**Q3 research question code explanation, output and inference (Simarleen)**

**Research Question:** Are the patterns of changes through time different for size than for other parameters indicating different selection pressure?

**Code Explanation:**

**1. Importing Libraries**

* **pandas**: Used for data manipulation and analysis.
* **matplotlib.pyplot**: Used for creating plots/graphs.
* **seaborn**: Used for statistical data visualization, like line plots.
* **sklearn.preprocessing.MinMaxScaler**: Used for normalizing/standardizing data (scaling it to a range of 0-1).
* **scipy.stats.linregress**: Used for calculating linear regression, which helps in determining the slope of a trend line.

**2. Loading and Processing Data**

* **df\_sorted['Age\_Ma']**: This converts the 'Age\_Ma' column into numeric values, ensuring there are no errors in the age data.
* **Group Data by 'Age\_Ma'**: The data is grouped by the age in millions of years (Age\_Ma), and the average of the parameters (like Max Diameter, Min Diameter, Elongation, etc.) is calculated for each age group.

**3. Smoothing the Data**

* **Rolling Mean**: The rolling() function with a window of size 10 is used to smooth out fluctuations in the data. This means for each point, the average of the 10 surrounding points is calculated. This makes the trend less jagged.

**4. Normalizing the Data**

* **MinMaxScaler**: This normalizes the other parameters (Elongation, Shape Factor, Sphericity) so that they all fall within the same range (0 to 1). This makes it easier to compare them fairly since they could have different units or scales.

**5. Plotting the Data**

* **Two Subplots**: The code creates two graphs:
  1. **Size Trends (Max and Min Diameters)**: This shows how the size of the organism changes over time (in Age\_Ma).
  2. **Other Parameters (Elongation, Shape Factor, Sphericity)**: This shows how the other parameters change over time.
  3. The second plot uses **two y-axes**: one for Sphericity and another for Elongation and Shape Factor. This helps in comparing them on the same graph, even if their values are different.

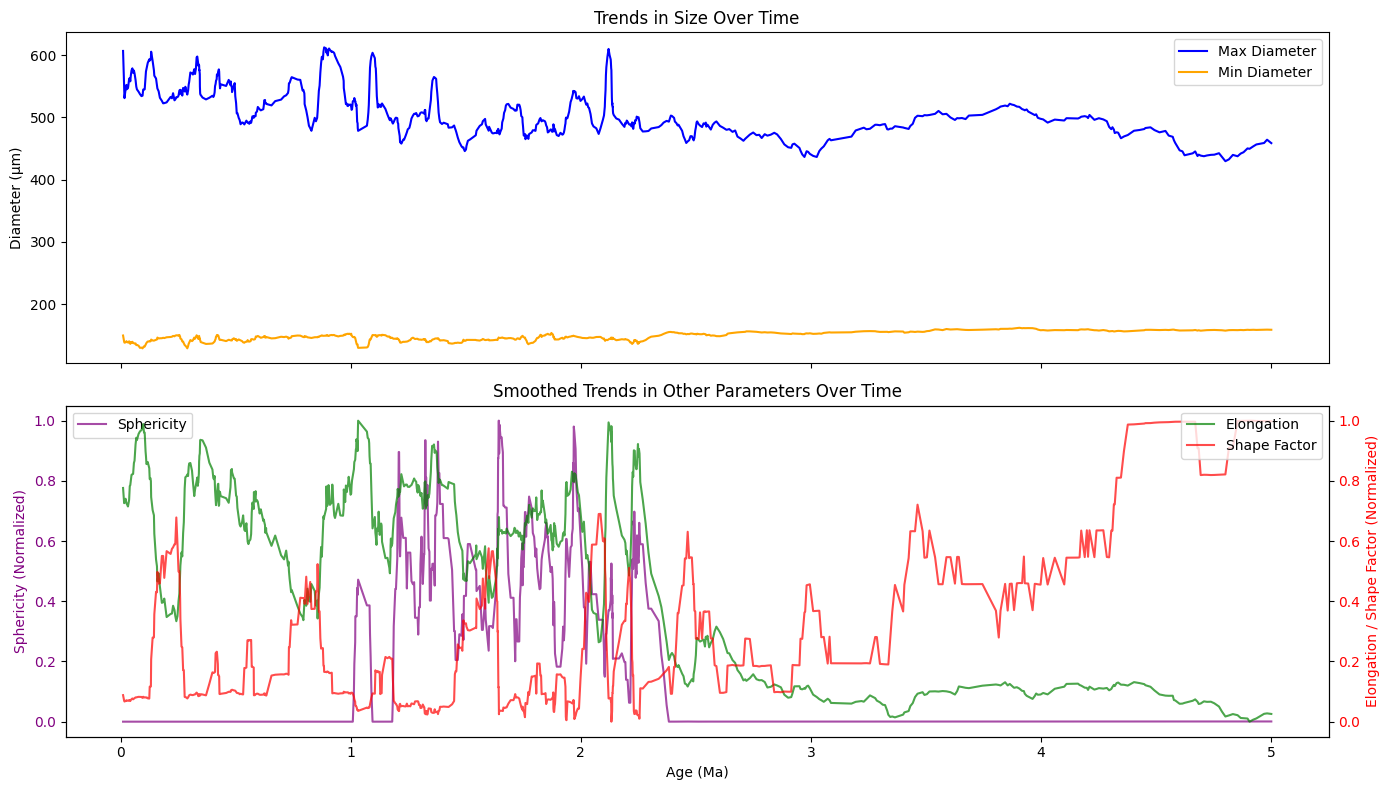
**6. Linear Regression to Calculate Slopes**

* **calculate\_slope function**: This function calculates the **slope** of the line (trend) for each parameter using linear regression. The slope tells us whether the parameter is increasing or decreasing over time.
  + A **positive slope** means the parameter is increasing over time.
  + A **negative slope** means the parameter is decreasing over time.

**7. Printing the Slopes**

* The slopes of **Size Parameters** (Max Diameter, Min Diameter) and **Other Parameters** (Elongation, Shape Factor, Sphericity) are printed, which tells us the trend (increase or decrease) of each parameter over time

**Output:**



Slopes of Size Parameters:

{'Max\_Diameter': -18.682392460939635, 'Min\_Diameter': 4.039869097392751}

Slopes of Other Parameters:

{'Elongation': -0.17716711864107568, 'Shape Factor': 0.12672516968406478, 'Sphericity': -0.01609504723843795}

**Inference:**

    1. Trends in Size Over Time (Top Plot)

        - Max Diameter (Blue Line)

            There are fluctuations in the maximum diameter of the specimens over time.

            Despite short-term variability, the overall trend appears relatively stable, with no

            clear long-term increase or decrease.

            There are a few spikes, possibly indicating short-lived events affecting foraminifera size.

        - Min Diameter (Orange Line)

            The minimum diameter remains relatively stable over time.

            Unlike the max diameter, there are fewer fluctuations, suggesting that while the largest

            specimens show variability, the smallest ones remain more consistent.

        Possible Interpretation:

            Environmental factors or evolutionary pressures might have influenced larger specimens more

            than smaller ones.

            The stability in the min diameter suggests a lower size threshold that remains relatively

            unaffected by these changes.

    2. Smoothed Trends in Other Parameters (Bottom Plot)

        - Sphericity (Purple Line - Left Axis)

            Sphericity shows irregular fluctuations, with no clear long-term trend.

            Some intervals show significant drops, indicating periods where specimens became more elongated

            rather than spherical.

        - Elongation (Green Line - Right Axis)

            Initially high, elongation decreases after ~2 Ma, showing a more stable trend afterward.

            This suggests that older specimens were more elongated, but modern ones might have evolved toward a

            more compact shape.

        - Shape Factor (Red Line - Right Axis)

            The shape factor is highly erratic, especially in the early period.

            A significant increase is seen around 4-5 Ma, possibly indicating environmental changes that favored

            more irregular or complex shapes.

        Possible Interpretation:

            A shift in foraminifera morphology over time, possibly due to climate fluctuations or oceanic conditions.

            The decrease in elongation could mean specimens adapted to a more hydrodynamic or efficient shape.

            The sharp increase in shape factor towards 4-5 Ma might indicate a transition period with more morphological diversity.

**### Slopes of Size Parameters:**

    Max Diameter: -18.68

        A negative slope for Max Diameter indicates a decreasing trend over time. The relatively steep

        magnitude of this negative slope suggests that the size of the organisms is shrinking over time.

    Min Diameter: 4.04

        A positive slope for Min Diameter indicates an increasing trend over time. This suggests that the

        minimum size of the organisms is growing over time, although the magnitude is relatively smaller compared to the Max Diameter.

**### Slopes of Other Parameters:**

    Elongation: -0.18

        The negative slope for Elongation suggests that elongation decreases slightly over time, but the magnitude

        of the slope is small, meaning the change is relatively minor.

    Shape Factor: 0.13

        The positive slope for Shape Factor indicates a slight increase over time, suggesting that the shape of

        the organisms is becoming more circular (or at least more uniform) over time.

    Sphericity: -0.02

        A very small negative slope for Sphericity indicates a negligible decrease over time, suggesting that

        the organisms' sphericity is relatively stable with a slight decline.

        Inference:

**### Differential Selection Pressures:**

The size parameters (Max Diameter and Min Diameter) exhibit more noticeable and contrasting trends: a large decrease

in Max Diameter (-18.68) and a smaller increase in Min Diameter (4.04). This suggests a potential directional selection or

environmental factors that are reducing the maximum size, while allowing or favoring slight growth in the minimum size.

Other parameters like Elongation, Shape Factor, and Sphericity show much smaller changes, particularly with Elongation

and Sphericity having small negative slopes. The Shape Factor increases slightly, indicating a shift toward a more circular

or uniform shape.

**### Different Rate of Change:**

The size parameters show a stronger and more pronounced change over time (especially Max Diameter), indicating that

size may be under stronger selective pressures. These pressures might be related to environmental or biological factors

favoring smaller maximum sizes or larger minimum sizes.

The changes in other parameters are relatively subtle, suggesting that shape, elongation, and sphericity might be

experiencing weaker or less direct selection pressure compared to size. The small slopes in these parameters imply

they are either stable or under milder selective pressures.

**### Conclusion:**

Different selection pressures seem to be acting on size and other shape-related parameters. The size parameters,

especially the Max Diameter, exhibit more significant and opposing trends, implying stronger directional selection

on size. In contrast, the other shape-related parameters change more slowly, suggesting that shape and sphericity

might not be under as intense selection pressure, or the selection pressures are more subtle.

Thus, the patterns of change over time for size are indeed different from those for other parameters, supporting the

idea that different selection pressures are acting on size versus shape and sphericity.