Part a :

URL of the webpage:
https://www.kaggle.com/datasets/divyansh22/flight-delay-prediction?resource=download

URL for the site:
https://www.kaggle.com/datasets/divyansh22/flight-delay-prediction?resource=download

Part -b

Brief description of data set:

--This dataset contains all the flights in the month of January 2020.

– There are more than 400,000 flights in the month of January throughout the United States. The features were manually chosen to do a primary time series analysis.

– There are several other features available on their website.

-- Objects: we have a data set of size and length (1191331, 23)

--Attributes: We have 23 attributes for the above data set. which is a combination of different types like DAY_OF_MONTH, DAY_OF_WEEK, **OP_UNIQUE_CARRIER, OP_CARRIER_AIRLINE_ID, OP_CARRIER, TAIL_NUM, OP_CARRIER_FL_NUM, ORIGIN_AIRPORT_ID**

	tipo
DAY_OF_MONTH	int64
DAY_OF_WEEK	int64
ORIGIN	object
DEST	object

DEP_TIME	float64
DEP_DEL15	float64
DEP_TIME_BLK	object
ARR_TIME	float64
ARR_DEL15	float64
CANCELLED	float64
DIVERTED	float64
DISTANCE	float64
year	int64

Part -c:

- This data could well be used to predict the flight delay at the destination airport specifically for the month of January in upcoming years as the data is for January only.
- This file contains all the flights starting from 1st January 2020 till 31st January 2020. There are around 1191331 rows in this file and 23 feature columns indicating the features of the flight including information about the origin airport, destination airport, airplane information, departure time, and arrival time.
- By using the current month data set related to flight delays and their arrival times, we can predict the upcoming flight delays or arrival times.

Part -D:-

- It will be pretty useful for the people who travel a lot from one place to another and plan their journey according to the predictions.
- From this data set, we can even tell which career is maintaining the timeline and arriving on time at the airport.
- We can even say when a flight is canceled due to what by checking the cancelled column in the data set
-