in GCP for scheduling and running bash command. After it done, deploying .py file to DAG folder so it can appear in Airflow Task.

## Apache Airflow DAGs (.py) here!




DAGs is appear in Airflow Task with the name **“run_ssis_btpns”** and scheduled to run everyday at 5 am

## Apache Airflow DAGs (.py) here!

```
Administrator: Command Prompt
Microsoft Windows [Version 10.0.19045.4046]
(c) Microsoft Corporation. All rights reserved.

C:\WINDOWS\system32>"C:\Program Files\Microsoft SQL Server\160\DTS\Binn\DTExec.exe" /File "C:\Users\GilangW\source\repos\btpns_final_project\Rakamin_BTPN_Syariah.dtsx"
Microsoft (R) SQL Server Execute Package Utility
Version 16.0.1000.6 for 64-bit
Copyright (C) 2022 Microsoft. All rights reserved.

Started:  15:06:16
Progress: 2024-03-01 15:06:16.10
   Source: CardCategory Data Flow Task
   Validating: 0% complete
End Progress
Progress: 2024-03-01 15:06:16.10
   Source: CardCategory Data Flow Task
   Validating: 50% complete
End Progress
Progress: 2024-03-01 15:06:16.10
   Source: CardCategory Data Flow Task
   Validating: 100% complete
End Progress
```

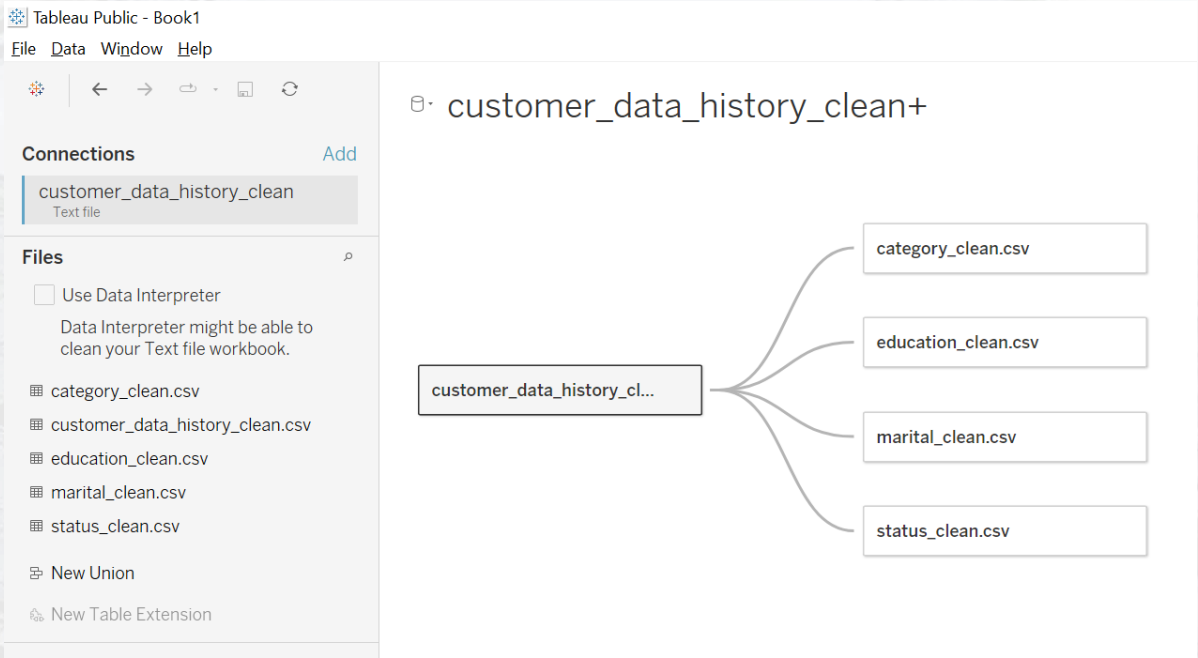
```
   Source: CustStatus Data Flow Task
   Cleanup: 0% complete
End Progress
Progress: 2024-03-01 15:06:16.70
   Source: CustStatus Data Flow Task
   Cleanup: 50% complete
End Progress
Progress: 2024-03-01 15:06:16.70
   Source: CustStatus Data Flow Task
   Cleanup: 100% complete
End Progress
DTExec: The package execution returned DTSER_SUCCESS (0).
Started:  15:06:16
Finished: 15:06:16
Elapsed:  0.703 seconds

C:\WINDOWS\system32>
```

Bash command to run **SSIS Package** (dtsx file) that executed using **Apache Airflow**.

Result of SSIS Package that have been executed. The ETL process is done and produce clean data to be further used for data **visualization** in **Tableau**.

# 3. Creating Visualization with Tableau

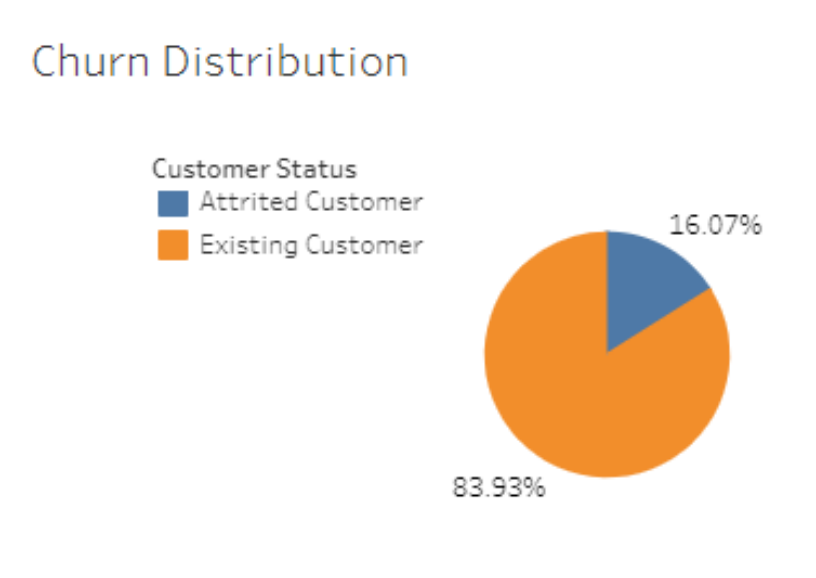



Add Text File then choose CSV file that has been cleaned before. Repeat this step until all CSV file is loaded to Tableau.

Set a relation between CSV file. Choose the foreign key between data to create a relation.

**Tableau Dashboard here!**

# Churn Distribution

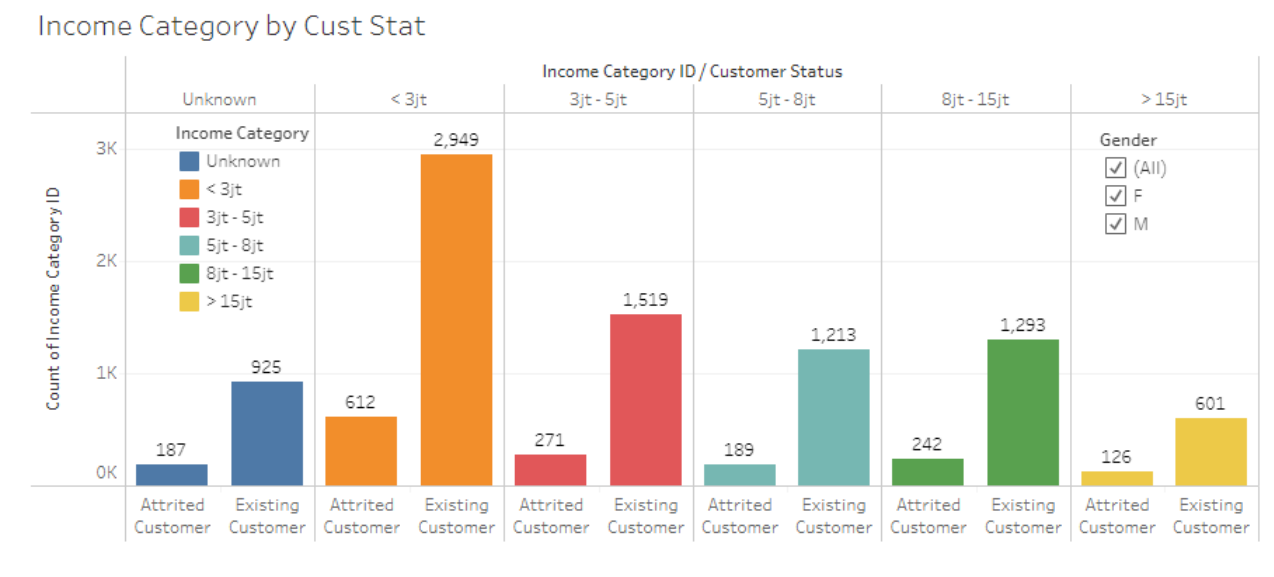



## Churn Rate

- Churn customer for BTPN Syariah Credit Card customer is at **16.07%**.
- Considered bad because according to forbes.com the **normal churn rate is 5-7%**.

**Tableau Dashboard [here!]**

# Income Category by Customer Churn




## Income Category

- Most of the credit card customer is in **low-income category**.
- **Most churn** customer is came from customer with **under 5m** per month.

**Tableau Dashboard here!**

# Marital Status by Customer Churn

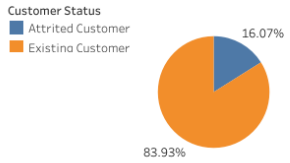

## Marital Status

- Most of the Credit Card customer is **Single** and **Married**.
- Percentage of customer churn in each status is about the same at **15-17%**.

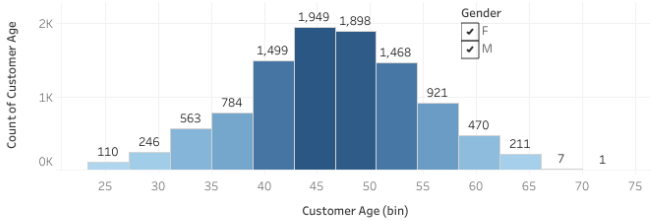
# Tableau Dashboard [here!]

## BTPNS Credit Card Customer

### Churn Distribution

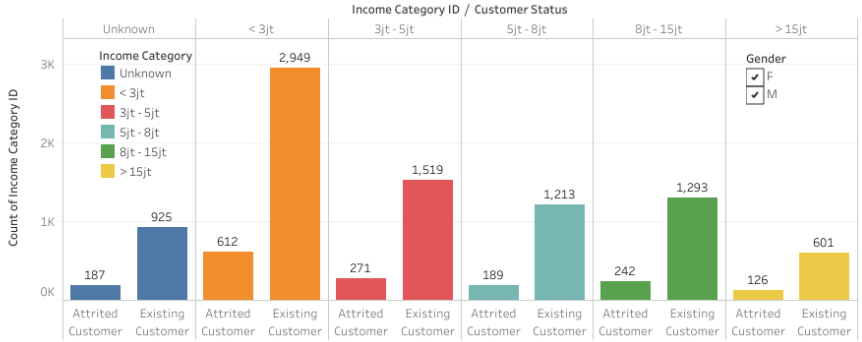



### Age Distribution

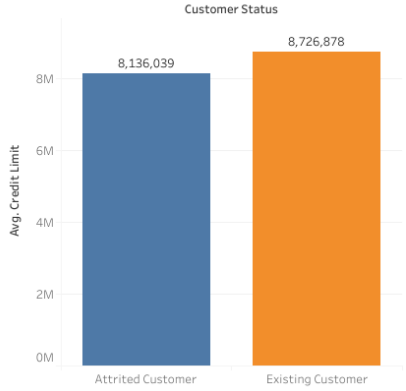



### Marital Stat by Cust Stat


### Income Category by Cust Stat




### Avg Credit Limit by Cust Stat



# Conclusion

- Based on the visualization we found that most of Bank BTPN Syariah Credit Card customer is came from people with **income less than 5m** per month.
- Also their most customer age is between **40 to 55** years old.
- The percentage of customer churn based on marital status is about the same. Means there is **no correlation** between **churn and marital** status.

# Suggestion

- For customer with **high-income**, giving a **reward** for using their credit card very often, so this customer will feel **appreciated and respected**. For high-income customer, **appreciating** them is more important rather than give them a discount or promotion.

- Whilst for **low-income customer** can give them a **promotion/discount** for a relatively **low-cost item** if they use credit card to attract them for using their credit card more.

- Do not make a promotion based on the **marital status** since there is **no correlation** between marital status and customer churn.

- Invest in providing exceptional **customer service experiences**. Ensure that customer support is easily accessible, responsive, and empathetic. Train representatives to **address customer concerns** effectively and resolve issues promptly.

# Project Result and Link

**Presentation Project | Video**

**SSIS Package (.dtsx) | GitHub**

**Apache Airflow DAGs (.py) | GitHub**

**Tableau Dashboard | Tableau**

# Thank You

Rakamin Academy x btpn syariah