1. Principal tensor analysis
   1. Manipulate the data to fit requirements of the analysis: specifically
      1. No missing data allowed
         1. Want sampling locations that occurs throughout the dataset
         2. Want fish species that are caught throughout the dataset
      2. Species rarity
         1. Super rare fish (many zeroes) will throw off the model
         2. Too little number of species though will cause uninformative groups
         3. For my proportion of zeroes analysis: only look at top 5 stations that the species have historically been caught to calculate days during which there was 0 catch
   2. Transform the catch data
      1. Convert to CPUE to standardize effort/volume
      2. Transform your data if required
         1. Intercept model can be without transformation, of which you can check various transformations against
         2. Ensure data is transformed into same distributions if possible
      3. All data input should be in an array format, which each array representing a tensor of the model
   3. Run the models
      1. Several filters to run the model:
         1. Year filter
         2. Station filter
         3. Species filter, proportions of zeroes
         4. Explicit species filter (for species that are not filtered out at B)
   4. Tune the model
      1. Sensitivity analysis of number of species to use (prop of zero)
         1. Optimize location in which eval metric is not changing much, maximizing the number of species included
      2. Find number of tensors to keep
         1. Scree plot
         2. Lowest number of tensors possible
   5. Extract projection DF
      1. Projections are data contained by the tensors selected
   6. Apply clustering analysis
      1. Optimize clusters based on various methods:
         1. Ward.D2
         2. Single
         3. Complete
         4. Average
         5. Evaluation metrics: which method to use
            1. Cophenetic correlation
            2. Gower’s distance
      2. Optimize number of clusters
         1. Evaluation metric: silhouette length
   7. Visualize the dendogram
      1. Uses distance matrix from clustering analysis of the PTA tensors
      2. Assess clusters by space and time (assuming a 3 tensor analysis)
      3. Check silhouette width across the clusters for cluster agreement
   8. Tanglegram
      1. Used to evaluate community structure between two dendograms
      2. Examples:
         1. Removal of sampling stations
         2. Two data sources