**Variations of the data life cycle**

You learned that there are six stages to the data life cycle. Here is 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.

**Warning:** Be careful not to mix up or confuse the six stages of the data life cycle (Plan, Capture, Manage, Analyze, Archive, and Destroy) with the six phases of the data analysis life cycle (Ask, Prepare, Process, Analyze, Share, and Act). They shouldn't be used or referred to interchangeably.

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](https://www.coursera.org/learn/foundations-data/supplement/WWlrt/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](https://www.fws.gov/data/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](https://www.usgs.gov/products/data-and-tools/data-management/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](https://sfmagazine.com/post-entry/july-2018-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](https://online.hbs.edu/blog/post/data-life-cycle).

**Key takeaway**

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.

**Variasi siklus hidup data**

Anda telah mempelajari bahwa ada enam tahapan dalam siklus hidup data. Berikut rekapnya:

1. **Rencana:** Putuskan jenis data apa yang dibutuhkan, bagaimana pengelolaannya, dan siapa yang akan bertanggung jawab.
2. **Capture:** Mengumpulkan atau membawa data dari berbagai sumber berbeda.
3. **Kelola:** Merawat dan memelihara data. Hal ini termasuk menentukan bagaimana dan di mana penyimpanannya serta alat yang digunakan untuk menyimpannya.
4. **Analisis:** Gunakan data untuk memecahkan masalah, membuat keputusan, dan mendukung tujuan bisnis.
5. **Arsip:** Menyimpan data yang relevan untuk referensi jangka panjang dan masa depan.
6. **Hancurkan:** Hapus data dari penyimpanan dan hapus semua salinan data yang dibagikan.

**Peringatan:** Berhati-hatilah untuk tidak mencampuradukkan atau membingungkan enam tahapan siklus hidup data (Rencanakan, Tangkap, Kelola, Analisis, Arsip, dan Hancurkan) dengan enam fase siklus hidup analisis data (Tanya, Siapkan, Proses, Analisis , Bagikan, dan Bertindak). Kata-kata tersebut tidak boleh digunakan atau dirujuk secara bergantian.

Siklus hidup data memberikan kerangka umum atau umum tentang bagaimana data dikelola. Anda mungkin ingat bahwa variasi siklus hidup analisis data telah dijelaskan di[Asal usul proses analisis data](https://www.coursera.org/learn/foundations-data/supplement/WWlrt/origins-of-the-data-analysis-process). Hal yang sama dapat dilakukan untuk siklus hidup data. Bacaan selanjutnya memberikan gambaran sekilas tentang bagaimana pemerintah, lembaga keuangan, dan pendidikan dapat memandang siklus hidup data dengan sedikit berbeda.

**Layanan Ikan dan Margasatwa AS**

US Fish and Wildlife Service menggunakan siklus hidup data berikut:

1. Rencana
2. Mendapatkan
3. Menjaga
4. Mengakses
5. Evaluasi
6. Arsip

Untuk informasi lebih lanjut, lihat[Siklus Hidup Pengelolaan Data Ikan dan Satwa Liar AS](https://www.fws.gov/data/life-cycle)halaman.

**Survei Geologi AS (USGS)**

USGS menggunakan siklus hidup data di bawah ini:

1. Rencana
2. Mendapatkan
3. Proses
4. Menganalisa
5. Melestarikan
6. Publikasikan/Bagikan

Beberapa aktivitas lintas sektoral atau menyeluruh juga dilakukan pada setiap tahap siklus hidup mereka:

* Jelaskan (metadata dan dokumentasi)
* Kelola Kualitas
* Cadangkan dan Amankan

Untuk informasi lebih lanjut, lihat[Siklus Hidup Data USGS](https://www.usgs.gov/products/data-and-tools/data-management/data-lifecycle)halaman.

**Lembaga keuangan**

Lembaga keuangan mungkin mengambil pendekatan yang sedikit berbeda terhadap siklus hidup data seperti yang dijelaskan dalam[Siklus Hidup Data](https://sfmagazine.com/post-entry/july-2018-the-data-life-cycle/), sebuah artikel di majalah Strategic Finance:

1. Menangkap
2. Memenuhi syarat
3. Mengubah
4. Memanfaatkan
5. Laporan
6. Arsip
7. Membersihkan

**Sekolah Bisnis Harvard (HBS)**

Satu siklus hidup data akhir yang diinformasikan oleh penelitian Universitas Harvard memiliki delapan tahap:

1. Generasi
2. Koleksi
3. Pengolahan
4. Penyimpanan
5. Pengelolaan
6. Analisis
7. Visualisasi
8. Penafsiran

Untuk informasi lebih lanjut, lihat[8 Langkah dalam Siklus Hidup Data](https://online.hbs.edu/blog/post/data-life-cycle).

**Pengambilan kunci**

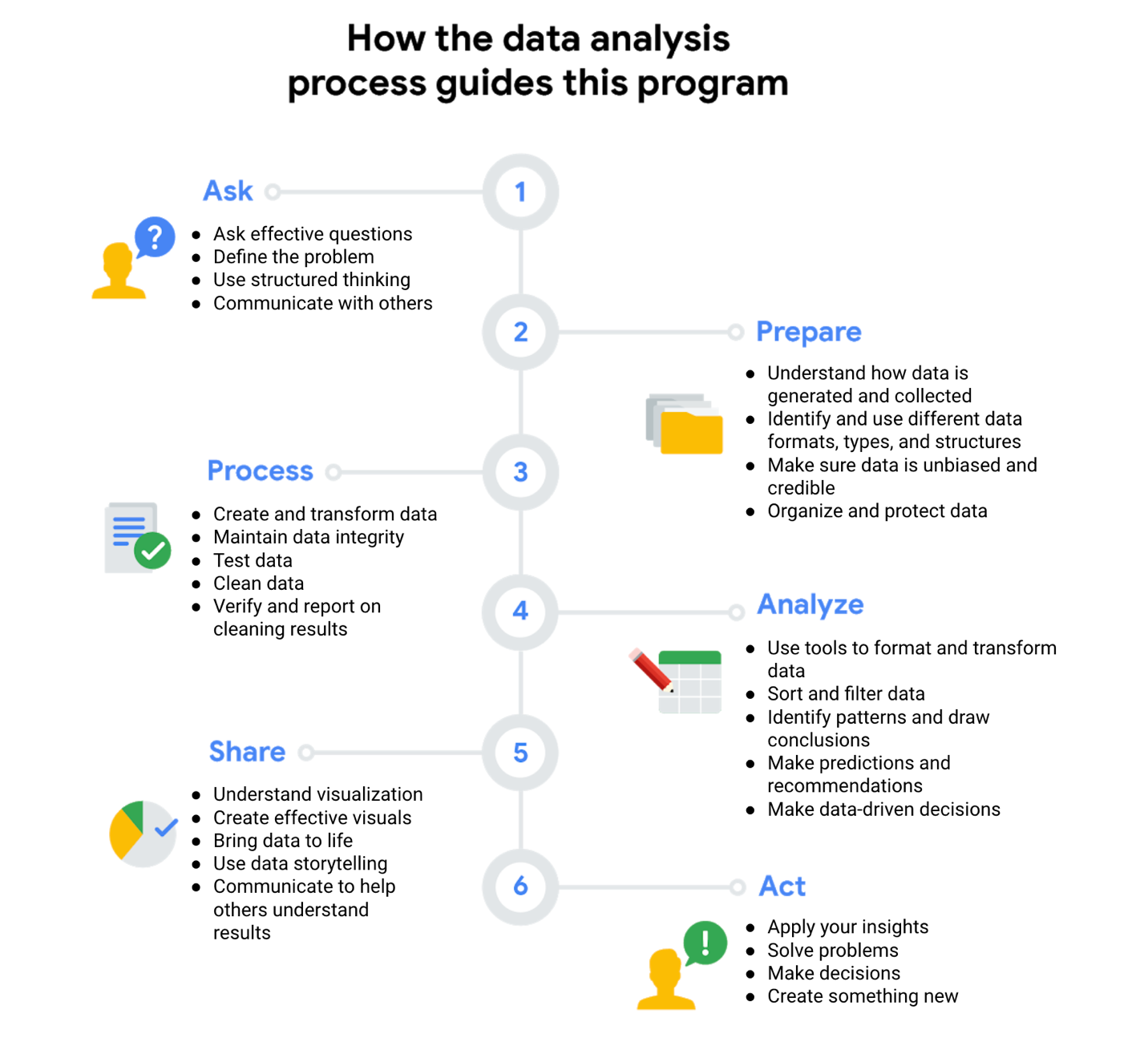
Memahami pentingnya siklus hidup data akan menyiapkan Anda untuk sukses sebagai analis data. Tahapan individu dalam siklus hidup data akan bervariasi dari satu perusahaan ke perusahaan lain atau berdasarkan industri atau sektor. Data historis penting bagi US Fish and Wildlife Service dan USGS, sehingga siklus hidup data mereka berfokus pada pengarsipan dan pencadangan data. Minat Harvard adalah dalam penelitian dan pengajaran, sehingga siklus hidup datanya mencakup visualisasi dan interpretasi meskipun hal ini lebih sering dikaitkan dengan siklus hidup analisis data. Siklus hidup data HBS juga tidak memerlukan tahapan untuk membersihkan atau menghancurkan data. Sebaliknya, siklus hidup data untuk keuangan dengan jelas mengidentifikasi tahapan pengarsipan dan pembersihan. Singkatnya, meskipun siklus hidup data berbeda-beda, ada satu prinsip pengelolaan data yang bersifat universal.

**Proses analisis data dan program ini**

*Anda dapat menyimpan bacaan ini untuk referensi di masa mendatang. Jangan ragu untuk mengunduh versi PDF dari bacaan di bawah ini*

[DAC1 Proses analisis data.pdf](https://d3c33hcgiwev3.cloudfront.net/XT8DBaLYTUC_AwWi2O1Asw_bed8d5f8f56e4f1c9ed1dbdc5917e4f1_DAC1-The-data-analysis-process.pdf?Expires=1693612800&Signature=Jo~vksaWKlxGfLKnW~UbsnxYxNyXkwtJWW6J85iUmvSVoGGF2ajPnXa5RID4TS6XCrtfSJAB81SUu5y3gPMN9slidWyhTOwHtpVzgtvuUU7801kdYGcxPScN25Q4PbhUhqZwuTZWJXx3EasGt~pKyuIpH7zJB57VqZsdKmZxjZs_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[File pdf](https://d3c33hcgiwev3.cloudfront.net/XT8DBaLYTUC_AwWi2O1Asw_bed8d5f8f56e4f1c9ed1dbdc5917e4f1_DAC1-The-data-analysis-process.pdf?Expires=1693612800&Signature=Jo~vksaWKlxGfLKnW~UbsnxYxNyXkwtJWW6J85iUmvSVoGGF2ajPnXa5RID4TS6XCrtfSJAB81SUu5y3gPMN9slidWyhTOwHtpVzgtvuUU7801kdYGcxPScN25Q4PbhUhqZwuTZWJXx3EasGt~pKyuIpH7zJB57VqZsdKmZxjZs_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

1. Ask: Ask effective questions, Define the problem, Use structure thinking, Communicate with others 2. Process: Create and transform data, Maintain data integrity, Test data, Clean data, Verify and report on cleaning results 3. Share: Understand visualization, Create effective visuals, Bring data to life, Use data storytelling, Communicate to help others understand results 4. Prepare: Understand how data is generated and collected, Identify and use different data formats, types, and structures, Make sure data is unbiased and credible, Organize and protect data 5. Analyze: Use tools to format and transform data, Sort and filter data, Identify patterns and draw conclusions, Make predictions and recommendations, Make data-driven decisions 6. Act: Apply your insights, Solve problems, Make decisions, Create something new

**Pelajari prosesnya melalui program:**

1. Pelajari lebih lanjut tentang fase **Tanya** dari proses di[Ajukan Pertanyaan untuk Membuat Keputusan Berdasarkan Data](https://www.coursera.org/learn/ask-questions-make-decisions/home/welcome)kursus.
2. Pelajari lebih lanjut tentang fase **Persiapan** dari proses di[Siapkan Data untuk Eksplorasi](https://www.coursera.org/learn/data-preparation/home/welcome)kursus.
3. Pelajari lebih lanjut tentang fase **Proses** dari proses di[Memproses Data dari Kotor menjadi Bersih](https://www.coursera.org/learn/process-data/home/welcome)kursus.
4. Pelajari lebih lanjut tentang fase **Analisis** proses di[Analisis Data untuk Menjawab Pertanyaan](https://www.coursera.org/learn/analyze-data/home/welcome)Dan[Analisis Data dengan Pemrograman R](https://www.coursera.org/learn/data-analysis-r/home/welcome)kursus.
5. Pelajari lebih lanjut tentang fase proses **Berbagi di**[Berbagi Data Melalui Seni Visualisasi](https://www.coursera.org/learn/visualize-data/home/welcome)Dan[Analisis Data dengan Pemrograman R](https://www.coursera.org/learn/data-analysis-r/home/welcome)kursus.
6. Pelajari lebih lanjut tentang fase **Tindakan** dari proses di[Batu Penjuru Analisis Data Google: Selesaikan Studi Kasus](https://www.coursera.org/learn/google-data-analytics-capstone/home/welcome)kursus.

**Catatan:** Tautan kursus ditujukan untuk Anda melihat pratinjau dan tidak menyelesaikan kursus saat ini. Anda dapat menandai aktivitas ini sebagai selesai setelah Anda memahami bagaimana kursus diselaraskan dengan proses analisis data.

**Alat analis data utama**

Saat Anda mempelajarinya, program dan solusi paling umum yang digunakan oleh analis data mencakup spreadsheet, bahasa kueri, dan alat visualisasi. Dalam bacaan ini, Anda akan mempelajari lebih lanjut tentang masing-masingnya. Anda akan membahas kapan menggunakannya, dan mengapa mereka sangat penting dalam analisis data.



**Spreadsheet**

Analis data mengandalkan spreadsheet untuk mengumpulkan dan mengatur data. Dua aplikasi spreadsheet populer yang mungkin akan sering Anda gunakan dalam peran Anda di masa depan sebagai analis data adalah Microsoft Excel dan Google Sheets.

Spreadsheet menyusun data dengan cara yang bermakna dengan membiarkan Anda

* Mengumpulkan, menyimpan, mengatur, dan mengurutkan informasi
* Identifikasi pola dan satukan data dengan cara yang sesuai untuk setiap proyek data tertentu
* Buat visualisasi data yang luar biasa, seperti grafik dan bagan.

**Basis data dan bahasa kueri**

Basis data adalah kumpulan data terstruktur yang disimpan dalam sistem komputer. Beberapa program Structured Query Language (SQL) yang populer mencakup MySQL, Microsoft SQL Server, dan BigQuery.

Bahasa kueri

* Izinkan analis untuk mengisolasi informasi spesifik dari database
* Memudahkan Anda mempelajari dan memahami permintaan yang dibuat ke database
* Izinkan analis memilih, membuat, menambah, atau mengunduh data dari database untuk dianalisis

**Alat visualisasi**

Analis data menggunakan sejumlah alat visualisasi, seperti grafik, peta, tabel, bagan, dan banyak lagi. Dua alat visualisasi yang populer adalah Tableau dan Looker.

Alat-alat ini

* Ubah bilangan kompleks menjadi cerita yang dapat dipahami orang
* Membantu pemangku kepentingan menghasilkan kesimpulan yang mengarah pada keputusan yang tepat dan strategi bisnis yang efektif
* Memiliki banyak fitur

- Fitur drag-and-drop Tableau yang sederhana memungkinkan pengguna membuat grafik interaktif di dashboard **dan**

lembar kerja

- **Looker** berkomunikasi langsung dengan database, memungkinkan Anda menghubungkan data langsung ke visual

alat yang Anda pilih

Karier sebagai analis data juga melibatkan penggunaan bahasa pemrograman, seperti R dan Python, yang banyak digunakan untuk analisis statistik, visualisasi, dan analisis data lainnya.

**Pengambilan kunci**

Anda memiliki banyak alat sebagai analis data. Ini adalah gambaran sekilas tentang berbagai kemungkinannya, dan Anda akan menjelajahi banyak alat ini secara mendalam sepanjang program ini.

**Memilih alat yang tepat untuk pekerjaan itu**

Sebagai seorang analis data, Anda biasanya harus memutuskan program atau solusi mana yang tepat untuk proyek tertentu yang sedang Anda kerjakan. Dalam bacaan ini, Anda akan mempelajari lebih lanjut tentang cara memilih alat yang Anda perlukan dan kapan.

Bergantung pada fase proses analisis data yang Anda jalani, Anda perlu menggunakan alat yang berbeda. Misalnya, jika Anda berfokus pada pembuatan visualisasi yang kompleks dan menarik perhatian, maka alat visualisasi yang telah kita bahas sebelumnya adalah pilihan terbaik. Namun jika Anda berfokus pada pengorganisasian, pembersihan, dan analisis data, Anda mungkin akan memilih antara spreadsheet dan database menggunakan kueri. Spreadsheet dan database menawarkan cara untuk menyimpan, mengelola, dan menggunakan data. Konten dasar untuk kedua alat tersebut adalah serangkaian nilai. Namun, ada juga beberapa perbedaan utama:

| **Spreadsheet** | **Basis Data** |
| --- | --- |
| Aplikasi software | Penyimpanan data - diakses menggunakan bahasa kueri (misalnya SQL) |
| Struktur data dalam format baris dan kolom | Struktur data menggunakan aturan dan hubungan |
| Mengatur informasi dalam sel | Atur informasi dalam koleksi yang kompleks |
| Memberikan akses ke sejumlah data terbatas | Memberikan akses ke data dalam jumlah besar |
| Entri data manual | Entri data yang ketat dan konsisten |
| Umumnya satu pengguna pada satu waktu | Banyak pengguna |
| Dikendalikan oleh pengguna | Dikendalikan oleh sistem manajemen basis data |

Anda tidak harus memilih salah satu karena masing-masing memiliki tujuannya masing-masing. Umumnya, analis data bekerja dengan kombinasi keduanya, karena kedua alat tersebut sangat berguna dalam analisis data. Misalnya, Anda bisa menyimpan data dalam database, lalu mengekspornya ke spreadsheet untuk dianalisis. Atau, jika Anda mengumpulkan informasi dalam spreadsheet, dan jumlahnya terlalu banyak untuk platform tertentu, Anda dapat mengimpornya ke database. Dan, nanti dalam kursus ini, Anda akan belajar tentang bahasa pemrograman seperti R yang memberi Anda kontrol lebih besar atas data, analisisnya, dan visualisasi yang Anda buat.

Saat Anda terus mempelajari alat penting ini, Anda akan memperoleh pengetahuan untuk memilih alat yang tepat untuk pekerjaan data apa pun.

We’ve covered a lot of terms—some of which you may have already known, and some of which are new. To make it easy to remember what a word means, we created this glossary of terms and definitions.

To use the glossary for this course item, click the link below and select “Use Template.”



Link to glossary: [Week 3 Glossary](https://docs.google.com/document/d/1OugS302nmLJzV98g6muMvzI44KLf8vTvOmP1DSl0t-w/template/preview)