AID-Project

I - Developing the data warehouse

1. Creation of the airports database

Once connected to the MySQL server, we execute the command to create the provided airport database:
`source airports.sql`

2. Creation of the airports data warehouse

We create 4 dimension tables and 1 fact table into the data warehouse airports_dw :

- dim_airline : Airline dimension
- dim_airplane : Airplane dimension
- dim_airport : Airport dimension for origin and destination
- dim_time : Time dimension for departure and arrival
- fact_flight : Fact table for flight

Here is the code:

```
DROP DATABASE IF EXISTS airports_dw;

CREATE DATABASE airports_dw;

USE airports_dw;

CREATE TABLE dim_airline (
    AIRLINE_ID INT,
    AIRLINE_NAME VARCHAR(255),
    PRIMARY KEY (AIRLINE_ID)
);

CREATE TABLE dim_airplane (
    AIRPLANE_ID INT,
    AIRPLANE_TYPE VARCHAR(255),
    PRIMARY KEY (AIRPLANE_ID)
);

CREATE TABLE dim_airport(
    AIRPORT_ID INT,
    AIRPORT_NAME VARCHAR(255),
    CITY VARCHAR(255),
    COUNTRY VARCHAR(255),
    PRIMARY KEY (AIRPORT_ID)
);

CREATE TABLE dim_time(
    TIME_ID DATETIME,
```

```
    YEAR_ID INT,
    MONTH_ID INT,
    MONTH_NAME VARCHAR(255),
    DAY_ID INT,
    PRIMARY KEY (TIME_ID)
);

CREATE TABLE fact_flight(
    FLIGHT_ID INT,
    PASSENGERS_NUMBER INT,
    RECEIVE_TOTAL DOUBLE,
    AIRLINE_ID INT,
    AIRPLANE_ID INT,
    ORIGIN_AIRPORT_ID INT,
    DESTINATION_AIRPORT_ID INT,
    DEPARTURE_TIME_ID DATETIME,
    ARRIVAL_TIME_ID DATETIME,
    PRIMARY KEY (FLIGHT_ID),
    FOREIGN KEY (AIRLINE_ID) REFERENCES dim_airline (AIRLINE_ID),
    FOREIGN KEY (AIRPLANE_ID) REFERENCES dim_airplane (AIRPLANE_ID),
    FOREIGN KEY (ORIGIN_AIRPORT_ID) REFERENCES dim_airport (AIRPORT_ID),
    FOREIGN KEY (DESTINATION_AIRPORT_ID) REFERENCES dim_airport
(AIRPORT_ID),
    FOREIGN KEY (DEPARTURE_TIME_ID) REFERENCES dim_time (TIME_ID),
    FOREIGN KEY (ARRIVAL_TIME_ID) REFERENCES dim_time (TIME_ID)
);
```

Then to create the airports data warehouse we execute the following command: `source airports_dw.sql`

II - Transformations developed in PDI

Airline dimension

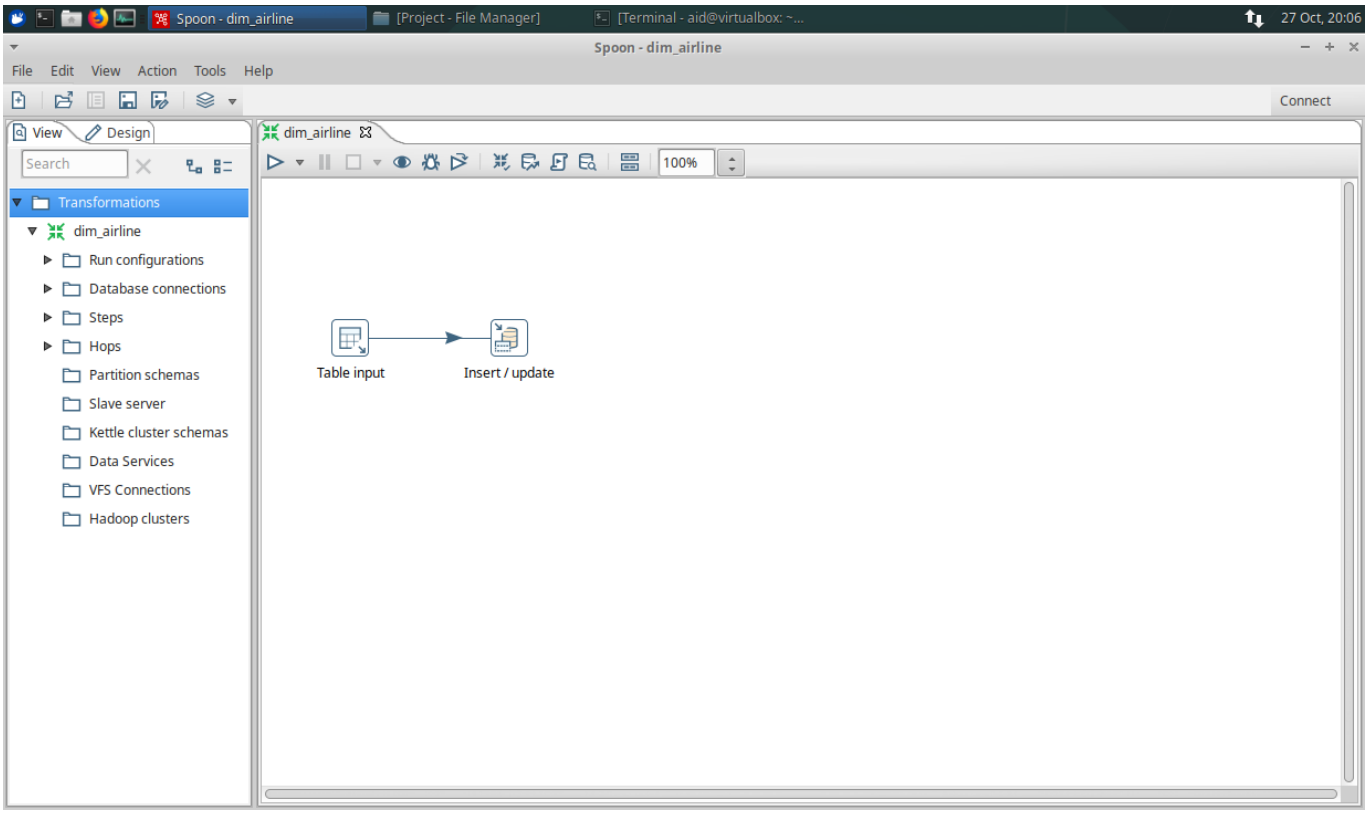

Figure 1 - dim_airline entire transformation

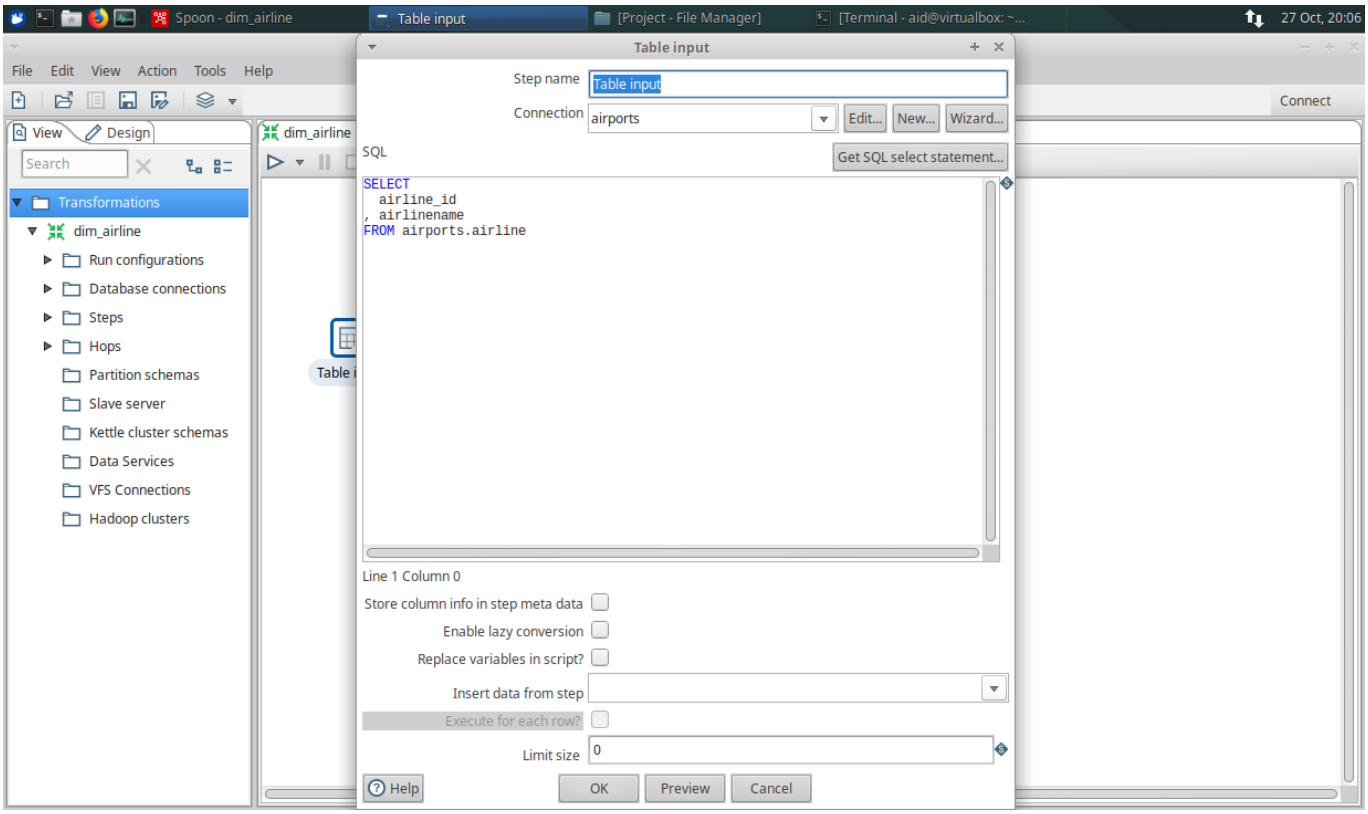

Figure 2 - dim_airline table input window

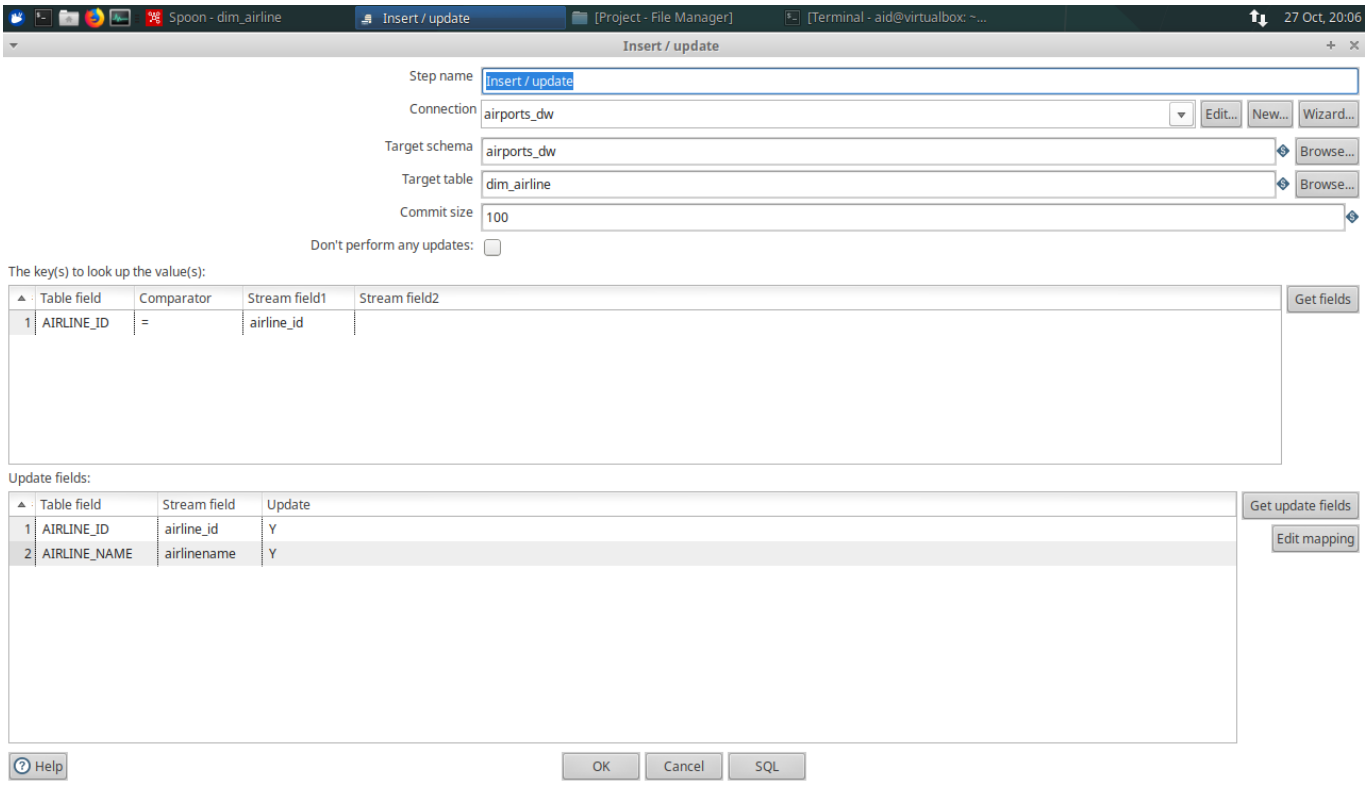

Figure 3 - dim_airline insert/update window

Airplane dimension

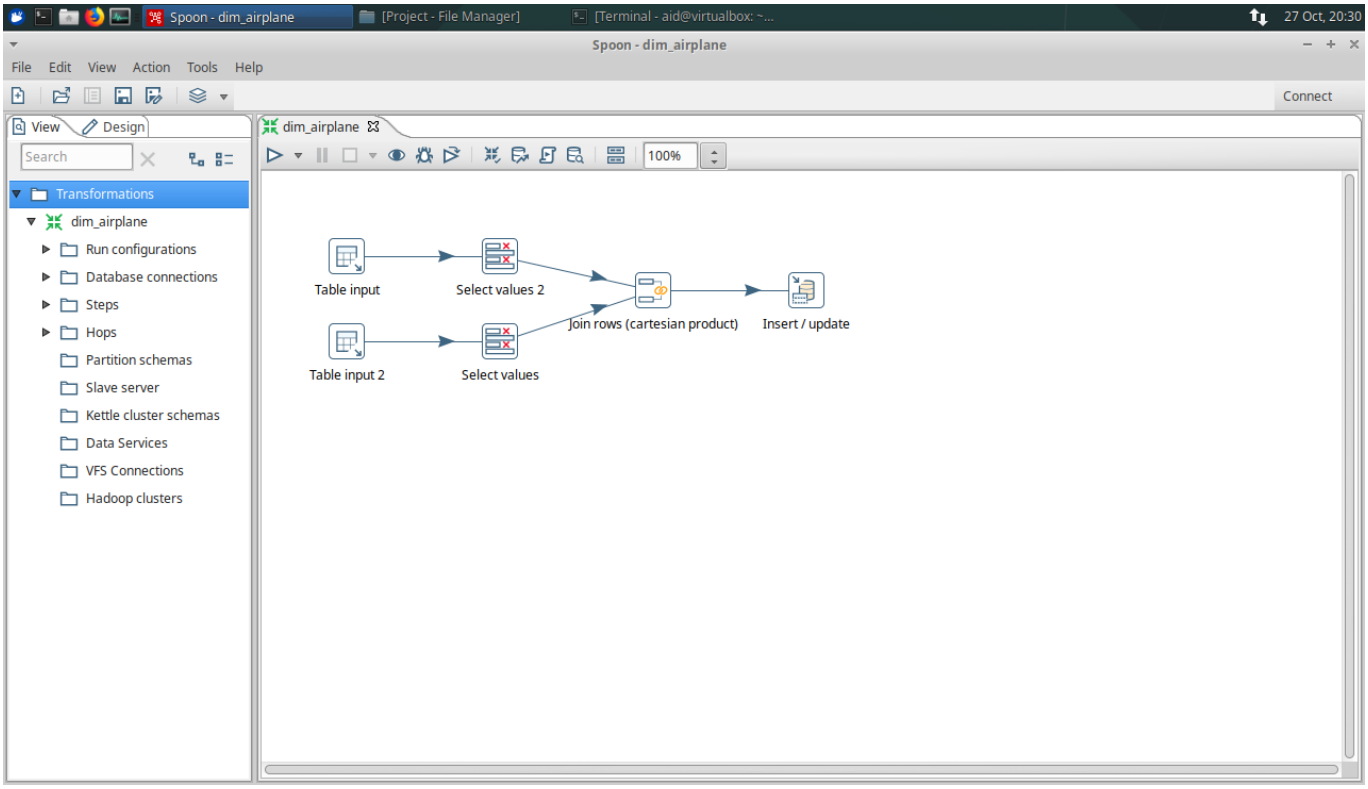

Figure 4 - dim_airplane entire transformation

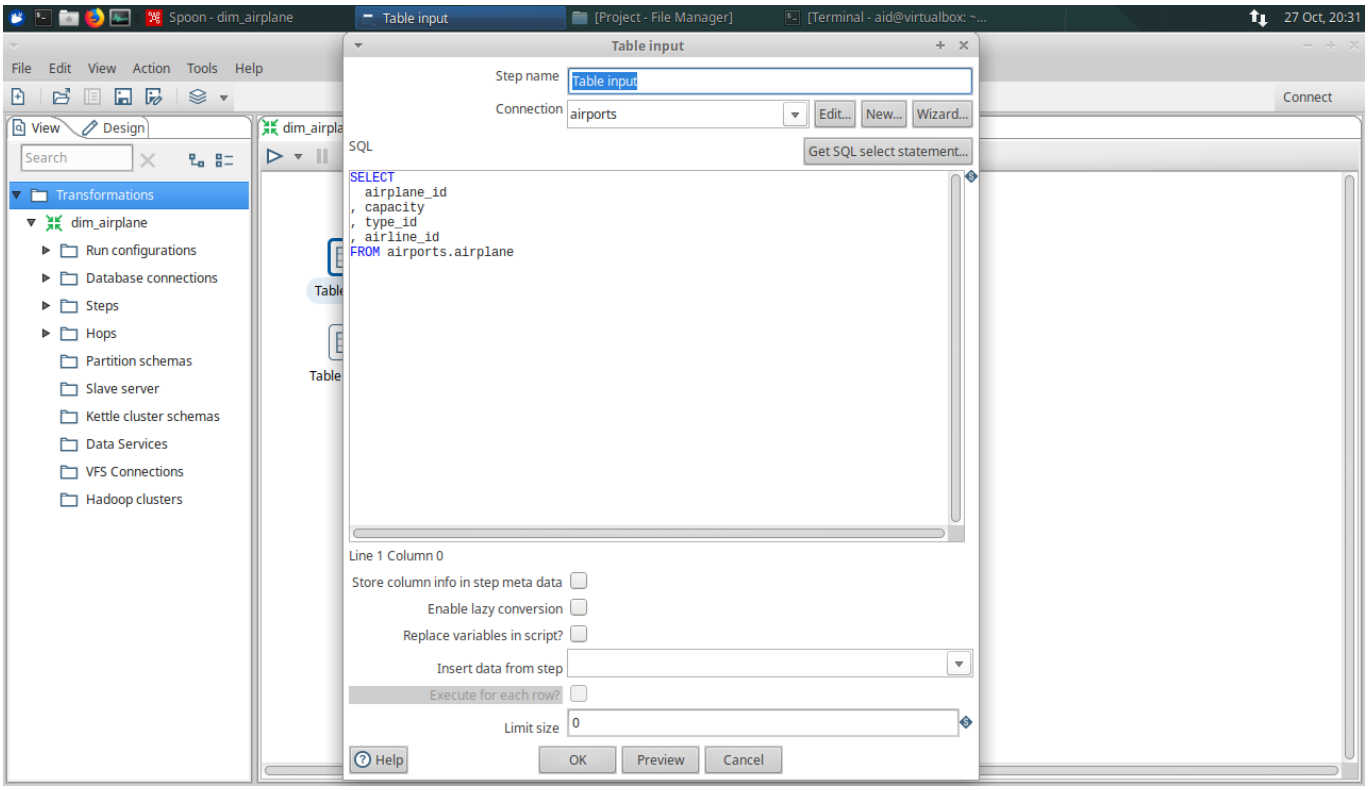

Figure 5 - dim_airplane table input window

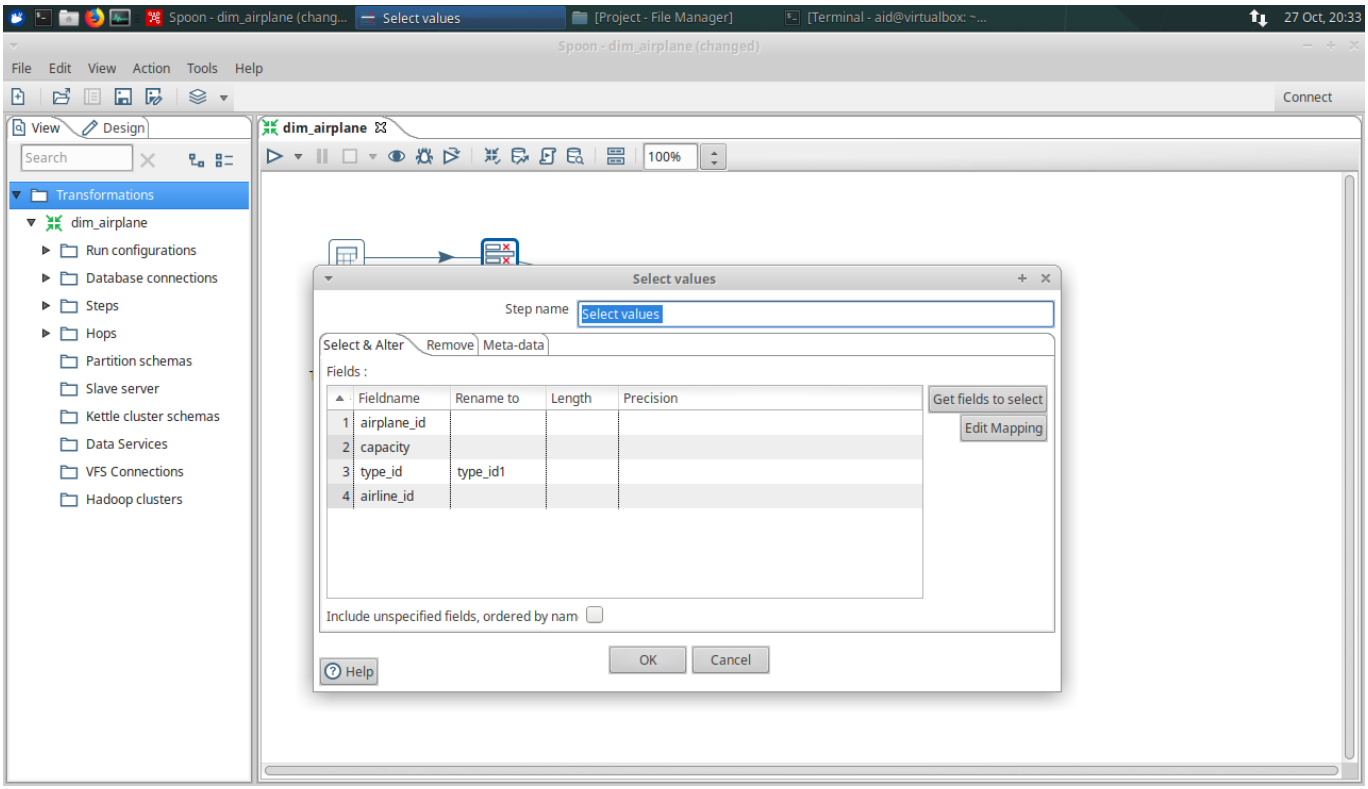

Figure 6 - dim_airplane select values window

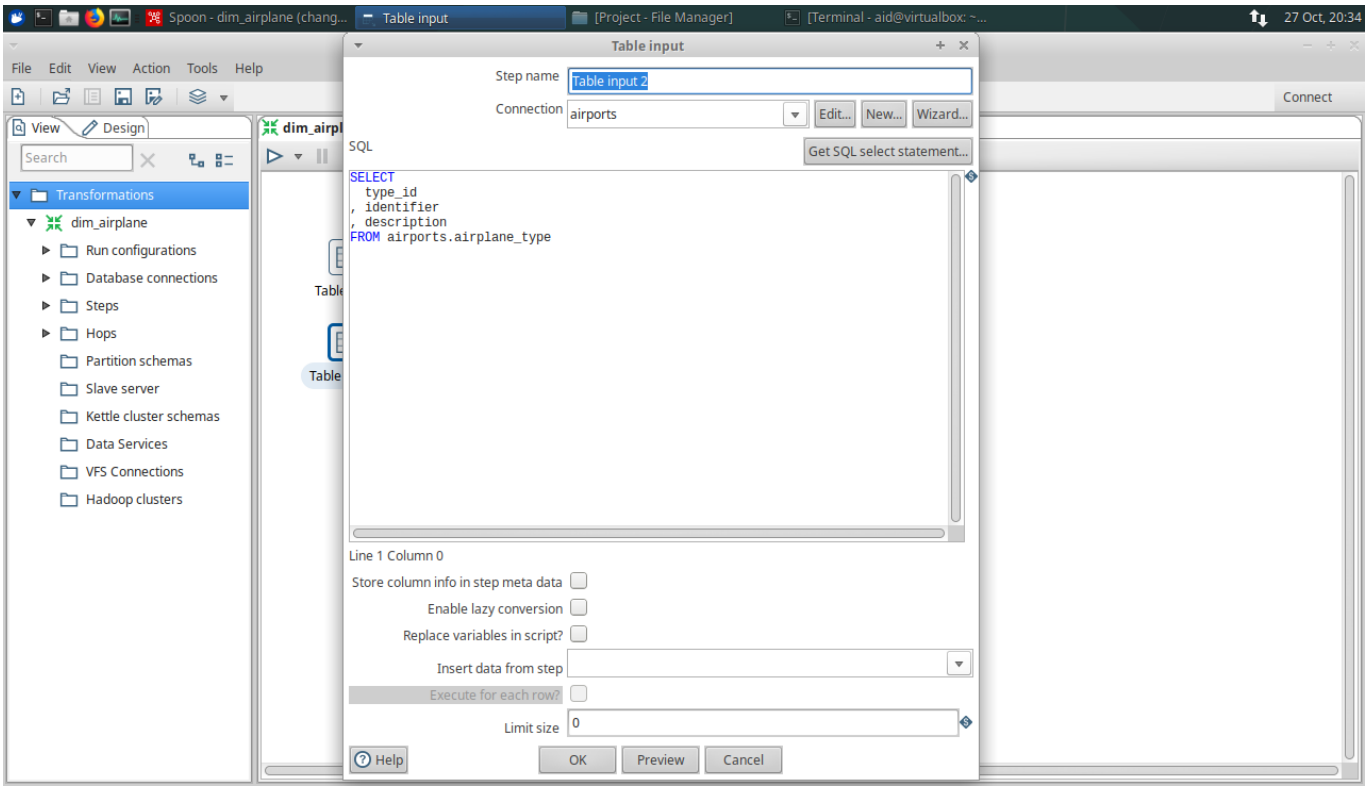

Figure 7 - dim_airplane table input 2 window

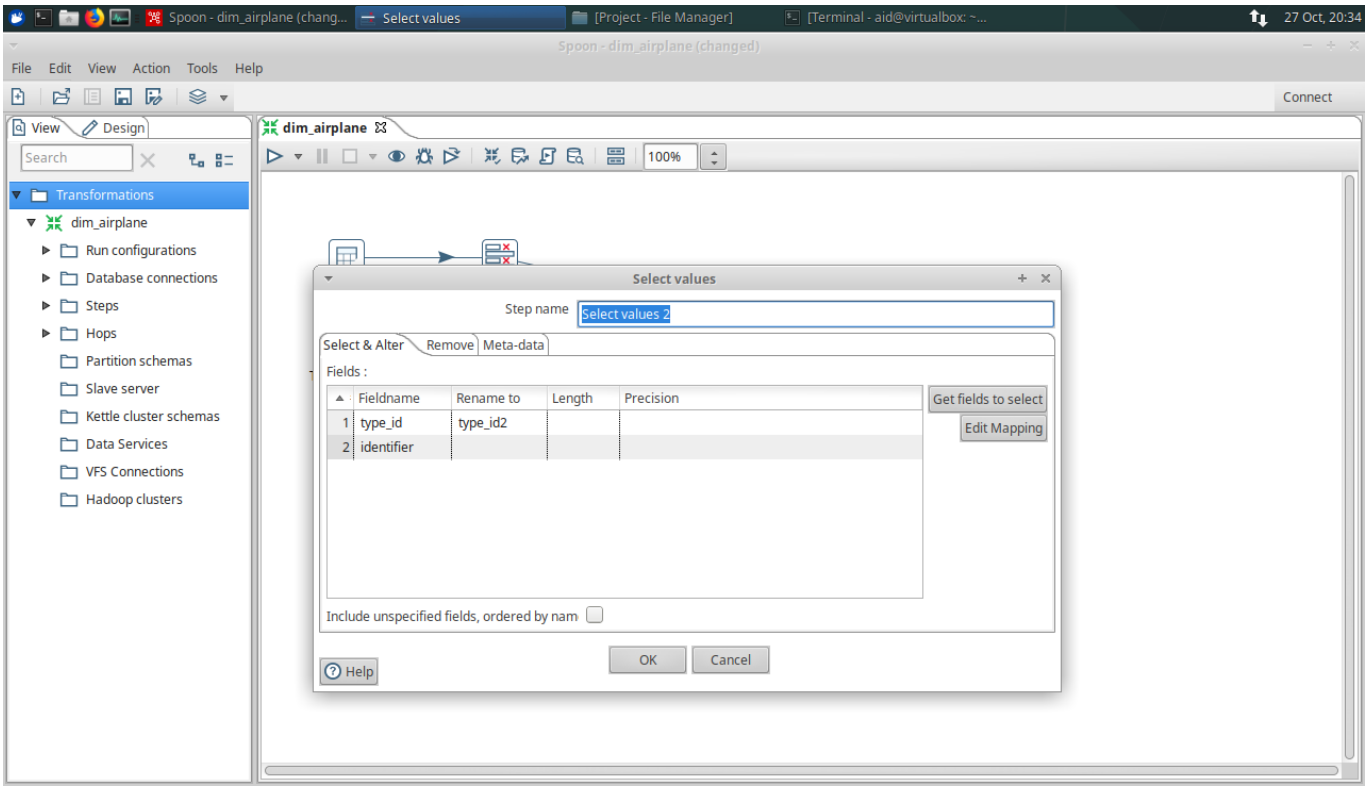

Figure 8 - dim_airplane select values 2 window

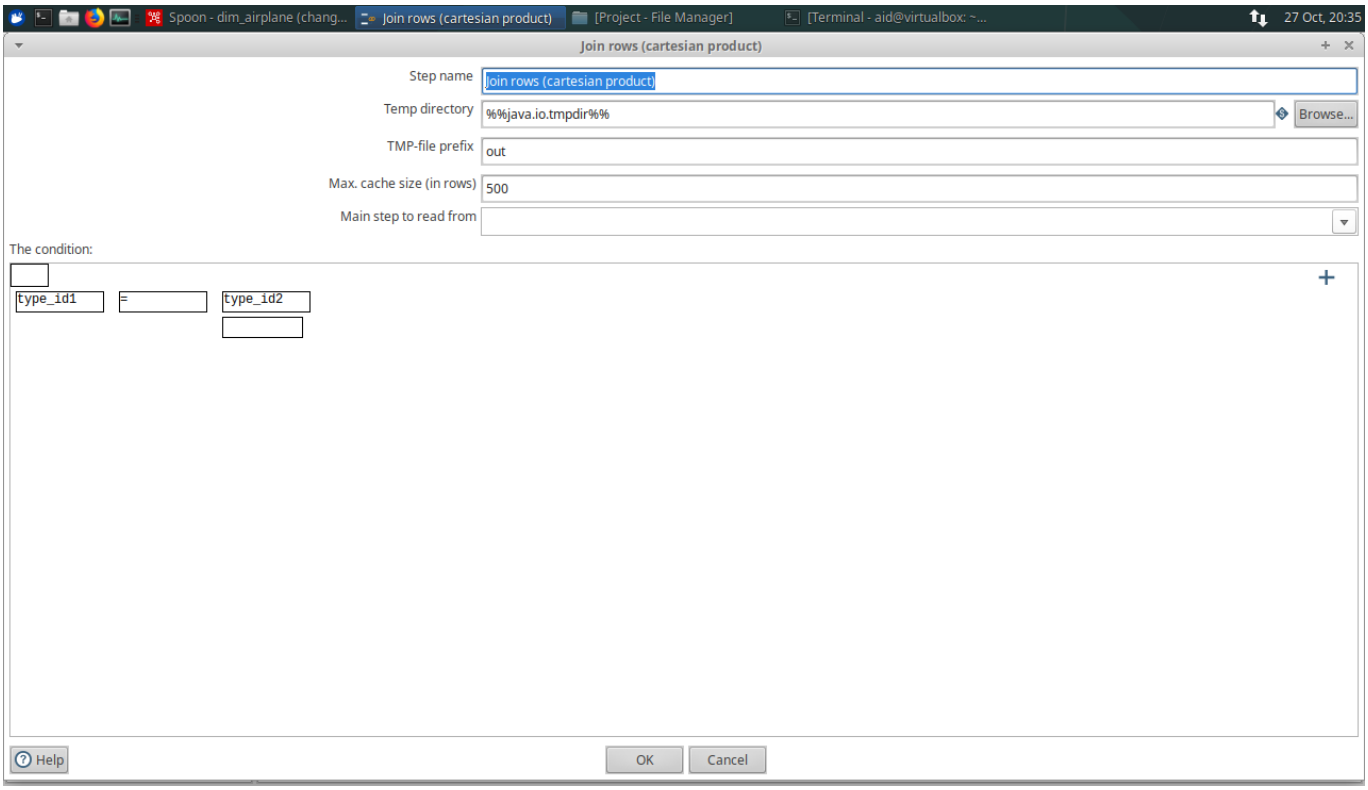

Figure 9 - dim_airplane join rows window

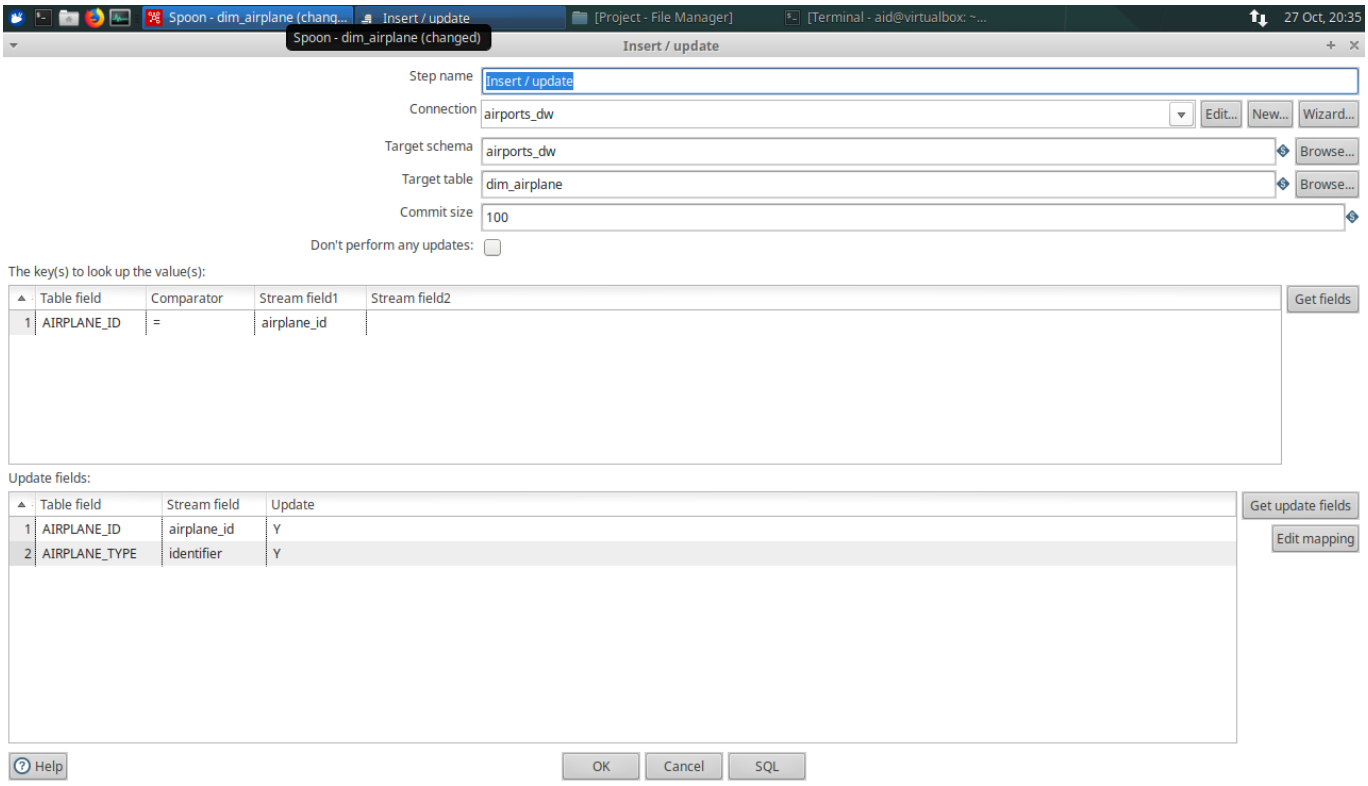

Figure 10 - dim_airplane insert/update window

Airport dimension

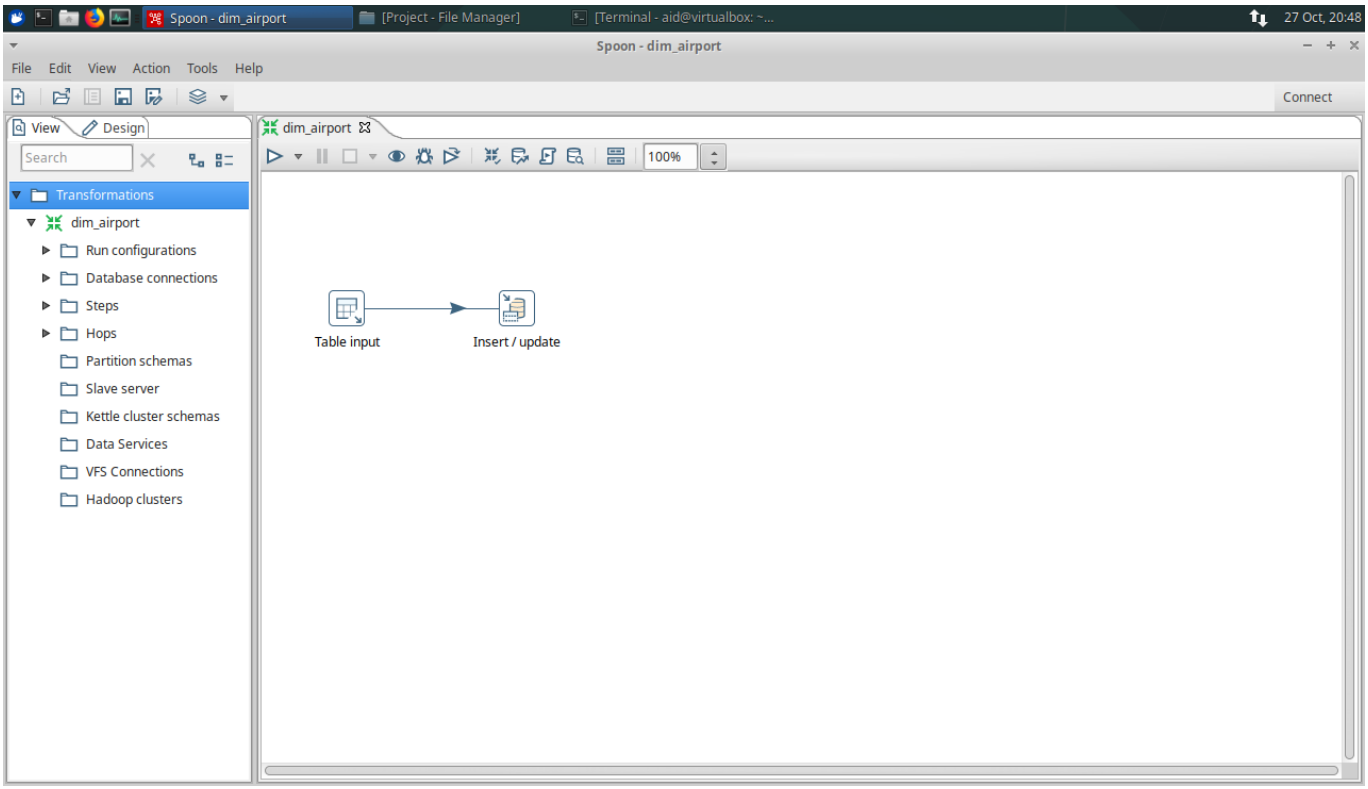

Figure 11 - dim_airport entire transformation

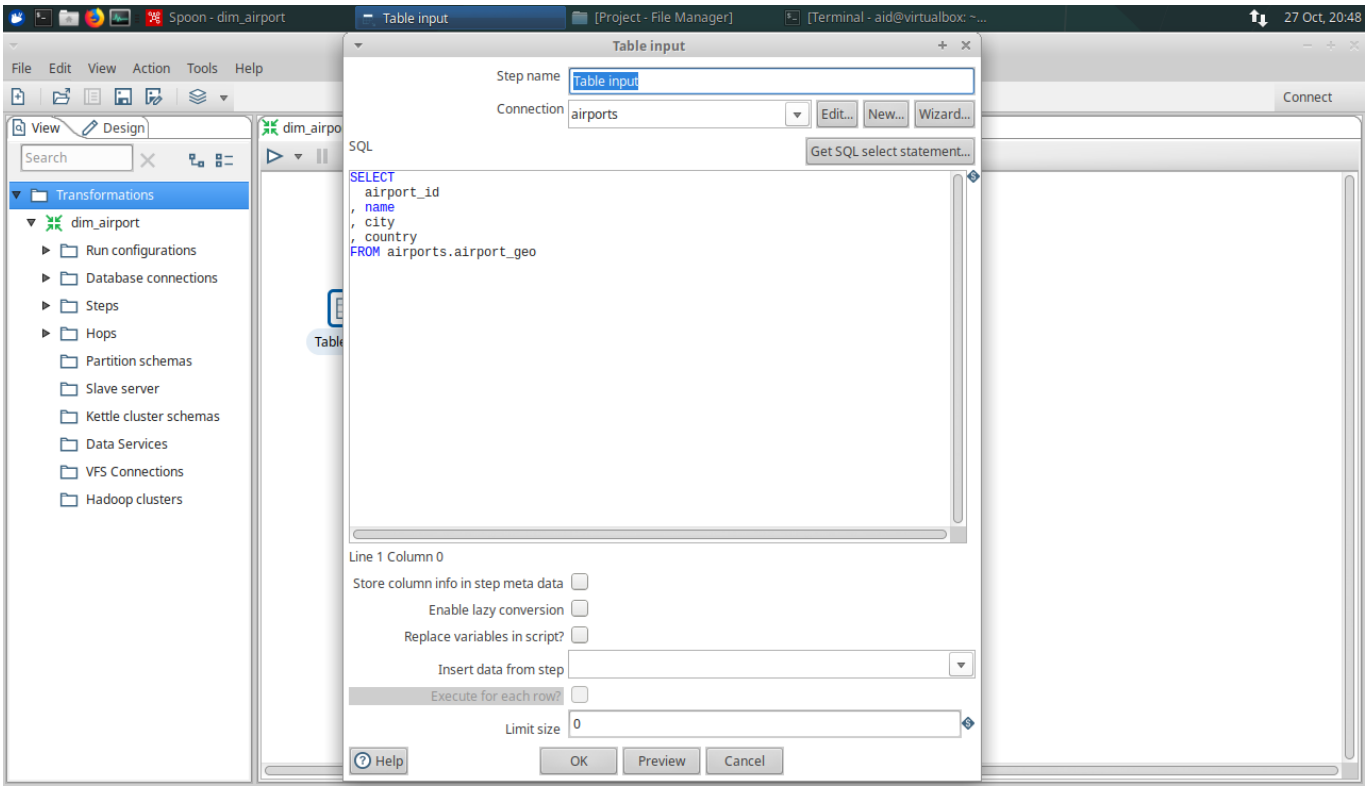

Figure 12 - dim_airport input table window

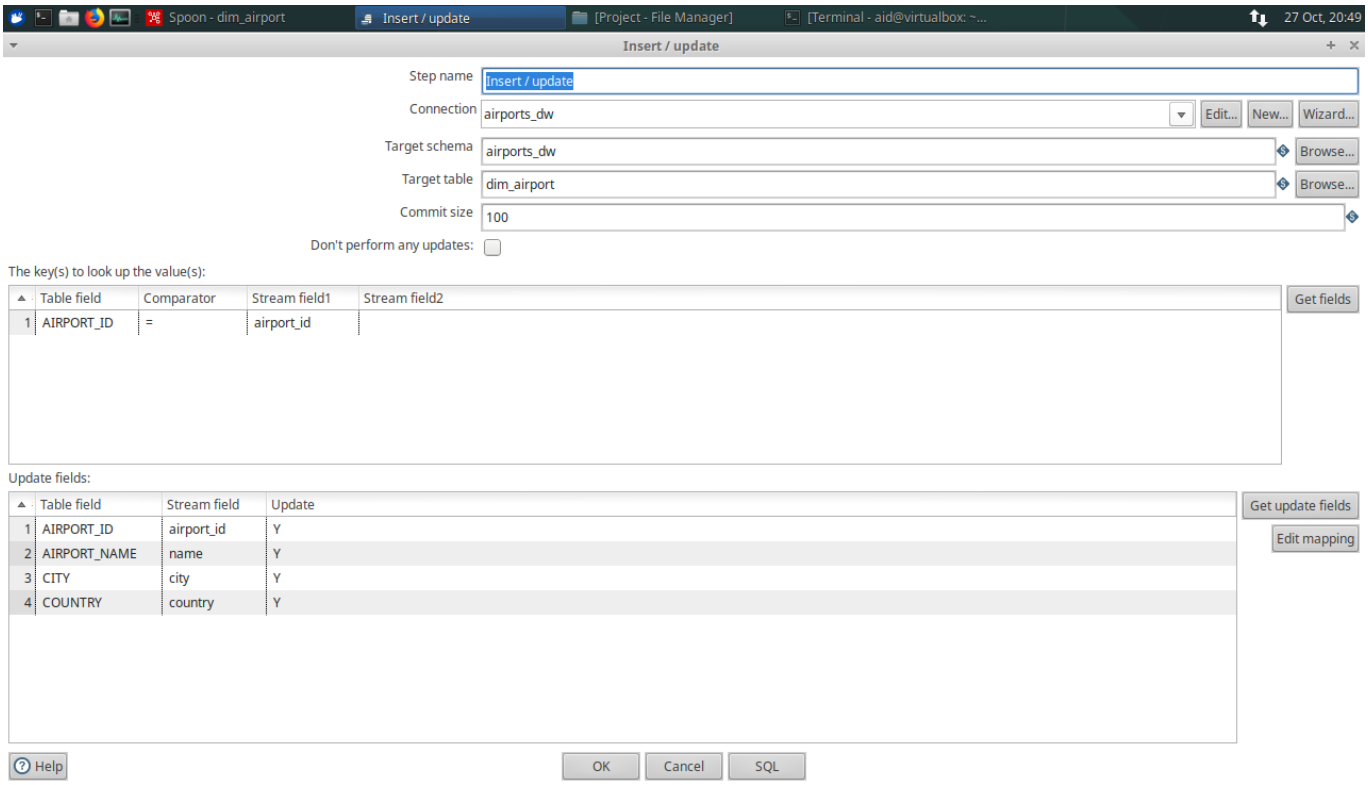

Figure 13 - dim_airport insert/update window

Time dimension

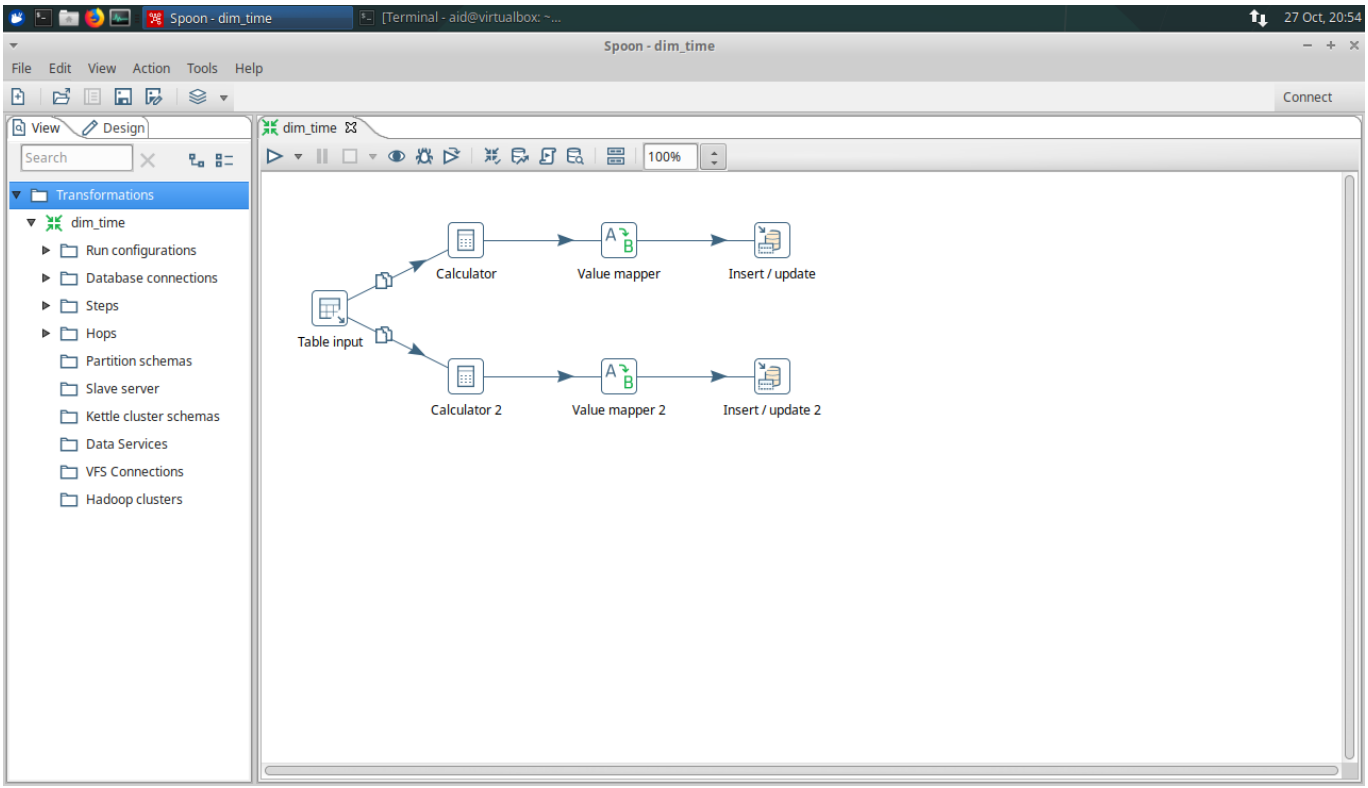

Figure 14 - dim_time entire transformation

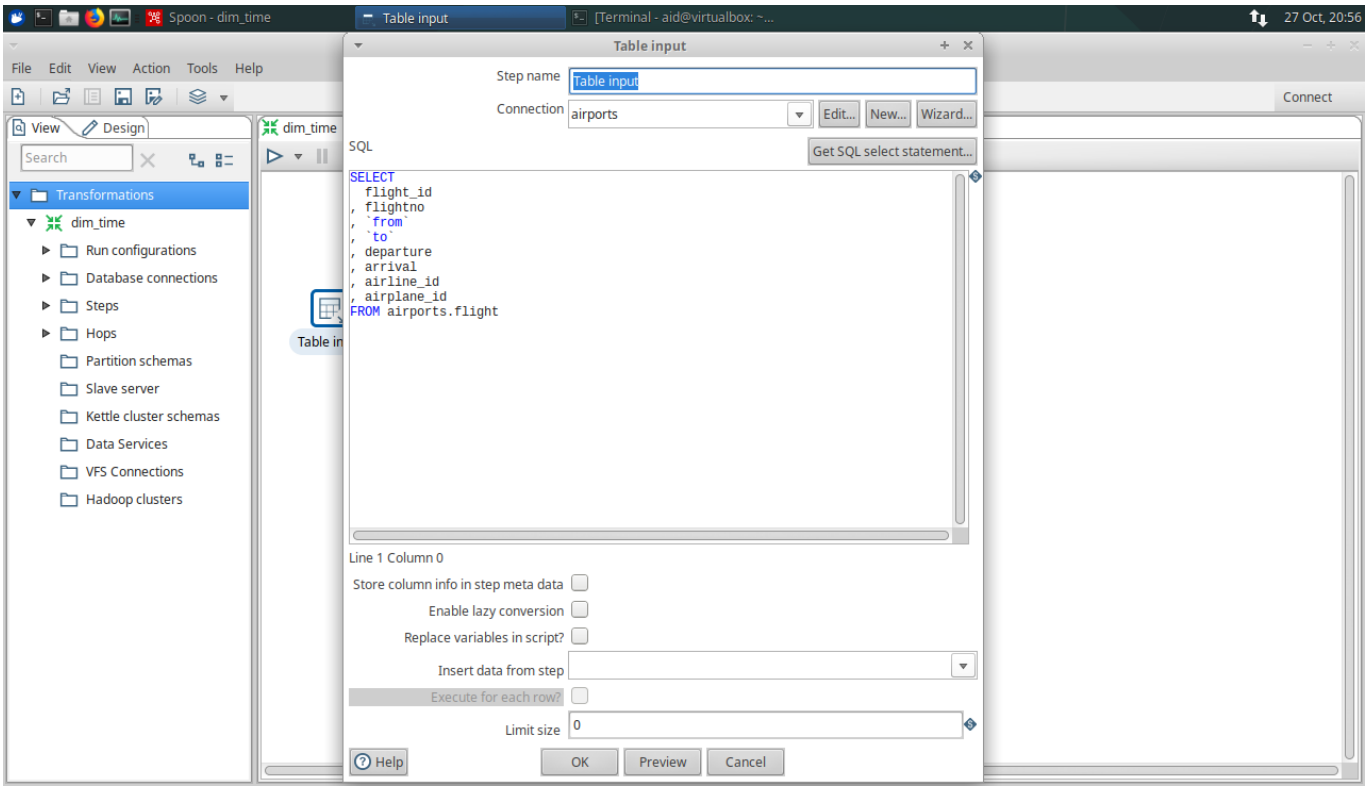

Figure 15 - dim_time table input window

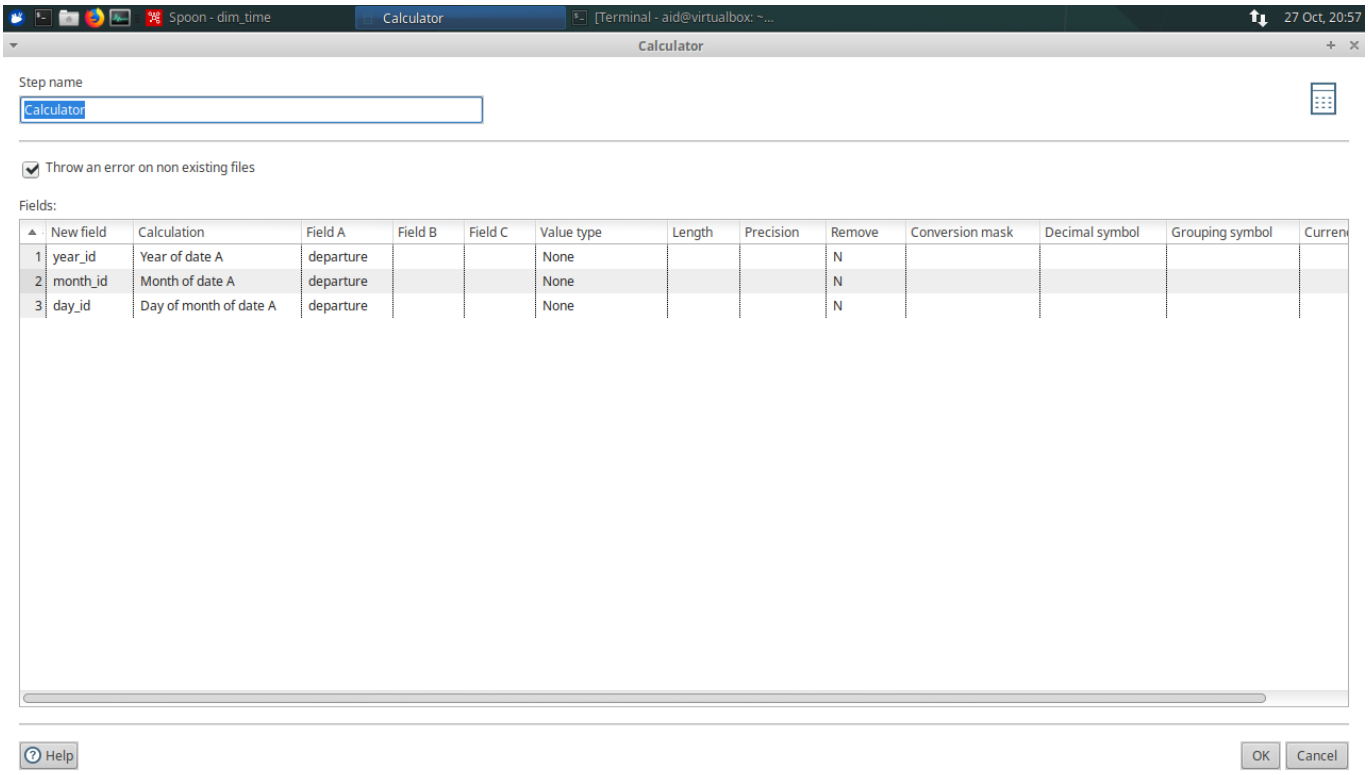

Figure 16 - dim_time calculator window

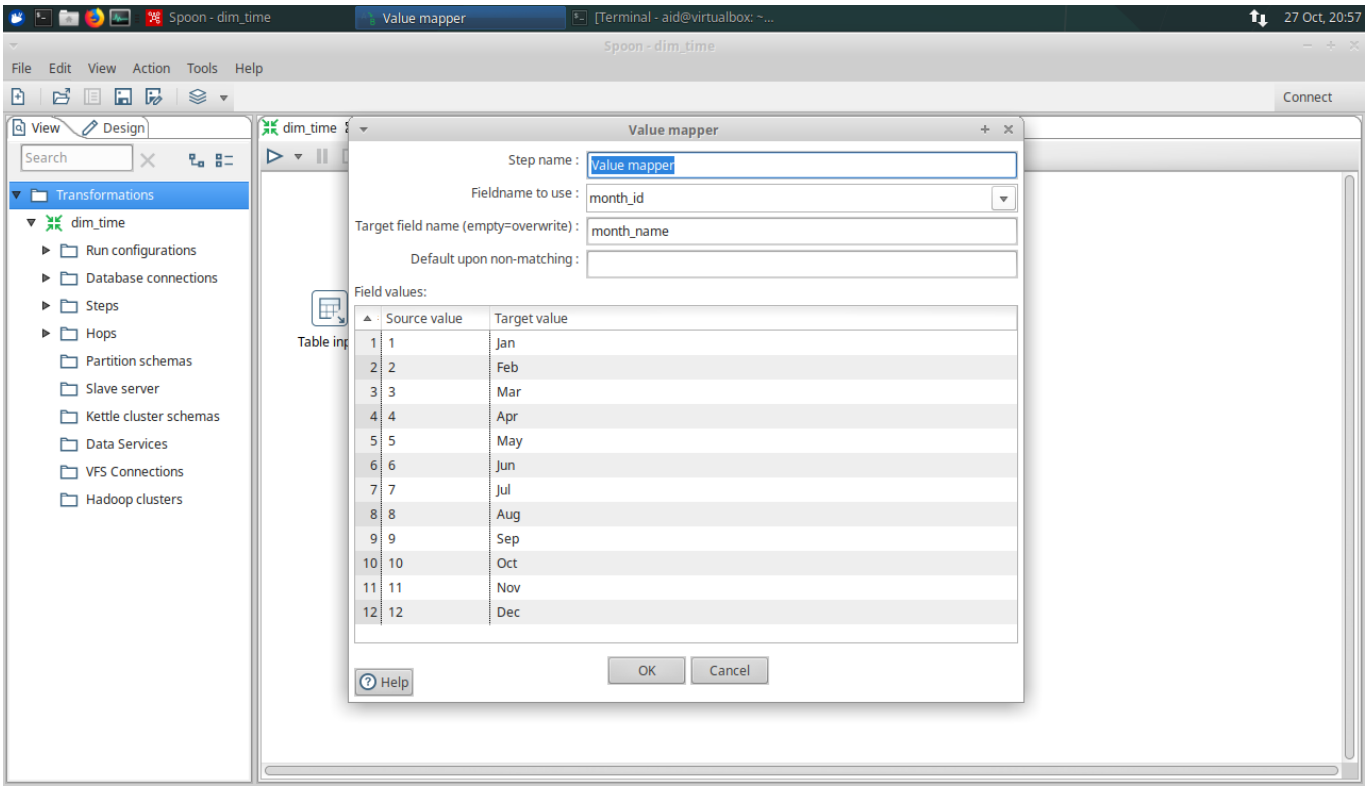

Figure 17 - dim_time value mapper window

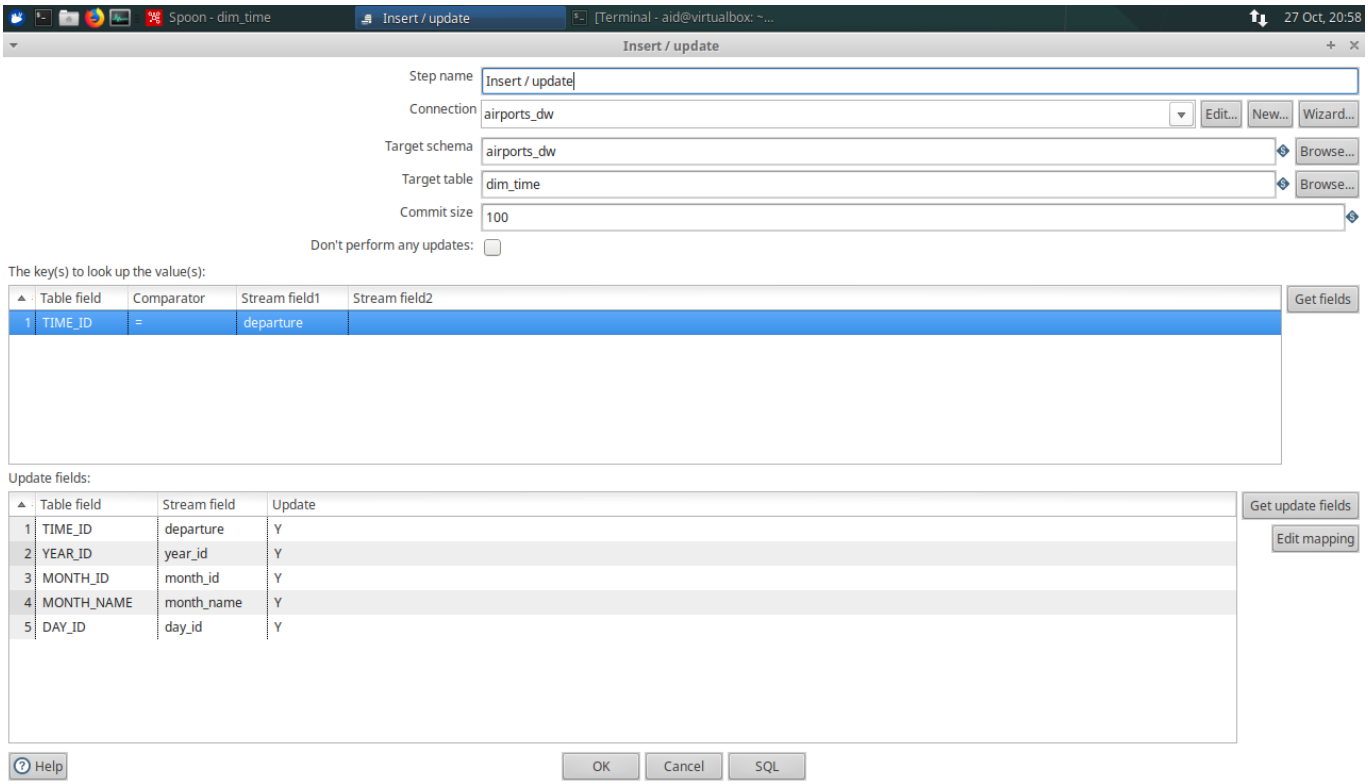

Figure 18 - dim_time insert/update window (departure time)

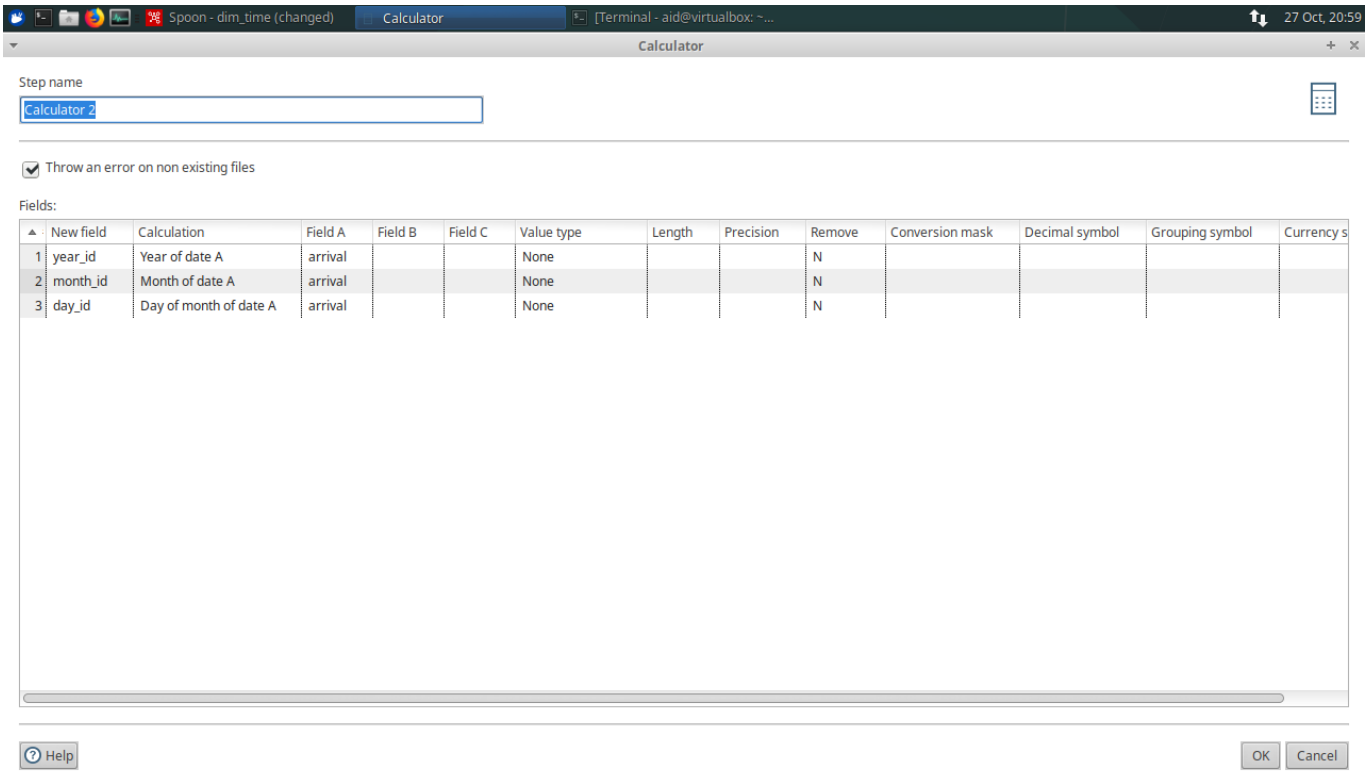

Figure 19 - dim_time calculator 2 window

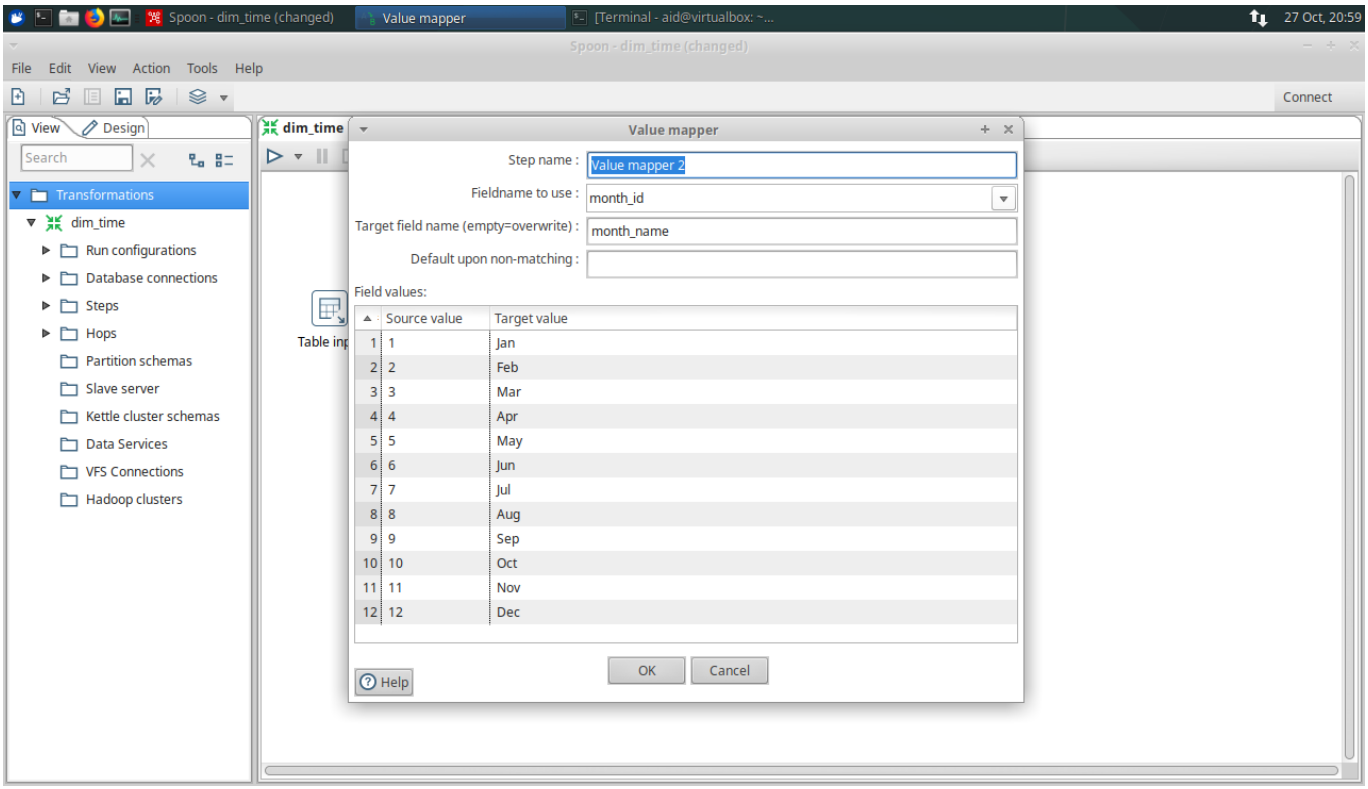

Figure 20 - dim_time value mapper 2 window

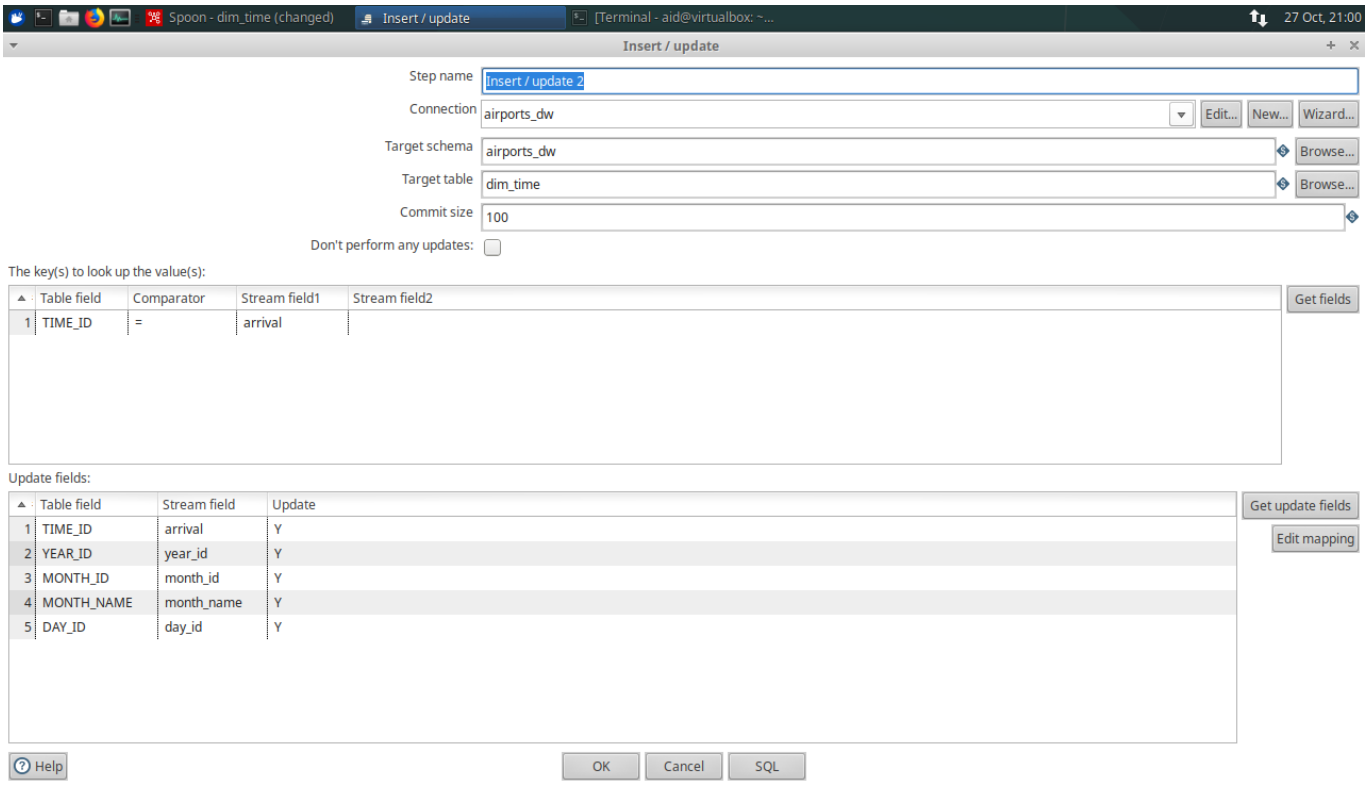

Figure 21 - dim_time insert/update 2 window (arrival time)

Flight fact

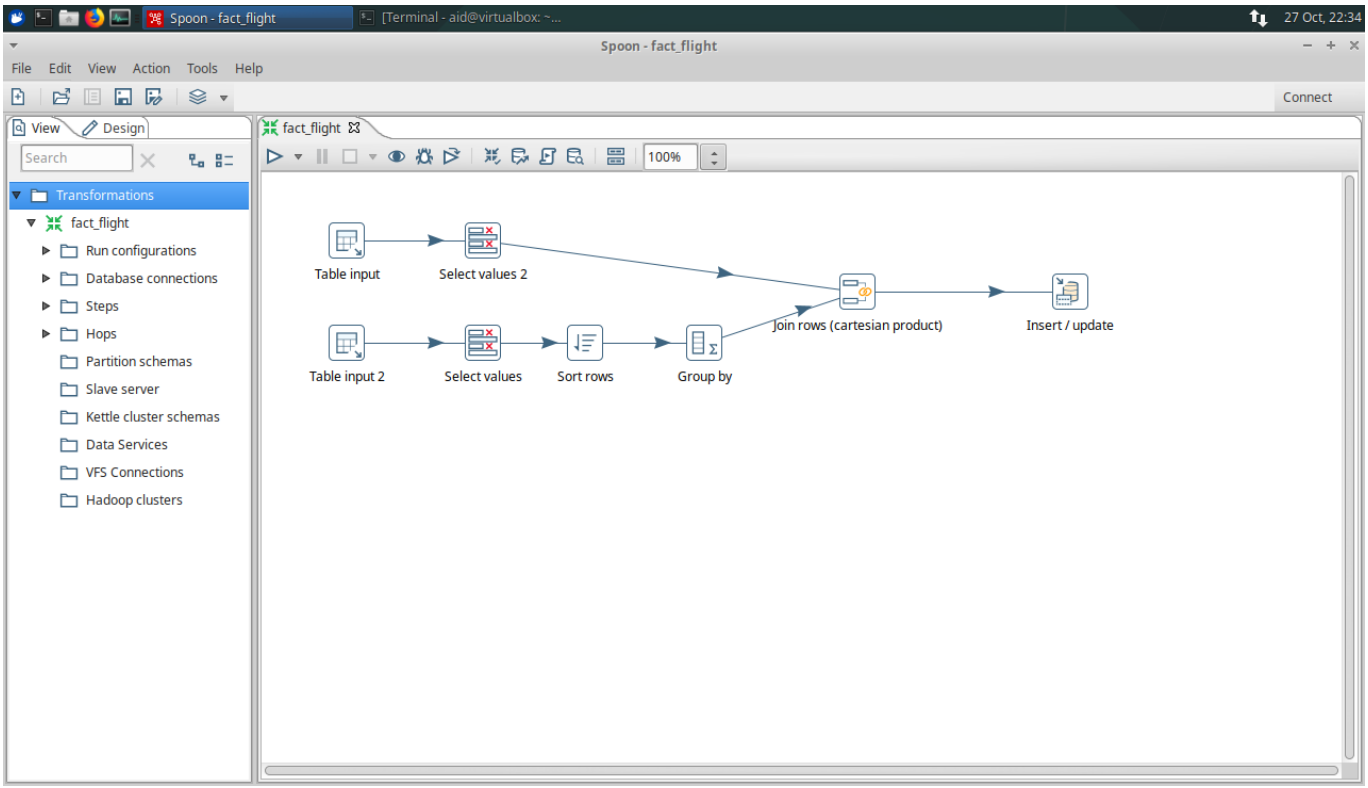

Figure 22 - fact_flight entire transformation

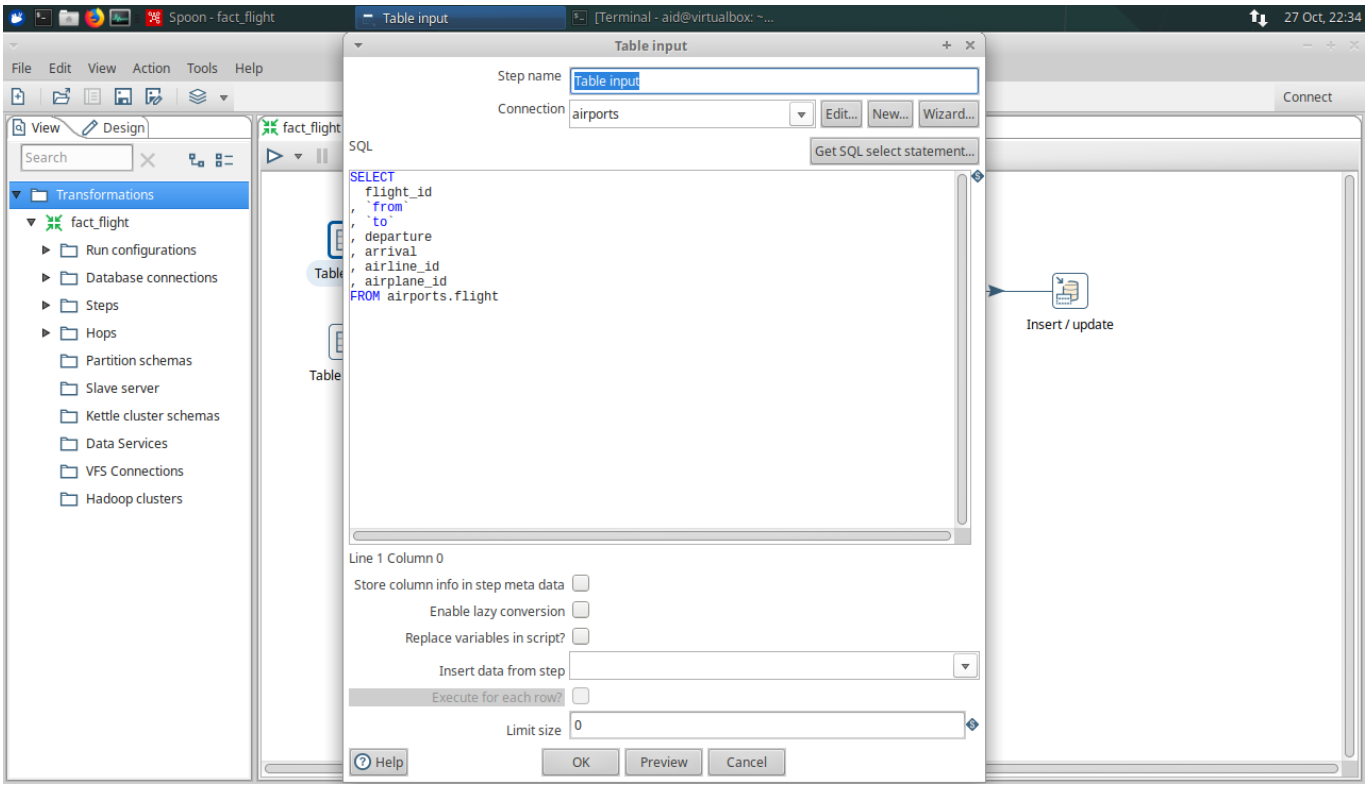

Figure 23 - fact_flight table input window

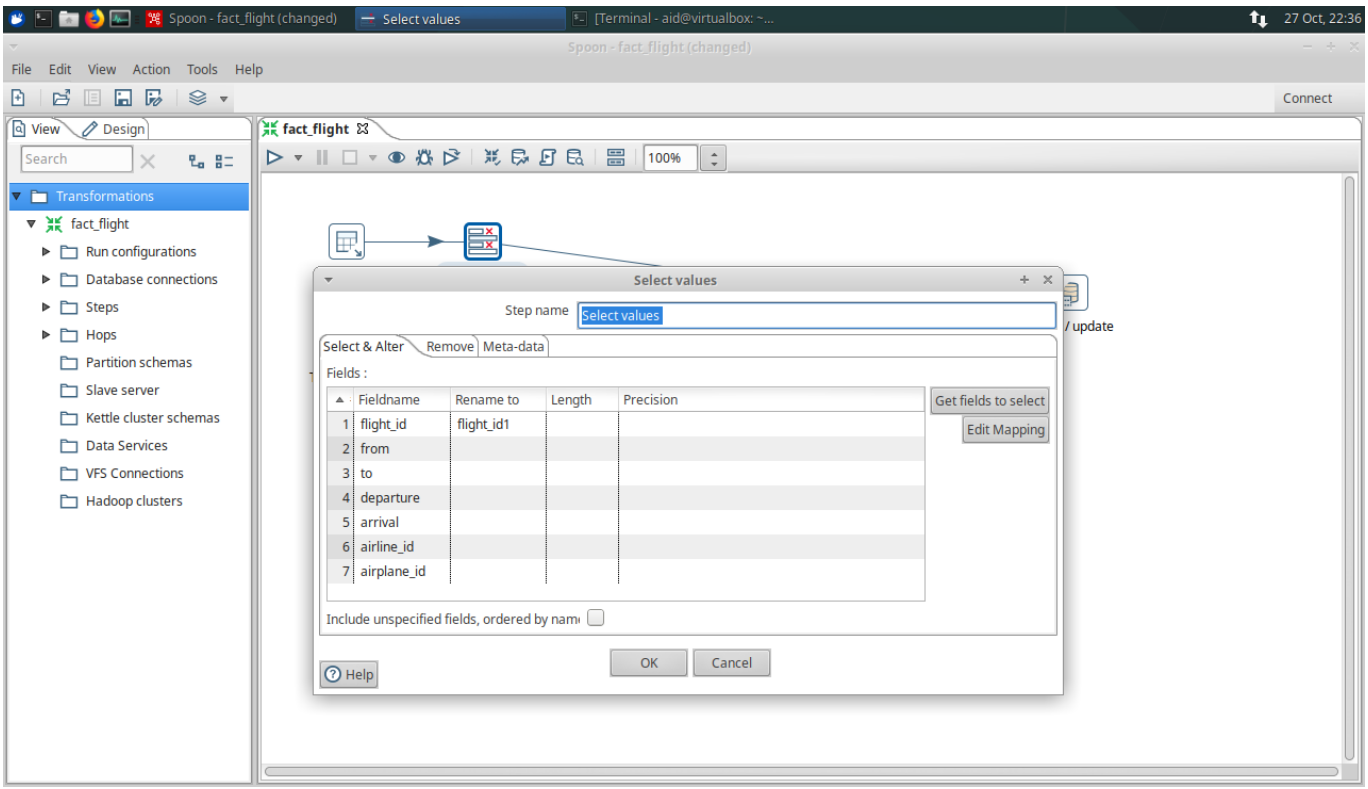

Figure 24 - fact_flight select values window

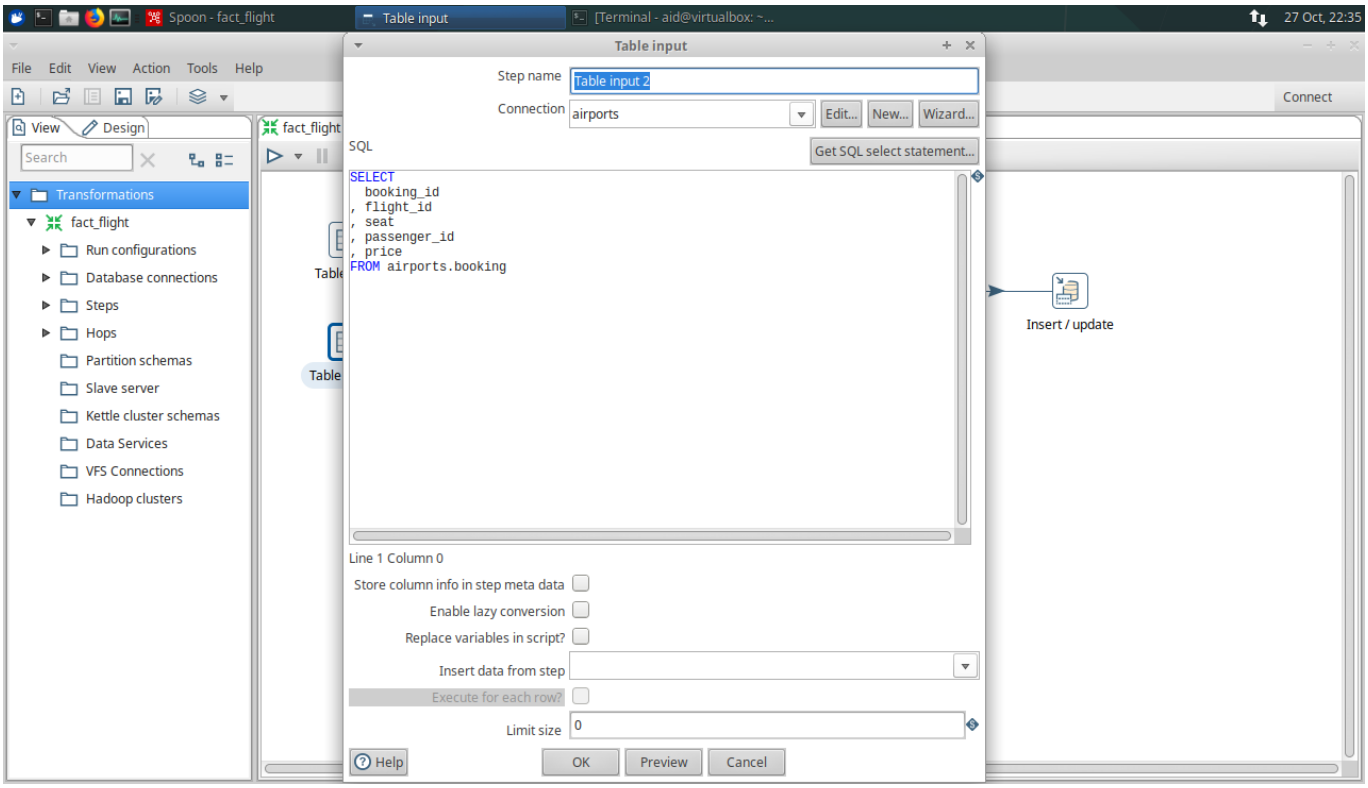

Figure 25 - fact_flight table input 2 window

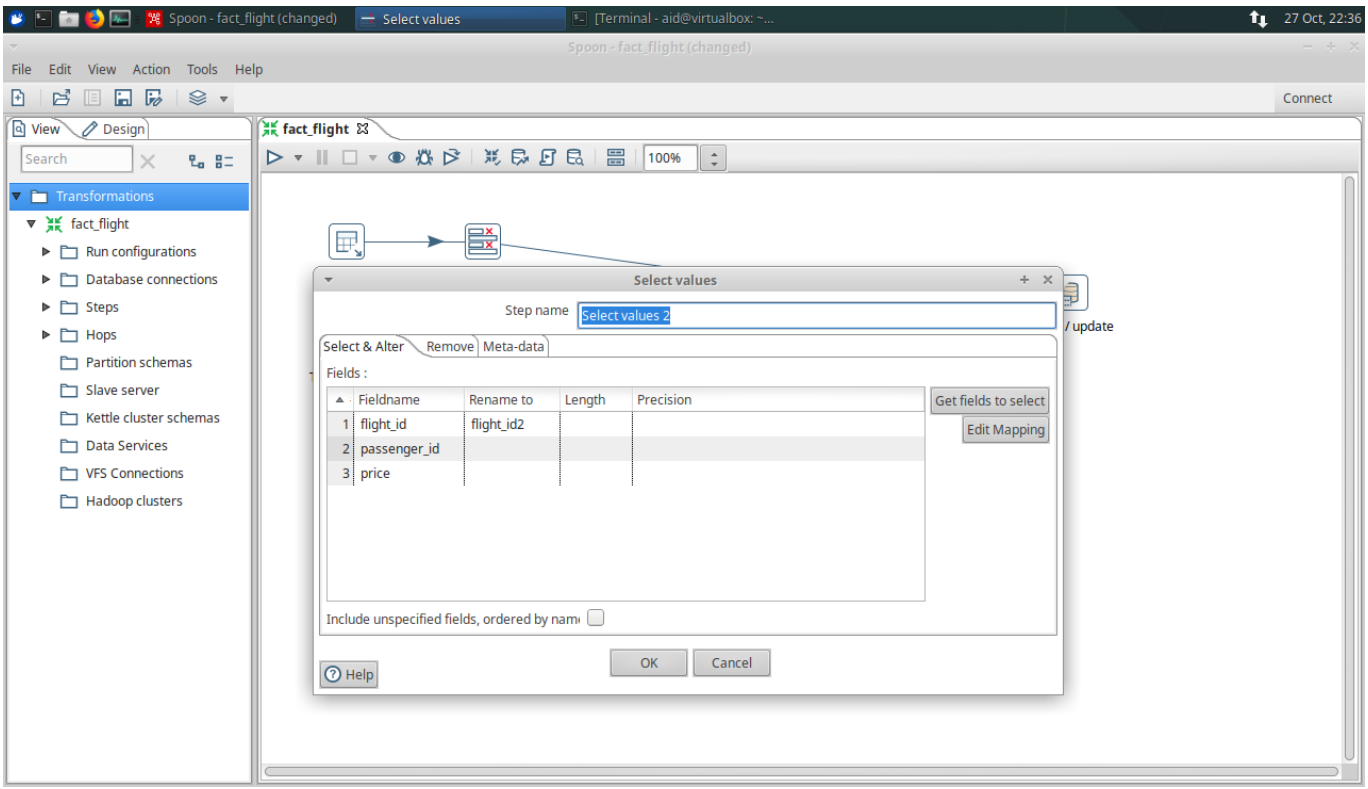

Figure 26 - fact_flight select values 2 window

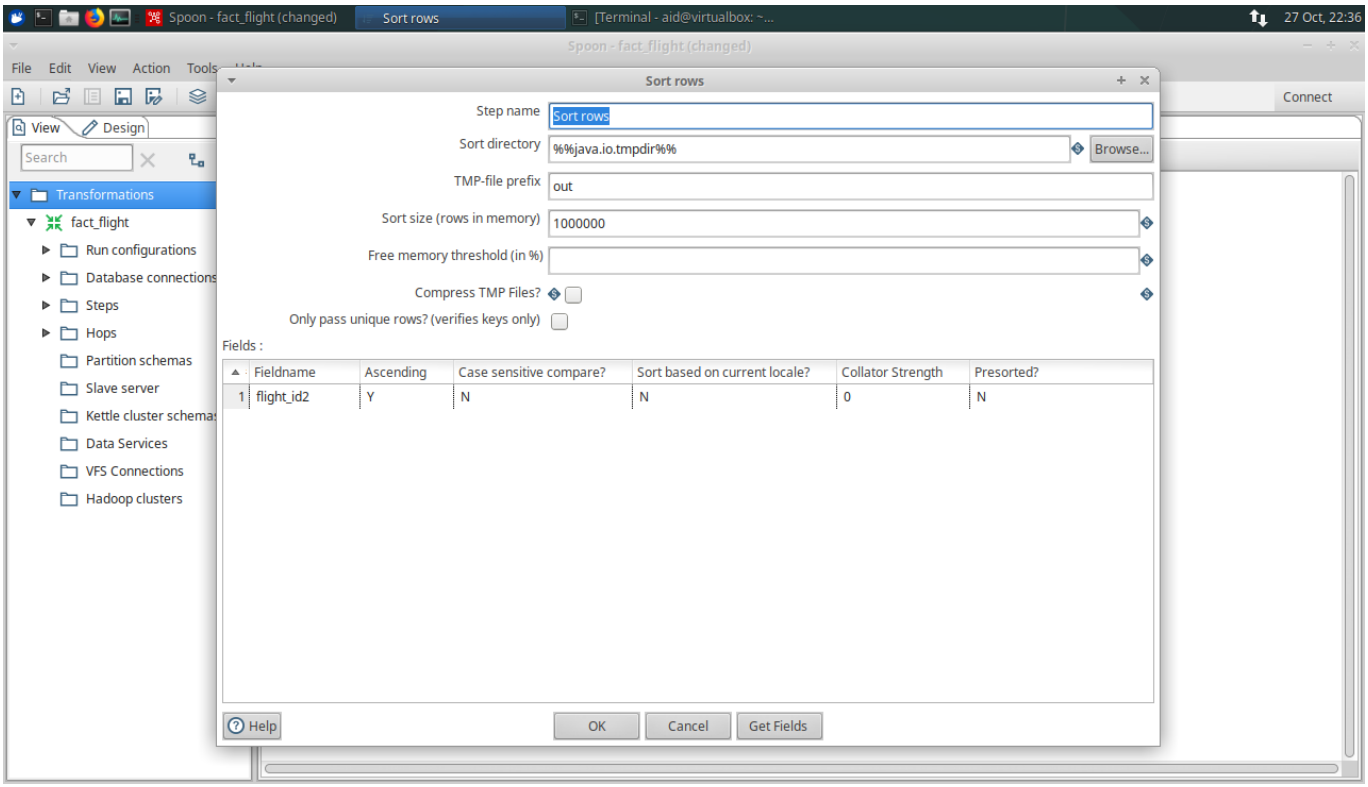

Figure 27 - fact_flight sort rows window

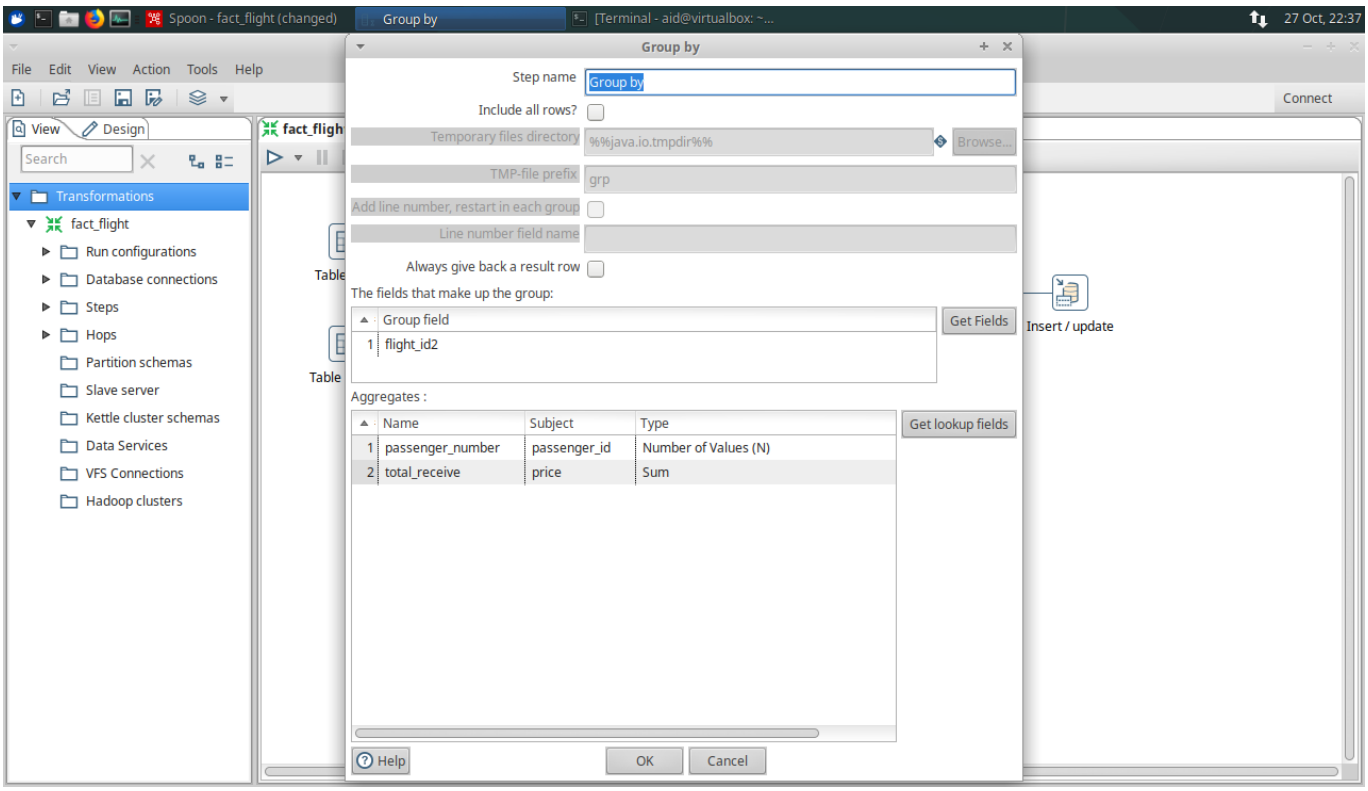

Figure 28 - fact_flight group by window (number of passengers and revenue)

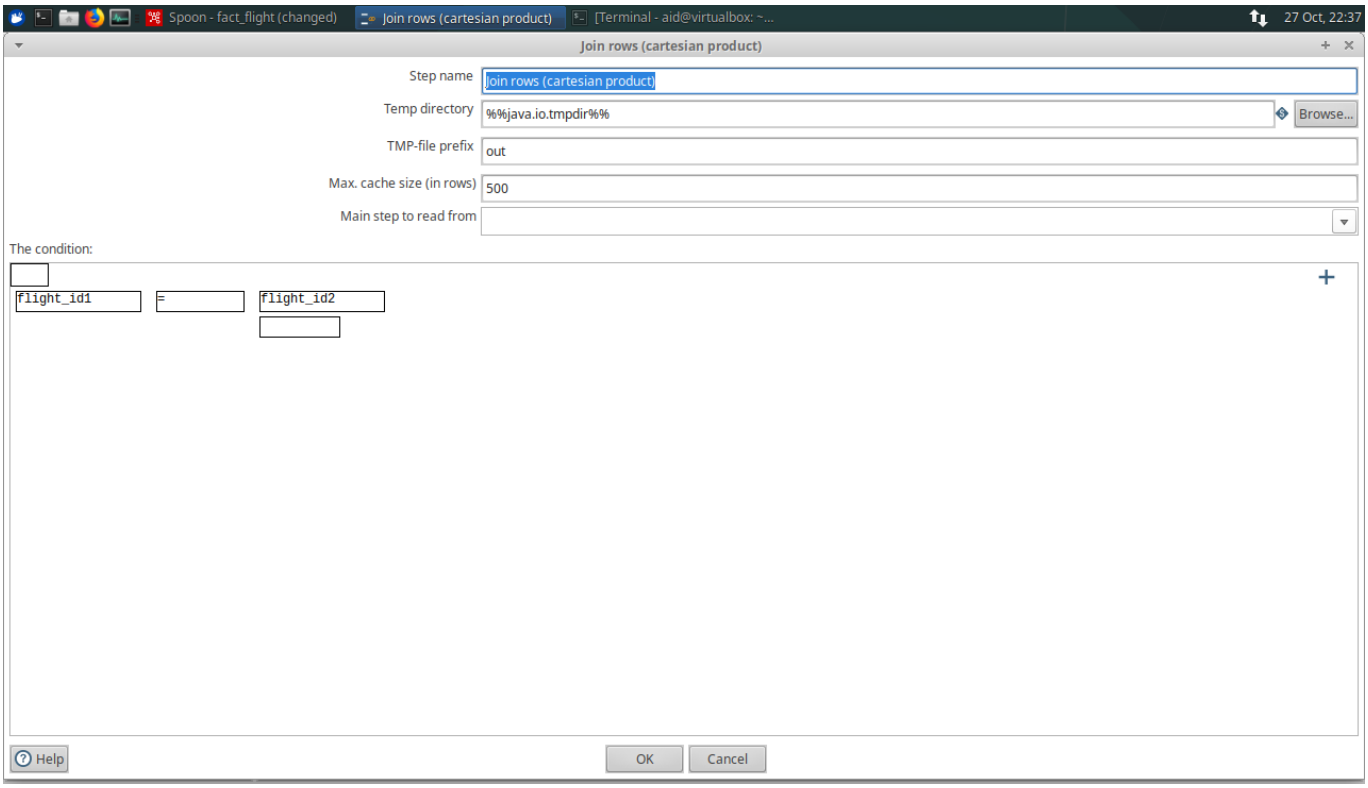

Figure 29 - fact_flight join rows window

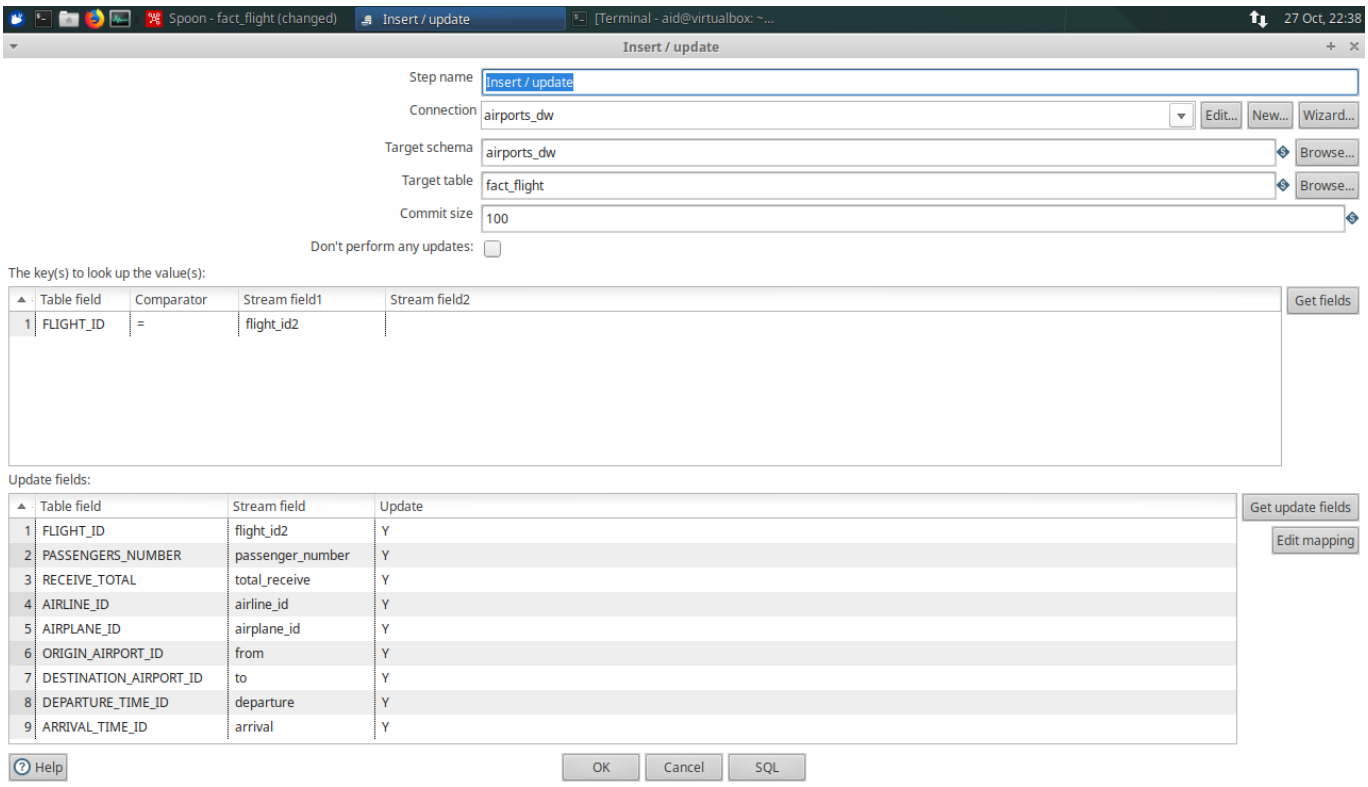

Figure 30 - fact_flight insert/update window

III - XML Code for the cube definition

```
<Schema name="airports_dw">
  <Cube name="Flights" visible="true" cache="true" enabled="true">
    <Table name="fact_flight">
    </Table>
```

```
    <Dimension type="StandardDimension" visible="true"
foreignKey="AIRLINE_ID" highCardinality="false" name="Airline">
      <Hierarchy name="Airline Hierarchy" visible="true" hasAll="true"
allMemberName="All Arlines" primaryKey="AIRLINE_ID">
        <Table name="dim_airline">
        </Table>
        <Level name="Name" visible="true" column="AIRLINE_NAME"
type="String" uniqueMembers="false" levelType="Regular"
hideMemberIf="Never">
        </Level>
      </Hierarchy>
    </Dimension>
    <Dimension type="StandardDimension" visible="true"
foreignKey="DESTINATION_AIRPORT_ID" highCardinality="false"
name="Airport">
      <Hierarchy name="Airport Hierarchy" visible="true" hasAll="true"
allMemberName="All Airports" primaryKey="AIRPORT_ID">
        <Table name="dim_airport">
        </Table>
        <Level name="Country" visible="true" column="COUNTRY"
type="String" uniqueMembers="false" levelType="Regular"
hideMemberIf="Never">
        </Level>
        <Level name="City" visible="true" column="CITY" type="String"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never">
        </Level>
    <Level name="Name" visible="true" column="AIRPORT_NAME" type="String"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never">
        </Level>
      </Hierarchy>
    </Dimension>
    <Dimension type="TimeDimension" visible="true"
foreignKey="DEPARTURE_TIME_ID" highCardinality="false" name="Time">
      <Hierarchy name="Time Hierarchy" visible="true" hasAll="true"
allMemberName="All Years" primaryKey="TIME_ID">
        <Table name="dim_time">
        </Table>
        <Level name="Year" visible="true" column="YEAR_ID" type="Integer"
uniqueMembers="false" levelType="TimeYears" hideMemberIf="Never">
        </Level>
        <Level name="Month" visible="true" column="MONTH_NAME"
ordinalColumn="MONTH_ID" type="String" uniqueMembers="false"
levelType="TimeMonths" hideMemberIf="Never">
        </Level>
        <Level name="Day" visible="true" column="DAY_ID" type="Integer"
uniqueMembers="false" levelType="TimeDays" hideMemberIf="Never">
        </Level>
      </Hierarchy>
    </Dimension>
    <Dimension type="StandardDimension" visible="true"
foreignKey="AIRPLANE_ID" name="Airplane">
      <Hierarchy name="Airplane Hierarchy" visible="true" hasAll="true"
allMemberName="All Airplanes" primaryKey="AIRPLANE_ID">
        <Table name="dim_airplane">
```

```
        </Table>
        <Level name="Type" visible="true" column="AIRPLANE_TYPE"
type="String" uniqueMembers="false" levelType="Regular">
        </Level>
      </Hierarchy>
    </Dimension>
    <Measure name="NumPassengers" column="PASSENGERS_NUMBER"
datatype="Integer" formatString="#,###" aggregator="sum" visible="true">
    </Measure>
    <Measure name="TotalReceive" column="RECEIVE_TOTAL" datatype="Numeric"
formatString="$ #,###.00" aggregator="sum" visible="true">
    </Measure>
  </Cube>
</Schema>
```