Initial Data Exploration:

The data contains a total of 2864056 entries of time series data with 33 columns

Entries are spaced by 0.01s

The variable we predict activityID, which has 13 different values.

We are able to find 193 task switch of various time lengths

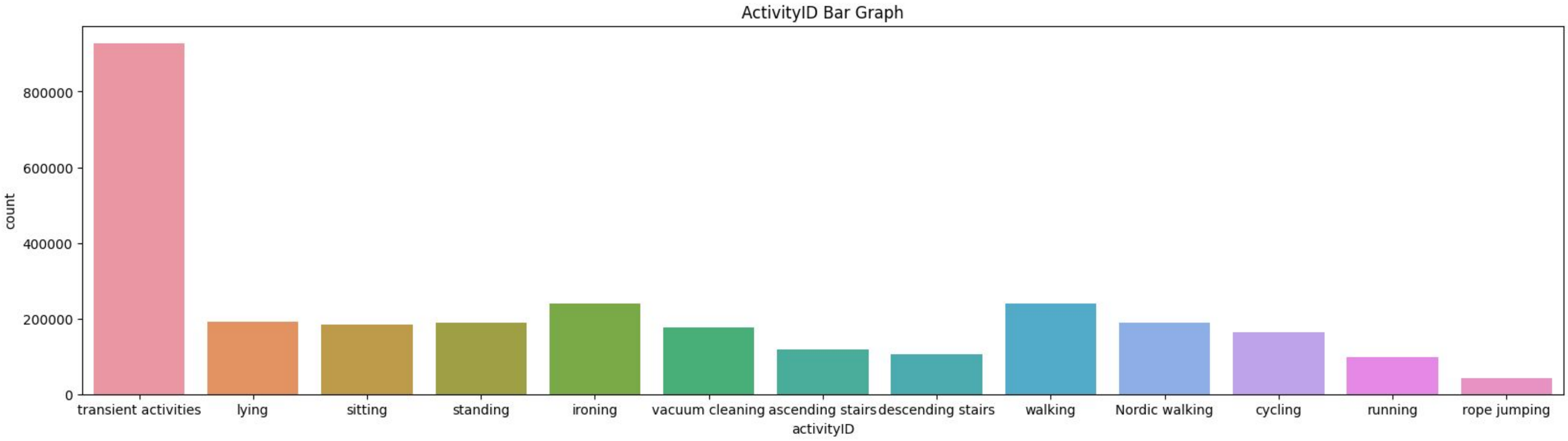
```
RangeIndex: 2864056 entries, 0 to 2864055
Data columns (total 33 columns):
```

Samples are float64

Initial Data Exploration:

The histogram below displays the counts of each ActivityID label

While ‘transient activities’ is the most common, most other labels have similar frequencies

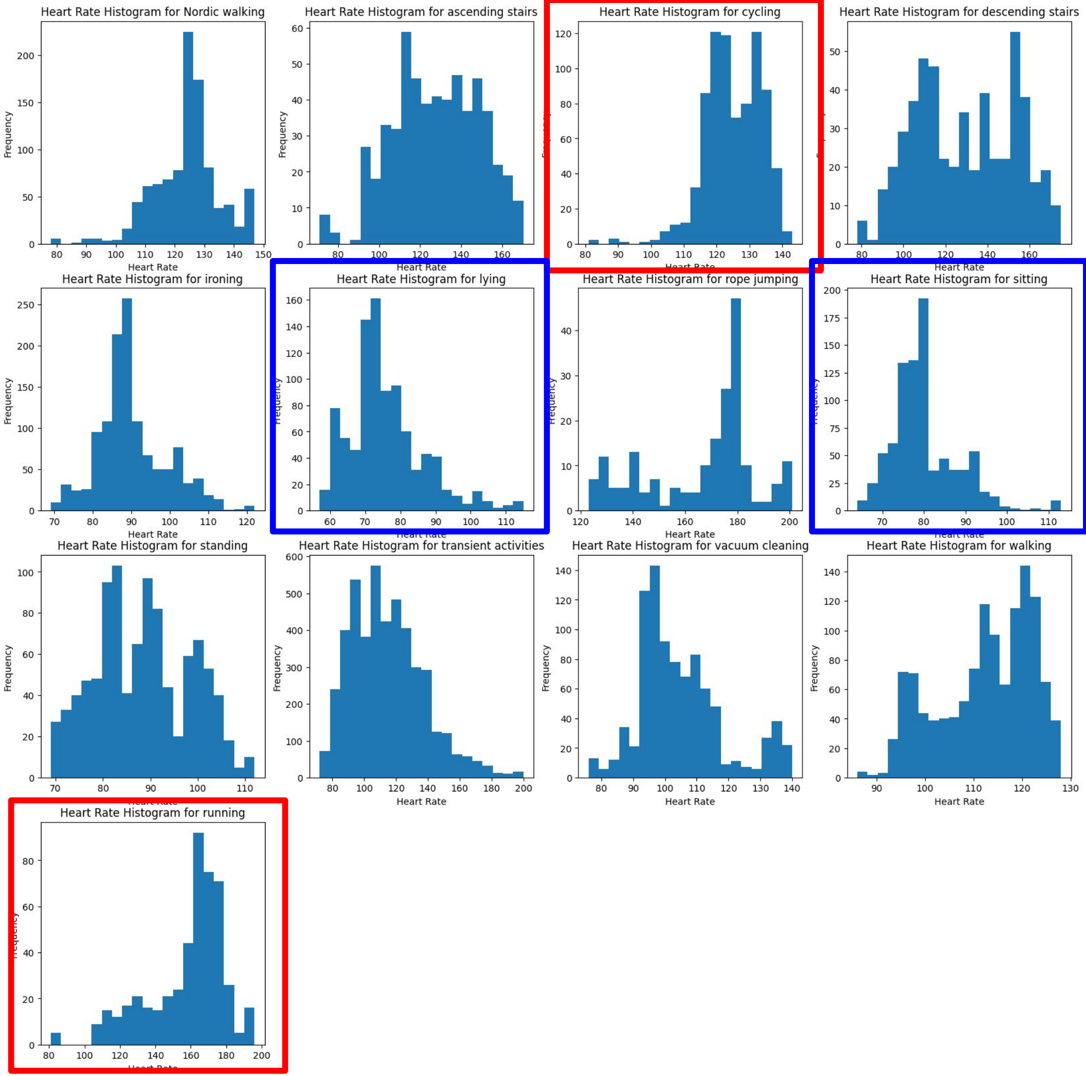


Initial Data Exploration:

We suspect heart rate has a high feature importance

The heart rates for cycling and running are much higher than the heart rates for sitting and laying down.

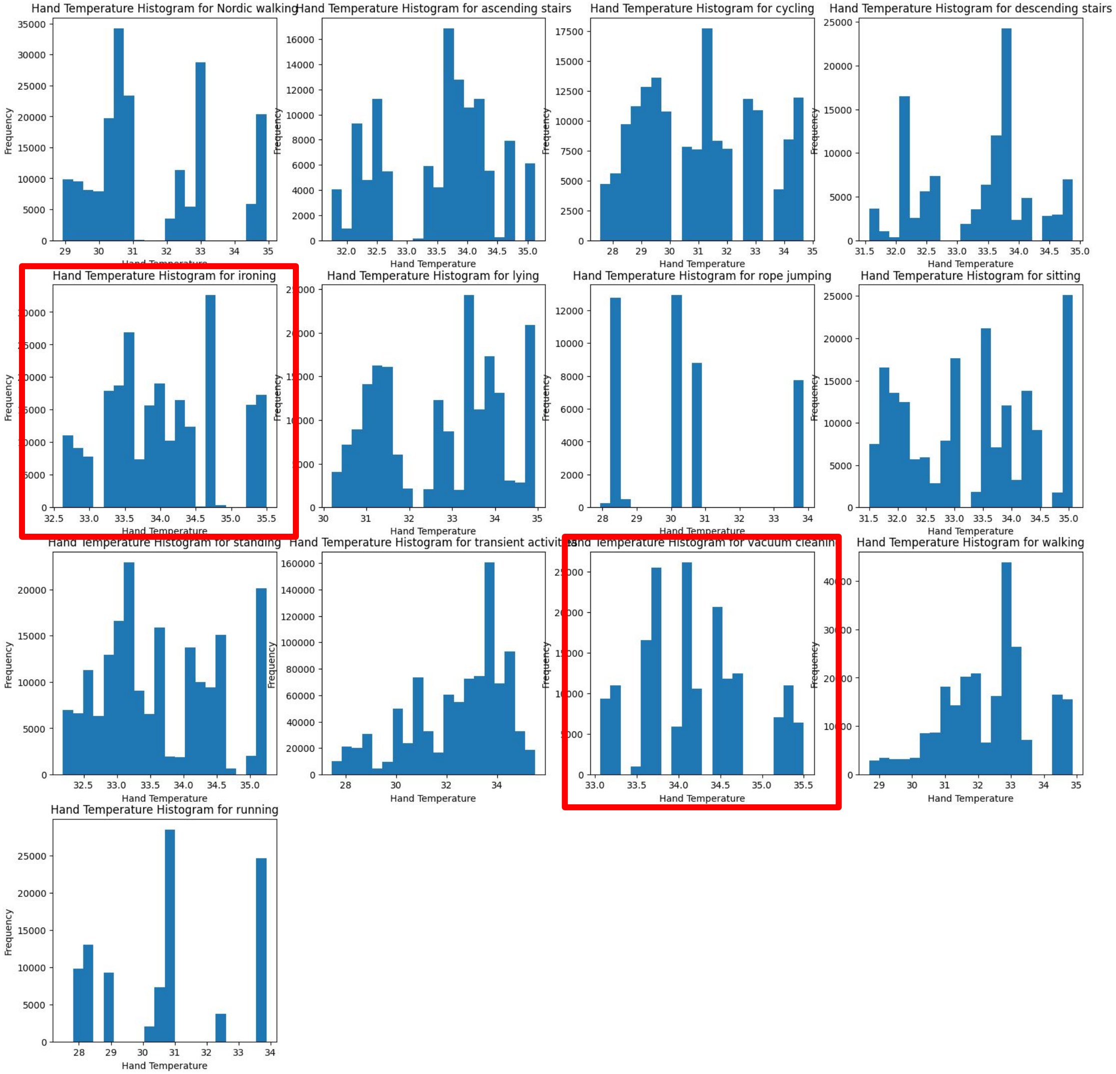
Other activities are in between



Initial Data Exploration:

In the case of hand temperature, most labels have ranges between 28ºC and 35ºC.

Hand temperatures for ironing and vacuuming are all greater than 32ºC

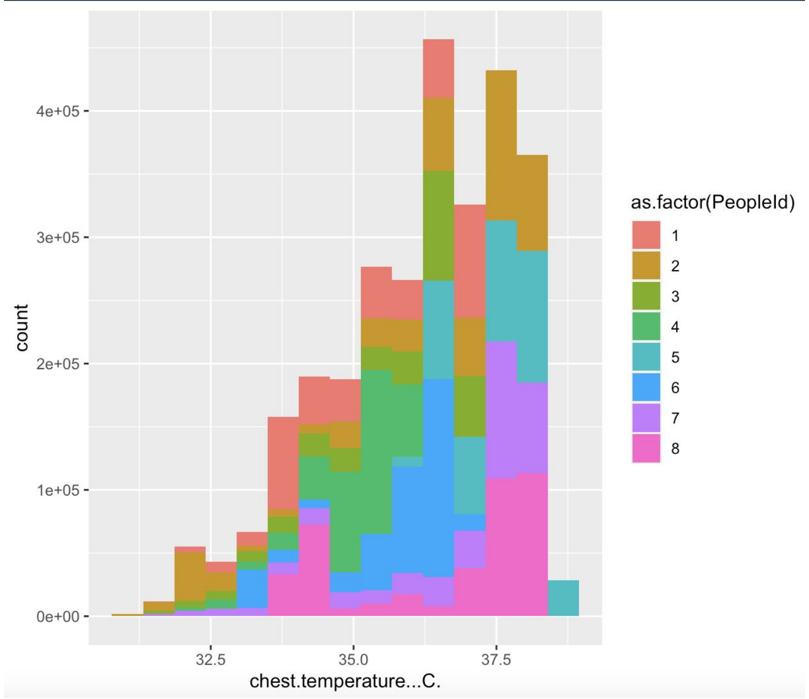


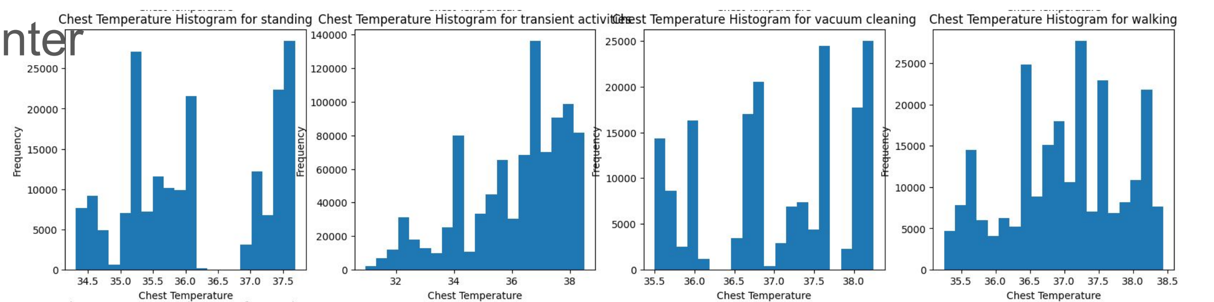
Initial Data Exploration:

For chest temperature, some labels have isolated peaks

One possible explanation is that different people have different average chest temperatures

We can see that people 2, 5, 7, and 8 have chest temperatures above 37ºC, while the rest center around 35ºC.





Initial Data Exploration:

Other features, like X, Y, and Z measurements of acceleration, gyroscope, and magnetosphere are less intuitive.

This graph explores X, Y, and Z chest acceleration in a time series of jump rope.

Notice high acceleration shown by the blue line, which is the Y acceleration: we can intuit this is the bouncing motion

Both X and Z are much lower, and signify slight motion forward/back and left/right

Data Cleaning:

A brief look at missing data →

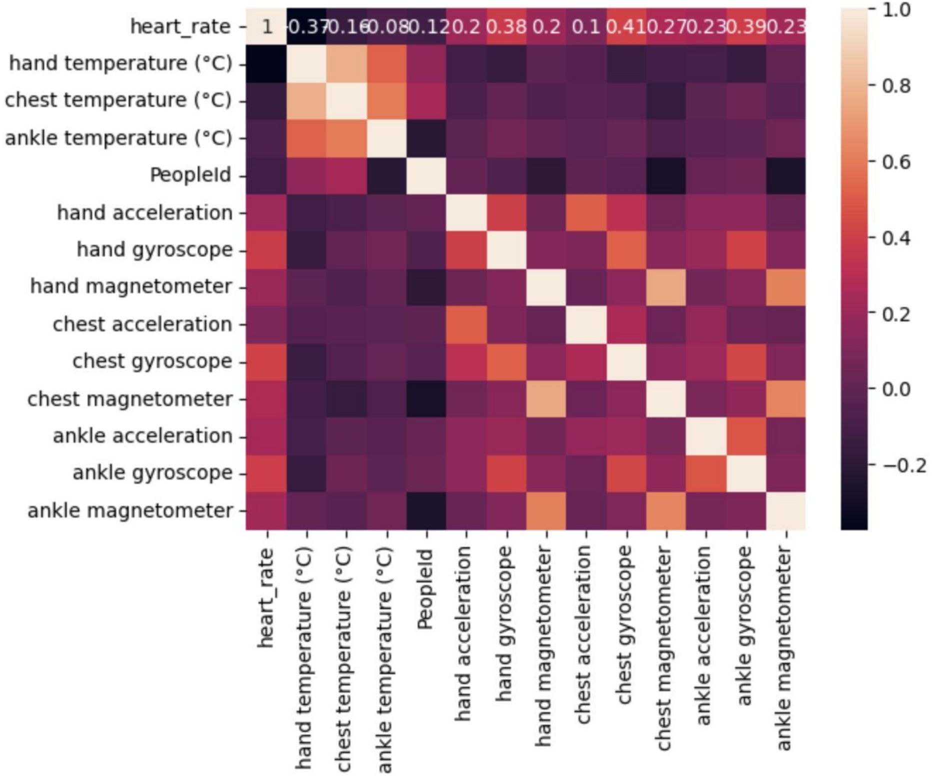
Tiny fraction of missing heart_rate data

We will drop these samples

```
activityID                  0
heart_rate                 46
hand temperature (°C)       0
hand acceleration X ±16g    0
hand acceleration Y ±16g    0
hand acceleration Z ±16g    0
hand gyroscope X            0
hand gyroscope Y            0
hand gyroscope Z            0
hand magnetometer X         0
hand magnetometer Y         0
hand magnetometer Z         0
chest temperature (°C)      0
chest acceleration X ±16g   0
chest acceleration Y ±16g   0
chest acceleration Z ±16g   0
chest gyroscope X           0
chest gyroscope Y           0
chest gyroscope Z           0
chest magnetometer X        0
chest magnetometer Y        0
chest magnetometer Z        0
ankle temperature (°C)      0
ankle acceleration X ±16g   0
ankle acceleration Y ±16g   0
ankle acceleration Z ±16g   0
```

Data Cleaning:

Plotting feature correlation we shows very little collinearity, suggesting that we do not need to drop features.



Data Sampling:

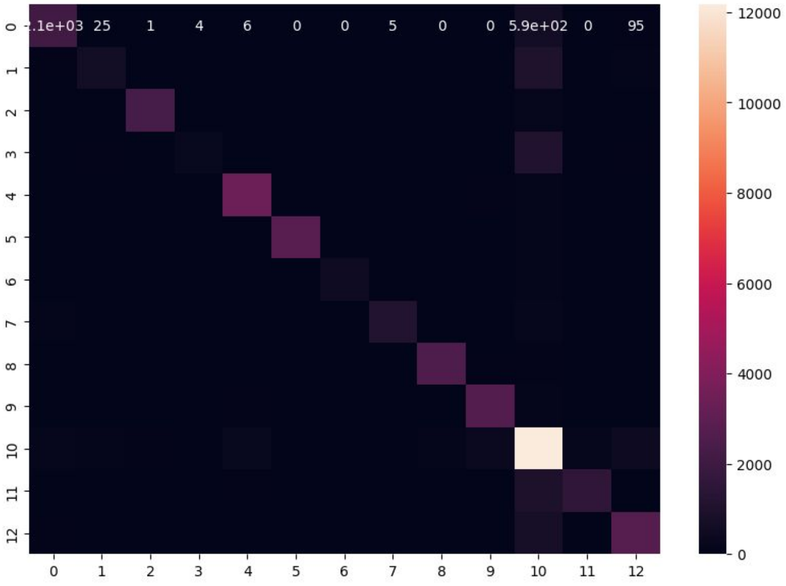
Because our cases are balanced, we do not expect to sample the data.

However, because we will be training a multiclass classifier using One vs. Rest, our data is naturally imbalanced.

To address this, we can oversample, sample with SMOTE, or apply weights to the ‘One’ class.

Machine Learning Techniques

SVM: accuracy 0.8041

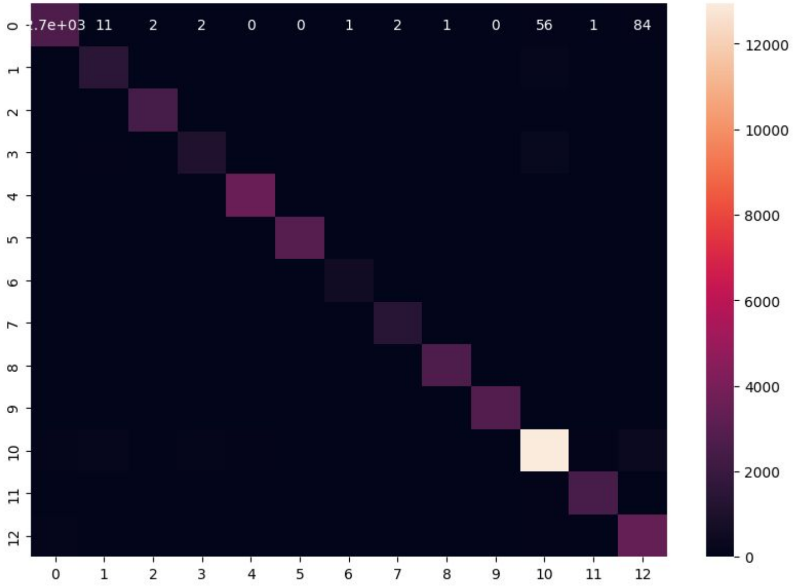


	precision	recall	f1-score	support
Nordic walking	0.80	0.74	0.77	2838
ascending stairs	0.71	0.32	0.44	1804
cycling	0.93	0.92	0.92	2440
descending stairs	0.81	0.17	0.29	1541
ironing	0.88	0.95	0.91	3537
lying	1.00	0.96	0.98	2933
rope jumping	0.87	0.73	0.80	634
running	0.96	0.75	0.84	1499
sitting	0.95	0.92	0.94	2691
standing	0.83	0.92	0.87	2850
transient activities	0.70	0.87	0.78	13936
vacuum cleaning	0.87	0.59	0.70	2658
walking	0.79	0.76	0.78	3599
accuracy			0.80	42960
macro avg	0.85	0.74	0.77	42960
weighted avg	0.81	0.80	0.79	42960

Machine Learning Techniques

Knn =3

accuracy : 0.9387



	precision	recall	f1-score	support
Nordic walking	0.87	0.94	0.91	2838
ascending stairs	0.82	0.82	0.82	1804
cycling	0.97	0.99	0.98	2440
descending stairs	0.86	0.71	0.78	1541
ironing	0.97	0.99	0.98	3537
lying	1.00	0.99	1.00	2933
rope jumping	0.97	0.94	0.95	634
running	0.98	0.94	0.96	1499
sitting	0.99	0.99	0.99	2691
standing	0.98	0.99	0.98	2850
transient activities	0.94	0.93	0.93	13936
vacuum cleaning	0.96	0.94	0.95	2658
walking	0.87	0.93	0.90	3599
accuracy			0.94	42960
macro avg	0.94	0.93	0.93	42960
weighted avg	0.94	0.94	0.94	42960