Predicting Delayed Flight

The file FlightDelays.csv contains information on all commercial flights departing the Washington, DC area and arriving at New York during January 2004. For each flight, there is information on: the departure and arrival airports, the distance of the route, the scheduled time and date of the flight, and so on. The variable that we are trying to predict is whether or not a flight is delayed. A delay is defined as an arrival that is at least 15 minutes later than scheduled.

Variables include:

CRS\_DEP\_TIME : Departure Time

CARRIER

DEP\_TIME: actual departure time

DEST: destination

DISTANCE: distance flown, in miles

FL\_DATE: flight date

FL\_NUM : flight number

ORIGIN :origin city, airport code

Weather: Code of 1 represents there was a weather-related delay

DAY\_WEEK

DAY\_OF\_MONTH

TAIL\_NUM : Tail number

Flight Status: delayed or on-time