



FOCUS ON GEODATABASES IN ARCGIS PRO

By David W. Allen

Esri Press, 2019

260 pages

Softcover: \$59.99, ISBN 978-1-5894-8445-0

Review by: Vincenza Ferrara

THE CREATION, MODIFICATION, AND STORAGE of spatial data are issues of prime importance in geospatial work, as are concerns about data integrity and data sharing. Esri's geodatabase format is a powerful and sophisticated system for storing geographic information in relational databases—either in the form of an enterprise relational database management system (RDBMS) or an Esri file geodatabase—and accessing that data with Esri GIS software. *Focus on Geodatabases in ArcGIS Pro*, written by David W. Allen and published by Esri Press, is a technical workbook that leads the user through the conceptual and practical aspects of designing, building, and working with geodatabases in Esri's ArcGIS Pro 2.3 environment.

The book is divided into seven chapters plus a one-page "Introduction." In my view, the chapters fall into two groups, with the first four covering the topic of geodatabase creation—design, data population, and facilitation of access—and the remaining three focusing on various techniques available in ArcGIS Pro for working with and editing data. Each chapter has a similar structure, including the following elements:

- a brief theoretical explanation of the topic covered
- two hands-on tutorials with detailed learning objectives and step-by-step instructions
- two more suggested exercises, provided without instructions or tutorial data, but with a list of potential deliverables that could result from completing them

- a concluding summary of the topic covered, methodology applied, and results obtained in each tutorial
- a list of study and review questions
- a list of ArcGIS Pro Help documents suggested for further reading

Data and other materials associated with each tutorial exercise are available for download from the ArcGIS online website, and the downloads include a free 180-day trial for ArcGIS Pro software.

Chapter 1 deals with designing the logical model of the geodatabase, and outlines four important aspects that must be considered: what sort of data is the geodatabase meant to store? what is the hierarchical relationship between the elements? how will the data be edited? how will the data be maintained over time?

In the Chapter 1 tutorials, the geodatabase design process is carried out in spreadsheets, a methodology that keeps readers focused on the careful thought and pre-planning required. The aim, at this stage, is to lead the reader to consider how their design will support an accurate portrayal of the geographic features and their relationships. A well structured logical model for the geodatabase will significantly reduce the number of problems and mistakes that might arise when it is later populated with data. Furthermore, because some geodatabase elements, once created, cannot be changed but only deleted and re-created, modifying or fixing geodatabase design errors can be



© by the author(s). This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

a very time-consuming task. If only for this reason, it is important to have a well thought-out logical model from the beginning.

It is in Chapter 2 that the user implements their design model by creating a geodatabase in the ArcGIS Pro environment. The explanations and tutorials guide the reader through the construction of each of the various hierarchical geodatabase data structure components—feature datasets, feature classes, tables, rasters, mosaics, and relationships—together with the establishment of data integrity rules through the specification of attribute domains and feature subtypes.

The loading of data into a new or existing geodatabase is covered in Chapter 3. The importance of understanding the data is emphasized—the structure of the data, how it will be represented, and how it will be manipulated—and an iterative procedure, centred on the Append geoprocessing tool, that can be adapted to account for and accommodate different feature types and feature subtypes, is proposed. Multiple import passes through the data may be required before every feature is “matched” (93) to its place in the database structure, but care at this stage pays dividends in data integrity.

Chapter 4 closes the first of the two parts that I have divided the book into (building the geodatabase), with methodologies for sharing geodatabase content online and for extending that content to three dimensions. ArcGIS Online (www.arcgis.com), with its range of tools for creating online applications such as story maps, is the author’s recommended tool for data sharing, and the chapter’s first tutorial demonstrates rules for using such applications. The second tutorial explains in detail how 2D data can be transformed into 3D format using elevation attribute values.

The rest of *Focus on Geodatabases in ArcGIS Pro*—Chapters 5, 6, and 7—explores ways to edit and enhance geodatabase data.

Chapter 5 deals with line feature creation and editing, and focuses largely on the establishment of the snapping environment and use of the feature snapping tools. Snapping is fundamental to establishing line connectivity and maintaining feature accuracy.

Polygon feature creation is covered in Chapter 6, which delves into the role of group templates in building a framework of features with a minimum amount of user interaction. The seventh and final chapter is dedicated entirely to the application of feature topologies in simplifying the feature editing process, and for ensuring data completeness and integrity. A geodatabase topology is a set of rules governing the spatial relationships that are permitted between point, line, and polygon features. For example, a topology might specify that every line feature (representing, say, a pipe) can only join another line feature at a point (representing, in this case, a valve), and that a valve point can only exist at the junction of two or more (pipe) lines. GIS topologies, if properly planned and applied, are powerful tools, and, consequently, it is important to enforce such relationships during feature editing.

In ArcGIS Pro one can apply topologies on either the map or the geodatabase level, and this chapter deals with both types. A map topology is an on-the-fly way to set up temporary relationships among layers in an ArcGIS Pro project in order to ensure the validity of the relationships while the features are edited. A geodatabase topology, on the other hand, is a more robust set of rules, permanently stored in the geodatabase, that are available for use in other projects. One great advantage of a geodatabase topology is that it can be updated—new rules can be added or others removed—as requirements, or the situation being represented, evolve. The final tutorial is dedicated to building and validating a geodatabase topology rule-base.

Focus on Geodatabases in ArcGIS Pro is a do-it-yourself workbook dedicated to showing how to build a geodatabase from scratch. It starts at the beginning, with an outline of the conceptual rationale of the needs the geodatabase must serve, and the structure it will require to meet those needs. The first two chapters and their tutorials cover only the concept and logic of sound geodatabase design, and each subsequent chapter builds on the work covered in its predecessor. The workbook’s focus moves progressively from how to structure a geodatabase, first on paper and then in ArcGIS Pro, through filling the empty framework with data, to activating the relational connections between records. This is an excellent training methodology that allows the reader to understand the relationship between a geodatabase and its geodata. However, because so many design decisions have to take into account specific aspects of the particular data involved, and its particular use and maintenance scenarios, a reader might well struggle at

first with reconciling the connection between the book's somewhat abstract concepts and the realities of the task at hand. In the end, of course, that reconciliation can only come through experience, not from a book.

With the structuring and data-loading groundwork laid, the workbook moves on to demonstrating how new data is interactively created, manipulated, and maintained with ArcGIS Pro's editing tools. The range of tools covered is wide and their efficiency in ensuring data integrity is very high, but one of the most important takeaways is the way these tools can be used within the logical framework of a well designed geodatabase.

Great attention and the right amount of emphasis is placed on data integrity issues throughout this book. Good quality control practices are recommended, and the exercises focus particular attention on the specifics of implementation. In particular, the supplemental exercise following Tutorial 2-2, "Adding complex geodatabase components," and Tutorial 7-2, "Working with geodatabase topology," are outstanding. The latter, for example, provides a step-by-step approach to effective quality control: establishing topology rules; building a geodatabase topology; examining topology errors; and correcting the errors.

The book gives some specific advice and tips to instructors—deliverables to assign, study questions to ask, and readings to suggest—but recommendations are also given at a general level. The helpful advice is not restricted to a particular chapter's topic, either. For example, the author often remarks on matters such as: the way schema or structural changes can affect data integrity; when one might create a subtype instead of a separate feature class; how to name group layers so as to reflect their content; the interpretation of error messages and their relative seriousness (can they be ignored?); clear reasons why good metadata is important; and the flexibility gained by combining feature creation tools.

The tutorials are well designed and, having been based on a single case study project, logically well connected, so the reader is guided through the training in an incremental manner. The materials provided for each tutorial are complete and sufficient, and the reader should, in most cases, gain a clear view of both the steps and the goals as they work through each of them. However, in some parts the instructions given are not very clear and do not match with what seems to result from working through the tutorial. In some instances it can be difficult to sort out which of several similarly named geodatabase elements is the one that should be used, or why or how a particular geoprocessing tool should be employed. Sometimes, explanations seem to come later in the book than where they might have been most useful. Occasionally, too, the instructions just seem obscure.

The text layout of the tutorial pages could also use some improvement. Some instructions are in bold text, but it is not always obvious why they are emphasized and others are not. Sometimes, too, the instructions seem to be ordered confusingly, with some paragraphs jumping ahead without any apparent logic.

Finally, the titles of Chapters 3 and 4 are slightly misleading. Chapter 3, "Populating and sharing a geodatabase," does not contain any content about sharing, a topic that is instead covered in Chapter 4, "Extending data formats" (which might be better named "Sharing a Geodatabase and Extending Data Formats").

These concerns notwithstanding, David Allen's *Focus on Geodatabases in ArcGIS Pro* is an excellent aid to understanding how to design a geodatabase in ArcGIS Pro and how to master the editing tools in it, and is a reasonably good investment for both GIS professionals and educators. With some improvements in layout, associated materials, and exercises, however, it could be even better.

