

Subject of database: airline services (flights, passengers, personnel)

1. Task analysis

1.1 Customers: the airlines.

1.2 Users: airlines employees with permissions, directors and owners; some part of database will be available for everyone through the company's website (information about scheduled flights).

1.3 Purpose of the database: storing and searching for information about flights and passengers, collecting data on aircrafts owned by the airlines, do analysis based on stored data and make improvements for the future, present available flights for next two years for customers of the airlines, controlling monthly profits from tickets.

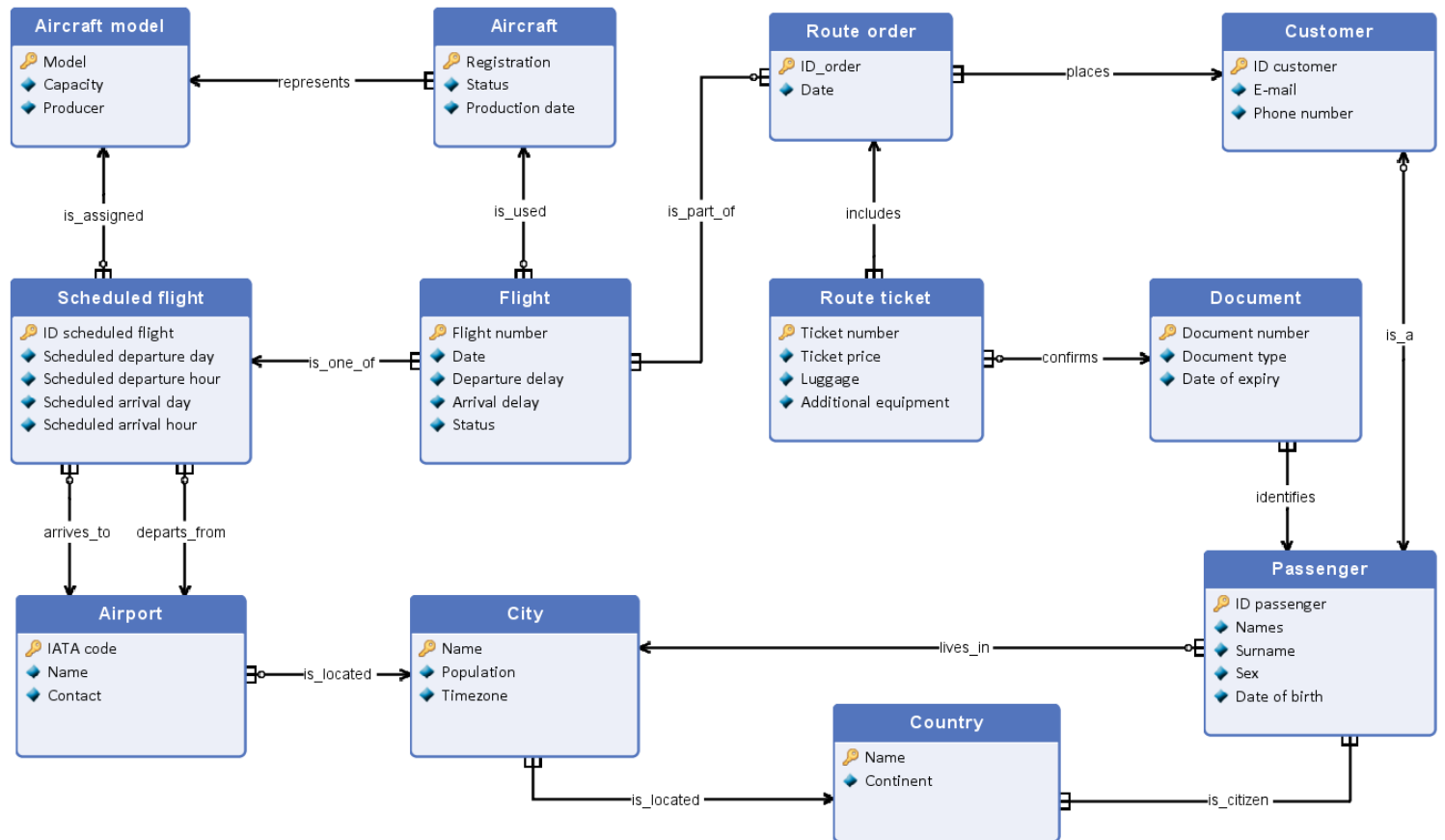
1.4 Possible scenarios of database use:

- Someone wants to find flight from Warsaw to Gdańsk between 17th and 20th September.
- Someone who wants to choose a vacation spot, checks the directions of flights offered by airlines from his country.
- Someone wants to check if he will be able to transport from Gdańsk to Los Angeles without stopovers. If not, how many changes and where he will have to do to make the whole journey as short and as cheap as possible.
- Airlines manager wants to know how many aircrafts from each model are available at the moment and how many are under repair.
- Airlines data analyst wants to compare previous year's monthly profits and present results for supervisor.
- The company's president wants to find out which routes are the most popular among customers during winter and summer seasons.
- The president of the airlines wants to know which scheduled flights have the most interest and consider introducing additional flights on a given route.
- The director of the company wants to know from which countries they have the most customers.
- Airlines data analyst examines profiles of passengers and tries to find dependencies between his characteristic and chosen flight. For example, he checks where students from Poland fly the most often.
- The company's president wants to know how many passengers were serviced by airlines last year.

1.5 Assumptions and limitations of the designed database

- system does not include booking cars, hotels or transport to/from the airport
- database does not contain ticket prices for individual flights and promotions (only the price paid by the passenger for the entire route on the day of purchase)
- system does not include complaints about lost or damaged luggage
- system does not include additional services (travel insurance, choosing class and seat in the plane, meals and drinks on board etc.)
- system does not include returning or changing tickets

2. ERD diagram



3. Description of ERD diagram

3.1 Definition of entities and explanation of attributes

Passenger (~millions entries) – set of current and past passenger of the airlines (current passenger is defined as a person who has ticket (as a part of order) for the route offered by the airlines (information about past passenger from last year are kept in case of any complaints). An entry (or several entries) is added when a new customer places an order and buys tickets for a group of people or only for one person. Entity might be deleted if someone has not had an active ticket for a year. Passenger attributes:

- ID passenger (primary key) – index of the passenger in the database, natural number from 1 to number of passengers
- Names and Surname – basic personal data, may be repeated between entities, consist of usually from 3 to 15 characters
- Sex – define gender of the person, possible value: male or female
- Date of birth – the day the person was born, in format of 8 digits (2 for day, 2 for month and 4 for year of birth) separated by dashes

Customer (~millions entries but less than number of passengers) – set of all passengers that place an order for a group of people or only for himself. Customer entity set extends passenger entity set with contact details because the customer is responsible for the order and receives the necessary information and ticket(s) for route. New entry is added when new customer places an order and is deleted when customer has not had an active order for a year. Customer additional attributes:

- ID customer (primary key) – index of the customer in the database, natural number from 1 to number of customers
- E-mail – e-mail address of the customer to communicate (information about flights, tickets etc.), is made up from a local-part, the symbol @ and a domain name
- Phone number – phone number of the customer to communicate (emergency situations), consist of call prefix (sign „+“ and maximum 4 digits) and phone number

Document (~millions entries – more than number of passengers) – set of all valid documents of passengers from database on the basis which passenger can buy his ticket and confirm his personal data. An entry is added when new customer appears in database or when current customer buys ticket using new document. Document is deleted from database when it becomes out-of-date or passenger that possesses particular document is deleted from the database. Document attributes:

- Document type – identify type of the document, possible options: identity card or passport
- Document number (primary key) – uniquely identifies the document, value according to the rules of country (for example in Poland: passport number consist of 2 letters and 7 digits, identity card consist of 3 letters and 6 digits)
- Date of expiry – day when document becomes invalid, in format of 8 digits separated by dashes (2 for day, 2 for month and 4 for year of birth)

Country (~less than 194 entries) – set of all countries where airports are located or passengers reside. New entity is added when new customer from new country appear in the database or when airlines introduces flight between airports in countries which are not in the database. Deleted when no more passengers come from and no more airports are located in the particular country. Attributes:

- Name (primary key) – name of the country, possible are only values from the official list of all countries of the World

- **Continent** – name of the continent in which country is located.

City (~tens of thousands entries) – set of all cities where passengers live and where airports are located. Entries are added and deleted similar to the Country entity set. Attributes:

- Name (primary key) – name of the city from a particular country
- Population - the number of residents living in the city (to determine the size of the city)
- Timezone – timezone that is valid in a given city, possible values: from UTC-12:00 to UTC+14:00

Aircraft model (~several dozen entries) – set of all types of aircraft owned by the airlines and currently used to transport (for example Boeing 787-9 Dreamliner, Embraer 195 etc.). New model is added when company buys newer model or model from other producer. An entry is deleted when there are no more aircrafts of this type in use. Attributes:

- Model (primary key) – the exact model of the plane, usually consist of name of the manufacturer and several digits that indentify the model among others aircraft of the producer
- Capacity – number of passengers that aircraft is able to carry on during one flight, for few to few hundreds of people
- Producer – the company which produced aircraft, mostly one of these: Boeing, Airbus, Embraer, Bombardier Aerospace, Cessna, Learjet, ATR.

Aircraft (~several hundred entries) – set of all aircrafts owned by the airlines, each aircraft is representative of one models from the Aircraft model entity set. In these entities there are information about particular aircraft, not about the model, because the company can own several aircrafts of the same model. New entry is added when company buy new aircraft and is deleted when particular aircraft if no longer used or sold. Attributes:

- Registration (primary key) – code unique to a single aircraft, it consist of prefix (fixed for each country) and usually 3 letters
- Status – information whether aircraft is at the airport or during the flight or under repair
- Production date – the date on which the aircraft was produced, in format of 8 digits seperated by dashes (2 for day, 2 for month and 4 for year of birth)

Airport (~hundreds entries) – set of all airports from which airplanes of the airlines depart and arrive to. A new entity is added when airlines launch new routes and is removed when airlines give up routes through the airport. Attributes:

- IATA code (primary key) – three-letter geocode designating airports defined by the International Air Transport Association.
- Name – name of the airport
- Contact – phone number to airport information, consist of call prefix (sign „+“ and maximum 4 digits) and phone number

Scheduled flight (~few thousands entries) – set of all flights offered by the airlines according to the weekly schedule. An entry is added when airlines introduce another flight to the schedule and removed when they resign from scheduled flight. Attributes:

- ID scheduled flight (primary key) – index of the flight in the database, natural number from 1 to number of scheduled flights.
- Scheduled departure day – day of the week of scheduled flight departure
- Scheduled departure time – time of scheduled flight departure (hour and minutes)

- Scheduled arrival day - day of the week of scheduled flight arrival
- Scheduled arrival time - time of scheduled flight arrival (hour and minutes)

Flight (~hundred of thousands entries) – set of all past (from last year) and future (for the next two years) airline flights. Information about past flights is kept in case of any complaints (for example about too long delay time). Flights from the next two years are in database because customer can place an order for a flight during this period. New entries are added (flights in two years) and deleted (flights from one year ago) every day. Attributes:

- Flight number (primary key) – number of the flight identifying it among other scheduled flights (several letters combined with digits)
- Date - the day the aircraft took off or is scheduled to take off, in format of 8 digits (2 for day, 2 for month and 4 for year of birth) separated by dashes
- Departure delay - by how many minutes the plane departed not as planned (positive if it was late, negative if it was ahead of schedule)
- Arrival delay – by how many minutes the plane arrived not as planned (positive if it was late, negative if it was ahead of schedule)
- Status - whether the flight has already taken place, or it is currently in progress or will take place in the future

Route order (~more than number of customers, more than millions entries) – set of all orders placed by customers (through the website) for route for groups of people or for themselves only. An entry is added when customer place new order and is deleted one year after the order is used up. Attributes:

- ID order (primary key) - index of the order in the database, natural number from 1 to number of orders
- Date - the day the order was placed, in format of 8 digits (2 for day, 2 for month and 4 for year of birth) separated by dashes

Route ticket (~more than number of passengers, more than millions entries) – set of all passenger tickets for a single route. New entry is added when new order is placed or someone added new ticket to his order and is deleted when order is removed from the database.

- Ticket number (primary key) – the number of ticket held by the passenger
- Ticket price – price paid for ticket by passenger with currency
- Luggage – type of purchased baggage
- Additional equipment – for instance a pram or a bicycle (empty values allowed)

3.2 Definition of relationships between entities

- Relationship **represents** between Aircraft and Aircraft_model entities – represents what model the plane is. Each Aircraft is representative of exactly one model (no optionality, since each Aircraft must be of some type). One or more aircrafts can represent each Aircraft Model (usually it is the case that airlines have many representatives of one kind, no optionality, since I assume that when there is no aircraft of particular model, this model is deleted from database).
- Relationship **is_assigned** between Aircraft_model and Scheduled_flight – represents the assignment of an aircraft model to a scheduled flight. To each Scheduled_flight is assigned exactly one fixed model of the aircraft, but each Aircraft_model can be assigned to zero, one or many scheduled flight (temporarily the model may not be used but is still in stock).
- Relationship **is_used** between Aircraft and Flight – represents exactly which aircraft is used for a particular flight. Each Flight is carried out using one plane (no optionality because it wouldn't make sense if there was no plane assigned to the flight), but each Aircraft can be used to zero, one or many flights (temporarily the plane may not be used as it can be under repair)
- Relationship **is_one_of** between Scheduled_flight and Flight – represents which of the planned flights this particular flight is. Each Flight is one of exactly one scheduled flight (no optionality since all flights are in the schedule). Since Scheduled_flight represents weekly flights, there are many flights from different weeks which represent them (no optionality as if the flight is in the schedule, the flights take place).
- Relationships **arrives_to** and **departs_from** between Scheduled_flight and Airport – represent from which airport, the plane takes off and to which it arrives. Each Scheduled_flight departs from and arrives to exactly one airport (no optionality as each flight must be interrupted). Each airport is used for multiple flights (optionality as the airport may not be in use for some time).
- Relationship **is_located** between Airport and City – represents in what city is the airport located. Each airport can be located in one city (no optionality since every airport is assigned to some city). In each city can be zero, one or more airports (since not in every city where the passengers live must be airport and there are some big cities where are more than one airport)
- Relationship **is_located** between City and Country – represent to which country the city belongs and is located. Each City is located only in one country, but each country can definitely have more than one city (no optionality since I assume that city-state is a country with one city).
- Relationship **lives_in** between City and Passenger – represent where each passenger lives. Of course in each city live many people (optionality since in some cities, where the airport is located, may not live any passenger). But each passenger can live only in one city.
- Relationship **is_citizen** between Passenger and Country – represent what countries passengers are citizens of. Each passenger must be a citizen of some country (may be even multiple countries). Every country must have many citizens (otherwise it does not exist so no optionality).
- Relationship **identifies** between Passenger and Document - represents the personal documents held by the passenger, which are required at the time of purchasing the ticket and to confirm their identity. Each passenger may have more than one document confirming his identity (ID card and passport) and there is no optionality as each passenger must enter the data from the document when purchasing a ticket and then show it at the airport. Each document can only belong to one person (as it is their personal document). There is no optionality because each document has its owner.

- Relationship **confirms** between Document and Route_ticket - represents from which document the data was given to purchase the corresponding ticket. Each ticket must be confirmed by exactly one document (ID card or passport), therefore there is no optionality and multiple. However, when a passenger has at same time several tickets, one document may be used many times. But the second document held by passengers may not currently be assigned to any ticket, that is why I added an optionality.
- Relationship **includes** between Route_order and Route_ticket - represents which tickets are included in each order. When a customer places an order, he can purchase tickets for a group of people, therefore, one order can have many tickets inside (no optionality as there is no order without a ticket). However, one ticket is part of only one order (also without the optionality).
- Relationship **is_part_of** between Flight and Route_order - represents which flights are included in the route for which the order was placed. When the route is with changes, many flights are assigned to the order (there is no optionality since the order is placed for at least one flight). Each flight can also be part of many orders, for example a flight from Gdansk to Frankfurt can be part of a route from Gdansk to the United States and from Gdansk to Brazil (I added an optionality, because there is a possibility of a situation when there is no person willing for a particular flight, although it is in the database).
- Relationship **places** between Route_order and Customer - represents which customer placed the order. Each customer can place several orders (no optionality because placing an order distinguishes a customer from passengers), but each order can be assigned to only one customer (no optionality because someone had to place an order).
- Relationship **is_a** between Customer and Passenger – this a built-in relationship. It means that each entity from Customer entity set has all the attributes of Passenger entity set and some additional ones absent in Passenger. In this relationship, the point is that customer is also a passenger but the special one who placed the order for a group of people and provided his contact details. That is why this relationship is with cardinality 0..1:1 since not all passengers are customers, but each customer is also passenger.

4. Relational database schema

Aircraft_model (Model, Capacity, Producer)

Aircraft (Registration, Status, Production_day, Model REF Aircraft_model)

Airport (IATA_code, Name, Contact, Location_city_name REF City)

Scheduled_flight (ID_scheduled_flight, Scheduled_departure_day, Scheduled_departure_hour, Scheduled_arrival_day, Scheduled_arrival_hour, Aircraft_model_assigned REF Aircraft_model, IATA_departure_airport REF Airport, IATA_arrival_airport REF Airport)

Flight (Flight_number, Date, Departure_delay, Arrival_delay, Status, Used_aircraft_registration REF Aircraft, ID_scheduled_flight REF Scheduled_flight)

Route (Flight_number REF Flight, ID_order REF Route_order)

Route_order (ID_order, Date, ID_customer REF Customer)

Route_ticket (Ticket_number, Ticket_price, Luggage, Additional_equipment, Confirmation_document_number REF Document, ID_route_order REF Route_order)

Document (Document_number, Document_type, Date_of_expiry, ID_passenger REF Passenger)

Customer (ID_customer, E-mail, Phone_number, ID_passenger REF Passenger)

Passenger (ID_passenger, Names, Surname, Sex, Date_of_birth, Residence_city_name REF City)

City (Name, Population, Timezone, Country_name REF Country)

Country (Name, Continent)

Citizenship (ID_passenger REF Passenger, Country_name REF Country)