**Exploratory Data Analysis (EDA)**

Several key factors affect the pipe's sound:

1. **Air pressure (wind)**: Determines the power of the tone (Pa).
2. **Frequency (pitch)**: Sets the note being played (Hz).
3. **Flow rate**: Depends on the size of the toe-hole (.
4. **Flue size**: Shapes the airflow and tone quality (m).
5. **Cut-up height**: Impacts the tone’s brightness and balance (m).
6. **Air density**: Influences sound production (.

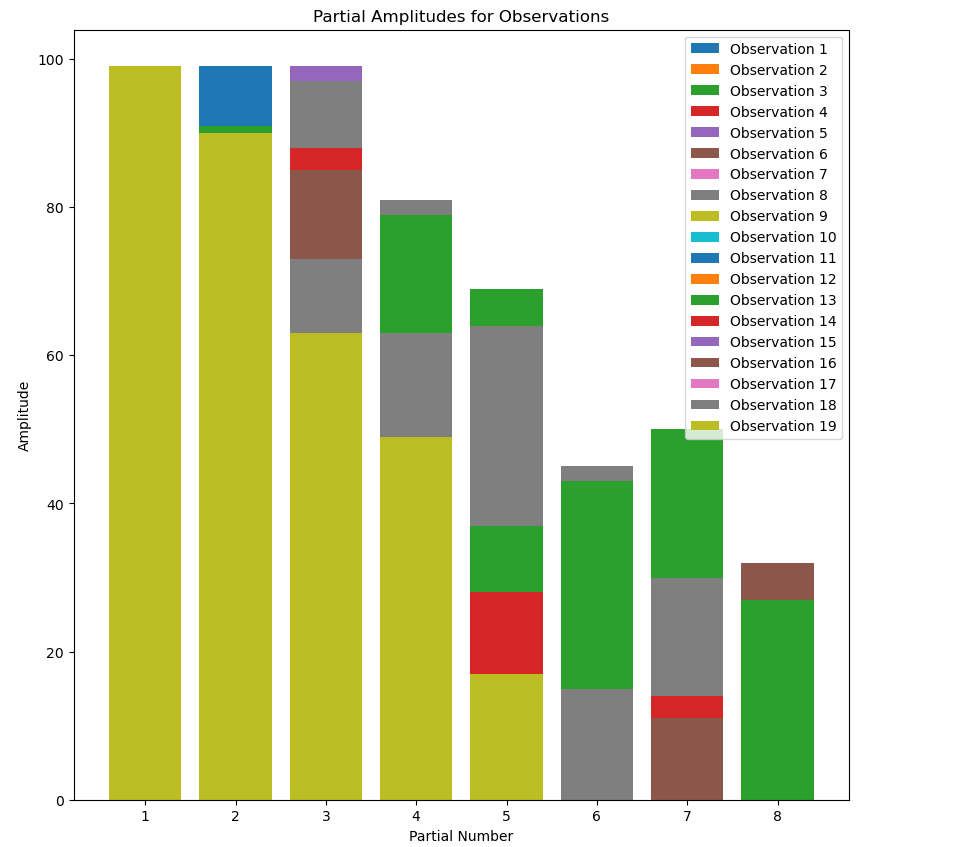
The Ising formula provides a way to calculate the ideal tone, but it doesn’t include a key element, which is flow rate (toe-hole size), which is crucial for controlling tonal results. Measurements show that pipes may overblow—producing higher pitches like octaves—earlier than predicted by the formula. This project aims to enhance the Ising formula by incorporating flow rate, thereby creating a more reliable method for producing and prevoicing organ pipes in industrial settings.

In the initial stage of data wrangling, we exclude constants and parameters that can be readily derived through computational methods. Instead, we prioritize key variables that require direct physical measurement, including flue depth, frequency, cut-up height, acoustic intensity, and toe diameter. Subsequently, we formulate equations to compute flow rate, Ising numbers, and modified Ising numbers incorporating flow rate.

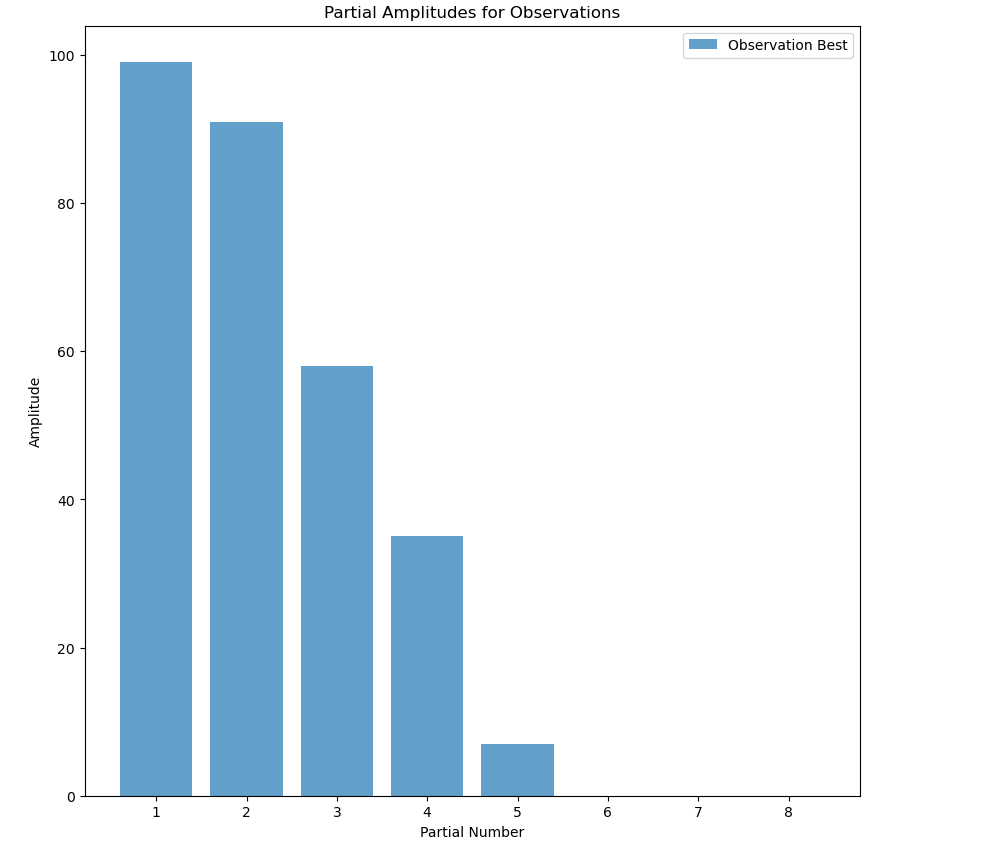
**Let use Matplotlib**, (Hunter, J. D. (2007). *Matplotlib: A 2D graphics environment*. Computing in Science & Engineering, *9*(3), 90-95. https://doi.org/10.1109/MCSE.2007.55

), a popular library for data visualization, to create bar charts for the amplitudes of each partial in the first 19 observations. This is to give us a general idea of what the amplitude generally looks like in most of the observations.

We extract data-- select specific columns (partial1 to partial8) from a Data Frame. Then, converting the selected portion into a NumPy array, set up the size for the bar plot, and create a loop that iterates i from 0 to 18, while it plots 8 partials for each observation.

****

An ideal pipe should produce amplitudes close to the following graph:



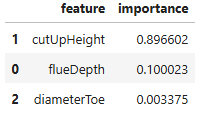
The second method for general analysis is the Gradient Boosting Regressor. (Scikit-learn Developers. (n.d.). GradientBoostingRegressor. Retrieved April 29, 2025, from https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingRegressor.html)

Let X (Independent Variables): flueDepth, cutUpHeight, and diameterToe.

Y (Dependent Variable) → The frequency produced by the pipe.

This method is used for predicting continuous values (like frequency). It learns patterns in the data by combining multiple decision trees in a sequential way (boosting). Once trained, the model evaluates which features (geometric parameters) contribute the most to predicting frequency.

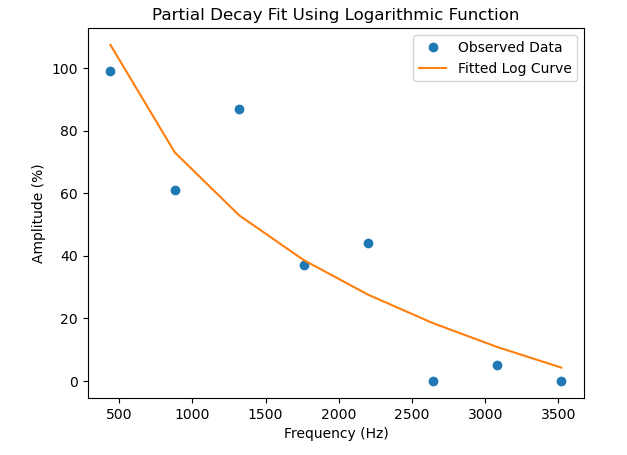
Based on the result below, we can detect that cut-up height is an important factor for frequency. The diameter toe, in the collected data, does not have any effect on the frequency. It is an early indication that the flow factor does not have a significant effect on tone.



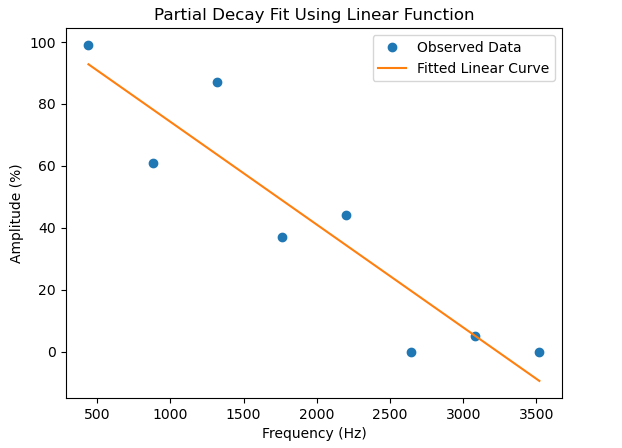
Using a linear regression strategy, I can fit the data following a log function and a linear function

**a \* log(x) + b**

Suppose fundamental frequency f0 = 440hz (A4 = 440 Hz)



**a \* x + b**



With the constant value of P = 0.77 Pa and RHO (air density) = 1.185 . We calculated the predicted Ising, toe area, and flow rate as follows:

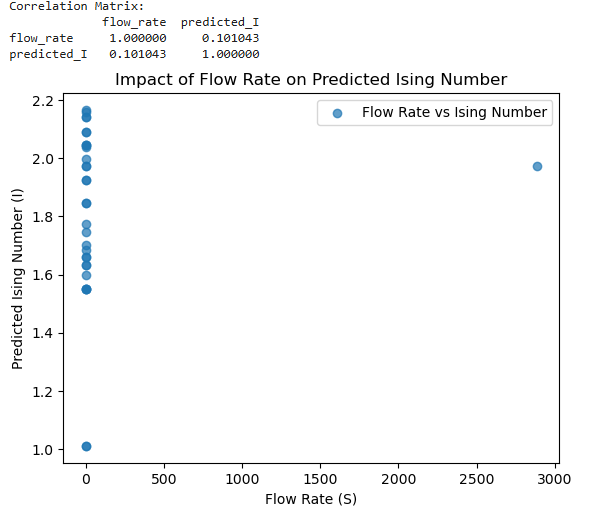
**Toe area =**

**Flow rate =**

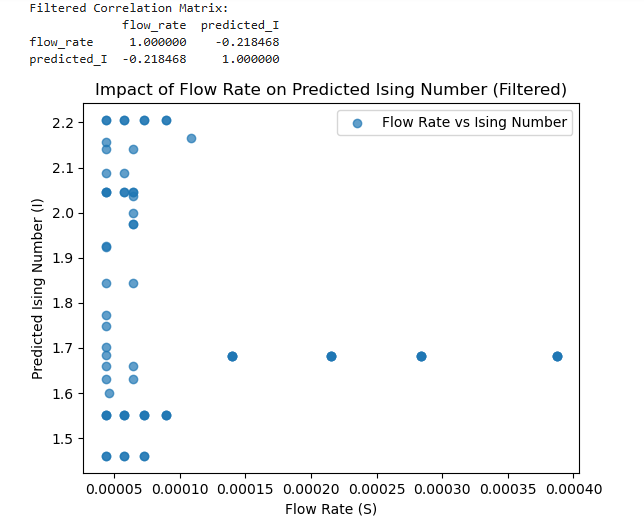
**Ising =** √ **(2P↓\*D)/ (ρ\*H3))/Q**

**Correlation Analysis**

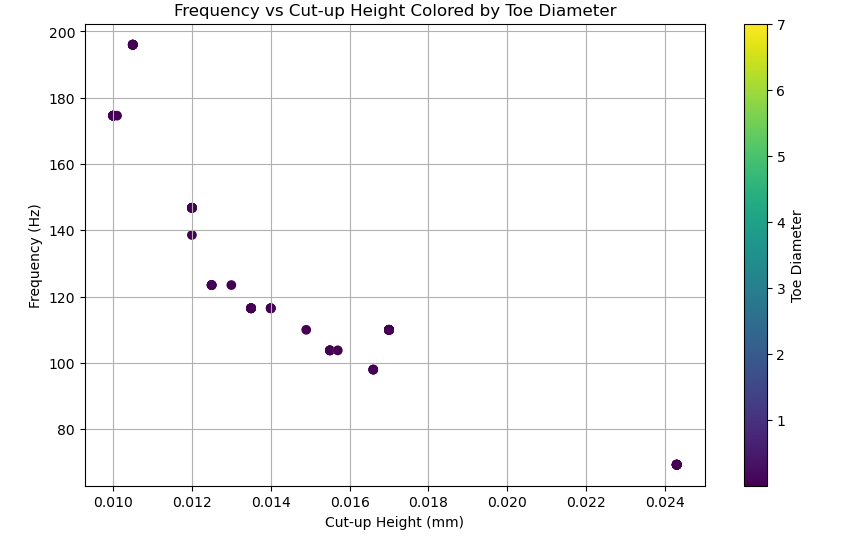
**Correlation matrix** ([Matplotlib 3.10.3 documentation](https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.xcorr.html)) can help us understand the **relationship between two or more variables** in the dataset. In this case, analyzing how **flow rate** affects **predicted Ising values**, I may see if airflow plays a major role in tonal stability. Then, we use a scatterplot to visualize the Correlation Matrix.



We detected an outlier in our scatterplot, which could be the result that makes our correlation faulty. However, using Z-score Filtering ([zscore — SciPy v1.15.3 Manual](https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.zscore.html)) to eliminate outliers, we can produce a Correlation Matrix as follows.



The Correlation Matrix and scatterplot show a weak negative correlation between flow rate and predicted I, with a correlation coefficient of -0.218. This suggests that as the flow rate increases, the predicted Ising number tends to decrease slightly, but the relationship isn’t very strong.



Here, we are using another graphing technique ([Matplotlib 3.10.3 documentation](https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.scatter.html)) showing the relationship between frequency and cut-up height. As the cut-up height increases, the frequency decreases. And the toe diameter stays as the smallest value below 1.

**Assumption**

The **Ising values** are a measure of tonal stability in an organ pipe. It considers the balance between **airflow dynamics, geometry, and acoustic behavior**. Essentially, a higher Ising value suggests a more stable sound, while a lower value indicates more fluctuations. But we should be careful for an overblow situation when the Ising number is over 2.5.

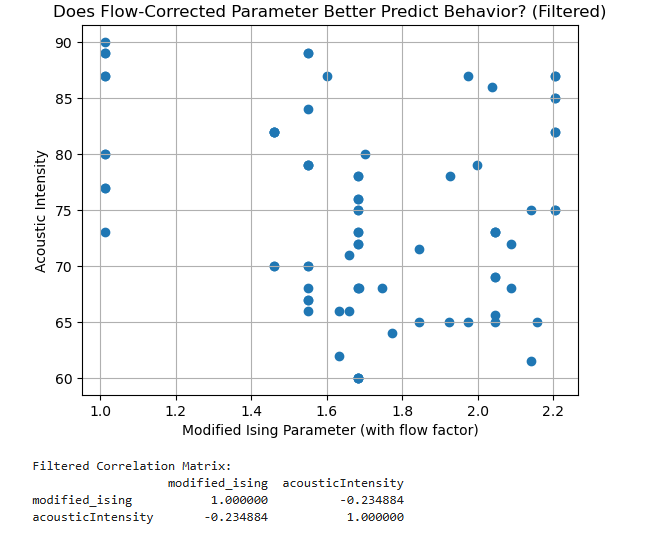
The assumption is that, let's suppose we want to scale the Ising Value. The purpose here is to create a linear positive relationship between modified Ising values and flow rate.

Instead of treating the Ising values as a fixed value, I allow the flow rate to enhance or suppress them.

If the flow rate is high, the term (1 + 0.5 × flow\_factor) increases the Ising parameter. This indicates overblow and bad voicing.

If the flow rate is low, the modification is minimal.

**Modified ising = predicted Ising \* (1 + 0.5\* flow rate)**

****

There is a weak relationship detected. When the acoustic intensity increases, the modified Ising parameter decreases. **The purpose of plotting Acoustic Intensity and Modified Ising Parameters** is to test **whether flow-corrected predictions align better with real-world sound measurements**. If the modified Ising value shows a stronger correlation **with acoustic intensity**, it suggests that **flow rate should be considered in tonal stability models**. **However, in this case, we failed to prove that since we did not have enough data.**