Case Study 1:

Problem: Flights to trop mapping using generic code which can work for any number of connecting flights i.e recursive merges.

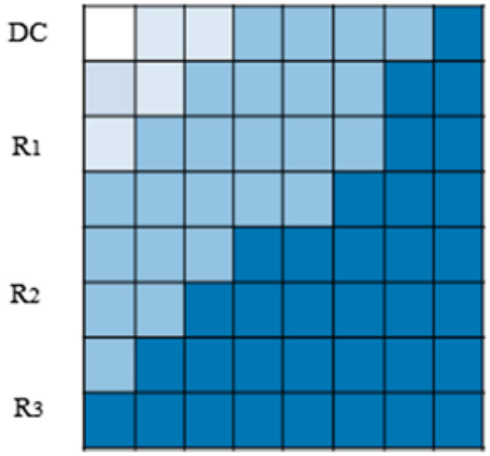
```
/*

Database of Airlines

Table1: Customer Id, Ticket ID, Origin, Destination, Start time and end Time

Table2: Flight Number, Date, Origin, Destination, Start time and end time of the flight


Assumption: There can be only one flight taking off from an airport at a given time


Desired Output – for given cust_id, ticket_id, need sequence of all flight_no used for the trip (all connecting flights)

PS: You need to get all possible sequences for the trip.

Table1 : cust_id, ticket_id, first_cust_port, last_cust_port, cust_start, cust_end

Table2 : flight_no, date, orig_port, dest_port, flight_start, flight_end

*/
```

```
data cust_data;

   input cust_id $ ticket_id first_cust_port $ last_cust_port $
cust_start cust_end;

   datalines;

      A123 9342312 Delhi Mumbai 815 1730

      A456 6542312 Delhi Mumbai 830 1015

      A789 7542312 Delhi Mumbai 845 1330

;


data flight_data;

   input flight_no $ date $ orig_port $ dest_port $ flight_start
flight_end;

   datalines;

      B99 04042014 Delhi Mumbai 830 1015

      B89 04042014 Delhi Jaipur 845 1045

      B69 04042014 Jaipur Mumbai 1130 1330

      B79 04042014 Delhi Agra 815 1130

      B59 04042014 Agra Surat 1230 1430

      B49 04042014 Surat Pune 1445 1530

      B39 04042014 Pune Mumbai 1615 1730

;
```

Case Study 2:

Problem: Take a random sample of 10 images from any standard dataset. Process the following steps:

1. For every 8x8 non-overlapping block in the image, apply Discrete cosine transform (DCT). Each block is then transformed into three **regions R1, R2, and R3** which represent the **low frequency (LF), medium frequency (MF) and high frequency (HF) regions**, respectively as shown in the image.



Representation of (8 × 8) DCT bloc Frequency bands

2. On each frequency range, local statistics (**mean, variance, kurtosis, skewness, entropy, and energy** of all AC components constituting each region) are calculated.

3. The vector descriptor of the image consists of **the averages of respective statistics found in all the blocks** of the image.

The descriptor vector will therefore be composed of **18 features.** (6 features/region × 3 Regions = 18 features)

4. Create a data frame which consists of the vector descriptor for all the images in the sample. Give appropriate names for each of the engineered features.

You are required to share the modular and efficient code for the above process along with the csv of the vector descriptor data frame.