

Visualisation des données : comment obtenir plus de valeur

Les responsables du traitement analytique et des données doivent investir davantage dans le processus de visualisation et d'analyse des données en tant que « parcours analytique » à partir duquel les utilisateurs obtiendront le résultat souhaité.

Introduction

La fourniture d'une expérience de visualisation des données est l'un des attributs fondamentaux des plates-formes modernes d'informatique décisionnelle. Dans une très large mesure, c'est l'essor de la fonctionnalité de visualisation intuitive et interactive des données qui a déclenché et conduit la transition de l'informatique décisionnelle qui s'est opérée sur plusieurs années, passant de la génération de rapports dirigée par l'entité informatique à un traitement analytique moderne et agile, dirigé par l'entreprise.

En 2019, la fonctionnalité de visualisation interactive des données est devenue la nouvelle norme, disponible dans presque toutes les plates-formes modernes d'informatique décisionnelle et de traitement analytique. Il existe même une norme émergente en Europe pour l'utilisation professionnelle de la représentation graphique.

Toutefois, ce n'est pas parce que la visualisation des données est courante dans les entreprises qu'elles en tirent toute la valeur possible. Souvent, ce ne sont que quelques parties des capacités d'un outil qui sont utilisées. Comme toutes les technologies, les plates-formes modernes d'informatique décisionnelle exigent un certain niveau de connaissance des données, mais celui-ci peut être très limité dans certaines entreprises. L'expérience de partage de la visualisation des données est en outre limitée par la gouvernance des données et la politique d'entreprise.

De plus, le contexte de l'application des plates-formes modernes d'informatique décisionnelle s'élargit pour inclure de nouveaux cas d'utilisation, dont la mise en récit des données. À l'instar de bien d'autres domaines, la visualisation des données est influencée positivement par l'essor d'approches intelligentes et assistées qui ont le potentiel d'enrichir et d'étendre l'expérience utilisateur. Cependant, le manque de bonnes pratiques en matière de visualisation des données peut avoir un impact sur le retour sur investissement.

En conséquence, comment les entreprises peuvent-elles maximiser la valeur de la visualisation des données ? En appliquant les pratiques décrites dans le présent article, les responsables du traitement analytique et des données peuvent s'assurer que l'utilisation de la visualisation des données par leur entreprise reste dynamique et ne stagne pas (voir la figure 1).



Formez les utilisateurs au traitement analytique visuel pour obtenir plus de perspectives

Les entreprises qui réussissent avec le traitement analytique et l'informatique décisionnelle et qui atteignent des niveaux de maturité plus élevés offrent une formation multicanaux dans les techniques de traitement analytique. La fourniture de contenu didactique pour aider les utilisateurs à comprendre la visualisation des données doit faire partie des initiatives continues de connaissance du traitement analytique.

Au minimum, les utilisateurs doivent savoir quand et comment utiliser les cartes thermiques, les graphiques en nuage de points et en bulles, les cartes proportionnelles, les diagrammes de Sankey et les graphiques en cascade. Chacun offre un moyen supplémentaire d'explorer et de communiquer les constatations à l'intérieur des données au-delà de celles qui sont disponibles sous forme de barres, courbes, cartes et secteurs. Le processus de création et de modification de diverses visualisations de données ad hoc à partir de jeux de données afin de répondre à des questions opérationnelles sous la forme d'un parcours est la clé pour en tirer de la valeur, en particulier pour les utilisateurs du libre-service.

La plupart des entreprises reconnaissent la valeur de la visualisation des données, en ce sens qu'elle est conviviale et a un impact visuel. Certaines entreprises ayant un faible niveau de connaissance des données demandent à des équipes centralisées ou à des prestataires de services de créer toutes leurs visualisations des données, de sorte que les utilisateurs internes puissent simplement les afficher pour exploiter les informations. Toutefois, cette méthode sous-exploite les outils et les données à leur disposition pour produire une sorte de résultat pro forma. Cela ne doit pas être la destination finale des investissements dans les outils de visualisation des données.

La grande majorité des visualisations de données utilisées aujourd'hui dans le traitement analytique d'entreprise reprend les formes que tout le monde connaît : histogrammes ou graphiques en barres, graphiques en courbes, graphiques en secteurs et cartes géographiques avec superposition de données (généralement des cartes choroplèthes utilisant des zones en couleur).

Cette situation peut paraître surprenante, étant donné que les logiciels de traitement analytique

capables de générer beaucoup plus de formes visuelles existent depuis des décennies. Elle reflète toutefois le fait que les graphiques les plus souvent utilisés sont ceux que les individus connaissent bien et qu'ils ont l'impression de savoir lire. La figure 2 présente les résultats d'une étude qualitative menée par Gartner auprès des fournisseurs en juin 2017, qui met en évidence les formes ou types de graphiques de visualisation des données les plus couramment utilisés.



Pour la plupart des tableaux de bord de gestion descriptifs, ces graphiques bien connus sont adéquats. Cependant, pour une exploration plus diagnostique des données, ou pour des données davantage multidimensionnelles, d'autres types de graphiques conviennent mieux. Par exemple, les cartes proportionnelles (voir l'exemple de la figure 3) sont utiles pour afficher les données hiérarchiques du type couramment utilisé dans les données dimensionnelles des produits.

C'est particulièrement vrai lorsqu'une petite proportion de nœuds contribue le plus à l'ensemble (sinon, un graphique à barres conviendrait). Un autre exemple serait le graphique en nuage de points (ou x/y) montrant deux ensembles de variables et la relation possible entre elles ; ce type de graphique est intuitif, mais encore sous-utilisé.



L'utilisation de graphiques familiers pour la plupart des tableaux de bord est acceptable, dans un premier temps. Toutefois, l'utilisation de ces formes présente le risque que la vue des données par le responsable ou l'analyste ne soit limitée et que les données ne reçoivent pas l'attention nécessaire. Une fois que les utilisateurs maîtrisent mieux le traitement analytique visuel avec plus de formes, ils sont en mesure de fournir plus de perspectives et de construire un état d'esprit analytique.

Utilisez de grands écrans pour que la visualisation des données soit un sport d'équipe

La visualisation des données ne doit pas se limiter à une activité individuelle ; elle doit également porter sur le partage d'équipe, favorisant la collaboration entre les équipes. L'utilisation d'affichages en grand format peut élever la visualisation des données à un autre niveau aussi bien en termes de visibilité que de degré de participation.

Les techniques de conception adaptative font partie intégrante des approches de visualisation des données basées sur HTML5. Elles sont couramment utilisées pour que les tableaux de bord créés sur un système de bureau puissent être affichés sans modification de leur conception sur divers formats d'écran. Cette approche de « conception unique » est généralement associée à la visualisation des données sur des appareils plus petits, tels que les tablettes et smartphones. Toutefois, la valeur potentielle de l'adaptabilité dans l'interface utilisateur va au-delà de la nécessité de « simplifier et réduire » le contenu.

Ce qui était autrefois une perturbation due à la visualisation est aujourd'hui un phénomène

courant. Les deux principales capacités qui utilisent la visualisation des données (les tableaux de bord analytiques et l'exploration visuelle interactive) sont désormais courantes dans les entreprises, ce qui reflète l'évolution de la maturité de l'informatique décisionnelle, de descriptive à diagnostique (voir la figure 4). Les tableaux de bord analytiques et l'exploration visuelle interactive peuvent être utilisés plus efficacement sur des écrans de grand format.



L'utilisation d'interfaces utilisateur adaptatives pour mettre à l'échelle le contenu sur des écrans en grand format permet de générer plus de valeur opérationnelle, du fait que les utilisateurs peuvent partager des récits de données, analyser conjointement et explorer les données en collaboration dans le cadre de réunions de gestion.

Certaines entreprises de biens de consommation courante ont réservé des pièces spécifiquement à cette fin et ont transféré leurs revues de gestion des catégories dans cet environnement afin d'encourager la focalisation sur les données. Il existe maintenant de grands écrans tactiles multipoints qui favorisent davantage encore la collaboration côte à côte et, potentiellement, le développement de contenu (par exemple, la programmation en binôme de style agile où les développeurs de contenu et les utilisateurs de l'entreprise travaillent en tandem).

Les pièces utilisant de grands écrans (même muraux), peut-être composés d'écrans en mosaïque, ont le potentiel de pousser plus loin cette tendance en offrant des expériences quasiment immersives. Elles permettent aux individus de concevoir les données d'une manière qui remet en question les attentes et stimule la discussion. Si vous ne tirez pas encore profit de la mise à l'échelle des formes de présentation dans ces nouvelles arènes, il est temps de le faire.


À l'avenir, ce type d'expérience immersive et analytique se produira probablement (en partie) par le biais de la réalité virtuelle ou, plus probablement, de formes de réalité augmentée. Attendez-vous à de futures notes de recherche sur le sujet.

Favorisez la mise en récit pour les cas d'utilisation à fort impact en utilisant les styles infographiques ciblés

Un certain nombre de plates-formes d'informatique décisionnelle et de traitement analytique prennent désormais en charge des formes de visualisation des données de type infographie pour soutenir les techniques de mise en récit des données en communiquant la signification des données aux groupes décisionnels. Il peut s'agir d'une fonction intégrée ou de bibliothèques de visualisation des données en code source ouvert. Ce type de représentation peut contribuer à attirer l'attention des individus et les aider à se remémorer les données ultérieurement. Le pilotage d'un ou deux cas d'utilisation afférents à l'aide de ces techniques de mise en récit permettra à l'auditoire de s'impliquer davantage dans son propre environnement métier, comme l'analyse des campagnes marketing ou de la planification des ventes.

Les collaborateurs ont ainsi plus confiance pour prendre des décisions fondées sur les données dans leur domaine. Des recherches universitaires publiées en 2015 par l'université Northwestern ont porté sur un type particulier de forme infographique, à savoir le système international d'éducation par les images typographiques (ISOTYPE). L'un des constats est le suivant.

« Lorsque la mémoire de travail est sollicitée sous charge, les données dans les visualisations ISOTYPE sont mémorisées avec plus de précision qu'avec de simples histogrammes. La présentation de visualisations successives de différentes informations (telles que la visualisation des ventes de différents produits ou des préférences alimentaires dans différentes régions) peut bénéficier des visualisations ISOTYPE. Malgré la complexité visuelle supplémentaire, les informations sont remémorées avec une plus grande précision.² »

La figure 5 montre un exemple simple de visualisation des données ISOTYPE. Elle présente des  informations similaires à celles de la figure 2, mais sous forme d'infographie.

En plus de l'exploration visuelle interactive et des tableaux de bord analytiques, les entreprises doivent envisager la mise en récit des données comme troisième pilier de la représentation visuelle. La mise en récit des données vise à communiquer les constats, plutôt qu'à surveiller la progression ou à analyser les causes. Pour ce faire, elle doit étendre la représentation visuelle des données pour être la plus efficace possible et différencier ses résultats des tableaux de bord de gestion standard. Cela exige un ensemble de compétences différent avec une compétence en conception qui peut se trouver à l'extérieur de l'équipe actuelle chargée de l'informatique décisionnelle. Envisagez de recruter ou de sous-traiter ces compétences pour ce type d'élément livrable.

Dans de nombreux cas, les infographies nécessitent l'utilisation de visualisations ou d'extensions personnalisées. Pratiquement toutes les plates-formes d'informatique décisionnelle et de traitement analytique peuvent désormais étendre leurs capacités de visualisation par le biais de l'utilisation de bibliothèques en code source ouvert (la bibliothèque JavaScript D3.js est la plus répandue). L'un des principaux avantages de cette approche est le niveau de contrôle qu'elle offre aux développeurs web pour créer et déployer des visualisations des données extrêmement personnalisées. Ils ne sont en effet pas restreints par les types de graphiques intégrés fournis par les éditeurs eux-mêmes, tout en continuant à utiliser les fonctionnalités de traitement des données de la plate-forme sous-jacente d'informatique décisionnelle et de traitement analytique.

Il y a toutefois un problème potentiel, qui est commun à tous les projets en code source ouvert : le support et la maintenance continus qui sont requis lors de l'utilisation de ces extensions, ainsi que le risque qu'un objet fourni par la bibliothèque cesse de fonctionner avec les mises à niveau des versions. En conséquence, il est préférable d'appliquer les visualisations infographiques à des éléments livrables de projet peu fréquents et très visibles, à titre exceptionnel (par exemple, dans le cadre de l'analyse des décisions de projets stratégiques).

Pour les visualisations de données qui seront utilisées fréquemment et de façon répétée dans le cadre de processus tactiques ou opérationnels de routine, il est préférable de s'en tenir aux formes visuelles intégrées des plates-formes d'informatique décisionnelle et de traitement analytique afin d'éviter ce problème.

Automatisez la visualisation et l'analyse des données à l'aide du traitement analytique augmenté

Les utilisateurs attendent de plus en plus de leurs logiciels de traitement analytique qu'ils en fassent plus pour eux. La capacité à aider automatiquement les utilisateurs à effectuer des analyses fait partie de la capacité de découverte des données augmentée que Gartner inclut dans sa définition d'une plate-forme d'informatique décisionnelle et de traitement analytique.

Les capacités de traitement analytique augmenté sont utiles dans la visualisation des données, lorsque les individus ne savent pas trop quelle forme de graphique est la mieux adaptée au jeu de données qu'ils veulent examiner ou analyser. Elles peuvent en partie contribuer à contrer la tendance des utilisateurs à n'utiliser que des types de graphiques bien connus en créant la visualisation qui leur convient le mieux à partir de toutes les options de visualisation disponibles. En revanche, s'ils ne sont pas en mesure de lire la visualisation qu'ils ne connaissent pas bien, celle-ci n'aura qu'une valeur marginale, de sorte que l'impératif de formation demeure. La combinaison de la visualisation automatique des données avec d'autres technologies favorisées par l'intelligence artificielle (comme la génération automatique de textes [GAT] en langage naturel) décrit les points saillants des formes visuelles (voir l'exemple à la figure 6).



Une description textuelle est utile en tant qu'outil supplémentaire de fourniture d'informations et d'apprentissage. Pour aller plus loin et augmenter encore la valeur analytique, les entreprises doivent étudier la capacité fonctionnelle émergente des plates-formes modernes d'informatique décisionnelle à générer automatiquement un contenu analytique plus avancé au sein des visualisations.

Cela inclut la capacité à identifier et à visualiser les corrélations ou les clusters dans un jeu de données, à afficher les arbres de décision et à mettre en évidence les tendances au moyen d'affichages de régression. Les capacités de visualisation augmentée des données peuvent idéalement être mises à la disposition des utilisateurs qui manifestent un intérêt à devenir scientifiques des données citoyens. Ils peuvent alors développer les compétences nécessaires pour exploiter le résultat autogénéré et communiquer son impact aux autres.

Cette approche générera plus de valeur de l'activité de visualisation des données et aidera à élever le niveau de maturité du traitement analytique plus largement à travers l'entreprise ; au lieu de simplement visualiser les données, les employés pourront mieux les analyser.