Analysis Process Phases

Montag, 30. Juni 2025 22:13

Analysis Process Phases
+ ask + prepare + process + analyse + share + act +

End

Datei auswählen Keine ausgewählt

Stages Of The Data Life Cycle

Montag, 30. Juni 2025 22:13

Phases Data Go through

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.

Plan
+ what data needs
+ how managed
+ who responsible

Capture
+ countless of ways
+ data from outside
+ or own of company

Manage
+ taking care of data
+ data cleansing
+ how and where stored
+ tools to keep it safe/secure

Analyse
+ where we shine
+ where we solve problems

Destroy
+ protecting private information
+ physically and digital

Database
+ collection of data stored in a computer system

End

Datei auswählen Keine ausgewählt

Variations Of Data Life Cycle

Montag, 30. Juni 2025 22:13

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

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:

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

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.

Datei auswählen Keine ausgewählt

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.

Detailed Phases Of Data Analysis

Montag, 30. Juni 2025 22:13

Phases Of The Data Analysis Process

Ask Phase
+ define a problem to be solved
L> difference from what it is and should be
L> sth wrong or obstacle to be fixed
+ fully understand stakeholder expectations

Prepare Phase
+ collect and store date for the analysis process
+ discover which data is correct
+ unbiased and objective data
+ my decisions should always be based on facts

Process Phase
+ eliminate all errors and problems in data
+ usually means cleaning data
+ transforming into useful format
+ combining datasets

Analyse Phase
+ using tools to transform/organise information
+ make conclusions and predictions
+ make with spreadsheets
+ make with structured query language
+ or SQL - pronounced sequel

Share Phase
+ inform others
+ so that we make effective data-driven decisions
+ visualisation is our best friend

Act Phase
+ business takes all of our insights
+ and puts it into work to solve problem

Stakeholders
+ people who invested time and resources into a project
+ are interested in the outcome

Outliers
+ data points that lie significantly outside the overall pattern
+ could potentially skew the information.

Datei auswählen Keine ausgewählt

Data Analyst tools (Video)

Montag, 30. Juni 2025 22:13

Most common ones:
+ spreadsheets
+ query languages
+ visualisation tools

Two popular spreadsheets
+ Microsoft excel
+ google sheets

+ usefulness of data = how well is structured

Formula
+ set of instructions
+ perform a specific calculation
+ using data in spreadsheet

Function
+ preset command
+ automatically perform specific process/task
+ using data in a spreadsheet

Query Language
+ computer programming language
+ allows to retrieve and manipulate data
+ from a database

Most popular Query Language
+ SQL = structured query language

Popular Visualisation Tools
+ Tableau (very easy to understand)
+ Looker (very complete visualisations)

End

Datei auswählen Keine ausgewählt

Data Analyst tools (Text)

Montag, 30. Juni 2025 22:13

Spreadsheets

Data analysts rely on spreadsheets to collect and organize data. Two popular spreadsheet applications you will probably use a lot in your future role as a data analyst are Microsoft Excel and Google Sheets.

Spreadsheets structure data in a meaningful way by letting you

- Collect, store, organize, and sort information
- Identify patterns and piece the data together in a way that works for each specific data project
- Create excellent data visualizations, like graphs and charts.

Databases and query languages

A database is a collection of structured data stored in a computer system. Some popular Structured Query Language (SQL) programs include MySQL, Microsoft SQL Server, and BigQuery.

Query languages

- Allow analysts to isolate specific information from a database(s)
- Make it easier for you to learn and understand the requests made to databases
- Allow analysts to select, create, add, or download data from a database for analysis

Visualization tools

Data analysts use a number of visualization tools, like graphs, maps, tables, charts, and more. Two popular visualization tools are Tableau and Looker.

These tools

- Turn complex numbers into a story that people can understand
- Help stakeholders come up with conclusions that lead to informed decisions and effective business strategies
- Have multiple features
  - **Tableau**'s simple drag-and-drop feature lets users create interactive graphs in dashboards and worksheets
  - **Looker** communicates directly with a database, allowing you to connect your data right to the visual tool you choose

A career as a data analyst also involves using programming languages, like R and Python, which are used a lot for statistical analysis, visualization, and other data analysis.

Key takeaway

You have a lot of tools as a data analyst. This is a first glance at the possibilities, and you will explore many of these tools in-depth throughout this program.

How To Choose The Right Tool

+ special visualisation tools = complex and eye-catching
+ besides spreadsheets and databases

Spreadsheets	Databases
Accessed through a software application	Database accessed using a query language
Structured data in a row and column format	Structured data using rules and relationships
Organizes information in cells	Organizes information in complex collections
Provides access to a limited amount of data	Provides access to huge amounts of data
Manual data entry	Strict and consistent data entry
Generally one user at a time	Multiple users
Controlled by the user	Controlled by a database management system

Key Takeaways

+ spreadsheets suitable for organising, cleaning, analysing
L> small to medium datasets
+ databases for storing, managing, analysing
L> large and complex datasets
+ data analysts use combinations of both

End

Datei auswählen Keine ausgewählt

Glossary

Montag, 30. Juni 2025 22:13

Terms and definitions for Course 1, Module 2

Database: A collection of data stored in a computer system

Formula: A set of instructions used to perform a calculation using the data in a spreadsheet

Function: A preset command that automatically performs a specified process or task using the data in a spreadsheet

Query: A request for data or information from a database

Query language: A computer programming language used to communicate with a database

Stakeholders: People who invest time and resources into a project and are interested in its outcome

Structured Query Language: A computer programming language used to communicate with a database

Spreadsheet: A digital worksheet

SQL: (Refer to Structured Query Language)

Datei auswählen Keine ausgewählt