**Author:** Divya Reddy Konda  
**Date:** 03/07/2025  
**Programming Language & Version:** Python 3.12

**Dependencies**

* numpy
* Pandas
* scipy

**Input/Assignment Files**

* DecayTimecourse (1).txt  
  This file contains the 60-minute time-series expression data for yeast genes. The data includes three replicates, with each replicate recorded at 9 time points.
* Assignment-2 (1).pdf

**Script Description**

This project performs two key tasks:

1. **Half-life Calculation:**  
   The script reads the expression data, converts non-numeric entries to numbers, and reshapes the data into three replicates (each with 9 time points). For each replicate, it calculates the transcript half-life by performing linear regression on the natural log of the expression values against time. The final half-life for each gene is obtained by averaging the half-life from the three replicates.
2. **Extreme Gene Identification:**  
   After computing the half-lives, the script identifies the genes with very high and very low half-lives. It does this by determining the 90th and 10th quantiles, thereby selecting the top 10% (longest half-lives) and bottom 10% (shortest half-lives) of the transcripts.

**Output Files**

* **calculated\_half\_lives.csv**  
  Contains each gene's averaged half-life.
* **top\_10\_percent\_half\_lives.csv**  
  Lists the genes with the highest 10% half-lives.
* **bottom\_10\_percent\_half\_lives.csv**  
  Lists the genes with the lowest 10% half-lives.