

The following is an edited version of the questions and answers given during the second Q&A session for the FDIC Academic Challenge on November 12, 2020. Thanks to all participants in the conference call for their engagement and interest.

Q: Is there a recommended format or length for the required reference list?

A: Under the Academic Challenge rules, the list of references does not count against the page limit for the written submission. Please be sure to cite all relevant sources that are referenced in or that helped inform the work presented in the written submission.

Any academic citation style is acceptable for your reference list. A popular choice in the social sciences is the author-date format detailed in *The Chicago Manual of Style*. This format involves a short reference to the relevant source placed in the text of the document, as well as a full identification of the source placed in the reference list in alphabetical order by the last name of the source's first author. For more details, including the formatting for citations for different types of sources, please visit *The Chicago Manual of Style* website at https://www.chicagomanualofstyle.org/tools_citationguide/citation-guide-2.html and <https://www.chicagomanualofstyle.org/book/ed17/part3/ch15/toc.html>.

Q: What are the units of measurement for dollar-denominated variables in the FDIC-provided dataset?

A: There are many variables in the data that are measured in dollars, and these data come from multiple sources. As part of the process of cleaning the data and preparing it for use, every variable measured in dollars has been scaled so that it is reported in units of thousands of dollars. For example, if an observation shows a value of 5000 for the dollar amount of commercial and industrial loans made by a bank in a given year, then this value signifies that the bank has made \$5,000,000 worth of loans in that year, not \$5,000. Teams should make sure that their analysis and discussion of results reference the correct dollar amounts as part of their written submissions.

Since all variables in the FDIC-provided dataset have been standardized in this way, comparing dollar values is straightforward. However, dollar values from any external dataset merged with the FDIC-provided data may express dollar-denominated variables in a different format. Teams should take care when comparing dollar amounts from the FDIC-provided data to variables in dollars obtained from external data.

Q: What is the SIM DESCRIPTION variable in the Summary of Deposits data?

A: Note that this variable is not included in the FDIC-provided data for the Challenge because of a difference in data structure. The SIM DESCRIPTION variable is a branch-level variable provided as part of the Summary of Deposits, while the FDIC-provided dataset for the Challenge aggregates the Summary of Deposits data to the county level within each financial institution.

The SIM DESCRIPTION variable is used for internal purposes to record information on the result of the match between location information provided by banks and the latitude and longitude. Teams that are interested in using the latitude and longitude variables at the branch level reported in the Summary of Deposits data should regard the latitude and longitude provided there as the exact location of the branch for purposes of the Academic Challenge.

Q: When grading submissions, how much weight is put towards methodology and analysis as compared to the results found? If a team focuses on a state where community banks play a different role in the local economy than they do in another state, does the difference between the states affect the grading of the submission?

A: The Grading Rubric for both the written and final rounds of the Academic Challenge is available online and explains the categories on which submissions will be graded. The results that a team finds are not considered as a component of the grade. Methodology and analysis are part of the grading rubric in multiple categories. For example, the “Research and Data Analysis” section category refers to both conclusions being supported by data and analysis methods being appropriate to address the question as grading criteria, while the “Written Response” category refers to both arguments being compelling and the response using appropriate visualizations as grading criteria. Having sound methodology and analysis, as well as being able to explain why the methodology and analysis are sound, are both parts of submission assessment. These grading criteria apply regardless of the state that a team chooses to focus on in its submission.

Q: How is the number of firm owners calculated in the BEA data included as part of the FDIC-provided dataset? How do firm employees affect this calculation?

A: The number of owners in the BEA data is computed as the number of sole proprietors plus the number of non-corporate business partners. For example, if a firm has one owner and five non-owner employees, then the number of owners contributed by this firm would be one. If the firm is a general partnership among three people and has five non-owner employees, then the number of owners contributed by this firm would be three.

Employees of firms are counted separately in the “jobs” variable. The jobs variable is the total employment in a county, which is determined as the number of wage and salary workers plus the number of sole proprietors plus the number of non-corporate business partners. For the two example firms described above, the contribution to the number of jobs would be six and eight, respectively.

Q: Loans in the FDIC-provided data are at the bank level, not the county level. Can teams make and justify reasonable assumptions to get approximations of loans at the county level?

A: Yes. There are a few similar assumptions that are made in the banking research literature. Teams should make whatever assumptions that they feel can be defended and supported.

There are several public datasets that contain data on the geography of lending. The FFIEC makes available both HMDA data on mortgage lending and CRA data on business lending at various levels of geography. These data are available by bank, so it would be possible to get a measure of where some of types of bank lending take place through these data.

Alternatively, the geographic distribution of deposit data can be used as a proxy for the geographic distribution of lending data. The Summary of Deposits data show how a bank’s deposits are distributed throughout the U.S. for all banks. A reasonable assumption might be that a bank’s loans are distributed in the same fashion. Therefore, if 1% of a bank’s deposits in a given year come from a given county, then assuming that 1% of the bank’s loans in the same year were made in the same county is reasonable.

Teams that make assumptions regarding the geographic distribution of bank loans should state such assumptions clearly and offer a justification and defense of the assumptions made.

Q: Why are some cells blank in the FDIC-provided data?

A: When opening the CSV data file in Excel, some cells are blank. Blank cells in the Excel should be interpreted as missing data, rather than interpreting this cell as having a value of zero. Not all variables have data in all years. For example, while the dataset begins in 1994, some lending variables from the Call Report data are not available until 1998, including data on small loans to businesses and farms. Teams that wish to use these data in their analysis should take in account that the data are not available in all years and tailor their analysis appropriately. The other provided data formats will mark empty cells differently than Excel does, but the program will still recognize that the data in these cells are missing.

Q: What data are available on financial regulations? Is there information on topics related to Dodd-Frank available to the public?

A: As a general source of information on regulation, the Federal Register (<https://www.federalregister.gov/>) is the official daily journal of the federal government. The sections of the Federal Register most likely to contain relevant information are “Rules and Regulations” and “Proposed Rules,” both of which are updated daily. Rules and Regulations includes policy statements and interpretations of rules by federal agencies, while Proposed Rules includes possible future regulations that are currently open for feedback from the public.

Published final rules from federal agencies are collected and published annually in a separate volume called the Code of Federal Regulations (<https://www.ecfr.gov/>). If only final regulations are of interest, then the Code of Federal Regulations may be easier to use because regulations are organized by topic and regulatory agency. Using stress testing regulation as an example relevant to Dodd-Frank, Title 12 of the Code of Federal Regulations is “Banks and Banking,” Chapter III contains regulations put forth by the FDIC, and subchapter B, parts 325.1 through 325.7 concern Stress Testing. Looking specifically at 325.5, Methodologies and Practices, the last line at the end of the regulation text shows the dates on which the regulation was first published in the Federal Register (Oct. 15, 2012) and the dates on which the regulation was amended (Apr. 24, 2018; Feb. 14, 2019; Oct. 24, 2019).

State regulations are more difficult to find because different states have different rules about what information is shared with the public and how that information is shared. Some states have websites similar to the Code of Federal Regulations that contain information on banking regulations. For example, the Massachusetts Division of Banks offers the text of individual regulations (<https://www.mass.gov/archive/division-of-banks-regulations>), as well as dates of repeal of older regulations. In general, teams will need to search for this information on a state-by-state basis.

Q: Is the UNINUMBR variable in the Summary of Deposits data a unique identifier for a bank branch? How does it differ from the BRNUM branch identifier?

A: Note that neither variable is included in the FDIC-provided data for the Challenge because of a difference in data structure. The UNINUMBR and BRNUM variables are both branch-level variables provided as part of the Summary of Deposits, while the FDIC-provided dataset for the Challenge aggregates the Summary of Deposits data to the county level within each financial institution.

The UNINUMBR variable is associated with a specific physical location of a bank branch regardless of ownership. It only changes over time in rare cases, including some physical

branch moves and some merger scenarios. The BRNUM variable identifies branches within a bank and may change for bank branches that change ownership in sale or merger scenarios.

Q: How should teams compare bank characteristics that vary by county with those that are the same in all counties?

A: The FDIC-provided data for the Challenge have one observation for each bank-county-year. A single observation contains information about a particular bank, in a particular county, in a particular year.

However, different data sources used to assemble the FDIC-provided dataset report information at different levels, and these differences may make some comparisons challenging. Summary of Deposits data vary at the county-bank-year level, so a given bank in a given year will have different amounts of deposits assigned to different counties. However, data on this same bank's loans taken from Call Report data vary only at the bank and year level, so the dollar amount of agricultural loans for the same given bank in the same given year does not differ across counties. Instead, this loan amount represents the dollar amount of agricultural loans made by the bank nationwide in that year. Comparing the bank's deposit amount for a single county against the loan amount that the same bank makes across all counties is not likely to be a useful comparison. In order to obtain the national amount of deposits for a given bank in a given year, first sum up the county-level deposit amounts for the bank in that year.

Q: Which variable in the Call Report data identifies a single bank over time?

A: In the FDIC-provided data as well as the raw Call Report data, the variable CERT uniquely identifies a bank over time. This is the variable that was used to identify banks when assembling the FDIC-provided data.

Q: Are written submissions due before a particular time on the due date?

A: Submissions for the written round of the Academic Challenge are due by midnight Eastern Standard Time on November 20, 2020.