1. General description for Enrolment process

a. general description of the business process and a description of the performance metrics generated by this process, possible current analytical problems.

The process is as follows: parent enters kindergarten's website and moves to enrolment section. Then he or she fulfils the whole survey which contains required data about the child and parent itself. Optionally additional activities can be chosen from the list. Then the parent receives e-mail whether the application was confirmed or not. Next parent needs to visit the facility and sign a contract. Then when everything is settled the identification number is assigned to the child.

The main goal would be fulfilled if the annual increase of application in the kindergarten is at least 5% and also the annual income growth by 10%.

b. Typical questions

1. Compare the number of application to each age group.
2. Which additional lessons are mostly chosen by parents during fulfilling the application.
3. In which month the number of applications is the highest and the lowest.
4. What is the percentage of parents who had the child in "Giraffe" before?
5. What is the average time between sending an application and signing the contract?
6. What is the percent of confirmed applications?(Confirmed application receives the email informing about appointment)
7. What is the percent of parents that their application was confirmed but they did not sign a contract.
8. Which additional activity was chosen by the smallest number of parents?
9. What is the number of children from each districts?
10. What is the percent of children attending the kindergarten from different district than the facility itself?

c. Data

All application data are extracted from the system – "GiraffeEnroll". Application system stores information about the parent filing the application, the application number, the kid sensitive data and the location where he wants to send the child. In addition it is known that the kindergarten signs the contract with the parent and all data about it is stored in the database as well.

2. Data sources structure

a.Giraffe Enrol

Table 1: Child				
Attributes				
Name	**Primary key**	**Type**	**Domain**	**description**
IDChild	yes	integer	12-digit number from interval (1 000 000 00, 9999 999 999 99)	Unique child identifier
name	no	text	Up to 20 chars. Roman alphabet with polish signs, any special characters not allowed(ex. @,!,_)	Name of child
surname	no	text	Up to 30 chars. Roman alphabet with polish signs, special sign as "-" is allowed the other no.	Surname of child
Birthday	No	Date	Allows to introduce date from the interval (1.01.2018 -today) The form: ('YYYYMMDD'	Date of birth of the child
PESEL	no	integer	A Polish number, consisting of 11 digits, out of which first 6 represent birth date.	Number contains 11 digits. Only polish citizens have it.
Medicines	No	Text	Up to 100 chars. Roman alphabet with polish signs.	Medicines taken by a child
FK_mother	no	Integer	13-digit number from interval (1 000 000 000, 9999 999 999 999)	Unique patient identifier. Contains 13 digits
FK_father	no	Integer	13-digit number from interval (1 000 000 000, 9999 999 999 999)	Unique patient identifier. Contains 13 digits
FK_address	No	Integer	12-digit number from interval (1 000 000 00, 9999 999 999 99)	Address
FK_group	no	integer	5-digit number	Unique group identifier

Table 2: CONTRACT				
Attributes				
Name	**Primary key**	**Type**	**Domain**	**description**
IDContract	yes	Integer	10-digit number	Unique contract identifier. Contains 12 digits
Issue_date	no	date	Allows to introduce date The form: ('YYYYMMDD')	Date of issue the contract.
Ending_date	no	date	Allows to introduce date from the interval (today-01.01.2050) The form: ('YYYYMMDD')	Date of ending the contract.
Signing_date	no	date	Allows to introduce date from the interval (1.01.2000-today) The form: ('YYYYMMDD')	Date of the signing the contract./ NULL meaning that contract was not signed.
FK_parent	no	Integer	13-digit number from interval (1 000 000 000, 9999 999 999 999)	Unique patient identifier. Contains 13 digits
FK_application	no	Integer	10-digit number	Unique application identifier

Table 3: PARENT				
Attributes				
Name	**Primary key**	**Type**	**Domain**	**description**
IDparent	yes	Integer	13-digit number from interval (1 000 000 000, 9999 999 999 999)	Unique parent identifier. Contains 13 digits
name	no	text	Up to 20 chars. Roman alphabet with polish signs, any special characters not allowed(ex. @,!,_)	Name of parent
surname	no	text	Up to 30 chars. Roman alphabet with polish signs, any special characters not allowed(ex. @,!,_)	Surname of parent

PESEL	no	integer	A Polish number, consisting of 11 digits, out of which first 6 represent birth date.	Number contains 11 digits. Only polish citizens have it.
phone	No	integer	Contains 9 digits. Any special signs are not allowed.	Phone number contains 9 digits.
Email	No	Text	Up to 30 chars. Roman alphabet with polish signs	Email of a parent

Table 4: APPLICATION				
Attributes				
Name	**Primary key**	**Type**	**Domain**	**description**
IDapplication	Yes	Integer	10-digit number	Unique application identifier
Issue_date	No	Date	Allows to introduce date from the interval (1.01.2000-today) The form: ('YYYYMMDD')	Date of the issue of application.
status	No	Text	confirmed/rejected	The status of application
Confirmation_date	No	Data	Allows to introduce date from the interval (1.01.2000-today) The form: ('YYYYMMDD')	Date of confirming the application.
FK_facility	No	Integer	3-digit number	Unique facility identifier.
FK_child	No	integer	12-digit number from interval (1 000 000 00, 9999 999 999 99)	Unique child identifier

Table 5 : Group				
Attributes				
Name	**Primary key**	**Type**	**Domain**	**description**
IDgoup	yes	integer	5-digit number	Unique group identifier
name	no	text	Up to 20 chars. Roman alphabet with polish signs, any special	Name of group

			characters not allowed(ex. @,!,_)	
age	no	integer	Number from interval (3,5)	Age of children
year	no	Integer	Number from interval (2002-today)	Year of existance

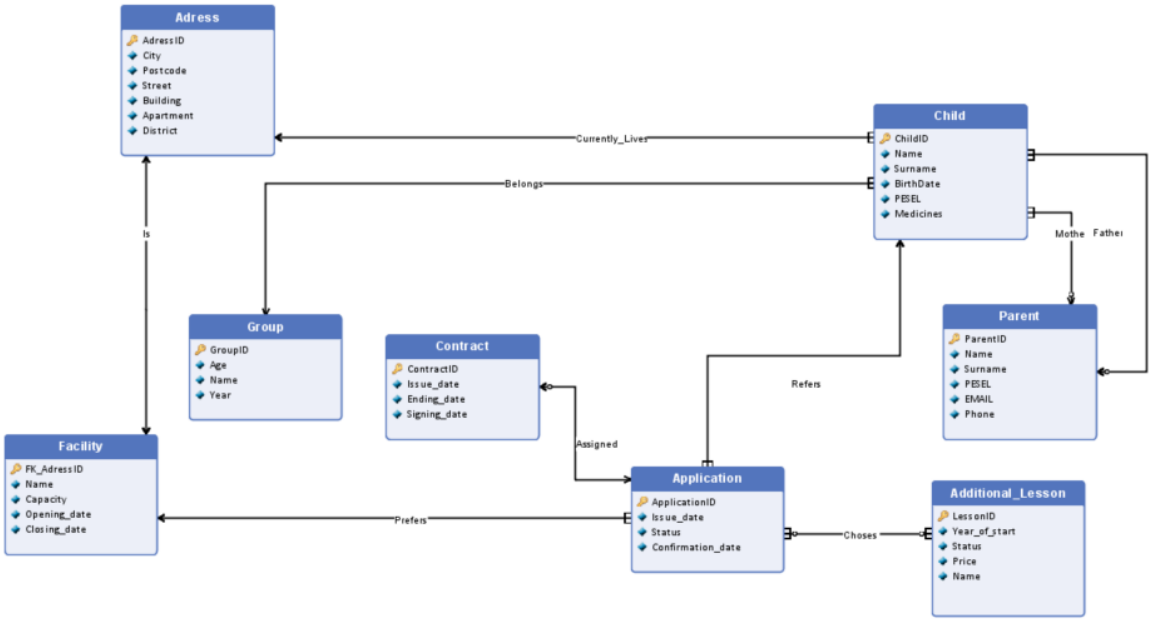
Table 6 : Facility				
Attributes				
Name	**Primary key**	**Type**	**Domain**	**description**
IDfacility	yes	Integer	3-digit number	Unique facility identifier.
name	no	text	Up to 20 chars. Roman alphabet with polish signs, any special characters not allowed(ex. @,!,_)	Name of facility
capacity	no	integer	Number from interval (100,300)	Number of children
Opening_date	no	Integer	Number from interval (2002-today)	Date of opening
Closing_date	no	Integer	Number from interval (2002-today)	Date of closing
FK_address	No	Integer	12-digit number from interval (1 000 000 00, 9999 999 999 99)	Address of facility

Table 7: Address				
Attributes				
Name	**Primary key**	**Type**	**Domain**	**description**
IDaddress	Yes	Integer	12-digit number from interval (1 000 000 00, 9999 999 999 99)	Address
Zip-code	No	integer	Contains 5 digits. Any special signs are not allowed.	Number which contains 5 digits
city	No	text	Up to 20 chars. Roman alphabet with polish signs, any special characters not allowed(ex. @,!,_)	Name of the city
street	No	text	Up to 30 chars. Roman alphabet with polish signs, any special	Street name

			characters not allowed(ex. @,!,_)	
Building	No	text	Number from interval (0,1000)	Number of patient's home
Apartment	No	integer	Number from interval (0,1000)	Number of patient's flat
District	No	Text	Up to 20 chars. Roman alphabet with polish signs, any special characters not allowed(ex. @,!,_)	District in the city

Table 8 : Additional_Lesson				
Attributes				
Name	**Primary key**	**Type**	**Domain**	**description**
IDLesson	Yes	Integer	5-digit number	Lesson unique identifier
name	no	text	Up to 20 chars. Roman alphabet with polish signs, any special characters not allowed(ex. @,!,_)	Name of subject
Year_of_start	no	Integer	Number from interval (2002-today)	Year of start
status				
Price	No	Integer	Integers in range (200-500)	Price for a class per month

Table 9:Chooses				
Attributes				
Name	**Primary key**	**Type**	**Domain**	**description**
IDLesson	Yes	Integer	5-digit number	Lesson unique identifier
IDapplication	Yes	Integer	10-digit number	Unique application identifier



b.MeetingRegister

Sheet 1 (Information about meeting):

- Column A– Identification number of application
- Column B – Identification number of application
- Column C - Identification number of a child (numeric, 12 digits)
- Column D - Identification number of mother
- Column E - Identification number of mother
- Column F – date of meeting (in format year - month – day)
- Column G – hour of meeting(in format hh:mm)
- Column H – facility (text)

3. Scenarios of analytical problems

Why was there increase/decrease in the number of applications in comparison to previous year?

1. Compare the number of application to each age group in current month of the year to the same moth in the previous year.
2. Which additional lessons are mostly selected by the parents during submitting the application in the current year.
3. In which month the number of applications was the highest and the lowest in the previous year.
4. If the number of additional classes increased in comparison to previous year, how many percent more or less of contracts were signed also in comparison to previous year?
5. What is the difference in percent of confirmed applications taking into account previous and current year?

Which facility and why was the most popular?

1. Which location is mostly selected by the parents during submitting the application in the current year?
2. What is the average number of children in each facility in the current year?
3. What percentage of contract were signed in the same day as meeting in the current year?(both)
4. If both parents attended the meeting, which additional lesson was the most frequently chosen in the current year?(both)
5. What is the percent of unrealised appointments in each facility in the last three months?(excel)

4. Data needed for analytical problems

1.Compare the number of application to each age group in current month of the year to the same moth in the previous year.

- number of groups - “GiraffeEnroll”, table Group, colum GroupID
- age of group - “GiraffeEnroll”, table Group, column Age
- number of applications - “GiraffeEnroll”, table Application, column ApplicationID
- Date of application - “GiraffeEnroll”, table Application, column Issue_date

2. Which additional lessons are mostly selected by the parents during submitting the application in the current year.

- name of course - “GiraffeEnroll”, table Additional_Lesson, column Name
- number of occurences - “GiraffeEnroll”, table Chooses, colums FK_IDLesson
- Date of application - “GiraffeEnroll”, table Application, column Issue_date

3. In which month the number of applications was the highest and the lowest in the previous year.

- month of application - “GiraffeEnroll”, table Application, column Issue_date
- number of applications - “GiraffeEnroll”, table Application, column ApplicationID
- Date of application - “GiraffeEnroll”, table Application, column Issue_date

4. If the number of additional classes increased in comparison to previous year, how many percent more or less of contracts were signed also in comparison to previous year?

- Number of additional classes - “GiraffeEnroll”, table Additional_lesson, column LessonID
- Date of the class - “GiraffeEnroll”, table Additional_lesson, column Issue_date
- Number of all contracts - “GiraffeEnroll”, table Contract, column ContractID
- Date of contract - “GiraffeEnroll”, table Contract, column Issue_date
- Whether was signed - “GiraffeEnroll”, table Contract, column Signing_date

5.What is the difference in percent of confirmed applications taking into account previous and current year?

- total number of application - “GiraffeEnroll”, table Application, column ApplicationID
- number of confirmed applications - “GiraffeEnroll”, table Application, column Status
- Date of application - “GiraffeEnroll”, table Application, column Issue_date

1. Which location is mostly selected by the parents during submitting the application in the current year?

- Date of application - “GiraffeEnroll”, table Application, column Issue_date
- address of facility- “GiraffeEnroll”, table Adress, all columns apart from AdressID
- number of applications to each facility - “GiraffeEnroll”, table Application, column FK_Facility

2. What is the average number of children in each facility in the current year?

- facility - “GiraffeEnroll”, table Facility, column FacilityID
- Number of children - “GiraffeEnroll”, table Application, column FK_FacilityID
- Active contract - “GiraffeEnroll”, table Contract, column Issue_date and Ending_date

3. What percentage of contract were signed in the same day as meeting in the current year?(both)

- day of signing the contract - “GiraffeEnroll”, table Contract, column Signing_date
- date of meeting – “MeetingRegister”, column date of meeting

4. If both parents attended the meeting, which additional lesson was the most frequently chosen in the current year?(both)

- Mother– “MeetingRegister”, column motherID
- Father - “MeetingRegister”, column fatherID
- Additional lesson - “GiraffeEnroll”, table Application, column FK_LessonID

5. What is the percent of unrealised appointments in each facility in the last three months?(excel) (unrealised mother and father are null)

- Mother– “MeetingRegister”, column motherID
- Father - “MeetingRegister”, column fatherID
- Number of appointments- “MeetingRegister”, column AppointmentID
- Date of appointment - “MeetingRegister”, Dateofmeeting

5.Additional data query(don’t demand changing the process)

- Does the number of friends liking our facebook profile increases person's willingsness to enroll its child to kindergarten?

6.Additional data query(demand changing the process)

- Does the income of parents affect the chance of the child to being accepted to kindergarten?