|  |  |
| --- | --- |
| PROGRAMA DE FORMACIÓN | Gestión analítica de datos |

|  |  |  |  |
| --- | --- | --- | --- |
| COMPETENCIA | 220501115 - Integrar datos según técnicas de visualización y metodologías de análisis. | RESULTADOS DE APRENDIZAJE | 220501115-04 Generar reportes de acuerdo a herramientas de visualización. |

|  |  |
| --- | --- |
| NÚMERO DEL COMPONENTE FORMATIVO | CF17 |
| NOMBRE DEL COMPONENTE FORMATIVO | Visualización de datos |
| BREVE DESCRIPCIÓN | La visualización de datos es la representación gráfica de la información, mediante el uso de herramientas visuales como gráficos, mapas y tableros de control, permitiendo así una mejor comunicación de los resultados. En la ciencia de datos las técnicas de visualización, son de gran importancia para el análisis de datos y la toma de decisiones. |
| PALABRAS CLAVE | *Dashboards*, gráficos, mapas de calor, tablero de control, visualización. |

|  |  |
| --- | --- |
| ÁREA OCUPACIONAL | Ciencias naturales, aplicadas y relacionadas |
| IDIOMA | Español |

# **Tabla de contenidos**

**Introducción**

1. **Ventajas y beneficios de la visualización de datos**
2. **Despliegue de visualizaciones en Internet**

**Introducción**

|  |
| --- |
| Cuadro de texto |
| Estimado aprendiz, bienvenido a este componente formativo, donde conocerá la fase de visualización de datos y su importancia en los procesos de ciencias de datos para la comunicación de los resultados, así como para la exploración inicial del conjunto de datos disponibles. Aquí conocerá cómo seleccionar los gráficos adecuados, de acuerdo con el tipo de datos disponibles, así como el diseño y creación de tableros de control, que permiten visualizar los resultados más importantes de los proyectos de ciencia de datos.  En el siguiente video conocerá, de forma general, la temática que se estudiará a lo largo del componente formativo. |

**Guion de video introductorio**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tipo de recurso** | Video spot animado | | | |
| **NOTA** | **La totalidad del texto locutado para el video no debe superar las 500 palabras aproximadamente** | | | |
| **Título** | Introducción a la visualización de datos | | | |
| **Escena** | **Imagen** | **Sonido** | **Narración (voz en off)** | **Texto** |
| **1** | Imagen visualizando datos    <https://www.pexels.com/es-es/foto/ordenador-portatil-gris-y-negro-265087/> | Música de fondo | La visualización de datos es la representación gráfica de la información y se ha convertido en una herramienta eficaz para comunicar los resultados de los procesos de ciencias de datos. | Visualización de datos  Representación gráfica de la información |
| **2** | Persona trabajando con herramienta de diseño de gráficas interactivas    <https://www.pexels.com/es-es/foto/mujer-telefono-inteligente-escritorio-oficina-6476566/> | Música de fondo | Existen en la actualidad diversas herramientas que permiten diseñar gráficas interactivas, tableros de control e informes visuales detallados llegando así con mayor claridad al público objetivo y despertando el interés en la información publicada. | Gráficas interactivas  Tableros de control  Informes visuales |
| **3** | Persona programando en un computador    <https://www.pexels.com/es-es/foto/codificacion-de-persona-en-portatil-574071/> | Música de fondo | Este campo de la ciencia de datos combina áreas como la programación, el diseño y la investigación para crear herramientas cada vez más dinámicas y parametrizables logrando así mejores resultados en los procesos de difusión y socialización de los resultados del proyecto. | Programación  Diseño  Investigación |
| **4** | Persona visualizando datos sobre papel o una pantalla    <https://www.pexels.com/es-es/foto/foto-de-primer-plano-de-la-hoja-de-calculo-de-la-encuesta-590022/> | Música de fondo | La visualización de datos permite contar historias con la información analizada y ayuda a entender mejor los datos, destacando tendencias o valores atípicos que solo con la utilización de análisis numéricos serían difíciles de detectar, logrando así un mejor entendimiento de los datos disponibles para su posterior análisis y aprovechamiento. | Permite contar historias |
| **5** | Imagen donde se muestren datos y gráficas    <https://www.pexels.com/es-es/foto/foto-de-primer-plano-del-portatil-gris-577210/> | Música de fondo | Para que este proceso funcione, debe existir un balance entre la forma y la función. Los datos y las gráficas seleccionadas deben trabajar en conjunto para que estos elementos visuales agreguen valor y sean de utilidad para los interesados. | Balance entre la forma y la función |
| **6** | Imagen que muestre histogramas, gráficos, geomapas, o mapas de color.    <https://www.pexels.com/es-es/foto/mujer-negocio-habitacion-altavoz-8761528/> | Música de fondo | Según el tipo de información que se esté procesando y el conocimiento o experticia del público objetivo, se pueden seleccionar diversos tipos de gráficas al momento de presentar la información para representarla mejor. Los histogramas, gráficos de dispersión, geomapas, mapas de calor, diagramas de cajas, nubes de palabras, diagramas de dispersión y escalas de tiempo son algunas de las opciones disponibles en las herramientas que se presentarán en este módulo para la elaboración de visualizaciones de datos más atractivas y entendibles. | Tipos de gráficas:  Histogramas  Gráficos de dispersión  Geomapas  Mapas de calor |
| **7** | Imagen que muestre lenguaje de programación    <https://www.pexels.com/es-es/foto/monitor-negro-que-muestra-textos-270373/> | Música de fondo | Estas herramientas se presentan en algunos casos como librerías para los lenguajes de programación utilizados en los proyectos de ciencias de datos como con *Python* o *R,* y, en otros casos como aplicaciones independientes que reciben los datos a visualizar y permiten la generación de los tableros de control y las gráficas seleccionadas como son *Tableau*, *Power BI* o el mismo *Excel*. | Herramientas:  Librerías para *Python* y *R*  *Tableau*  *Power BI*  *Excel* |
| **8** | Imagen que muestre a un profesional programando    <https://www.freepik.com/free-photo/professional-programmer-working-late-dark-office_5698342.htm> | Música de fondo | En este módulo se conocen los fundamentos de este campo de la ciencia de datos y los criterios para seleccionar los tipos de gráficos adecuados según la información a presentar. Así mismo, las herramientas libres como la librería *GGPLOT* para el lenguaje de programación *R* o *SEABORN* y *PLOTLY* para *Python* y herramientas pagas (aunque en algunos casos con funcionalidades básicas en su plan gratuito de prueba) como son *Power BI*. | Librerías:  *GGPLOT*  *SEABORN*  *PLOTLY* |
| **9** | Imagen que muestre a un docente compartiendo conocimiento al aprendiz.    <https://www.freepik.com/free-photo/team-stockbrokers-are-having-conversation-dark-office-with-display-screens-analyzing-data-graphs-reports-investment-purposes-creative-teamwork-traders_9277155.htm> | Música de fondo | Con estas herramientas, se implementan gráficos estáticos y dinámicos que permitirán explorar con mayor detalle los datos disponibles y extraer de estos, información valiosa. Se aprende cómo descargarlas, configurarlas y utilizarlas en los proyectos de ciencia de datos que esté ejecutando y cómo desplegar los resultados para que sean accesibles. |  |
| **10** | Imagen que muestre al aprendiz creando tableros de control    <https://www.freepik.com/free-photo/businessman-pointing-his-finger-growth-graph_12195356.htm> | Música de fondo | Finalizando el módulo, se pueden crear visualizaciones y tableros de control según sus necesidades y los requerimientos del proyecto de ciencia de datos para publicarlos o distribuirlos con los interesados y facilitar así la difusión de la información procesada y analizada. |  |

|  |  |
| --- | --- |
| **Nombre del archivo** | 228131\_V\_01 |

**Desarrollo de contenidos**

1. **Ventajas y beneficios de la visualización de datos**

|  |  |
| --- | --- |
| Tipo de recurso | Cajón de texto de color |
| En la actualidad, la visualización de datos es un tema fundamental cuando se habla de procesos de ciencia de datos, *Big data* e inteligencia de negocios, porque esta representación gráfica de información y datos, brinda una manera accesible de ver y entender tendencias, identificar valores atípicos y patrones en los datos utilizados. Algunos ejemplos comunes de visualizaciones gráficas son: gráficos de barras, líneas de tiempo, gráficos de línea, diagramas de dispersión, diagramas de Venn, entre otros.  **Figura 1**  *Tipos de visualización de datos*    Nota. Adaptada de ejemplos de gráficas para un *dashboard* ( 2021).  **Imagen:** 228131\_i\_01 | |

|  |  |
| --- | --- |
| **Tipo de recurso** | Acordeón tipo 1 |
| **Introducción** | Las siguientes son algunas de las ventajas al procesar los datos a través de gráficos: |
| Juego de iconos gráficos y gráficos. Gráficos y gráficos. Pie , Línea , Candlestick Chart. Planificación y visualización de estadísticas. Iconos 3d aislados, objetos sobre un fondo transparente  Fuente: <https://www.shutterstock.com/es/image-vector/charts-diagram-icon-set-graphs-pie-2155105343>  **Imagen:** 228131\_i\_02 | |
| **Ahorro de tiempo**  El uso de visualizaciones de datos, en los procesos de ciencia de datos, permite a las organizaciones ahorrar tiempo (y dinero) pues permite la toma de decisiones acertadas y agiliza los procesos de análisis y divulgación. | |
| **Mejor entendimiento de la información**  Las estrategias de visualización de datos, desde etapas tempranas del proceso de ciencia de datos, permiten a los analistas entender mejor los datos disponibles, las posibles relaciones existentes entre variables, la detección de datos atípicos que puedan generar problemas en la interpretación de la información, entre otros. | |
| **Apoyo para la toma de decisiones**  La visualización de datos permite dar manejo a las grandes cantidades de información que tienen que administrar las organizaciones, consiguiendo tener un panorama más amplio y una mejor interpretación de los datos para tomar decisiones acertadas soportadas por información y no por la intuición. | |
| **Optimización de procesos de divulgación de la información**  Las diversas herramientas disponibles para la visualización de datos, permiten con gran facilidad la difusión de las representaciones visuales de datos con todos los actores de la compañía. Esto sumado al uso de las aplicaciones móviles y las aplicaciones web, favorece la movilidad y la interacción en tiempo real. | |
| **Reducción de complejidad / carga de trabajo del equipo**  Gracias a las herramientas de visualización, el análisis de datos deja de ser una tarea exclusiva del equipo tecnológico y cada vez más personas pueden participar de este proceso, pues dichas visualizaciones le permiten a los interesados, utilizar la información de un modo más intuitivo. | |

|  |
| --- |
| Cuadro de texto |
| A medida que los datos se convierten en una parte cada vez más importante de la forma en la que se entiende el mundo, la visualización de datos y las infografías se están convirtiendo en palabras de uso más común, que se refieren a la forma en que se aprenden y se comparten estos datos con los demás. Cabe destacar que las visualizaciones y las infografías no son lo mismo, pero están estrechamente relacionadas.  Las visualizaciones de datos por lo general necesitan algo de narración como complemento o contexto para que los datos se vuelvan procesables y entendibles, lo cual es una de las principales características de las infografías. Por otro lado, ayudan a complementar y soportar las historias contadas en las infografías. Es importante que las visualizaciones de datos creadas sean concisas, claras y convincentes y eso se logra eliminando cualquier elemento extra que pueda interferir con el análisis de la información y seleccionando el tipo de gráfica adecuado. |

|  |  |
| --- | --- |
| **Tipo de recurso** | Infografía estática |
| **Texto introductorio** | A continuación, se presenta una infografía que le permitirá entender la relación entre los datos, su funcionalidad y diseño, para la realización de excelentes visualizaciones de datos: |
| Fuente: <https://es.venngage.com/blog/visualizacion-grafica/>  Nota: | |
| **Código de la imagen** | 228131\_i\_03 |

**Selección de gráficos según el tipo de datos**

|  |  |  |
| --- | --- | --- |
| **Tipo de recurso** | | Pestañas o tabs Verticales |
| **Introducción** | | Existen incontables tipos diferentes de gráficos, tablas y visualizaciones de datos, disponibles para que analistas y científicos de datos puedan representar y transmitir información importante. Los diversos tipos de gráficos sirven para agilizar y mejorar la interpretación de los datos, pero no todos son apropiados para el mismo trabajo y, por esta razón, es importante diferenciar estos y sus usos recomendados.  Para poder decidir de manera acertada, qué tipo de visualización de datos utilizar, se hace necesario comprender las principales funciones que este proceso puede cumplir. A continuación, se mencionan algunas de estas funciones: |
| Ingeniero de software, estadístico, visualizador y analista trabajando en un proyecto. Conferencia de Big Data, presentación de big data, concepto de ciencia de datos. Ilustración aislada del vector violeta brillante brillante brillante brillante  Fuente: <https://www.shutterstock.com/es/image-vector/software-engineer-statistician-visualizer-analyst-working-1199913829>  **Imagen:** 228131\_i\_04 | | |
| **Función 1** | **Comparación de valores**  Se utiliza cuando se desean comparar los valores máximos y mínimos de un conjunto de datos, así como las diferencias entre los valores, brechas existentes y posibles tendencias. Algunas visualizaciones o gráficas adecuadas para esta función son: gráficos de columnas, gráficos de pastel, gráficos de barras y gráficos de dispersión. | |
| **Función 2** | **Mostrar composición**  Esta función permite separar los conjuntos de valores, para mostrar cómo las unidades individuales afectan la imagen general. Por ejemplo, una organización puede necesitar visualizar desde qué ciudades o regiones del país proceden sus ventas o, qué categoría de productos es la más comprada del total de ventas realizadas en un periodo de tiempo. En este caso, son de gran utilidad las gráficas de pastel, gráficos de barras apiladas y gráficas de área. | |
| **Función 3** | **Determinar la distribución**  En muchos casos la organización necesitará comprender la distribución general de sus datos, los intervalos o posibles valores del conjunto de datos y la frecuencia con la que ocurren. Utilizando este tipo de visualizaciones, la organización podrá identificar tendencias además de valores atípicos que pueda estar interfiriendo con dichas tendencias. Para este caso, se recomienda utilizar los gráficos de dispersión. | |
| **Función 4** | **Identificar las relaciones entre diferentes tipos de datos**  Muchas de las variables de las organizaciones se ven afectadas positiva o negativamente por el comportamiento de otras variables (internas o externas a la organización) y es fundamental que dichas relaciones se puedan identificar. Para representar visualmente dicha relación entre variables se pueden usar: gráficos de dispersión y gráficos de líneas. | |

|  |  |  |
| --- | --- | --- |
| **Tipo de recurso** | Slider Imagen | |
| **Introducción** | A continuación, se describen algunos de los tipos de visualizaciones mencionadas y se muestran algunos ejemplos de éstas. Con la práctica, se puede seleccionar, cada vez con mayor facilidad, el gráfico adecuado según el tipo de dato que esté manejando y el público al que se está dirigiendo. | |
| **Tablas**  Las tablas son una de las formas de visualización más empleadas, permiten presentar los datos en filas y columnas para ver detalles y comparar valores. Por lo general, las columnas corresponden a atributos o propiedades del elemento descrito y las filas corresponden a registros independientes de dicho elemento. Este tipo de visualización, permite también realizar filtros para mostrar solo la información de interés, o de ordenar los valores, según algunos de los atributos seleccionados. | | **Imagen:** 228131\_i\_05 |
| **Gráficos de barras**  Es una forma de representar gráficamente un conjunto de datos o variables, mediante el uso de barras rectangulares (verticales u horizontales) con longitudes proporcionales a los valores que representan. Se utilizan como una forma de resumir un conjunto de datos por categorías (variable cualitativa), generalmente ubicadas en el eje horizontal “X” y su frecuencia de aparición ubicada en el eje vertical “Y”. | | **Imagen:** 228131\_i\_06 |
| **Gráficos de líneas**  Son especialmente útiles para mostrar tendencias a lo largo de un periodo de tiempo. Un gráfico de línea, permite mostrar una o varias series de datos en la misma figura, posibilitando la comparación entre sus comportamientos. Se representan dibujando primero, los puntos de los datos según sus valores y luego conectando una línea entre dichos puntos. Se utilizan generalmente, para representar datos continuos. | | **Imagen:** 228131\_i\_07 |
| **Gráfico circular o de torta**  Los gráficos circulares se dividen en sectores donde cada uno muestra el tamaño de un fragmento de información relacionado (el tamaño del sector es proporcional a la cantidad que representa). Por lo general, este tipo de gráficos se utiliza para mostrar tamaños relativos de partes de un todo. Es útil cuando se comparan entre 3 y 5 categorías diferentes con valores diferentes y que sean fáciles de diferenciar. | | **Imagen:** 228131\_i\_08 |
| **Mapas de calor**  Mapa de calor o *heatmap,* es una forma de visualización en formato tabular en donde, además de los números, se utiliza el color de las celdas para transmitir la magnitud relativa de los valores. En los mapas de calor, se puede usar la saturación del color para proporcionar pistas visuales que ayuden a localizar más fácilmente los puntos de interés (valores límites, por ejemplo). Por lo general, se utilizan los colores más oscuros para indicar baja actividad y los colores más brillantes para indicar alta actividad. | | **Imagen:** 228131\_i\_09 |
| **Gráficos de dispersión**  Gráfico de dispersión o *scatterplot,* permite mostrar la relación entre dos variables porque se pueden configurar de tal forma, que una variable se muestre sobre el eje horizontal **X** y, la otra sobre el eje vertical **Y** y, de esta manera, visualizar si existe una relación. Si los elementos graficados forman una franja que se extiende desde la parte inferior izquierda hasta la parte superior derecha, lo más probable es que haya una correlación positiva entre esas dos variables. Si es difícil ver un patrón, es probable que no se tenga una correlación entre las variables graficadas. | | **Imagen:** 228131\_i\_10 |
| ***Treemap***  Un gráfico de árbol, proporciona una visión jerárquica de los datos y facilita la detección de patrones. Por ejemplo, en una tienda, cuáles artículos son los más vendidos. Las ramas del árbol están representadas por rectángulos y cada subrama se muestra como un rectángulo más pequeño. El gráfico de árbol muestra las categorías por color y proximidad y puede mostrar fácilmente, muchos datos que serían difíciles de ver con otros tipos de gráficos. Los gráficos de mapa de árbol son buenos para comparar las proporciones dentro de la jerarquía. Sin embargo, no son buenos para mostrar los niveles jerárquicos entre las categorías más grandes y cada punto de datos. | | Nota. Tomado de <https://support.content.office.net/en-us/media/b3bfff05-380e-4dd9-b49b-7a2bfd128870.png>  **Imagen:** 228131\_i\_11 |
| **Diagrama de caja**  El diagrama de caja o *box plot*, es un tipo de gráfico que permite visualizar y comparar la distribución y la tendencia central de valores numéricos mediante sus cuartiles, los cuales son los tres valores de la variable que dividen a un conjunto de datos ordenados, en cuatro partes iguales. | | Nota. Tomado de [**https://pro.arcgis.com/es/pro-app/latest/help/analysis/geoprocessing/charts/GUID-0E2C3730-C535-40CD-8152-80D794A996A7-web.png**](https://pro.arcgis.com/es/pro-app/latest/help/analysis/geoprocessing/charts/GUID-0E2C3730-C535-40CD-8152-80D794A996A7-web.png)    **Imagen:** 228131\_i\_12 |
| ***Geomaps***  Los gráficos de mapa o *geomaps*, se utilizan, habitualmente, para comparar valores y mostrar categorías entre regiones geográficas. Los gráficos de mapa, son más útiles cuando los datos contienen información geográfica (países, regiones, estados, ciudades, códigos postales, etc.). Utilizando herramientas gratuitas como Google Maps*,* se pueden elaborar visualizaciones de datos geográficos en mapas interactivos y responsivos, teniendo mayor control sobre cómo se ubican los puntos en el mapa. | | Nota. Adapatao de Maptive (2020). <https://www.maptive.com/wp-content/uploads/2020/11/sales-density-covid-19-590x355.jpg>  **Imagen:** 228131\_i\_13 |

|  |
| --- |
| Cuadro de texto |
| Para mayor información relacionada con los tipos de visualizaciones, se recomienda revisar la página de consulta de Microsoft Learn (2022). <https://learn.microsoft.com/es-es/training/> |

**Herramientas de creación de visualización (estáticas e interactivas)**

|  |
| --- |
| Cuadro de texto |
| Para poder sacar partido de las grandes cantidades de información que produce la organización y extraer de estas, conclusiones e *insights* de valor, se hace necesario aprovechar al máximo las herramientas que la ciencia de datos pone a disposición, como la visualización de datos. Estas, permiten interpretar los datos de manera más sencilla y visual, facilitando así su entendimiento y comunicación. La visualización de datos es utilizada en diferentes fases del proceso de ciencia de datos; al principio, como herramienta para explorar y analizar los datos recibidos y empezar a encontrar en estos, relaciones o comportamientos y, en las fases finales, como un instrumento que permite mostrar con mayor facilidad, los resultados obtenidos o poner a disposición de los *stakeholders*, gráficos dinámicos, tablas y controles que pueda manipular para explorar de cuenta propia la información.  Según sea la etapa del proyecto de ciencia de datos o el objetivo que se busca con la visualización, existen diferentes herramientas (gratuitas y pagas) que los científicos de datos y los analistas pueden aprovechar para realizar este proceso. Algunas de estas herramientas permiten realizar gráficas estáticas adecuadas para la generación de reportes, o cuándo la información no cambia en el tiempo. Otras permiten hacer gráficas dinámicas que se pueden editar y que, en algunas ocasiones, toman sus datos de fuentes de información dinámica y, por lo tanto, variable en el tiempo. |

|  |  |  |
| --- | --- | --- |
| **Tipo de recurso** | Slider pasos | |
| **Introducción** | Algunas de estas herramientas para la visualización de datos son: | |
| **Slide 1** | **Excel***:* es un programa de hoja de cálculo, lanzado por la empresa Microsoft en el año de 1987. Es un *software* multiplataforma (funciona en sistemas operativos como Windows, MacOS, Android, iOS), que permite crear tablas, gráficos y otras visualizaciones. Facilita, de igual manera, de una forma sencilla y rápida, crear visualizaciones a partir de los datos existentes y personalizarlos, a gusto del usuario, pero tiene en su contra, el no estar diseñado ni optimizado para manipular grandes cantidades de datos, como puede ser requerido en la mayoría de los proyectos de ciencia de datos.  En sus últimas versiones el número total de filas y columnas en una hoja de cálculo es de 1.048.576 filas por 16.384 columnas. En caso de requerirse más filas (más registros), es necesario usar otra hoja u otro archivo de Excel, volviendo más complejo el procesamiento de la información. | **Imagen:** 228131\_i\_14 |
| **Slide 2** | **Tableau**: es una de las herramientas de visualización más completas y utilizadas que existen en la actualidad. Gracias a su interfaz, permite generar fácilmente visualizaciones a partir de grandes volúmenes de datos. Aunque es una herramienta paga con versiones para móvil, escritorio y soluciones en la nube, Tableau también cuenta con una versión pública gratuita que le permite al aprendiz explorar la herramienta y sacar provecho de la gran variedad de funciones y tipos de visualizaciones que tiene disponibles, además de compartirlas en línea.  La versión académica puede ser consultada en el siguiente enlace: <https://public.tableau.com/app/discover>  Para crear una visualización básica de datos en Tableau*,* se deben seguir los siguientes pasos:   1. **Conectar los datos:** se deben seleccionar y recopilar los datos que se utilizan para la visualización. La plataforma Tableau permite conectar diversas fuentes de datos como hojas de cálculo, bases de datos y archivos planos, los cuales pueden ser visualizados directamente desde la herramienta, para seleccionar los atributos requeridos. 2. **Crear una vista:** una vista es una visualización creada en Tableau, la cual puede ser un cuadro, un gráfico, un mapa o una tabla de texto. La creación de vistas debe partir de la pregunta ¿qué desea saber la organización? 3. **Explorar los datos:** una vez generada la vista, esta puede ser explorada modificando sus parámetros, aumentando o disminuyendo el nivel de detalles y aplicando filtros. | Nota. Tomada de Tableau Desktop (s.f).<https://www.tableau.com/sites/default/files/2022-05/Products_Desktop_Intro.png> **Imagen:** 228131\_i\_15 |
| **Slide 3** | **Power BI:** es una herramienta de inteligencia de negocios, basada en la nube, que permite enlazar diferentes fuentes de datos analizarlos y presentar dicho análisis a través de informes y paneles. Power BI, está disponible a través de aplicación de escritorio, como SaaS (*Software as a Service*) y como aplicación móvil para teléfonos y tabletas.  Entre las muchas ventajas de la plataforma Power BI, se destacan:   * Su interfaz intuitiva permite visualizar datos e interpretarlos de una forma fácil y ágil. * Se integra fácilmente con plataformas de la empresa Microsoft como son Sharepoint, Office 365 y Dynamics 365. * El fabricante de la plataforma Microsoft garantiza la seguridad y privacidad de los datos, ofreciendo controles de accesibilidad a nivel interno y externo.   ,  La plataforma es ofrecida con diferentes tipos de licencias: licencia gratuita (Power BI Desktop y Power BI Mobile), Licencia Pro y Licencia Premium (estas últimas dos con costo). La opción gratuita, permite recopilar y analizar datos hasta de 70 fuentes diferentes y ofrece las mismas visualizaciones que la opción paga con la opción de exportar los reportes realizados en diferentes formatos. El principal limitante de esta licencia, es que no permite compartir las visualizaciones directamente con otros usuarios. | Nota. Tomada de Khanna (s.f). [*https://powerbicdn.azureedge.net/cvt-9394a7615e76e3565238e1688022efbee201671e2cbb78c8d1b1df694b985908/pictures/pages/index/blade2/powerbi-001.jpg*](https://powerbicdn.azureedge.net/cvt-9394a7615e76e3565238e1688022efbee201671e2cbb78c8d1b1df694b985908/pictures/pages/index/blade2/powerbi-001.jpg)  **Imagen:** 228131\_i\_16 |
| **Slide 4** | **Google Charts:** proporciona una manera eficaz de visualizar datos en un sitio web. Desde gráficos de líneas simples hasta mapas de árboles jerárquicos complejos, la galería de gráficos de GoogleCharts proporciona una gran cantidad de tipos de gráficos listos para usar. La forma más común de usar GoogleChartses con JavaScript, incrustado en la página web, cargando algunas bibliotecas de gráficos de Google, enumerando los datos que se van a graficar, seleccionando opciones para personalizar el gráfico y, finalmente, creando un objeto de gráfico con la identificación que elija. Luego, más adelante en la página web, creando un <div> con esa identificación para mostrar el gráfico de Google.  El código debe ser incluido en un archivo .html y visualizado desde un navegador web.  <html>  <head>  <!--Load the AJAX API-->  <script type="text/javascript" src="https://www.gstatic.com/charts/loader.js"></script>  <script type="text/javascript">  // Carga la API de visualización y el paquete corechart.  google.charts.load('current', {'packages':['corechart']});  // Set a callback to run when the Google Visualization API is loaded.  google.charts.setOnLoadCallback(drawChart);  // Función que crea y rellena una tabla de datos,  // instancia el gráfico circular, pasa los datos y  // lo dibuja.  function drawChart() {  // Creación de tabla de datos.  var data = new google.visualization.DataTable();  data.addColumn('string', 'Topping');  data.addColumn('number', 'Porciones');  data.addRows([  ['Champiñones', 3],  ['Cebolla', 1],  ['Aceitunas', 1],  ['Zucchini', 1],  ['Pepperoni', 2]  ]);  // Configuración de las opciones del gráfico  var options = {'title':'¿Cuánta pizza comí anoche?',  'width':400,  'height':300};  // Instanciar y dibujar el gráfico, pasando algunas opciones.  var chart = new google.visualization.PieChart(document.getElementById('chart\_div'));  chart.draw(data, options);  }  </script>  </head>  <body>  <!--Div that will hold the pie chart-->  <div id="chart\_div"></div>  </body>  </html>  Se incluye el siguiente enlace para descarga directa del archivo en html con el código. (Archivo: chart.html).  <https://developers-dot-devsite-v2-prod.appspot.com/chart/interactive/docs/quick_start>  El gráfico generado con este código, es interactivo (el usuario puede ubicar el puntero del mouse sobre las zonas del gráfico o las etiquetas para ver más información).  Para cambiar el tipo de gráfico y usar por ejemplo uno de barras, solo se debe modificar la línea google.visualization.PieChartdonde se especifica que es una gráfica circular y usar google.visualization.BarChart, para usar un gráfico de barras. La selección del tipo de gráfico dependerá del tipo de información a presentar, buscando que las visualizaciones seleccionadas faciliten la comprensión de la información.  Para conocer más de esta herramienta, consultar el sitio web Google Charts (<https://developers-dot-devsite-v2-prod.appspot.com/chart>). En la parte superior derecha se puede modificar el idioma del sitio web para facilitar su consulta. | **Imagen:** 228131\_i\_17    **Imagen:** 228131\_i\_18 |
| **Slide 5** | **Qlik Desktop:** es una aplicación para el sistema operativo Windows*,* que permite crear visualizaciones, gráficos, cuadros de mando interactivos (*dashboards*) e informes para uso local y sin conexión. Esta herramienta, está desarrollada con el lenguaje de programación R, lenguaje de código abierto usado para el análisis estadístico y procesos de ciencias de datos.  Su aplicación va desde los temas relacionados con finanzas y recursos humanos, hasta de marketing y ventas. Procesos que diariamente generan datos valiosos y susceptibles de ser explorados y aprovechados para hacer más competitiva la organización. | Nota. Tomada de QliK (2017). <https://learning.qlik.com/pluginfile.php/98829/mod_resource/content/9/Create_Visualizations/Create_Visualizations.html>  **Imagen:** 228131\_i\_19 |
| **Slide 6** | **Plotly:** es una librería de código abierto, disponible para varios lenguajes entre ellos Python, Java, R y Matlab, entre otros. Por lo general, es más utilizada en Pythonal ser este uno de los lenguajes más populares para desarrollo, relacionados con *machine learning* e inteligencia artificial. Esta librería, soporta más de 40 tipos diferentes de gráficos para uso estadístico, financiero, geográfico y científico.  Plotlypermite a los usuarios de Python crear visualizaciones interactivas basadas en web que pueden ser usadas en un cuaderno Jupyter, guardadas como archivos HTML independientes o incluidas en aplicaciones web construidas con Python y la librería DASH.  Permite también la creación de gráficos de mapas usados en proyectos de ciencias de datos cuando se tienen datos geográficos; por ejemplo, para representar frecuencias (mapas térmicos). | Nota. Tomada de Geomapik (2019). [**http://www.geomapik.com/wp-content/uploads/2019/08/Screenshot\_20190806\_110520.png**](http://www.geomapik.com/wp-content/uploads/2019/08/Screenshot_20190806_110520.png)  **Imagen:** 228131\_i\_20    **Imagen:** 228131\_i\_21 |
| **Slide 7** | **Looker Studio:** conocida anteriormente como “Google Data Studio”, Looker Studio es una herramienta que permite conectar fuentes de datos para crear cuadros de mando interactivos (*dashboards*) orientada a la inteligencia de negocios y a la toma de decisiones a nivel empresarial. | Nota. Tomada de Google Analytics Marketing Website (s.f). <https://datastudio.google.com/u/0/reporting/0B_U5RNpwhcE6SF85TENURnc4UjA/preview/>  **Imagen:** 228131\_i\_22 |

**Creación de *dashboards***

|  |  |
| --- | --- |
| **Tipo de recurso** | Acordeón tipo 2 |
| **Introducción** | Un *dashboard* (o tablero de control) es una herramienta para la gestión de la información, que permite monitorear y analizar, de forma visual y centralizada, los indicadores claves de desempeño (KPIs), métricas y demás datos importantes para la organización. Es entonces, una especie de “resumen” que recopila los principales datos y los presenta en *stakeholders* para que se pueda consultar solo la información necesaria. Entre las principales características de un *dashboard* se encuentran: |
| Interfaz de usuario de paneles, UX, KIT, gran diseño para cualquier propósito del sitio. Plantilla de infografía de negocios. Ilustración plana vectorial. Concepto de gran tamaño Diseño de plantillas de panel de administración de usuarios. Panel de administración de análisis.  Fuente: <https://www.shutterstock.com/es/image-vector/dashboard-ui-ux-kit-great-design-1938829057>    **Imagen:** 228131\_i\_23 | |
| **Personalización**  Los *dashboards* muestran únicamente la información relevante para el departamento o el *stakeholder* que la recibe. Una organización podrá tener varios tableros de control según sus áreas o proyectos y el contenido de estos será diferente según las necesidades y preguntas a responder. | |
| **Visual**  Los tableros de control presentan su información de una forma visual y atractiva aprovechando la diversidad de gráficas y distintivos existentes (flechas, colores, cifras resaltadas) para facilitar al usuario la comprensión y aprovechamiento de la información allí presentada. | |
| **Tiempo real**  En algunos casos, los *dashboards* están conectados directamente con fuentes de información activas (las cuales continúan recibiendo datos) permitiendo que la información presentada esté actualizada y mostrando los valores de los indicadores en tiempo real. | |

|  |  |  |
| --- | --- | --- |
| **Tipo de recurso** | | Pestañas o tabs verticales |
| **Introducción** | | A continuación, se presentan los pasos sugeridos para la creación de un *dashboard* útil y eficaz para la organización. |
| Plantilla de panel de información. Diseño azul verde simple de la interfaz, panel administrador con gráficos, diagramas de gráficos. Infografías de vectores  Fuente: <https://www.shutterstock.com/es/image-vector/infographic-dashboard-template-simple-green-blue-1222098511>  **Imagen:** 228131\_i\_24 | | |
| **Establecer el objetivo del *dashboard*** | Para iniciar la creación del tablero de control es necesario saber cuál es el objetivo de éste, qué pregunta (o preguntas) de la organización queremos responder. Por ejemplo, ¿la organización quiere analizar las ventas para mejorarlas e incrementarlas para mejorar los ingresos?, ¿se quiere hacer seguimiento al avance y ejecución de los proyectos? | |
| **Identificar el público objetivo** | Es importante tener claridad sobre el objetivo del *dashboard* a diseñar. Cada área de la organización y sus colaboradores, tendrá diferentes conocimientos e intereses, y si esto no se tiene en cuenta, se corre el riesgo de presentar en ellos información diferente a la requerida o en términos que no sean fáciles de entender. | |
| **Definir las métricas** | Se deben seleccionar las principales métricas que ayuden a dar respuesta a las preguntas formuladas o que estén relacionadas. Si el indicador seleccionado no aporta a la temática en cuestión, posiblemente genere ruido en la visualización y deba ser eliminado. Entre los indicadores seleccionados es importante incluir los KPI *(*Key Performance Indicators*)* o indicadores claves de desempeño, utilizados por la organización, para evaluar su éxito en alcanzar diversos objetivos. | |
| **Diseñar e implementar el *dashboard*** | Una vez definidas las métricas e información a presentar, y teniendo en cuenta la herramienta a emplear, se definirán las herramientas gráficas adecuadas según el tipo de información y el público objetivo. Es muy importante, antes de publicar el acceso al tablero de control, validar la calidad y veracidad de la información presentada, confirmar que las visualizaciones seleccionadas se muestran correctamente y que la información presentada responde a los objetivos definidos en el primer paso. | |

|  |  |
| --- | --- |
| **Tipo de recurso** | Infografía estática |
| **Texto introductorio** | Para seleccionar la gráfica adecuada según el tipo de variable(s), se sugiere revisar la siguiente infografía: |
| Fuente: <https://extremepresentation.com/wp-content/uploads/6a00d8341bfd2e53ef0148c699cc96970c.jpg> | |
| **Código de la imagen** | 228131\_i\_25 |

**Tipos de fuentes de información según su origen**

|  |  |
| --- | --- |
| **Tipo de recurso** | Rutas / Pasos. Verticales 1 |
| **Introducción** | Los datos a utilizar en los proyectos de ciencias de datos, pueden tener diferentes orígenes o fuentes, siendo necesario para cada una, utilizar diversas metodologías de captura y procesamiento, según sean estructurados o no. Los datos estructurados son aquellos que tienen un formato definido y específico (como el que se encuentra en las bases de datos), mientras que los datos no estructurados no tienen una estructura o formato fijo haciendo más complicada su gestión (cabe anotar que más del 80 % de los datos de una organización son de este tipo).  Según su origen, se pueden clasificar las fuentes de información como: |
| Trabajadores de oficina que organizan el almacenamiento de datos y el archivo de archivos en el servidor o la computadora. ilustración de dibujos animados  Fuente: <https://www.shutterstock.com/es/image-vector/infographic-dashboard-template-simple-green-blue-1222098511>  **Imagen:** 228131\_i\_26 | |
| **Botón 1** | **Biométricas:** hace referencia a la identificación automática de personas basada en sus características anatómicas (huella, retina, ADN). |
| **Botón 2** | **Máquina a máquina:** son aquellos datos compartidos directamente entre dos o más dispositivos como lo permite el Internet de las cosas (IoT). Por ejemplo, los *chips* NFC (Near Field Communication) que permiten la transmisión de datos de forma inalámbrica y segura con dispositivos como el teléfono celular para la realización de pagos sin contacto. |
| **Botón 3** | **Datos de transacciones:** son los datos que se generan y registran en los procesos de venta y facturación de la organización y demás operaciones habituales. Pueden ser registrados usando *software* diseñado para tal fin, o todavía en algunos casos, de forma física y manual. |
| **Botón 4** | **Generados por los humanos:** en esta categoría se incluyen las comunicaciones realizadas por las personas usando diferentes canales: llamadas, grabaciones, servicio de atención al cliente, correos electrónicos, registros médicos y demás. |
| **Botón 5** | **Medios sociales:** son aquellos datos que se originan en la red y los medios digitales disponibles como redes sociales (Facebook, Instagram, YouTube, Twitter, LinkedIn), buscadores (Google, Bing, Yahoo) y sitios web. |

1. **Despliegue de visualizaciones en Internet**

|  |
| --- |
| Cuadro de texto |
| La fase de despliegue en un proyecto de ciencia de datos, se refiere al momento en el que se lleva el modelo de *machine* *learning*, el *dashboard* o el informe final, a un ambiente de producción, para que pueda ser accedido y utilizado por los *stakeholders,* dándole así valor al trabajo realizado. Esta fase de despliegue puede ser tan sencilla, como la generación de un informe para comunicarlo y compartirlo con los interesados, o tan compleja, que implique la implementación de un entorno independiente, donde esté alojada la solución y que tenga los controles y permisos adecuados para que solo las personas autorizadas puedan acceder. Por lo general, este despliegue es realizado por equipos de Tecnologías de la información TI, que se encargan de configurar y administrar la infraestructura requerida para el proceso.  Cada vez es más importante que los científicos de datos se familiaricen con el proceso requerido para el despliegue de sus proyectos y que tengan la capacidad de adaptar el desarrollo de su modelo o visualización a una arquitectura de *software* con los requerimientos necesarios para producción. Los científicos de datos, pueden, además de diseñar sus modelos, implementarlos y desplegarlos por cuenta propia, porque ellos tienen una enorme ventaja competitiva frente a los que no. |

|  |  |  |
| --- | --- | --- |
| **Tipo de recurso** | Rutas / Pasos. Verticales 2 | |
| **Introducción** | Muchas organizaciones inician sus procesos o proyectos de ciencias de datos, sin tener claro un plan que puedan pasar, finalmente, a producción. Esto genera grandes complicaciones al momento del despliegue, pues no se tienen en cuenta elementos tan importantes como la capacidad de cómputo requerida o el espacio de almacenamiento y la seguridad, siendo muy costoso el proceso de implementación.  Algunos elementos a tener en cuenta desde el principio del proyecto son: | |
| **1** | **Datos (cantidad, uso y almacenamiento):** los procesos de ciencia de datos dependen de los datos y durante su ejecución, se utilizan para entrenar, evaluar, probar o predecir. Debe, entonces, la persona encargada del proceso, preguntarse entre otras cosas: ¿qué tamaño tienen los datos a utilizar?, ¿cómo se van a obtener?, ¿dónde se van a almacenar?, ¿si los datos serán modificados? Todas estas preguntas ayudan a definir elementos tan importantes como las herramientas a utilizar (por ejemplo, según el tamaño o la cantidad de datos se pueden descartar herramientas básicas o que no estén optimizadas para gestionar grandes volúmenes de información como puede ser Excel). | Ilustración de concepto informe de datos  Fuente: <https://www.freepik.es/vector-gratis/ilustracion-concepto-informe-datos_6195527.htm#query=datos&position=0&from_view=search&track=sph>    **Imagen:** 228131\_i\_27 |
| **2** | **Herramientas (*frameworks*, librerías, lenguajes)**: son utilizadas para diseñar e implementar los proyectos de ciencias de datos y requieren de una combinación de *frameworks* y herramientas adecuadas y, de igual manera, *hardware* y *software,* que permitan desarrollar e implementar los modelos y/o visualizaciones. Para una adecuada selección, se hace necesario investigar y comparar los resultados del uso de diferentes herramientas utilizadas para una misma labor, teniendo en cuenta su eficiencia, soporte y popularidad. Esta selección, también impacta el proceso de despliegue porque se debe evaluar si las elegidas son compatibles con las plataformas o servicios disponibles (entornos en la nube, tales como AWS (Amazon WebServices), Azure y GCP (Google CloudPlatform), qué sistema operativo requieren y si son de código abierto o no. | Servicio de mantenimiento de servidores. transferencia de información, configuración de hardware. idea de servidor de red. tecnología de hospedaje, almacenamiento de bases de datos, equipos de programación. ilustración de metáfora de concepto aislado de vector  Fuente: <https://www.freepik.es/vector-gratis/servicio-mantenimiento-servidores-transferencia-informacion-configuracion-hardware-idea-servidor-red-tecnologia-hospedaje-almacenamiento-bases-datos-equipos-programacion-ilustracion-metafora-concepto-aislado-vector_12470270.htm#query=hardware%20software&position=12&from_view=search&track=sph>  **Imagen:** 228131\_i\_28 |

**Servicios para despliegue de *dashboards* en Internet**

|  |
| --- |
| Cuadro de texto |
| En la actualidad, el almacenamiento y la gestión de datos en la nube permite a las organizaciones almacenar, compartir y sincronizar archivos y aplicaciones de forma segura y rápida. El uso de la computación en la nube permite a los usuarios y organizaciones escalar los servicios que utilizan según sus necesidades, acceder a estos servicios desde cualquier lugar y personalizar las aplicaciones y contenidos allí almacenados (flexibilidad). Además, posibilita desplegar las soluciones y aplicaciones rápidamente para su uso y aprovechamiento, sin que la organización tenga que preocuparse por los costos asociados al adquirir y mantener una infraestructura tecnológica, que algunas veces no es aprovechada en su totalidad y, otras, no es suficiente para dar respuesta a los requerimientos de los proyectos (eficiencia).  A continuación, se presentan algunos de los servicios y herramientas más utilizados en la actualidad: |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tipo de recurso** | Video spot animado | | | |
| **NOTA** | **La totalidad del texto locutado para el video no debe superar las 500 palabras aproximadamente** | | | |
| **Título** | Servicios de computación en la nube | | | |
| **Escena** | **Imagen** | **Sonido** | **Narración (voz en off)** | **Texto** |
| **01** | Persona en un computador viendo almacenamiento en la nube    <https://img.freepik.com/free-photo/saas-concept-collage_23-2149399294.jpg> | Música de fondo | Los servicios de computación en la nube son infraestructuras, plataformas o sistemas de *software* que alojan los proveedores externos y que se ponen a disposición de los usuarios a través de Internet. Los siguientes son algunos ejemplos de estos servicios: | Servicios de computación en la nube:    Infraestructuras  Plataformas  Sistemas de *software* |
| **02** | Imagen que muestre modelo de computación en la nube  Concepto de servicio en la nube con el dedo del hombre tocando la pantalla digital con iconos de aplicación de servicio en la nube en el fondo abstracto de la ciudad  <https://www.shutterstock.com/es/image-photo/cloud-service-concept-man-finger-touching-1912951960> | Música de fondo | *AWS (Amazon Web Services)* es un proveedor de servicios en la nube, que permite disponer de almacenamiento, recursos de computación, aplicaciones móviles, bases de datos y demás servicios de computación en la nube.    AWS, lanzó sus primeros servicios *web* en 2002 a partir de la infraestructura interna que la empresa *Amazon* construyó para gestionar sus operaciones de venta *online*.    En 2006, comenzó a ofrecer sus servicios definidos como SaaS de IaaS (*Infraestructure as a Service*). AWS, fue una de las primeras empresas en introducir un modelo de computación en la nube de pago por uso, que se amplía para proporcionar a los usuarios computación, almacenamiento o rendimiento según sus necesidades.  AWS, tiene presencia en 190 países alrededor del mundo y cuenta con más de 10.000 *partners* asociados. | Amazon Web *Services*    Hace parte de la empresa *Amazon*    Lanza sus primeros servicios en 2002 |
| **03** | Persona en un computador accediendo a los servicios de *Google Cloud*  Servidor de nube de datos de analista en ilustración isométrica  <https://www.freepik.es/vector-premium/servidor-nube-datos-analista-ilustracion-isometrica_20899560.htm#query=cloud&from_query=google%20cloud&position=15&from_view=search&track=sph> | Música de fondo | GCP (*Google Cloud Platfom*) es un conjunto de servicios de computación en la nube pública ofrecido por *Google*. La plataforma incluye una serie de servicios alojados para la computación, el almacenamiento y el desarrollo de aplicaciones que se ejecutan en el *hardware* de *Google*.    Los desarrolladores de *software*, los administradores de la nube y otros profesionales de TI de las organizaciones, pueden acceder a los servicios de *Google Cloud* a través de la Internet pública o de una conexión de red dedicada. | *Google Cloud Platform*    Servicio ofrecido por Google    Incluye:    Servicios alojados en la nube    Almacenamiento |
| **04** | Persona en un pc utilizando la plataforma *Microsoft Azure*  POZNAN, POL - 12 DIC. 2021: Equipo portátil que muestra el logotipo de Microsoft Azure, un servicio de computación en nube para la creación, prueba, implementación y administración de aplicaciones y servicios  <https://www.shutterstock.com/es/image-photo/poznan-pol-dec-12-2021-laptop-2121429641> | Música de fondo | *Azure* (*Microsoft Azure*), antes conocida como *Windows Azure*, es la plataforma de computación en la nube pública de *Microsoft*. Ofrece una amplia gama de servicios en la nube, como computación, análisis, almacenamiento y redes. Los usuarios pueden elegir entre estos servicios para desarrollar y escalar nuevas aplicaciones o ejecutar aplicaciones existentes en la nube pública.    La plataforma *Azure* tiene como objetivo ayudar a las organizaciones a gestionar los retos y cumplir sus objetivos organizativos. Ofrece herramientas que dan soporte a todos los sectores y es compatible con las tecnologías de código abierto. Esto da a los usuarios la flexibilidad de utilizar sus herramientas y tecnologías preferidas.    *Azure* ofrece cuatro formas diferentes de computación en la nube: infraestructura como servicio (IaaS), plataforma como servicio (PaaS), *software* como servicio (SaaS) y funciones sin servidor. | *Microsoft Azure*    Ofrece una amplia gama de servicios en la nube    Tiene servicios de:    Infraestructura como servicio (IaaS)    Plataforma como servicio (PaaS)    *Software* como servicio (SaaS)  Funciones sin servidor |
| **05** | Persona en un *pc* programando  Sistema de código fuente Concepto de código abierto PHP  <https://www.shutterstock.com/es/image-photo/source-code-system-php-open-concept-395931805> | Música de fondo | *Docker*, es una plataforma de código abierto que permite a los desarrolladores crear, desplegar, ejecutar, actualizar y gestionar contenedores, es decir, componentes estandarizados y ejecutables que combinan el código fuente de las aplicaciones con las bibliotecas del sistema operativo (SO) y las dependencias necesarias para ejecutar ese código en cualquier entorno.    Los contenedores simplifican el desarrollo y la entrega de aplicaciones distribuidas. Se han vuelto cada vez más populares a medida que las organizaciones cambian al desarrollo nativo de la nube y a entornos híbridos de múltiples nubes. Los desarrolladores, pueden crear contenedores sin *Docker,* trabajando directamente con las capacidades integradas en *Linux* y otros sistemas operativos. *Docker* hace que la creación de contenedores sea más rápida, fácil y segura. | *Docker*    Plataforma de código abierto    Permite crear, desplegar, ejecutar, actualizar y gestionar    Contenedores |
| **Nombre del archivo** | 228131\_V\_02 | | | |

**Proceso de despliegue de *dashboards***

|  |  |  |
| --- | --- | --- |
| **Tipo de recurso** | Slider pasos | |
| **Introducción** | A continuación, se presentan los pasos sugeridos para el proceso de despliegue de proyectos de ciencias de datos o de *dashboards*, para la presentación de los resultados. | |
| **Paso 1** | **Selección del entorno de trabajo adecuado:** según las herramientas utilizadas en la ejecución del proyecto de ciencia de datos, se debe elegir la plataforma correspondiente para su despliegue. Se puede optar por usar las soluciones presentadas en el numeral anterior (AWS, GCP, Azure) o también la infraestructura propia de la organización, si esta da respuesta a las necesidades del proyecto y cumple, además, con los requerimientos de seguridad y accesibilidad. | ilustración vectorial, asistente en línea en el trabajo. promoción en la red. gerente de trabajo remoto, buscando nuevas ideas soluciones, trabajando juntos en la empresa, brainstorming  Fuente: <https://www.shutterstock.com/es/image-vector/vector-illustration-online-assistant-work-promotion-1110532412>  **Imagen:** 228131\_i\_29 |
| **Paso 2** | **Configuración del *hardware*:** se debe configurar el entorno de trabajo seleccionado para que cumpla con los requerimientos del proyecto a desplegar. Esto incluye, aspectos como el sistema operativo que soporte la aplicación, librerías requeridas, la configuración de acceso a dicho entorno, los permisos y roles. | Fuente: <https://www.freepik.es/vector-gratis/idea-almacenamiento-nube-computacion-linea-base-datos-internet-servidor-respaldo-equipo-programacion-acceso-limitado-pase-control-configuracion-privacidad-ilustracion-metafora-concepto-aislado-vector_12470272.htm#query=configuraci%C3%B3n%20de%20hardware&position=2&from_view=search&track=sph>  **Imagen:** 228131\_i\_30 |
| **Paso 3** | **Instalación / migración de la aplicación:** se alojan en el entorno ya configurado los archivos relacionados con el proyecto de ciencia de datos, los modelos y/o los informes o *dashboards* listos para su uso y explotación por parte de los interesados. En este caso, solo se usan los archivos finales, las bases de datos a usar y los modelos optimizados omitiendo conjuntos de datos de entrenamiento o archivos no requeridos para el uso de la aplicación o de la visualización. | Ilustración del concepto de estadísticas del sitio  Fuente: <https://www.freepik.es/vector-gratis/ilustracion-concepto-estadisticas-sitio_7140739.htm#query=dashboard&position=4&from_view=search&track=sph>  **Imagen:** 228131\_i\_31 |
| **Paso 4** | **Pruebas:** se debe verificar que la aplicación o la visualización funcione correctamente en el entorno seleccionado. Para esto, se ejecuta un plan de pruebas que abarque aspectos funcionales y de diseño, validando que se cumplan con todos los requerimientos iniciales del proyecto y que el resultado es el esperado por los *stakeholders*. | Empresario isométrico con lista de verificación y lista de tareas. portapapeles con una lista de verificación. gestión de proyectos, planificación y mantenimiento de la puntuación del concepto de tareas completadas.  Fuente: <https://www.freepik.es/vector-premium/empresario-isometrico-lista-verificacion-lista-tareas-portapapeles-lista-verificacion-gestion-proyectos-planificacion-mantenimiento-puntuacion-concepto-tareas-completadas_22992900.htm#query=pruebas&position=22&from_view=search&track=sph>  **Imagen:** 228131\_i\_32 |

|  |
| --- |
| Recuerde explorar los demás recursos que se encuentran disponibles en este componente formativo; para ello, diríjase al menú principal, donde encontrará la síntesis, una actividad didáctica para reforzar los conceptos estudiados, material complementario, entre otros. |

**Síntesis**

|  |  |
| --- | --- |
| **Tipo de recurso** | Síntesis |
|  | |
| **Introducción** | En el siguiente mapa conceptual se encuentran, de manera sintetizada, todos los temas vistos y aprendidos en este componente formativo. Le servirá como guía para recordar cada contenido y así cumplir de manera satisfactoria las respectivas actividades a desarrollar. |
|  | |

**Actividad didáctica**

|  |  |  |
| --- | --- | --- |
| **Tipo de recurso** | Actividad didáctica. Verdadero y falso | |
| Esta actividad permite determinar el grado de apropiación de los contenidos abordados en este recurso de aprendizaje. Lea la afirmación de cada ítem y señale si es verdadera o falsa según corresponda. | | Data inform illustration concept    **Fuente:** <https://img.freepik.com/free-vector/data-inform-illustration-concept_114360-864.jpg>  **Imagen:** 228131\_i\_33 |
| 1. La visualización de datos brinda una manera accesible de ver y entender tendencias, identificar valores atípicos y patrones en los datos utilizados. | | Online video training or virtual conference with people characters cartoon vector  **Fuente:** <https://img.freepik.com/premium-vector/online-video-training-virtual-conference-with-people-characters-cartoon-vector_605858-171.jpg>  **Imagen:** 228131\_i\_34 |
| **Verdadero (Correcto)**  Retroalimentación:  ¡Correcto! La visualización de datos es un tema fundamental cuando se habla de procesos de ciencia de datos, *big data* e inteligencia de negocios pues ayuda a comprender mejor la información y los resultados. | | Falso  Retroalimentación:  ¡Incorrecto! Se recomienda revisar nuevamente la sección de: ventajas y beneficios de la visualización de datos. |
| 2. Algunos ejemplos comunes de visualizaciones gráficas son: gráficos de barras, líneas de tiempo, gráficos de línea, diagramas de dispersión, diagramas de *Venn*. | | Set of business infograph vectors  **Fuente:** <https://img.freepik.com/free-vector/set-business-infograph-vectors_53876-64601.jpg>  **Imagen:** 228131\_i\_35 |
| **Verdadero (Correcto)**  Retroalimentación:  ¡Correcto! Estos tipos de gráficos, así como otros más complejos como los *geomaps* y los *dashboards* son ejemplos de herramientas que se pueden utilizar en la visualización de datos. | | Falso  Retroalimentación:  ¡Incorrecto! Se recomienda revisar nuevamente la sección de: ventajas y beneficios de la visualización de datos. |
| 1. Algunos beneficios de utilizar la visualización de datos son:  * Ahorro de tiempo * Mejor entendimiento de la información * Apoyo para la toma de decisiones * Almacenamiento de datos más eficiente | | Illustration of data analysis graph  **Fuente:** <https://img.freepik.com/free-vector/illustration-data-analysis-graph_53876-20421.jpg>  **Imagen:** 228131\_i\_36 |
| **Verdadero**  Retroalimentación:  ¡Incorrecto! Se recomienda revisar nuevamente la sección de: ventajas y beneficios de la visualización de datos. | | **Falso (correcto)**  Retroalimentación:  El almacenamiento de datos más eficiente está relacionado con las herramientas y procesos de ETL utilizados en el proyecto, no con los procesos de visualización de datos. |
| 4. Las gráficas de barras son especialmente útiles para mostrar tendencias a lo largo de un periodo de tiempo. Un gráfico de línea permite mostrar una o varias series de datos en la misma figura permitiendo así comparar sus comportamientos. | | Bar graph  **Fuente:** <https://img.freepik.com/free-vector/bar-graph_52683-9732.jpg>  **Imagen:** 228131\_i\_37 |
| **Verdadero**  Retroalimentación:  ¡Incorrecto! Se recomienda revisar nuevamente la sección de: selección de gráficos según el tipo de datos. | | **Falso (correcto)**  Retroalimentación:  La descripción corresponde a los gráficos de línea que se utilizan generalmente para representar datos continuos. |
| 5. Un mapa de calor (*heatmap* en inglés) es una forma de visualización en formato tabular donde, además de los números, se utiliza el color de las celdas para transmitir la magnitud relativa de los valores. | | Background of stock heatmap on the monitor in predominantly red  **Fuente:** <https://img.freepik.com/premium-photo/background-stock-heatmap-monitor-predominantly-red_98862-3889.jpg?w=740>  **Imagen:** 228131\_i\_38 |
| **Verdadero (Correcto)**  Retroalimentación:  ¡Correcto! en los mapas de calor se puede usar la saturación del color para proporcionar pistas visuales que ayuden a localizar fácilmente los puntos de interés (valores límites, por ejemplo). | | **Falso**  Retroalimentación:  ¡Incorrecto! Se recomienda revisar nuevamente la sección de: selección de gráficos según el tipo de datos. |
| 6. Esta gráfica representa un tipo de visualización llamado “diagrama de caja”. | | **Fuente:** <https://pro.arcgis.com/es/pro-app/latest/help/analysis/geoprocessing/charts/GUID-0E2C3730-C535-40CD-8152-80D794A996A7-web.png>  **Imagen:** 228131\_i\_39 |
| **Verdadero (Correcto)**  Retroalimentación:  ¡Correcto! El diagrama de caja (*box plot* en inglés) es un tipo de gráfico que permite visualizar y comparar la distribución y la tendencia central de valores numéricos mediante sus cuartiles. | | **Falso**  Retroalimentación:  ¡Incorrecto! Se recomienda revisar nuevamente la sección de: selección de gráficos según el tipo de datos. |
| 7. Microsoft Excel NO es una herramienta utilizada para la visualización de datos. | | Spreadsheets concept illustration  **Fuente:** <https://img.freepik.com/free-vector/spreadsheets-concept-illustration_114360-945.jpg>  **Imagen:** 228131\_i\_40 |
| **Verdadero**  Retroalimentación:  ¡Incorrecto! Se recomienda revisar nuevamente la sección de: herramientas de creación de visualización (estáticas e interactivas). | | **Falso (correcto)**  Retroalimentación:  ¡Correcto! Excel permite de una forma sencilla y rápida crear visualizaciones a partir de los datos existentes y personalizarlos a gusto del usuario, pero tiene en su contra, que no está diseñado ni optimizado para manipular grandes cantidades de datos como puede ser requerido en la mayoría de los proyectos de ciencia de datos. |
| 8. La herramienta GoogleCharts permite la creación de gráficos para usar de forma rápida en un sitio web propio. | | Dashboard business user panel  **Fuente:** <https://img.freepik.com/free-vector/dashboard-business-user-panel_23-2148368431.jpg>  **Imagen:** 228131\_i\_41 |
| **Verdadero (Correcto)**  Retroalimentación:  ¡Correcto! La galería de gráficos de Google Charts proporciona una gran cantidad de tipos de gráficos listos para usar. La forma más común de usar Google Charts es con JavaScript incrustado en la página web. | | **Falso**  Retroalimentación:  ¡Incorrecto! Se recomienda revisar nuevamente la sección de: herramientas de creación de visualización (estáticas e interactivas). |
| 9. Un *dashboard* (o tablero de control) es una herramienta para la gestión de la información que permite monitorear y analizar de forma visual y centralizada los indicadores claves de desempeño (KPIs), métricas y demás datos importantes para la organización. | | Interfaz de usuario del panel. Software de datos simple, diagramas de gráficos y HUD, paneles de administración. Infografía moderna de la interfaz de aplicaciones financieras. Estadística de visualización del diagrama de informes de ilustraciones  **Fuente:** [**https://www.shutterstock.com/es/image-illustration/dashboard-ui-simple-data-software-chart-1910372830**](https://www.shutterstock.com/es/image-illustration/dashboard-ui-simple-data-software-chart-1910372830)  **Imagen:** 228131\_i\_42 |
| **Verdadero (Correcto)**  Retroalimentación:  ¡Correcto! Los *dashboards* recopilan los principales datos y los presenta a los *stakeholders* para que puedan consultar solo la información necesaria. | | **Falso**  Retroalimentación:  ¡Incorrecto! Se recomienda revisar nuevamente la sección de: creación de *dashboards.* |
| 10. El diseño no es importante en los *dashboards*, únicamente los datos. | | Hombre trabajando con big data y bases de datos, usando laptops y tablas de excel. Trabajador de oficina haciendo análisis e informando con hojas de cálculo en el equipo. Ilustración vectorial plana aislada en fondo blanco  Fuente: <https://www.shutterstock.com/es/image-vector/man-working-big-data-databases-using-2015415122>  **Imagen:** 228131\_i\_43 |
| **Verdadero**  Retroalimentación:  ¡Incorrecto! Se recomienda revisar nuevamente la sección de: creación de *dashboards.* | | **Falso (Correcto)**  Retroalimentación:  ¡Correcto! Los tableros de control presentan su información de una forma visual y atractiva aprovechando la diversidad de gráficas y distintivos existentes (flechas, colores, cifras resaltadas) para facilitar al usuario la comprensión y aprovechamiento de la información allí presentada. La selección de un diseño adecuado es fundamental para la comprensión de la información presentada. |

**Material complementario**

|  |  |  |  |
| --- | --- | --- | --- |
| Tipo de recurso | Material complementario | | |
| Tema | Referencia APA del material | tipo | Enlace |
| Ventajas y beneficios de la visualización de datos | Microsoft Learn. (2022). *Tipos de visualización en Power BI*. | Sitio *web* | <https://learn.microsoft.com/es-es/power-bi/visuals/power-bi-visualization-types-for-reports-and-q-and-a> |
| Ventajas y beneficios de la visualización de datos | Tableau public. (2022). *Discover*. | Sitio *web* | <https://public.tableau.com/app/discover> |
| Ventajas y beneficios de la visualización de datos | Google Developers. (2022). *Charts*. | Sitio *web* | <https://developers-dot-devsite-v2-prod.appspot.com/chart> |
| Ventajas y beneficios de la visualización de datos | IBM. (2021). *Conceptos básicos de ayuda de CRISP DM*. | Sitio *web* | <https://www.ibm.com/docs/es/spss-modeler/saas?topic=dm-crisp-help-overview> |

**Glosario**

|  |  |
| --- | --- |
| **Tipo de recurso** | Glosario |
| Computación en la nube: | (*cloud computing*) es una tecnología que permite acceder de manera remota a recursos de almacenamiento, procesamiento, servidores y demás según las necesidades de la organización permitiendo así ahorrar costos en infraestructura y mantenimiento. |
| Cuartiles: | son valores que dividen un grupo de datos en cuatro grupos que contienen aproximadamente la misma cantidad de observaciones. |
| *Dashboard*: | es una herramienta de visualización de datos que permite monitorizar, analizar y mostrar de manera visual los indicadores clave de desempeño. |
| *Framework*: | es un esquema o marco de trabajo, un conjunto de prácticas y conceptos estandarizados que brindan la estructura base para elaborar un proyecto específico. |
| *Geomaps*: | son visualizaciones que se utilizan habitualmente para comparar valores y mostrar categorías entre regiones geográficas. Los gráficos de mapa son más útiles cuando los datos contienen información geográfica (países, regiones, estados, ciudades, códigos postales, etc.). |
| HTML: | lenguaje de marcado de hipertexto, del inglés HyperText Markup Language es el código que se utilizar para estructurar una página *web* y sus contenidos. |
| KPI: | indicador clave de desempeño, del inglés Key Perfomance Indicator es una métrica que se emplea para resumir la información sobre la eficacia y productividad de las acciones que se llevan a cabo en una organización. |
| Librería: | es un conjunto de archivos compuestos de código y datos, utilizados para desarrollar *software*. |
| *SaaS*: | *software* como servicio, del inglés *Software as a Service,* permite a los usuarios conectarse y utilizar aplicaciones basadas en la nube utilizando el Internet. El *software* o la aplicación no está instalada en los equipos de la organización. |
| *Stakeholders*: | son actores internos o externos a la organización que tienen algún tipo de relación o interés con esta o que son afectados positiva o negativamente por las acciones y proyectos ejecutados. |

**Referentes bibliográficos**

|  |  |
| --- | --- |
| **Tipo de recurso** | Bibliografía |
| Fernández, A. (2020). *Ciencia de datos para la ciberseguridad*. RAMA. <https://www.alphaeditorialcloud.com/reader/ciencia-de-datos-para-la-ciberseguridad-1628020600?location=eyJjaGFwdGVySHJlZiI6IngwMiIsImNmaSI6Ii80W3gwMl0vMi8yW19pZENvbnRhaW5lcjAwMl0vMiJ9> | |
| García, J. (2018). *Ciencia de datos. Técnicas analíticas y aprendizaje estadístico en un enfoque práctico*. Editorial Altaria. <https://www-alphaeditorialcloud-com.bdigital.sena.edu.co/reader/ciencia-de-datos?location=1> | |
| Geomapik. (2019). *Graphing libraries*. http://www.geomapik.com | |
| Jones, H. (2019). *Ciencia de los datos: la guía definitiva sobre análisis de datos, minería de datos, almacenamiento de datos, visualización de datos.* <https://biblioteca.sena.edu.co/F/QK1UI7RDS4Q5XCEJHTEG681M74XTARHJQABMVRV1F1YV8U8ELM-13705?func=full-set-set&set_number=003407&set_entry=000001&format=999> | |
| Joyanes, L. (2015). *Sistemas de información en la empresa*. Alfaomega | |
| Khanna (s.f). *Power BI Desktop*. [*https://powerbicdn.azureedge.net/cvt-9394a7615e76e3565238e1688022efbee201671e2cbb78c8d1b1df694b985908/pictures/pages/index/blade2/powerbi-001.jpg*](https://powerbicdn.azureedge.net/cvt-9394a7615e76e3565238e1688022efbee201671e2cbb78c8d1b1df694b985908/pictures/pages/index/blade2/powerbi-001.jpg) | |
| Microsoft Learn. (2022). *Descubra su ruta de aprendizaje*. <https://learn.microsoft.com/es-es/training/> | |
| Maptive. (2020). *Apping software for every professional.* https://www.maptive.com | |
| Orbegozo, B. (2019). *Gestión de bases de datos con SQL, MySQL y Access*. Alfaomega. | |
| Piattini, M. y Ruiz, F. (2020). *Gobierno y gestión de las tecnologías y los sistemas de información*. Ra-Ma. <https://www-alphaeditorialcloud-com.bdigital.sena.edu.co/reader/gobierno-y-gestion-de-las-tecnologias-y-los-sistemas-de-informacion-1591724977> | |
| QliK. (2917). *How do I create a visualization* [Video]. <https://learning.qlik.com/pluginfile.php/98829/mod_resource/content/9/Create_Visualizations/Create_Visualizations.html> | |
| TuDashboard. (2 de agosto de 2021). *Ejemplos de gráficas para un dashboard.* <https://tudashboard.com/tipos-de-visualizacion-de-datos/> | |