# DORA Manager Take-Home Assignment

Completed by Frederick (Fritz) Zuhl

Dec. 16, 2021

Melodie:

Thank you for giving me this opportunity to solve a problem common among all financial institutions-the problem of the rare, yet costly, charge off.

I used Python 3.9 to complete the exercise, as well as Jupyter Notebooks. I saved my code in three forms:

1. The original Jupyter Notebook
2. The Python code in text.
3. An HTML printout of the Jupyter notebook.

The code is in 2 seperate files for each format (1-3) above:

**Intake** This code file looks over data, doing common EDA and feature preparation.

**Model**  This code file does some data preparation for the modeling, fits the model, ranks features in order of importance, and solves for the 'Active' accounts.

In addition, I included the accounts given to me, along with the probability scores that my model provided. The file is mrm5\_model\_data\_scored.csv.

About the modeling

I choose to use Python because Python has the best algorithm commony available that I thought would be best in this exercise, Extream Gradient Boosting or XGBoost. The Python implementation also comes with the best documentation commonly available. I know there is an implementation of XGBoost in R, which your team uses, but it is not as well documented, and not as current as the one in Python.