Copy of Bernardo MLOps: Prepare for Production

short line



|  |  |  |
| --- | --- | --- |
| **Version** | **Author** | **Date** |
| 1.0 | bernardoyavhelopez | 2024-10-22 21:18:43 |

Table of contents

[I. Project general description 4](#_Toc137204120)

[II. Graphical Representation of the Flow 5](#_Toc137204121)

[III. Datasets 6](#_Toc137204122)

[1. {{$dataset.name}} 6](#_Toc137204123)

[a) Status of the dataset {{$dataset.name}} 6](#_Toc137204124)

[b) Related Recipes 7](#_Toc137204125)

[c) Schema 7](#_Toc137204126)

[IV. Managed folders 9](#_Toc137204127)

[1. {{$folder.name}} 9](#_Toc137204128)

[a) Status 9](#_Toc137204129)

[b) Derived datasets 10](#_Toc137204130)

[c) Related Recipes 10](#_Toc137204131)

[V. Recipes 11](#_Toc137204132)

[1. {{ $recipe.name }} 11](#_Toc137204133)

[a) Status 11](#_Toc137204134)

[b) Inputs / Outputs 11](#_Toc137204135)

[c) Prepare Steps 12](#_Toc137204136)

[d) Join 12](#_Toc137204137)

[e) Data Relationships 13](#_Toc137204138)

[f) Feature Transformations 14](#_Toc137204139)

[g) Code 14](#_Toc137204140)

[h) Windows 15](#_Toc137204141)

[i) Sorted Columns 15](#_Toc137204142)

[j) Top rows and sorting direction 15](#_Toc137204143)

[k) Parameters 16](#_Toc137204144)

[VI. Saved Models 17](#_Toc137204145)

[1. {{ $model.name }} 17](#_Toc137204146)

[a) Status 17](#_Toc137204147)

[b) Related Recipes 17](#_Toc137204148)

[VII. Model Evaluation Stores 19](#_Toc137204149)

[1. {{ $mes.name }} 19](#_Toc137204150)

[a) Related Recipes 19](#_Toc137204151)

[VIII. Labeling Tasks 21](#_Toc137204152)

[1. {{ $labeling\_task.name }} 21](#_Toc137204153)

[a) Status 21](#_Toc137204154)

[b) Inputs / Outputs 21](#_Toc137204155)

[IX. Deployment and Monitoring 23](#_Toc137204156)

[A. Implementation Details 23](#_Toc137204157)

[B. Version Control 23](#_Toc137204158)

# Project general description

Project references for the created flow:

* Project name: **Copy of Bernardo MLOps: Prepare for Production**
* Project key: **COPYOFBERNARDOMLOPSPREPAREFORPRODUCTION**
* Project short description:

The project \*STARTERAdvanced Designer\* was created by sylvain.franco on Mar 20th 2023  
The project \*STARTERMLOPSPREPAREFORPRODUCTION\* was copied from \*STARTERADVANCEDDESIGNER\* by alejandro.defrutos on Apr 11th 2023  
The project \*HITAMLOPSPREPAREFORPRODUCTION\* was copied from \*STARTERMLOPSPREPAREFORPRODUCTION\* by ivan\_hita on Oct 22nd 2024  
The project \*COPYOFBERNARDOMLOPSPREPAREFORPRODUCTION\* was copied from \*HITAMLOPSPREPAREFORPRODUCTION\* by bernardoyavhelopez on Oct 22nd 2024

* Description:

# Backgroud  
  
Your are working for ELO bank which is the Major Brazilian debit and credit card brand. Elo Back has issued more than 80 million cards.  
  
In ELO back there is data enginering department reponsible for providing datasets to the different department. In this case, you have access to three types of datasets:  
- Credit card transactions  
- Information of the merchant  
- Information of the cardholder  
  
# The problem  
  
Recently, the number of fraudulent transactions in the U.S. is increasing an, as consequence, the insurance cost of Elo Bank  
  
# Objective  
  
In this training you will need to:  
  
- Create metrics and checks to monitor the status of objects like datasets and models  
- Design robust scenarios that include metrics, checks, triggers, and reporters  
- Understand how code empowers users to customize functionality beyond the built-in metrics, checks, and scenarios  
- Perform other essential tasks required to move a project into such as: project refactoring and documentation techniques

* Status:
* Current git branch: **master**
* Tags: **duplicated, imported**
* Created by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 15:08:27
* Last Modification by bernardoyavhelopez on the 2024-10-22 16:44:41

# Graphical Representation of the Flow

# Datasets

List all the datasets in the flow.

### TRANSACTIONS\_2021

Summarized information about your dataset TRANSACTIONS\_2021:

* Tags**: origin:sql\_import, managed-snowflake**

#### Status of the dataset TRANSACTIONS\_2021

Characteristics of the dataset TRANSACTIONS\_2021:

* Type: **Snowflake**
* Connection: **managed-snowflake**
* Total Size: **Not computed**
* Record Count: **Not computed**
* Created by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 15:10:17
* Last Modification by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 17:12:43

#### Related Recipes

The dataset TRANSACTIONS\_2021 relates to the following recipes.

Successor Recipes

List of recipes with TRANSACTIONS\_2021 as an input:

* compute\_TRANSACTIONS\_stacked (Stack)

#### Schema

Schema details of the dataset.

|  |  |  |
| --- | --- | --- |
| ****Column Name**** | Column Type | Description |

|  |  |  |
| --- | --- | --- |
| ****transaction\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****authorized\_flag**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_category\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****item\_category**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_amount**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****signature\_provided**** | string |  |

### TRANSACTIONS\_2022

Summarized information about your dataset TRANSACTIONS\_2022:

* Tags**: origin:sql\_import, managed-snowflake**

#### Status of the dataset TRANSACTIONS\_2022

Characteristics of the dataset TRANSACTIONS\_2022:

* Type: **Snowflake**
* Connection: **managed-snowflake**
* Total Size: **Not computed**
* Record Count: **Not computed**
* Created by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 15:10:17
* Last Modification by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 17:12:34

#### Related Recipes

The dataset TRANSACTIONS\_2022 relates to the following recipes.

Successor Recipes

List of recipes with TRANSACTIONS\_2022 as an input:

* compute\_TRANSACTIONS\_stacked (Stack)

#### Schema

Schema details of the dataset.

|  |  |  |
| --- | --- | --- |
| ****Column Name**** | Column Type | Description |

|  |  |  |
| --- | --- | --- |
| ****transaction\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****authorized\_flag**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_category\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****item\_category**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_amount**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****signature\_provided**** | string |  |

### TRANSACTIONS\_stacked

Summarized information about your dataset TRANSACTIONS\_stacked:

#### Status of the dataset TRANSACTIONS\_stacked

Characteristics of the dataset TRANSACTIONS\_stacked:

* Type: **Snowflake**
* Connection: **managed-snowflake**
* Total Size: **Not computed**
* Record Count: **Not computed**
* Created by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 15:11:29
* Last Modification by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 15:11:29

#### Related Recipes

The dataset TRANSACTIONS\_stacked relates to the following recipes.

Parent Recipe

List of recipes with TRANSACTIONS\_stacked as an output:

* compute\_TRANSACTIONS\_stacked (Stack)

Successor Recipes

List of recipes with TRANSACTIONS\_stacked as an input:

* compute\_TRANSACTIONS\_stacked\_joined (Join)

#### Schema

Schema details of the dataset.

|  |  |  |
| --- | --- | --- |
| ****Column Name**** | Column Type | Description |

|  |  |  |
| --- | --- | --- |
| ****transaction\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****authorized\_flag**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_category\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****item\_category**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_amount**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****signature\_provided**** | string |  |

### MERCHANT\_INFO

Summarized information about your dataset MERCHANT\_INFO:

* Tags**: origin:sql\_import, managed-snowflake**

#### Status of the dataset MERCHANT\_INFO

Characteristics of the dataset MERCHANT\_INFO:

* Type: **Snowflake**
* Connection: **managed-snowflake**
* Total Size: **Not computed**
* Record Count: **Not computed**
* Created by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 15:10:17
* Last Modification by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 15:12:54

#### Related Recipes

The dataset MERCHANT\_INFO relates to the following recipes.

Successor Recipes

List of recipes with MERCHANT\_INFO as an input:

* compute\_TRANSACTIONS\_stacked\_joined (Join)

#### Schema

Schema details of the dataset.

|  |  |  |
| --- | --- | --- |
| ****Column Name**** | Column Type | Description |

|  |  |  |
| --- | --- | --- |
| ****merchant\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_category\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****subsector\_description**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****latitude**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****longitude**** | string |  |

### CARDHOLDER\_INFO

Summarized information about your dataset CARDHOLDER\_INFO:

* Tags**: origin:sql\_import, managed-snowflake**

#### Status of the dataset CARDHOLDER\_INFO

Characteristics of the dataset CARDHOLDER\_INFO:

* Type: **Snowflake**
* Connection: **managed-snowflake**
* Total Size: **Not computed**
* Record Count: **Not computed**
* Created by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 15:10:17
* Last Modification by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 17:12:27

#### Related Recipes

The dataset CARDHOLDER\_INFO relates to the following recipes.

Successor Recipes

List of recipes with CARDHOLDER\_INFO as an input:

* compute\_TRANSACTIONS\_stacked\_joined (Join)

#### Schema

Schema details of the dataset.

|  |  |  |
| --- | --- | --- |
| ****Column Name**** | Column Type | Description |

|  |  |  |
| --- | --- | --- |
| ****internal\_card\_mapping**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****first\_active\_month**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****reward\_program**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****latitude**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****longitude**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****fico\_score**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****age**** | string |  |

### TRANSACTIONS\_joined

Summarized information about your dataset TRANSACTIONS\_joined:

#### Status of the dataset TRANSACTIONS\_joined

Characteristics of the dataset TRANSACTIONS\_joined:

* Type: **Snowflake**
* Connection: **managed-snowflake**
* Total Size: **Not computed**
* Record Count: **Not computed**
* Created by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 16:39:20
* Last Modification by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 16:39:22

#### Related Recipes

The dataset TRANSACTIONS\_joined relates to the following recipes.

Parent Recipe

List of recipes with TRANSACTIONS\_joined as an output:

* compute\_TRANSACTIONS\_stacked\_joined (Join)

Successor Recipes

List of recipes with TRANSACTIONS\_joined as an input:

* compute\_TRANSACTIONS\_joined\_prepared (Prepare)

#### Schema

Schema details of the dataset.

|  |  |  |
| --- | --- | --- |
| ****Column Name**** | Column Type | Description |

|  |  |  |
| --- | --- | --- |
| ****transaction\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****authorized\_flag**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_category\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****item\_category**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_amount**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****signature\_provided**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_subsector\_description**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_latitude**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_longitude**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_first\_active\_month**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_reward\_program**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_latitude**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_longitude**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_fico\_score**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_age**** | string |  |

### TRANSACTIONS\_joined\_prepared

Summarized information about your dataset TRANSACTIONS\_joined\_prepared:

* Short description**:**

This dataset contains enriched features on credit card transactions.

* Tags**: credit-card-fraud**

#### Status of the dataset TRANSACTIONS\_joined\_prepared

Characteristics of the dataset TRANSACTIONS\_joined\_prepared:

* Type: **Snowflake**
* Connection: **managed-snowflake**
* Total Size: **Not computed**
* Record Count: **Not computed**
* Created by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 16:39:43
* Last Modification by sylvain.franco@dataiku.com (deleted) on the 2023-03-09 16:33:50

#### Related Recipes

The dataset TRANSACTIONS\_joined\_prepared relates to the following recipes.

Parent Recipe

List of recipes with TRANSACTIONS\_joined\_prepared as an output:

* compute\_TRANSACTIONS\_joined\_prepared (Prepare)

Successor Recipes

List of recipes with TRANSACTIONS\_joined\_prepared as an input:

* split\_TRANSACTIONS\_joined\_prepared (Split)

#### Schema

Schema details of the dataset.

|  |  |  |
| --- | --- | --- |
| ****Column Name**** | Column Type | Description |

|  |  |  |
| --- | --- | --- |
| ****transaction\_id**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****authorized\_flag**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****authorized\_flag\_description**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date\_parsed**** | date |  |

|  |  |  |
| --- | --- | --- |
| ****card\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_category\_id**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****item\_category**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_amount**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****signature\_provided**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****risk**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_subsector\_description**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****spending\_type**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_latitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_longitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****card\_first\_active\_month**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_reward\_program**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_latitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****card\_longitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****card\_fico\_score**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****card\_age**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_location**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_state**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_state\_enName**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_location**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_state**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_state\_enName**** | string |  |

### TRANSACTIONS\_known

Summarized information about your dataset TRANSACTIONS\_known:

#### Status of the dataset TRANSACTIONS\_known

Characteristics of the dataset TRANSACTIONS\_known:

* Type: **Snowflake**
* Connection: **managed-snowflake**
* Total Size: **Not computed**
* Record Count: **Not computed**
* Created by sylvain.franco@dataiku.com (deleted) on the 2023-03-02 14:03:25
* Last Modification by sylvain.franco@dataiku.com (deleted) on the 2023-03-09 16:34:18

#### Related Recipes

The dataset TRANSACTIONS\_known relates to the following recipes.

Parent Recipe

List of recipes with TRANSACTIONS\_known as an output:

* split\_TRANSACTIONS\_joined\_prepared (Split)

Successor Recipes

List of recipes with TRANSACTIONS\_known as an input:

* train\_Predict\_authorized\_flag\_\_binary\_ (Prediction Training)

#### Schema

Schema details of the dataset.

|  |  |  |
| --- | --- | --- |
| ****Column Name**** | Column Type | Description |

|  |  |  |
| --- | --- | --- |
| ****transaction\_id**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****authorized\_flag**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****authorized\_flag\_description**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date\_parsed**** | date |  |

|  |  |  |
| --- | --- | --- |
| ****card\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_category\_id**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****item\_category**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_amount**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****signature\_provided**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****risk**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_subsector\_description**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****spending\_type**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_latitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_longitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****card\_first\_active\_month**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_reward\_program**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_latitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****card\_longitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****card\_fico\_score**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****card\_age**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_location**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_state**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_state\_enName**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_location**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_state**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_state\_enName**** | string |  |

### TRANSACTIONS\_unknown

Summarized information about your dataset TRANSACTIONS\_unknown:

#### Status of the dataset TRANSACTIONS\_unknown

Characteristics of the dataset TRANSACTIONS\_unknown:

* Type: **Snowflake**
* Connection: **managed-snowflake**
* Total Size: **Not computed**
* Record Count: **Not computed**
* Created by sylvain.franco@dataiku.com (deleted) on the 2023-03-02 14:03:46
* Last Modification by sylvain.franco@dataiku.com (deleted) on the 2023-03-09 16:34:18

#### Related Recipes

The dataset TRANSACTIONS\_unknown relates to the following recipes.

Parent Recipe

List of recipes with TRANSACTIONS\_unknown as an output:

* split\_TRANSACTIONS\_joined\_prepared (Split)

Successor Recipes

List of recipes with TRANSACTIONS\_unknown as an input:

* score\_TRANSACTIONS\_unknown (Prediction Scoring)

#### Schema

Schema details of the dataset.

|  |  |  |
| --- | --- | --- |
| ****Column Name**** | Column Type | Description |

|  |  |  |
| --- | --- | --- |
| ****transaction\_id**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****authorized\_flag**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****authorized\_flag\_description**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date\_parsed**** | date |  |

|  |  |  |
| --- | --- | --- |
| ****card\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_category\_id**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****item\_category**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_amount**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****signature\_provided**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****risk**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_subsector\_description**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****spending\_type**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_latitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_longitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****card\_first\_active\_month**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_reward\_program**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_latitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****card\_longitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****card\_fico\_score**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****card\_age**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_location**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_state**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_state\_enName**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_location**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_state**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_state\_enName**** | string |  |

### TRANSACTIONS\_unknown\_scored

Summarized information about your dataset TRANSACTIONS\_unknown\_scored:

#### Status of the dataset TRANSACTIONS\_unknown\_scored

Characteristics of the dataset TRANSACTIONS\_unknown\_scored:

* Type: **Snowflake**
* Connection: **managed-snowflake**
* Total Size: **Not computed**
* Record Count: **Not computed**
* Created by alejandro.defrutos@dataiku.com (deleted) on the 2023-04-11 08:33:07
* Last Modification by alejandro.defrutos@dataiku.com (deleted) on the 2023-04-11 08:33:07

#### Related Recipes

The dataset TRANSACTIONS\_unknown\_scored relates to the following recipes.

Parent Recipe

List of recipes with TRANSACTIONS\_unknown\_scored as an output:

* score\_TRANSACTIONS\_unknown (Prediction Scoring)

#### Schema

Schema details of the dataset.

|  |  |  |
| --- | --- | --- |
| ****Column Name**** | Column Type | Description |

|  |  |  |
| --- | --- | --- |
| ****transaction\_id**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****authorized\_flag**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****authorized\_flag\_description**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date\_parsed**** | date |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date\_parsed\_month**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_date\_parsed\_day**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****puchase\_dow**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_weekends**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_id**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_category\_id**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****item\_category**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****purchase\_amount**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****signature\_provided**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****risk**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_subsector\_description**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****spending\_type**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_latitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_longitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****card\_first\_active\_month**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_first\_active\_month\_parsed**** | date |  |

|  |  |  |
| --- | --- | --- |
| ****days\_active**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****card\_reward\_program**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_latitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****card\_longitude**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****card\_fico\_score**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****card\_age**** | bigint |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_location**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****distance\_card\_merchant**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_state**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****merchant\_state\_enName**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_location**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_state**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****card\_state\_enName**** | string |  |

|  |  |  |
| --- | --- | --- |
| ****proba\_0**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****proba\_1**** | double |  |

|  |  |  |
| --- | --- | --- |
| ****prediction**** | string |  |

# Recipes

List all the recipes in the flow.

### compute\_TRANSACTIONS\_stacked

Summarized information about the recipe compute\_TRANSACTIONS\_stacked.

#### Status

Characteristics of the recipe compute\_TRANSACTIONS\_stacked:

* Type: **Stack**
* Created by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 15:11:29
* Last Modification by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 15:11:32

#### Inputs / Outputs

List of the inputs of the recipe:

* Dataset TRANSACTIONS\_2021 (Snowflake)
* Dataset TRANSACTIONS\_2022 (Snowflake)

List of the outputs of the recipe:

* Dataset TRANSACTIONS\_stacked (Snowflake)

Stacking mode: Union of input schemas

### compute\_TRANSACTIONS\_stacked\_joined

Summarized information about the recipe compute\_TRANSACTIONS\_stacked\_joined.

#### Status

Characteristics of the recipe compute\_TRANSACTIONS\_stacked\_joined:

* Type: **Join**
* Created by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 15:12:00
* Last Modification by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 16:39:25

#### Inputs / Outputs

List of the inputs of the recipe:

* Dataset TRANSACTIONS\_stacked (Snowflake)
* Dataset MERCHANT\_INFO (Snowflake)
* Dataset CARDHOLDER\_INFO (Snowflake)

List of the outputs of the recipe:

* Dataset TRANSACTIONS\_joined (Snowflake)

#### Join

Left TRANSACTIONS\_stacked - MERCHANT\_INFO And

|  |  |  |
| --- | --- | --- |
| Left input | Condition | Right input |

|  |  |  |
| --- | --- | --- |
| merchant\_id | = | merchant\_id |

|  |  |  |
| --- | --- | --- |
| merchant\_category\_id | = | merchant\_category\_id |

Left TRANSACTIONS\_stacked - CARDHOLDER\_INFO And

|  |  |  |
| --- | --- | --- |
| Left input | Condition | Right input |

|  |  |  |
| --- | --- | --- |
| card\_id | = | internal\_card\_mapping |

### compute\_TRANSACTIONS\_joined\_prepared

Summarized information about the recipe compute\_TRANSACTIONS\_joined\_prepared.

#### Status

Characteristics of the recipe compute\_TRANSACTIONS\_joined\_prepared:

* Type: **Prepare**
* Created by sylvain.franco@dataiku.com (deleted) on the 2022-12-08 16:39:43
* Last Modification by sylvain.franco@dataiku.com (deleted) on the 2023-03-09 16:33:52

#### Inputs / Outputs

List of the inputs of the recipe:

* Dataset TRANSACTIONS\_joined (Snowflake)

List of the outputs of the recipe:

* Dataset TRANSACTIONS\_joined\_prepared (Snowflake)

#### Prepare Steps

|  |  |
| --- | --- |
| Type | Comment |

|  |  |
| --- | --- |
| FilterOnBadType |  |

|  |  |
| --- | --- |
| FindReplace |  |

|  |  |
| --- | --- |
| DateParser |  |

|  |  |
| --- | --- |
| GeoPointCreator |  |

|  |  |
| --- | --- |
| GeoPointCreator |  |

|  |  |
| --- | --- |
| CityLevelReverseGeocoder |  |

|  |  |
| --- | --- |
| CityLevelReverseGeocoder |  |

|  |  |
| --- | --- |
| SwitchCase |  |

|  |  |
| --- | --- |
| VisualIfRule |  |

### split\_TRANSACTIONS\_joined\_prepared

Summarized information about the recipe split\_TRANSACTIONS\_joined\_prepared.

#### Status

Characteristics of the recipe split\_TRANSACTIONS\_joined\_prepared:

* Type: **Split**
* Created by sylvain.franco@dataiku.com (deleted) on the 2023-03-02 14:04:07
* Last Modification by alejandro.defrutos@dataiku.com (deleted) on the 2023-04-11 08:29:50

#### Inputs / Outputs

List of the inputs of the recipe:

* Dataset TRANSACTIONS\_joined\_prepared (Snowflake)

List of the outputs of the recipe:

* Dataset TRANSACTIONS\_known (Snowflake)
* Dataset TRANSACTIONS\_unknown (Snowflake)

Split mode: Specific filter for each output dataset

### train\_Predict\_authorized\_flag\_\_binary\_

Summarized information about the recipe train\_Predict\_authorized\_flag\_\_binary\_.

#### Status

Characteristics of the recipe train\_Predict\_authorized\_flag\_\_binary\_:

* Type: **Prediction Training**
* Created by sylvain.franco@dataiku.com (deleted) on the 2023-03-02 16:03:26
* Last Modification by alejandro.defrutos@dataiku.com (deleted) on the 2023-04-11 08:29:50

#### Inputs / Outputs

List of the inputs of the recipe:

* Dataset TRANSACTIONS\_known (Snowflake)

List of the outputs of the recipe:

* Saved Model Predict authorized\_flag (binary) (Random forest (Feature engineering) - v2)

### score\_TRANSACTIONS\_unknown

Summarized information about the recipe score\_TRANSACTIONS\_unknown.

#### Status

Characteristics of the recipe score\_TRANSACTIONS\_unknown:

* Type: **Prediction Scoring**
* Created by alejandro.defrutos@dataiku.com (deleted) on the 2023-04-11 08:33:07
* Last Modification by alejandro.defrutos@dataiku.com (deleted) on the 2023-04-11 08:33:09

#### Inputs / Outputs

List of the inputs of the recipe:

* Dataset TRANSACTIONS\_unknown (Snowflake)
* Saved Model Predict authorized\_flag (binary) (Random forest (Feature engineering) - v2)

List of the outputs of the recipe:

* Dataset TRANSACTIONS\_unknown\_scored (Snowflake)

# Saved Models

List all the saved models in the flow.

### Predict authorized\_flag (binary)

Summarized information about the saved model Predict authorized\_flag (binary):.

* Short Description:

We added feature to the model

#### Status

Model id: Y1aHspjf

Active version: Random forest (Feature engineering) - v2

#### Related Recipes

List of the recipes related to the saved model.

Parent Recipe

Recipe with Predict authorized\_flag (binary) as an output:

* train\_Predict\_authorized\_flag\_\_binary\_ (Prediction Training)

Successor Recipes

List of recipes with Predict authorized\_flag (binary) as an input:

* score\_TRANSACTIONS\_unknown (Prediction Scoring)

# Deployment and Monitoring

## Implementation Details

* The flow can be found here: https://dss-649c8ce9-81d9904e-dku.us-east-1.app.dataiku.io/projects/COPYOFBERNARDOMLOPSPREPAREFORPRODUCTION/
* The name of the generated file is: Dataiku Flow Documentation - COPYOFBERNARDOMLOPSPREPAREFORPRODUCTION.docx

## Version Control

* The flow was executed with the following version of DSS: 13.2.0