

```
# Pseudo code for function teamEM
```

```
# Inputs:
```

```
# data - data frame with N observations and 3 columns: FishID, Length, Age (k numbered groups for known, and NA for unknown).  
# epsilon - desired tolerance value for convergence.  
# maxit - maximum number of iterations.
```

```
# Outputs:
```

```
# inits - dataframe with mu, sigma, lambda for k age groups.  
# posterior - posterior probabilities of N observations in k age groups.  
# estimates - new mu, sigma, lambda for k age groups.  
# converged - TRUE if convergence is met before the maximum iteration, FALSE otherwise.  
# likelihood - vector of length equals to number of iterations, maximum  
# length equals to maxit.
```

```
teamEM <- function(data, epsilon = 1e-08, maxit = 1000) {
```

```
  # 1. Initialisation
```

```
  # Assign labels to fish with latent age variables based on data from fish
```

```
  # with known age:
```

```
  # Take the the values of the probability density functions for each fish length, compare them, and assign the age groups  
  in
```

```
  # accordance with the highest pdf value.
```

```
  # Take the data with the known and assigned values and calculate mean (mu), sd (sigma) for each age group.
```

```
  # Calculate probability (lambda) for a fish to be in each age group.
```

```
  # Repeating expectation, maximization and testing of convergence.
```

```
  # Create a for-loop with maximum iterations of maxit times.
```

```
  for (i in 1:maxit) {
```

```
    # 2. Expectation
```

```
    # Calculate probabilities for fish belonging to each age group given the observed lengths.
```

```
    # Use inits as a start and then use estimates from previous iterations
```

```
    # 3. Maximization
```

```
    # Calculate new estimates for mean, sd, and probability for each age group based on the posterior probabilities.
```

```
    # 4. Testing Convergence
```

```
    # Calculate the log-likelihood with parameters of this iteration.
```

```
    # If the difference between log-likelihood of this iteration and the previous is smaller than tolerance value, then  
    convergence is met.
```

```
    # Break the for-loop if convergent
```

```
    if (log(ith likelihood) - log((i-1)th likelihood) < epsilon)
```

```
      break.
```

```
    # If not convergent, reassign Age groups to latent variables in the dataframe the same way it has been done at the
```

```
    # Initialisation step.
```

```
    # Go back to the E step. Maximum amount of iterations - maxit.
```

```
  }
```

```
}
```