Specification of requirements for the process of Choosing the new destinations of a trip

1. Business process description

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

The process of choosing the new destination is as follows: a company worker that is responsible for this process needs to log in the system and open the Excel file which contains all of the information about the sold trips. He or she needs to use the program formulas to find the most frequently chosen places. They also analyze the offers of other companies using the Trivago or other services. After this analysis they come up with 4 potential travel destinations report. On its basis CEO of company makes the final decision.

b. *Typical questions*

- Which of the trips are the most popular one?
- How the willingness to choose the trip depend on the transport form?
- Compare the amount of sold trips per country.
- How the popularity of places depends on a time of a year.
- Compare prices of trips with our main competitors.
- What are the destinations recommended by our previous customers

c. *Data*

Data from the Excel file provided by the company, which contains the data from each of the trips (cost & number of sold one) and database where we can find the data about the travel destinations and all of our workers. The data can be also collected from the internet.

2. Data sources structure

Company database:

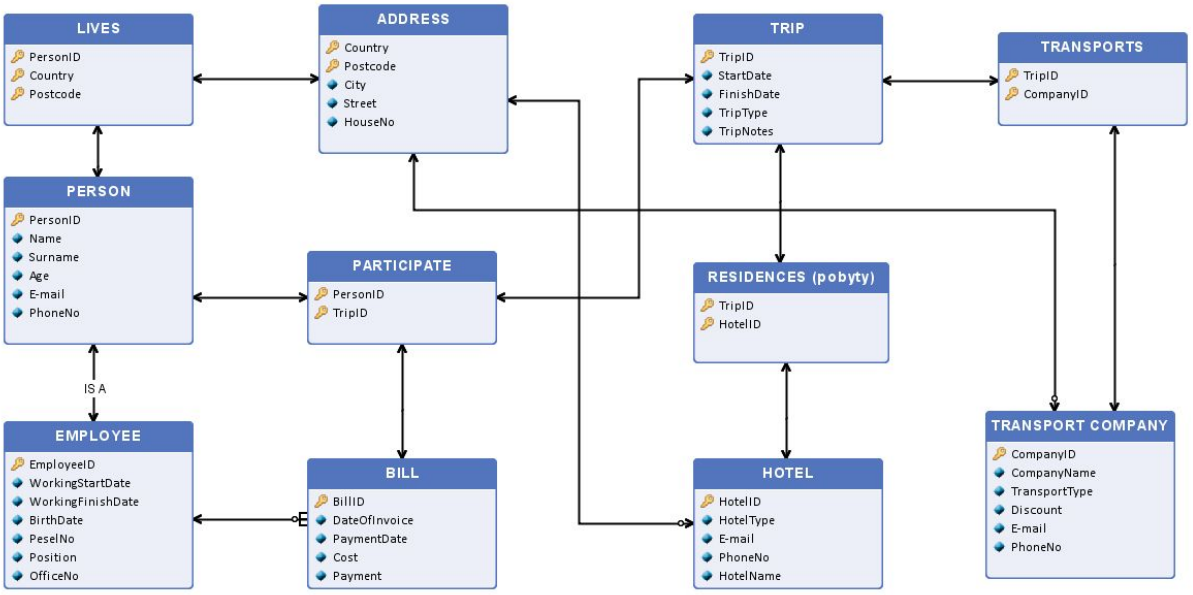

TABLE NAME	*ATTRIBUTE*	*ATTRIBUTE TYPE*	*DESCRIPTION*
PERSON	*The table contains information about the people we need to store information about (customers, employees)*		
	PersonID	*int*	*PK*

	Name	*varchar*	*The name of the person*
	Surname	*varchar*	*Surname of the person*
	Age	*int*	*The age of the person*
	E-mail	*varchar*	*Person's e-mail address*
	PhoneNo	*varchar*	*Person's phone number*
TRIP	*The table contains informations about trips we offered to Our clients*		
	TripID	*int*	*PK*
	StartDate	*date*	*The start date of the trip*
	FinishDate	*date*	*The end date of the trip*
	TripType	*varchar*	*Type of the trip*
	TripNotes	*varchar*	*Our comments about the trip*
ADDRESS	*The table contains informations about addresses that We have to store (customer's address, employee addresses, addresses of companies we work with)*		
	Country	*varchar*	*Part of the PK*
	Postcode	*varchar*	*Part of the PK*

	City	*varchar*	*City*
	Street	*varchar*	*Street*
	HouseNo	*varchar*	*House number*
TRANSPORT COMPANY	*The table contains informations about all transport companies that We cooperate with*		
	CompanyID	*int*	*PK*
	CompanyName	*varchar*	*Name of the company*
	TransportType	*varchar*	*Transport type that this company offer (allowed values: ship, bus, plane, train)*
	Discount	*int*	*Number of people from which we get discount*
	E-mail	*varchar*	*E-mail address*
	PhoneNo	*varchar*	*Phone number*
HOTEL	*The table contains informations about hotels all over the world that We cooperate with*		
	HotelID	*int*	*PK*
	HotelType	*varchar*	*Type of the hotel (allowed values: 1/2/3/4/5 stars, motel, hostel)*

	E-mail	*varchar*	*E-mail address*
	PhoneNo	*varchar*	*Phone number*
	HotelName	*varchar*	*Name of the hotel*
BILL	*The table contains informations about bills that have been issued in Our company.*		
	BillID	*int*	*PK*
	DateOfInvoice	*date*	*Date of issuing the bill*
	PaymentDate	*date*	*Date of bill payment*
	Cost	*int*	*The cost of the bill (in PLN)*
	Payment	*varchar*	*Payment type (allowed values: cash, card, bank transfer)*
EMPLOYEE	*The table contains information about all Our employees*		
	EmployeeID	*int*	*PK*
	WorkingStartDate	*date*	*Date of commencement of work*
	WorkingFinishDate	*date*	*Date of finishing work*
	BirthDate	*date*	*Date of birth*

	PeselNo	*int*	*Employee's PESEL number*
	Position	*varchar*	*Employee's position in the company*
	OfficeNo	*int*	*Starts with the Voivodeship TERYT number and 3 numbers. Example : 22001 - a person no. 1 from pomeranian Voivodeship)*
TRANSPORTS	*Implementation of relationship n to n between TRIP and TRANSPORT COMPANY. It is identified by two foreign keys of TRIP and TRANSPORT COMPANY tables. The table contains informations about all transports that had occurred during specific trip*		
	TripID	*int*	*FK TRIP, part of the PK*
	CompanyID	*int*	*FK TRANSPORT COMPANY, part of the PK*
RESIDENCES (pobyty)	*Implementation of relationship n to n between TRIP and HOTEL. It is identified by two foreign keys of TRIP and HOTEL tables. The table contains informations about accommodations during a specific trip*		
	TripID	*int*	*FK TRIP, part of the PK*
	HotelID	*int*	*FK HOTEL, part of the PK*

LIVES	*Implementation of relationship n to n between ADDRESS and PERSON. It is identified by three foreign keys of ADDRESS and PERSON tables. The table contains informations of people's addresses (addresses of our clients'/employees' residence)*		
	PersonID	*int*	*FK PERSON, part of the PK*
	Country	*int*	*FK ADDRESS, part of the PK*
	Postcode	*int*	*FK ADDRESS, part of the PK*
PARTICIPATE	*Implementation of relationship n to n between TRIP and PERSON. It is identified by two foreign keys of TRIP and PERSON tables. The table contains informations about people who took part in a specific trip*		
	PersonID	*int*	*FK PERSON, part of the PK*
	TripID	*int*	*FK TRIP, part of the PK*

Company EXCEL file:

Sheet 1 (All of the costs of the trips)

- Column A - Trip ID (number)
- Column B - Profit from the trip in PLN (currency)
- Column C - Total travel revenue in PLN (currency)
- Column D - Total cost of a trip in PLN- the total amount We paid for the travel (currency)
- Column E - Total cost of transport in PLN (currency)
- Column F - Total cost of hotel in PLN (currency)
- Column G - Total food cost in PLN (currency)
- Column H - Total additional costs in PLN (currency)
- Column I - Price per person (currency)
- Column J - Total number of the participants (number)
- Column K - The number of remaining places (number)
- Column L - What kind of additional costs were there (text)
- Column M - Comments about the costs (text)
- Column N - City name (text)
- Column O - Date of the beginning (date)
- Column P - Date of the end (date)

3. Scenarios of analytical problems

What influences the willingness to buy a trip?

1. Compare destinations chosen by our clients month by month.
2. Compare the number of people per trip grouped by the destination.
3. Which destination is recommended by our clients?
4. Compare no. of bought trips according to the type of transport
5. How the number of sold trip is influenced by time of the year?
6. How the age of clients influence willingness to buy a trip?
7. How does the travel time influence appreciation the of the trip?

Which of the trip is the most profitable?

1. What is a profit from each trip?
2. What kind of transport gives the biggest discount for a trip
3. What are the groups of age of the customers and what is the number of people traveling in each group?
4. Which destination produces the smallest amount of dissatisfied clients?
5. Which of the tips destinations created the biggest percent of unsold trips?

4. Data needed for analytical problems

Analytical problem: What influences the willingness to buy a trip?

1. Compare destinations countries chosen by our clients month by month.
    - the total number of people travelling to every destination - data from Excel file, sheet 1, Column J (Total number of the participants)
    - city name - data from Excel file, sheet 1, Column N (City name)
    - date of the beginning - data from Excel file, sheet 1, Column O (Date of the beginning)
    - date of the end - data from Excel file, sheet 1, Column P (Date of the end)
2. Compare the number of people per trip grouped by the destination.
    - the total number of people travelling to every destination - data from Excel file, sheet 1, Column J (Total number of the participants)
    - city name - data from Excel file, sheet 1, Column N (City name)
3. Which destination is recommended by our clients?
    - ◆ there's no such information available in both data sources. The proposals for acquiring such information:
    - provide the additional survey to collect necessary data
4. Compare no. of bought trips according to the type of transport
    - trip ID - data from the database, table TRIP, column TripID

- company ID - data from the database, table TRANSPORTS, column CompanyID
- type of transport - data from the database, table TRANSPORT COMPANY, column TransportType

5. How the number of sold trip is influenced by time of the year?
    - date of the beginning - data from Excel file, sheet 1, Column O (Date of the beginning)
    - date of the end - data from Excel file, sheet 1, Column P (Date of the end)
6. How the age of clients influence willingness to buy a trip?
    - person ID - data from the database, table PARTICIPATE, column PersonID
    - age of a customer - data from the database, table PEOPLE, column Age
    - customer trips - data from the database, table PARTICIPATE, column TripID
7. How does the travel time influence appreciation of the trip?
    - *Travel time* - no such an information
        ➔ proposition to solve this problem:
        - use the average time from the google maps
        - Add a column in Excel
        - Add an attribute in database
    - *Trip evaluation* - no such an information
        ➔ the proposals for acquiring such information:
        - provide the additional survey to collect necessary data

Analytical problem: Which of the trip is the most profitable?

1. What is a profit from each trip?
    - profit from each trip - data from Excel file, Column B (Profit from the trip in PLN)
2. What kind of transport gives the biggest discount for a trip?
    - transport type - data from the database, table TRANSPORT COMPANY, column TransportType
    - number of people from which we get discount - data from the database, table TRANSPORT COMPANY, column Discount
    - a trip for which we need to find a transport - tripID from Trip
3. What are the groups of age of the customers and what is the number of people traveling in each group?
    - person ID - data from the database, table PARTICIPATE, column PersonID
    - age of a customer - data from the database, table PEOPLE, column Age

4. Which destination produces the smallest amount of dissatisfied clients?
    - *Clients opinion* - there's no such information available in both data sources.
        - ➔ the proposals for acquiring such information:
            - provide the additional survey to collect necessary data
    - City name - data from Excel file, sheet 1, Column N (City name)

5. Which of the tips destinations created the biggest percent of unsold trips?
    - city name - data from Excel file, sheet 1, Column N (City name)
    - number of remaining places - data from Excel file, sheet 1, Column K (The number of remaining places)

It is not possible to build a BI system to support “Martola” in solving some analytical problems without introducing additional activities. We suggest to introduce a survey that will be send to every client after the trip.
Such questionnaire should contain as minimum the following set of questions:

- How would you evaluate your trip? (from 1(horrible) to 5 (amazing))
- What are Your recommendations for the next trip destination?