

Many questions within the commercial lending space take the form “What is X for a bank like us?” For example:

- “What is the average loan-to-value ratio on a typical owner-occupied real estate loan, for banks like us?”
- “What is the average rate on a typical 5-year term, owner-occupied real estate loan, for banks like us?”

To answer these questions, we need to be able to identify a group of banks that are “like” a given bank. We use a combination of our own data and publicly available data (such FDIC data found here: <https://cdr.ffiec.gov/public/PWS/DownloadBulkData.aspx>). We would like to see your solution to this problem. We would like you to solve the “banks like this one” problem.

Use your imagination, your training, the Internet, and any other sources you’d like, to answer the following questions.

- 1: Where would you start? What information do you wish you had? In the absence of perfect knowledge and access to perfect data, what can you do to make a first attempt?
- 2: Now do it. Build a model that classifies banks. Explain the steps you took and the decisions you made along the way, as if you were leaving notes for the next person who might tackle this problem.
- 3: Now that you’re done, suppose a co-worker is eager to use your results & ideas in our business, starting immediately. What would you advise and why?
- 4: When should your solution NOT be used?
- 5: If you had more time and resources, what would you do next, to improve or refine your work?
- 6: Tell us what you think of this homework assignment. What would you do differently, if you were designing it?

Our goal is to gain insight into your problem-solving approach as well as your Data Science skills so please include both a narrative and a technical / code response. For delivery formats we suggest MS Word, PowerPoint, or PDF for the narrative. For code samples Jupyter notebooks are preferred, though text files containing well commented code are also welcome.