



Unidad I: Introducción a la visualización de datos

Maestría en Ciencia de Datos e Inteligencia de Negocios

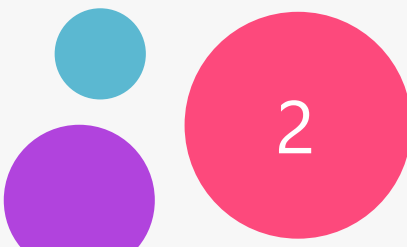


Dra. Argelia Berenice Urbina Nájera

argeliaberenice.urbina@upaep.mx – orcid: <http://orcid.org/0000-0002-3700-7287>

Contenido

- 1 Importancia de la visualización de datos
- 2 Usar un código para la visualización de datos
- 3 Interacción en la visualización de datos
- 4 Principios de la visualización de datos
- 5 Referencias
- 6

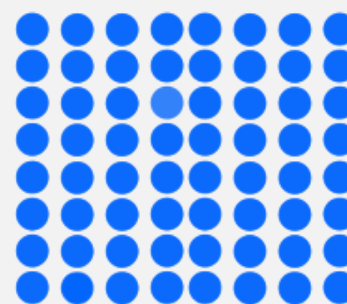


1

Importancia de la visualización de datos

David McCandless: La belleza de la visualización de datos

Sobrecarga de Información



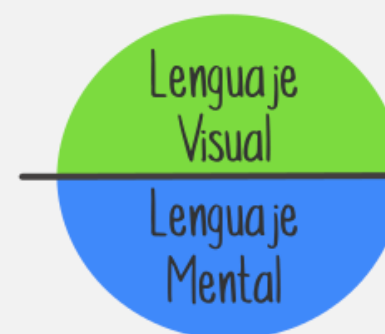
Ver patrones



Visualizar



Estamos conscientes del 0.7% de la información que percibimos



Así estructuramos el contenido en nuestra mente



Problemas de información

Introducción



Infoxicación

Recibir centenares de información cada día a los que no se dedica tiempo para analizarlos y procesarlos, no se profundiza en ellos, posponiendo la toma de decisiones o bien, tomando decisiones equivocadas bajo el estrés que provoca tanta información.



Introducción



Somos seres visuales

La información visual organiza, asocia ideas e información usando palabras clave o imágenes tal y como las procesa el cerebro humano, ayudando a retener la información en la memoria puesto que activa varias zonas del cerebro.

Lóbulo frontal → lóbulo parietal → lóbulo temporal → cerebelo → tálamo → tallo cerebral → y estos a todas partes



Definición 2

Es una forma de explorar, analizar y contextualizar datos para razonar sobre ellos.



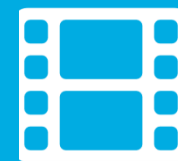
Definición 4

La visualización ideal no solamente debe comunicar con claridad, sino también debe estimular la participación y atención del espectador



Definición 1

El uso de representaciones visuales interactivas de datos abstractos para amplificar la cognición



Definición 3

La visualización de datos explica historias, simplifica, mide, compara, explora, descubre, permitiendo ser comprendidos de manera efectiva.



Tarea

La labor de un diseñador de datos es interpretar, aclarar, sintetizar, explicar, transmitir contenidos y ponerlos en contexto para que sean significativos.

2



2

Usar un código para la visualización de datos



Introducción



Importancia

La visualización de datos es una parte esencial tanto para la analítica digital como para elaboración de reportes ejecutivos. Es útil para identificar patrones o tendencias o valores atípicos que merecen un análisis posterior con profundo detalle.



Técnicas

Existen multitud de técnicas y aproximaciones para la visualización según sea la naturaleza del dato de la información. Desde el punto de vista de los datos, especialmente datos estructurados (o semi-estructurados) y su explotación visual, se puede establecer una clasificación según complejidad y elaboración de la información.

Antecedentes de la Visualización de datos

Pioneros y antecedentes históricos.

1983

Edward Tufte

Considerado como pionero de la visualización de datos con su trabajo *<The Visual Display of Quantitative Information>*

Excelencia

Tufte la define como *<define como la comunicación de ideas complejas con claridad, precisión y eficiencia>*

1983

2006

Michael Friendly

Impulsor del sitio web DataVis, en el cual se pueden encontrar muchos otros recursos relacionados, incluyendo artículos, libros y software.

Evolución

La base de la necesidad de resolver problemas concretos, relacionados con el deseo de visualizar fenómenos y relaciones entre elementos de forma diferente, y esto último es posible gracias al desarrollo de las metodologías (análisis estadístico) y las tecnologías (ordenadores).

2006

2017

Tarea de un Diseñador de datos

Es interpretar, aclarar, sintetizar, explicar, transmitir contenidos y ponerlos en contexto para que sean significativos.



Clasificación de las visualizaciones de datos



Visualizaciones básicas

Gráficas

Barras

Mapas



Cuadros de mando

Composición compleja de visualizaciones individuales



Infografías

Construcción narrativa de datos



Aplicaciones/código

Utilización de software especializado para generar interacción con el usuario ya sea mediante gráficos, gráficos en tiempo real, mallas, realidad aumentada, móviles, lenguajes de programación, entre otros,

Gráficos

Estándares para generar gráficos y bibliotecas que facilitan la creación de gráficos mediante código

Gráficos 2D

D3.js, Google Charts, chart.js, HighchartsJS,
Javascript INFOVIS Toolkit

Gráficos 3D

ThreeJS, BabylonJS

Mapas

Polymaps.JS, leaflet

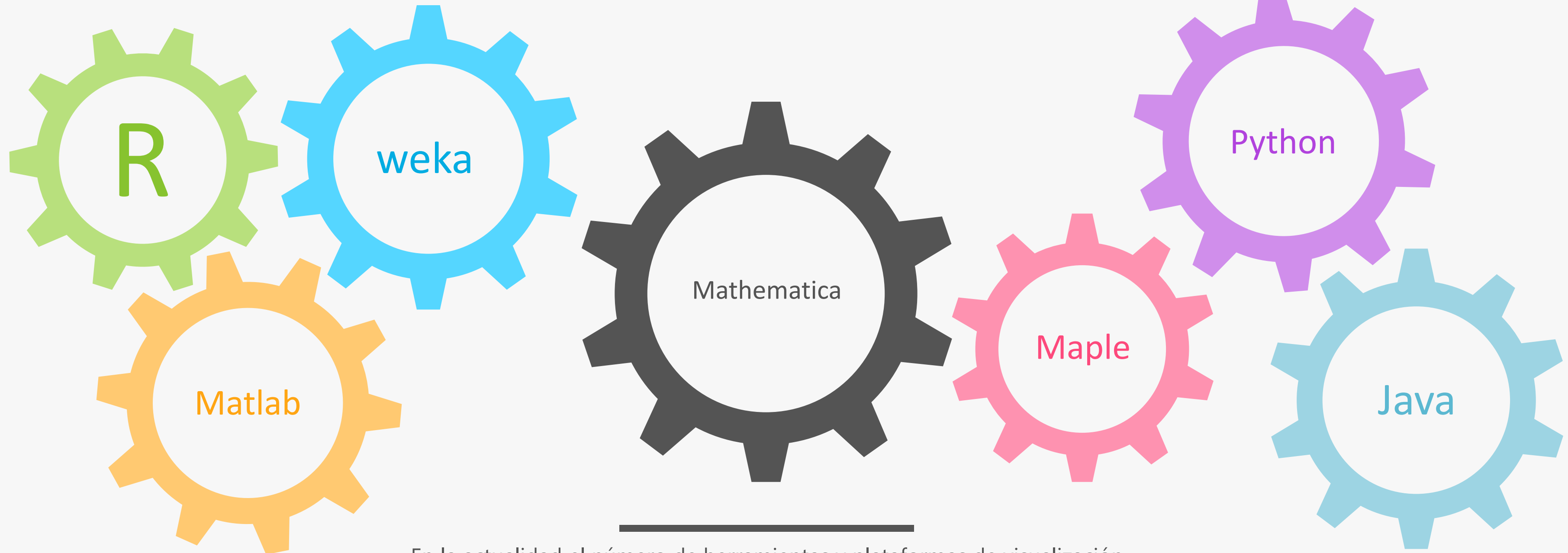
Líneas del tiempo

TimelineJS, Sigma.js



Gráficos

Lenguajes de programación para crear gráficos



En la actualidad el número de herramientas y plataformas de visualización de datos al que se puede acceder es muy extenso. En función de las necesidades del usuario se puede elegir entre un gran número de opciones de vanguardia.



Ejemplos



Evolución de la
web



Historia de la
música



Mejores
visualizaciones
de 2018

Datos abiertos



Gobierno de México



Gobierno de España




Otros

Bases de datos libres



1. Machine Learning Repository
<https://archive.ics.uci.edu/ml/datasets.html>
2. Gapminder <https://www.gapminder.org/data/>
3. Instituto Nacional de Estadística
España <http://ine.es/>
4. INEGI <https://www.inegi.org.mx/datos/>
5. Datos abiertos México <https://datos.gob.mx/>
6. Datos Banco Mundial
<https://datos.bancomundial.org/>
7. Varios/para estudiantes
<https://www.springboard.com/blog/free-public-data-sets-data-science-project/>
8. Varios
<https://bigdata-madesimple.com/70-websites-to-get-large-data-repositories-for-free/>
9. Kaggle.com <https://www.kaggle.com/datasets>

Interacción en la visualización de datos



- Como se ha visto anteriormente, en cada minuto se generan millones de datos, si estos datos son presentados en una gráfica, tabla o infografía es posible que en menos de una semana ya sean **obsoletos**.

Por ello, los **datos interactivos** están tomando gran importancia, dado que muchos de ellos son generados en tiempo real.



Interacción en la Visualización de datos

Técnicas para representar datos de manera interactiva



Dashboard

Es una representación gráfica de los principales indicadores de desempeño de una empresa que intervienen en la consecución de objetivos del negocio para la toma de decisiones que ayuden a generar ventajas competitivas y estrategias de negocio.

Herramientas de software

Power BI	Qlik
ClicData	Cyfe
Tableau	



Infografía interactiva

Se combinan grandes cantidades de datos **en una sola página**, ya que el usuario puede navegar entre ellos y, la mayoría de las veces, decidir en cual focalizarse. Se suele utilizar el color para distinguir fácilmente las tendencias.

Herramientas de software

Canva	Piktochart
Info.gram	Easel.ly
Genially	Visme



Interacción en la Visualización de datos

Técnicas para representar datos de manera interactiva



Mapas interactivos

Es una herramienta que muestra información geográfica.

Herramientas de software

Tripline	MyHistro	Animaps
Educaplay	Mapme	
Mango	ArcGis	
MapHub	StepMap	



Gráficos interactivos

Permiten modificar sus características después de haber sido creados, ejecutando los cambios en tiempo real; es decir, se pueden alterar los parámetros hasta obtener los resultados deseados.

Dentro de los parámetros que se facilitan modificar son el color, la textura, la forma, el tipo de gráfico, los títulos, insertar o editar funciones de resumen, intercambiar ejes para las variables, ingresar nuevas variables, etc.

Herramientas de software

Stat Planet	TableauPublic
GapMinderWorld	ViewShare
Prefuse	Flare

Interacción en la Visualización de datos

Ejemplos



Dashboard



Infografía



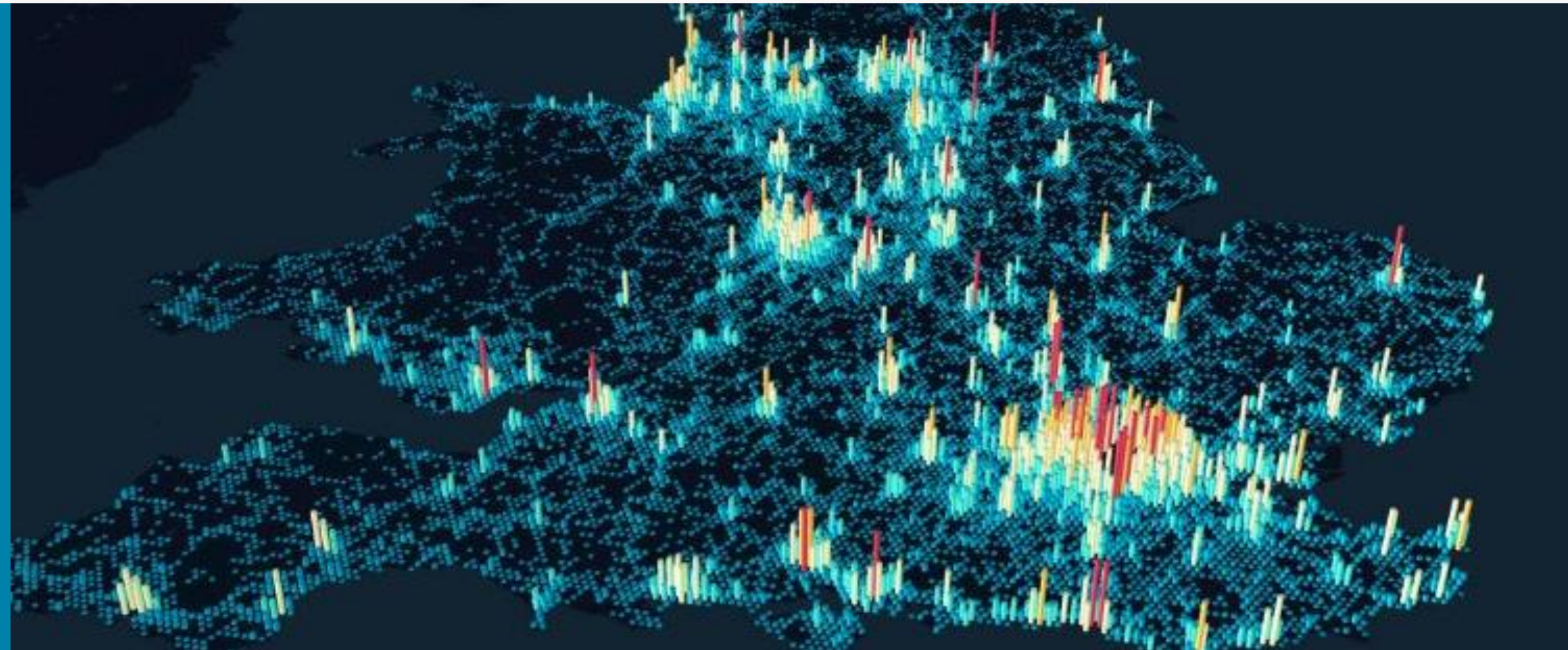
Mapas



Gráficos

Principios de la **visualización**

Como lo ha dicho Stephen Few, la visualización de datos es la visualización gráfica de información abstracta con dos propósitos: **Creación de sentido (análisis de datos) y comunicación.**



La información es abstracta ya que describe cosas que no son físicas. Ya sea que se trate de ventas, incidencias en enfermedades, rendimiento atlético o cualquier otro tema, se puede mostrar visualmente, aunque para hacerlo **se debe encontrar la manera de dar forma a lo que no lo tiene.**

Principios de la visualización

La visualización de datos solo tiene éxito en la medida en que se codifica la información de una manera que nuestros ojos pueden discernir y nuestros cerebros pueden entender.

Hacer esto bien es mucho más una ciencia que un arte, lo cual solo se puede lograr estudiando la **percepción humana**.

El objetivo es traducir la información abstracta en representaciones visuales que pueden decodificarse de manera fácil, eficiente, precisa y significativa.

Principios Visualización

Cinco principios básicos

Claridad

1

Indicar claramente cómo se relacionan los valores entre sí

2

Precisión

Representar las cantidades con precisión

Comparación

3

Hacer que sea fácil comparar las cantidades

4

Orden

Hacer que sea más fácil ver el orden de los valores clasificados

Motivación

5

Hace obvio cómo las personas deberían usar la información, para qué deberían usarla, y motivarlos a hacerlo.



Principios de la Visualización

Hoy en día, con tecnologías y metodologías nuevas y mejoradas para la exploración cerebral, abundan las oportunidades para mejorar la efectividad perceptiva de la visualización de datos. Dos áreas de estudio en particular son especialmente útiles:

El procesamiento visual preatento

- El procesamiento visual preatento es esa parte que ocurre automáticamente en el cerebro antes de la conciencia.

Mecanismos y limitaciones de atención y memoria

- Los estudios en atención y memoria están revelando nuestra capacidad sorprendentemente limitada para mantener múltiples objetos simultáneamente en la conciencia.



Principios de la Visualización

-
- La aplicación de la visualización de datos más allá de las estadísticas descriptivas en el ámbito del análisis predictivo, como el uso de modelos visuales interactivos.
 - Integración más estrecha de algoritmos de minería de datos para encontrar patrones significativos con visualización de datos para proporcionar una mejor manera de revisar y explorar esos patrones.
 - Dispositivos mejorados de interfaz hombre-computadora para interactuar con la visualización de datos de una manera más rápida y sin problemas.
 - La integración de pantallas geoespaciales y de red (como diagramas de nodo y enlace) con otras formas de visualización para una interacción perfecta y uso simultáneo.



Referencias de la Unidad

-
1. Grant R. (2018). Data visualization. Chapman and Hall/CRC
 2. Sosulski K. (2018). Data Visualization Made Simple. Routledge
 3. Evergreen S. (2017). **Effective Data Visualization: The Right Chart for the Right Data** 1st Edition. **SAGE. UK.**
 4. Magnuson L. (2016). Data visualization. Rowman & Littlefield Publishers
 5. Julià Minguillón A. (2016). Introducción a la visualización de datos. Recuperado en enero 2018. Disponible en: <http://openaccess.uoc.edu/webapps/o2/handle/10609/57624>
 6. Berinato S. (2016). **Good Charts: The HBR Guide to Making Smarter, More Persuasive Data Visualizations. HBR. Boston.USA**
 7. Alcalde Perea, I. (2015). Visualización de la información: de los datos al conocimiento. Editorial UOC.
 8. **Few S. (2013). Information Dashboard Design: Displaying Data for At-a-Glance Monitoring** Second Edition, Second edition Edition.