

Origins of the Data Analysis Process

6 Steps of Data Analysis

1. **Ask:** Business Challenge/Objective/Question
2. **Prepare:** Data generation, collection, storage, and data management.
3. **Process:** Data Cleaning/Data Integrity
4. **Analyze:** Data exploration, visualization, and analysis.
5. **Share:** Communicating and interpreting results.
6. **Act:** Putting your insights to work to solve the problem.

Dell/EMC Data Analysis Life Cycle (Model Created by David Dietrich)

EMC Corps data analytics life cycle is cyclical with six steps:

1. Discovery
2. Pre-processing data
3. Model Planning
4. Model Building
5. Communicate results
6. Operationalize

While different from the methodology taught in this course by Google. The core ideas are very similar in that each step leads to the next and require enough thought is put into each step before ultimately returning to the beginning.

- IE. You have to discover and ask questions before you can prepare and analyze the data which is done before you model, share and act on the data.

SAS's Iterative Life Cycle of Data

1. Ask
2. Prepare
3. Explore
4. Model
5. Implement
6. Act
7. Evaluate

The SAS model emphasized the cyclical nature of their model by visualizing it as an infinity symbol. Their life cycle has seven (7) steps, many of which we have seen in the other models, like **Ask, Prepare, Model, and Act**. But this life cycle is also a little different; it includes a step after the **Act** phase, designed to help analysts evaluate their solutions and potentially return to the ask phase again.

Project-based data analytics life cycle

1. Identifying the Problem
2. Designing Data Requirements

3. Pre-Processing Data
4. Preforming Data Analysis
5. Visualizing Data

This data analytics project life cycle was developed by *Vignesh Prajapati*. It doesn't include the sixth phase, or what we have been referring to as the **Act** phase. However, it still covers a lot of the same steps as the life cycles we have already described, It begins with *Identifying the problem, preparing and processing data before analysis, and ends with data visualization*.

Big Data Analytics Life Cycle

- From the Book **Big Data Fundamentals: Concepts, Drivers & Techniques**

1. Business Case Evaluation
2. Data Identification
3. Data Acquisition and Filtering
4. Data Extraction
5. Data Validation and Cleaning
6. Data Aggregation and Representation
7. Data Analysis
8. Data Visualization
9. Utilization of Analysis Results

This Life Cycle appears to have three or four more steps than the previous life cycle models. But in reality, they have just broken down what we have been referring to as **Prepare and Process** into smaller steps. It emphasizes the individual tasks required for gathering, preparing, and cleaning data before the analysis phase.