

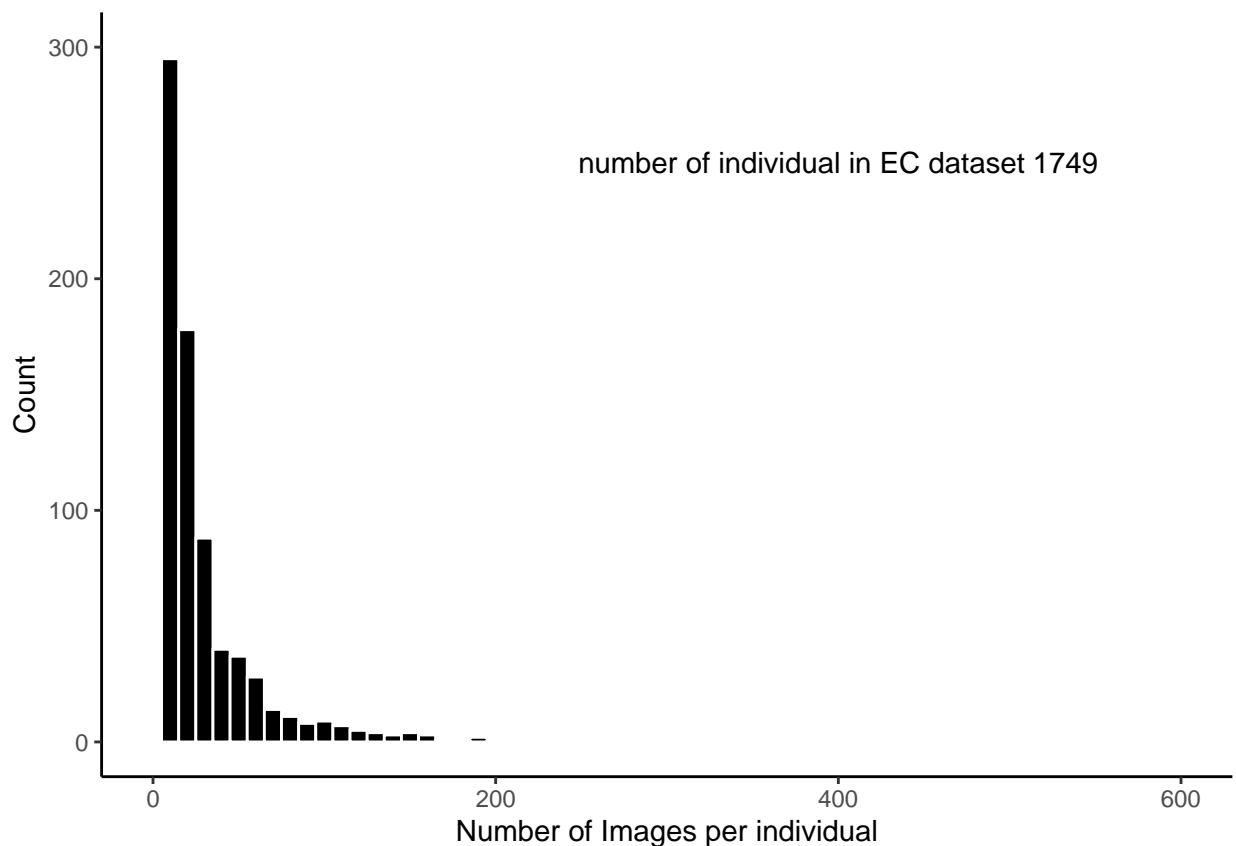
Descriptive analysis of Images in the database

Chandan Kumar Pandey

This report is the descriptive result of the tiger database which we will be using for training the machine learning network.

The data can be loaded with the Running the source-code “Fetch_data_from_database.R”. This source code can be downloaded from the git repo. You can change the **query** in the code to extract different information. The method to connect to MS-Access is described in “*Connect_to_database_odbc.pdf*”. For now I have extracted the information earlier and saved as CSV.

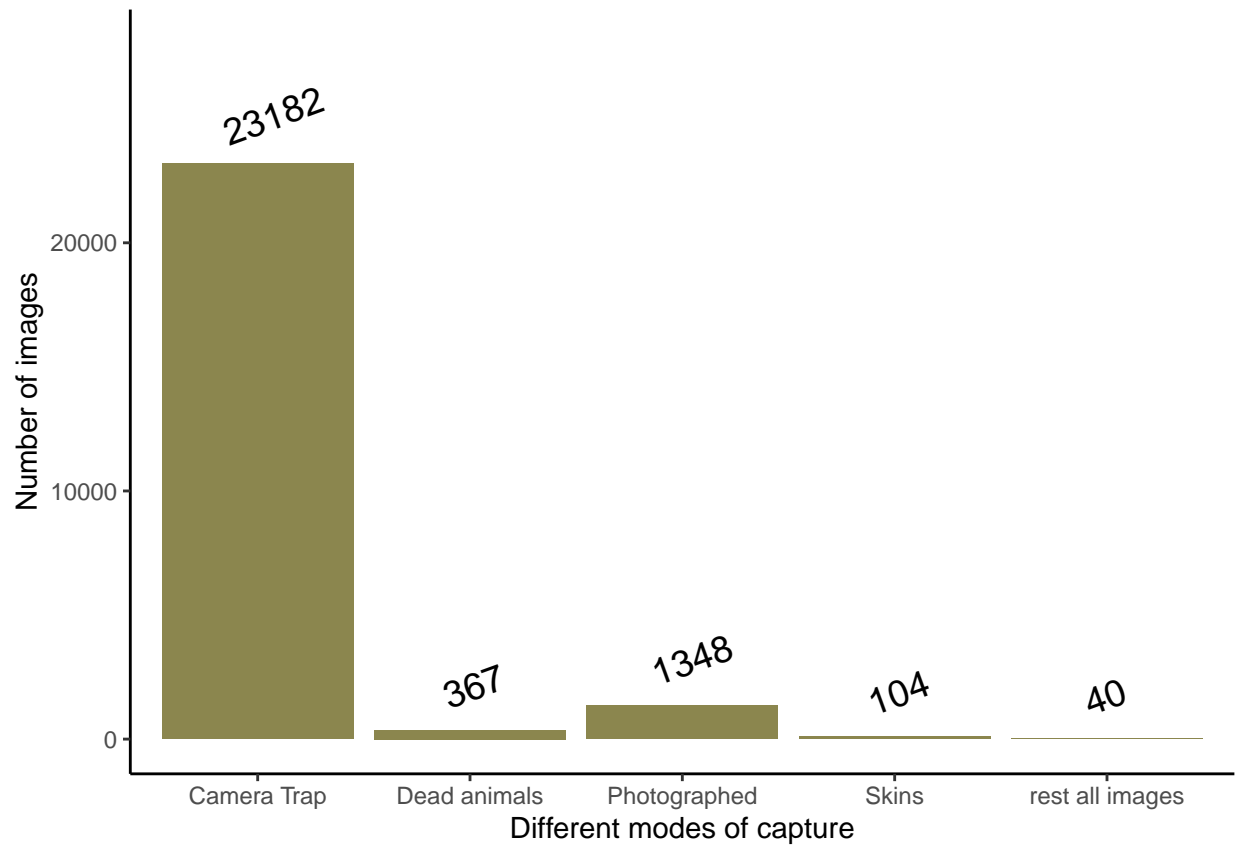
Histogram of number of image per individuals.



From the figure above it is clear that most of the individuals have few images. For better we can also see the quantile distribution.

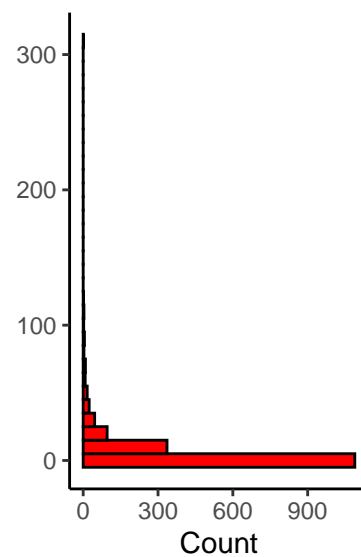
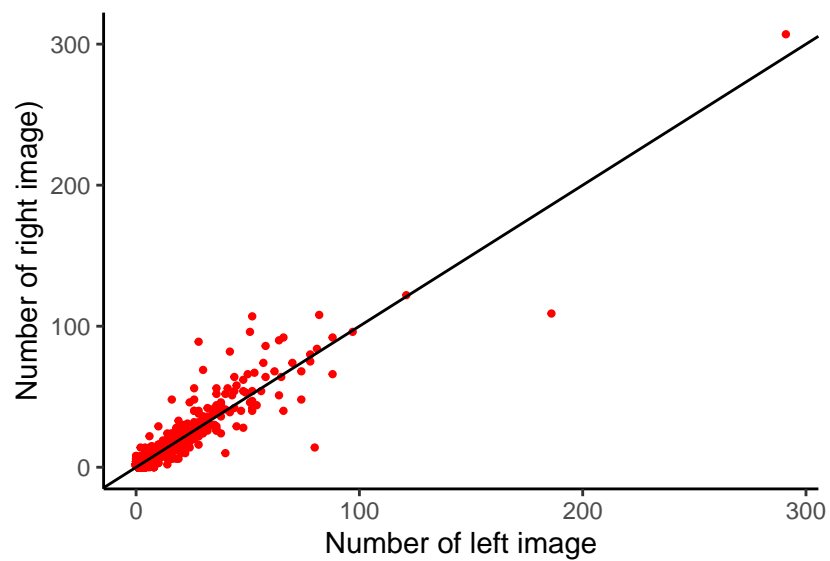
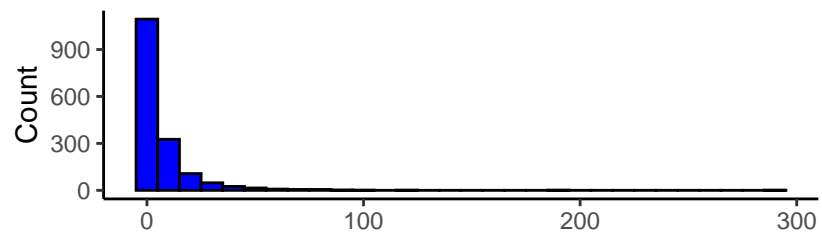
quantile	value
0%	1
25%	2
50%	4
75%	16
100%	598

It is also important to understand how the image were taken in the database. There are various methods by which the tiger had been captures, viz. “Camera Trap”, “Dead animal photo”, “Photographed animals” etc



It is clear that **93%** of data is camera trap followed by photographed images which together makes **98%** of total images.

Further camera trap images are take in left and right flank. The number of left and right flanks might not be same for all the tigers.



As from the discussion it was clear that night and day image was important in classification of the image. It is also important which training the network