

Variations of the data life cycle

You have learned that there are six stages to the data life cycle. Here's a recap:

1. **Plan:** Decide what kind of data is needed, how it will be managed, and who will be responsible for it.
2. **Capture:** Collect or bring in data from a variety of different sources.
3. **Manage:** Care for and maintain the data. This includes determining how and where it is stored and the tools used to do so.
4. **Analyze:** Use the data to solve problems, make decisions, and support business goals.
5. **Archive:** Keep relevant data stored for long-term and future reference.
6. **Destroy:** Remove data from storage and delete any shared copies of the data.

Note: Be careful not to confuse the six stages of the data life cycle (plan, capture, manage, analyze, archive, and destroy) with the six phases of the data analysis process (ask, prepare, process, analyze, share, and act). They are not interchangeable.

The data life cycle provides a generic or common framework for how data is managed. You may recall that variations of the data analysis life cycle were described in [Origins of the data analysis process](#). The same can be done for the data life cycle. The rest of this reading provides a glimpse of how government, finance, and education institutions can view data life cycles a little differently.

U.S. Fish and Wildlife Service

The U.S. Fish and Wildlife Service uses the following data life cycle:

1. Plan
2. Acquire
3. Maintain
4. Access
5. Evaluate
6. Archive

For more information, refer to [U.S. Fish and Wildlife's Data Management Life Cycle](#) page.

The U.S. Geological Survey (USGS)

The USGS uses the data life cycle below:

1. Plan
2. Acquire
3. Process
4. Analyze
5. Preserve
6. Publish/share

Several cross-cutting or overarching activities are also performed during each stage of their life cycle:

- Describe (metadata and documentation)
- Manage quality

- Backup and secure

For more information, refer to the [USGS Data Lifecycle](#) page.

Financial institutions

Financial institutions may take a slightly different approach to the data life cycle as described in [The Data Life Cycle](#), an article in Strategic Finance magazine:

1. Capture
2. Qualify
3. Transform
4. Utilize
5. Report
6. Archive
7. Purge

Harvard Business School (HBS)

One final data life cycle informed by Harvard University research has eight stages:

1. Generation
2. Collection
3. Processing
4. Storage
5. Management
6. Analysis
7. Visualization
8. Interpretation

For more information, refer to [8 Steps in the Data Life Cycle](#).

Key takeaways

Understanding the importance of the data life cycle will set you up for success as a data analyst. Individual stages in the data life cycle will vary from company to company or by industry or sector. Historical data is important to both the U.S. Fish and Wildlife Service and the USGS, so their data life cycle focuses on archiving and backing up data. Harvard's interests are in research and teaching, so its data life cycle includes visualization and interpretation even though these are more often associated with a data analysis life cycle. The HBS data life cycle also doesn't call out a stage for purging or destroying data. In contrast, the data life cycle for finance clearly identifies archive and purge stages. To sum it up, although data life cycles vary, one data management principle is universal: Govern how data is handled so that it is accurate, secure, and available to meet your organization's needs.

Mark as completed

Like

Dislike

1. There are six phases in the data life cycle, which culminate with analyze, archive, and destroy. What is the correct order of the first three phases?

- ☐ Plan, manage, capture
- ☒ Plan, capture, manage
- ☐ Manage, capture, plan
- ☐ Manage, plan, capture

✓ **Correct**

The correct order is plan, capture, manage. The progression of the data life cycle is: plan, capture, manage, analyze, archive, and destroy.

2. Fill in the blank: During the _____ phase, data teams decide what kind of data is needed and who will be responsible for it.

- ☐ capture
- ☒ planning
- ☐ analyze
- ☐ manage

✓ **Correct**

During the planning phase, data teams decide what kind of data is needed and who will be responsible for it.

3. A data team considers how and where data will be stored, what tools are required to safeguard it, and how to maintain it properly. What phase of the data life cycle does this scenario describe?

- ☐ Capture
- ☐ Archive
- ☒ Manage
- ☐ Analyze

✓ **Correct**

Considering how and where data will be stored, what tools are required to safeguard it, and how to maintain it properly occurs during the manage phase.

4. What is the term for the computer system in which organizations store their data?

- ☐ Cache
- ☒ Database
- ☐ Data ecosystem
- ☐ Catalog

✓ **Correct**

A database is the computer system in which organizations store their data.

More on the phases of data analysis and this program

Each step in the data analysis process—ask, prepare, process, analyze, share, and act—plays a crucial role in extracting meaningful insights from data. As you navigate through each phase, from asking the right questions to taking informed actions, you harness the true power of data. In this reading, you'll explore how the data analysis process guides this program.

The ask phase

At the start of any successful data analysis, the data analyst:

- Takes the time to fully understand stakeholder expectations
- Defines the problem to be solved
- Decides which questions to answer in order to solve the problem

Qualifying stakeholder expectations means determining who the stakeholders are, what they want, when they want it, why they want it, and how best to communicate with them. Defining the problem means looking at the current state and identifying the ways in which it's different from the ideal state. With expectations qualified and the problem defined, you can derive questions that will help achieve these goals.

In an upcoming course, you'll learn how to ask effective questions and define the problem by working with stakeholders. You'll also cover strategies that can help you share what you discover in a way that keeps people interested.

The prepare phase

In the prepare phase, the emphasis is on identifying and locating data you can use to answer your questions. In an upcoming course, you'll learn more about the different types of data and how to identify which kinds of data are most useful for solving a particular problem. You'll also discover why it's so important that data and results are objective and unbiased. In other words, any decisions made from an analysis should always be based on facts and be fair and impartial.

The process phase

In this phase, the aim is to refine the data. Data analysts find and eliminate any errors and inaccuracies that can get in the way of results. This usually means:

- Cleaning data
- Transforming data into a more useful format
- Combining two or more datasets to make information more complete
- Removing outliers (data points that could skew the information)

After data analysts process data, they check the data they prepared to make sure it's complete and correct. This phase is all about getting the details right. Accordingly, the data analyst will refine strategies for verifying and sharing their data cleaning with stakeholders. In an upcoming course, you'll use spreadsheets and structured query language, or SQL, to clean data.

The analyze phase

With a solid foundation of well-defined questions and clean data, you'll delve into the analyze phase. This is when you turn the data you've gathered, prepared, and processed into actionable information. Data analysts use many powerful tools in their work. In one upcoming course you'll continue using two of them: spreadsheets and SQL. In another upcoming course you'll explore using the programming language R to work with and analyze data.

The share phase

This phase is exactly what it sounds like: It's time to share what you've learned with your stakeholders! In this part of the program, you'll learn how data analysts interpret results and share them with others to help stakeholders make effective, data-driven decisions. In the share phase, visualization is a data analyst's best friend. So, an upcoming course will highlight why visualization is essential to getting others to understand what your data is telling you. In another upcoming course, you'll learn how to visualize data with R.

The act phase

The data analysis journey culminates in the act phase, when data insights are put to work. For you, this action involves preparing for your job search and having the chance to complete a case study project. It's a great opportunity for you to bring together everything you've worked on throughout this course. Plus, adding a case study to your portfolio helps you stand out from other candidates!

1. Which of the following activities are part of the manage phase of the data life cycle? Select all that apply.

- ☒ Select tools to keep data secure
- ☒ Decide where to store data
- ☐ Use the data solve problems
- ☐ Reflect on how to care for data

2. Which spreadsheet feature uses a set of instructions to perform calculations, such as multiplication or division?

- ☐ Format
- ☐ Function
- ☐ Filter
- ☒ Formula

3. Which of the following statements correctly describe the archive and the destroy phases of the data life cycle? Select all that apply.

- ☒ A key reason for destroying data is to protect sensitive customer information.
- ☒ Shredders may be used to destroy data that is on paper.
- ☒ Archiving means storing data, although it may not be used again.
- ☐ The best way to protect private data at the end of a project is to archive it for future use.

4. What tasks may occur during the ask step of the data analysis process? Select all that apply.

- ☒ Use the five whys
- ☒ Determine whether something needs to be fixed
- ☐ Properly maintain a database
- ☐ Communicate with stakeholders

5. Fill in the blank: Data is _____ during the prepare step of the data analysis process.

- ☐ analyzed and interpreted
- ☐ filtered and sorted
- ☐ presented and shared
- ☒ collected and stored

6. Company leaders at a business in Egypt want to ensure all supplies are sourced within the country. So, they ask the data team to investigate the performance of local suppliers. Team members consider how to manage the data for this project and who will be responsible for it. What phase of the data life cycle does this scenario describe?

- ☒ Plan
- ☐ Manage
- ☐ Destroy
- ☐ Analyze

7. A car rental company is interested in improving the customer experience. A data professional fixes typos and inaccuracies from a dataset containing feedback and ratings. They also verify and share their data-cleaning procedures with stakeholders. What step of the data analysis process does this scenario describe?

1 point

- ☒ Prepare
- ☐ Analyze
- ☐ Ask
- ☐ Process

8. Fill in the blank: During the _____ step of the data analysis process, data team members use tools such as SQL to make predictions.

1 point

- ☐ act
- ☒ analyze
- ☐ process
- ☐ share

9. What are some key benefits of data visualizations? Select all that apply.

1 point

- ☒ Complex data can be shared more effectively with nontechnical audiences in visualizations.
- ☐ Data visualizations do not require labels or descriptions.
- ☒ Data visualizations help people gain insights from data more quickly.
- ☒ Maps help people identify patterns in data.

10. Fill in the blank: A university system gathers historical academic data from an outside database and information from the internal records of student applications. This work occurs during the _____ phase of the data life cycle.

1 point

- ☐ analyze
- ☒ capture
- ☐ manage

