

Interactive Data Visualizations using JS lib D3

Kantha Cikaiahgari
Pitss GmbH
Stuttgart

Keywords:

Visualizations, REST (Representational State Transfer) web services, Scalable Vector Graphics(SVG), tool tip

Introduction

Data Visualization is Analyze & communicates information clearly and effectively through graphical means. Data visualization is the pictorial or graphical representation of the data. It deals in figuring out the ways to translate data into visuals that are more readily understood. It is a key feature in the success of data analysts and businesses. It helps the analysts and decision makers in effectively understanding the difficult concepts and in identifying the new patterns.

Why Visualizations need to be interactive

Interactive Visualizations provide more information, helps in easier perception of more complex data, promotes exploration and provides support in decision making.

Good User Experience relies on 3 primary rules: Overview first, Zoom, filter and details-on-demand

Business Usecases for Data Visualization

The key objective of a business usecase is to illustrate how the business processes are used by its partners and users. However, the density of the data makes it difficult for decision makers to use the data effectively. Here plays the data visualization a key role. It presents large data in an easy to understand and graphical format to provide decision makers perceptive information.

Different types of Data can have different requirements or usecases where each type of data has to be visualized in different ways. For example, a business process flow can be better visualized using BPMN (Business Process Modeling Notation) than any kind of graphs.

The probable primary usecases for the data visualization include

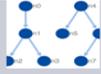
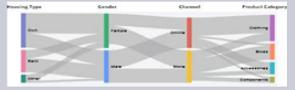
1. Navigation Flow, how a user could run through a system with different applications.
 - a. Code Flow, how the algorithm could run through different units of the application. E.g.: Call Stack
 - b. Data Lineage, how data could run through a system of different ETL (extract - transform - load) components.
2. The process flow, how user navigate in real for a single process through the system. That could be aggregate for different user and different objects inside the system.

It also helps to visualize

- Dependency of more than 1000+ different packages and how they interact with each other. Eg: Dagre Layout

- Process flows like how a user runs through different forms in an application
Eg: CoSE (Compound Spring Embedder) Layouts
- How data flows through different ETL (Extract Transform and Load) Components of the Database. Eg: Sankey Diagrams

Different Data Models - supported Layouts

Model	description	Supported Graphs or Layouts	Examples
Network	Node/link Diagram(uses link based layout algorithms)	Dagre, CoSE, cola	
Tree / Hierarchal	General tree Visualization	Dendrogram, treemap	
Temporal	Visualizing Data lineage	Alluvial, <u>Sankey</u>	
Arbitrary Graphs	Node-Link Graphs, Matrix	force directed graphs	
Processes	Process diagrams	<u>CoSE</u> (compound Spring Embedder) layout	

Sankey Diagram

Sankey Diagram depicts flows of any kind, where the width of each flow pictured is based on its quantity. Sankey diagrams are very good at showing particular kinds of complex information - where money came from & went to (budgets, contributions), flows of energy from source to destination and flows of goods from place to place and potentially many more.

The width of the arrows is proportional to the flow quantity. The arrows show flows from one node to another node - ideal for production systems or for value chains, as this cannot be achieved using pie or bar charts. Sankey diagrams are more suitable for visualizing energy or material flows when compared to conventional bar or pie charts and even flow charts. It also provides effective communication of messages.

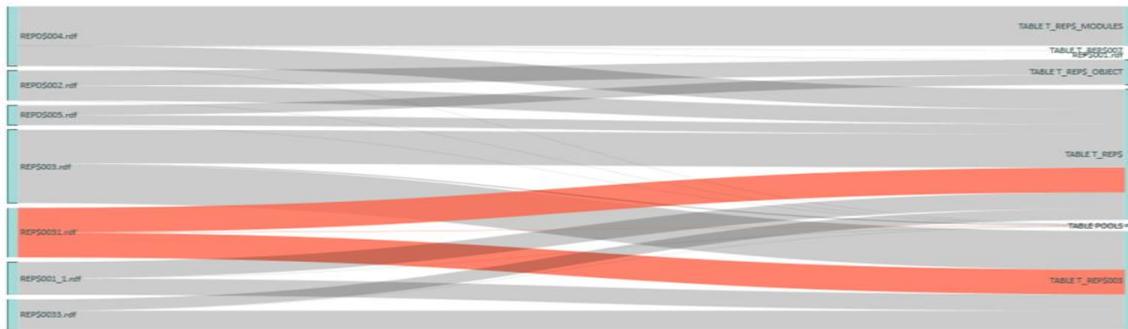


Figure 1. Sample Sankey Diagram

It gives overall view of the complete system and at the same time allows interactions to know how the nodes are connected in the visualization. Sankey diagrams can be drawn using D3 visualization.

CoSE Bilkent Layout

CoSE is an algorithm based on *force-directed layout* scheme with extensions to handle

- Multi-level nesting
- Edges between nodes of arbitrary nesting level
- Varying node sizes
- Other possible application-specific constraints.

JavaScript Libraries: D3 and Cytoscape

JavaScript libraries are the best platforms which provide built in functions to be utilized. D3 stands for Data Driven Documents. D3.js is a JavaScript library for manipulating documents based on data. D3's emphasis on web standards provides the full capabilities of modern browsers without binding to a fixed framework, combining powerful visualization components and a data-driven approach to DOM (Document Object Model) manipulation. DOM model is explained in the below section. D3.js focuses on binding data to DOM elements. D3.js supports in data exploration by providing control over data's representation using interactivity.

D3 Features

D3.js is an open source project. It requires very less code and can be used to produce a variety of visualizations from simple to complex including user interaction and transition effects. Some of its salient features include:

- *Data Driven:* D3 is data driven. It can use static data or fetch it from the remote server in different formats such as Arrays, Objects, CSV, JSON, XML etc. to create different types of charts.
- *DOM Manipulation:* D3 allows manipulation of the Document Object Model (DOM) based on given data.
- *Data Driven Elements:* It authorize data to dynamically generate elements and apply styles to the elements, be it a table, a graph or any other HTML element and/or group of elements.
- *Dynamic Properties:* D3 gives the flexibility to provide dynamic properties to most of its functions. Properties can be specified as functions of data.

- *Types of visualization:* With D3, there are no standard visualization formats. But it enables creations from an HTML table to a Pie chart, from graphs and bar charts to geospatial maps.
- *Custom Visualizations:* Since D3 works with web standards, it offers complete control over the visualization features.
- *Transitions:* D3 provides the transition () function. This is quite powerful because internally, D3 works out the logic to interpolate between values and find the intermittent states.
- *Interaction and animation:* D3 provides great support for animation with functions like duration (), delay () and ease (). Animations from one state to another are fast and responsive to user interactions.

Cytoscape Features

Cytoscape is an open source software platform for visualizing complex networks and integrating these with any type of attribute data.

- Multi-graph layouts support (including traditional graphs, directed graphs, undirected graphs, multigraphs and hypergraphs)
- Interactive graph manipulation and analysis possible (CRUD)
- Several well-known graph theory algorithms are included, such as connectivity search, shortest path, minimum span tree, minimum cut, ranking and centrality measures.
- As a Java Script application, the use of CSS and HTML elements is of particular importance.
- Import graph and data via Jason & Export (PNG , PDF or JPG) Incl. High Resolution for Publications
- animations usability and automatic layout design

How Javascript libraries integrated into APEX?

Get the Java Script Library from the Javascript library (D3, Cytoscape) Home Page, customize Java script with graph properties and data definition and standard integration of scripts into Apex like all other javascript files.

Customize Java Script with graph properties and data definition

- Data is predefined in the SQL database
- Nodes, Edges & Relations are provided by APEX RESTFUL SERVICES.
- Data retrieved using a JSON URL call can be intercepted.
- The URL contains RESTFUL SERVICES address, which allows CRUD operations in the database.

Benefit Analysis from the Apex Perspective

Free Java Script Library

One application instead of 2 applications Apex, JS

- Same Look and Feel for the application
- Use of Integrated CSS (Theme Roller from Apex)Page Templates, Skins
- Declarative Components, Utility Classes
- Optimal database calls and operations for dynamic design and presentation of data on the surface using Apex

- Easy integration for new developers
- Avoid rework

Contact address:

Kantha Cikaiahgari
Pitss GmbH
Industrie Straße 3
70565 Stuttgart

Phone: +49 (0) 711/ 91 40 12-25
Email: kcikaiahgari@pitss.com
Internet: www.pitss.de