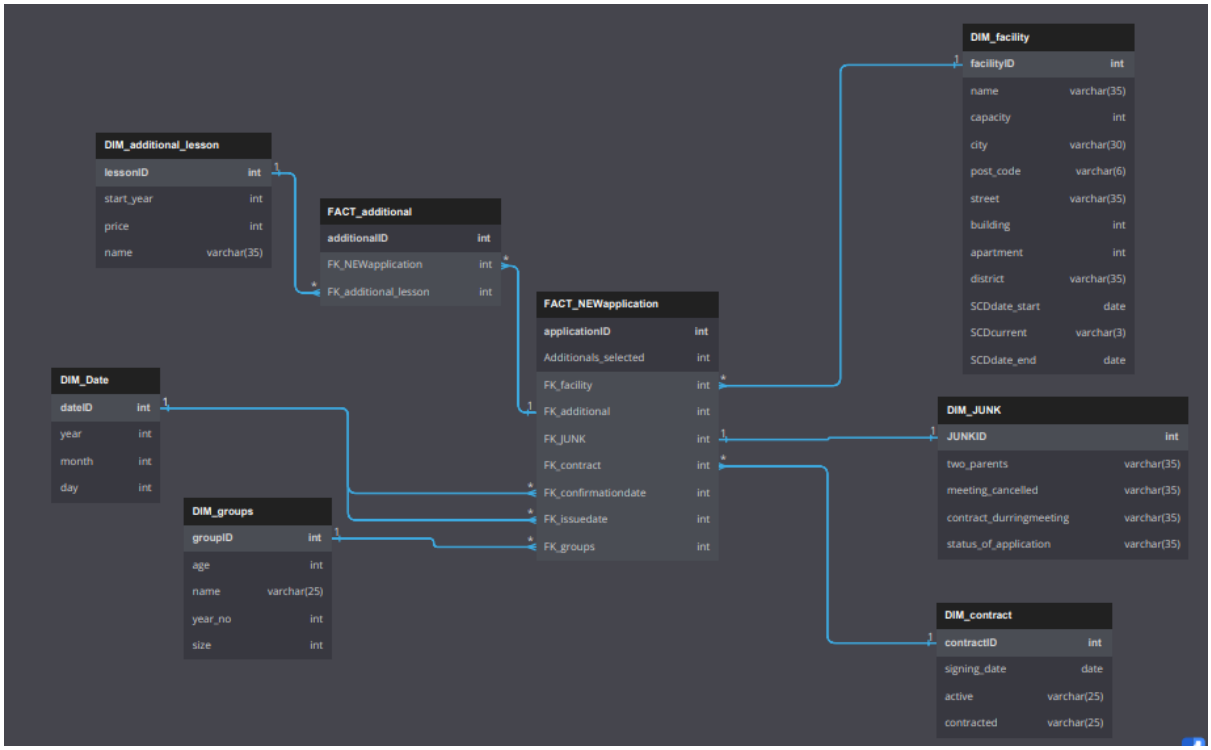
Kamil Dędza
Agnieszka Kuleta

GIRAFFE KINDERGARDEN DATAWAREHOUSE DESIGN

Business process

The Date warehouse is designed for enrolment to the kindergarten business process. This process is described in the document Specification of business processes in Giraffe's kindergarten network.

Relational Database schema



FACT TABLE : FACT_NEWapplication		
NAME	TYPE	DESCRIPTION
applicationID	numeric	Unique application identifier **PK**
Additionals_selected	Numeric	Number of additional lessons selected
FK_facility	numeric	Unique facility identifier
FK_contract	numeric	Unique contract identifier
FK_confirmationdate	numeric	Date of receiving the confirmation unique identifier
FK_issuedate	numeric	Date of issue the application unique identifier
FK_groups	numeric	Unique group identifier
FK_JUNK	Numeric	Unique junk identifier

DIMENSION TABLE : DIM_JUNK		
NAME	TYPE	DESCRIPTION
JUNKID	numeric	Unique application identifier **PK**
Two_parents	Varchar	Mark whether both parents attended the meeting (Present/absent)
Meeting_cancelled	varchar	Mark whether meeting was cancelled (cancelled/not)

Kamil Dędza
Agnieszka Kuleta

Contract_duringmeeting	varchar	Mark whether contract was signed during the first meeting (signed/not)
Status_of_application	Varchar	Mark whether application is rejected or active

DIMENSION TABLE : DIM_facility (implementation of SCD2)		
NAME	TYPE	DESCRIPTION
facilityID	numeric	Unique facility identifier. **PK**
Name	Varchar	Name of facility **– business key**
City	Varchar	Name of the city
Post-code	Varchar	Number which contains 5-digits
Street	Varchar	Name of the street
Building	Varchar	Number of house
Apartment	Varchar	Number of apartment
District	Varchar	Name of district
Date_start	Date	Date of opening the facility in the given localisation **(SCD2)**
Date_end	Date	Date of closing the facility in the given localisation**(SCD2)**
Current	Varchar	Inform whether the location of facility is current or changed **(SCD2)**

DIMENSION TABLE : DIM_contract		
NAME	TYPE	DESCRIPTION
IDContract	numeric	Unique contract identifier. **PK**
Signing_date	date	Date of the signing the contract./ NULL meaning that contract was not signed.
Active	Varchar	Mark whether contract is active or inactive
Contracted	Varchar	Mark whether contract is signed or not

DIMENSION TABLE : DIM_groups		
NAME	TYPE	DESCRIPTION
groupsID	numeric	Unique group identifier. **PK**
Name	varchar	Name of group.
Age	varchar	Age of children.
Year	date	Year of existance
Size	varchar	Number of children in the group

DIMENSION TABLE : DIM_Date		
dateID	numeric	Unique date identifier. **PK**
Day	numeric	Day in the month
Month	numeric	Month in the year
Year	numeric	Year

FACT TABLE : FACT_additional		
NAME	TYPE	DESCRIPTION

Kamil Dędza
Agnieszka Kuleta

FK_NEWApplication	numeric	Unique application identifier **PK**
FK_additional_lesson	numeric	Lesson unique identifier **PK**

DIMENSION TABLE: DIM_additional_lesson		
NAME	TYPE	DESCRIPTION
lessonID	numeric	Unique lesson identifier **PK**
Name	Varchar	Name of the class
Start_year	date	Year of start

Dimensional model

Fact definitions

Fact 1 New application fact: New application allows enrolment to the kindergarten , issued on a

specified day, at a specified time. Enrolment is for specific facility and specific group, additional activities are also selected.

Fact table: FACT_NEWapplication

Granularity:

- A specified date of issue
- A specified date of confirmation
- A specified contract
- A specified facility
- A specified group

Measures and aggregation functions:

- Number of new application facts – COUNT (1)
- Number of additional_lessons selected – COUNT (1)

Fact 2 Choosing additional classes fact: Choosing additional classes during enrolment process.

Fact table: FACT_additional

Granularity:

- a specified lesson chosen during in one application,
- a specific lesson

Measures and aggregation functions:

- Number of choosing additional classes facts – COUNT(1)

Dimension definitions

Dimensions for Fact 1 Sale fact:

DIMENSION	TYPE	TABLE/Column
Confirmation date	**Hierarchical dimension**	**• Date.Year •• Date.Month ••• Date.Date**
Confirmation date	**Date**	**Dimension**
Confirmation YEAR	Date.Year	Dimension attribute
Confirmation MONTH	Date.Month	Dimension attribute
Confirmation DAY	Date.Day	Dimension attribute
Issue date	**Hierarchical dimension**	**• Date.Year**

Kamil Dędza
Agnieszka Kuleta

		•• Date.Month **••• Date.Date**
Issue date	**Date**	**Dimension**
Issue YEAR	Date.Year	Dimension attribute
Issue MONTH	Date.Month	Dimension attribute
Issue DAY	Date.Day	Dimension attribute
Meetings	**Dimension**	**DIM_meeting_type**
New application	Dimension attribute	DIM_meeting_type.FK_NEWApplication
Two Parents	Dimension attribute	DIM_meeting_type.Two_parents
Meeting cancelled	Dimension attribute	DIM_meeting_type.Meeting_cancelled
Contract signed during the first meeting	Dimension attribute	DIM_meeting_type.Contract_duringmeeting
Status_of_application	Dimension attribute	DIM_meeting_type.Status_of_application
Facility	**Dimension**	**DIM_Facility**
ID of facility	Dimension attribute	DIM_Facility.facilityID
City	Dimension attribute	DIM_Facility .City
Post-code	Dimension attribute	DIM_Facility .Post-code
Street	Dimension attribute	DIM_Facility .Street
Building	Dimension attribute	DIM_Facility .Building
Apartment	Dimension attribute	DIM_Facility .Apartment
District	Dimension attribute	DIM_Facility .district
Current	Dimension attribute(SCD2)	DIM_Facility.current
Date_of_end	Dimension attribute(SCD2)	DIM_Facility.date_of_end
Date _start	Dimension attribute(SCD2)	DIM_Facility.date_start
Contract	**Dimension**	**DIM_Contract**
ID of contract	Dimension attribute	DIM_Contract .IDContract
Signing_date	**Hierarchical dimension**	**• Date.Year** **•• Date.Month** **••• Date.Date**
Contract signed	Dimension attribute	DIM_Contract .Contracted
Active	Dimension attribute	DIM_Contract .active
Groups	**Hierarchial Dimension**	**• Groups.year_no** **•• groups.age** **••• Groups.size**
Groups	**Dimension**	**DIM_groups**
ID of groups	Dimension attribute	DIM_Groups.groupsID
Name of group	Dimension attribute	DIM_Groups.name
Age of children in group	Dimension attribute	DIM_Groups.age
Year	Dimension attribute	DIM_Groups.year
Size of group	Dimension attribute	DIM_Groups.size

Dimensions for Fact 2 Choosing additional classes fact:

DIMENSION	TYPE	TABLE/Column
Additional lesson	**Dimension**	**DIM_additional_lesson**
Start year	Dimension attribute	DIM_additional_lesson.start_year
Name	Dimension attribute	DIM_additional_lesson.name

Checking the feasibility of queries based on the multidimensional model

1. Compare the number of application to each age group in current month of the year to the same moth in the previous year.
    - Measure: Number of applications
    - Dimension: Groups (*dimension attributes groups age*)

- Dimensions: Date (*dimension attributes date year*)

2. Which additional lessons are mostly selected by the parents during submitting the application in the current year.
   - Measure Number of additional
   - Dimensions: Additional_lesson (*dimension attributes additional lesson lessonID*)
   - Dimensions: Date (*dimension attributes date year*)

3. In which month the number of applications was the highest and the lowest in the previous year.
   - Measure: Number of applications
   - Dimensions: Date (*dimension attributes date year*)

4. If the number of additional classes increased in comparison to previous year, how many percent more or less accepted application with signed contract were in comparison to previous year?
   - Measure: Number of applications
   - Dimension: Meeting (dimension attribute status_of_application)
   - Dimension: Contract (dimension attributes contracted)
   - Dimensions: Date (dimension attributes date year)

5. What is the difference in percent of confirmed applications taking into account previous and current year?
   - Measure: Number of applications
   - Dimension: Meeting (dimension attribute status_of_application)
   - Dimensions: Date (*dimension attributes date year*)

6. How many applications were send to the most frequent location in the current year?
   - Measure: Number of application
   - Dimension: facility (dimension attribute name)
   - Dimensions: Date (*dimension attributes date year*)

7. What is the average number of applications with active contract in each facility in the current year?
   - Measure: Number of applications
   - Dimension: contract(dimension attribute active)
   - Dimension: Date (*dimension attributes date year*)
   - Dimension: facility(dimension attribute name)

8. What is the percentage of accepted applications if the contract was signed in the same day as meeting?
   - Measure: Number of applications
   - Dimension: Meeting (dimension attribute contract_duringmeeting)
   - Dimensions: Date (*dimension attributes date year*)

9. If both parents attended the meeting, which additional lesson was the most frequently chosen in the current year?
   - Measure: Number of choosing additional classes
   - Dimension: Meeting (dimension attribute two_parents)
   - Dimension: additional_lesson (dimension attribute lessonID)
   - Dimensions: Date (*dimension attributes date year*)

10. What is the percentage accepted applications if the meeting was cancelled?
    - Measure: Number of applications
    - Dimension: Meeting (dimension attribute meeting_cancelled)
    - Dimensions: Date (*dimension attributes date moth*)

Kamil Dędza
Agnieszka Kuleta

11. What is average number of additional classes selected in the most selected facility in the application?
    - Measure: Number of selected additional classes
    - Dimension: Facility (dimension attribute name)
    - Dimensions: Date (dimension attributes date moth)

Checking if there are Date in the Date sources needed to fill the Date warehouse

FACT_NEWapplication – one tuple describes one application	
Application ID	Based on Application ID stored in Application table in GiraffeEnroll source.
FK_facility	Facility where the kid is wanted to be enrolled. Foreign key from dimension table. Based on Facility ID stored in Facility table in GiraffeEnroll source.
FK_contract	Contract ID. Foreign key from dimension table. Based on Contract ID stored in Contract table in GiraffeEnroll source.
FK_confirmationdate	Date, when the application is confirmed. Foreign key from dimension table. Based on Confirmation_date stored in Application table in GiraffeEnroll source.
FK_issuedate	Date, when the application isissued. Foreign key from dimension table. Based on Facility ID stored in Application table in GiraffeEnroll source.
FK_groups	Group ID. Foreign key from dimension table. Based on Group ID stored in Group table in GiraffeEnroll source.
Additionals_selected	Based on counting all records in Chooses table in GiraffeEnroll source, which has the application ID the same as our application.

DIM_JUNK– one tuple describes one meeting	
Two_parents	Marked present, when none of the parent is NULL in Meeting register COLUM Identification number of mother or of father. Otherwise absent.
Meeting_canceled	Mark cancelled, if in the meeting register date of the meeting connected with given application is null.
Contract_duringmeeting	Contract issue during first meeting, selects all meetings connected with application
Status_of_application	Application status. Based on status stored in Application table in GiraffeEnroll source

DIM_facility – one tuple describes one facility	
facilityID	Facility ID. Based on Idfacility stored in Facility table in GiraffeEnroll source.
Name	Facility Name. Based on Facility Name stored in Facility table in GiraffeEnroll source.
City	Facility City localisation. Based on Facility FK_address stored in Facility table in GiraffeEnroll source.FK_addres connects facility with Address table and by joining them we can find city in city columnn

Kamil Dędza
Agnieszka Kuleta

Post-code	Facility localisation. Based on Facility FK_address stored in Facility table in GiraffeEnroll source.FK_addres connects facility with Address table and by joining them we can find post- code in post-code column
Street	Facility localisation. Based on Facility FK_address stored in Facility table in GiraffeEnroll source.FK_addres connects facility with Address table and by joining them we can find street in street column
Building	Facility localisation. Based on Facility FK_address stored in Facility table in GiraffeEnroll source.FK_addres connects facility with Address table and by joining them we can find building in building column
Apartment	Facility localisation. Based on Facility FK_address stored in Facility table in GiraffeEnroll source.FK_addres connects facility with Address table and by joining them we can find appartment in appartment column
District	Facility localisation. Based on Facility FK_address stored in Facility table in GiraffeEnroll source.FK_addres connects facility with Address table and by joining them we can find district in district column
Date_start	Facility's day of opening. Based on Facility opening_date stored in Facility table in GiraffeEnroll source
Date_end	Facility's day of closing . Based on Facility closing_date stored in Facility table in GiraffeEnroll source
Current	Facility's current address. Based on Facility closing_date stored in Facility table in GiraffeEnroll source, if it is null the current is turn to yes.

DIM_contarct one tuple describes one contract	
IDContract	Contract ID. Based on Contarct ID stored in Contarct table in GiraffeEnroll source
Active	Contract Activeness. Based on Contract issue_date and ending_date stored in Contract table in GiraffeEnroll source.
Contracted	Contract ID. Based on Contract signing date stored in Contract table in GiraffeEnroll source. If not null then true.

DIM_groups-one tuple describes one group	
groupsID	GROUP ID. Based on GROUP ID stored in Groups table in GiraffeEnroll source
Name	GROUP name. Based on column name stored in Groups table in GiraffeEnroll source
Age	GROUP age. Based on column age stored in Groups table in GiraffeEnroll source. The three, four and five years-old
Year	GROUP year. Based on column year stored in Groups table in GiraffeEnroll source.
Size	GROUP ID Based on column FK_groups stored in Child table in GiraffeEnroll source. Counts all children whose FK_groups is the same. If the count is smaller than 10, then the group is small, if is between 10 and 15 hen is medium, bigger than 15 is big

DIM_Date-one tuple describes one day. All the data in this table are generated tuple by tuple based on any calendar, before ETL process

FACT TABLE : FACT_additional -one tuple describes one additional class chosen	
FK_NEWApplication	Application ID. Based on Application ID stored in Application table in GiraffeEnroll source

Kamil Dędza
Agnieszka Kuleta

FK_additional_lesson	Lesson ID. Based on Lesson ID stored in Additional_lesson table in GiraffeEnroll source

DIM_additional_lesson- one tuple one class offered by school	
lessonID	Lesson ID. Based on Lesson ID stored in Additional_lesson table in GiraffeEnroll source
Name	Lesson name. Based on Lesson.name stored in Additional_lesson table in GiraffeEnroll source
Start_year	Lesson.Start_year. Based on Lesson Start_year stored in Additional_lesson table in GiraffeEnroll source