

مصورسازی داده

در دنیایی که به سرعت در حال تغییر و تحول است، درک وقایع باید بلادرنگ صورت گیرد، در غیر این صورت ممکن است زبان‌های جبران ناپذیری بوجود آید. ذهن و تفکر انسانی در درک و شناخت اجسام و اشکال قوی‌تر از اعداد و ارقام عمل می‌کند. البته این موضوع به این علت است که انتزاع اعداد برای بیان پدیده‌ها، پردازش بیشتری در ذهن انسان احتیاج دارد. برای مثال درک بلندی و کوتاهی قد دو نفر با چشم بسیار بهتر و سریع‌تر درک می‌شود تا دیدن رقم مربوط به ارتفاع قد آن‌ها. حتی ممکن است با توجه به موقعیت‌های جغرافیایی متفاوت و در کشورهای مختلف، مقیاس‌های اندازه‌گیری طول با یکدیگر متفاوت بوده و امکان مقایسه صحیح داده‌های حاصل از اندازه‌گیری قد آن‌ها نیز به راحتی میسر نباشد. برای مثال ارتفاع قد یک نفر برابر با ۵ فوت و قد نفر دوم ۱۶۷ سانتی‌متر باشد. واضح است که برای مقایسه قد این دو نفر فقط براساس ارقام دچار مشکل خواهیم شد. در این نوشتار با تکنیک و شیوه‌های مصورسازی داده آشنا شده و روش‌هایی برای توصیف اعداد و ارقام با سریع‌ترین و البته دقیق‌ترین میزان انتقال اطلاعات، ارائه خواهیم کرد. در این بین نیم‌نگاهی هم به انواع نمودارهای آماری و اطلاعاتی داشته و خصوصیات هر یک را مرور می‌نماییم.

فهرست مطالب این نوشته

مصورسازی داده ۱.

مصورسازی داده چیست؟ ۱.۱.

ویژگی‌های نمایش تصویری موثر ۱.۲.

مصورسازی داده با نمایش اطلاعات عددی ۱.۳.

درک تصویری و مصورسازی داده ۱.۴.

ادارک، شناخت و تجسم در انسان ۱.۵.

تاریخچه مصورسازی داده ۱.۶.

واژه‌شناسی ۱.۷.

نمونه‌ای از نمودارهای مورد استفاده در مصورسازی داده ۱.۸.

مقایسه ویژگی‌های نمودارهای مختلف در مصورسازی داده ۱.۸.۱.

خلاصه و جمع‌بندی ۲.

رسم و نمایش و رسم نمودار برای داده‌ها — معرفی و کاربردها از آنجایی که در اغلب موارد از نمودارها برای مصورسازی داده استفاده می‌شود، مطالعه نوشتارهای دیگر مجله فرادرس با عنوان‌های

— Matplotlib رسم نمودار در پایتون با و رسم آن در پایتون — به زبان ساده (Boxplot) نمودار جعبه ای پیشنهاد می‌شود. همچنین خواندن مطالب راهنمای کاربردی — SPSS نمودار در

نیز خالی از لطف نیست راهنمای کاربردی

مصورسازی داده

می‌نامند. در این شیوه، اطلاعات و داده‌ها، به واسطه تصاویر و شکل‌ها، قابل نمایش شده و بیننده قادر به درک سریع‌تر و بهتر (Data Visualization) «نمایش تصویری اطلاعات را» «مصورسازی داده اطلاعات نهفته در داده‌ها خواهد شد.

، آن‌ها را به (Graphs) ارتباط تصویری، نگاشتی سیستمی، بین تصاویر و مقادیر متغیرها (کمی یا کیفی) است که در مصورسازی داده به بهترین نحو صورت می‌گیرد. ویژگی‌های متنوع در نمودارها ابزاری مهم برای مصورسازی داده تبدیل کرده است. وجود رنگ، اندازه‌های مختلف برای هر دنباله از داده‌ها و همچنین نمایش روند تغییرات، از مواردی است که نمودارها را برای ارتباط تصویری بهتر، نسبت به جداول و گزارشات متنی، متمایز می‌کند.

ابزارهایی مهمی در مصورسازی داده محسوب می‌شوند. به این ترتیب، داده‌های عددی (کمی) یا اطلاعات کیفی، به صورت تصاویر، خطوط، (Icons) نمودارهای آماری، اینفوگرافیک، شکل‌ها و نمادها میله یا قطاع‌هایی از دایره، تبدیل شده و اطلاعات مربوطه را منتقل می‌کنند.

ارائه یک روش موثر برای مصورسازی داده کاربران و محققین حوزه تحلیل داده را در شناخت و کشف روابط بین پدیده‌ها و داده‌های حاصل از اندازه‌گیری آن‌ها، یاری می‌رساند. به این ترتیب مجموعه داده‌های پیچیده، به شکلی ساده و موثر، نمایش داده می‌شوند و امر استخراج قواعد یا تحلیل آن‌ها آسان می‌گردد. البته اغلب برای مقایسه یا خلاصه‌کردن داده‌ها از جدول نیز استفاده می‌شود ولی زمانی که تعداد گروه‌ها برای مقایسه زیاد باشند، جدول‌ها نیز کارایی خود را از دست می‌دهند.

قرار [\(Descriptive Statistics\) آمار توصیفی](#) از دیدگاه بسیاری از صاحب‌نظران، مصورسازی داده یک هنر و در عین حال یک علم است و به عقیده بعضی، تکنیک‌های مصورسازی داده در شاخه می‌گیرد.

مصورسازی داده چیست؟

مصورسازی داده به روش و تکنیک‌های گفته می‌شود که به منظور برقراری ارتباط داده یا اطلاعات با بیننده با سرعت و دقت، مورد استفاده قرار می‌گیرد. تبدیل مقادیر عددی به صورت اشیاء بصری (به عنوان مثال، نقاط، خطوط یا میله‌ها) مصورسازی داده نامیده می‌شود. در مصورسازی داده هدف این است که اطلاعات را به روشنی و با میزان اثرگذاری چشمگیر، به کاربران ارائه کنیم.

، مصورسازی داده و ویژگی‌ها آن اینطور (Vitaly Friedman) «محسوب می‌شود. «ویتالی فریدمن (Data Science) «مصورسازی داده اولین گام در مراحل تجزیه و تحلیل داده‌ها در «علم داده توصیف می‌کند:

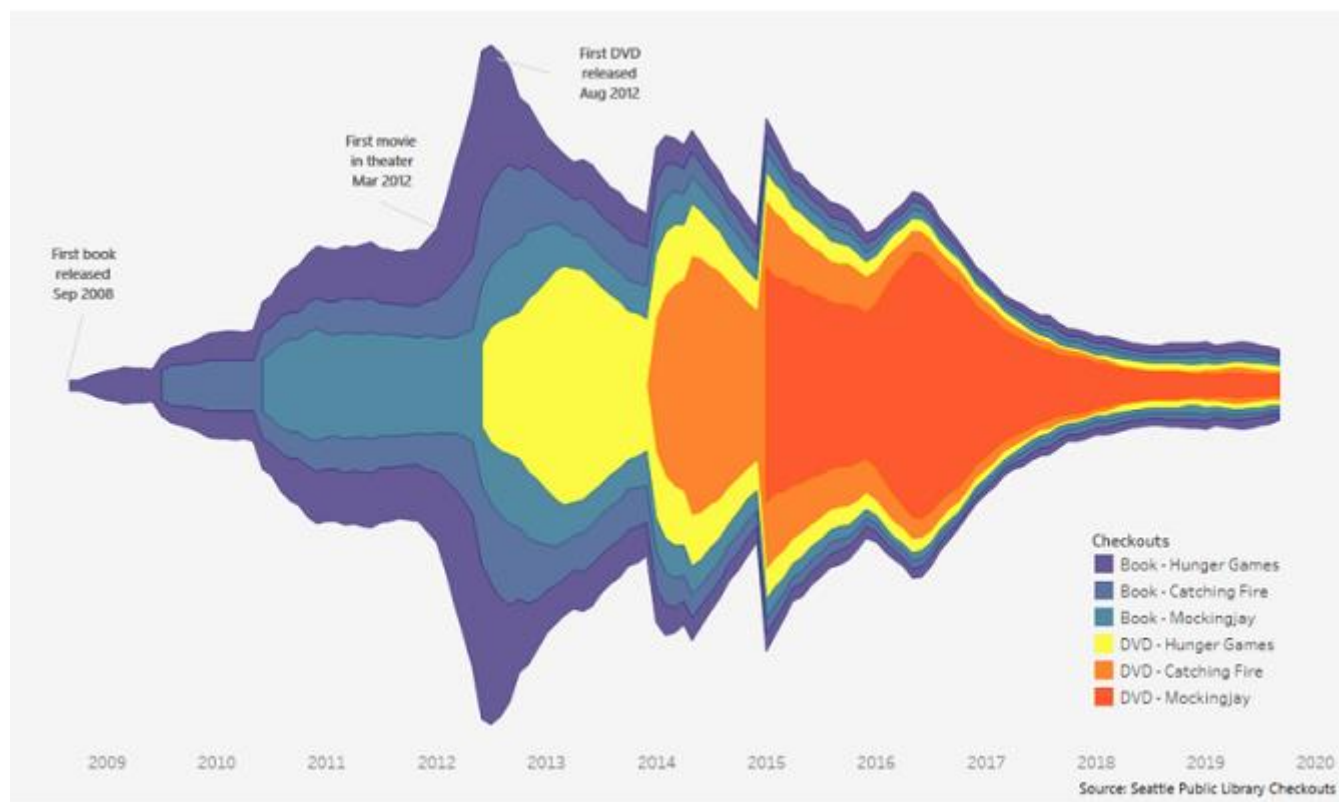
هدف اصلی مصورسازی داده یا تجسم آن‌ها، برقراری ارتباط واضح و موثر از طریق ابزارهای گرافیکی است. این بدان معنا نیست که لزوماً مصورسازی داده باعث ایجاد یک تصویر زیبا شود بلکه «درک اطلاعات به شیوه ساده و راحت منظور این روش توصیفی محسوب می‌شود. به طور مؤثر، هم فرم زیبا شناختی و هم عملکرد باید دست به دست هم دهند و با برقراری ارتباط با جنبه‌های اصلی آن به روشی بصری، اطلاعات نهفته در داده‌های نسبتاً پراکنده و پیچیده را ارائه دهند. مصورسازی داده که بدون هدف تولید شده و فقط جاذبه‌های بصری داشته باشند، منظور نظر مصورسازی داده نخواهد «بود»

اظهار می‌دارند که مصورسازی داده در حالت ایده‌آل نه تنها باید به طور واضح با مخاطب و (Martin M. Wattenberg) «و» «مارتین واتنبرگ (Fernanda Viegas) «در واقع، «فرناندا ویگاس کاربر ارتباط برقرار کند، بلکه باید توجه و حساسیت بیننده را هم بر بیانگیزند.

، (Scientific Visualization) «تجسم علمی»، (Information Visualization) «تجسم اطلاعات»، (Information Graphics) «مصورسازی داده در ارتباط نزدیک با «گرافیک اطلاعات است (Statistical Graphics) «و» «نمودارهای آماری (Exploratory Data Analysis) «تجزیه و تحلیل داده‌های اکتشافی»

در قرن حاضر، با توجه به حجم بسیار زیاد اطلاعات و لزوم به شناخت رفتار آن‌ها در زمان کوتاه، مصورسازی داده به یک حوزه فعال تحقیق، تدریس و توسعه تبدیل شده است، بطوری که این تکنیک، تجسم علمی و اطلاعاتی را متحد کرده است.

در تصویر ۱، نمونه‌ای از شیوه موفق نمایش اطلاعات در مصورسازی داده را مشاهده می‌کنید. استفاده از رنگ، راهنمای نمودار و همچنین یادداشت‌گذاری در لابلای نمودار، توانسته روند تغییرات در طول سال‌ها را به خوبی نشان دهد.



تصویر ۱: نمایش اطلاعات کتابخانه و موضوعاتی که اعضای یک باشگاه بازی رایانه‌ای بیشتر به آن توجه دارند، برحسب سال

ویژگی‌های نمایش تصویری موثر

توضیح می‌دهد که کاربرانی که از روش‌های نمایش اطلاعات استفاده می‌کنند، وظایف تحلیلی خاصی مانند انجام مقایسه را رهبری می‌کنند. طراحی گرافیکی (Edward Tufte) «پروفسور» ادوارد توفت نشان دادند که عناصر گرافیکی (Robert McGill) «و» «رابرت مک گیل (William Cleveland)» اطلاعات باید از وظیفه این گروه که همان تحلیل داده‌ها است، پشتیبانی کند. البته «ویلیام کلیولند مختلف، کم و بیش این کار را انجام می‌دهند

و (Dot Plot) البته باید این موضوع را هم در نظر گرفت که میزان کارایی هر یک از عوامل یا عناصر گرافیکی با یکدیگر متفاوت است. به عنوان مثال برای داده‌های چند متغیره، نمودارهای نقطه عملکرد بهتری دارند (Pie Chart) از نمودارهای دایره‌ای (bar Chart) نمودارهای میله‌ای

، که در سال ۱۹۸۳ منتشر شد، اصولی را (In his ۱۹۸۳ book The Visual Display of Quantitative Information) «ادوارد توفت» در کتابی با عنوان «نمایش تصویری اطلاعات کمی» برای نمایشگرهای گرافیکی مؤثر در عبارت زیر تعریف می‌کند

«تعالی در گرافیک آماری، شامل ایده‌های پیچیده‌ای است که با وضوح، دقت و کارایی ارتباط برقرار می‌کند»

از نظر او ویژگی‌های زیر برای مصورسازی داده و ابزارهای آن باید در نظر گرفته شود

- نمایانگر داده‌ها باشد
- برای بیننده میزان تمرکز و درک را به حداکثر ممکن برساند و ذهن او را مشغول تکنولوژی رنگ، چارچوب‌های طراحی گزارش تصویری نکند
- شامل عناصر مزاحم در درک اطلاعات نباشد
- بیشترین حجم اطلاعات را در کمترین فضا نشان دهد
- جمع‌بندی حجم زیادی از اطلاعات را به همراه داشته باشد
- به بیننده، امکان مقایسه و تشخیص اختلافات را بدهد
- داده‌ها را در چند سطح (از جزئیات به کلیات) در یک نمای کلی و ساختار مناسب، نشان دهد

- هدفدار و معقول باشد. شامل شرح، جدول‌بندی و جلوه‌های بصری مناسب باشد.

- به‌طور کامل با آمار و توصیف شاخص‌های عددی، همگام باشد.

با این ویژگی‌ها، گرافیک در مصورسازی داده باید ماهیت اصلی اطلاعات را به ما نشان دهد. حتی می‌توان گفت که مصورسازی داده می‌تواند دقیق‌تر و شفاف‌تر از آمار، رفتار داده‌ها را برایمان بیان کند.

برای میزان تلفات ارتش ناپلئون در دوره ۱۸۱۲ تا ۱۸۱۳ و مربوط به جنگ جهانی اول، هنگامی که ارتش او به روسیه حمله کرد، (Minard) «به تصویر شماره ۲ دقت کنید. این نمودار توسط «مینارد را نشان می‌دهد. در این نمودار پنج متغیر هم‌زمان نشان داده شده‌اند

۱. تعداد هنگ‌های ارتش

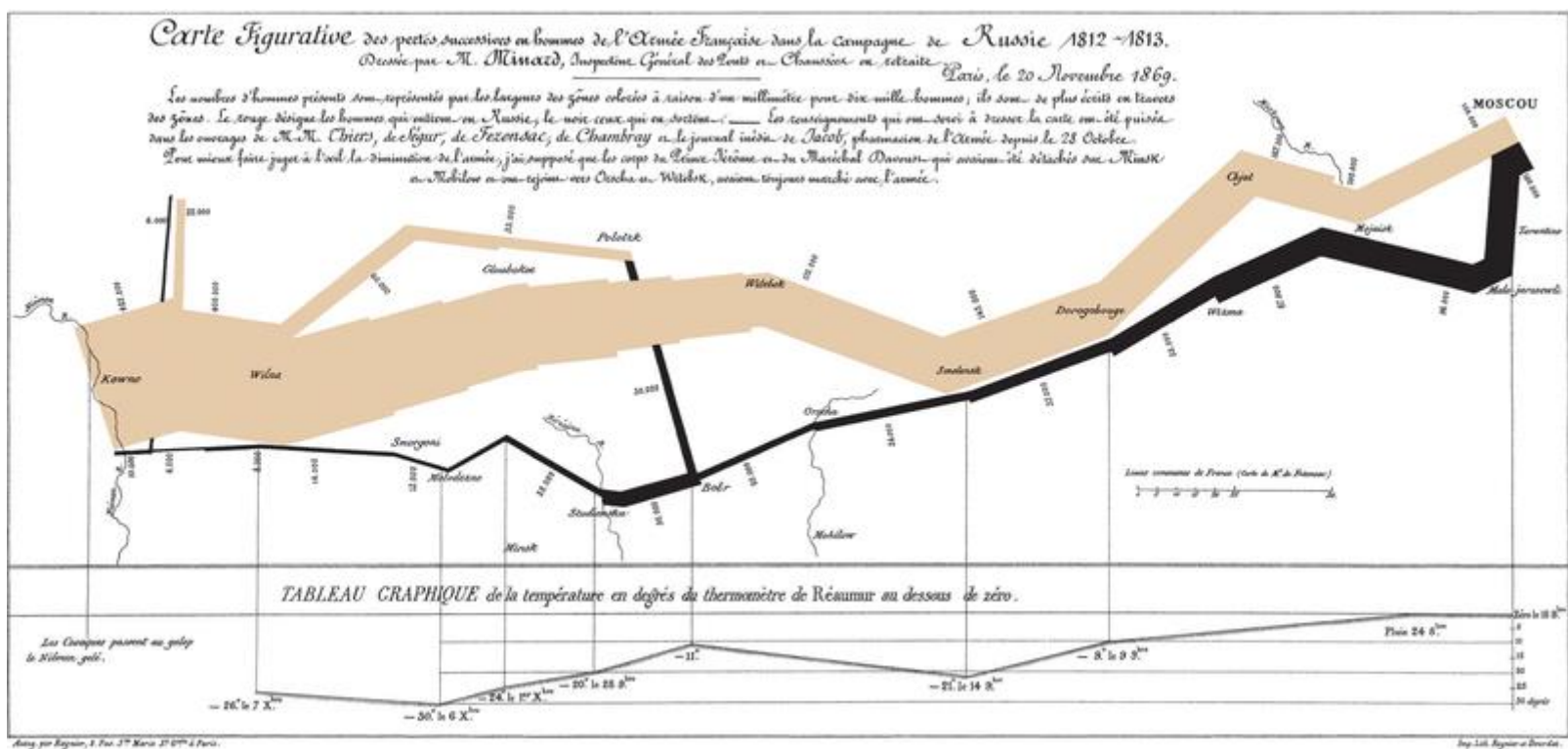
۲. (X,Y) متغیر دو بعدی مکان قرارگیری آن‌ها، شامل طول و عرض جغرافیایی

۳. زمان

۴. جهت حرکت

۵. درجه حرارت

پهنای خط، اندازه ارتش در نقاط جغرافیایی مختلف برحسب زمان را نشان می‌دهد. از طرفی خطوط پایینی نمودار، ارتباط میزان دما با تغییر اندازه ارتش را مشخص می‌کند. این نمایشگر چند متغیره بر روی یک سطح دو بعدی داستانی را روایت می‌کند که هنگام شناسایی داده‌ها توسط راهنما و توضیحات همراه نمودار، می‌تواند فوراً قابل درک باشد. «توفت» این نموداری را به این شکل توصیف می‌کند: “این نمودار را می‌توان بهترین گرافیک آماری دانست که تاکنون ترسیم شده است



تصویر ۲: نمایش اطلاعاتی از حمله ناپلئون به روسیه در جنگ جهانی اول

عدم استفاده از اصول ذکر شده در بالا، ممکن است کاربران را از واقعیت گمراه کند و باعث تحریف پیامی شود که داده‌ها به همراه دارند. این امر مدیران و رهبران جامعه را به سمت و سوی نتایج می‌نامند. معمولاً این گونه نمایش اطلاعات به همراه گرافیک عالی و رنگ‌های اغوا کننده بوده، بطوری که (Chartjunk) «نادرست سوق می‌دهد. چنین مصورسازی داده را گاهی «نمودارهای بی‌ثمر هدف اصلی را کنار زده و بیننده را محو تماشای تکنیک‌های ترسیمی می‌نمایند.

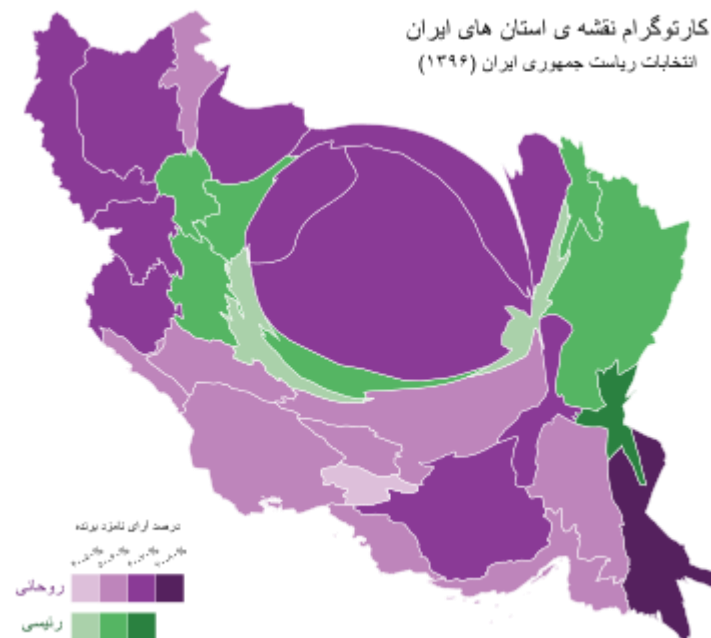
تفکیک راهنمای نمودارها از خود نمودار و ایجاد جلوه‌های تصویر بی‌ثمر فقط باعث گمراهی و صرف زمان بیشتر برای درک حجم کم اطلاعات خواهد شد. مشخص است که باید اطلاعات نمایش داده شده نسبت به فضای مصرفی برای نمایش بیشینه شده و از ارائه توضیحات و اطلاعات غیرضروری حذر داشت.

مصورسازی داده با نمایش اطلاعات عددی

به هشت رویکرد یا نوع اطلاعات عددی اشاره می‌کند که ممکن است کاربران را با مصورسازی داده مواجه کند. این موضوع نشانگر اهمیت ارتباط و تبادل اطلاعات بین (Stephen Few) «استفان فیو» ببیننده و تصویری است که مصورسازی داده ایجاد کرده است.

۱. برای مجموعه اطلاعاتی به کار می‌رود که وابسته به زمان بوده و بر اساس پدیده‌هایی متغیر برحسب زمان ایجاد می‌شوند. برای مثال می‌توان (Time Series) اصطلاحاً سری زمانی: **سری زمانی** ۱. تغییرات (Trend) بسیار موثر است. با این کار، روند (Line Chart) میزان یا نرخ بیکاری در طول یک دوره ۱۰ ساله در یک کشور خاص اشاره کرد. نمایش چنین اطلاعاتی توسط نمودار خطی چنین پدیده‌ای در طول زمان مشخص می‌شود.
۲. مرتب کرد. برچسبی که به این مقادیر مرتب شده، به صورت شماره سریال، داده می‌شود، رتبه (Descending) یا نزولی (Ascending-) می‌توان مقادیر عددی را برحسب ترتیب صعودی: **رتبه‌بندی** ۲. گفته می‌شود. البته دسته‌ها یا گروه‌هایی که از طبقه‌بندی داده‌های کیفی (ترتیبی) ساخته می‌شوند نیز دارای رتبه خواهند بود. برای مثال برحسب شعبه‌های مختلف یک فروشگاه زنجیره‌ای (Rank) نمایش داد (Bar Chart) یا نمودار میله‌ای (Column Chart) (متغیر طبقه‌ای) می‌توان رتبه فروش را محاسبه کرد و در یک نمودار ستونی ساخته می‌شوند. به این ترتیب داده‌های دو بعدی را می‌توان به صورت نسبت در آورده و در یک نمودار (Part-to-Whole) مقادیر نسبت یا نسبی از حاصل تقسیم یک جزء به کل: **جزء نسبت به کل** ۳. نمایش داد. برای مثال نمایش سهم از بازار در بین چند فروشگاه زنجیره‌ای بواسطه چنین نمودارهایی، به خوبی امکان مقایسه را به بیننده می‌دهد (Pie Chart) دایره‌ای گاهی لازم است که یک مقدار کمی را در بین چند گروه با یکدیگر مقایسه کنیم. بخصوص زمانی که هدف مقایسه گروه‌ها با یک استاندارد یا گروه مرجع مطرح باشد. برای مثال: **نمایش اختلاف‌ها** ۴. برای مقایسه و نمایش اختلافات بودجه محقق شده و برنامه‌ریزی شده در حوزه‌های مختلف صنعت، خدمات و بانک‌داری، بسیار کارآمد عمل می‌کند (Bar Chart) نمودار میله‌ای معروف است، مقادیر مختلف عددی را به صورت دسته یا گروه‌هایی که شامل فاصله‌های عددی (Histogram) که اغلب به هیستوگرام (Frequency Chart) یک نمودار فراوانی: **توزیع فراوانی** ۵. مشخص و طبقه‌بندی کرده و تعداد اعضای گروه‌ها یا فراوانی آن‌ها را شمارش می‌کند. شکل این نمودار به صورت مستطیل‌هایی با ضخامت فاصله‌های عددی و ارتفاع برحسب فراوانی ترسیم می‌شود. که (Boxplot) «نمودار جعبه‌ای» ایجاد می‌شوند. به این ترتیب گروه‌ها از لحاظ اعضایشان قابل مقایسه می‌شوند. همچنین (Frequency Table) «جدول فراوانی» معمولاً این نمودارهای براساس ابداع شده، روشی دیگری برای نمایش توزیع فراوانی داده‌ها است. این نمودار، شاخص‌های دیگری مانند میانه و چارک‌ها را هم نشان می‌دهد. (John Tukey) «توسط آمار شناس بزرگ» «جان توکی از کاربردهای مهم این نمودار محسوب می‌شود. او اعتقاد دارد یک نمودار باید آن چیزی که شما با نگاه کردن به داده‌ها متوجه نمی‌شوید را در دیدتان قرار داده و شما را (Outlier) **نقاط پرت** نمایش متعجب سازد.

۶. بین دو متغیر به خوبی **(Correlation) همبستگی** است. معمولاً در چنین نموداری، (Scatter/Dot Plot) یکی از روش‌های نمایش وابستگی بین دو متغیر، رسم نمودارهای پراکندگی: **همبستگی** و متغیر وابسته در (X) دیده می‌شود. مشخص است که در صورت همبستگی شدید بین دو متغیر، با دانستن مقدار یکی از آن‌ها، دیگری را می‌توان پیش‌بینی کرد. معمولاً متغیر مستقل در محور افقی و در چنین نمودارهایی به خوبی، ظاهر خواهند شد **(Regression Model) رابطه رگرسیونی** یا **(Functions) توابع ریاضی**. مشخص می‌شود (Y) محور عمودی ایجاد شده‌اند، هدف محقق است. از آنجایی که این گونه مقادیر دارای (Nominal Measure) گاهی مقایسه بین دسته‌های مختلف که بوسیله داده‌هایی با مقیاس اسمی: **مقایسه‌های متغیرهای اسمی** ۷. ترتیبی نیستند، نمی‌توان برای آن‌ها نمودارهای یک شکلی در نظر گرفت. ولی اغلب براساس بیشترین فراوانی می‌توان گروه‌ها را مرتب و در یک نمودار ستونی یا میله‌ای نمایش داد. برای یکی از این گونه روش‌ها در مصورسازی داده محسوب می‌شود **(Pareto Chart) «نمودار پارتو»** مثال
۸. امروزه نمایش گزارشات مرتبط با حوزه‌های جغرافیایی مختلف در نمودارهایی که شامل نقشه یک ناحیه است، تکمیل می‌شوند. به این ترتیب علاوه بر نمایش: **نمودارهای جغرافیایی یا نقشه‌ای** یک نمونه از این گونه نمودارها است (Cartogram) شاخص‌هایی مورد نظر، نواحی و مختصات طول و عرض جغرافیایی نیز به همراه نمودار دیده می‌شود. نمودار نقشه آماری



(Cartogram) تصویر ۳- نمونه یک نقشه آماری

تحلیلگرانی که مجموعه‌ای از داده‌ها را بررسی می‌کنند، ممکن است بیش از یک نمودار را به منظور مصورسازی داده به کار گیرند. به این ترتیب انتخاب بهترین شیوه یا روش برای نمایش داده‌ها، یک هنر محسوب شده که البته تکنیک‌های فنی نیز در آن نقش خواهد داشت. به همین علت مصورسازی داده را یک هنر و علم قلمداد کرده‌اند.

درک تصویری و مصورسازی داده

ذهن انسان می‌تواند تفاوت در طول خط، شکل، جهت‌گیری و رنگ (شدت رنگ) را به راحتی تشخیص دهد، بدون آنکه تلاش قابل توجهی برای پردازش آن‌ها صورت گیرد. این ویژگی‌ها به «خصوصیات معروف هستند. به عنوان مثال، برای شناسایی تعداد دفعات تکرار رقم «۵» در یک سری اعداد، به زمان و تلاش یا «پردازش توجه» زیادی نیاز است، اما (Pre-attentive Attributes) «پیش از توجه». اگر این رقم در اندازه، جهت یا رنگ با بقیه اعداد متفاوت باشد، تشخیص آن با سرعت توسط ذهن صورت خواهد گرفت.



مصورسازی داده به شرطی موثر و مورد توجه است که از خاصیت و ویژگی‌های «قبل از توجه» بیشترین بهره را برده باشد. برای مثال، از آنجا که انسان‌ها راحت‌تر می‌توانند اختلاف در طول خط را نسبت به اختلاف در سطح، پردازش کنند، ممکن است استفاده از نمودار میله‌ای (که از طول خط برای نشان دادن مقادیر استفاده می‌کند) مؤثرتر از نمودار دایره‌ای (که برای نشان دادن مقایسه از سطح استفاده می‌کند) در بیان داده‌ها عمل کرده و با سرعت بیشتری اطلاعات را منتقل کند.

ادراک، شناخت و تجسم در انسان

می‌توان گفت که تقریباً همه روش‌های تجسمی برای نمایش داده‌ها به مصرف انسانی می‌رسند (مگر آنکه حیوانات و حتی ربات‌ها را هم وادار به دیدن و استنباط از روی نمودار بکنیم). به همین علت آگاهی از نحوه ادراک و شناخت انسان، هنگام طراحی مصورسازی داده ضروری است.

«شکل‌گیری مفهوم»، «Thought» «اندیشه»، «Memory» «حافظه»، «Learning» «یادگیری»، «Attention» «توجه»، «Perception» «شناخت نسبت به فرآیندهای یادگیری در انسان، مانند «ادراک ، کلیدهای اصلی برای انتقال اطلاعات به شکل‌های مختلف مانند متنی، تصویری یا صوتی، هستند. پردازش (Problem Solving) «و «حل مسئله» (Reading) «خواندن»، (Concept Formation) بینایی انسان در تشخیص تغییرات و انجام مقایسه بین مقادیر، اندازه، شکل نسبت به بقیه حس‌ها، بسیار کارآمد است

هنگامی که متغیرهای عددی به صورت نمادین به خصوصیات بصری تبدیل می‌شوند، انسان می‌تواند حجم زیادی از داده‌ها را در زمان کوتاه پردازش کند. به همین علت کارآمدی ابزار مصورسازی داده به شکلی قابل قبول مورد توجه است

تخمین زده می‌شود که حدود ۶۰ درصد از نورون‌های مغز می‌توانند در پردازش بینایی نقش داشته باشند. تجسم خلاق، رویکرد متفاوتی را برای نشان دادن ارتباطات و روابط بین مقادیر به صورت بلقوه به انسان می‌دهد که بدون استفاده از ابزارهای بصری امکان ایجاد آن وجود ندارد. به همین دلیل تجسم می‌تواند به ابزاری برای اکتشاف شهودی و تحلیل داده تبدیل شود

تاریخچه مصورسازی داده

هیچ تاریخچه‌ای از شیوه یا روش‌های مصورسازی داده وجود ندارد زیرا کل پیشرفت تفکر بصری و بازنمایی بصری داده‌ها به صورت تدریجی و با رشد و بلوغ قدرت‌های ذهنی بشر، ایجاد شده است. خطوط تصویری، نقاشی‌های درون غارها، همگی روش‌هایی برای انتقال اطلاعات از فردی به فرد دیگر یا نسلی به نسلی دیگر محسوب می‌شوند

برخلاف تصور عمومی، مصورسازی داده یک پیشرفت مدرن و امروزی نیست. تصاویر مربوط به اجرام سماوی و ستاره‌ها روی دیوارهای غارها (مانند اطلاعات موجود در غار لاساکس در جنوب فرانسه) بیانگر قدمت مصورسازی داده است. آثار باستانی قابل لمس همچون لوح‌های گلی و سفالی بین النهرین (۵۵۰۰ سال قبل از میلاد)، میله‌های باستانی اینکاها (۲۶۰۰ قبل از میلاد مسیحی) و نمودارهای چوب در جزایر مارشال نیز می‌توانند به عنوان تجسم و تصویرسازی از اطلاعات کمی در نظر گرفته شوند

که به طور دقیق، توزیع منابع زیرمینی را نشان می‌دهد، یکی از (Turin Papyrus Map) «اولین تجسم داده‌های مستند شده را می‌توان در قرن ۱۱۶۰ قبل از میلاد مشاهده کرد. «نقشه تورین پاپیروس این گونه مستندات است. این نقشه‌ها را می‌توان به عنوان جد علم نقشه‌برداری در نظر گرفت. نقشه تورین پاپیروس نوعی تجسم داده است که اطلاعات خاص را از طریق یک تصویر جغرافیایی طراحی شده، با یک موضوع خاص در یک منطقه جغرافیایی مشخص، مرتبط می‌کند.

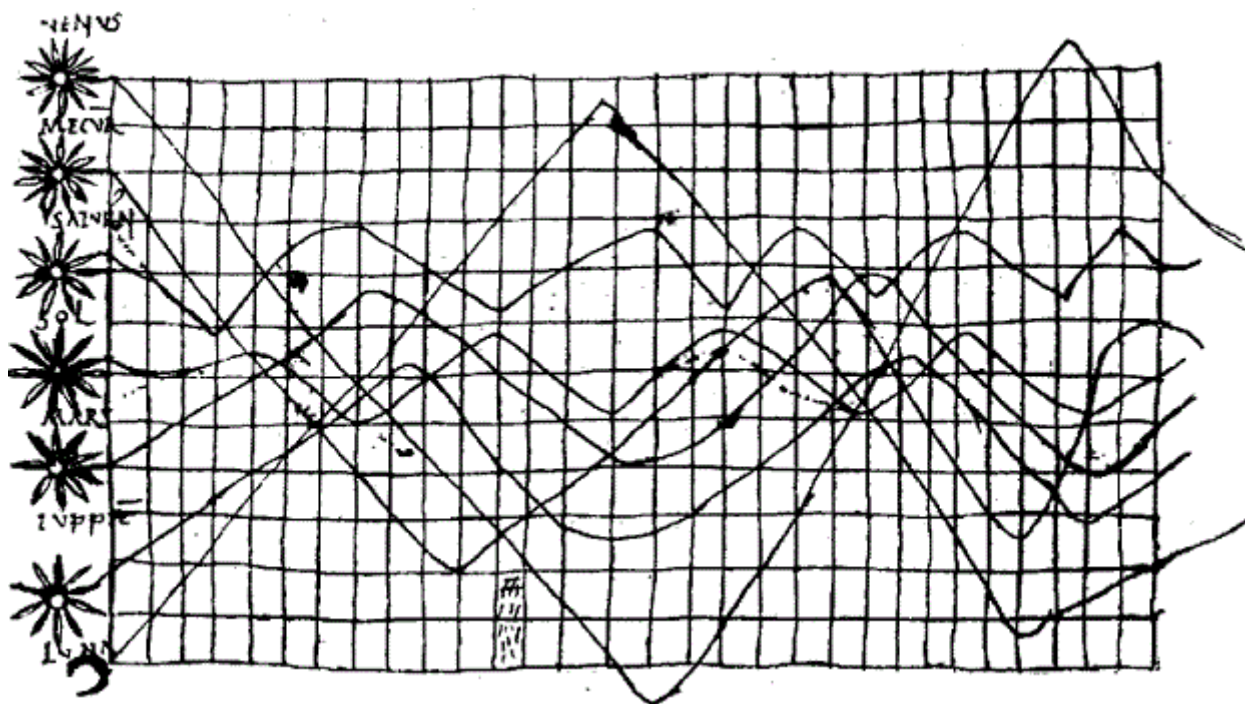


تصویر ۴: نقشه تورین روی پاپیروس، نمایش منابع و بهره‌گیری از آن‌ها توسط مصریان

، (Linear B Table) از یونان باستان B نخستین اشکال مستندسازی تجسم نقشه‌ها، موضوعی متشکل از فرهنگ‌ها و ایدئولوژی‌ها و سلسله مراتب اجتماعی است. به عنوان مثال، جداول سنگی خطی تجسم اطلاعات مربوط به معاملات دوره عصر برنز در حوزه جغرافیایی مدیترانه را ارائه می‌دهد. همچنین ایده به کارگیری از مختصات توسط نقشه برداران مصری باستان در شهرسازی مورد استفاده قرار گرفت. موقعیت‌ها و مختصات خاکی و آسمانی، توسط چیزی شبیه به عرض و طول جغرافیایی حداقل تا سال‌ها قبل از میلاد مسیح، مطرح بود. تشکیل نقشه یک زمین کروی با عرض و طول در اسکندریه ابداع شده بود تا قرن ۱۴ میلادی به عنوان معیارهای اصلی و مرجع به کار می‌رفت (Claudius Ptolemy) «جغرافیایی که توسط «کلودوس بطلمیوس

اما با اختراع کاغذ همه چیز متحول و با سرعت بیشتری رشد کرد. کاغذ، امکان توسعه بیشتر تجسم و تصویر سازی را در طول تاریخ فراهم کرد. اشکال و نمودارها از قرن دهم یا احتمالاً یازدهم برای نمایش تصویری حرکت سیاره‌ها در پیوست کتاب‌های درسی در مدارس صومعه مورد استفاده قرار می‌گرفت.

ظاهراً چنین ترسیم‌هایی، برای نشان دادن نقشه‌ای از مدارهای سیاره‌ای به عنوان تابعی از زمان به کار می‌رفته‌اند. برای مثال در علم هیئت و اجرام سماوی، ناحیه منطقه‌البروج را در یک صفحه با یک خط افقی تقسیم شده به سی قسمت به عنوان محور زمان یا طولی نمایش می‌دادند. محور عمودی نیز در این نقشه‌ها، پهنای منطقه‌البروج را تعیین می‌کند. به نظر می‌رسد مقیاس افقی برای هر سیاره به صورت جداگانه انتخاب شده. متن همراه این گونه نقشه‌ها فقط به مقادیر اشاره دارد و منحنی‌ها ظاهراً به زمان ارتباطی ندارند. نمونه‌ای از این نقشه را در تصویر ۵، مشاهده می‌کنید.



تصویر ۵: نمایش حرکت سیارات و اجرام سماوی

تا قرن شانزدهم میلادی، تکنیک‌ها و ابزارهای دقیق برای مشاهده و اندازه‌گیری مقادیر اجسام و موقعیت‌های جغرافیایی و سماوی به خوبی توسعه یافته بودند. این ابزارها با کمک علم مثلثات و سایر روش‌های محاسباتی، قادر به تعیین دقیق مکان‌های نقشه‌برداری شده بودند.

، هندسه تحلیلی و سیستم مختصات دو بعدی را توسعه دادند که به طور کامل بر روش‌های (Pierre de Fermat) «و» «پیر د فرما» (René Descartes) «فیلسوف و ریاضیدان فرانسوی» «رنه دکارت» و ابزارهای به کارگیری آن شد (Data) در مورد آمار و نظریه احتمال، زمینه ساز مفهوم داده (Blaise Pascal) «علمی و عملی نمایش و محاسبه مقادیر تأثیر گذاشت. کار «فرما» و «بلیز پاسکال»

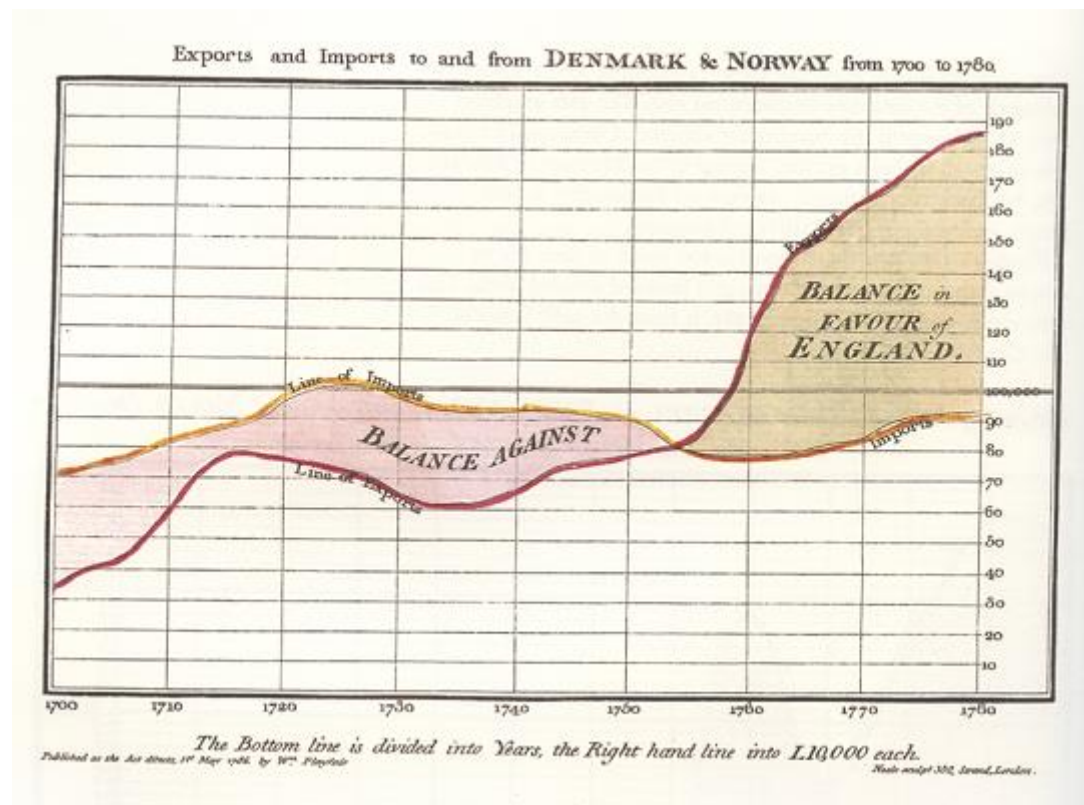


تصویر ۶: رنه دکارت فیلسوف و دانشمند فرانسوی



تصویر ۷: فرما، دانشمند و ریاضیدان فرانسوی

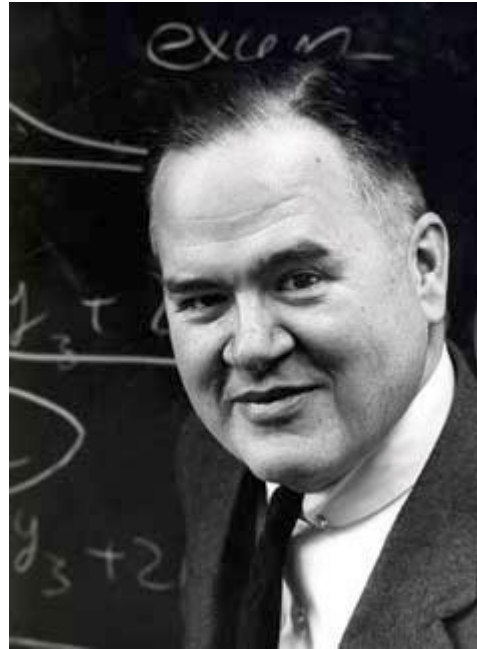
کمک کرد که روش‌های گرافیکی آماری را برای نمایش ارتباط داده‌های کمی ایجاد کند. در تصویر ۸، نمونه‌ای از نمودار سری زمانی ترسیم شده (William Playfair) «این تحولات به «ویلیام پلیفر توسط او دیده می‌شود.



تصویر ۸: نمایش داده‌های سری زمانی مربوط به واردات و صادرات سال‌های ۱۷۰۰ تا ۱۷۸۰، ترسیم شده توسط پلیفیر

جان توکی» و «ادوارد توفت»، مرزهای مصورسازی داده را به جلو بردند. «توکی» با رویکرد جدید و با نگرش آماری خود، در تجزیه و تحلیل داده‌های اکتشافی گام‌های موثری برداشت. همچنین «

، راه و روش‌هایی برای تصحیح تکنیک‌های مصورسازی داده‌ها، برای کسانی که با آمار آشنایی نداشتند، معرفی کرد «نمایش تصویری اطلاعات کمی» «توفت» با کتاب



، ریاضیدان و آمارشناس آمریکایی (John Tukey) تصویر ۹: جان توکی

با پیشرفت تکنولوژی، تکنیک‌های موثر و جدیدتری نیز در مصورسازی داده‌ها بوجود آمد. به این ترتیب استفاده از رایانه‌ها به جای استفاده از روش‌های ترسیم دستی، برنامه‌های کاربردی بسیاری را وارد بازار کار کرده بطوری که سرعت و دقت و همچنین تنوع نمودارها و روش‌های مصورسازی وارد مرزهای جدیدی گردید.

و موارد دیگر، امکان تجسم و مصورسازی داده‌ها در زمینه‌های مختلف بخصوص آمار را فراهم کرده‌اند. سایر برنامه‌های Cornerstone، Minitab، R، SOFA، SAS برنامه‌های کاربردی مانند هستند. بسیاری از رشته‌های دانشگاهی و حتی مدارس نیز دوره‌هایی برای آموزش مصورسازی داده JavaScript و Python، D3 مصورسازی داده بیشتر متمرکز بر زبان‌های برنامه‌نویسی مانند تدارک دیده‌اند.

یک برنامه سالانه در زمینه تجسم داده‌ها را به صورت تعاملی، اجرا کردند. این Pasadena در JPL و Caltech، ArtCenter با شروع سمپوزیوم «کشف داده» در سال ۲۰۱۳، کالج طراحی برنامه‌ها سعی دارند به پرسش‌هایی در زمینه مصورسازی داده پاسخ دهند. بعضی از سوالاتی که هدف اجرای چنین برنامه‌ای است، در ادامه به صورت فهرست‌وارد ذکر شده‌اند.

• چگونه تجسم تعاملی داده‌ها، می‌تواند به دانشمندان و مهندسين کمک کند تا رفتار داده‌های خود را به طور مؤثرتر کشف کنند؟

• چگونه محاسبات و تفکر طراحی می‌تواند به نتیجه بخش بودن نتایج تحقیقات کمک کند؟

• برای کسب دانش در این زمینه چه روش‌هایی مؤثر است؟

این برنامه آموزش و کارگاه‌های مربوطه، با کدگذاری اطلاعات رابطه‌ای با خصوصیات دیداری و تعاملی مناسب، برای به دست آوردن بینش جدید به داده‌ها، رویکردهای بین رشته‌ای جدیدی را برای حل مشکلات علمی پیچیده، طراحی کرده است.

واژه‌شناسی

تجسم داده‌ها شامل اصطلاحات خاصی است که برخی از آنها از علوم آماری گرفته شده است. به عنوان مثال، مقیاس‌های اندازه‌گیری که در آمار تعریف می‌شوند در اینجا هم به کار برده خواهند شد. این مقیاس‌ها و واژه‌های مرتبط به منظور هماهنگی با اصطلاحات آماری به کار می‌روند. در ادامه اصطلاحات اصلی که مربوط به نوع مقادیر و نحوه نمایش آن‌ها است، ارائه می‌شود.

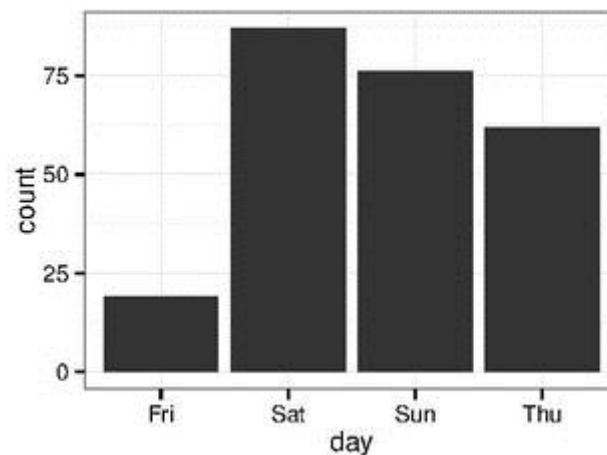
- برچسب‌های متنی که ماهیت اطلاعاتی دارند، مانند اسم، جنس یا رده شغلی، از نوع مقادیر طبقه‌ای هستند. به کمک این گونه مقادیر، می‌توانیم جامعه را دسته‌بندی یا گروه‌بندی **(Categorical) طبقه‌ای** کنیم. این داده‌ها معمولاً به صورت غیر عددی بوده و به کمک کدگذاری، به اعداد تبدیل می‌شوند. باید توجه داشت که این اعداد قابلیت انجام محاسبات ریاضی نداشته و نمی‌توان آن‌ها را مرتب کرد.
- مقدارهای عددی حاصل از اندازه‌گیری یا شمارش، در گروه داده‌های مقداری قرار می‌گیرند. برای مثال دستمزدها، میزان جمعیت و فشار خون بیماران از نوع داده‌های مقداری **(Quantitative) مقداری** هستند.
- جدول حاوی داده‌های کمی و عددی است که بر اساس نظم خاصی در ردیف‌ها و ستون‌های قرار گرفته‌اند. ردیف‌ها یا ستون‌ها به کمک برچسب‌هایی طبقه‌ها را مشخص می‌کنند و محل **(Table) جدول** تقاطع سطرها با ستون‌ها شاخص‌های عددی مانند فراوانی یا میانگین را برای گروه‌ها، نمایش می‌دهند. در درجه اول از جدول برای جستجوی مقادیر خاص استفاده می‌شود. جدول ممکن است دارای برچسب‌های ستونی دسته‌ای باشد که نام متغیرهای کمی (مثل سن و وزن) بوده و سطرها نیز بیانگر متغیر کیفی مثلاً نام افراد باشد. مقادیر درون خانه‌های جدول نیز سن و وزن هر یک از افراد را نشان می‌دهد.

- از نمودارها معمولاً برای نمایش روابط بین داده‌ها و مقادیر متغیرهای طبقه‌ای استفاده می‌شود. این شیوه‌های نمایشی با تکیه بر ویژگی‌های بصری انسان مثلاً تشخیص خطوط، ارتفاع (Graph) نمودار میله‌ها یا فاصله نقاط از یکدیگر ایجاد می‌شوند. مقادیر عددی در ناحیه‌ای مشخص می‌شوند که توسط یک یا چند محور محصور شده‌اند. این محورها مقیاس‌های (کمی و طبقه‌ای) را برای برچسب زدن و اختصاص مقادیر به اشیاء بصری فراهم می‌کنند. به این نمودار گاهی نگاره نیز گفته می‌شود.

نمونه‌ای از نمودارهای مورد استفاده در مصورسازی داده

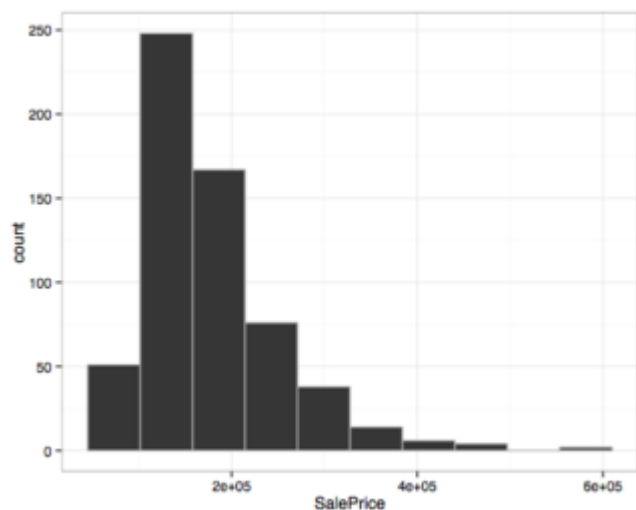
در ادامه شکل ظاهری بعضی از نمودارهایی که بخصوص در مصورسازی داده به کار می‌روند، اشاره شده است. واضح است که هر یک از آن‌ها، برای بیان ویژگی خاصی به کار می‌روند.

- مقایسه تعداد نامه‌های الکترونیکی ارسال شده (محور عمودی – داده کمی) برحسب روزهای هفته (محور افقی – داده طبقه‌ای): نمودار میله‌ای



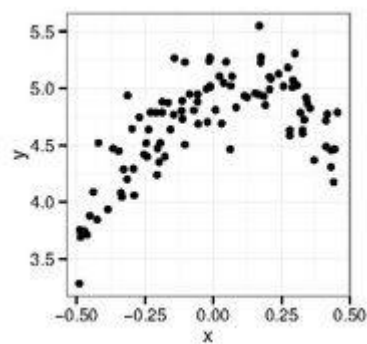
تصویر ۱۰: یک نمونه از نمودار میله‌ای یا ستونی

- این نمودار، فراوانی گروه‌های مختلف را نمایش می‌دهد. برای مثال نمودار فراوانی قیمت خانه‌ها، شامل فراوانی (محور عمودی- شمارشی) و محور افقی (Histogram) – نمودار فراوانی (هیستوگرام دسته یا گروه‌های قیمتی (کیفی- ترتیبی) است.



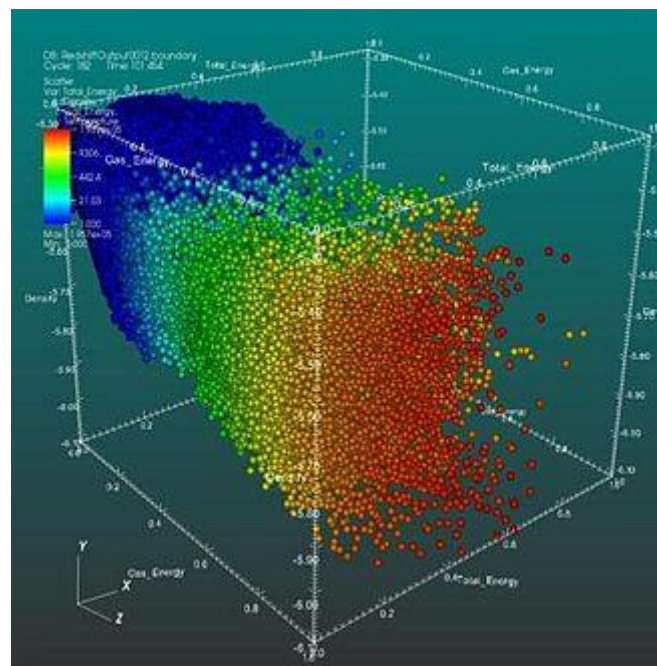
(Histogram) تصویر ۱۱: یک نمونه از نمودار فراوانی یا هیستوگرام

نمایش ارتباط بین دو متغیر کمی (عددی) در این نمودار صورت می‌گیرد. هر دو محور افقی و عمودی عددی بوده و مدل رابطه (همبستگی) بین دو متغیر نمایش داده: **نمودار پراکندگی (Scatter Plot)** نمودار پراکندگی می‌شود.



تصویر ۱۱: یک نمونه از نمودار نقطه‌ای برای نمایش رابطه بین دو متغیر عددی

توسط این گونه نمودارها ترسیم می‌شود $z=f(x,y)$ هر سه محور به صورت عددی بوده و برای نمایش توابع دو متغیره مناسب است. مثلاً: **نمودار پراکندگی سه بُعدی (3D Scatter Plot)**



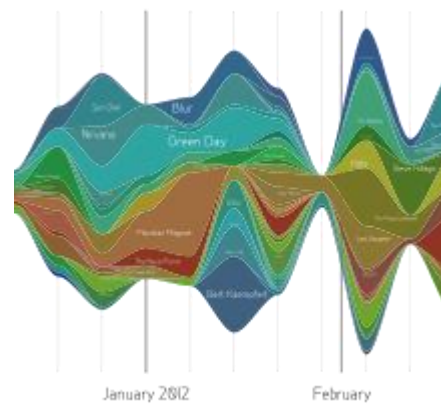
تصویر ۱۲: نمونه‌ای از نمودار نقطه‌ای سه بعدی و رنگی

در این گونه نمودارها، ارتباط بین نقطه‌ها مد نظر بوده و برای نمایش گراف‌ها و نمایش راس و اضلاع یک گراف یا جنگل، که از مفاهیم ریاضیات گسسته هستند، به کار **(Network)** نمودار شبکه‌ای می‌رود. با رشد شبکه‌های اجتماعی، استفاده از این گونه نمودارها گسترش یافته و به این ترتیب نحوه ارتباط بین افراد و اجتماعات انسانی و شبکه‌ها، به خوبی نمایش داده می‌شود.



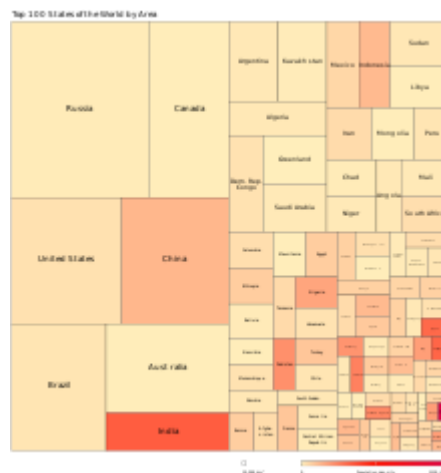
تصویر ۱۳: نمایش یک نمودار شبکه‌ای و ارتباط بین نقاط یک جنگل

در این نمودار روند تغییرات برای چندین متغیر (هم واحد) روی یک نمودار دو بُعدی ترسیم می‌شود. به همین دلیل استفاده از این نوع نمودار امکان مقایسه بین مقادیر **نمودار جریان (Streamgraph)** در طول زمان را میسر خواهد ساخت.



تصویر ۱۴: نمودار جریان برای نمایش روند تغییرات چند متغیره

این نمودارها برای نمایش دو ویژگی همزمان در بین چندین گروه مناسب است. البته تفاوت آن با نمودار دایره‌ای در تعداد زیاد گروه‌ها است که این نمودار را **(Treemap) نمودار نقشه درختی** بخصوص در نمایش نحوه عملکرد شرکت‌ها در بورس مورد پسند کاربران کرده است.



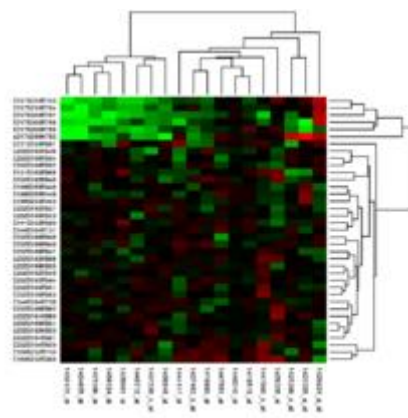
تصویر ۱۵: نمودار نقشه درختی برای نمایش جمعیت و مساحت کشورهای مختلف جهان

از نمودار گانت برای نمایش انجام فعالیت‌ها برحسب زمان صورت می‌گیرد. ارتباط بین فعالیت‌ها نیز در چنین نموداری ظاهر شده و اطلاعات زیادی در رابطه با تاریخ **(Gantt Chart) نمودار گانت** شروع، خاتمه و پیش‌نیازهای یک فعالیت در آن دیده می‌شود.



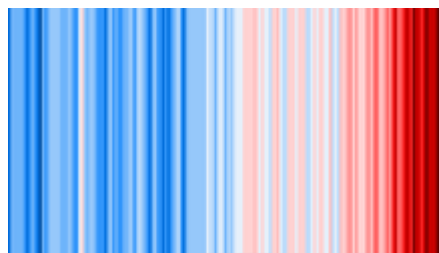
تصویر ۱۶: نمودار گانت برای نمایش رابطه بین فعالیت‌ها برحسب زمان

برای نمایش همزمان چند متغیر و رابطه بین گروه‌ها، این نمودار به خوبی عمل می‌کند. استفاده از این نمودار درست به مانند یک جدول توافقی است که می‌تواند بین دو یا **(Heat Map) نقشه گرمایی** چند گروه، یک یا چند شاخص محاسباتی را مقایسه کند.



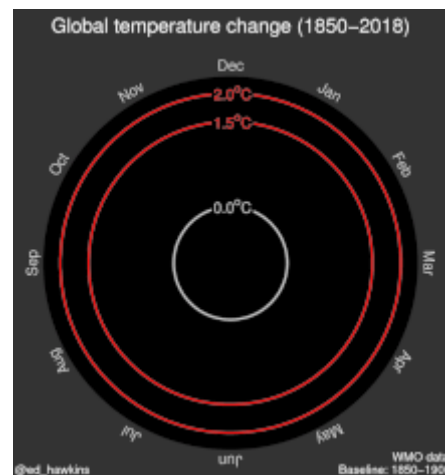
تصویر ۱۷: نمودار نقشه گرمایی برای نمایش ارتباط بین دو یا چند متغیره گروهی و شاخص‌های محاسباتی بین گروه‌ها

این نمودار دارای یک محور افقی است و مقادیر براساس اندیسی که روی محور افقی دارند، مرتب شده‌اند. رنگ‌ها نیز بیانگر شدت یا ضعف (کوچکی یا بزرگی) یک (Strip Graphic) گراف نواری ویژگی عددی هستند.



تصویر ۱۸: نمایش یک نمودار نواری برای مقایسه دمای یخ‌های قطبی نسبت به میانگین دوره

نمایش تغییرات یک پدیده در طول زمان در این نمودار توسط دایره‌های که مرکز آن مبدا یا مجموعه مرجع را نشان می‌دهد، توسط نمودار مارپیچی صورت (Spiral Graph) نمودار مارپیچی می‌گیرد.



تصویر ۱۹: نمودار متحرک مارپیچی برای نمایش تغییرات گرمایی کره زمین در سال‌های ۱۸۵۰ تا ۲۰۱۸

مقایسه ویژگی‌های نمودارهای مختلف در مصورسازی داده

در جدول زیر ویژگی‌های نمودارهای مطرح در مصورسازی داده، مانند کاربرد، متغیرهای مورد استفاده و مثال‌های واقعی با یکدیگر مقایسه شده است.

مثال کاربردی	ویژگی	متغیرها و ابعاد به کار رفته در نمودار	نام نمودار
مقایسه فروش به تفکیک شعب یک فروشگاه زنجیره‌ای با در نظر گرفتن شیفت صبح و شب	مقایسه بین شاخص‌هایی مانند میانگین یک اندازه یا شمارش اعضای یک دسته	اندازه / شمارش متغیر گروهی یا طبقه‌ای	(Bart / Column Chart) نمودار میله‌ای / ستونی
		رنگ برای تمایز بین دسته‌های متفاوت به عنوان متغیر دوم گروه	

بندی

مقایسه بین تعداد دختران و پسران دانش‌آموز در رده‌های سنی مختلف	مقایسه بین تعداد یا فراوانی بین گروه یا دسته‌ها	تعداد / متغیر شمارشی دسته یا گروه‌ها	نمودار فراوانی / هیستوگرام (Histogram)
		رنگ برای تمایز گروه‌ها	
نمایش رابطه طول قد و وزن افراد برحسب جنسیت (رنگ) و همچنین نژاد (نماد) با توجه به متوسط کالری (دریافتی) (اندازه)	به منظور نمایش رابطه بین دو متغیر وابسته (با در y مستقل) با x عددی نظر گرفتن متغیر سوم (اندازه) به عنوان متغیر عددی مستقل در دسته یا گروه‌هایی که توسط نماد یا رنگ مشخص شده‌اند	x محور y محور نماد یا نشانه‌گذاری نقاط اندازه (به صورت بزرگی یا قطر هر نقطه) (رنگ) برای نمایش دسته‌ها	(Scatter / Dot Plot) نمودار نقطه‌ای / پراکندگی
رسم توابع ریاضی دو یا چند متغیره در فضای سه بُعدی	به منظور نمایش رابطه بین دو متغیر با در نظر (x, y) عددی مستقل و در (z) گرفتن متغیر سوم وابسته دسته یا گروه‌هایی که توسط رنگ مشخص شده‌اند	x محور y محور z محور	نمودار نقطه‌ای سه بُعدی (3D Scatter Plot)

(رنگ) برای نمایش دسته‌ها

(Network Plot) نمودار شبکه

(گره یا راس‌ها) موقعیت

(رنگ گره‌ها) (گروه یا دسته

بزرگی یا کوچکی اندازه گره‌ها

اضلاع یا خطوط ارتباطی

مقدار عددی برای ضخامت خطوط

ارتباطی

دوری و نزدیکی گره‌ها

ارتباط بین متغیرهای گروهی براساس
شده رابطه یا بزرگ‌ری جمعیتی در هر
گروه

نمایش شبکه‌های اجتماعی و ارتباط
کاربران با یکدیگر یا نمایش تعداد
اعضای گروه‌های طرفدار یا
محبوب در شبکه‌های اجتماعی
ایجاد خوشه‌های همگن و هم شکل

تشخیص نقاط دور افتاده یا پرت که به
هیچ دسته یا گروهی تعلق ندارند

نمایش گروه یا افرادی که نقش
ارتباطی دارند

تعیین افراد یا گروه‌های تاثیرگذار در
شبکه‌های اجتماعی مناسب برای
شرکت‌ها و نهادهای خدماتی برای
تاثیر بیشتر روی جامعه هدف

(Streamgraph) نمودار جریان

(پهنای نوارها) متغیر عددی

پیگیری تغییرات یک پدیده براساس

میزان مرگ و میر (ارتفاع) براساس

عوامل مختلف سوانح رانندگی و مقایسه ماه‌ها (ضخامت) مشابه در سال‌های مختلف	چندین متغیر در بین چندین گروه	(رنگ) دسته‌ها
		(محور افقی) زمان یا جریان
نمایش میزان حجم، ارزش، تعداد سهام فروخته شده در بازار بورس	مقایسه بین چندین گروه براساس یک متغیر کمی	(اندازه) (متغیر کمی) (رنگ) (متغیر طبقه‌ای)
نمودار روند پیشرفت کار پیمانکاران و تحلیل عملیات توسط کارفرما	تحلیل روند فعالیت‌ها (تقدم و تاخیر) به همراه طول دوره فعالیت	(زمان) (محور افقی) (رنگ) (دسته یا فعالیت‌ها)
نمودار مقایسه روش‌های خوشه یا دسته‌بندی	تشخیص دسته‌های همسان و نمایش همسانی براساس رنگ‌ها	(سطر) (دسته‌ها) (ستون) (دسته‌ها)
		(خوشه‌ها) (گروه‌های هم شکل)
		(رنگ‌ها) (فاصله بین اعضای خوشه)
نمایش تغییرات دما نسبت به یک تاریخ (خاص) (مرجع)	مقایسه بین مقادیر مختلف و اختلاف آن‌ها از مرکز نمودار به عنوان مقدار مرجع	(محور افقی) (دسته‌ها) (رنگ‌ها) (شدت متغیر عددی)
Corona (میزان رشد ویروس کرونا در یک دوره زمانی Virus)		

نمایش میزان گرمایش زمین برحسب ماه و سال‌های مختلف رنگ‌ها نشانگر سال‌ها و قطعات نیز ماه‌ها را نشان می‌دهند.	نمایش روند تغییرات یک متغیر وابسته به زمان به تفکیک و تقسیم ماه یا دوره‌های تناوب مشخص در سال‌های گوناگون	(شعاع دایره‌ها (متغیر وابسته (زاویه چرخش (متغیر تناوب – ماه (رنگ (سال‌های مختلف	(Spiral Graph) نمودار مارپیچی
---	--	---	-------------------------------

قطر هر یک از دایره نیز میزان
گرما را به عنوان یک متغیر وابسته
به زمان نشان می‌دهد. (رجوع کنید
(به تصویر ۱۹

خلاصه و جمع‌بندی

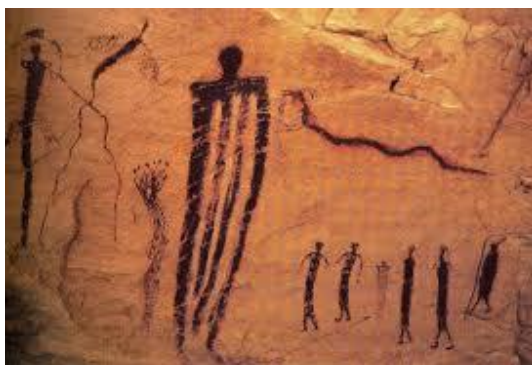
در این نوشتار با شیوه‌های مختلف مصورسازی داده آشنا شده و همچنین تاریخچه و ویژگی‌های آن را یادآوری کردیم. ابزارهای آماری و بهره‌گیری از ترسیم‌های آماری نیز در مصورسازی داده نقش مهمی دارند بطوری که شاید بتوان مصورسازی داده را برگرفته از آماری توصیفی دانست. معرفی نمودارهای مهم و عملکرد آن‌ها در مصورسازی داده نیز در این نوشتار نیز از مواردی بود که به آن پرداختیم.

نمودارها، منحنی ها، دیاگرام ها و بصری سازی



مقدمه:

از دیرباز استفاده از نمادها و اشکال گوناگون برای انتقال اطلاعات و دانش در میان انسان ها رواج داشته است. از نقش و نمادهای کنده کاری شده روی دیواره های غارها گرفته تا استفاده از نمادهای قبیله ای دلالت بر جایگاه تصاویر و دیدن در میان بشر دارد. بعضی از این تصاویر انصافاً شباهت زیادی به نمودارها دارند! خط (مانند خطی میخی) هم نوعی بصری سازی به شمار می رفت که طی سالیان دراز به شکل امروزی درآمده است.



در طول تمامی ادوار تاریخی، بصری سازی داده ها به علت محدود بودن خود دادگان، با محدودیت مواجه بود. به مرور با افزایش حجم اطلاعات، نیاز مردم به روش نمایش بهتر و تحلیل اطلاعات افزایش یافت. انتهای قرن ۱۸، Willialm Playfair تعداد زیادی از نمودارها و هیستوگرام ها را در کتاب خود معرفی نمود. او از این نمودارها برای نمایش داده های اقتصادی بهره جست. در آن زمان بیشتر داده ها در قالب جدول نشان داده می شدند اما Playfair آنها را به شکل اینفوگرافیک تبدیل کرد. از همان زمان، نمودارها و گراف ها به بینندگان برای اتخاذ تصمیمات مناسب تر بر اساس داده ها کمک نمودند. متخصصان آمار ابزارهای بیشتری برای نمایش بصری داده ها طراحی کردند و توسعه دادند و هم اکنون فناوری نوین روش ها و ابزارهای در اختیار ما قرار داده که تا به حال قابل تصور نبوده است. رواج استفاده از یک استاندارد جهانی برای نمایش مفاهیم، پیام ها، داده ها و اطلاعات به قرن هفدهم میلادی بر می گردد. در آن زمان نمودارهای دایره ای (شبیه شیرینی پای) توسعه پیدا کردند. پس از آن به مرور انواع نمودار طراحی و به کار گرفته شدند. برای درک کاربرد نمودارها و علت توجه زیاد به آنها کافی است به دنبال پاسخی برای این سؤال مهم گشت:

چه باید کرد اگر قرار باشد حجم عظیمی از اطلاعات در کمترین زمان ممکن در اختیار مخاطب قرار بگیرد؟

دیدن و تماشا کردن اشکال و نمادها، راحت تر و سریع تر از خواندن خط است. بنابراین مطمئناً ابزاری گرافیکی که بتواند اطلاعات را به گویاترین و کم حجم ترین شکل ارائه دهد کاربرد خواهد داشت. از این رو امروزه نمودارها جایگاه ویژه ای در گزارش دهی و اطلاع رسانی دارند.

نمودارها طوری طراحی شده و مورد استفاده قرار می گیرند که به راحتی درک مناسبی از داده و ارتباط میان آن ها در مخاطب ایجاد کنند. درک پراکندگی گونه های گیاهی، اطلاع از نقشه های هواشناسی، پیدا کردن مسیر روی نقشه مترو و پیش بینی قیمت کالاها در روزهای آتی با بررسی روند تغییر قیمت آن ها همگی اشاره به کاربردهای گوناگون ابزار گرافیکی دارند.



نمودار و بصری سازی کشورهای جهان

دسته بندی

در یک دسته بندی نمودارها را به سه دسته تقسیم می کنند. مبنای این دسته بندی هدف استفاده از نمودار است:

الف) نمودارهای مقایسه ای:

برای مقایسه بین دو متغیر (یا بیشتر) یا دو عنصر به کار می روند.

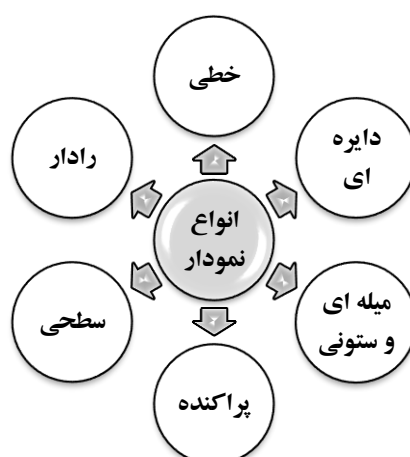
ب) نمودارهای پراکندگی:

پراکندگی داده ها را بر روی یک صفحه نمایش می دهند.

ج) نمودارهای رابطه ای:

رابطه ی متغیرها را با یکدیگر نشان می دهند.

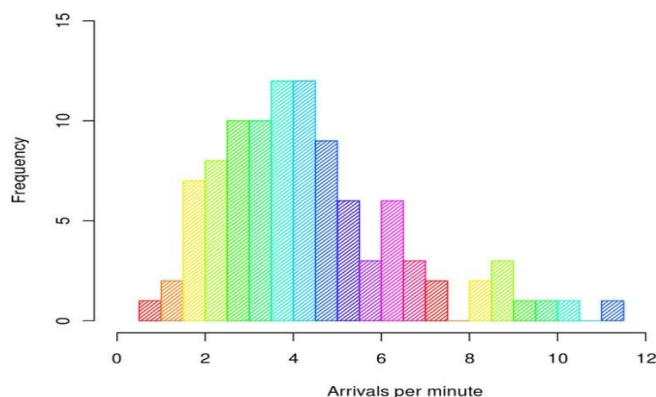
اما نمودارها را بر اساس ظاهر نیز دسته بندی بندی می کنند که عبارتند از:



نمودارها دارای تنوع عملکردی زیادی هستند که در ادامه به بررسی اغلب آنها پرداخته می شود.

۱- هیستوگرام (Histogram)

هیستوگرام یا بافت نگار، یکی از انواع نمودارها و برای نمایش گرافیکی توزیع داده های عددی می باشد. این نوع نمودار که توسط کارل پیرسن معرفی شده، تقریبی از توزیع احتمالاتی یک متغیر کمی و پیوسته است. هیستوگرام ها کاربردهای متنوعی دارند و در اکثر تحلیل های آماری نیز به کار می روند. یکی از جالب ترین کاربردهای آن ها در پردازش و تحلیل تصاویر است.



نمونه نمودار هیستوگرام

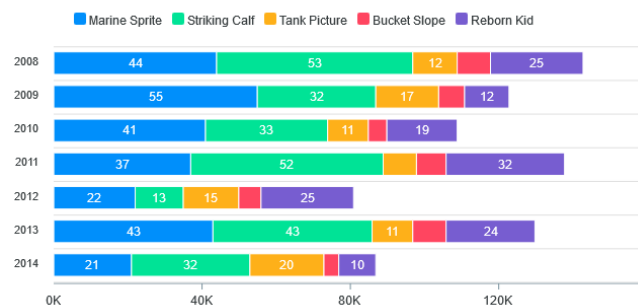
برای رسم این نمودار می بایست محدوده مقادیر متغیر مربوطه را به تعدادی بازه (به طور پیش فرض هم اندازه) تقسیم کرده، سپس فراوانی داده هایی که در محدوده هر بازه قرار می گیرند شمارش شود. ارتفاع هر ستون بیانگر فراوانی محاسبه شده است. بازه ها اغلب متوالی، مجاور و ناهمپوشان هستند. محور افقی در شکل فوق، مدت زمان انجام فرایند و محور عمودی تعداد فرایندهایی است که مدت زمان آن ها در بازه مورد نظر واقع شده است.

۲- نمودار میله ای (Bar chart)

نمودار میله ای، ساده ترین راه مقایسه دسته های متفاوت است. از نمودارهای میله ای برای نمایش و مقایسه فراوانی مقادیر، نسبت ها و درصدها استفاده می شود. این نوع نمایش امکان مقایسه گروه های مختلف داده و کلی گویی درباره داده را به دست می دهد. یک محور آن دسته هایی هستند که مقایسه می شوند و محور دیگر مقدار هریک را نمایش می دهد. این نمودار تقریباً برای نشان دادن هر نوع داده ای مناسب است.

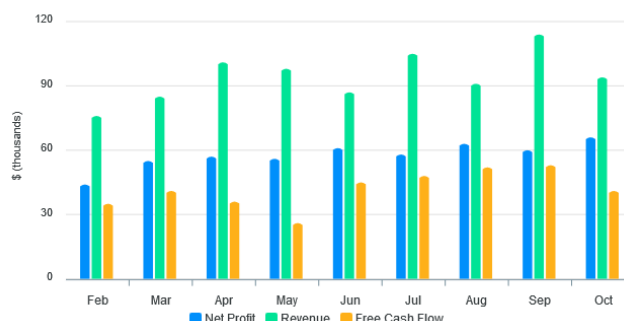
این نمودار، نموداری است که برای نمایش دادن داده های گروه بندی شده از میله های مستطیلی شکلی استفاده می کند که طول هر یک بیانگر مقدار متناسب با هر گروه داده است. بدین ترتیب شباهت زیادی بین این نوع از نمودارها با نمودارهای هیستوگرام وجود دارد. در این نمودار، میله ها می توانند مستطیلی، استوانه ای، هرمی یا مخروطی باشند.

Fiction Books Sales



نمونه نمودار میله ای افقی

نمودار میله ای می تواند عمودی هم ترسیم گردد که به آن نمودار میله ای ستونی می گویند. در شکل زیر مقادیر (درصد) مربوط به چهار گزینه مختلف در ارتباط با چهار کشور مختلف در قالب نمودار میله ای عمودی نشان داده شده است. این نمودار از آن جهت که در قبال هر یک از کشورها، مقادیر گروهی از اقلام مختلف را نشان می دهد گروهی نامیده می شود.



نمونه نمودار میله ای ستونی - گروهی

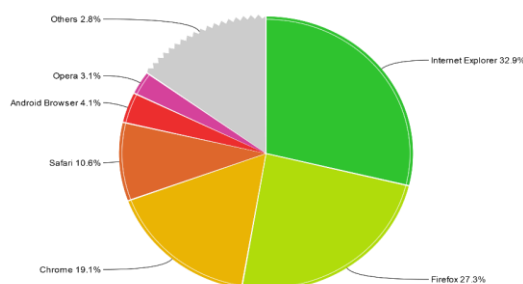
نمودارهای میله ای در رسم از ساده ترین انواع نمودارها هستند. برای رسم تنها کافی است مقادیر (چه مقدار واقعی، چه درصد و چه مقدار تجمعی) مربوط به هر یک از اقلام محاسبه شده و در قالب یک میله مشخص، در جای معین خود ترسیم شود. گاهی اوقات هنگام مطالعه و بررسی گروهی از افراد، مقایسه چندین متغیر در یک زمان متداول است. علاوه بر آن، بسیار کاربردی است که پیشینه های قومی و نژادی، سن و جنسیت را در برابر کل جمعیت مورد آزمایش قرار داد. در این موارد با تلفیق کاربرد دو نمودار ستونی و دایره ای، نسبت ها و روندها را در کنار هم بررسی می کند. نمودار حاصله، نمودار میله ای انباشته (Stacked Bar Graph) نام دارد. این نمودارها فقط برای نشان دادن بخش هایی از یک کل مفید نیستند. آنها همچنین می توانند برای نمایش متغیرهای اضافی استفاده شوند. در حالی که نمودار میله ای پایه می تواند نشان دهد که چه بخشی از یک جمعیت در یک دوره زمانی، به عنوان افرادی که اضافه وزن دارند طبقه بندی شده اند، با استفاده از یک نمودار میله ای جمع شده علاوه بر این مورد همچنین می توانید ببینید چه مقداری از کل آن افراد چاق، فربه هستند.



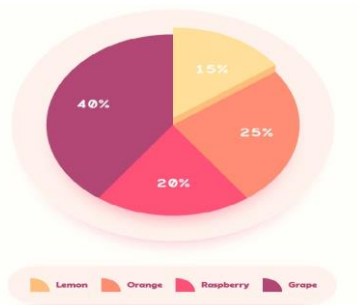
نمونه نمودار میله ای انباشته

۳- نمودار دایره ای (Circle chart / Pie chart)

نموداری بسیار کارآمد برای مقایسه بخش هایی از یک کل است. نمودار دایره ای، پای یا کیک، نمایشی از نسبت های عددی را در قالب برش هایی روی یک دایره ارائه می دهد. برای رسم این نمودار ابتدا می بایست مجموع مقادیر را به دست آورد، سپس سهم هر برش (نسبت مقدار متناظر با آن به مقدار کل) را حساب کرد. بر اساس همین نسبت ها، محیط دایره به قطعات مختلف تقسیم می شود که طول هر یک برابر نسبت مربوطه در محیط دایره است. نمودارهای جذاب دایره ای شکل انواع مختلفی دارد که نمودار کیک با قطعات جدا شده و نمودار چندسطحی (دوایر هم مرکز) از جمله آن ها هستند.



نمونه نمودار دایره ای



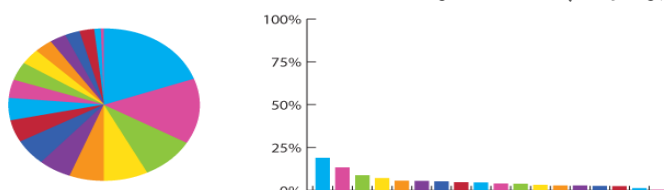
نمونه نمودار دایره ای با قطعات جدا شده

نمودار Sunburst، نوعی از نمودار دایره ای چند سطحی است و نمایانگر داده های سلسله مراتبی با استفاده از دایره های متحدالمرکز است. هر حلقه از این نمودار، یک سطح از سلسله مراتب را مشخص می کند. حلقه ها همچنین می توانند بیشتر به تقسیم بندی های چندگانه در یک سطح سازمانی جدا شوند.



نمونه نمودار Sunburst (نمودار دایره ای چند سطحی)

نقطه ضعف نمودار دایره ای این است که مقایسه مساحت برش ها یا مقایسه طول کمان ها در مواردی که تقریباً هم اندازه باشند بسیار دشوار و گمراه کننده است. از این رو غالباً نمودارهای دیگری نظیر نمودار میله ای را به آن ترجیح می دهند. در شکل زیر می توان کارایی نمودار میله ای را در برابر نمودار پای مقایسه کرد.

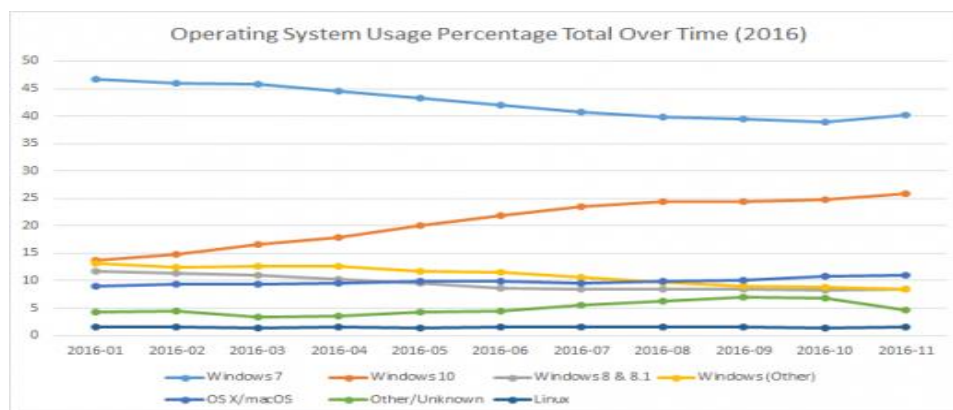


نمودار دایره ای و میله ای متناظر

۴- نمودار خطی (Line Graph)

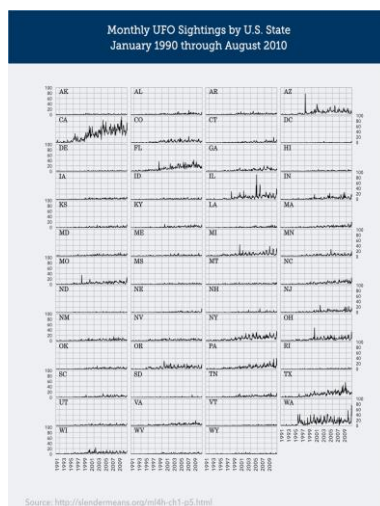
نمودار خطی گونه ای از نمودارها است که اطلاعات را در قالب دنباله ای از نقاط داده ای (نشانگر) که با خطوط شکسته به ترتیب به هم متصل شده اند نشان می دهد. این نمودارها ابزار بصری سازی قدرتمندی هستند که در ردیابی تغییرات در دوره های زمانی کوتاه یا بلندمدت و یا میزان همبستگی خاصی نشان می دهند. مثلاً یک محور ممکن است یک مقدار متغیر را نشان دهد در حالی که محور دیگر تعیین کننده بازه زمانی باشد. زمانی که تغییرات کوچکی وجود دارد، نمودارهای خطی بهتر از نمودارهای میله ای می توانند آن را نشان دهند. از این نمودارها همچنین برای مقایسه تغییرات در دوره های زمانی یکسان برای بیشتر از یک گروه داده می توان استفاده کرد. در این نمودار، ابتدا هر مقدار به طور مجزا مشخص می شود و در آخر نقاط به هم متصل می گردند؛ حتی بیش از یک خط هم برای نمایش نمودارهای مختلف بارنگ های متفاوت قابل نمایش است.

نمودار خطی در اغلب حوزه ها به عنوان نمودار پایه ای مورداستفاده قرار می گیرد. موضوع مهمی که درباره نمودارهای خطی مطرح است یافتن تابعی است که بیشترین تطبیق را به نقاط داده داشته باشد. چنین فرایندی فرایند برازش منحنی نامیده می شود. انواع روشهای رگرسیون و شبکه عصبی مصنوعی در این زمینه کاربرد دارند.



نمونه نمودار خطی

نمودار خطی مشبک (Trellis Line Graph)، نوعی نمودار خطی است که به دانشمندان اجازه می دهد تا مجموعه داده های پیچیده و چند متغیری را بررسی نمایند، یعنی سنجش میزان بیشتری اطلاعات در یک زمان. نمودار بر مبنای اصل تشابه با هم تایی ساده تر خود است



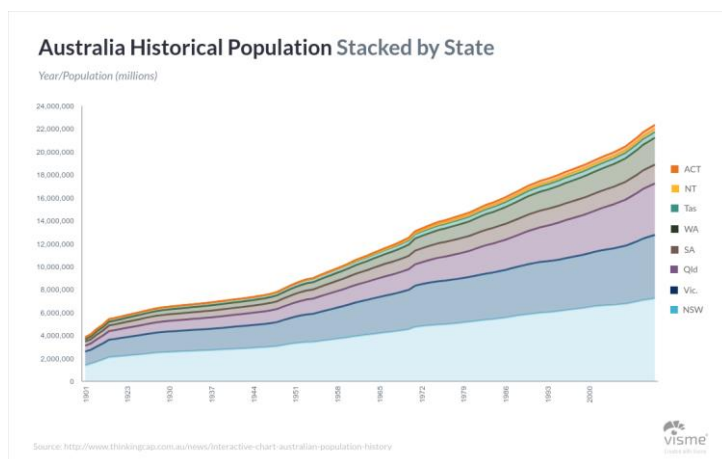
نمونه نمودار خطی مشبک

نوع دیگری از نمودار خطی، نمودار سهام یا Stock Charts است که به سرمایه گذاران کمک می کند تا بازارها را به منظور تعیین سود و زیان و همچنین تصمیم گیری های خرید و فروش، دنبال کنند. این نمودار به سادگی تغییرات در ارزش سهام خاص یا کلی بازار را طی یک دوره زمانی نشان می دهد. چندین سهام را می توان همزمان ردیابی و به سادگی با استفاده از خطوط مختلف رنگهای مختلف مقایسه کرد.



نمونه نمودار سهام

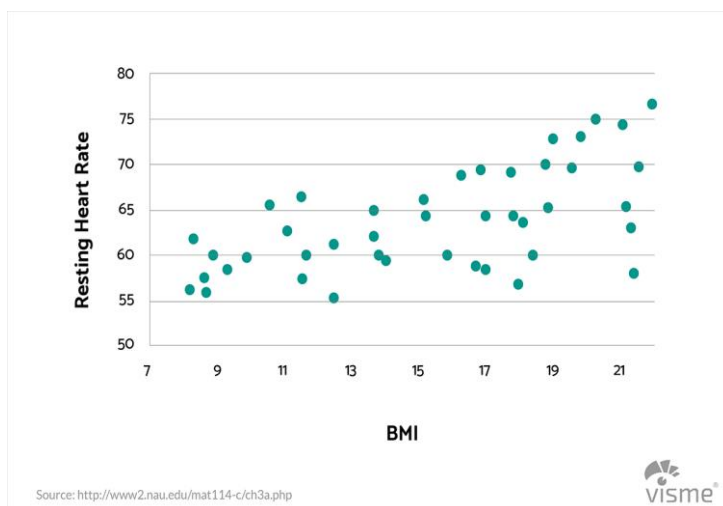
نمودار سطح انباشته یا مساحت (Stacked Area Chart)، همانند نمودارهای خطی هستند، با این تفاوت که ناحیه ی زیر خطوط کاملاً پر شده است. همچنین تمرکز نمودارهای خطی بر روی تغییرات مقدار نسبت به زمان است در حالی که نمودار مساحت بیشتر رشد در طول زمان را نمایش می دهد. تفاوت این دو بسیار جزئی است، ولی زمانی که بخواهیم برای مثال تفاوت تعداد درختان قطع شده بین سالهای ۱۳۸۰ تا ۱۳۸۵ و بین سالهای ۱۳۹۰ تا ۱۳۹۵ را به دست بیاوریم، نمودار مساحت است که این کار را به بهترین نحو انجام می دهد.



نمونه یک نمودار سطح انباشته

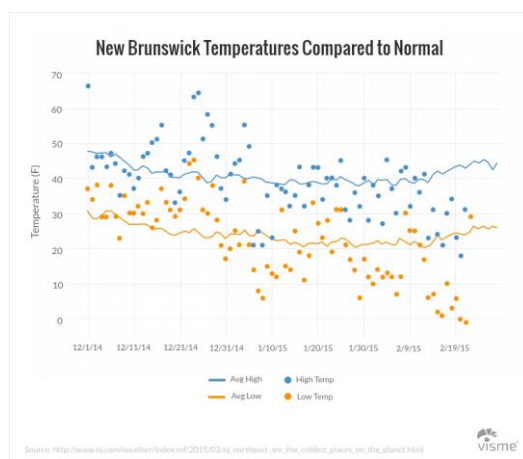
۵- نمودار پراکندگی (Scatter Plot Chart)

این نمودار ارتباط میان دو یا چند متغیر را نشان می دهد. داده ها در این نمودار به عنوان مجموعه ای از نقاط مشخص شده اند، هر کدام نمایانگر مقدار متغیری که در محورهای افقی یا عمودی تعیین شده است. مثلاً یک دکتر با مشاهده نمودار نرخ ضربان قلب در برابر شاخص توده بدن (BMI) می تواند به این نتیجه برسد که هرچه نرخ ضربان قلب بالاتر باشد، مقدار BMI بیشتر است.



نمونه نمودار پراکندگی

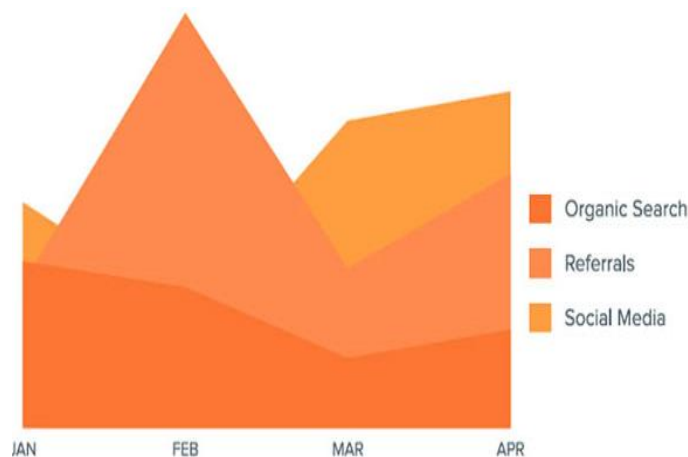
با ترکیب یک نمودار خطی با یک طرح پراکنده، متخصصان آمار می توانند رابطه بین دو مجموعه داده را بیان کنند. به این نمودار، نمودار Scatter-Line Combo می گویند. نمودار از دو محور تشکیل شده که هر کدام مجموعه ای از داده ها را نمایش می دهند. مثلاً یک محور ممکن است درصد اشغال تخت مربوط به بخش های بالینی به تفکیک باشد و محور دیگر چرخه اشغال تخت همان بخش ها را مشخص کند. برای هر بخش، میانگین درصد اشغال تخت بر حسب چرخه اشغال تخت با نقاطی روی نمودار نمایش داده می شود. نمودار پابون لاسو نمونه ای از این مدل نمودارهاست.



نمونه نمودار ترکیبی پراکنده - خطی

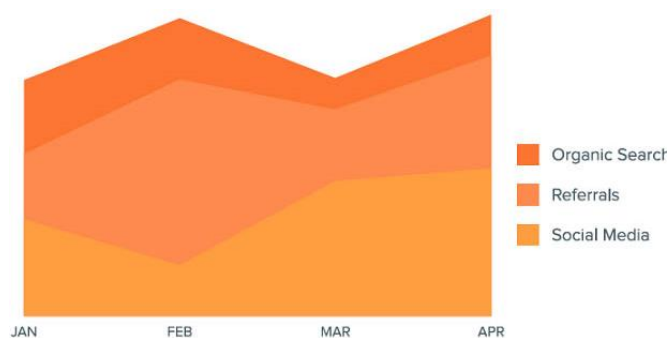
۶- نمودار منطقه ای (Area chart) انباشته یا تجمعی

نوع دیگری از نمودارهای مقایسه ای نمودار منطقه ای تجمعی است که حالت خاصی از نمودار منطقه ای است. برای همین ابتدا نمودار منطقه ای را معرفی می کنیم و در خلال آن به معرفی نمودار منطقه ای انباشته یا تجمعی می پردازیم. نمودار منطقه ای جهت نمایش روابط میان سری های زمانی بکار می رود؛ اما برخلاف نمودارهای خطی، آنها حجم را نیز می توانند به خوبی نمایش دهند. در اکثر موارد این نوع نمودارها دو یا تعداد بیشتری موضوعات را با یکدیگر مقایسه می کنند. این نوع نمودارها زمانی که با روندهای کلی در تعامل هستیم بسیار سودمند هستند. بلکه بهترین شکل کاربرد آنها زمانی است که می خواهیم روند چندین سری زمانی را که در میان آنها رابطه جزء به کل مطرح است در یک شکل نمایش دهیم.

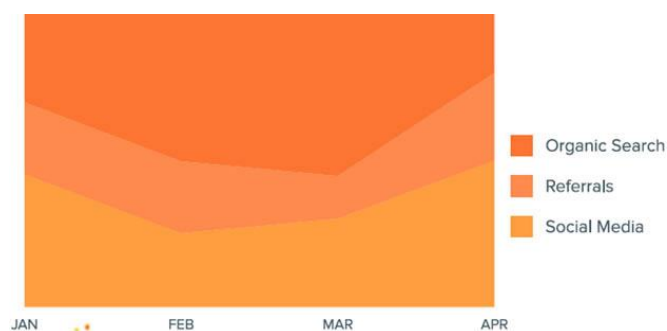


نمونه نمودار منطقه ای

رسم این نمودارها شبیه نمودارهای خطی است با این تفاوت که مساحت زیر منحنی هم می بایست نمایش داده شود و بدین ترتیب تقدم و تأخر سری های زمانی (متغیرها) نیز اهمیت می یابد؛ زیرا سطح زیر یک منحنی می تواند به کل، منحنی دیگر را بپوشاند. این نوع نمودارها، بسته به کاربرد، سه گونه مختلف دارند. شکل بالا نوع ساده (استاندارد) آنها را نمایش می دهد. نوع تجمعی و تجمعی ۱۰۰٪ به ترتیب در زیر آمده است.



نمونه نمودار منطقه ای تجمعی



نمونه نمودار منطقه ای تجمعی ۱۰۰٪

۷- نمودار عنکبوتی یا رادار (Spider / Radar Chart)

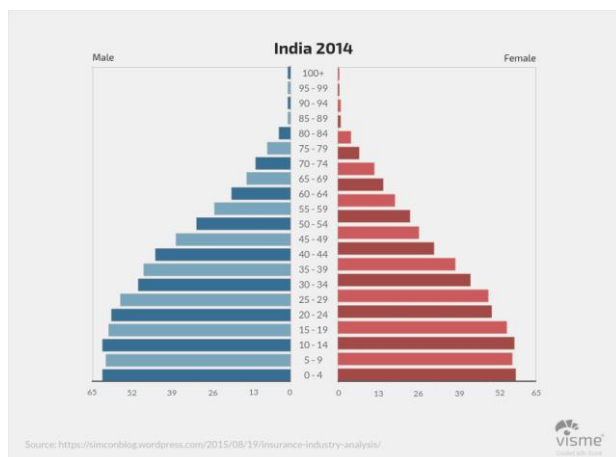
نمودار رادار که معمولاً به عنوان یک نمودار عنکبوتی یا نمودار ستاره ای نامیده می شود، مجموعه داده هایی را که شامل سه یا چند متغیر در یک گراف دو بعدی است، نمایش می دهد. مقدار کمی هر متغیر در طول محور که معمولاً از نقطه مرکزی نمودار شروع می شود، منعکس می شود. بعد از اینکه متغیرها روی نمودار مشخص شوند، با استفاده از خطوطی به یکدیگر وصل شده و یک چندضلعی نامنظم شکل می گیرد که ممکن است شباهتی به یک ستاره یا عنکبوت داشته باشد. مجموعه داده های چندگانه را می توان در یک گراف رادار، با رنگ های مختلف از هم تفکیک کرد و با استفاده از برجسب ها مشخص کرد که هر رنگ مربوط به کدام متغیر موردنظر می باشد. برای مثال، یک نمودار رادار می تواند به طور واضح هزینه ها و پیامدهای رویه های پزشکی مختلف را که به شرایط متنوعی وابسته می باشند به خوبی مقایسه کرده و نشان دهد.



نمونه نمودار عنکبوتی

۸- نمودار هرم جمعیتی (Population Pyramid Chart)

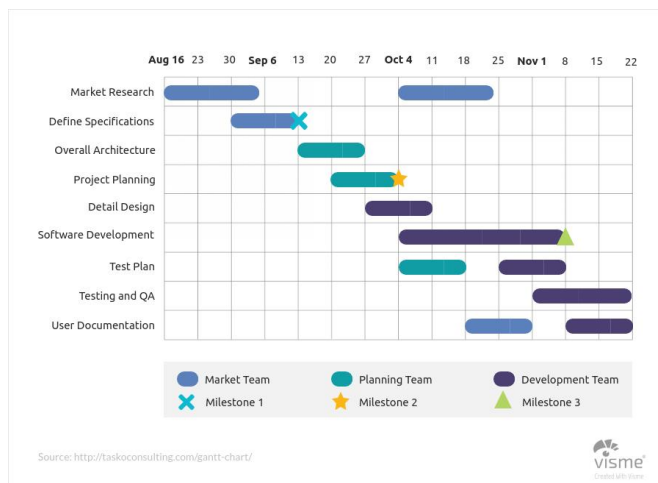
هرم جمعیتی، برای نمایش سن و جنسیت در یک جمعیت مشخص است. در حالت معمولی که جمعیت سالم و رو به رشد باشد، این نمودار شکل هرم به خود می گیرد - بیشترین گروه ها جوان ترین ها هستند و هر جنسیتی با افزایش سن افراد رفته رفته کاهش می یابد. در مواقع قحطی یا رونق اقتصادی ممکن است نمودار به خاطر افزایش مرگ و میر یا زاد و ولد، از این شمایل خارج شود.



نمونه هرم جمعیتی

۹- نمودار گانت (Gantt Charts)

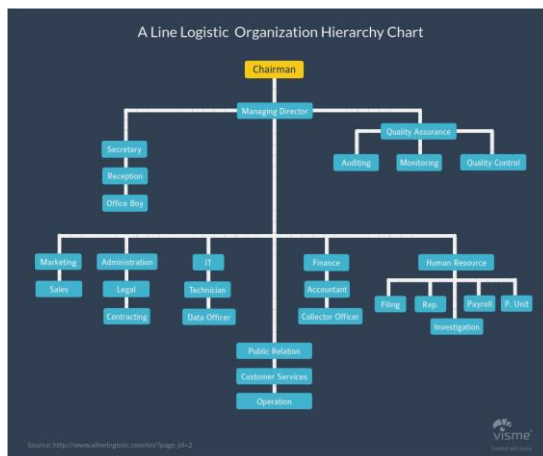
نمودار گانت نوع خاصی از نمودارهای میله‌ای است که برای طراحی و زمانبندی پروژه ها استفاده می‌شود. استفاده از میله های رنگی با طول های مختلف، نه تنها تاریخ شروع و پایان پروژه، بلکه رویدادهای مهم، وظایف، نقاط قوت و زمان بندی آنها را نیز نشان می دهد. نمودارهای مدرن گانت همچنین می توانند روابط وابستگی فعالیت ها را نشان دهند. به عنوان مثال، اتمام کار C، توسط تیم ۳ به تکمیل قبلی کار B توسط تیم ۲ بستگی دارد، نمودار نه تنها این رابطه را نشان می دهد، بلکه تاریخ و مهلت برنامه ریزی شده برای آنها را نیز نشان می دهد.



نمونه نمودار گانت

۱۰- نمودار سلسله مراتبی (Hierarchy Diagram)

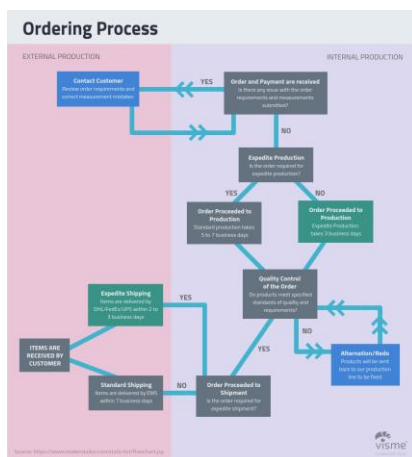
یک نمودار سلسله مراتبی، به عنوان یک نمودار سازمانی شناخته می شود که ساختار یک سازمان و همچنین روابط درون آن را نمایش می دهد. برای نمونه در تصویر، یک شرکت سازمانی معمولی مدیر عامل شرکت را در بالای فهرست و به دنبال آن مدیران، معاونان و سایر اعضای سازمان نشان داده شده است. یک نمودار سازمانی می تواند زنجیره فرمان را از هر کارمند به بالا به تصویر بکشد. نمودارهای سلسله مراتبی به طور مشابه برای نشان دادن نژادشناسی، طبقه بندی علمی، جمعیت شناسی و هر مجموعه داده ای که به صورت مشابه تفکیک شده استفاده می شود.



نمونه نمودار سازمانی

۱۱- نمودار جریان (Flow Chart)

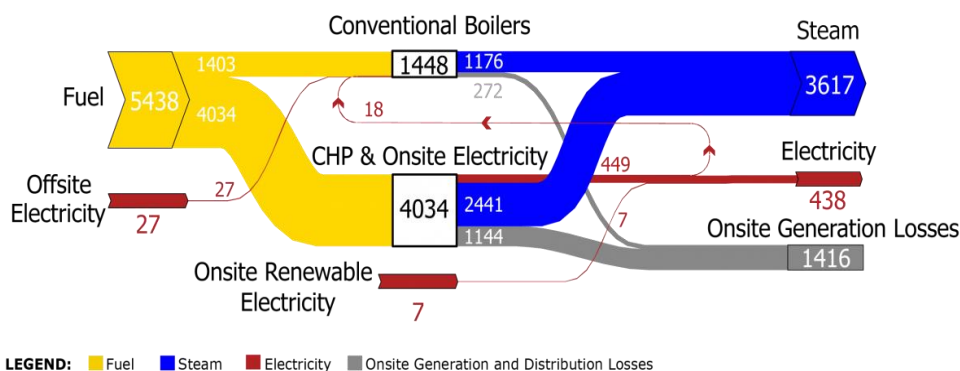
اغلب اوقات در سازمان، فرایندهای کاری باید در قالب نمودار نشان داده شوند. یک فلوچارت فرآیند را به صورت گام به گام از ابتدا تا انتهای آن به منظور تجزیه و تحلیل، طراحی، مستندسازی یا مدیریت آن دنبال می کند. نمودار جریانی معمولاً توسط بیمارستان ها، کلینیک ها و سایر اماکن پزشکی مورد استفاده قرار می گیرد تا اطمینان حاصل شود که رویه های یکسانی دنبال می شوند. در این نمودارها حتی ممکن است چندین خطوط آغازین و پایانی را برای یک گام از فرآیند رسم کرد. این نمودارها بیشتر برای نشان دادن توالی پیچیده تر با تصمیمات یا شرایط مختلف در طول مسیر استفاده می شوند.



نمونه نمودار جریان (فلوچارت)

نمودارهای سنکی (Sankey diagram) گونه خاصی از نمودارهای جریان هستند که عرض یا ضخامت هر کمان نسبت مقدار جریان را نشان می دهد. این نمودارها عمدتاً برای تصویر کردن تبادلات انرژی، مواد یا هزینه میان فرایندها و گزارش جریان انرژی و مواد در سطوح کلان به کار می رود. نمودارهای سنکی بر روی جریان ها یا نقل و انتقالات بزرگ درون یک سیستم تأکید می کنند و می توان با کمک آنها بخش های عمده یک جریان کلی را شناسایی کرد.

Onsite Generation (TBtu), 2010

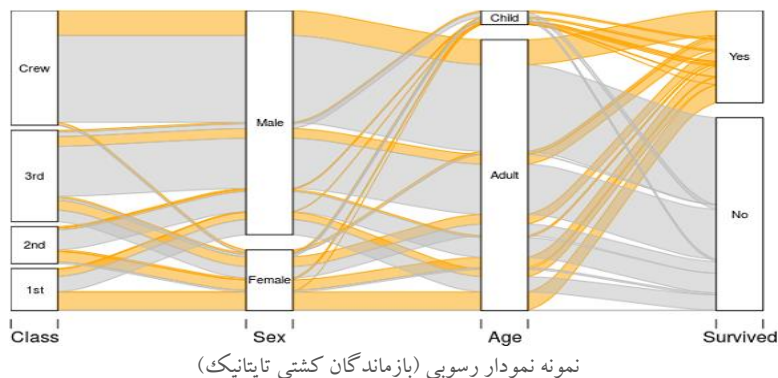


نمونه نمودار سنکی فعالیت دیگ بخار

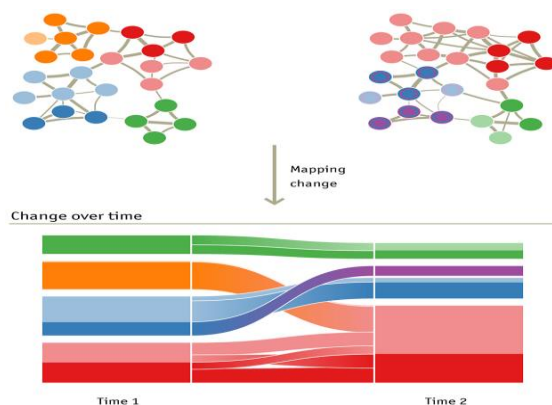
رسم دستی نمودار سنکی (انرژی) شامل نکات زیر است:

- ورودی ها در سمت چپ نمودار واقع می شوند.
- خروجی های مفید در سمت راست واقع می شوند.
- خروجی های غیر مفید (ضایعات) به صورت کمان های عمودی و رو به سمت پایین کشیده می شوند.
- عرض هر کمان بیانگر (نسبت) مقدار انرژی مربوط به آن کمان است درحالی که طول کمان ها ارتباطی با مقدار انرژی ندارند.

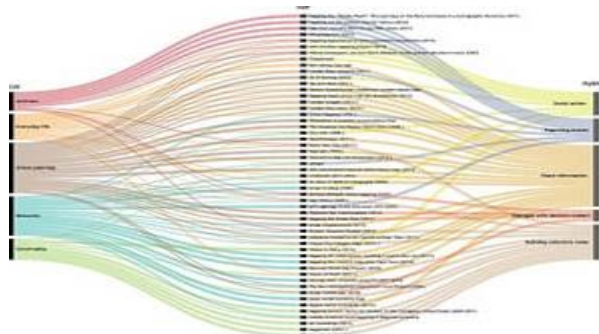
دیاگرام های رسوبی (Alluvial diagram) گونه ای دیگر از دیاگرام های جریان هستند که هدف از توسعه آنها نمایش تغییرات در ساختار شبکه در طول زمان بوده است. وجه تسمیه این نمودارها پیدایش مخروط های رسوبی است که به طور طبیعی توسط خاکی که در آب جاری رسوب می کند، تشکیل می گردد. نمونه ای از این نمودارها در شکل زیر نمایش داده شده است. یکی از اصلی ترین کاربردهای دیاگرام های رسوبی در نمایش تغییرات ساختاری در شبکه های پیچیده است. این نمودارها می توانند جهت نمایش هرگونه تغییرات در ترکیب گروه ها در حالات یا زمان های مختلف به کار گرفته شوند و اطلاعات آماری را نیز به همراه داشته باشند. همچنین آن ها می توانند جهت نشان دادن الگوهای جریان در یک شبکه ثابت در طول زمان بکار روند. به عنوان نمونه می توان از این نمودار جهت نمایش چگونگی حرکت بازدید کنندگان سایت از صفحه ای به صفحه دیگر استفاده نمود.



رسم این نمودار، همان طور که از شکل آن نیز مشخص است، به سادگی نمودارهای دیگر نیست. البته نسخه های ساده تری (از نظر رسم) برای این نمودار وجود دارد که از نظر گرافیکی در سطح نازل تری از شکل های این قسمت است. برای رسم این نمودار دو مرحله اصلی عبارت اند از: ۱- تعیین گروه های اصلی که (در طول زمان) تغییر یافته اند. ۲- رسم تغییرات. شکل زیر این موضوع را نشان داده است.



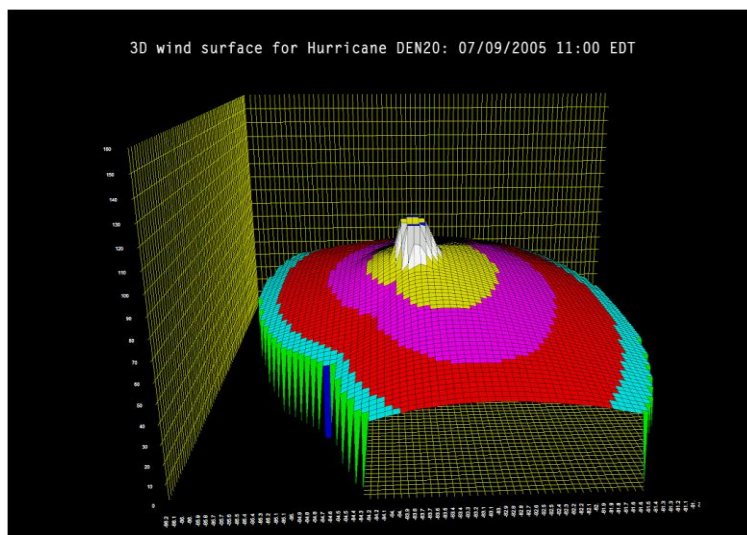
یکی از ساده ترین روش های رسم دیاگرام رسوبی نوشتن اقلام در ستون های مجزا و وصل کردن آن ها به حالت تغییر یافته خود با رنگ معین است. نمونه ای از این نحوه رسم را در شکل زیر می بینید.



نمونه ساده رسم دیاگرام رسوبی

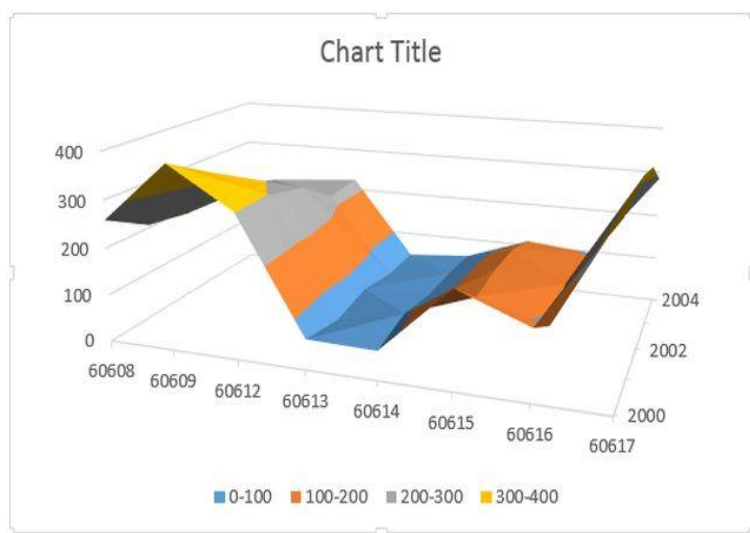
۱۲- نمودار ۳ بعدی (3D Graph)

تکنولوژی در حال حاضر اجازه می دهد تا متخصصان آمار مجموعه داده های چند بعدی را در فرم درست نمایش دهند. نمودار سه بعدی که توسط نرم افزارهای تخصصی ایجاد شده، منعکس کننده رابطه بین سه متغیر در سه محور است. برای مثال یک هواشناس می تواند فیلد باد و طوفان را در قالب نمودار رسم کند.



نمونه نمودار سه بعدی

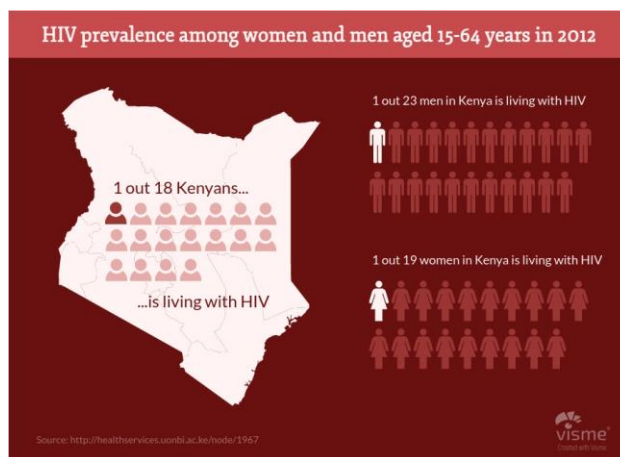
نمودار سطحی که نوعی نمودار سه بعدی است به شما این امکان را می دهد که از دو سری نقاط داده ای در داخل چند بعد استفاده کنید. ممکن است استفاده از این نمودار پیچیده باشد، ولی اگر نقاط داده ای درستی را استفاده کرده باشید (دو مجموعه داده ای که رابطه ای مشخصی با یکدیگر داشته باشند)، به یک نمودار گرافیکی بسیار زیبا خواهید رسید.



نمونه نمودار سطحی

۱۳- نمودار Pictograph

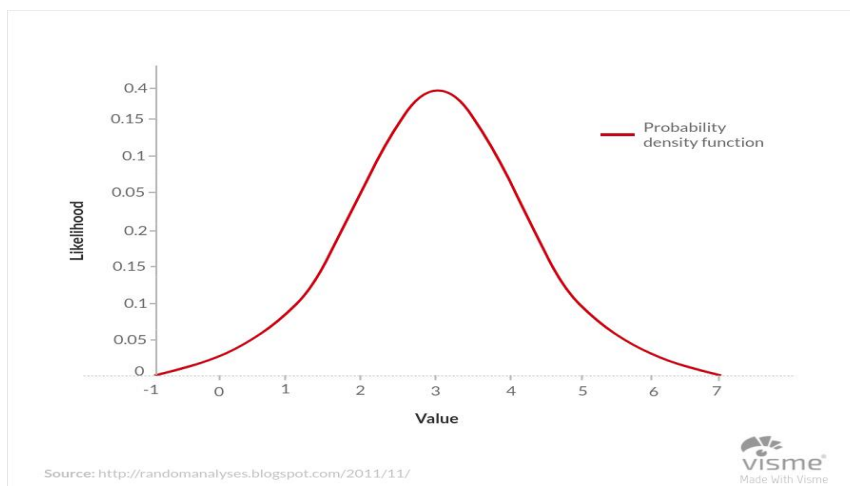
در این نمودار، تصاویر و نمادها برای نشان دادن داده ها استفاده می شوند. به عنوان مثال، یک نمودار تصویری پایه ممکن است یک تصویر از خورشید را برای نشان دادن روزهایی که هوا آفتابی است و یک ابر بارانی برای نماد هر روز طوفانی استفاده کند. از آنجا که ما از تصاویر احساس بیشتری نسبت به داده های خام دریافت می کنیم، اغلب نمودارهای تصویری برای ارائه اطلاعات پزشکی استفاده می شوند. زمانی که در یک تصویر از تعداد ۲۰ نفر، ۵ نفر را به منظور نشان دادن نرخ مرگ و میر ۲۰ درصدی رنگی می کند، این پیام اثر گذارتر از به عنوان مثال یک نمودار میله ای یا دایره ای که همین داده های مشابه را نشان می دهند، می گذارد.



نمونه نمودار پیکتوگراف

۱۴- نمودار تابعی یا منحنی (Function Plot)

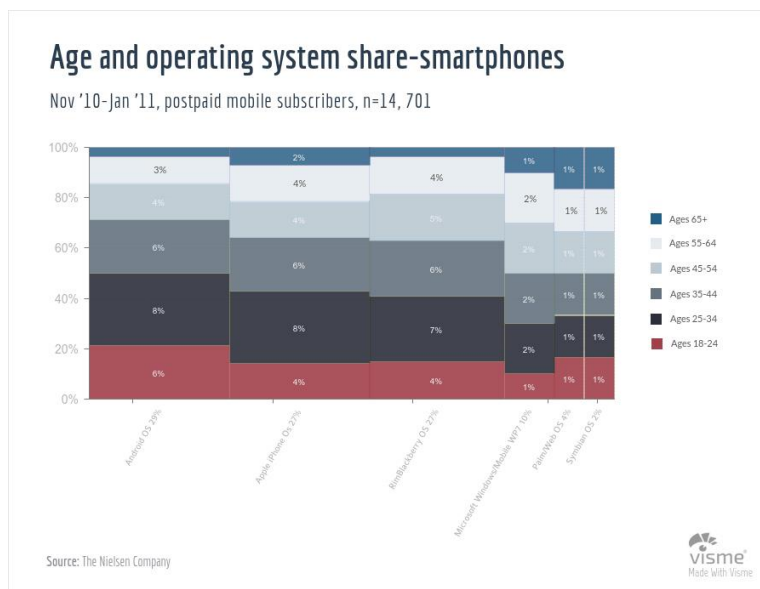
ریاضیدانان، مهندسان و متخصصان آمار اغلب برای تعیین مقدار یک معادله از نمایش نموداری آن استفاده می کنند. این نمودار مجموعه تمام نقاطی است که مقادیر ممکن معادله را برآورده می کنند. در نتیجه تابع یک معادله با مقادیر X و Y ، محورهای با همین نامها دارد. همچنین اگر متغیر سوم Z اضافه شود باید نمودار به حالت سه بعدی تبدیل شود.



نمونه نمودار تابعی (منحنی)

۱۵- نمودار موزاییکی (Mosaic or Mekko Chart)

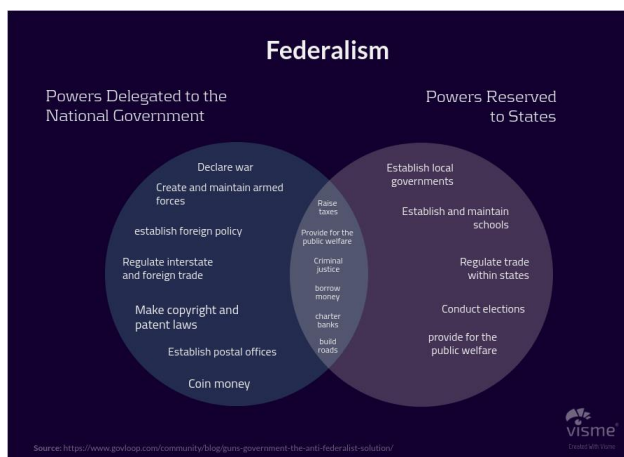
نمودارهای معرفی شده تا اینجا برای مقایسه یک یا دو متغیر در دسته های محدودی قابل استفاده هستند. اما در مواردی که چندین متغیر و دسته در یک زمان داشته باشیم و یا اگر تمام آن مقادیر از جنس عدد نباشند، نمودار موزاییکی بهترین گزینه است. مثلاً اگر بخواهیم فروشگاه های چند شهر را همزمان با تعداد مشتریان آن رصد کنیم از این نمودار استفاده می شود.



نمونه نمودار موزاییکی

۱۶- نمودار ون (Venn Diagram)

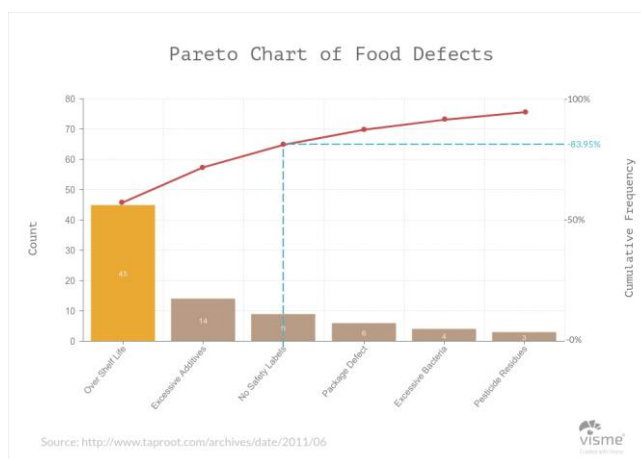
نمودار کلاسیک و معروف ون که تحت عنوان نمودار منطقی هم شناخته می شود، تمام روابط منطقی ممکن میان مجموعه ای از دادگان را نشان می دهد. مثلاً فضای مشترک میان دو نمودار بالا، قدرت هایی را نشان می دهد که در هر دو حزب حفظ شده است. مجموع تمام نمودارها در کنار هم اجتماع را مشخص می کند. نمودار ون که اندازه و فضای هر شکل آن متناسب است با اندازه گروهی که نشان می دهد.



نمونه نمودار ون

۱۷- نمودار پارتو (Pareto Chart)

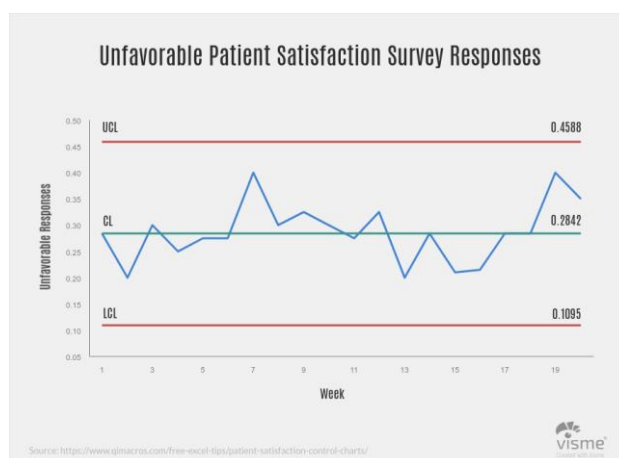
گاهی اوقات یک نمودار ساده، اطلاعات کافی را برای نتیجه گیری لازم در اختیار نمی گذارد. پارتو، دو نمودار ستونی و خطی را برای نمایش نه تنها مقدار مختلف طبقه ها بلکه مقدار تجمعی کل مجموعه ترکیب کرده است. در مثال بالا پارتو به بررسی فراوانی انواع مختلف مشکلات غذایی پرداخته است. نتیجه یک نمودار است که به روشنی رایج ترین مشکلات غذایی و درصدی از کل که به خود اختصاص داده اند را نشان می دهند.



نمونه نمودار پارتو

۱۸- نمودار کنترل (Control Chart)

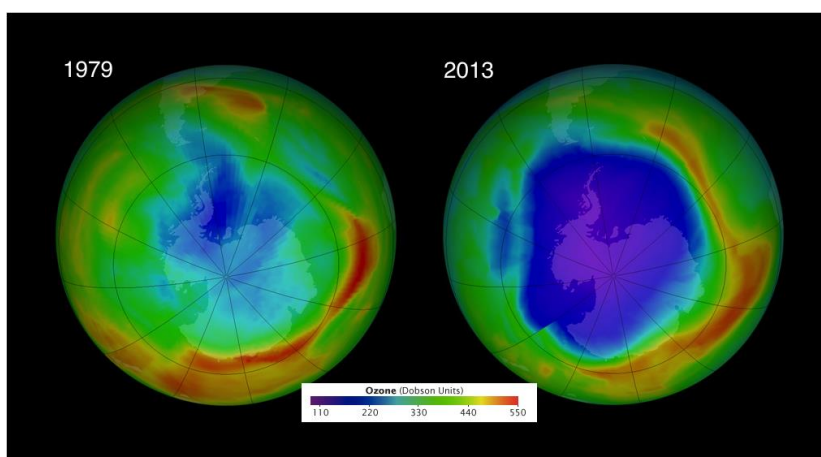
یک نمودار کنترل که اغلب در فرآیندهای کنترل کیفیت استفاده می شود کمک می کند تا تعیین شود که آیا یک مجموعه داده در یک محدوده کنترل یا از پیش تعیین شده قرار می گیرد یا خیر. به طور معمول یک نمودار کنترل شامل نقاطی است که بر روی دو محور قرار گرفته اند و نشان دهنده اندازه گیری های نمونه (sample measurement) می باشند. بر اساس میانگین محاسبه شده، خطی در امتداد گراف در این مقدار به دست آمده، در نظر گرفته می شود. سپس یک انحراف استاندارد از میانگین با استفاده از هر نمونه محاسبه می شود. در نهایت، محدودیت های بالایی و پایینی تر تعیین و تشریح می شوند تا نشان دهنده نقاطی باشند که انحراف بیش از حد انتظار است.



نمونه نمودار کنترل

۱۹- نمودار کانونی کروی (Spherical Contour Graphs)

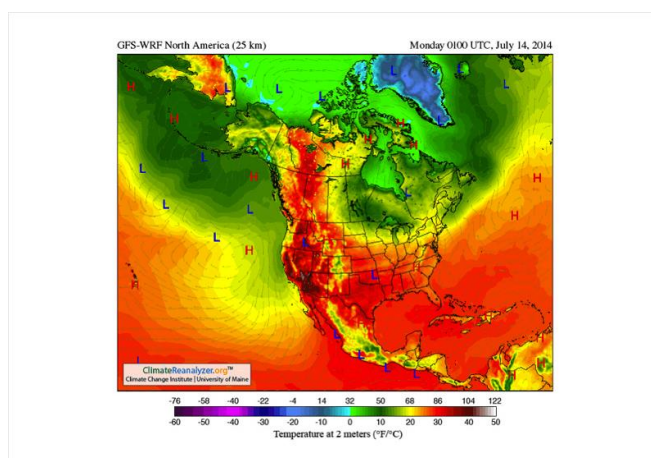
از آنجایی که زمین کروی شکل می باشد، ثبت یا رسم مسیر حرکت سیاره در یک گراف که شامل دو محور است، مشکل به وجود می آورد. به همین منظور میتوان، داده ها را در سه محور با استفاده از متغیرهای X ، Y و Z ترسیم کرد. طرح نهایی، اگر به درستی ترسیم شود، به صورت یک کره در می آید. به عنوان مثال، یک طرح کروی می تواند درجه حرارت جهانی یا بارندگی را با تعیین هر محدوده ارزش با یک رنگ مشخص نشان دهد، سپس داده ها را با نقاط رنگ مربوطه ترسیم می کند.



نمونه نمودار کروی

۲۰- نمودار Contour Plots

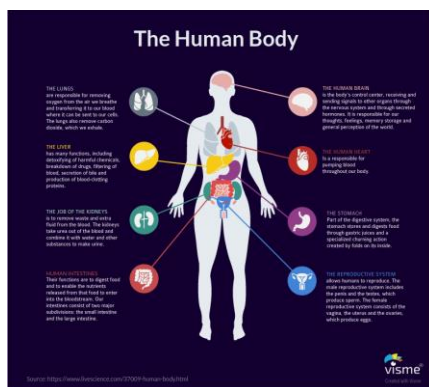
قطعه های Contour این امکان را فراهم می کند تا بتوانیم به تجزیه و تحلیل سه متغیر در فرمت دو بعدی بپردازیم. ترسیم داده ها در دو محور اصلی، بر اساس رنگ یا سایه است. درست همانطور که یک نقشه توپوگرافی، طول، عرض و ارتفاع را در یک طراحی دو بعدی نشان می دهد. نقشه های حرارتی (Heat Map Chart) در هواشناسی، نمونه ای از این نمودارهاست.



نمونه نمودار کانتور (نقشه حرارتی)

۲۱- نمودار تشریحی (Anatomical Diagram)

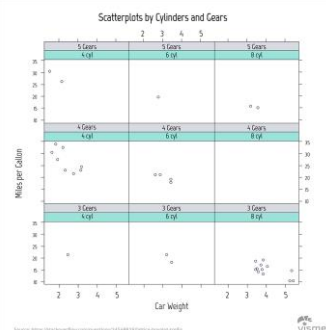
نمودارهای پزشکی اغلب برای نشان دادن آناتومی، طرزعمل یا پاتولوژی بیماری به منظور توضیح روش درمان برای بیماران و دیگر افرادی که پیش زمینه پزشکی را ندارند استفاده می شود. مسلماً بسیار ساده است زمانی که اثر اعصاب بر روی سیستم های مختلف بدن را با دیاگرام نشان دهیم تا اینکه بخواهیم شیوه ها و روش های صحیح را بدون تصویر توضیح دهیم. در حالی که نمودارهای پزشکی ترکیبی از علم و هنر در نظر گرفته می شود، آنها می توانند به همان اندازه مانند سایر نمودارهای کمی، تخصصی نیز باشند. و مهم نیست که تصویر چقدر با جزئیات تهیه شده باشد، نمودارهای تشریحی به منظور ارائه داده ها به صورت روشن و کارآمد طراحی شده اند. و این نمودارها بر اطلاعات کلیدی تمرکز می کنند، حتی اگر از مقادیر زیادی اطلاعات پزشکی یا علمی انتخاب شده باشد.



نمونه نمودار تشریحی

۲۲- نمودار داربستی یا مشبک (Trellis Plot)

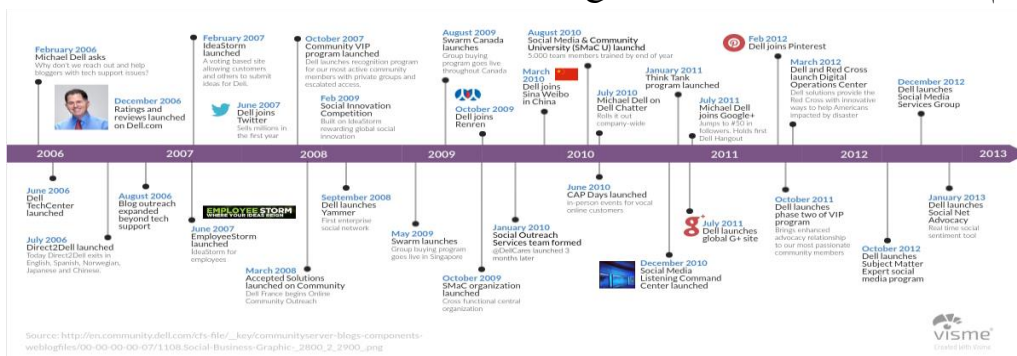
گاهی اوقات متخصصان آمار نیاز به مقایسه چندین مجموعه داده ها دارند که در یک گراف قابل نمایش باشد. مثلاً اگر علاوه بر درصد اشغال تخت و چرخه اشغال تخت، مدت اقامت بیمار و تعداد بیمار در هر بخش هم باید بررسی شود، کاربرد این نمودار هویدا می شود.



نمونه نمودار مشبک

۲۳- نمودار گاه شمار (Timelines)

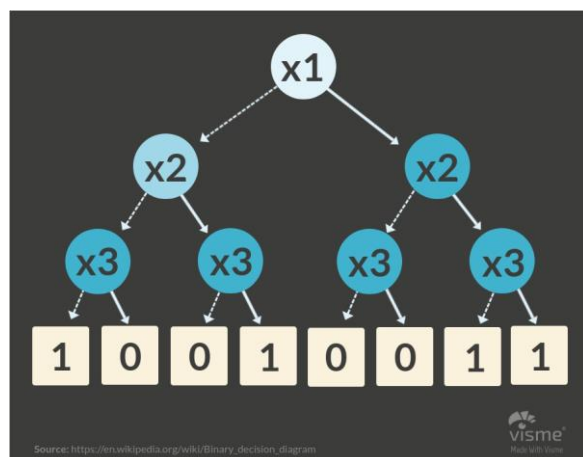
خود توصیف کننده ترین نمودار ممکن را می توان این نمودار به شمار آورد که داده ها را در بازه زمانی خاصی بررسی می کند. داده ها و اتفاقات مهم در نقاط مشخص مربوط به بازه زمانی وقوع، نمایش داده می شوند.



نمونه نمودار گاه شمار

۲۴- نمودار تصمیم دودویی (Binary Decision Diagram)

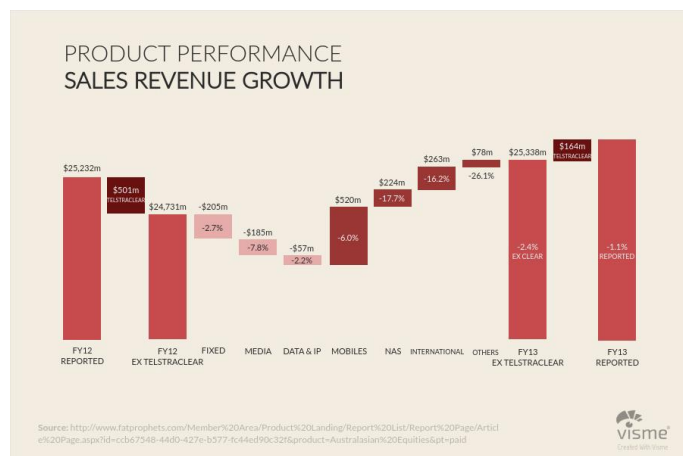
تصمیم دودویی، انتخابی میان دو راهکار متفاوت است، در نتیجه نمودار تصمیم دودویی مسیر ممکن میان یک تصمیم تا تصمیم بعدی را نشان می دهد. در علم کامپیوتر، تصمیم های دودویی، نوع داده بولی (Boolean) را تشکیل می دهند که اقدام های منتج از دو مقدار متفاوت را در یک جریان پردازشی مشخص می کنند. خارج از حوزه کامپیوتری هم برای هر فرایندی که منجر به اتخاذ تصمیم میان دو مقدار، بله/خیر، درست/نادرست و... می شود، کاربرد دارد.



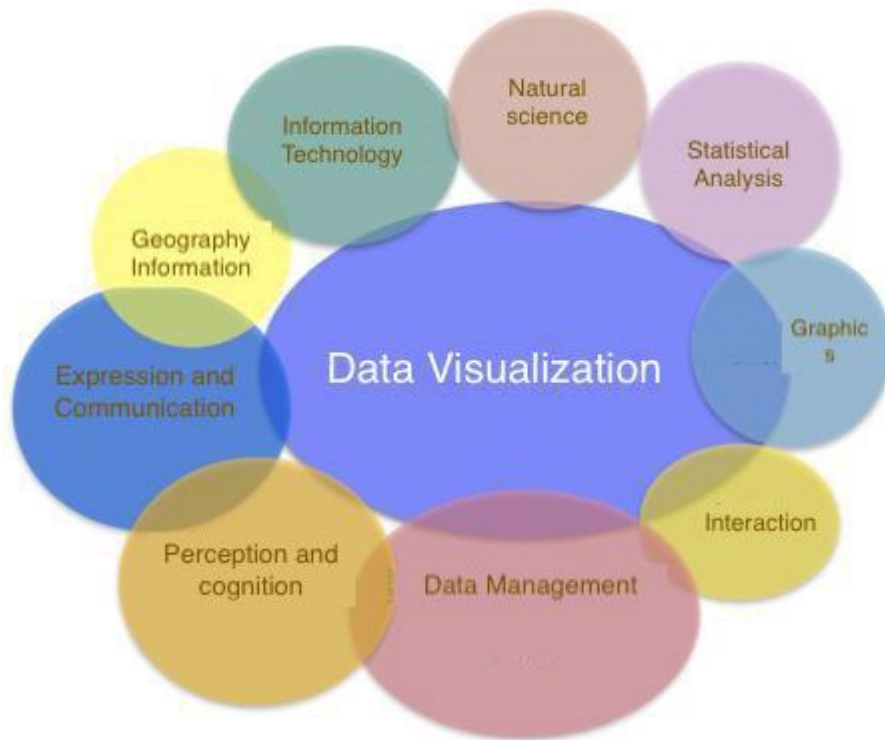
نمونه نمودار دودویی

۲۵- نمودار آبشاری (Waterfall Charts)

نمودارهای آبشاری به ویژه در حسابداری و تجزیه و تحلیل کیفی نشان می دهد که چگونه یک مقدار اولیه توسط فاکتورهای مختلف تاثیر مثبت و منفی می گیرد. به عنوان مثال، یک نمودار آبشاری می تواند به وضوح و به طور کارآمد به تصویر بکشد که چگونه یک بازده آغازین به طور ماهانه در طی یک سال تغییر می کند. از آنجا که اغلب به نظر می رسد که میله ها در سراسر نمودار شناور می شوند، نمودار آبشار گاهی اوقات به عنوان آجرهای شناور یا نمودار ماریو شناخته می شود.



نمونه نمودار آبشاری



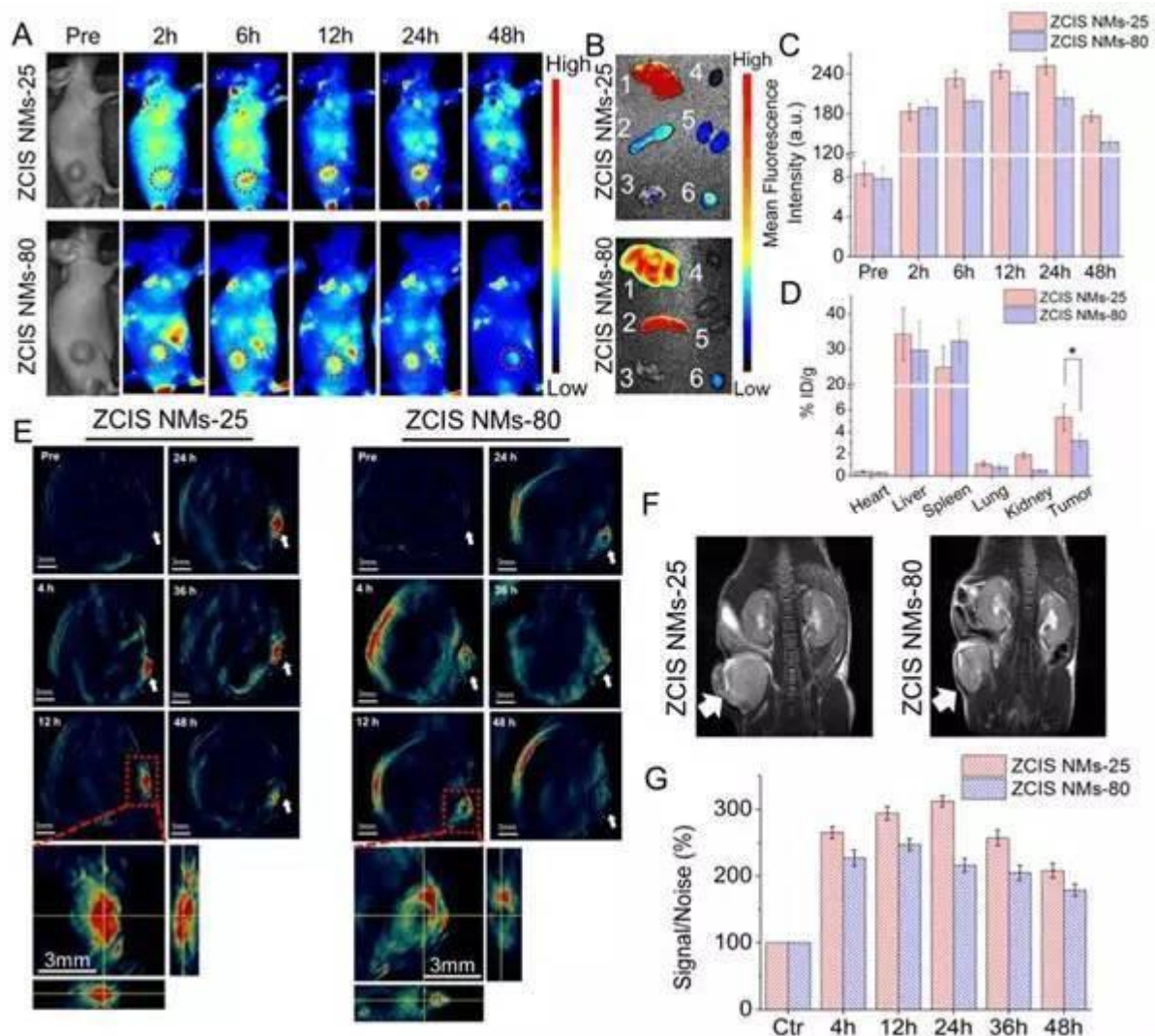
بدون شک امروزه در زمینه علم داده‌ها، مصورسازی داده‌ها واژه برتر است. مهم نیست که چه داده‌هایی را می‌خواهید آنالیز کنید، به نظر می‌رسد که انجام مصورسازی داده‌ها یک گام ضروری باشد. اما بسیاری از افراد مفهوم خاصی از مصورسازی داده‌ها ندارند و نمی‌دانند چگونه آن را اجرا کنند. بنابراین، امروز قصد دارم تا شما را با تعریف، مفهوم، فرآیند اجرا و ابزارهای مصورسازی داده‌ها آشنا کنم.

۱. مصورسازی داده‌ها چیست؟

مصورسازی علمی، مصورسازی اطلاعات و تجزیه و تحلیل بصری اغلب به عنوان سه شاخه اصلی مصورسازی دیده می‌شوند. "مصورسازی داده"، که ترکیبی از این سه شاخه است، یک نقطه شروع جدید در زمینه تحقیقات بصری است. مصورسازی عمومی داده‌ها شامل رشته‌های مختلفی مانند فن‌آوری اطلاعات، علوم طبیعی، تحلیل آماری، گرافیک، تعامل و اطلاعات جغرافیایی است.

۱.۱ مصورسازی علمی

مصورسازی علمی یک حوزه مطالعاتی و کاربردی میان رشته‌ای در علم است که بر تجسم پدیده‌های سبب‌مندی مانند معماری، هواشناسی، پزشکی یا سیستم‌های بیولوژیکی تمرکز دارد. هدف آن نشان دادن گرافیکی داده‌های علمی و قادر ساختن دانشمندان به درک، توضیح و جمع‌آوری الگوها از داده‌ها است.



۱.۲ مصورسازی اطلاعات

مصورسازی اطلاعات مطالعه بازنمایی‌های بصری تعاملی داده‌های انتزاعی برای افزایش شناخت انسان است. چکیده داده‌ها شامل داده‌های دیجیتال و غیر دیجیتال مانند اطلاعات جغرافیایی و متن هستند. گرافیک مانند هیستوگرام‌ها، نمودارهای روند، نمودارهای جریان و نمودارهای درختی همگی به تجسم اطلاعات تعلق دارند و طراحی این گرافیک‌ها مفاهیم انتزاعی را به اطلاعات بصری تبدیل می‌کند.

۱.۳ تحلیل بصری

تجزیه و تحلیل بصری یک زمینه جدید است که با توسعه مصورسازی علمی و مصورسازی اطلاعات، با تاکید بر استدلال تحلیلی از طریق یک رابط بصری تعاملی، تکامل یافته است.

on of sales amount

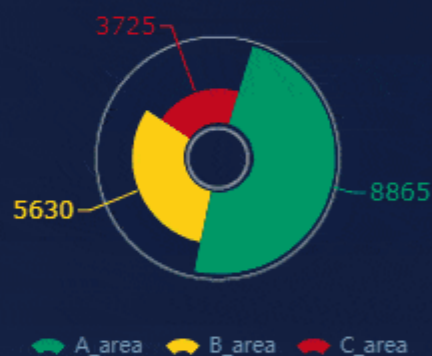


es Volume



User Statistics

Users	256 unit
Market share	26 %
Total sales	9163 set
Sales Amount	45195 10M



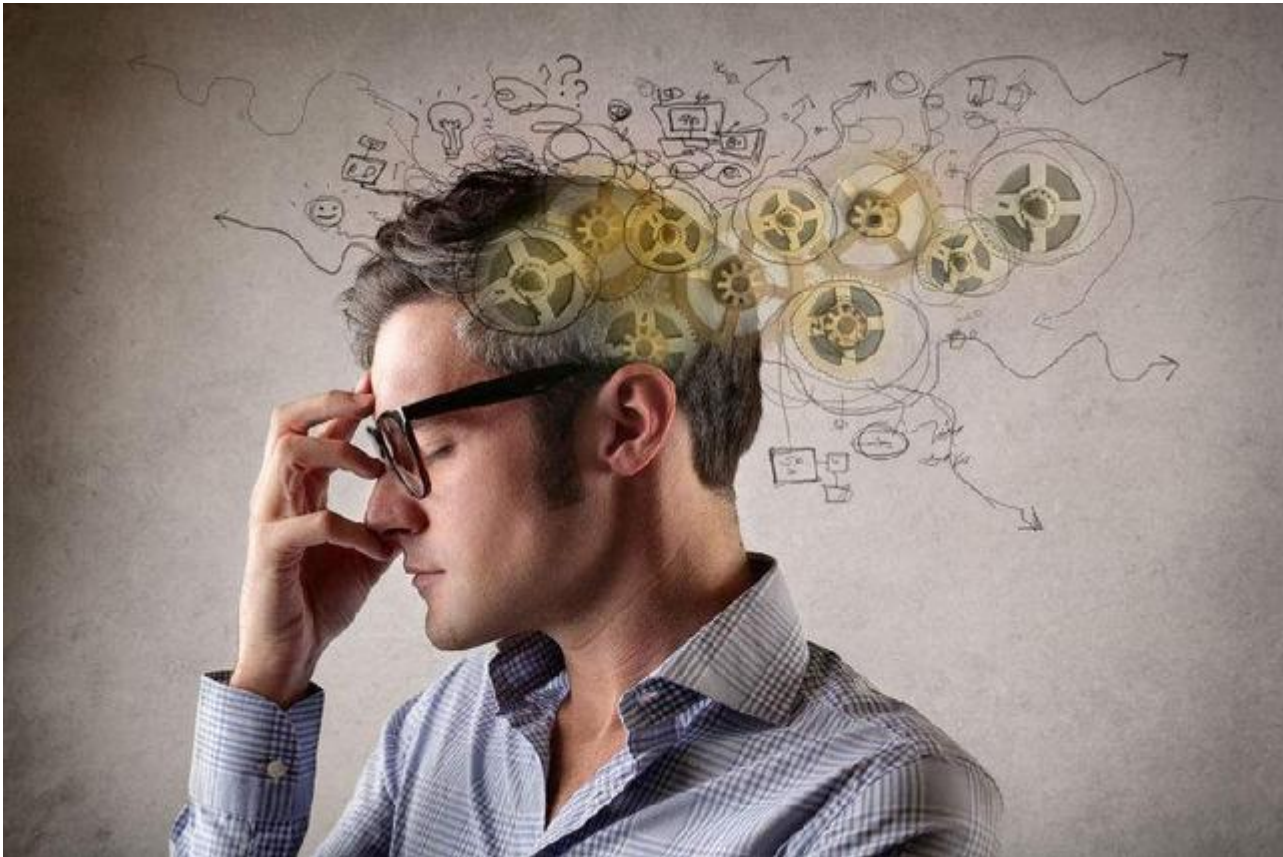
Market Analysis

■ Sale ● growth_rate



۲. چرا ما نیاز به مشاهده داده‌ها داریم؟

میزان اطلاعاتی که انسان‌ها از طریق بینایی به دست می‌آورند بسیار فراتر از دیگر اندام‌ها است. مصورسازی داده‌ها استفاده از مهارت‌های طبیعی انسان برای افزایش پردازش داده‌ها و کارایی سازمان است.

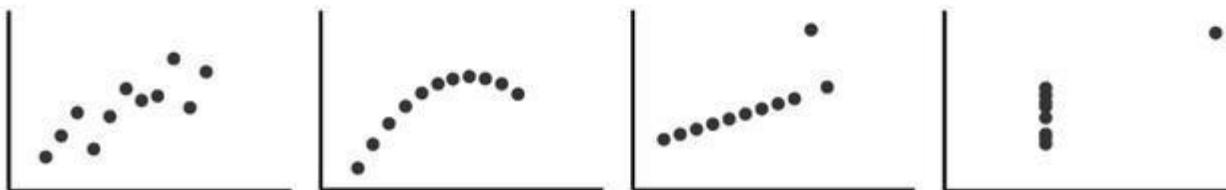


تجسم کردن می‌تواند به ما کمک کند تا با اطلاعات پیچیده‌تر سر و کار داشته باشیم و حافظه را افزایش دهیم. اغلب افراد اطلاعات آماری زیادی ندارند و روش‌های آماری پایه (میانگین، میانه، دامنه، و غیره) با ماهیت شناختی انسان هماهنگ نیستند. یکی از معروف‌ترین نمونه‌ها، [چهارتایی آنسکوم](#) است. مشاهده قانون با توجه به روش‌های آماری دشوار است، اما قوانین بسیار واضح هستند زمانی که داده‌ها به تصویر کشیده می‌شوند.

a

I		II		III		IV	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.10	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.10	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

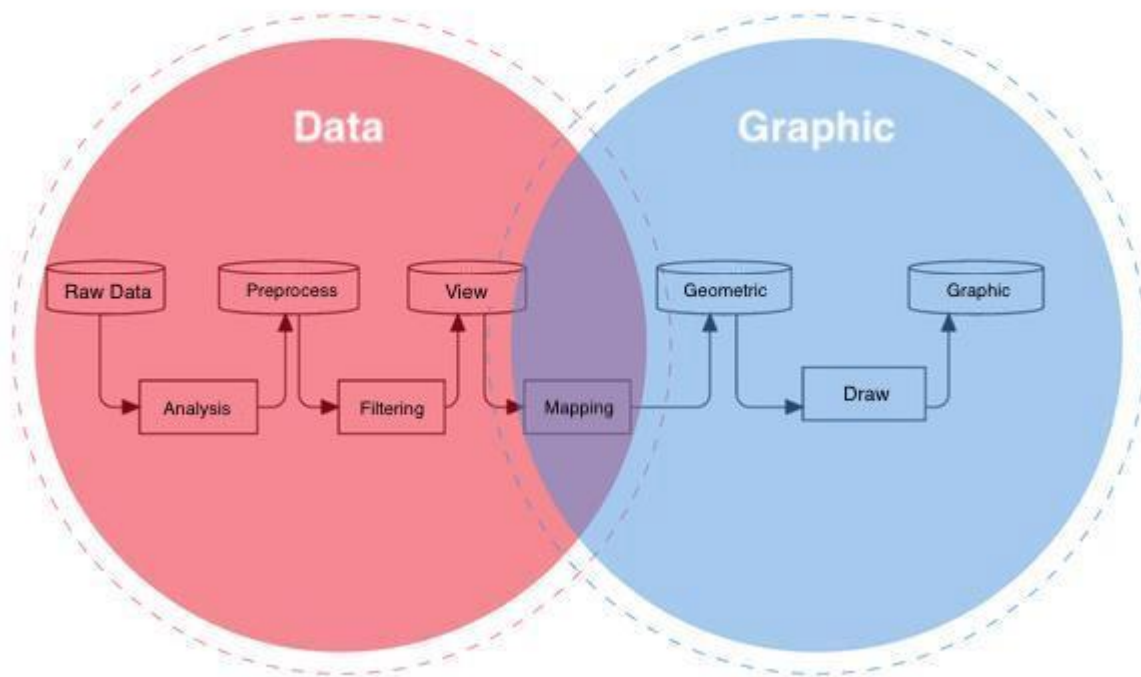
b



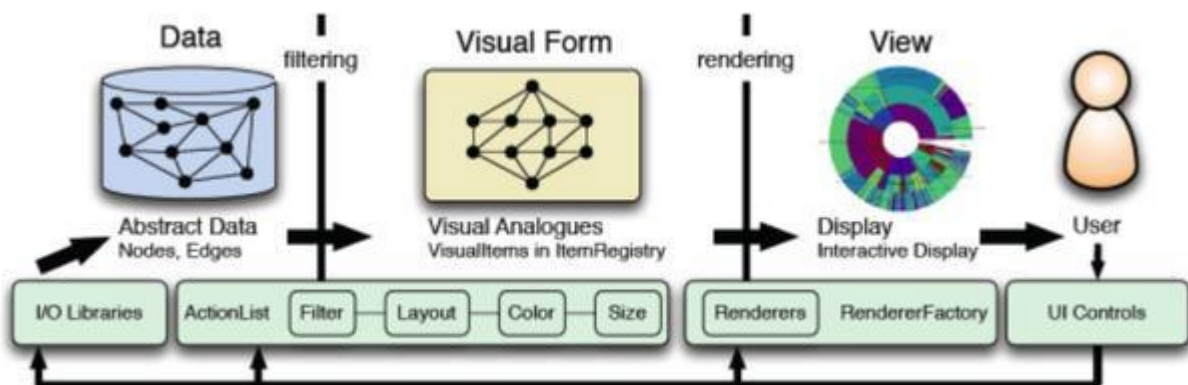
۳. چگونه می توان به مصورسازی داده ها دست یافت؟

به طور فنی، ساده ترین درک از مصورسازی داده ها، نگاشت از فضای داده ها به فضای گرافیکی است.

Mapping from data space to graphic space



یک روش پیاده‌سازی بصری کلاسیک پردازش و فیلتر کردن داده‌ها، تبدیل آن به یک شکل بصری قابل‌بیان، و سپس تبدیل آن به یک نمای کاربر - قابل‌مشاهده است.



پشته فن‌آوری تجسم

به طور کلی، مهندسان حرفه‌ای مصورسازی داده‌ها باید پشته فن‌آوری زیر را مدیریت کنند:

- ریاضیات پایه: توابع مثلثاتی، جبر خطی، الگوریتم هندسی
- گرافیک: کانس، SVG، WebGL، گرافیک محاسباتی، نظریه گراف
- الگوریتم‌های مهندسی: الگوریتم‌های پایه، الگوریتم‌های آماری، الگوریتم‌های چیدمان مشترک
- تجزیه و تحلیل داده‌ها: پاکسازی داده‌ها، آمار، مدل‌سازی داده‌ها

- **زیبایی‌شناسی طراحی:** اصول طراحی، قضاوت زیبایی‌شناسی، رنگ، تعامل، شناخت
- **مبانی بصری:** کدگذاری بصری، تحلیل بصری، تعامل گرافیکی
- **راه‌حل‌های تجسم:** استفاده صحیح از نمودارها، تجسم سناریوهای کسب‌وکار مشترک

۴. ابزارهای مصورسازی داده رایج

به طور کلی در دانشگاه‌ها، زبان‌های R و پایتون و ggplot۲ مورد استفاده قرار می‌گیرند. آشناترین ابزار برای کاربران معمولی اکسل است. محصولات تجاری شامل Tableau، FineReport، Power BI و غیره هستند.

(۱) ابزار جاوااسکریپت ۳D

ابزار ۳D.js یک کتابخانه جاوااسکریپت بر پایه اسناد دستکاری داده‌ها است. ۳D مولفه‌های تجسم قدرتمندی را با روش‌های دستکاری DOM برگرفته از داده ترکیب می‌کند.

ارزیابی: ۳D قابلیت عملیات SVG قدرتمندی دارد. این الگوریتم به راحتی می‌تواند داده‌ها را به تصاویر SVG نگاشت کند و تعداد زیادی ابزار و روش برای پردازش داده‌ها، الگوریتم‌های چیدمان و محاسبه گرافیکی را با هم ترکیب می‌کند. جامعه قوی و دموهای عالی و کامل دارد. با این حال API آن بسیار سطح پایین است. قابلیت استفاده مجدد زیادی وجود ندارد در حالی که هزینه یادگیری و استفاده از آن بالا است.

(۲) ابزار HighCharts

کتابخانه HighCharts یک کتابخانه نقشه نوشته شده در جاوااسکریپت خالص است که اضافه کردن نمودارهای تعاملی به برنامه‌های کاربردی وب را برای کاربران آسان و راحت می‌سازد. این کتابخانه پرکاربردترین ابزار نمودار در وب است و استفاده تجاری از آن نیازمند خرید مجوز تجاری است.

ارزیابی: آستانه استفاده بسیار پایین است. نقاط برجسته سازگاری خوبی دارند و بالغ هستند و به طور گسترده مورد استفاده قرار می‌گیرند. با این حال، این سبک قدیمی است و گسترش نمودارها دشوار است. و استفاده تجاری نیازمند خرید مجوز است.

(۳) ابزار Echarts

ابزار Echarts یک ابزار ترسیم نمودار حرفه‌ای است از تیم تجسم داده بایدو. این یک کتابخانه ترسیم نمودار جاوااسکریپت خالص است که به سادگی بر روی کامپیوترهای شخصی و دستگاه‌های موبایل اجرا می‌شود و با اکثر مرورگرهای فعلی سازگار است.

ارزیابی: Echarts دارای انواع غنی نمودار است که نمودارهای معمول آماری را پوشش می‌دهند. اما به انعطاف Vega و دیگر کتابخانه‌های نمودار بر اساس دستور زبان گرافیکی نیست، و تنظیم برخی نمودارهای ارتباطی پیچیده آن برای کاربران دشوار است.

(۴) ابزار Leaflet

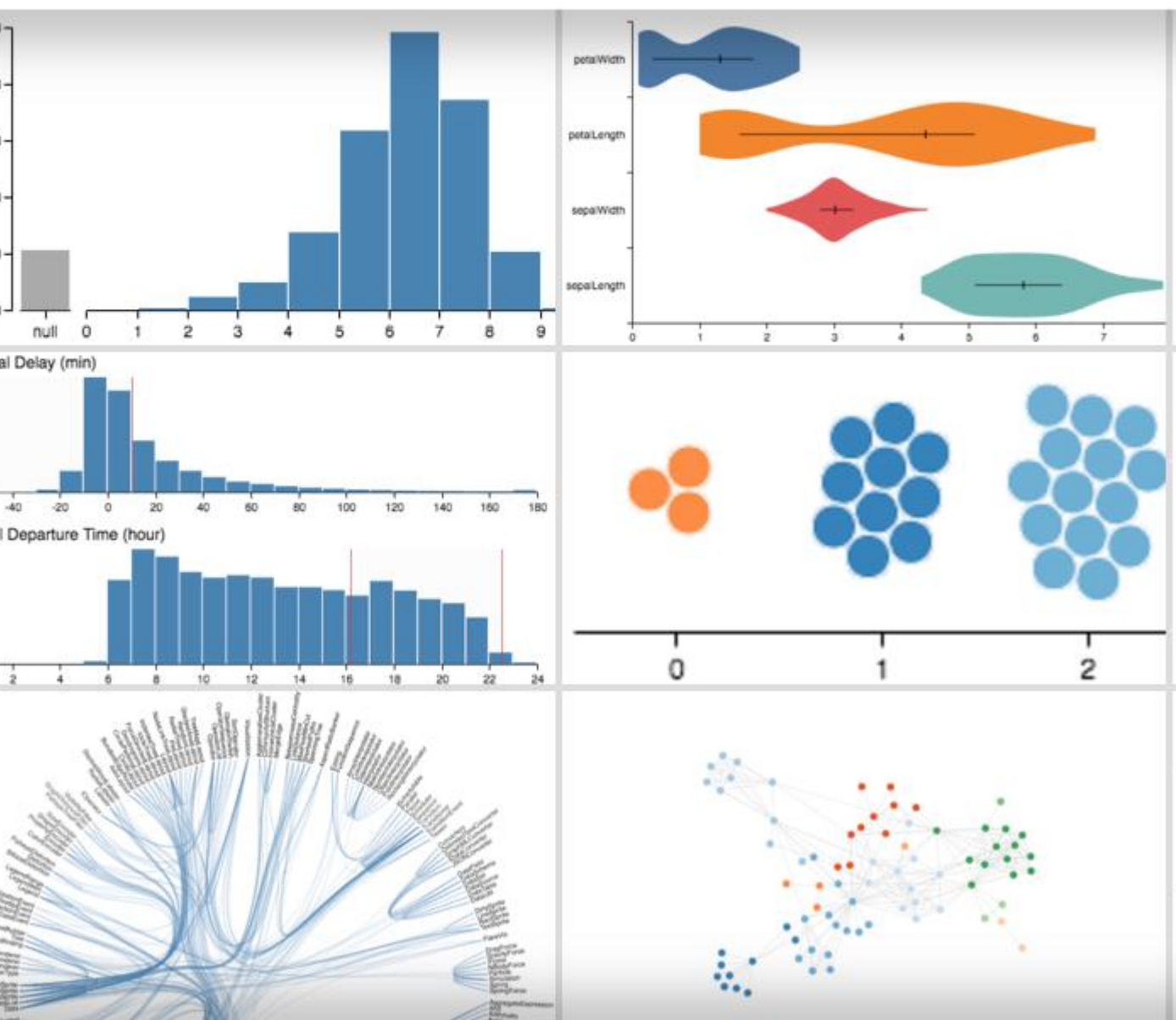
ابزار Leaflet یک کتابخانه جاوااسکریپت از نقشه‌های تعاملی است که برای دستگاه‌های تلفن همراه ساخته شده. همه ویژگی‌های نگاشت مورد نیاز بیشتر توسعه دهندگان را دارد.

ارزیابی: این ابزار می‌تواند به طور خاص برای کاربردهای نقشه مورد هدف قرار گیرد و سازگاری خوبی با موبایل دارد. API از مکانیزم پلاگین پشتیبانی می‌کند اما فعالیتش نسبتاً ساده است. کاربران باید قابلیت‌های توسعه ثانویه داشته باشند.

(۵) ابزار Vega

وگا مجموعه‌ای از گرامرهای گرافیکی تعاملی است که قوانین نگاشت را از داده به گرافیک، گرامرهای تعامل مشترک و عناصر گرافیکی مشترک تعریف می‌کند. کاربران می‌توانند به راحتی گرامرهای وگا را با هم ترکیب کنند تا نمودارهای مختلفی را بسازند.

Grammar



ارزیابی: از آنجایی که وگا اساس بر پایه گرامر JSON ایجاد شده، وگا قوانین نگاشت را از داده‌ها به گرافیک‌ها ارائه می‌دهد و از دستورهای تعاملی مشترک پشتیبانی می‌کند. اما طراحی دستور زبان پیچیده است، و هزینه استفاده و یادگیری بالاست.

۶ ابزار deck.gl

ابزار deck.gl یک کتابخانه کلاس تصویری بر پایه WebGL برای تجزیه و تحلیل داده‌های بزرگ است. این ابزار توسط تیم مصورسازی Uber ایجاد شده است.



ارزیابی: deck.gl بر روی تجسم نقشه سه‌بعدی تمرکز می‌کند. تصویرسازی اطلاعات جغرافیایی داخلی زیادی وجود دارد. این روش از تجسم داده‌های مقیاس بزرگ پشتیبانی می‌کند. اما کاربران باید از WebGL اطلاعات داشته باشند و توسعه لایه‌های آن پیچیده‌تر است.

۷) ابزار [Power BI](#)

ابزار Power BI مجموعه‌ای از ابزارهای تحلیل کسب‌وکار است که بینش‌هایی را در سازمان فراهم می‌کند. این ابزار می‌تواند صدها منبع داده را به هم متصل کند، آماده‌سازی داده‌ها را ساده کند و تحلیل فوری ارائه دهد. سازمان‌ها می‌توانند گزارش‌های تولید شده توسط پاور BI را بر روی دستگاه‌های تلفن همراه و وب مشاهده کنند.

ارزیابی: Power BI شبیه ابزار BI اکسل است، در حالی که عملکرد آن قدرتمندتر از اکسل است. این ابزار از چندین منبع داده پشتیبانی می‌کند. قیمت آن بالا نیست. اما تنها می‌تواند به عنوان یک ابزار BI جداگانه استفاده شود، و هیچ راهی برای ادغام آن با سیستم‌های موجود وجود ندارد.

۸) ابزار Tableau

ابزار Tableau یک ابزار هوش تجاری برای تحلیل بصری داده‌ها است. کاربران می‌توانند داشبوردهای تعاملی و قابل اشتراک ایجاد و توزیع کنند، روندها، تغییرات و تراکم داده‌ها را در نمودار و نمودار نشان دهند. Tableau می‌تواند به فایل‌ها، منابع داده رابطه‌ای و منابع داده بزرگ برای گرفتن و پردازش داده‌ها متصل شود.

ارزیابی: Tableau ساده‌ترین ابزار هوش تجاری در سیستم رومیزی است. این ابزار کاربران را مجبور به نوشتن کد سفارشی نمی‌کند. نرم‌افزار امکان ترکیب داده‌ها و همکاری در زمان واقعی را فراهم می‌کند. اما این ابزار گران است و در خدمات سفارشی سازی و پس از فروش عملکرد ضعیفی دارد.

۹) ابزار FineReport

ابزار FineReport یک ابزار گزارش دهی حرفه‌ای برای وب که به صورت خالص با جاوا نوشته شده است. این سیستم براساس مفهوم «توسعه بدون کد» طراحی شده است. با این گزارش، کاربران می‌توانند گزارش‌های پیچیده و داشبوردهای جذاب ایجاد کنند و یک سکوی تصمیم‌گیری با عملیات‌های ساده کشیدن و انداختن بسازند.

ارزیابی: FineReport می‌تواند به طور مستقیم به تمام انواع پایگاه‌های داده متصل شود و برای سفارشی کردن گزارش‌های پیچیده مختلف و داشبوردهای جذاب مناسب و سریع است. رابط کاربری آن شبیه به رابط اکسل است. این برنامه ۱۹ دسته‌بندی و بیش از ۵۰ سبک از نمودارهای اچ‌تی‌ام‌ال ۵ خود توسعه یافته، با اثرات سه‌بعدی و دینامیکی سرد را فراهم می‌کند. مهم‌ترین چیز این است که نسخه شخصی آن کاملاً رایگان است.

نتیجه‌گیری

مصورسازی داده‌ها یک زمینه بزرگ برای بسیاری از رشته‌ها است. این دقیقاً به خاطر طبیعت میان رشته‌ای حوزه تجسم است که پر از نشاط و فرصت است.