Email:

Dear Dr. Imaginary Collaborator,

I have made some headway into the Starfish Detection project and I am currently experimenting with some image processing steps. Since this is my first time working with underwater image data, I wanted to ask you the following questions:

1. Does this species of starfish have any distinctive features that might be worth highlighting via image processing steps? For example, patterns on its body and/or prominent colours?
2. The training images are typically of corals. Are the starfish typically in a special form while feeding on coral? For example, flat or wrapping around a surface?
3. There isn't much colour variance in the images (i.e., blue water, grey-esque images) so I am considering a first attempt by transforming the images to grayscale. Do you know if this approach has been tried already and/or results? Do you have any thoughts on this?

I appreciate your time and look forward to hearing about any insights you may have!

Best regards,
Bushra

Interview:

Isaac:
- *Data*: time series data that captures the brightness of 3 different stars
- Goal: given measurements of a star's brightness at some time interval, detect whether a stellar flare occurred
- *Approaches*: non-parametric model fitted to unsupervised data; simulating synthetic data for known stars with no flares to establish a baseline
- *Questions*: how would you simulate the data (e.g. a parametric distribution)? What types of models might be appropriate for time series data?

Vishu
- *Data*: sonar data obtained by blasting 3 species of fish (trout, white fish, small-mouth bass) with 400 frequencies and capturing the reflected sound waves
- *Goal*: classify and distinguish between the 3 species of fish based on the reflected sounds
- *Approaches*: tree-based models (e.g. decision tree, XGBoost) or a neural network
- *Questions*: related to EDA, how would you compare the distributions of the 3 species over the 400 different frequencies (since that's a lot of plots)? Would binning the frequencies or random sampling frequencies help to compare?