

10

Advanced Data Visualizations

Spark your inspiration and
take your visual delivery to the next level

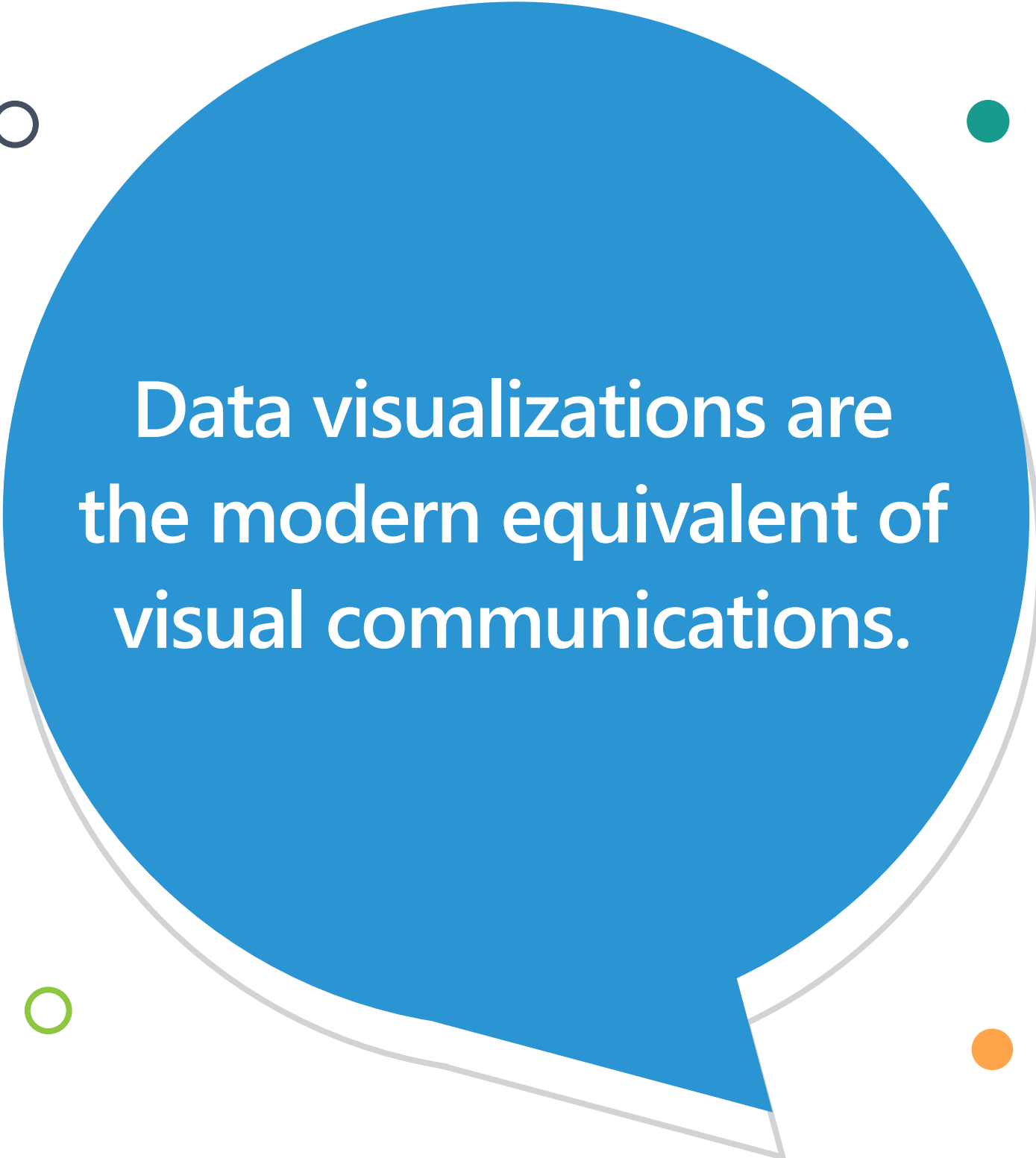
BY

Jordan Zenko

Community & Content Manager, Dundas Data Visualization

Ariel Pohoryles

Director of Product Marketing, Dundas Data Visualization



**Data visualizations are
the modern equivalent of
visual communications.**

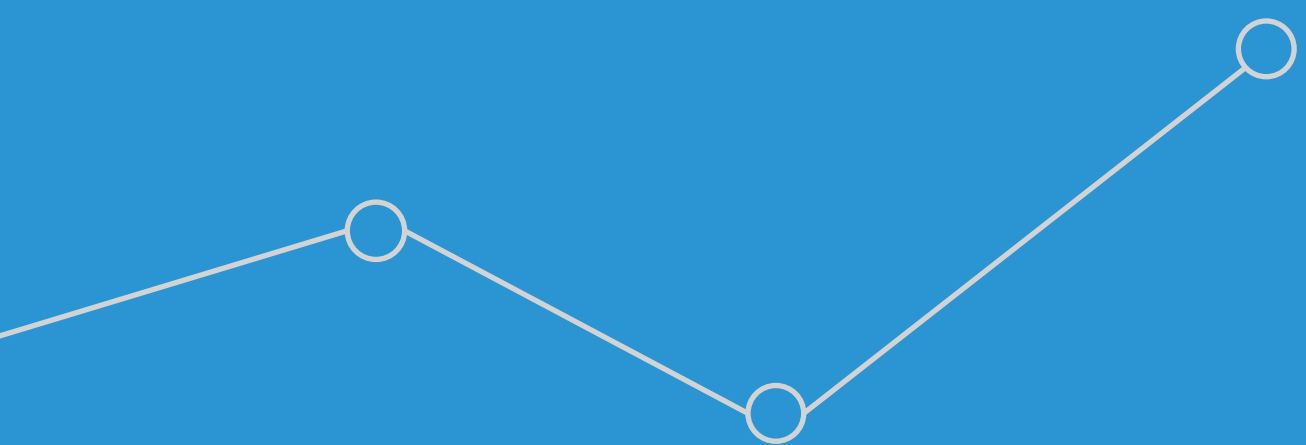


In the context of business intelligence, data visualizations are interactive representations and techniques that are displayed as visual objects, whose purpose is to clearly communicate the significance of data to those who view them.

Using data visualizations, patterns, trends and correlations within data that would normally go un-detected in non-graphic renditions (i.e., text-based data), are more likely to be revealed and identified. Business intelligence platforms such as Dundas BI, Dundas Data Visualization's business intelligence, dashboarding and data visualization software, enable its users to discover greater insights within their data and to step beyond the confines of traditional reporting tools with extremely interactive and engaging data visualizations.

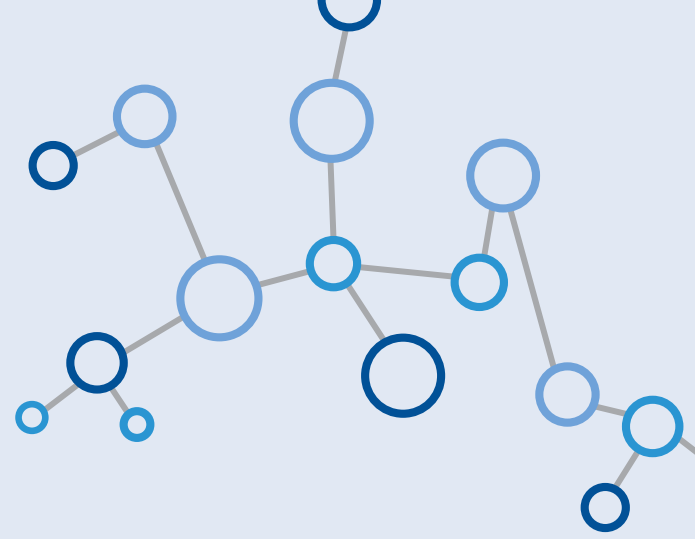
Virtually all business intelligence platforms feature data visualization capabilities, however, a distinct attribute of Dundas BI, is its hallmark, advanced smart visualizations that represent more than 20 years of technology leadership and innovation. These data visualizations are rich, extremely interactive, flexible, aesthetically stunning, and their extreme customization properties arm users with staggering granular control over almost all visual design elements.

Let's take a closer look at 10 advanced data visualizations that Dundas BI offers out-of-the-box, that'll take your visual delivery to the next level, and get people engaged and interested in the message you're trying to convey. While some of these visuals are unique and won't be found in other business intelligence platforms, all of them will fundamentally help you to create best practice, interactive dashboards and rich reports for complete data insight with higher user adoption rates.



**Relationship visualizations
are used to show relationships
and connections between
the data.**

Relationship Diagram



What is it?

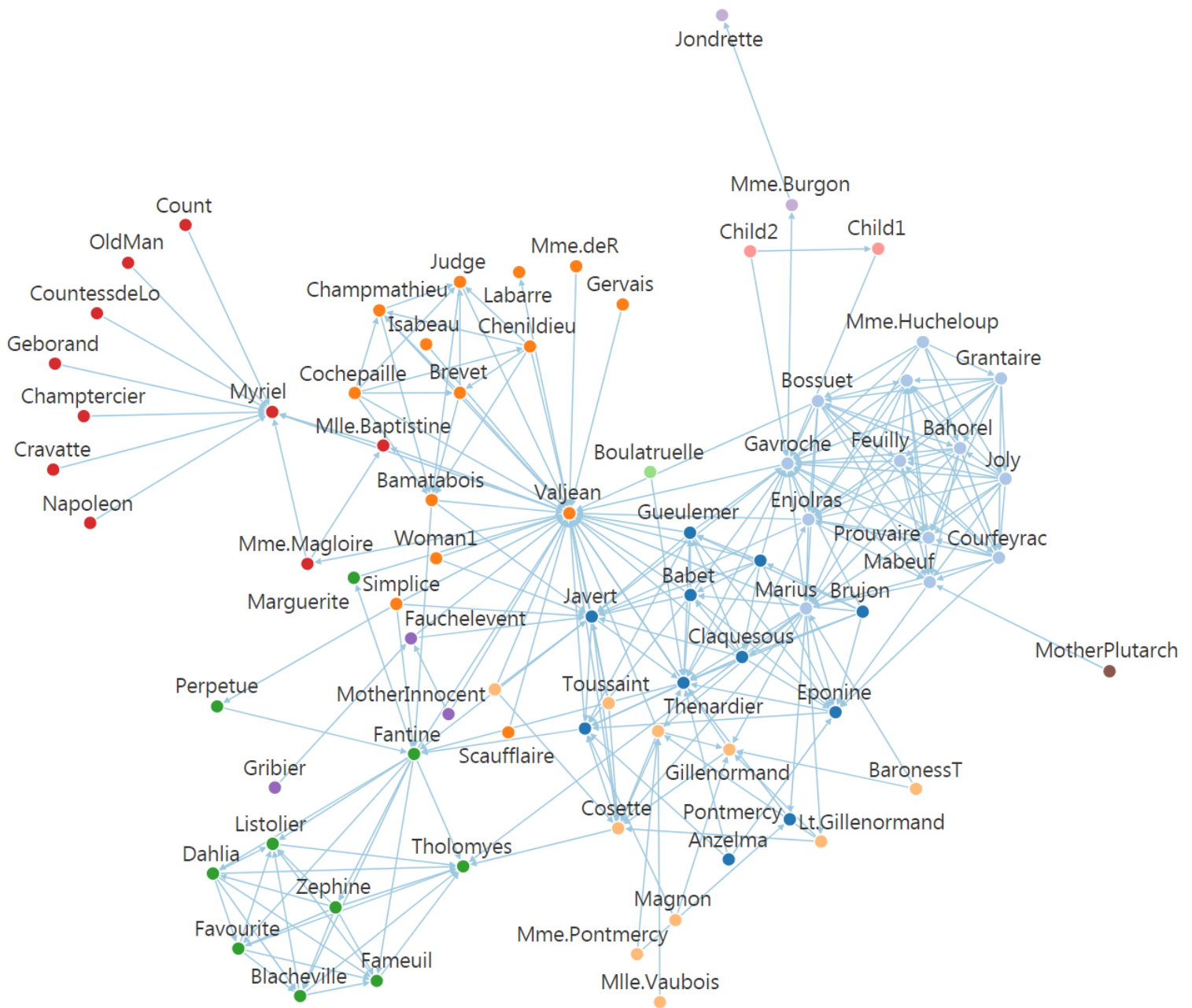
The Relationship Diagram, also known as a Force directed Graph, displays a collection of nodes (i.e., circles) and links (i.e., straight or curved lines) that connect the different nodes together.

The history of the Relationship Diagram can be traced to psychiatrist, psychosociologist, and educator, Jacob L. Moreno. Moreno, one of his eras leading social scientists, and the foremost pioneer of group psychotherapy, developed Sociograms, which helped him analyze choices or preferences within a group. The Sociogram's purpose is to diagram both the structure and patterns of various group interactions. Much like the Sociogram, Relationship Diagrams show interconnectivity, and help illuminate the types of relationships between groups of entities.

When can I use it?

This visual is great at analyzing and uncovering hidden relationships within data that other visualizations easily miss. While other visualizations show many of the relationships implicitly within the visualization, the Relationship Diagram displays explicit connections between many different data points at a glance.

You'll want to use the Relationship Diagram when analyzing social media connections and impact, customer engagement experience, directions of connections or other forms of hierarchy relationships.



 This relationship diagram shows character co-occurrence in Les Misérables. A physical simulation of charged particles and springs places related characters in closer proximity, while unrelated characters are farther apart.

Data based on character co-appearance in Victor Hugo's Les Misérables, compiled by Donald Knuth.

Sankey Diagram



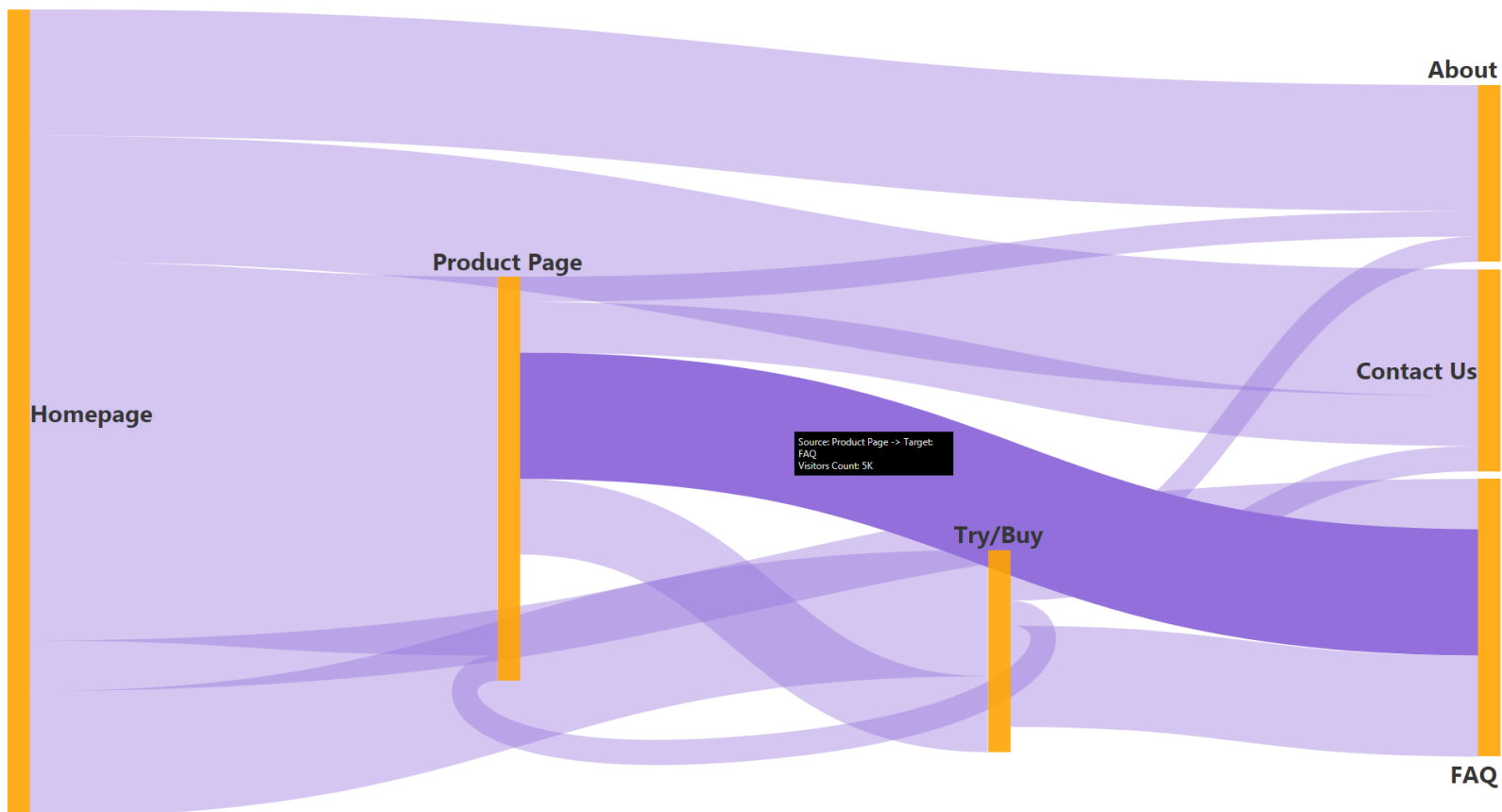
What is it?

The Sankey Diagram, often considered a sister-visual to the aforementioned Relationship Diagram, is designed to show relationships in a flow (defined breakdown) as opposed to by connections without any specific structure. This diagram is comprised of links, where the width of a link is proportional to the magnitude of the flow.

Officially fashioned in 1898 by Irish Captain Matthew Sankey, the first charts were simply used to display one type of flow (For example, Captain Sankey's diagram of steam engine efficiency that used arrows with widths that were proportional to heat loss), with color being added later to denote different types of flows and add more freedom to the Sankey Diagram. Surprisingly, Charles Minard's Map of Napoleon's Russian Campaign of 1812, which is widely-considered the "best statistical graphic ever drawn", is one of the most famous Sankey Diagrams, and actually pre-dates Captain Sankey's original Sankey Diagram.

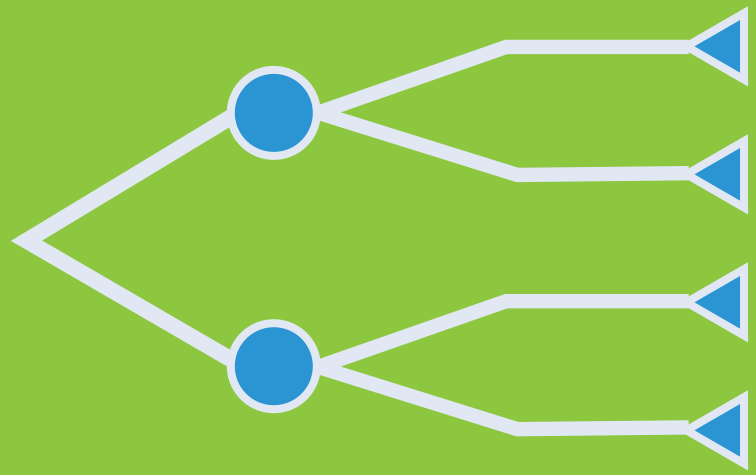
When can I use it?

Today, these attention-grabbing flow charts are most typically used to visualize distribution and loss of values such as energy, money, and/or materials across a flow (typically from left to right).



- ✎ This Sankey Diagram shows the flow of a website's visitors as they navigate their way around the site from one page to another. It helps to highlight popular flows and validate the website design is indeed meeting the way visitors are using the site. Note that since the navigation within a website isn't always linear, some of the flows may be going "backwards" (i.e., from left to right), such as the flow from the "Try/Buy" web page back to the "Product Page".

Tree Diagram



What is it?

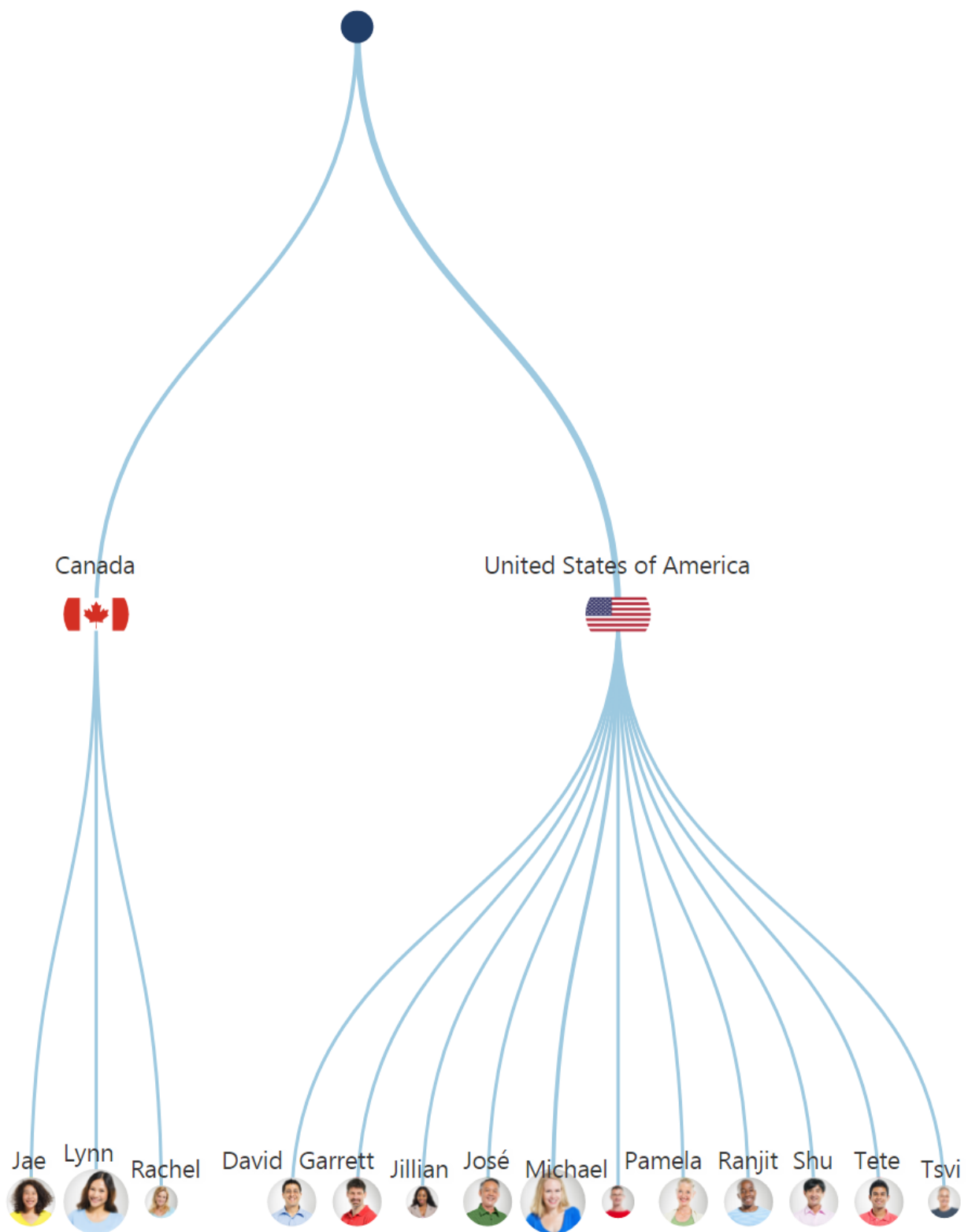
The Tree Diagram – dramatically statuesque and ever lush with its far-reaching branches of data – displays information about the ranked structure of the data, allowing us to visualize hierarchical data breakdowns and inter-relationships between categorical items.

Like the Sankey Diagram, the Tree Diagram is designed to show relationships in a flow. However, unlike the Sankey Diagram, whose link widths are proportionate to the magnitude of the flow, the Tree Diagram focuses on the different paths of data that are available. Typically, the structure of a Tree Diagram consists of elements such as nodes and branches, whereby nodes are linked by branches that represent the relationships and connections between members.

When can I use it?



Presented like a tree (hence the name), this relationship visualization is often used to display organizational charts and other family descent structures (such as to display the origins and classifications of species).



- ✎ This Tree Diagram shows the breakdown of sales across countries and the sales executives within those countries. The size of the sales executive's images varies based on their sales performance, helping to visually highlight the most successful sales executives. This view helps to quickly review the sales team organization.

Chord Diagram



What is it?

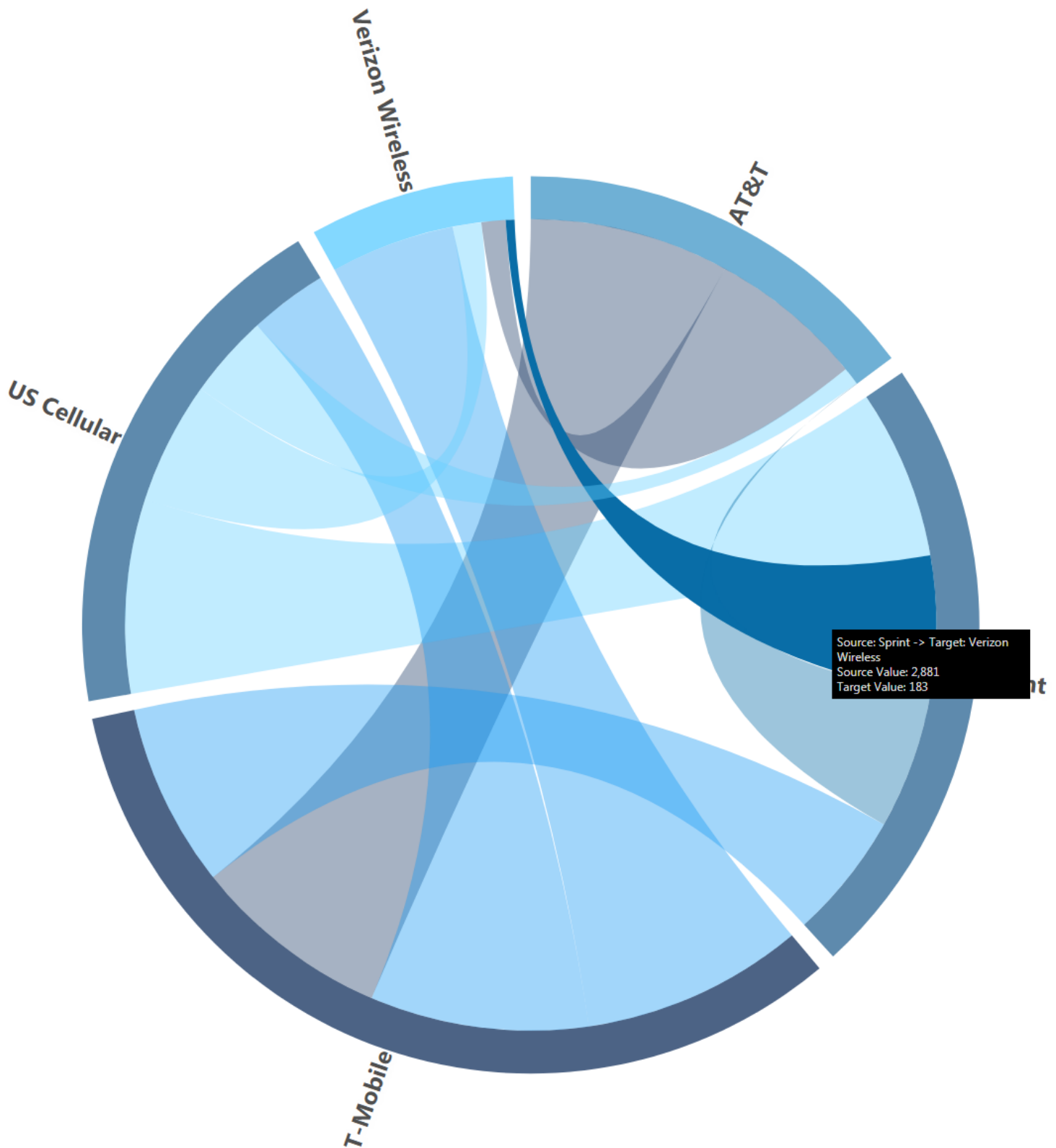
With Chord Diagrams, relationships in data can be seen from an entirely unique perspective. Often a difficult visualization to properly explain – or perhaps better said – often a misunderstood visualization, the Chord diagram displays the inter-relationships between categorical items. With a Chord Diagram, data is positioned radially around a circle with each node (or slice) corresponding to each category. The relationships between nodes are drawn together with sweeping arcs, their widths proportionate to the magnitude of the flow. What we're left with, is a remarkably aesthetically appealing visualization that is much easier to digest than say a pivot table, when attempting to measure movement between categories.

While other relationship visualizations show hierarchical relationships between data (think Sankey and Relationship Diagram), the Chord Diagram visualizes the relationships within the same level. That is to say, this type of visualization is most powerful when visualizing the relationships between 1-level hierarchy members.

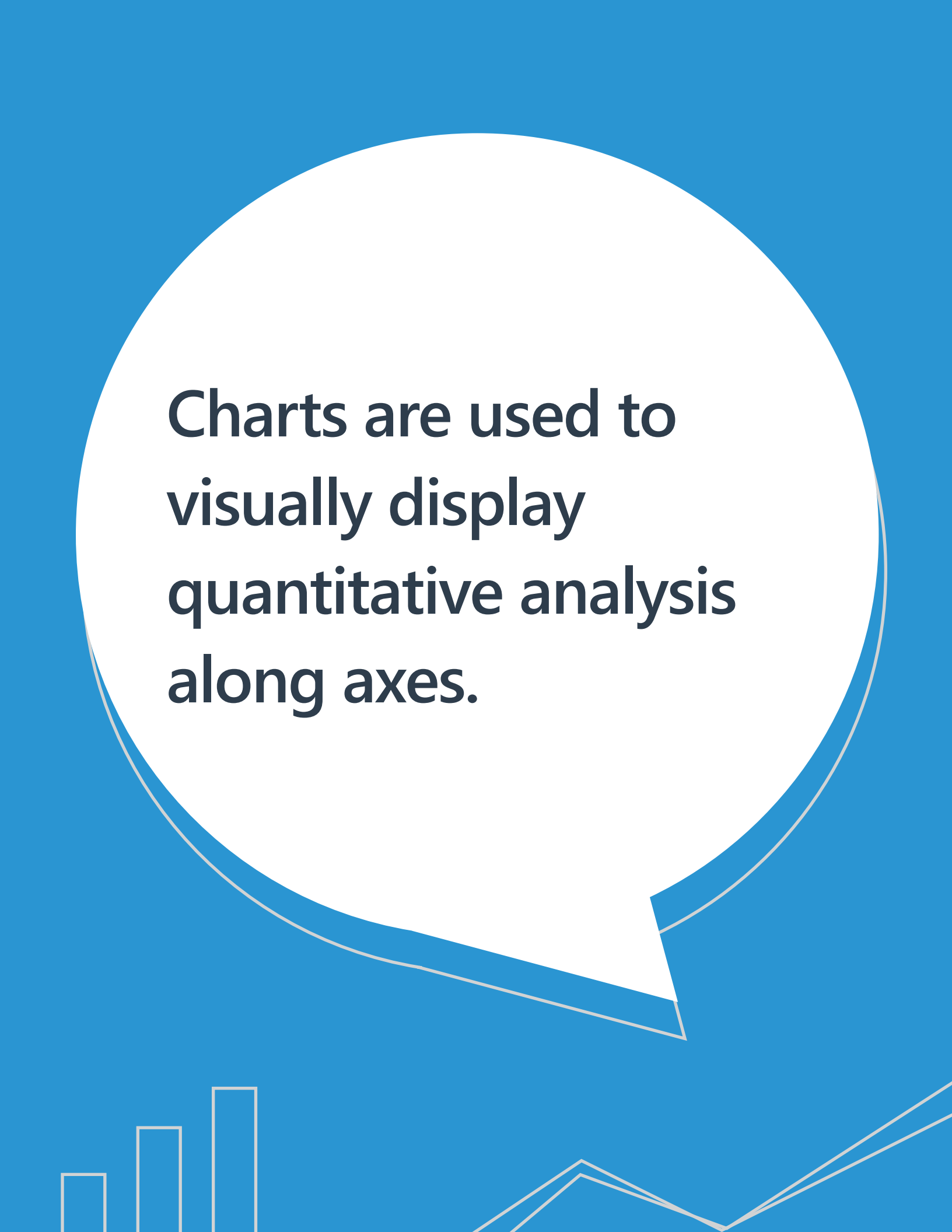
When can I use it?



An example use case for the Chord Diagram might be when monitoring churn rates of subscribers moving from one telecom network to another. By using a Chord Diagram in this scenario, we'd be able to immediately see the bi-directional paths of customers moving between networks, which would be much more tedious and tremendously difficult to conclude from other table-type visualizations.



-  This Chord Diagram shows the movement of network subscribers from one telecom provider to another. It helps to identify the leading telecom provider in regards to churn rates and where they are getting the new subscribers from to validate the success of specific campaigns promoting the movement from one provider to another.



**Charts are used to
visually display
quantitative analysis
along axes.**

Sunburst Chart



What is it?

Sunburst Charts are simply stunning data visualizations, and are great at both representing hierarchical data structures and increasing engagement of users. Often hidden in the simplicity of a circle, is a multitude of intriguing, distinctive qualities. What is lacked in precision, is often gained with an intuitiveness and aesthetic beauty that comes from its circular nature, which feeds our unique fascination with this shape.

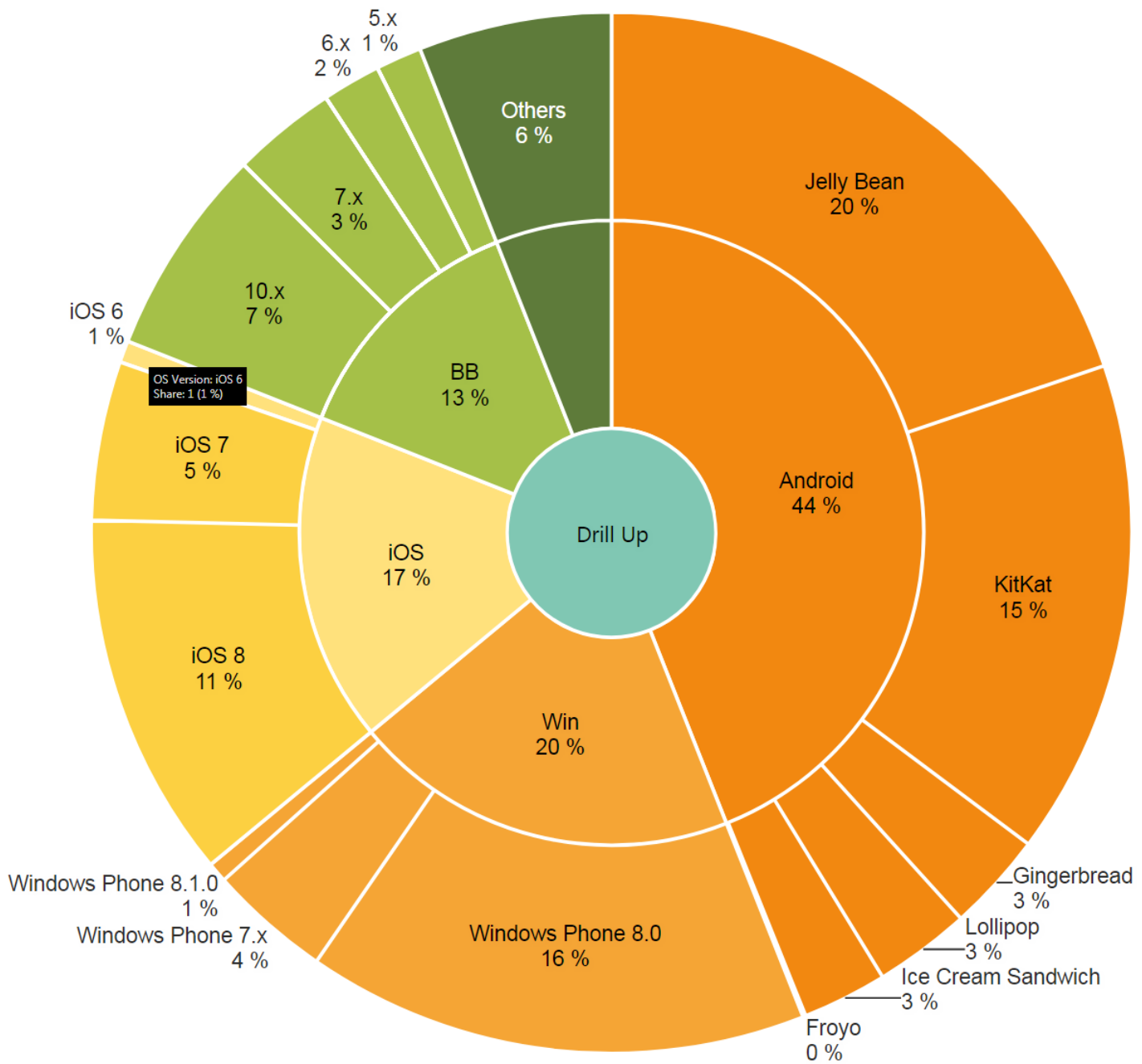
A Sunburst Chart, otherwise known as a Radial Treemap or Multi-level Pie Chart, shows a hierarchy breakdown through a series of concentric circles, which are sliced and diced depending on their hierarchical relationship to the root node. The root node, or parent node, is the first part of a data structure, which links to multiple nodes, containing data fields. Essentially, a Sunburst Chart allows the user to drill down into their data without filtering the original members.

And while this chart may not be as space-efficient as some alternatives, it allows users to easily interpret it at a glance, and can elicit an immediate response. It's simple, tasteful, ideal when representing hierarchical data structures, and above all else, effective.

When can I use it?



Sunburst Charts are particularly helpful when displaying multiple hierarchy levels at once, and allow for easier visualization of data at various levels. Furthermore, they're extensible to numerous events and solutions that have anything to do with hierarchy.



This Sunburst Chart shows the breakdown of the different mobile OS market share. It is grouped to the type of OS (iOS, Android, etc.) and then broken down to specific versions of those OS.

Waterfall Chart



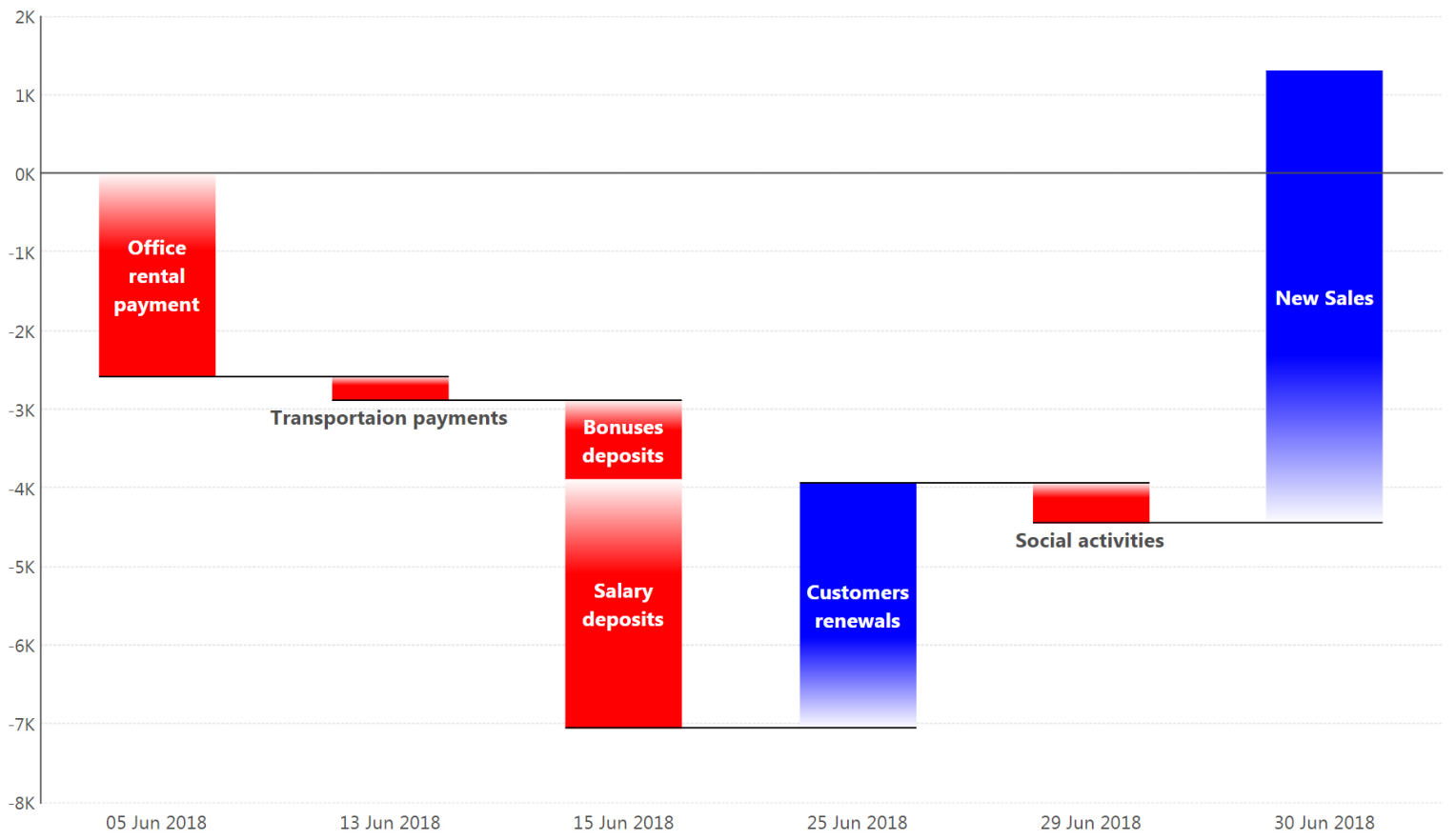
What is it?

The Waterfall Chart, popularized by consulting firm, McKinsey & Company, is a sterling way to delineate how a measure changes over a particular sequence, and how those changes impact the total. The Waterfall Chart clearly illustrates the cumulative effect of a sequence of events, or a process that has different transactions, some positive, some negative. In other words, it allows us to understand how we got from point 'A', to point 'B'.

Often times, when observing incremental change of individual values over time, a Bar Chart is the go-to visualization. Unfortunately, a Bar Chart will only offer a decomposition of each value, and will not emphasize the status of the aggregate balance. This is where the Waterfall Chart shines. It is a precise and versatile visualization that depicts static and active data, whether or not a value is positive or negative, and its final total.

When can I use it?

What's unique about the Waterfall Chart, is that its display changes in the aggregate as an abrupt increase or decrease from its prior value, which is more akin to the nature of this shift. Some of the most widely used examples of Waterfall Charts are representations of balance sheets, income statements, cash flows and other sequences, which ebb and flow quantitatively.



- ✎ This Waterfall Chart shows the monthly balance of office transactions. Positive transactions are marked in blue, negative in red and there is a highlighted grid line at 0, helping to show if the cumulative value of the transactions is positive or negative.



Pareto Chart



What is it?

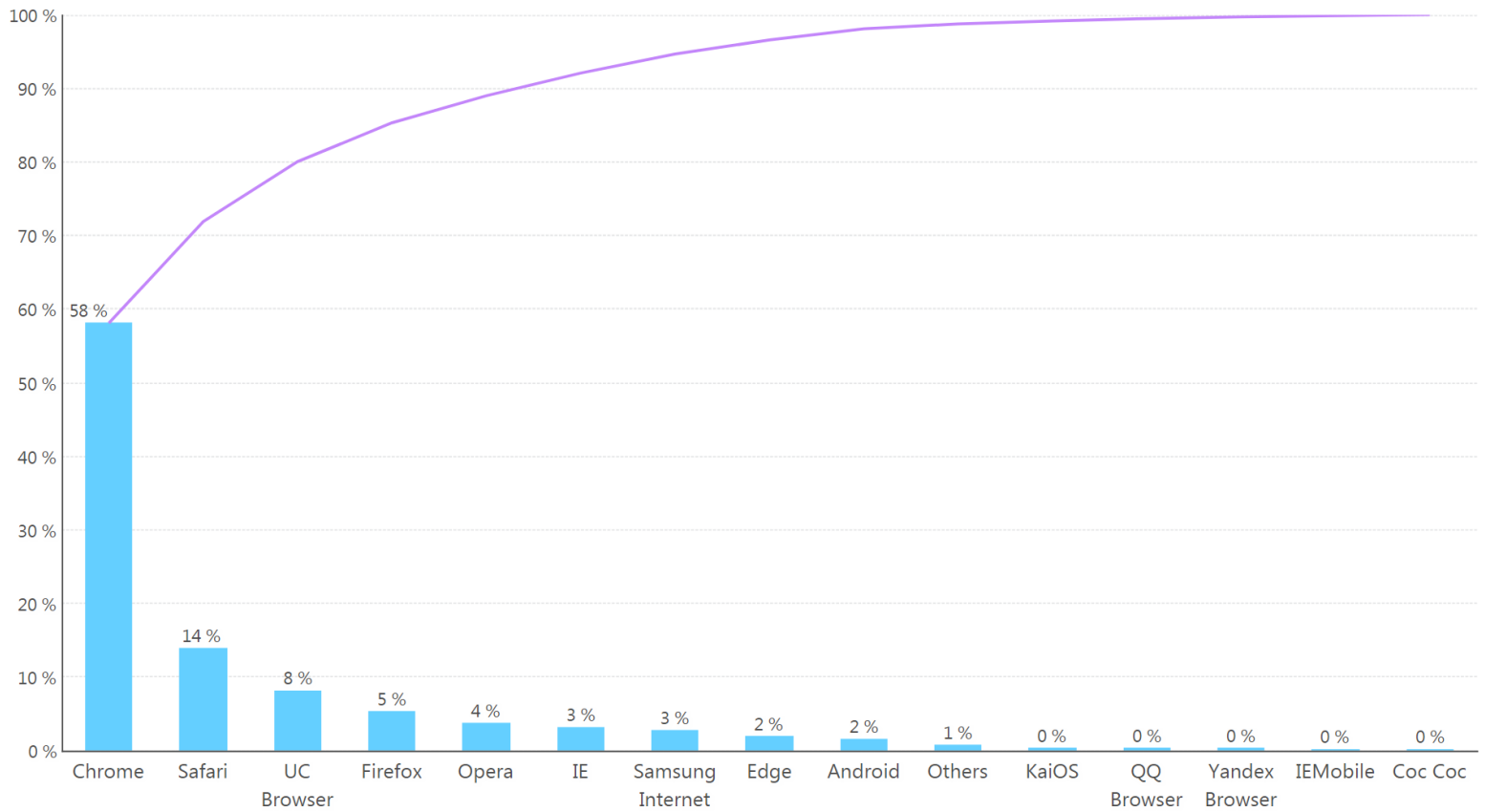
The Pareto Chart is a unique visualization in that it contains both bars and a line graph, where individual values are represented in descending order by bars, and the cumulative total of those values is represented by the line from left to right. For instance, the line begins at the top of the first bar, and as we move right, towards the second bar, the line displays the combined total of both bars as a percentage, representing the total share of these 2 items (i.e. products) out of the entire group of items. Upon reaching the final bar, the line's total will always be equal to 100%, representing the cumulative percentage share of all the bars (items).

The chart is a derivative of the Pareto principle, which states that, for many events, roughly 80% of the effects come from 20% of the causes. This principle is named after Italian engineer, Vilfredo Pareto, who observed that 80% of the land in Italy was owned by 20% of the population. The principle was further developed, as Vilfredo observed that roughly 20% of the peapods in his garden contained 80% of the peas.

When can I use it?



The Pareto Chart, derived from a histogram, views causes of a problem in order of severity from largest to smallest, ultimately demonstrating the Pareto principle. With a Pareto Chart, you are able to highlight the most important among a large set of factors. The most prevalent use case is for quality control, be it identifying the most common source of defects, the highest occurring type of defect, the most frequent reason for customer complaints, and so on. However, this visualization can easily be applied to other events, such as business, where 80% of profits are regularly elicited from 20% of products.



This Pareto Chart shows the 2018 breakdown of browsers market share. You can clearly see that 80% of the market is dominated by 3 browsers alone.

Parallel Coordinates Chart



What is it?

The Parallel Coordinates Chart is, put simply, not pretty. From afar, it often looks cluttered and indecipherable with its overlapping lines. But aesthetics is not its forte. When dissecting the value of a Parallel Coordinates Chart, you must set aside any prior reservations you may have about its resemblance to a Line Chart. Unlike a Line chart, which indicates change over time, the Parallel Coordinates Chart connects a series of values, each associated with a different value, that measure multiple aspects of something.

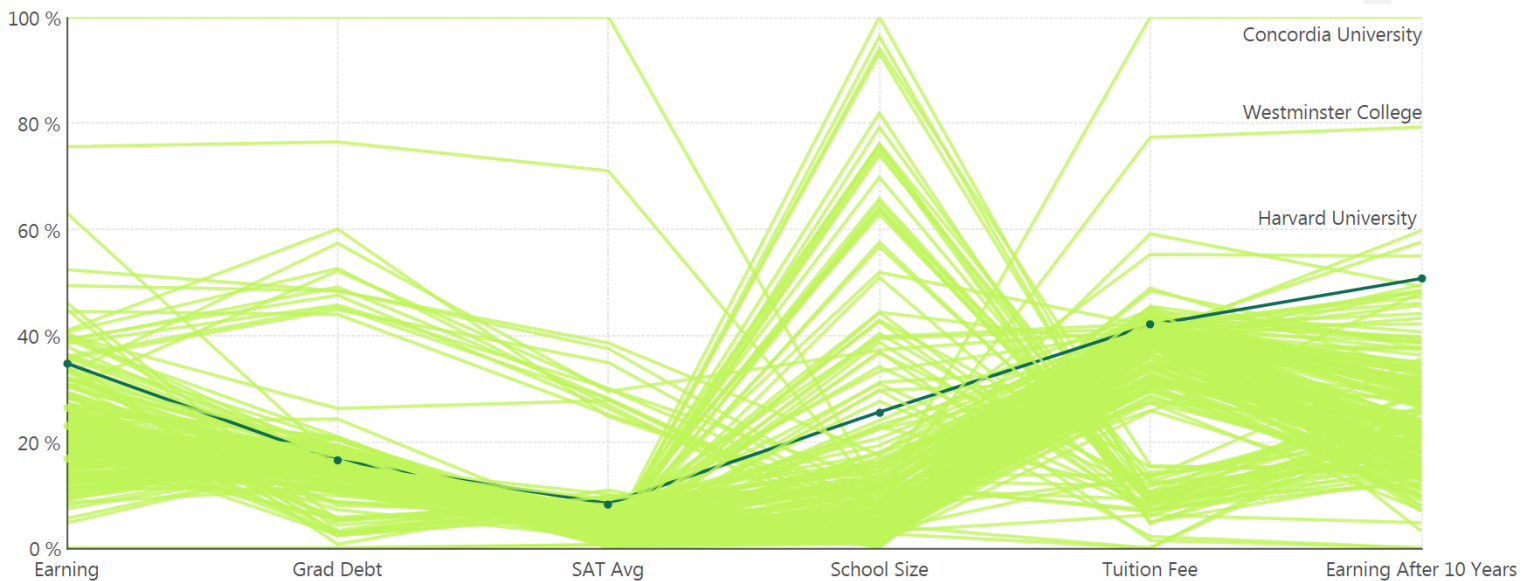
In other words, its strengths lay not in its ability to communicate the truth of data, but rather to bring meaningful multivariate patterns and comparisons to light.

Now, the challenge of analyzing multivariate data often boils down to the analyst's ability to see all of the different combinations displayed together at once, to quickly identify patterns and make comparisons without having to isolate two variables at a time, out of the many variables the dataset contains. Why the Parallel Coordinates chart is such a powerful and handy visualization, is because it allows us to look at a lot of data points at the same time, and see how close specific measures are in regards to patterns or how they stack up against measures in the same category.

When can I use it?

The best use case for the Parallel Coordinates Chart is when comparing an array of products with same attributes (i.e., comparing computer or car specifications across different models).

University	Earning	Grad Debt	SAT Avg	School Size	Tuition Fee	Earning After 10 Years
University of North Carolina at Chapel Hill	51,900	14,500	1,305	17,565	5,625	62,600
University of Notre Dame	69,400	17,125	1,433	8,367	38,477	83,100
University of Pennsylvania	75,700	17,125	1,436	10,842	38,970	106,900
University of Pittsburgh-Pittsburgh Campus	47,500	19,125	1,256	17,682	14,154	54,000
University of Puerto Rico-Mayaguez	26,300	5,500	1,205	12,474	2,008	31,400



 This Parallel Coordinates Chart shows US universities across multiple variables, helping you to benchmark and distinguish some universities from the rest based on different variables all normalized as percentages so they can be compared. In this example, you can see how the University of Pennsylvania graduates are gaining higher earnings after 10 years compared to most schools, but the SAT Avg required to get into the school isn't as high as for some of the other schools.

Range Bar Chart



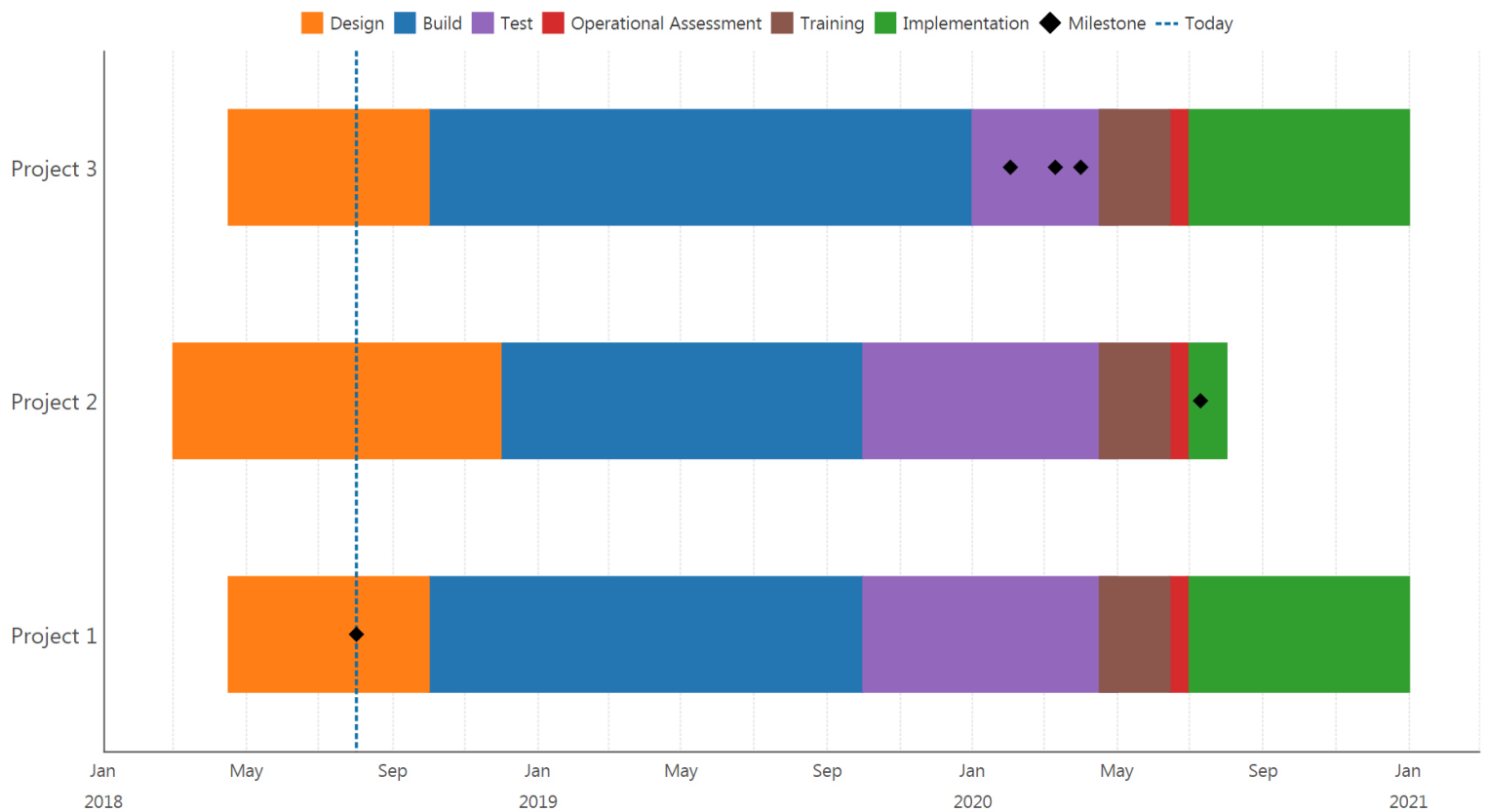
What is it?

The Range Bar Chart is a bar chart that displays bars for each dimension (category), ranging between a start value and an end value. A Range Bar Chart must have at least two measures (to be used as the start and end of the range bars) and one hierarchy. For example, the measures may represent low and high temperature values for a day, and the hierarchy will represent specific dates.

The Range Bar Chart was developed in the mid 1990's by Vicente Nicolellis, a Brazilian trader and broker who recognized how unstable the markets were at the time, and became interested in developing a way to capitalize on the volatility. His belief was that by taking solely price into consideration – thereby eliminating time from the equation – he could better interpret volatility and predict opportune times for trades.

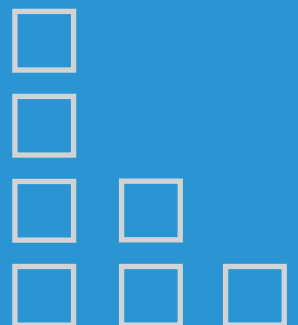
When can I use it?

While Range Bar Charts are most commonly used to provide a unique way to view and utilize the volatility of markets, they are extremely versatile, and can be used to represent the beginning and end of operations/tasks along a time axis, similar to a Gantt Chart. The Range Bar Chart is a great tool for tracking measurement ranges, tasks and project schedule adherence. Projects often have multiple phases (such as Planning, Design and Implementation), with each phase having milestones and dependencies. The Range Bar Chart can be used to show these milestones and dependencies within a project phase.



➤ This Range Bar Chart shows project development plans using different project phases and milestones, so at a glance, you can see the current state a project is in and if we are hitting the different timelines and specific milestones.

Tables are designed for organizing and displaying information, whereby data is arranged in rows and columns.



Correlation Matrix



What is it?

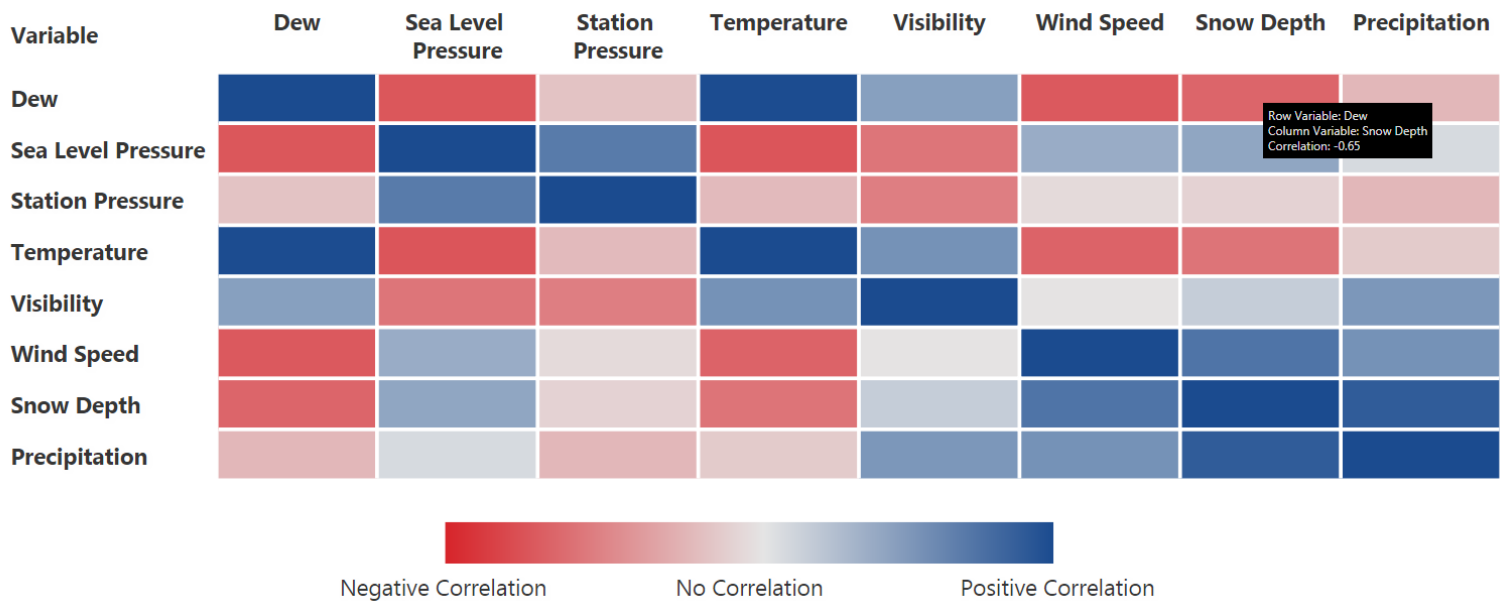
The Correlation Matrix is a table that displays the same variables across its rows as it does its columns, and whose correlation coefficients are mapped from -1 to +1 in respective colors (Red = Negative, Blue = Positive). The coefficients (values) that are situated diagonally on the table from top-left to bottom-right will always be blue (a value of +1) as this is where the variable is always correlated to itself. Conversely, with the Correlation Matrix being a grid, its other diagonal values will be symmetrical. This is because the correlation between variables x and y (row and column) are the same between variable y and x (column and row).

The Correlation Matrix will ultimately give you a complete view of the bi-variate correlations that exist in whatever dataset you're looking at.

When can I use it?

The Correlation Matrix is a great visualization to use when highlighting the positive and negative dependencies between multiple variables within your data in a single visual, rather than having to analyze and look for a hidden dependency one variable at a time. It can be used for any number of situations where measuring meaningful relationships is the objective, such as identifying the next steps in a Marketing campaign, to improve the design of a website, or to continue in-depth analysis on customers for additional correlation dependencies.

Day	Dew	Sea Level Pressure	Station Pressure	Temperature	Visibility	Wind Speed	Snow Depth	Precipitation
1	41	1,015	967	52	12	7	2,190,851	248,148
2	41	1,015	967	52	12	7	2,207,824	245,217
3	41	1,015	967	52	12	7	2,217,451	246,295
4	41	1,015	967	52	12	7	2,215,504	247,920



- ✎ This Correlation Matrix shows which weather variables impact (or don't) each other. For example, this matrix suggests there may be a negative correlation between the Wind Speed and the Dew quantity.

There are countless types of data visualizations, each varying in complexity and intent. What you choose to use depends heavily on the data you possess, and ultimately on the story you wish to tell.

Dundas Data Visualization's customizable and smart data visualizations are famous in the industry, and while it's unlikely you'll need to (or even want to) use all the visualizations discussed in this article, we're certain they'll satisfy your end users' needs. Using your own datasets, we invite you to get your hands dirty and explore for yourself what discoveries you can make with **Dundas BI**™

www.dundas.com