

Common issues

- Checks failing
 - Checks associated to inputs in the hands of the business users
 - Technical checks
- Schema changes in the data sources
- Imputer error
- StackOverflowError
- Best params dataset is not ready

Checks failing

Checks associated to inputs in the hands of the business users

TBD

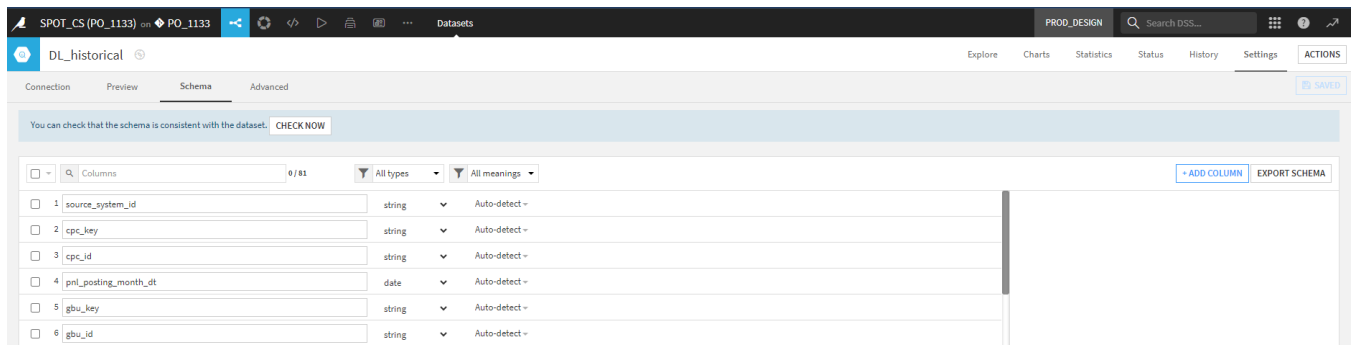
Technical checks

If any technical check is failing, please refer to the [dedicated page](#) to better understand how to resolve the situation.

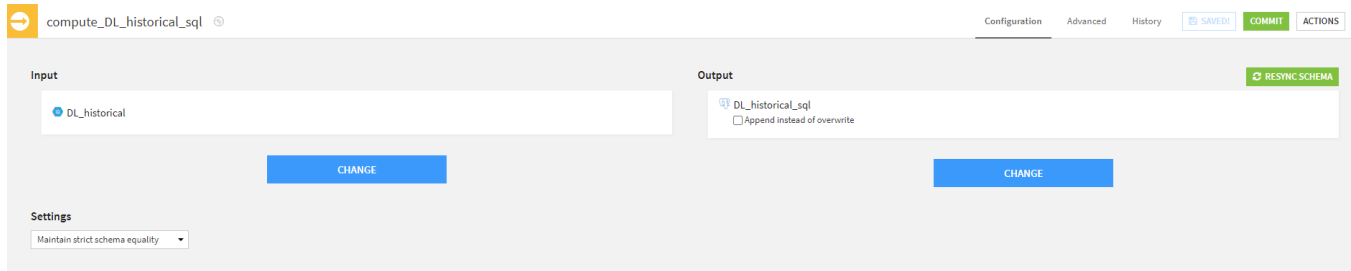
Schema changes in the data sources

If some of the recipes of the "data extracts" Flow zone fail, it is certainly due to a change of schema in one of the sources. It can especially happen on the Big Query Data Lake sources as they might evolve over time.

To resolve it, access the failing dataset and go in the "Settings" tab. From here, you can access the "Schema" tab, check it and update the schema from the source table if needed.



Once done, access the following Sync recipe and click the "Resync schema" button, then click "Run" at the bottom left of the screen.



Imputer error

The following error can happen if the imputer is called on an empty dataset. The imputer being called product family by product family, this can happen if one of the family does not have any CPC.

This probably means there is another issue at a previous stage of the Flow that went undetected. In this case, find the root of the issue in the data preparation steps.

1 Activity
✖ 1 HIDE FLOW

STATUS

✖ compute_price_drivers_features 13s

Activity Log
CHOOSE LOG
SHOW SUMMARY RECIPES OUTPUTS

✖ Oops: an unexpected error occurred

Error in Python process: At line 237: <class 'ValueError': Found array with 0 sample(s) (shape=(0, 1)) while a minimum of 1 is required by SimpleImputer. Please see our options for getting help

HTTP code: , type: <class 'ValueError'>

```

Status: Executing Python script
[12:35:27] [INFO] [diku] running compute_price_drivers_features_MP - .....
[12:35:27] [INFO] [diku] running compute_price_drivers_features_MP - OS: startup: 2d6 version:12.1.0
[12:35:27] [INFO] [diku] running compute_price_drivers_features_MP - OS Name: /dataiku/dss_data
[12:35:27] [INFO] [diku] running compute_price_drivers_features_MP - OS: Linux 5.15.0-1062-gcp amd64 - Java: Ubuntu 11.0.24
[12:35:27] [INFO] [diku.flow.jobrunner] running compute_price_drivers_features_MP - Allocated a slot for this activity!
[12:35:27] [INFO] [diku.flow.jobrunner] running compute_price_drivers_features_MP - Run activity
[12:35:27] [INFO] [diku.flow.activity] running compute_price_drivers_features_MP - Executing default pre-activity lifecycle hook
[12:35:27] [INFO] [diku.provider.connection.prepared] running compute_price_drivers_features_MP - Setup connection when encrypted fileds GCS_prj-data-pricing-dev
[12:35:27] [INFO] [diku.managefolders.handler] running compute_price_drivers_features_MP - Create provider for SPOT_SPP_17qmh251 with path /dataiku/runs/dev/SP17qmh251/
[12:35:27] [INFO] [diku.koprodviders.google] running compute_price_drivers_features_MP - GCS provider bucket:cs-hub-prj-data-pricing-dev:dataiku-effects@dataiku.com/dev/SP17qmh251
(connection:rootnull)
[12:35:27] [INFO] [diku.flow.activity] running compute_price_drivers_features_MP - Checking if sources are ready
[12:35:27] [INFO] [diku.flow.activity] running compute_price_drivers_features_MP - Will check readiness of SPOT_SPP_sales_evolution_features_pmp
[12:35:28] [INFO] [diku.dataset.hash] running compute_price_drivers_features_MP - managed non-PS dataset, computed new hash for SPOT_SPP_sales_evolution_features_MP: 11w0qWkdyb2pzhvNMQD
#P123262vFq6e1u6w6w, based on 2 propagated from upstream
[12:35:28] [INFO] [diku.flow.activity] running compute_price_drivers_features_MP - Checked source readiness SPOT_SPP_Sales_evolution_features -> true
[12:35:28] [INFO] [diku.flow.activity] running compute_price_drivers_features_MP - Will check readiness of SPOT_SPP_ALL_features_prices_21_31ined_pmp
[12:35:28] [INFO] [diku.dataset.hash] running compute_price_drivers_features_MP - managed non-PS dataset, computed new hash for SPOT_SPP_ALL_features_prices_21_31ined_MP: q5QmX14w-wq355k6d4d5w
#7Mx0d0yvh2p2e7mg, based on 9 propagated from upstream
[12:35:28] [INFO] [diku.flow.activity] running compute_price_drivers_features_MP - Checked source readiness SPOT_SPP_ALL_features_prices_21_31ined -> true
[12:35:28] [DEBUG] [diku.flow.activity] running compute_price_drivers_features_MP - computing hashes to propagate BEFORE activity
[12:35:28] [DEBUG] [diku.flow.activity] running compute_price_drivers_features_MP - Recorded 4 hashes before activity run
[12:35:28] [DEBUG] [diku.flow.activity] running compute_price_drivers_features_MP - Building recipe number of type
          
```

StackOverflowError

This error happens when Dataiku fails at computing the dependencies between all of the elements.

✖ Oops: an unexpected error occurred

StackOverflowError

Please see our options for getting help

HTTP code: , type: java.lang.StackOverflowError

To avoid this, when building a dataset it is advised to select the "Not recursive" option instead of the "Recursive upstream" one.

▶ Build dataset "pricing_features_dataset"
✖

← Recursive upstream

Build upstream dependencies and then this dataset

Not recursive

Build only this dataset

Recursive downstream →

Run recipes starting from this dataset

Update output schemas Before the recipe runs, update the schema of the output dataset(s)

BUILD DATASET

Best params dataset is not ready

✖ Input dataset is not ready (table does not exist)

Input dataset `SPOT_SPP_PO_1286.best_params_dataset` is not ready, caused by: `CodedSQLException: in running compute_8OoB6E18_NP: SQL table does not exist: SPOT_SPP_PO_1286_best_params_dataset`

[More information on this error in DSS documentation](#)

Additional technical details: Code: ERR_JOB_INPUT_DATASET_NOT_READY_NO_TABLE, type: com.dataiku.dip.exceptions.SourceDatasetNotReadyException

When creating a new branch for a development on one of the projects, it is necessary to rebuild some tables as shown [here](#) (example for SpP GBU). This error will appear if they are not rebuilt before running the scenarios.

