Data sources: accelerometer & gyroscope
75 total activities (squats, lunges, etc)
- For each activity: there are varying number of
  data signals (sets)

• For each set (stored as numpy array):
  · the amount of time varies from set to set
  · each set has 4 (x, y, z, t) features for each data point
  · so set shape is (seconds, 4)
                              ↑
                           varies

Segmentation Pre-Processing

· Data pre-processing for each.
  ① apply butterworth filter to normalize
     ↳ for each (x, y, z, t)
  ② create windows from the data for each (x, y, z)

idk if
we gotta
do it for t
(prob not)

| 3 | 4 | 5 | 4 | 3 | 4 | 5 | . . . . | n |

0.02   4.8sec
1 point   249 points

Resulting data format:

activity 1 :
  Set 1

$X_1 X_2 X_3 X_4 \ldots X_{250}$

	window 1	
$\otimes$	window 2	
	⋮	
	window T/s	

$Y_1, Y_2, Y_3, Y_4, \ldots Y_{250}$

	window 1	
$\textcircled{Y}$	window 2	
	⋮	
	window T/s	

Dataframe we want: (schema)

Activity name, set #, $t_1 \ldots t_{250}$, $x_1 \ldots x_{250}$, $y_1 \ldots y_{250}$, $z_1 \ldots z_{250}$, $\underbrace{t_1 \ldots t_{250}, x_1 \ldots x_{250}, y_1 \ldots y_{250}, z_1 \ldots z_{250}}$

$\underbrace{}$
1000 from each (gyroscope & accelerometer)

Segmentation Feature Computation

Added features (8)

(1000 columns from acc)
- X raw data (250)
- PC1: x, y, z for every 5 second interval (250)
- PC1: y, z " (250)
- xmag: $(x^2 + y^2 + z^2)^{1/2}$ (250)

(1000 columns from gyro)
- X raw data
- PC1: x, y, z for every 5 second interval
- PC1: y, z "
- xmag: $(x^2 + y^2 + z^2)^{1/2}$

$250 \times 4 \times 2 = 2000$ embeddings for each window

x 250
↓
5 secs

window | x | y | z
t_1
t_2
t_{250}
250 X 3

250
PCA

Next ──►

Then transpose:
| PCA_1 XYZ | PCA_2 YZ |
← 250 → ← 250 →
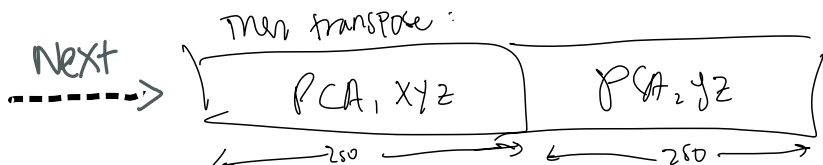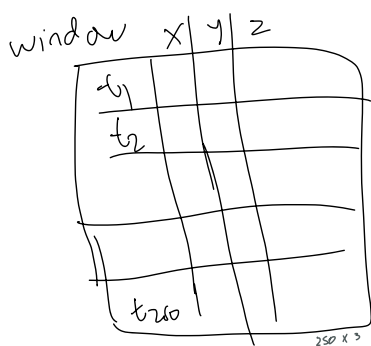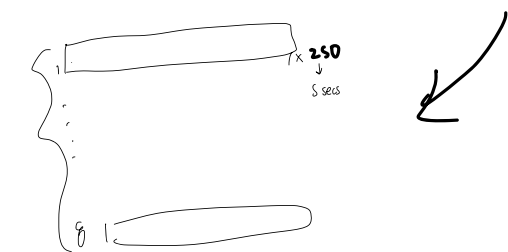
Resulting schema:

set-num (int),
activity (str),
g-x0-gyroscope, ... g-x2 49-gyroscope (float),