Sure, here are short and easy definitions for each topic:

1. **Hadoop**: A software framework for storing and processing big data across clusters of computers using simple programming models.

2. **Commands of Hadoop**: Specific instructions or actions used to interact with and manipulate data within the Hadoop ecosystem, such as HDFS commands for managing files or MapReduce commands for processing data.

3. **Big Data**: Extremely large and complex data sets that cannot be easily managed or processed using traditional data processing applications.

4. **MapReduce**: A programming model and processing technique for parallel and distributed computing of large data sets in the Hadoop framework, where data processing tasks are divided into smaller parts and processed independently across multiple nodes.

5. **HDFS (Hadoop Distributed File System)**: A distributed file system that provides high-throughput access to application data and is designed to store large files and data sets across multiple machines in a Hadoop cluster.

6. **R**: A programming language and software environment for statistical computing and graphics, commonly used for data analysis and visualization.

7. **Hive**: A data warehouse infrastructure built on top of Hadoop for providing data summarization, query, and analysis.

8. **Data Cleaning**: The process of detecting and correcting (or removing) inaccurate, incomplete, or irrelevant data in a dataset.

9. **Data Transformation**: The process of converting data from one format or structure into another to prepare it for analysis or other downstream processes.

10. **Data Integration**: The process of combining data from different sources into a single, unified view to facilitate analysis and decision-making.

11. **Data Model Building**: The process of designing and creating a structure or framework to organize and represent data for analysis or processing.

12. **R - Line Graph**: A type of graph that displays data points connected by straight lines, typically used to show trends or relationships over time.

13. **R - Bar Charts**: A type of graph that represents categorical data with rectangular bars, where the length of each bar corresponds to the value of the category it represents.

14. **R - Box Plots**: A graphical representation of the distribution of a dataset through quartiles, showing the median, quartiles, and outliers.

15. **R - Histograms**: A graphical representation of the distribution of numerical data, where data values are grouped into intervals and plotted as bars.

16. **R - Scatterplots**: A type of graph that displays the relationship between two numerical variables with points on a Cartesian plane.

17. **Scatterplot Matrices**: A grid of scatterplots showing the relationships between pairs of variables in a dataset, useful for identifying patterns and correlations.

18. **Tableau**: A data visualization tool that allows users to create interactive and shareable dashboards and reports from various data sources.

19. **1D (Linear) Data Visualization**: Visualizations representing data along a single dimension, such as line graphs or histograms.

20. **2D (Planar) Data Visualization**: Visualizations representing data in two dimensions, such as scatterplots or bar charts.

21. **3D (Volumetric) Data Visualization**: Visualizations representing data in three dimensions, such as 3D scatterplots or surface plots.

22. **Temporal Data Visualization**: Visualizations representing data over time, such as time series plots or Gantt charts.

23. **Multidimensional Data Visualization**: Visualizations representing data in more than three dimensions, often using techniques like parallel coordinates or multidimensional scaling.

24. **Tree/Hierarchical Data Visualization**: Visualizations representing hierarchical relationships in data, such as tree maps or dendrograms.

25. **Network Data Visualization**: Visualizations representing relationships between entities in a network, such as social networks or organizational charts.

Hadoop shell commands are simple instructions used to manage files and interact with data in a Hadoop distributed file system (HDFS). Some common ones include:

1. **ls**: Lists files and directories in the current directory.

2. **mkdir**: Creates a new directory.

3. **put**: Copies files from the local file system to HDFS.

4. **get**: Copies files from HDFS to the local file system.

5. **rm**: Deletes files or directories from HDFS.

6. **cat**: Displays the contents of a file.

7. **chmod**: Changes permissions of files or directories.

8. **chown**: Changes ownership of files or directories.

9. **mv**: Moves files or directories from one location to another.

10. **cp**: Copies files or directories within HDFS.

11. **du**: Shows disk usage of files and directories.

12. **df**: Displays disk space information of the filesystem.

Sure, here are simplified explanations for each topic:

1. **HBase**: A distributed, scalable, and column-oriented database management system designed to handle large volumes of structured data in a distributed environment, typically used in conjunction with Hadoop.

2. **NoSQL**: A type of database management system that provides a mechanism for storage and retrieval of data that is modeled in forms other than the tabular relations used in relational databases.

3. **ETL in Big Data**: ETL (Extract, Transform, Load) refers to the process of extracting data from various sources, transforming it into a usable format, and loading it into a target system, commonly used in big data environments to prepare data for analysis.

4. **Difference between SQL and NoSQL**: SQL databases are relational databases that use structured query language for defining and manipulating data, while NoSQL databases are non-relational and offer a flexible schema design and scalability for handling large volumes of unstructured data.

5. **Different kinds of big data analytics**: Big data analytics includes descriptive analytics (summarizing data to understand its characteristics), diagnostic analytics (analyzing data to determine why something happened), predictive analytics (using data to forecast future outcomes), and prescriptive analytics (suggesting actions based on analysis).

6. **CSV (Comma-Separated Values)**: A simple file format used to store tabular data, where each line corresponds to a row of data with the values separated by commas.

7. **JSON (JavaScript Object Notation)**: A lightweight data-interchange format that is easy for humans to read and write and easy for machines to parse and generate, commonly used for transmitting data between a server and a web application.

8. **Data Visualization**: The graphical representation of data to provide insights and facilitate understanding, typically through charts, graphs, and other visual elements.

9. **Types of Data Visualization**: Common types include line charts, bar charts, pie charts, scatter plots, histograms, heatmaps, and more, each suited to represent different types of data and relationships.

10. **Techniques of Big Data Visualization**: Techniques include aggregation (combining data into summary representations), filtering (removing unwanted data), drill-down (navigating from high-level summaries to detailed data), and interactivity (allowing users to manipulate visualizations).

11. **Tools for Data Visualization**: Tools include Tableau, Google Charts API, D3.js, Matplotlib, Plotly, and many others, each offering various features and capabilities for creating visualizations.

12. **Data Visualization using Tableau**: Tableau is a popular data visualization tool that allows users to create interactive and shareable dashboards and reports from various data sources with a user-friendly interface.

13. **Google Chart API**: Google Chart API is a tool provided by Google to create customizable, interactive charts and graphs for web pages, using simple JavaScript code to embed visualizations into web applications.

14. **Candela**: Candela is a library for building scalable, web-based visualizations using the Vega and Vega-Lite grammar, designed to simplify the process of creating complex visualizations in a web environment.

[Fill Missing Values: Mode or Mean (openai.com)](openai.com)

[Data Challenges: Big Solutions (openai.com)](openai.com)