Here is an overview of key concepts and techniques for using **Tableau** for data visualization and dashboard creation.

## Connecting Tableau to Data Sources 🔗

Tableau can connect to a wide variety of data sources, allowing it to act as a powerful analysis tool regardless of where your data resides.

| Connection Type | Description | Examples |
| --- | --- | --- |
| **File-Based** | Local files easily accessible on your computer. | Excel, CSV, JSON, Access. |
| **Relational Databases** | Requires credentials and a server name to connect to structured tables. | MySQL, PostgreSQL, SQL Server, Oracle. |
| **Cloud Data Sources** | Connectors for cloud-based data warehouses or applications. | Amazon Redshift, Google BigQuery, Salesforce, Google Sheets. |
| **Web Data Connectors (WDC)** | Used to connect to data accessible over HTTP, such as internal or public APIs. | Custom APIs, JSON data feeds. |

### Live vs. Extract

When connecting to a database, you choose between two primary modes:

1. **Live Connection:** Tableau queries the database directly, and the dashboard reflects real-time data changes. This is fast for small queries but can slow down dashboard performance if the underlying data source is large or slow.
2. **Extract:** Tableau pulls a snapshot of the data and saves it in a highly optimized file format (.hyper). This offers **faster performance** as calculations are done locally, but the data is not real-time unless the extract is refreshed.

## Creating Worksheets and Interactive Dashboards 🎨

The visualization process in Tableau moves from individual charts (Worksheets) to integrated displays (Dashboards).

### Worksheets (The Building Blocks)

A **Worksheet** is the canvas where you build individual charts. You create a visualization by dragging **Fields** (columns) onto the **Shelves** (Rows and Columns), the **Marks Card** (Color, Size, Detail, Text), and the **Filter Shelf**.

* **Pills and Shelves:** Fields dragged onto the canvas appear as "pills." Blue pills are **Discrete** (categories), and Green pills are **Continuous** (numerical values). Their placement determines the axis:
  + **Rows Shelf:** Defines the Y-axis.
  + **Columns Shelf:** Defines the X-axis.
* **Marks Card:** Controls the visual properties of the data points (e.g., using a field on the **Color** card to color data points by category, or on the **Size** card to represent magnitude).

### Dashboards (The Presentation Layer)

A **Dashboard** is a collection of several related worksheets, filters, parameters, and other objects presented in a single view to provide a holistic, interactive story.

## Incorporating Filters, Parameters, and Calculated Fields ✨

These tools provide interactivity and extend the analytical power of the raw data.

### 1. Filters

**Filters** are used to restrict the data shown in a view, worksheet, or dashboard.

| Type of Filter | Description | Example |
| --- | --- | --- |
| **Dimension Filter** | Selects or excludes specific members of a dimension (e.g., categories). | Showing data only for the states of **'New York'** and **'Texas'**. |
| **Measure Filter** | Filters a view based on the aggregate value of a measure. | Showing only categories where the **SUM(Sales) is greater than $100,000**. |
| **Context Filter** | An independent filter executed *before* other dimension filters. Useful for limiting the data set size for better performance. |  |

### 2. Parameters

A **Parameter** is a **dynamic placeholder** that replaces a constant value in formulas, filters, or reference lines. Unlike filters, parameters are **manual inputs** set by the user or designer.

* **How they work:** The user selects a value (e.g., from a drop-down list), and this value is passed to a **Calculated Field** to dynamically change the visualization.
* **Example:** A parameter named Top N Customers allows the user to select 5,10, or 20. This parameter is then used in a rank calculation to show only the top N customers.

### 3. Calculated Fields

**Calculated Fields** are new fields created using formulas based on existing data fields. They allow you to perform transformations, aggregations, and logic that don't exist in the original source.

| Type of Calculated Field | Description | Example Formula |
| --- | --- | --- |
| **Basic Calculation** | Row-level math and logic (executed for every row). | [Profit] / [Sales] (to calculate the Profit Ratio). |
| **Aggregated Calculation** | Aggregating data across a group of rows. | SUM([Sales]) / COUNTD([Customer ID]) (Average Sales per Customer). |
| **Level of Detail (LOD) Expression** | Advanced calculations to control the level of aggregation independently of the visualization's structure. | \{FIXED [Region] : SUM([Sales])\} (Calculates total sales for the region, regardless of what's shown on the Row/Column shelves). |