

Департамент образования и науки города Москвы
Государственное автономное образовательное учреждение
высшего образования города Москвы
«Московский городской педагогический университет»
Институт цифрового образования
Департамент информатики управления и технологий

Нургалеева Гузель Рустэмовна БД-241м

Data Engineering. Практическая работа 4-1.

Вариант задания: 18

Направление подготовки/специальность
38.04.05 - Бизнес-информатика
Бизнес-аналитика и большие данные
(очная форма обучения)

Москва

2024

Datacert: <https://www.kaggle.com/datasets/rakeshrau/social-network-ads?resource=download>

General

Advanced

Options

Pooling

Clustering

Connection tested successfully

Connection to lab_3_NG was successful.

Hostname: 95.131.149.21

Port: 3306

Database name: mgpu_m_etl_18

OK

Settings

Host Name: 95.131.149.21

Database Name: mgpu_m_etl_18

Port Number: 3306

Username: mgpu_m_etl_18

Password: ●●●●●●●●

☒ Use Result Streaming Cursor

Oracle RDB

Palo MOLAP Server

PostgreSQL

Redshift

Remedy Action Request System

SAP ERP System

Access:

Native (JDBC)

ODBC

Table output

Step name: Table output

Connection: lab_3_NG

Target schema: mgpu_m_etl_18

Target table: ads_results

Commit size: 1000

Truncate table: ☐

Ignore insert errors: ☐

Specify database fields: ☐

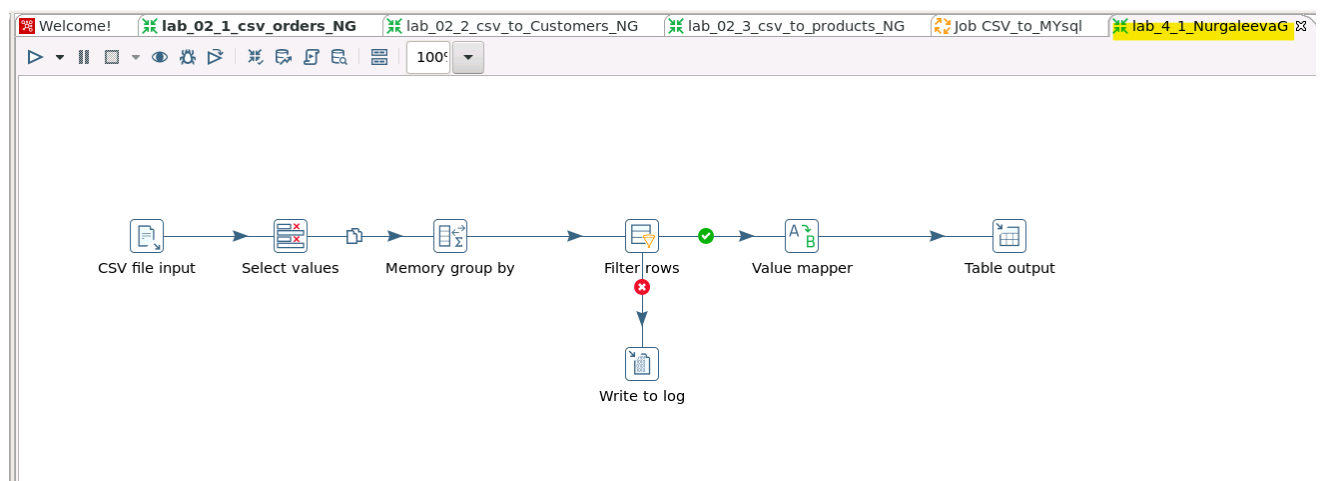
Database fields

phpMyAdmin

Server: localhost:3306 > Database: mgpu_m_etl_18 > Table: products

MySQL returned an empty result set (i.e. zero rows). (Query took 0.0003 seconds)

CREATE TABLE ads_results (user_id INT PRIMARY KEY, gender INT DEFAULT 0, -- 1 = Female, 2 = Male age INT, estimated_salary INT, purchased INT);



lab_02_1_csv_orders_NG lab_02_2_csv_to_Customers_NG lab_02_3_csv_to_products_NG Job CSV_to_Mysql lab_4_1_NurgaleevaG

CSV file input → Select values → Memory group by → Filter rows → Value mapper → Table output

Write to log

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

2025/03/18 11:59:54 - Select values.0 - Finished processing (I=0, O=0, R=400, W=400, U=0, E=0)
2025/03/18 11:59:54 - Memory group by.0 - Finished processing (I=0, O=0, R=400, W=400, U=0, E=0)
2025/03/18 11:59:54 - Filter rows.0 - Finished processing (I=0, O=0, R=400, W=400, U=0, E=0)
2025/03/18 11:59:54 - Value mapper.0 - Finished processing (I=0, O=0, R=400, W=400, U=0, E=0)
2025/03/18 11:59:56 - Table output.0 - Finished processing (I=0, O=0, R=400, W=400, U=0, E=0)
2025/03/18 11:59:56 - Spoon - The transformation has finished!!

phpMyAdmin

Server: localhost:3306 » Database: mgpu_m_etl_18 » Table: ads_results

Showing rows 0 - 24 (400 total, Query took 0.0002 seconds.)

SELECT * FROM `ads_results`

1 > >> Show all Number of rows: 25 Filter rows: Search this table Sort by key: None

	user_id	gender	age	estimated_salary	purchased
<input type="checkbox"/>	15566689	1	35	57000	0
<input type="checkbox"/>	15569641	1	58	95000	1
<input type="checkbox"/>	15570769	1	26	80000	0
<input type="checkbox"/>	15570932	2	34	115000	0
<input type="checkbox"/>	15571059	1	33	41000	0
<input type="checkbox"/>	15573452	1	21	16000	0
<input type="checkbox"/>	15573926	2	40	71000	1
<input type="checkbox"/>	15574305	2	35	53000	0
<input type="checkbox"/>	15574372	1	58	47000	1
<input type="checkbox"/>	15575002	1	35	60000	0
<input type="checkbox"/>	15575247	1	48	131000	1

phpMyAdmin

Server: localhost:3306 » Database: mgpu_m_etl_18 » Table: ads_results

Show query box

Current selection does not contain a unique column. Grid edit, checkbox, Edit, Copy and Delete features are not available.

Showing rows 0 - 0 (1 total, Query took 0.0003 seconds.)

SELECT count(user_id) as users FROM `ads_results` WHERE gender = 1;

Show all Number of rows: 25 Filter rows: Search this table

Extra options

users

204

Show all Number of rows: 25 Filter rows: Search this table