



Team 24

Flight Delay Prediction via Machine Learning

Teoman Berkay AYAZ

1800004169

Dennis BREZINA

1700004948

Kerem Safa DİRİCAN

1800002205

Zeynep Simge SEDEF

1700003227



Our end goals

The main purpose of the program will be to predict delays and cancellations. If possible, we aim to develop an API that can be implemented into any project in the future. (figs. 1.1, 1.2)

The diagram illustrates a transition between two states of a web form titled "Airline Delays". A blue arrow points from the left state (empty form) to the right state (filled form).

Left State (Empty Form):

- Title: Airline Delays
- Fields:
 - Airline
 - Weather (dropdown menu)
 - Wind Speed
 - Wind Direction
 - Takeoff Airport
 - Destination Airport
- Submit Button

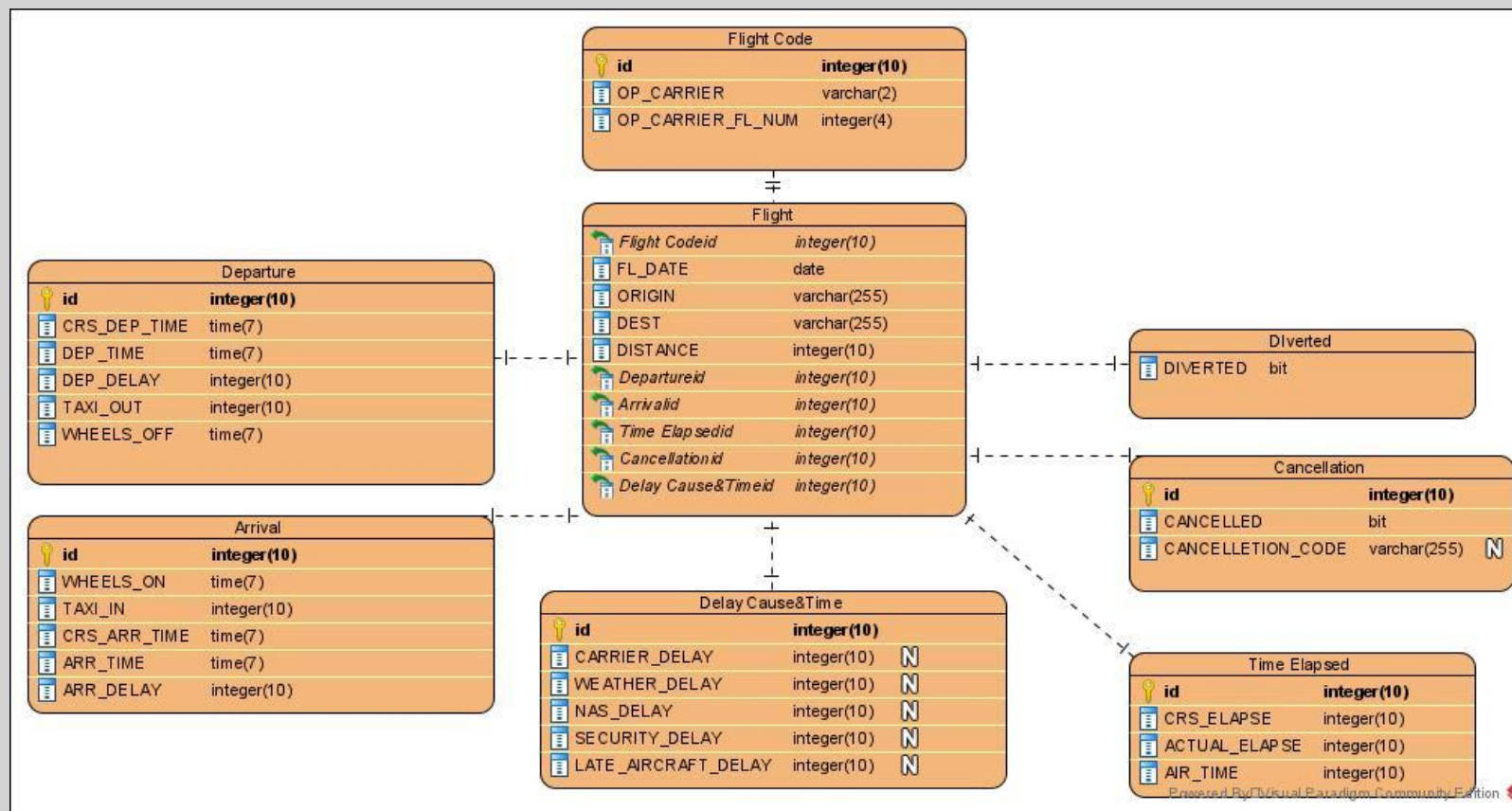
Right State (Filled Form):

- Title: Airline Delays
- Fields:
 - THY
 - Weather (dropdown menu)
 - 5 km/h
 - north
 - IST
 - LOS
- Submit Button

(figure 1.1)



(figure 1.2)

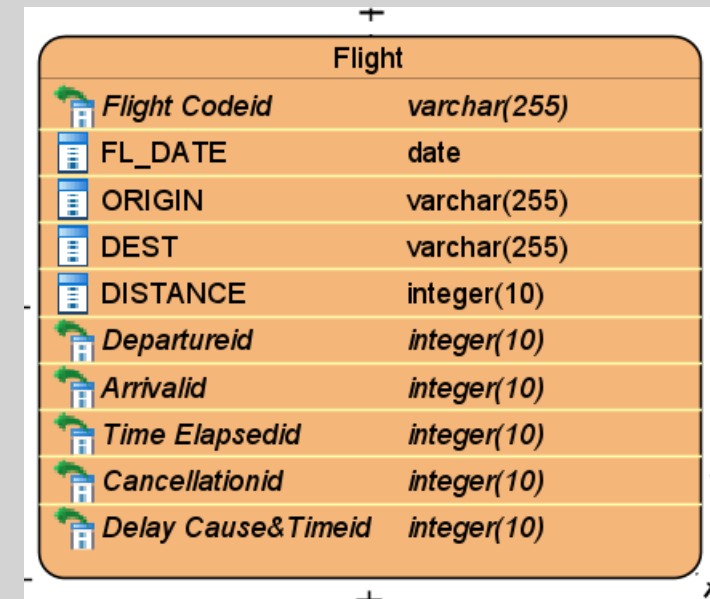












(figure 2)

As we have been instructed we have designed and created a database for our project(figure 2). Our database has 8 tables each with their respective data types and relationship.

Vpp and Db link > https://drive.google.com/file/d/1ZxF1XK3Tx2zzGwKHwbBVCdmlh8_FU-LX/view?usp=sharing

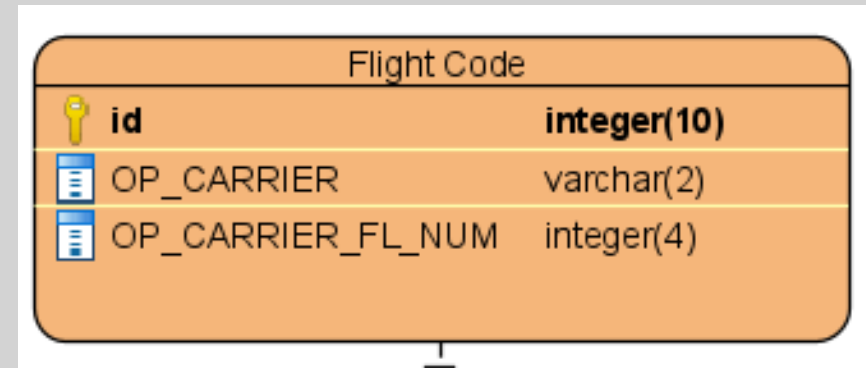
Our main table in our database that is connected to all other tables is the Flight table (fig 3.1). It contains data on our destination, flight date, distance, origin.

A diagram of a database table named 'Flight'. The table is represented as an orange rounded rectangle with a black border. Inside, there are ten rows, each representing a column. Each row starts with a small icon: a green house for 'Flight Codeid', a blue document for 'FL_DATE', 'ORIGIN', and 'DEST', and a green house with a red 'X' for 'Departureid', 'Arrivalid', 'Time Elapsedid', 'Cancellationid', and 'Delay Cause&Timeid'. The columns are: 'Flight Codeid' (varchar(255)), 'FL_DATE' (date), 'ORIGIN' (varchar(255)), 'DEST' (varchar(255)), 'DISTANCE' (integer(10)), 'Departureid' (integer(10)), 'Arrivalid' (integer(10)), 'Time Elapsedid' (integer(10)), 'Cancellationid' (integer(10)), and 'Delay Cause&Timeid' (integer(10)). The table is titled 'Flight' at the top center. There are small black crosshair symbols at the top and bottom center of the table diagram.

Flight	
 <i>Flight Codeid</i>	varchar(255)
 FL_DATE	date
 ORIGIN	varchar(255)
 DEST	varchar(255)
 DISTANCE	integer(10)
 <i>Departureid</i>	integer(10)
 <i>Arrivalid</i>	integer(10)
 <i>Time Elapsedid</i>	integer(10)
 <i>Cancellationid</i>	integer(10)
 <i>Delay Cause&Timeid</i>	integer(10)

(fig 3.1)

Our second table is Flight Code(fig 3.2).It contains the code of our flight, And since every flight code is Unique we gave it the id data to separate the codes.



(fig 3.2)

Our next tables are the Departure (fig 3.3) and Arrival (fig 3.4) tables. We keep our info about the departures and arrivals on these tables. Each one has their respective IDs to access data about individual departures and arrivals.

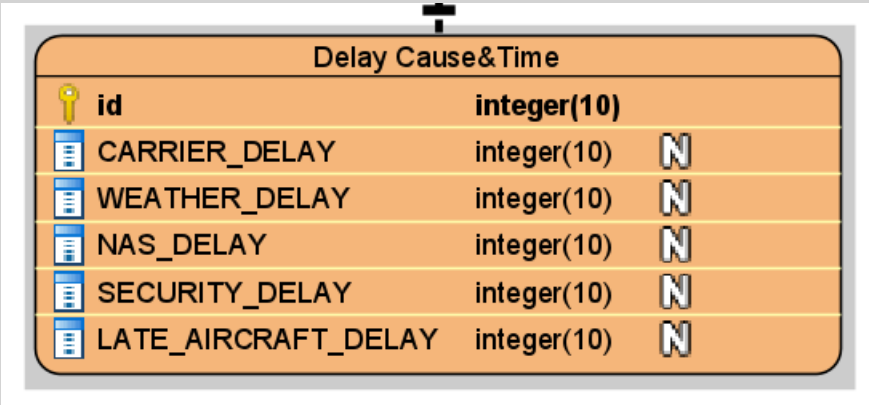
Arrival	
id	integer(10)
WHEELS_ON	time(7)
TAXI_IN	integer(10)
CRS_ARR_TIME	time(7)
ARR_TIME	time(7)
ARR_DELAY	integer(10)







(fig 3.3)

Departure	
id	integer(10)
CRS_DEP_TIME	time(7)
DEP_TIME	time(7)
DEP_DELAY	integer(10)
TAXI_OUT	integer(10)
WHEELS_OFF	time(7)

(fig 3.4)

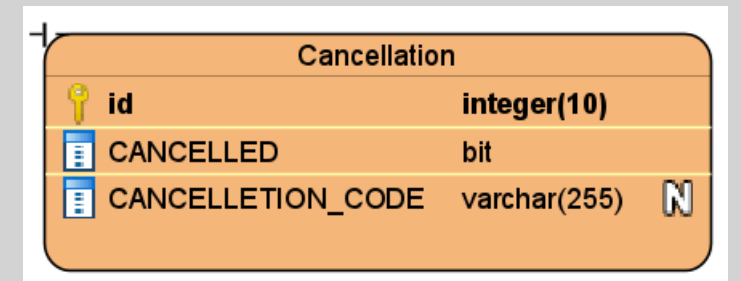
Our next table will be the Delay Cause&Time (fig 3.5) . We keep info about whether there has been a delay because of a unique issue like security, carrier caused delay etc. or not.



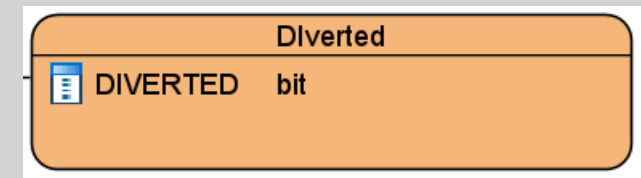
Delay Cause&Time			
	id	integer(10)	
	CARRIER_DELAY	integer(10)	N
	WEATHER_DELAY	integer(10)	N
	NAS_DELAY	integer(10)	N
	SECURITY_DELAY	integer(10)	N
	LATE_AIRCRAFT_DELAY	integer(10)	N

(fig 3.5)

Following that, we have our Diverted (fig3.6) and Cancellation (fig 3.7) tables. We keep the info of whether the flight has been cancelled, and if cancelled the cancellation code and if the flight has been diverted or not.

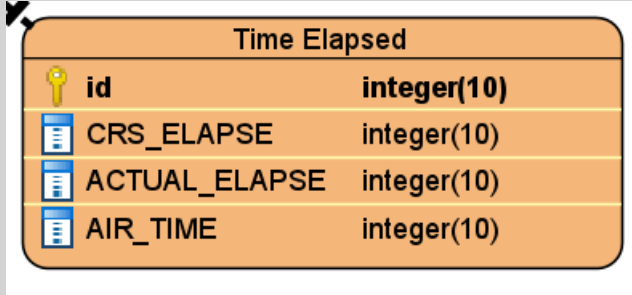






(fig 3.6)



(fig 3.7)

Our last table is Time elapsed (fig 3.8) table. It keeps the info of the elapsed time and air time of the carrier.



Time Elapsed	
 id	integer(10)
 CRS_ELAPSE	integer(10)
 ACTUAL_ELAPSE	integer(10)
 AIR_TIME	integer(10)

(fig 3.8)