**1. Ventajas del Promedio Móvil como medio de pronóstico frente a promedio total o ultimo resultado.**

METODOS DE EXTRAPOLACIÓN: Son modelos para hacer pronósticos de valores que pueda tomar una variable a futuro, se conocen también como modelos de SERIES DE TIEMPO ya que utilizan los valores que históricamente ha tomado una variable aleatoria X = (xt-n, xt-n+1, …., xt-1, xt, xt+1,….), para con ellos predecir los valores futuros. Se basan en el supuesto de que las cosas seguirán ocurriendo en el futuro, como lo han venido haciendo en el pasado.

Ello implica que nunca van a predecir crisis, ni cambios violentos, sino solo continuidad de lo que está ocurriendo.

Necesitan de datos anteriores para armar los modelos y otros datos para probarlos. Habitualmente se llama a los últimos post-sample de modo de distinguirlos de los originales con los cuales se los construye. Ello es muy importante pues sino se hace así se corre el riesgo de hacer modelos para un conjunto de datos, que luego no sirven para proveer información sobre comportamientos futuros.

Los valores que ha adoptado la variable en el pasado y eventualmente el actual, se producen en períodos que pueden ser igualmente espaciados en el tiempo o no.

Al pronóstico se lo designa como Ft,n, donde el primer subíndice expresa el tiempo en que se ha realizado el pronóstico y el segundo es para cuantos periodos hacia adelante se lo realiza.

El más obvio es el modelo del último valor en que el pronóstico se hace en el tiempo t y es para el período siguiente Ft,1 = xt. Como éste deja la duda respecto a la permanencia del valor, es mucho más utilizado en pronóstico por promedio de los valores pasados:

Ft+1 pero como naturalmente es muy objetable y a veces poco práctico, usar datos muy antiguos, se utiliza más el pronóstico por medias móviles:

Ft,1 = Promedio de las últimas N observaciones :

El problema de éste último, es saber cuál es el valor más adecuado para N. Para resolver eso definimos el error del pronóstico como la diferencia entre el valor realmente tomado por la variable y el valor que se había pronosticado para la misma (por ejemplo: ) y la Desviación Absoluta Media (MAD en ingles) como el promedio de los errores absolutos de las últimas N mediciones:

haciendo una pequeña cantidad de ensayos se puede determinar el valor de N más adecuado para minimizar esa desviación que además guarda una relación con la desviación estándar de los errores como , lo cual, fijado un intervalo de confianza, nos va a permitir predecir el error que tendrá nuestro pronóstico, asumiendo que éstos tienen una distribución normal.

**2. Describa los Componentes de una serie de tiempo.**

Se supone 4 componentes separados que se combinan para dar valores a la serie de tiempo:

- Tendencia: Cambio gradual creciente, decreciente o lineal de la serie de tiempo en el largo plazo.

- Cíclico: *es la fluctuación con forma de onda alrededor de la tendencia y, por lo común, se ve afectada por las condiciones económicas generales. Ciclos de varios años.*

* - Estacional: Patrón de cambio de la serie que se repite año tras año o menos. Ej. Venta de estufas que se incrementa en invierno o de aire acondicionado en verano.

- Irregular: Es la variabilidad aleatoria en la serie, resultado de factores de corto plazo, imprevistos y no recurrentes.

**3. Describa que es pronosticar y cuáles son los pasos en el proceso de pronosticar.**

Pronosticar es predecir qué sucederá en el futuro con los valores de variables que son de interés para una organización.

Los métodos de predicción pueden ser:

- Cuantitativos:

o *Métodos de pronósticos de series temporales o de extrapolación:* para pronosticar los valores futuros de series de tiempo a partir de valores anteriores de una serie temporal

o *Métodos de predicción causal:* se basan en el supuesto de que la variable que tratamos de pronosticar muestra una relación de causa y efecto con una o más variables

- Cualitativos: no requieren de una abierta manipulación de datos sino que hacen uso del juicio de quien pronostica, por su naturaleza éstos suelen ser subjetivos y no utilizan modelos matemáticos.

Pasos en el proceso de pronosticar

1. Formulación del problema y recopilación de datos

2. Manipulación y limpieza de datos: eliminar no válidos, cambio de unidades, procesar datos, etc.

3. Construcción y evaluación del modelo: ajustar datos a modelo de pronóstico

4. Implementación del modelo (el pronóstico real): generación del modelo

5. Evaluación del pronóstico: comparar pronósticos con históricos reales

**4. Ventajas del Método de Suavizado exponencial frente al Promedio Móvil.**

La media móvil da igual importancia a todos los N valores anteriores, sin embargo normalmente se piensa que las cosas van sufriendo pequeños cambios y que los valores más recientes son más importantes que los pasados; entonces se concibió el pronóstico por SUAVIZAMIENTO EXPONENCIAL, en el cual se les da a los valores pasados una importancia decreciente, que además evita tener que determinar una cantidad de valores N conveniente

**5. Para que se mide la precisión del método de pronóstico. Describa dos métodos, indique la fórmula que utilice y de un ejemplo de su uso.**

La precisión se mide para determinar que tan bien se ajusta el pronóstico a los valores reales de la serie. Pudiendo comparar métodos y ajustar parámetros.

2 métodos: error cuadrado medio (EMC) y desviación media absoluta (MAD).

EMC --> promedio de la suma de los errores (*et*) al cuadrado.



N = número de observaciones pronosticadas

- Este método penaliza los desvíos más grandes.

MAD --> promedio de la suma de los valores absolutos de los errores



N = número de observaciones pronosticadas

- Los desvíos están expresados en las mismas unidades de la serie.

Ejemplo

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Mes** | **X** | **F** | **|X-F|** | **|X-F|^2** |
| **1** | **10** |  |  |  |
| **2** | **5** | **7,5** | **2,5** | **6,25** |
| **3** | **15** | **10** | **5** | **25** |
| **4** | **10** | **12,5** | **2,5** | **6,25** |
| **5** | **20** | **15** | **5** | **25** |
|  |  |  |  |  |
|  |  | **MAD** | **3,75** |  |
|  |  | **EM**C | **15,625** |  |

**6. ¿Cómo incorpora el error en el pronóstico en el Método de Suavizado Exponencial?**

Otra forma de ponerlo, derivada de la original es: lo que la muestra como el pronóstico anterior más un término representativo del error obtenido en el pronóstico anterior afectado de un coeficiente de atenuación.

Esta bien esto de arriba??? alguien tiene la posta??

Esta resp es algo asi: como la formula es a(alpha) . Xt + (1-a). At-1 podria decirse que el error se incorpora al tener en la función incorporado tanto el valor real del mes t, como el pronostico para dicho mes dado en el mes anterior.. de esta manera el error se ve incluido en el calculo del proximo pronostico.

Aquí La forma en la que se incluye el error en el pronóstico en suavizado exponencial se puede observar reordenando los términos de la ecuación.

Ft = α \* Xt-1 + (1 - α) \* Ft-1

(distribuyendo 2do término)

Ft = α \* Xt-1 + Ft-1 - α \* Ft-1

(reordenando)

Ft = α \* (Xt-1 - Ft-1) + Ft-1

(reemplazando et = Xt - Ft)

Ft = Ft-1 + α \* et-1

**7. Que es** α **en el Método de Suavizado exponencial, que valores puede asumir, que relaciones se pueden observar con el MAD mínimo.**

α es la constante de suavizado y los valores que puede asumir son entre 0 y 1.

Generalmente se utiliza un α 0.1, 0.3 o 0.5. Si el valor de α que minimiza al MAD es superior a 0.5 entonces la serie presenta tendencia, estacionalidad o variación cíclica por lo que el suavizado exponencial no sería el adecuado.

**8. Valores recomendables para el coeficiente de suavizamiento α y porque.**

Los libros sugieren utilizar valores de α inferiores a 0,3 y como máximo 0,5, eligiéndolo de modo de minimizar el valor de la MAD. Como los datos con que se cuenta siempre serán acotados y en ocasiones pocos de los antiguos se hallarán confiablemente, habrá que tener o hacer una estimación inicial para calcular un primer pronóstico.

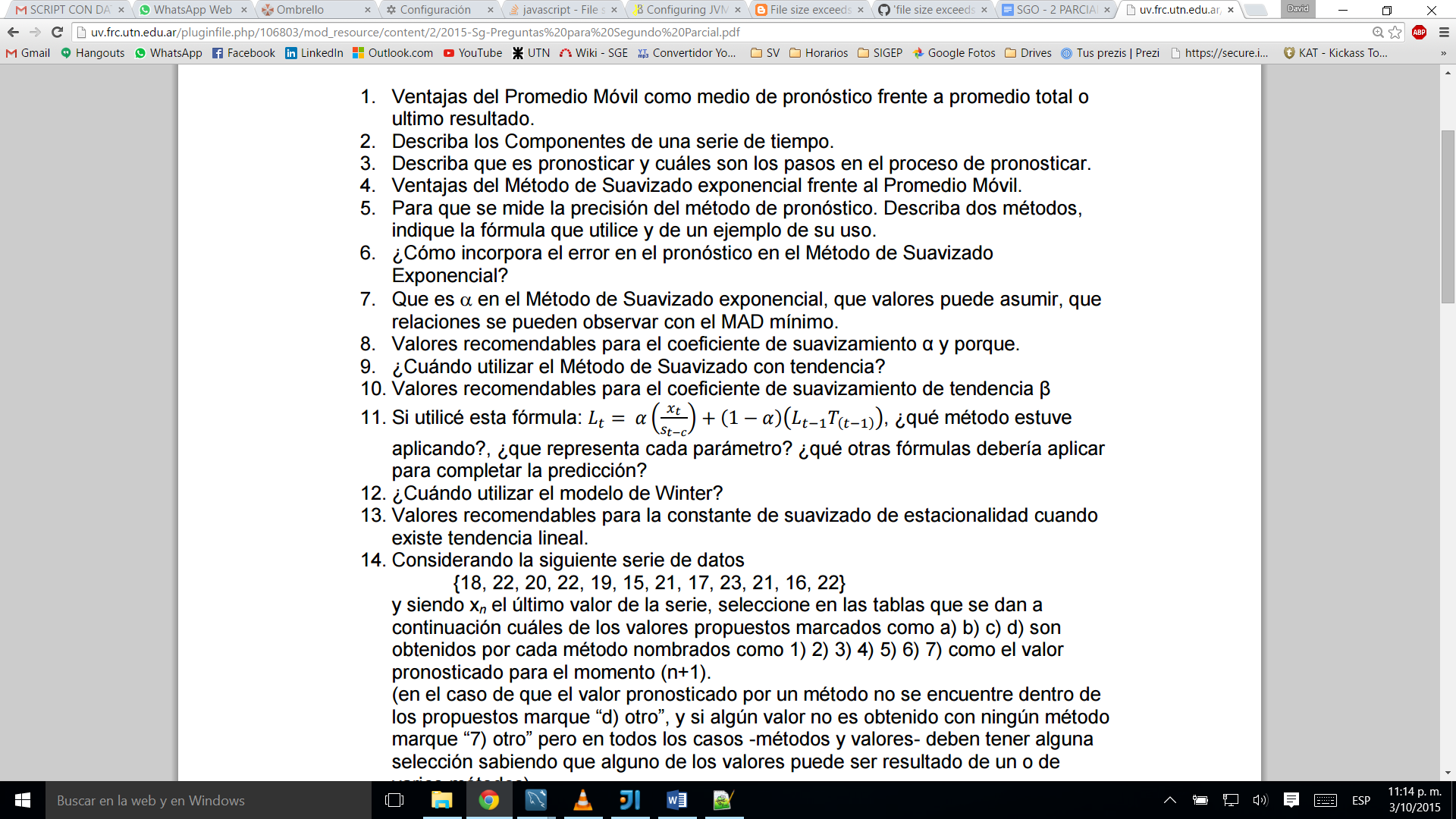
En la práctica se utiliza α igual a 0.10, 0.30 o 0.50. Si el valor de α que minimiza el MAD sobrepasa 0.5, entonces tendencia, o estacionalidad o variación cíclica está presente, por lo que no se recomienda el suavizamiento exponencial simple como técnica de pronóstico.

**9. ¿Cuándo utilizar el Método de Suavizado con tendencia?**

Este modelo funciona muy bien cuando el nivel de los resultados es constante; puede ocurrir que los resultados guarden una cierta tendencia, creciente o decreciente, como por ejemplo aumento de precios por inflación y por incorporación de un producto al mercado, o decreciente porque se está dejando de usar; también puede ocurrir que los resultados sean estacionales, como las ventas de estufas que se incremente en invierno o de los aire acondicionado que se incrementa en verano. En ambos casos, el suavizamiento exponencial deja de comportarse eficientemente y debe ser corregido, por tendencia (Método de Holt) o por estacionalidad (Método de Winter, que incluye también tendencia)

**10. Valores recomendables para el coeficiente de suavizamiento de tendencia β**

Mismas restricciones que para alfa, para los problemas en que el método de suavizamiento con tendencia es eficiente los valores de alfa y beta están por debajo de 0.5, es decir que si ambos valores no son menores a 0.5, entonces la estacionalidad o el comportamiento cíclico podría estar presente, por lo cual se tendría que aplicar otro método de pronóstico.

**11. Si utilicé esta fórmula:**

, ¿qué método estuve aplicando?, ¿que representa cada parámetro? ¿qué otras fórmulas debería aplicar para completar la predicción?

La fórmula corresponde a la estimación de base del método de Winter de suavizado exponencial con estacionalidad.

α es la constante de suavizado de base.

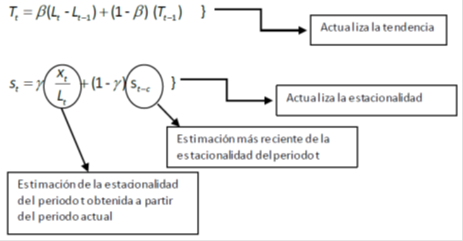
c = número de periodos en la duración del patrón estacional.

Por ejemplo si es trimestral c = 4, si es mensual c = 12

st = factor multiplicativo estacional para el periodo t, obtenido después de observar xt.

Lt : base

Tt : tendencia



β es la constante de suavizado de tendencia.

Por último el pronóstico ft,k del periodo t+k:



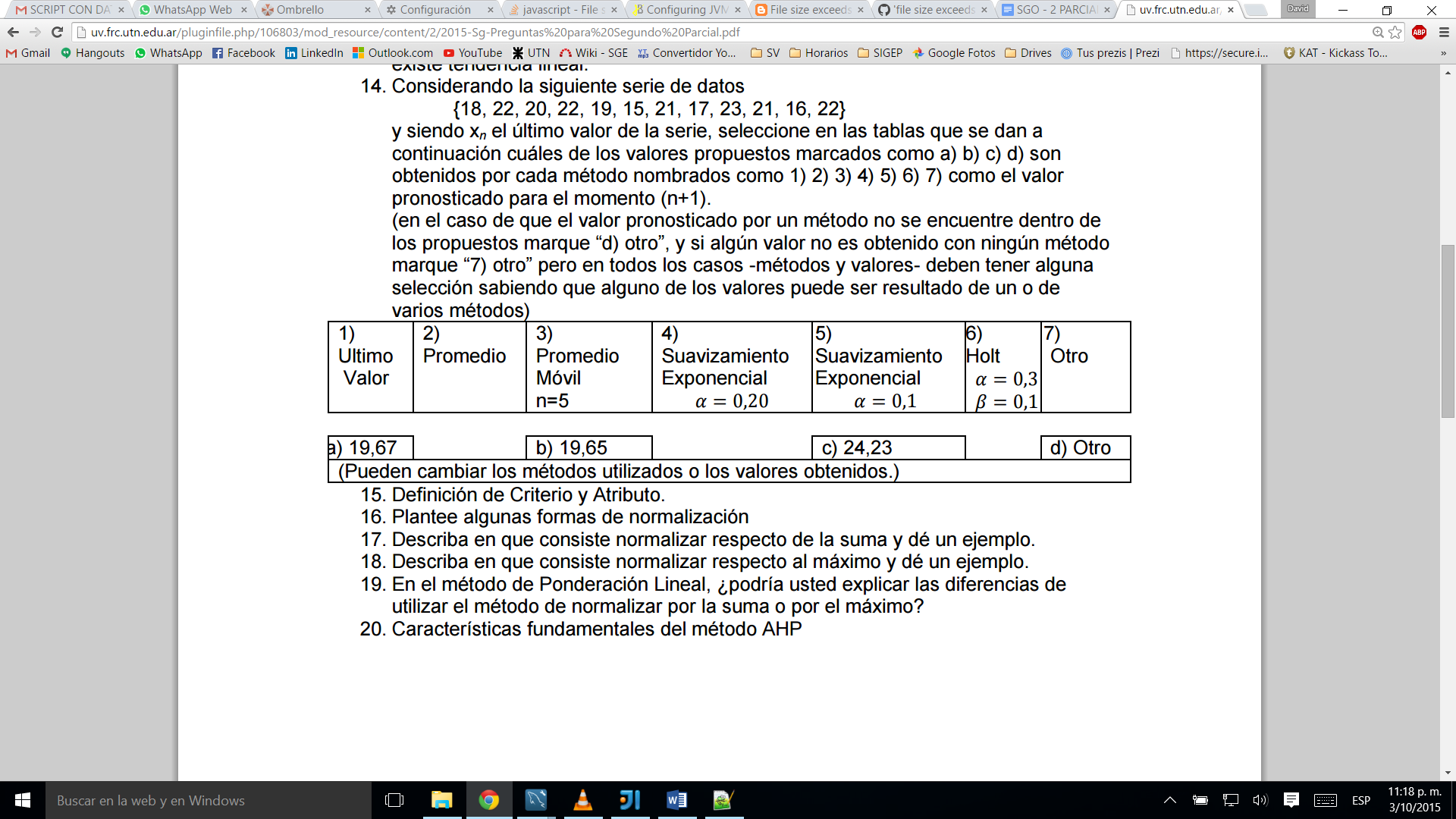
**12. ¿Cuándo utilizar el modelo de Winter?**

Cuando además de la tendencia (o en presencia de tendencia nula) hay estacionalidad, es decir los datos fluctúan en periodos iguales de tiempo (todos los veranos: son estacionales) se aplica el método de Winter en el cual aparece un factor de estacionalidad st con distintos valores para cada período t, que multiplica a un valor intermedio haciéndolo más grande o más pequeño. Como todas son estimaciones, ahora también debe estimarse ese factor

**13. Valores recomendables para la constante de suavizado de estacionalidad cuando existe tendencia lineal.**

Aunque los valores de α y β que minimizan MAD no deben ser mayores que 0.5 (como el método de Holt), es común, por lo que toca al mejor valor de g, que ésta sea mayor que 0.5. Esto se debe a que, para los datos mensuales, cada factor estacional mensual está actualizado sólo durante 1/12 de todos los períodos. Puesto que los factores de estacionalidad se actualizan tan pocas veces, necesitaríamos dar más peso a cada observación, de modo que g>0.5 no es imposible.

14. Considerando la siguiente serie de datos {18, 22, 20, 22, 19, 15, 21, 17, 23, 21, 16, 22} y siendo xn el último valor de la serie, seleccione en las tablas que se dan a continuación cuáles de los valores propuestos marcados como a) b) c) d) son obtenidos por cada método nombrados como 1) 2) 3) 4) 5) 6) 7) como el valor pronosticado para el momento (n+1). (en el caso de que el valor pronosticado por un método no se encuentre dentro de los propuestos marque “d) otro”, y si algún valor no es obtenido con ningún método marque “7) otro” pero en todos los casos -métodos y valores- deben tener alguna selección sabiendo que alguno de los valores puede ser resultado de un o de varios métodos)



1 - 3 - 5 - D

2 - A

4 - B

HOLT me dio 21,68 != C

por ende 5=D y C=7

alguien lo hizo?

SIII NOS DIO IGUAL, no hay chance que lo tome, es muy largo para hacer en el parcial.

15. ¿Además de la probabilidad a priori, cual es la información necesaria para el cálculo de la probabilidad a posteriori?

16. ¿Como se calcula el valor esperado de la información perfecta?

17. Para que se utiliza la siguiente formula y que representan los términos y parámetros de la misma: P (Ai / B) = P (Ai) P (B / Ai) / Σi [ P (Ai) P (B / Ai) ]

**18. Cuáles de las siguientes expresiones representan probabilidad a priori y probabilidad a posteriori (las expresiones podrían tener otros valores o referirse a otras variables):**

**a. cuando hubo Ventas Altas, la consultora pronosticó ventas altas el 80% de las veces.**

**b. La probabilidad de que las ventas sean altas dado que la consultora pronosticó que serían altas es de 0.80.**

**c. cuando hubo Ventas Altas, la consultora pronosticó ventas altas el 80% de las veces, Ventas Medias el 10% y Ventas Bajas un 10%.**

**d. La probabilidad de que se pronostiquen ventas altas es de 0.30**

**e. cuando hubo Ventas Medias, la consultora pronosticó Ventas Bajas un 10% de las veces.**

**f. La probabilidad de que se ocurran ventas altas es de 0.40**

**g. La probabilidad de que las ventas sean bajas dado que se pronosticó venta media es de 0.10.**

**19. Al analizar el costo de la prueba, ¿qué diferencias existen entre un problema con objetivo maximizante y uno con objetivo minimizante?**

**20. Suponga un caso en el que los estados de la naturaleza representan “Lluvia leve” y “Caída de granizo” y se tiene la posibilidad de contratar una consultora para que lo asesore que cobra $2000 por cada estudio. Este asesor tiene un grado de certeza alto respecto a los resultados entregados, es decir que ha predicho “Lluvia” un 25% de las veces en las que realmente ocurrió “Caída de Granizo” y, por otra parte, ha predicho “Granizo” el 10% de las veces que ha habido en realidad “Lluvia Leve”. Describa las variables y las probabilidades a priori que describen el caso. (las expresiones podrían tener otros valores o referirse a otras variables)**

**21. Considerando que en un caso los estados de la naturaleza representan “Ocurre Año Normal” y "Ocurre Año de escasez” y además en los pronósticos, el profesional los describe de la siguiente manera:**

**“Se debe notar que las 15 veces que se pronosticó año normal siempre se acertó y cuando hubo año de escasez el pronóstico siempre acertó. Se produjo año de escasez 2 veces en los 20 años analizados”. Describa las variables y la probabilidad condicional que describe a la prueba. (Las expresiones podrían tener otros valores o referirse a otras variables)**

**22. Definición de Criterio y Atributo**.

· Atributo: representan propiedades, características, capacidades de satisfacer necesidades y/o deseos, etc. que poseen las alternativas de decisión

· Criterio: es una función que refleja las preferencias del decisor en relación a un atributo.

**22. Plantee algunas formas de normalización**

· Procedimiento 1: normalización con respecto al máximo

· Procedimiento 2: normalización con respecto al rango

· Procedimiento 3: normalización con respecto a la suma



**24. Que significa normalizar?**

**25. Describa en que consiste normalizar respecto de la suma y dé un ejemplo.**

· La normalización respecto de la suma lleva los valores a una valuación con respecto al total o sumatoria de todos los valores

· Las valuaciones quedan comprendidas entre 0 y 1

· Con mejor valuación el valor más cercano a 1

· ri = ai / ∑ ai

Ejemplo:

|  |  |  |
| --- | --- | --- |
|  | **Ganancia** | **Valor normalizado** |
| **proyecto 1** | 8 | 0,4211 |
| **proyecto 2** | 3 | 0,1579 |
| **proyecto 3** | 5 | 0,2632 |
| **proyecto 4** | 3 | 0,1579 |
| Suma | 19 | 1 |

**26. Describa en qué consiste normalizar respecto al máximo y dé un ejemplo.**

· La normalización respecto al máximo lleva los valores a una valuación con respecto al mayor valor de todo el conjunto

· Las valuaciones quedan comprendidas entre 0 y 1

· Con mejor valuación el valor más cercano a 1

· Al máximo valor le corresponde el valor 1

· ri = ai / max ai

Ejemplo:

|  |  |  |
| --- | --- | --- |
|  | **Ganancia** | **Valor normalizado** |
| **proyecto 1** | 8 | 1,0000 |
| **proyecto 2** | 3 | 0,3750 |
| **proyecto 3** | 5 | 0,6250 |
| **proyecto 4** | 3 | 0,3750 |
| Max | 8 |  |

**27. Características fundamentales del método AHP**

* Estructura el problema en una jerarquía interrelacionada de los elementos de decisión:

La principal característica del AHP es que el problema de decisión se modeliza mediante una jerarquía en cuyo vértice superior está el principal objetivo del problema, meta a alcanzar y, en la base, se encuentran las posibles alternativas a evaluar. En los niveles intermedios se representan los criterios (los cuales a su vez se pueden estructurar también en jerarquías) en base a los cuales se toma la decisión.

* La asignación de pesos se realiza mediante un método de comparaciones binarias de los criterios:

La segunda característica del método es que, en cada nivel de la jerarquía, se realizan comparaciones entre pares de elementos de ese nivel, en base a la importancia o contribución de cada uno de ellos al elemento de nivel superior al que están ligados. Este proceso de comparación conduce a una escala de medida relativa de prioridades o pesos de dichos elementos.

* Resuelve problemas de tipo γ : para generar un ordenamiento de las alternativas

**28. ¿Por qué en AHP, a diferencia del método de Ponderación Lineal, no es necesario que todos los criterios sean a maximizar o a minimizar?**

· Por qué utiliza el método de comparación de a pares entre elementos de decisión donde el decisor asigna un valor de preferencia con respecto a un par de elementos de decisión sin tener en cuenta la medida real del atributo.

**29. Significado del término Consistencia en AHP**

· Los juicios del decisor en la comparación por pares deben tener un cierto grado de consistencia que cumple con la transitividad cardinal de los juicios de la matriz

**30. ¿Qué representan los valores de aij en la matriz de comparación de pares?**

· Representa el juicio de preferencia del decisor ante un par de elementos de decisión de un nivel del árbol jerárquico que contribuye al objetivo del nivel inmediato superior. Utilizando una escala de 1 a 9 refleja la respuesta a 2 preguntas:

o ¿Cuál de los elementos es más importante con respecto a un criterio de nivel superior?

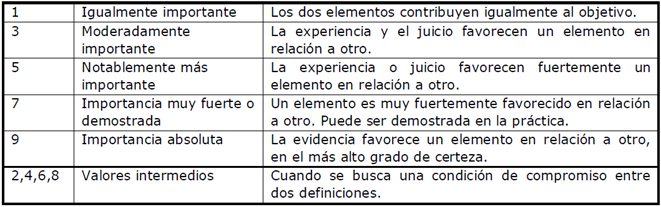
o ¿Cuánto más importante es?

**31. ¿Qué representan los valores de W en AHP?**

· Representan los pesos relativos de los criterios para la selección de la mejor opción

**32. ¿Cuál (o cómo es?) es la escala de valorización del AHP?**

· Es la escala fundamental propuesta por Saaty con valores que crecen en importancia relativa de 1 a 9.



**33. ¿Qué interpretación le puede dar al Índice de Consistencia ( IC ) en AHP?**

· Es una medida de la consistencia de los juicios del decisor y se utiliza para calcular la relación de consistencia y poder determinar así si la consistencia es aceptable o no.

· Los juicios del decisor en la comparación por pares deben tener un cierto grado de consistencia que cumple con la transitividad cardinal de los juicios de la matriz

**34. Como pueden sintetizarse en etapas el proceso analítico jerárquico propuesto por Saaty.**

Etapa 1: Estructurar el problema como un árbol jerárquico.

Etapa 2: Descubrir las preferencias del decisor mediante comparaciones de a pares entre los elementos de decisión - criterios y alternativas -.

Etapa 3: Estimar los pesos relativos de los elementos de decisión.

Etapa 4: Comprobar la consistencia de los juicios del decisor.

Etapa 5: Generar una evaluación global de cada alternativa a través de una ponderación lineal.

**35. Que representa la sigla AHP y como es interpretado en español.**

**·**  Analytic Hierarchy Process: Proceso analítico jerárquico

**36. Explique cómo se determina el vector de pesos de los criterios en AHP.**

**37. Explique qué significa que una matriz de comparaciones no sea consistente.**

**38. ¿Por qué es necesario el análisis de consistencia?**

**39. A que se denomina Aprendizaje Organizacional.**

**40. ¿Cuando decimos de la Administración del Conocimiento nos referimos a?.**

**41. Hay tres tipos principales de sistemas de administración del conocimiento. Explique cada uno de ellos.**

**42. ¿Cómo funcionan los sistemas expertos?**

**43. Definición de Minería de Datos.**

La minería de datos es un campo de las ciencias de la computación. Es el proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos. Utiliza los métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de bases de datos. La tarea de minería de datos real es el análisis automático o semi-automático de grandes cantidades de datos para extraer información, transformándola en una estructura comprensible, para su uso posterior. Se busca encontrar patrones de comportamiento oculto, tendencias o ciertos comportamientos que eran desconocidos en un principio, con el fin de crear nuevas asociaciones, predecir futuros resultados, etc. Esto quiere decir, convertir datos en (información) conocimiento para tomar decisiones. El entorno de la minería de datos suele tener una arquitectura cliente/servidor.

"Las herramientas de la minería de datos ayudan a extraer el mineral de la información enterrado en archivos corporativos o en registros públicos, archivados. El minero es, muchas veces un usuario final con poca o ninguna habilidad de programación, facultado con poderosas herramientas indagatorias para efectuar preguntas y obtener rápidamente respuestas."

**44. Principales características y objetivos de la Minería de Datos.**

**45. El proceso de la minería de datos se suele componer de cinco etapas principales. Explique.**

**46. En que escenarios se pueden aplicar los modelos de minería de datos**

**47. ¿Cuáles son las técnicas de la Minería de Datos más representativas?**

**48. Ejemplifique el uso de minería de datos**

**49. Defina Inteligencia de Negocios**

**50. ¿Qué es inteligencia de Negocios como solución tecnológica y como ventaja competitiva?**

**51. ¿Cuáles son beneficios más importantes al usar Inteligencia de Negocios?.**

**52. Mencionar algunas de las desventajas al realizar Inteligencia de Negocios**

**53. Explicar la Extracción -Transformación y Carga en la Inteligencia de Negocios.**

**54. ¿Qué es el Modelado Multidimensional? Explicar los componentes del modelado multidimensional**

**55. Defina DataWareHouse.**

**56. ¿Cuáles son las características principales del DataWareHouse?**

* *Integrado*: los datos almacenados en el Data Warehouse deben integrarse en una estructura consistente, por lo que las inconsistencias existentes entre los diversos sistemas operacionales deben ser eliminadas.
* *Temático*: Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del Data Warehouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.
* *Histórico*: el tiempo es parte implícita de la información contenida en un Data Warehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el Data Warehouse sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el Data Warehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.
* *No* *volátil*: el almacén de información de un Data Warehouse existe para ser leído, y no modificado. La información es por tanto permanente, significando la actualización del Data Warehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

**57. ¿Dónde se puede utilizar DatawareHouse?**

**58. ¿Cómo se clasifican los sistemas OLAP?. Explique cada uno de ellos.**

**59. ¿Cuáles son las razones para usar OLAP? Explique beneficios de OLAP**

**60. ¿Cuáles son las utilidades que presentan las aplicaciones OLAP?**

**61. ¿Qué significa que en OLAP los datos están estructurados en modelos estrellas o copo de nieve?**

**62. ¿Cuáles son las utilidades de OLAP?**

**63. Defina Data Mart**

**64. ¿Cuáles son las Razones para crear Data Mart?**

**65. Explique cuál es el Proceso General para crear Data Mart.**

**66. ¿Qué significa Big data?**

**67. ¿Cuál es la Importancia de Big Data?**

**68. Defina Dinámica de Sistemas**

**69. ¿Qué son los Modelos Mentales?**

**70. Desarrolle Diagrama Causales.**

**71. ¿Cuál es la Metodología de Aplicación de la Dinámica de Sistemas?**

**72. Desarrolle Sistemas Estables e Inestables y Sistemas con factor externo.**

**73. Desarrolle Sistemas Hiperestables, Oscilantes y Sigmoidales.**

**74. Defina Arquetipos Sistémicos.**

**PREGUNTAS VIEJAS**

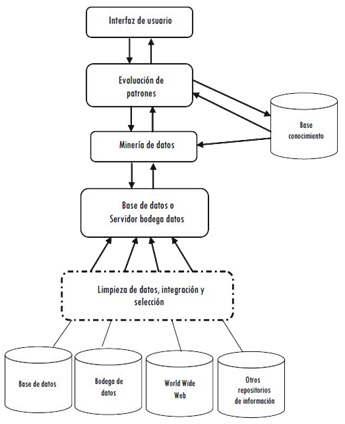
**29. Minería de Datos: Concepto.**

La minería de datos es un campo de las ciencias de la computación. Es el proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos. Utiliza los métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de bases de datos. La tarea de minería de datos real es el análisis automático o semi-automático de grandes cantidades de datos para extraer información, transformándola en una estructura comprensible, para su uso posterior. Se busca encontrar patrones de comportamiento oculto, tendencias o ciertos comportamientos que eran desconocidos en un principio, con el fin de crear nuevas asociaciones, predecir futuros resultados, etc. Esto quiere decir, convertir datos en (información) conocimiento para tomar decisiones. El entorno de la minería de datos suele tener una arquitectura cliente/servidor.

"Las herramientas de la minería de datos ayudan a extraer el mineral de la información enterrado en archivos corporativos o en registros públicos, archivados. El minero es, muchas veces un usuario final con poca o ninguna habilidad de programación, facultado con poderosas herramientas indagatorias para efectuar preguntas y obtener rápidamente respuestas."

30. Desarrolle la Arquitectura para Data Mining

* *Base de datos:* Puede ser de tipo base de datos normal, data warehouse, hoja de cálculo u otra clase de repositorio. A estos datos se le aplican *técnicas de limpieza e integración.*
* *Servidor de bases de datos*: Utilizado para obtener la información relevante según el proceso de minería de datos.
* *Base de conocimiento:* Conocimiento del dominio para guiar la búsqueda y evaluar los patrones. Se tienen en cuenta las creencias de los datos, los umbrales de evaluación y el conocimiento previo. Con el fin de dirigir búsquedas y evaluar patrones se tiene un completo conocimiento del dominio.
* *Minería de datos (algoritmos):* Modular para realizar distintos tipos de análisis de los datos en busca de conocimiento: Caracterización, Asociación, Clasificación, Análisis de grupos, Evolución (en espacio o tiempo) y Análisis de desviaciones. A través de estos se pueden realizar distintos tipos de análisis de los datos en búsqueda del conocimiento. Los algoritmos de minería de datos son procedimientos bien definidos, es decir, procesos codificados como un conjunto finito de reglas que toman los datos como entradas y sus salidas como modelos o patrones. De acuerdo con los métodos de aprendizaje, se pueden clasificar los algoritmos en: Algoritmos de tipo predictivo. Predicen el valor de un atributo del conjunto de datos a partir de otros atributos conocidos. Algoritmo de descubrimiento de conocimiento. Descubren patrones y tendencias en los datos actuales y permite obtener beneficio de estos.
* *Módulo de evaluación:* Mide que tan interesante es un patrón. Interactúa con el algoritmo de Minería de Datos para guiar la búsqueda hacia patrones interesantes. En conjunto con los algoritmos de minería, se busca evaluar los diversos patrones con el fin de encontrar el más interesante.
* *Interfaz gráfica*: Interacción con el usuario. Elección de la tarea de minería de datos. Provee información para enfocar la búsqueda. Ayuda a evaluar los patrones. Explora los patrones encontrados y la base de datos original. Visualiza los patrones en distintas formas.



**31. Explique las técnicas para la Minería de Datos.**

Las técnicas de la minería de datos provienen de la [inteligencia artificial](https://es.wikipedia.org/wiki/Inteligencia_artificial) y de la [estadística](https://es.wikipedia.org/wiki/Estad%C3%ADstica). Dichas técnicas, no son más que [algoritmos](https://es.wikipedia.org/wiki/Algoritmo), más o menos sofisticados que se aplican sobre un conjunto de datos para obtener unos resultados. Las técnicas más representativas son:

* Redes neuronales.- Son un [paradigma](https://es.wikipedia.org/wiki/Paradigma) de [aprendizaje](https://es.wikipedia.org/wiki/Aprendizaje) y procesamiento automático inspirado en la forma en que funciona el [sistema nervioso](https://es.wikipedia.org/wiki/Sistema_nervioso) de los animales. Se trata de un sistema de interconexión de [neuronas](https://es.wikipedia.org/wiki/Neurona) en una [red](https://es.wikipedia.org/wiki/Red_neuronal_artificial) que colabora para producir un estímulo de salida. Algunos ejemplos de red neuronal son:

o El [perceptrón](https://es.wikipedia.org/wiki/Perceptr%C3%B3n).

o El [perceptrón multicapa](https://es.wikipedia.org/wiki/Perceptr%C3%B3n_multicapa).

o Los [mapas auto organizados](https://es.wikipedia.org/wiki/Mapa_autoorganizado), también conocidos como redes de [Kohonen](https://es.wikipedia.org/wiki/Mapa_autoorganizado).

* Regresión lineal.- Es la más utilizada para formar relaciones entre datos. Rápida y eficaz pero insuficiente en espacios multidimensionales donde puedan relacionarse más de 2 variables.
* Árboles de decisión.- Un árbol de decisión es un modelo de predicción utilizado en el ámbito de la [inteligencia artificial](https://es.wikipedia.org/wiki/Inteligencia_artificial) y el [análisis predictivo](https://es.wikipedia.org/wiki/An%C3%A1lisis_predictivo), dada una base de datos se construyen estos diagramas de construcciones lógicas, muy similares a los sistemas de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones (alternativas de decisión) que suceden de forma sucesiva, para la resolución de un problema.
* Modelos estadísticos.- Es una expresión simbólica en forma de igualdad o ecuación que se emplea en todos los diseños experimentales y en la regresión para indicar los diferentes factores que modifican la variable de respuesta.
* [Agrupamiento o Clustering](https://es.wikipedia.org/wiki/Algoritmo_de_agrupamiento).- Consiste en agrupar un conjunto de datos basándose en la similitud de los valores de sus atributos
* [Reglas de asociación](https://es.wikipedia.org/wiki/Reglas_de_asociaci%C3%B3n).- Se utilizan para descubrir hechos que ocurren en común dentro de un determinado conjunto de datos.

**32. Desarrolle el Proceso de Minería de Datos.**

Un proceso típico de minería de datos consta de los siguientes pasos generales:

1. *Selección del conjunto de datos*, tanto en lo que se refiere a las variables objetivo (aquellas que se quiere predecir, calcular o inferir), como a las [variables independientes](https://es.wikipedia.org/wiki/Variable_dependiente) (las que sirven para hacer el cálculo o proceso), como posiblemente al [muestreo](https://es.wikipedia.org/wiki/Muestreo_en_estad%C3%ADstica) de los [registros](https://es.wikipedia.org/wiki/Registro_%28base_de_datos%29) disponibles.

2. *Análisis de las propiedades de los datos*, en especial los [histogramas](https://es.wikipedia.org/wiki/Histograma), diagramas de [dispersión](https://es.wikipedia.org/wiki/Dispersi%C3%B3n_%28matem%C3%A1tica%29), presencia de [valores atípicos](https://es.wikipedia.org/wiki/Valor_at%C3%ADpico) y ausencia de datos (valores [nulos](https://es.wikipedia.org/wiki/Null)).

3. *Transformación del conjunto de datos de entrada*.Se realizará de diversas formas en función del análisis previo, con el objetivo de prepararlo para aplicar la técnica de minería de datos que mejor se adapte a los datos y al problema. A este paso también se le conoce como pre procesamiento de los datos.

4. *Seleccionar y aplicar la técnica de minería de datos*. Se construye el modelo predictivo, de clasificación o segmentación.

5. *Extracción de conocimiento*. Mediante una técnica de minería de datos, se obtiene un modelo de conocimiento, que representa patrones de comportamiento observados en los valores de las variables del problema o relaciones de asociación entre dichas variables.

6. *Interpretación y evaluación de datos*.Una vez obtenido el modelo, se debe proceder a su validación comprobando que las conclusiones que arroja son válidas y suficientemente satisfactorias.

**33. Data Mart: Concepto y Componentes.**

Concepto:

Es un subconjunto de datos corporativos válidos para una unidad de negocio específica, sobre un tema puntual o para un grupo reducido de personas. Este subconjunto consiste en datos históricos, sumarizados y posiblemente detallados. Es una solución menos costosa y mucho más reducida que la implementación de un data warehouse. Normalmente comprende un modelo multidimensional para usuarios con perfil de granjeros.

Componentes:

1. Fuentes de Datos: Son las que alimentan de información al DataMart, están diseñadas para registrar grandes cantidades de transacciones.
2. Procesos de Extracción, transformación y carga de datos (ETL): Los datos se encuentran almacenados en bases de datos destinadas al registro de las transacciones. Es necesario extraer y transformar los datos antes de cargar los resultados en el DataMart.
3. Datamart: Es una solución que, con contenidos específicos, volumen de datos más limitado y un alcance histórico menor que el de un DataWarehouse permite dar soporte a una empresa pequeña, un departamento o área de negocios de una empresa grande.
4. Herramientas de Explotación: El Datamart está preparado para ser explotado mediante herramientas específicas que permiten la extracción de información significativa y patrones de comportamiento que permanecen ocultos en un repositorio de datos.

**34. Tipos de Data Mart**

Un datamart puede ser alimentado desde los datos de un datawarehouse, o integrar por sí mismo un compendio de distintas fuentes de información. Se pueden plantear dos tipos de datamarts:

* Datamart OLAP: Se basan en los populares cubos OLAP.
* Datamart OLTP: Los datamarts que están dotados con estas estructuras óptimas de análisis presentan las siguientes ventajas:

-Poco volumen de datos

-Mayor rapidez de consulta

-Consultas SQL y/o MDX sencillas

-Validación directa de la información

-Facilidad para la historización de los datos

Otra clasificación conocida divide a los data marts en dos tipos:

1. Data Marts Dependientes: Los datos que se utilizan para poblar el data mart provienen del data warehouse.

2. Data Marts Independientes: Los datos que se utilizan para poblar el data mart provienen de los sistemas operacionales y/o fuentes externas.

**35. Explique las fases de construcción de una Data Mart.**

Las Fases de Construcción de un Data mart son las siguientes :

1.- Construcción del Datamart: Esta actividad tiene el objetivo de construir el modelo de datos, la metadata de la herramienta de Explotación y La Arquitectura del Modelo Multidimensional en la herramienta de explotación.

2.- Construcción de los Procesos de Cargas: En ésta actividad es cuando se debe desarrollar los procesos de carga de datos, las rutinas de limpieza, los flujos de cargas de datos, las interfaces de acceso, los importadores e integradores de datos, los programas de entrada de datos, etc.

3.- Construcción de los reportes analíticos: Consiste en construir los reportes, tableros de control, dashboard, scorecard, etc .

4.- Construcción de los procesos de prueba: Se debe construir los programas, reportes, informes que permita probar los procesos de cargas y los reportes entregados.

36. Diferencias entre un Data Mart Top Down y Bottom Up

Arquitectura para Data Mart

|  |  |
| --- | --- |
| — Top-Down: el cual tiene como base un sistema DW para toda la empresa y a partir de este se desarrollan los DMs para las divisiones o departamentos | —Bottom-Up: en esta arquitectura, se definen previamente los DM y luego se integran en un DW centralizado |



**37. ¿Qué diferencia hay entre un Data Warehouse y un Data Mart?**

Data Warehouse: Se trata del lugar donde toda la data de una compañía es almacenada. Consiste en un sistema computarizado con una gran capacidad de almacenamiento, esencial para reunir y organizar la información proveniente de los distintos departamentos de la organización.

Data Mart: Esta herramienta se ocupa de almacenar información de un departamento o grupo de trabajo específico. Funciona como una aplicación del Data Warehouse o una alternativa para empresas medianas que no pueden afrontar los costos de implementar un sistema tan amplio de almacenamiento de data. Las Data Marts pueden ser dependientes o independientes del Data Warehouse. Sin embargo, cabe mencionar que contar con sistemas independientes que no se encuentren integrados entre sí puede dificultar las tareas de administración y mantenimiento.

Un Data Mart cumple los mismos principios que un Data Warehouse, construir un repositorio de datos único, consistente, fiable y de fácil acceso. Entonces ¿Qué diferencia hay entre un data warehouse y un data mart? Su alcance.

|  |  |
| --- | --- |
| DATA MART | DATAWAREHOUSE |
| Se centra solamente en los requerimientos de usuarios asociados con un departamento o función de negocio | Tiene como objetivo agrupar los datos de toda la empresa con el fin de facilitar su análisis, de forma que sean útiles para acceder y analizar información sobre la propia empresa. |
| Normalmente no contienen datos operacionales detallados. | Almacena datos de acuerdo a categorías o estructurándolos de forma que favorezcan el análisis histórico de los datos |
| Debido a que los data marts contienen menos información son más fácilmente entendibles y navegables | No está orientado a procesos relacionados con la operatividad de la empresa. |
| Normalmente usan menos datos de tal manera que la tarea de carga de datos es más fácil, y de esta manera la implementación es más simple | Usan datos de toda la empresa de manera que la tarea de carga de datos y su construcción e implementación es más compleja |
| El costo de implementación es menor. | El costo de implementación es mayor. |
| Los potenciales usuarios de un data mart son más claramente definidos y puede ser más fácilmente involucrados para obtener soporte para un proyecto de data mart | Los potenciales usuarios de un datawarehouse no están claramente definidos y es más complicado involucrarlos para obtener soporte para un proyecto |

**38. Ventajas y Desventajas de un Data Mart**

VENTAJAS:

* Da a los usuarios acceso a los datos que ellos necesitan para analizarlos más a menudo.
* Provee los datos en una forma que concuerda la vista colectiva de los datos por un grupo de usuarios en un departamento o función de negocio.
* Mejora el tiempo de respuesta al usuario final debido a la reducción en el volumen de información a ser accedido.
* Provee datos apropiadamente estructurados para satisfacer los requerimientos de las herramientas de acceso de usuario final.
* Estructura los datos para su adecuado acceso por una herramienta.
* La tarea de carga de datos es más fácil, y de esta manera la implementación de un data mart es más simple comparado con un datawarehouse corporativo
* El costo de implementación de un Data Mart normalmente es menor que el requerido para establecer un Datawarehouse
* Permiten segmentar los datos en diferentes plataformas de hardware.
* Los potenciales usuarios de un Data Mart son más claramente definidos y puede ser más fácilmente involucrados para obtener soporte para un proyecto de Data Mart

DESVENTAJAS: La principal desventaja de los Data Mart es que no permiten el manejo de grandes volúmenes de información.

**39. Defina Datawarehousing.**

El DataWareHousing (DWH) es un proceso que posibilita la extracción de datos de sistemas operacionales y de fuentes externas (una gran cantidad de datos recolectados durante años) para consolidar, integrar, homogenizar y centralizar los datos que la empresa genera en todos los ámbitos de una actividad de negocios (Compras, Producción, Ventas, etc), facilitando posteriormente una amplia gama de posibilidades de análisis multivariable que brindan soporte a la toma de decisiones estratégicas y tácticas.

Un Data Warehouse (DW) es un conjunto de datos integrados orientados a una material que varían con el tiempo y que no son transitorios, los cuales soportan el proceso de toma de decisiones de una administración.De acuerdo con algunas organizaciones, el Data Warehouse es una arquitectura. Para otras, es un depósito semánticamente consistente en datos (separados y que no interfieren con los sistemas operativos y de producción existentes) que Ilenan por completo los diferentes requerimientos de acceso y reporte de datos. Para algunos otros, el Data Warehouse es un proceso continuo que mezcla los datos de varias fuentes heterogéneas, incluyendo datos históricos y adquiridos para soportar la constante necesidad de consultas estructuradas y/o ad hoc, reportes analíticos y soporte de decisiones.

**40. ¿Cuáles son las características más importantes en un Dataware?**

* *Integrado*: los datos almacenados en el Data Warehouse deben integrarse en una estructura consistente, por lo que las inconsistencias existentes entre los diversos sistemas operacionales deben ser eliminadas.
* *Temático*: Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del Data Warehouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.
* *Histórico*: el tiempo es parte implícita de la información contenida en un Data Warehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el Data Warehouse sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el Data Warehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.
* *No* *volátil*: el almacén de información de un Data Warehouse existe para ser leído, y no modificado. La información es por tanto permanente, significando la actualización del Data Warehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

**41. Describa los pasos involucrados en el proceso ETL:**

1. (DataValidation) Extracción o validación: La primera parte del proceso ETL consiste en localizar y extraer los datos, ya que en la mayoría de los proyectos de almacenamiento se fusionan datos provenientes de diferentes sistemas de origen tales como ERP, CRM, bases de datos, planillas Excel, etc.

La extracción convierte los datos a un formato homogéneo y consolidado para iniciar el proceso de transformación.

2. (DataCleaning) Limpieza: este proceso recupera los datos en bruto, elimina los duplicados.

3. (Data Transformation) Transformación: este proceso recupera los datos limpios y de alta calidad y los estructura y recopila en los distintos modelos de análisis. El resultado de este proceso es la obtención de datos limpios, consistentes, sumarizados y útiles.

4. (Data Integration) Integración: valida que los datos que se cargan en el dw son consistentes con las definiciones y formatos del dw, los integra en los distintos modelos de las distintas áreas de negocio que hemos definido en el mismo.

5. (Data Loading) Actualización o carga: El objetivo de esta fase consiste en tomar los registros formateados por la fase de transformación y cargarlos en el Data Warehouse, que es el contenedor para todos los datos informativos (actuales e históricos) requeridos por las operaciones del Data Warehouse

**42. Ventajas y Desventajas de un Dataware.**

Ventajas:

1. Proporciona información clave para la toma de decisiones empresariales.

2. Mejora la calidad de las decisiones tomadas.

3. Especialmente útil para el medio y largo plazo.

4. Son sistemas relativamente sencillos de instalar si las fuentes de datos y los objetivos están claros.

5. Muy útiles para el almacenamiento de análisis y consultas de históricos.

6. Proporciona un gran poder de procesamiento de información.

7. Permite una mayor flexibilidad y rapidez en el acceso a la información.

8. Facilita la toma de decisiones en los negocios.

9. Las empresas obtienen un aumento de la productividad.

10. Proporciona una comunicación fiable entre todos los departamentos de la empresa.

11. Mejora las relaciones con los proveedores y los clientes.

12. Permite conocer qué está pasando en el negocio, es decir, estar siempre enterado de los buenos y malos resultados.

13. Transforma los datos en información y la información en conocimiento

14. Permite hacer planes de forma más efectiva.

15. Reduce los tiempos de respuesta y los costes de operación.

Desventajas:

1. Requieren una revisión del modelo de datos, objetos, transacciones y además del almacenamiento.

2. Tienen un diseño complejo y multidisciplinario.

3. Tienen un alto coste.

4. Requieren sistemas, aplicaciones y almacenamiento específico.

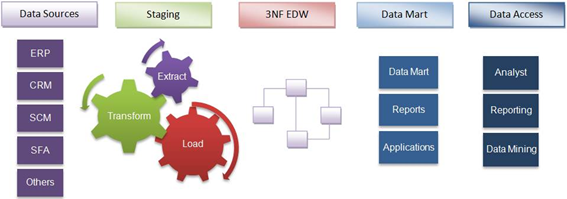
**43. Desarrolle los 2 (dos) enfoques para la arquitectura de un Dataware:**

Enfoques para la arquitectura de DW: Inmon y Kimbal

Cuando hay que construir un Data Warehouse para un negocio los dos métodos más utilizados son los de Kimball e Inmon. Tienen ventajas y desventajas pero ambos se utilizan a tal fin.

Inmon

define el Data Warehouse como un repositorio centralizado de toda la empresa completa. Y Datamarts multidimensionales son creados después de que el Data Warehouse en tercera forma normal fue completado.

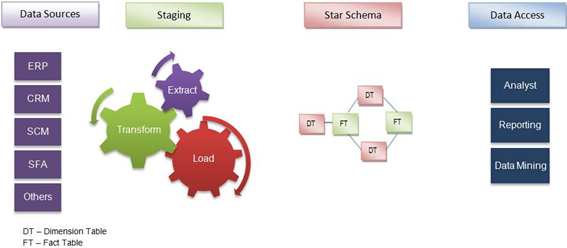


Kimball

Crea los Datamarts multidimensionales con la información más importante para la empresa o por áreas de negocio primero. Luego cuando se requiere estos se pueden combinar en un gran Data Warehouse (también multidimensional).

Kimball define un Data Warehouse como “Una copia de los datos transaccionales estructurada para consulta y análisis”.

La arquitectura de Data Warehouse de Kimball se conoce como “Data Warehouse Bus”.



|  |  |  |
| --- | --- | --- |
| **Kimball vs Inmon** | **Kimball** | **Inmon** |
| **Respuesta a requerimientos** | Inmediata | Escala de tiempo mayor |
| **Ámbito** | Áreas de negocio | Empresa |
| **Presupuesto** | Pequeño | Mayor |
| **Cliente** | Usuario | Corporativo |
| **Fuentes** | Estables | Cambiantes |
| **Costo inicial** | Bajo | Alto |
| **Proyectos** | Costo igual al inicial | Costo menor al inicial |
| **Enfoque** | Bottom-Up | Top-Down |

**44. Describa los logros del Dataware en la Gestión Comercial:**

Enfoque al cliente: las bases de datos ahora son temáticas y entre sus temas de interés se encuentran los clientes, además de productos, campañas, ventas, competencia, etc.

Inteligencia de clientes: quiénes son para la corporación los clientes fieles, ocasionales, potenciales o cuáles reportan la mayor parte de los ingresos.

Interactividad: la combinación DW con Internet/Intranet da lugar a un nuevo concepto los cyber warehouse que posibilitan una nueva comunicación on-line entre la empresa y el cliente.

Fidelización de clientes: un mejor conocimiento del cliente permite que se mejoren las características de las ofertas, el enfoque y los servicios consiguiendo “atrapar al cliente”.

Marketing directo: se pasa a desarrollar campañas basadas en perfiles con productos, ofertas y mensajes dirigidos específicamente a ciertos tipos de clientes.

Personalización: a través de segmentación de clientes y de los patrones de comportamiento que se pueden anticipar con Data Mining, se puede llegar a la personalización del mensaje.

**45. Definición de OLAP (On Line Analytical Processing, Procesamiento Analítico en Línea).**

OLAP (On Line Analytical Processing, Procesamiento Analítico en Línea) es una categoría de software que, a través de una interfaz sencilla y ágil, permite a analistas, administradores y ejecutivos analizar datos corporativos desde diferentes puntos de vista, con niveles de granularidad y agregación variables, ya sean datos históricos o proyecciones, y mostrando los resultados en términos familiares al usuario, con el objeto de tomar decisiones. Esta categoria de software, basa su potencia de consultas en una estructura de almacenamiento de datos multidimensional, la cual almacena datos cuantitativos (hechos) en categorias descriptivas (dimensiones).

**46. Ventajas de OLAP.**

La principal ventaja es la velocidad de respuesta, en tiempo real, a las consultas complejas con diversa cantidad de parámetros y extracción de datos de diversas fuentes.

Otras de las ventajas son:

* Un modelo de datos intuitivo y multidimensional que facilita la selección, recorrido y exploración de los datos.
* Un lenguaje analítico de consulta con la capacidad de explorar las complejas relaciones existentes entre los datos empresariales.
* Un pre-cálculo de los datos consultados con más frecuencia que permite una rápida respuesta a las consultas ad hoc.
* Uso más eficaz de los almacenes de datos para el análisis en línea proporcionando respuestas rápidas a consultas analíticas, complejas e iterativas.

**47. Diferencias entre OLAP (On Line Analytical Processing, Procesamiento Analítico en Línea) y OTLP (Procesamiento de Transacciones en Línea).**

La relación entre estas tecnologías es que una depende de la otra. OTLP (Procesamiento de Transacciones en Línea) se refiere a los sistemas transaccionales que utilizan bases de datos relacionales, en las cuales se almacenan entidades en tablas discretas y normalizadas. Éstos sistemas se ocupan de la captura de datos en tiempo real de transacciones, y por el contrario, los sistemas OLAP, se basan en las consultas de dichos datos para obtener informes y resultados que ayuden a la toma de decisiones. Es por eso que para la consulta de dichos datos en tiempo real se necesita de un modelo de datos multidimensional que desnormalice la estructura de los datos.

A continuación se presenta un cuadro con las diferencias más importantes:

|  |  |  |
| --- | --- | --- |
| **Características** | **OLTP** | **OLAP** |
| Objetivo | Control de procesos operacionales | Toma de decisiones |
| Cliente | Personal Operacional  Nivel Operacional | Gestores del Negocio  Nivel Estratégico |
| Datos | Atómicos, actualizados, y dinámicos | Consolidados, históricos y estables |
| Estructuras | Normalizada | Dimensional - Desnormalizada |
| Almacenamiento | Bases de datos Relacionales | Bases de datos Multidimensionales |
| Tiempo de Respuesta | Segundos | De segundos a minutos |
| Orientación | Orientado a Aplicaciones | Orientado a Información |
| Acceso | Alto | Moderado a bajo |
| Actualización | Alto | Moderado a bajo |
| Aplicación | Estructuras y procesos repetitivos | No estructurados y procesos analíticos |
| Tipos de acciones | Insert, update, delete (Transacciones) | Select (Consultas de datos) |

**48. Cubos OLAP: Definición y Características**

El cubo OLAP es una base de datos, en la cual el almacenamiento físico de los datos se realiza en un vector multidimensional. Esto quiere decir que se almacenan datos históricos integrados de una manera simple, concisa y comprensible, que permite ver la información desde distintos puntos de vista, variables y relaciones entre ellas. Se diferencia de las bases relacionales ya que éstas muestran datos normalizados, atómicos y las relaciones entre entidades. En cambio, los cubos tratan de desnormalizar relaciones, agrupar datos y resumirlos, para mejorar la performance de las consultas, definiendo hechos y dimensiones.

Algunas características:

* Representa una entidad diferente del negocio.
* Son creados a partir de un esquema de base de datos Estrella o Copo de Nieve (forma del modelo de datos multidimensional), y se caracterizan por tener tablas de *Hechos* y tablas de *Dimensiones*.
* Las tablas de Hechos guardan los datos históricos y generalmente son medidas numéricas que describen una transacción del negocio, por ejemplo: ventas, transacciones, costos, etc.
* Las tablas de Dimensiones dan significado de los datos contenidos en la tabla de hechos, son los distintos puntos de vista por los cuales pueden consultarse los datos, por ejemplo: regiones, tiempo, tipo de producto, estados, etc

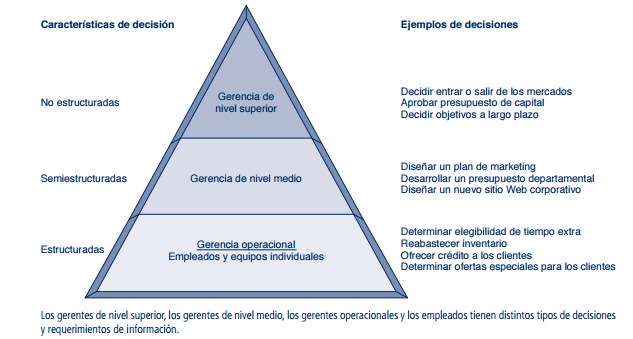
**49. Toma de Decisiones: Concepto**

Schein, plantea: "la toma de decisiones es el proceso de identificación de un problema u oportunidad y la selección de una alternativa de acción entre varias existentes, es una actividad diligente clave en todo tipo de organización."

**50. Explique Tipos de Decisiones**

Hay distintos niveles en una organización, cada uno tiene distintos requerimientos de información para el soporte de decisiones y responsabilidad para distintos tipos de decisiones.

Las decisiones se clasifican como: estructuradas, semiestructuradas y no estructuradas.



* No estructuradas: son aquellas en las que el encargado de tomarlas debe proveer un juicio, una evaluación y una perspectiva para resolver el problema. Cada una de estas decisiones es novel, importante y no rutinaria, por lo que no hay un procedimiento bien comprendido o acordado para tomarlas. *Por ejemplo:* establecer los objetivos a cinco o 10 años de la firma o decidir sobre los nuevos mercados en los que se debe participar.
* Estructuradas:son repetitivas y rutinarias; además se requiere un procedimiento definido para manejarlas, de modo que, cada vez que haya que tomarlas, no se consideren como si fueran nuevas.. *Por ejemplo*: los gerentes rara vez tienen que preocuparse por el ramo salarial de un empleado recién contratado porque, por regla general, las organizaciones cuentan con una escala de sueldos y salarios para todos los puestos.
* Semiestructuradas: Muchas decisiones tienen elementos de ambos tipos de decisiones en donde sólo una parte del problema tiene una respuesta clara proporcionada por un procedimiento aceptado. *Por ejemplo:* Una típica decisión gerencial de nivel medio podría ser “¿por qué el informe de cumplimiento de los pedidos muestra un declive durante los últimos seis meses en un centro de distribución en Minneapolis?” Este gerente de nivel medio obtendrá un informe del sistema empresarial sobre la actividad de los pedidos y la eficiencia operacional en el centro de distribución de Minneapolis. Esta es la parte estructurada de la decisión. Sin embargo, antes de llegar a una respuesta, este gerente de nivel medio tendrá que entrevistar empleados y recopilar más información no estructurada de fuentes externas en relación con las condiciones económicas locales o las tendencias de ventas.

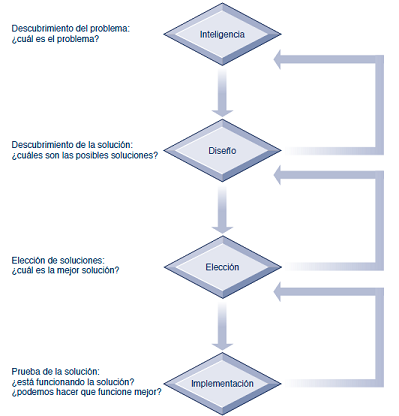
En general, las decisiones estructuradas son más prevalentes en los niveles más bajos de la organización, los problemas no estructurados son más comunes en los niveles más altos de la firma, y la gerencia de nivel medio se enfrenta a escenarios de decisiones semi- estructuradas.

**51. Desarrolle las Etapas en el Proceso de Toma de Decisiones.**

Simon (1960) describió cuatro distintas etapas en la toma de decisiones:

* La *inteligencia* consiste en descubrir, identificar y comprender los problemas que ocurren en la organización: por qué existe un problema, en dónde y qué efectos tiene sobre la firma.
* El *diseño* implica identificar y explorar varias soluciones para el problema. Se identifican y enumeran todas las alternativas, estrategias o vías de acción posibles.
* La *elección* consiste en elegir una de varias alternativas de solución.Para ello, se evalúan todas las líneas de acción teniendo en cuenta la concordancia de los objetivos de la empresa y los recursos. Además, la alternativa elegida debe ser factible y contribuir a la resolución del problema.
* La *implementación* implica hacer que funcione la alternativa elegida y continuar monitoreando qué tan bien funciona esa solución.

¿Qué ocurre si la solución que ha elegido no funciona? Se puede regresar a una etapa anterior en el proceso de toma de decisiones y repetirla si es necesario.



**52. Explique el modelo Racional Modelo de Racionalidad Limitada y el Modelo Político.**

MODELO RACIONAL.

Es aquel en el cual se sigue una serie de pasos para que aquella toma de decisiones sean lógicas y bien fundamentadas. La toma de decisiones está sustentada en un proceso lógico. Este tipo de modelo se aplica en la toma de decisión en la que se tiene el tiempo para llevar a cabo dicha toma de decisiones.  
  
 RACIONALIDAD LIMITADA.

Este modelo describe todas las limitaciones que tenemos dentro de nuestra organización y por lo tanto el proceso de toma de decisiones es limitado a las alternativas de solución.

MODELO POLÍTICO.  
 Describe la toma de decisiones de las personas para satisfacer sus propios intereses. Requiere de una solución casi de forma inmediata al problema y tiende a dar resultados favorables en el proceso de toma de decisiones inmediato. El modelo político tiende a ser egoísta ya que la decisión se basa en solo unos cuantos.

**53. Sistema de Soporte a las Decisiones (DSS): Definición y Características**

Un Sistema de Soporte a las Decisiones (DSS) es un sistema informático utilizado para servir de apoyo, más que automatizar, el proceso de toma de decisiones. La decisión es una elección entre alternativas basadas en estimaciones de los valores de esas alternativas. El apoyo a una decisión significa ayudar a las personas que trabajan solas o en grupo a reunir inteligencia, generar alternativas y tomar decisiones. Apoyar el proceso de toma de decisión implica el apoyo a la estimación, la evaluación y/o la comparación de alternativas.

Estos sistemas de apoyo son del tipo OLAP (Procesamiento Analítico en Línea) o de minería de datos, que proporcionan información y soporte para tomar una decisión. Los sistemas OLAP se utilizan para facilitar la consulta de grandes cantidades de datos, haciendo uso de estructuras multidimensionales (cubos OLAP) que contienen datos resumidos de grandes bases de datos o sistemas transaccionales (OLTP). Se utilizan en informes de negocios de ventas, marketing, informes de dirección, minería de datos y áreas similares.

Características principales

· Brindan apoyo a la toma de decisiones que no es rutinaria. Se enfocan en problemas que son únicos y cambian con rapidez, para los cuales el proceso para llegar a una solución tal vez no esté por completo predefinido de antemano.

· Su principal característica es la capacidad de análisis multidimensional (OLAP) que permite profundizar en la información hasta llegar a un alto nivel de detalle, analizar datos desde diferentes perspectivas, realizar proyecciones de información para pronosticar lo que puede ocurrir en el futuro, análisis de tendencias, análisis prospectivo, etc.

· Permite extraer y manipular información de una manera flexible.

· Permite al usuario definir interactivamente qué información necesita y cómo combinarla.

· Suele incluir herramientas de simulación, modelización, etc.

**54. ¿Cuáles son las herramientas para la Toma de Decisiones?**

Permite manejar grandes cantidades de información, clasificarla, interrelacionarla y tomar decisiones inteligentes.

1. Tablero de Control

Facilita la toma de decisiones a los socios y ejecutivos de una empresa ya que se tiene la información de manera inmediata de las diferentes áreas y permite detectar inmediatamente las desviaciones de los planes, programas y estrategias y decidir las medidas correctivas.

1. Cuadro de marco Integral/ Balance Scorecard (BSC)

Sugiere que veamos a la organización desde cuatro perspectivas, cada una de las cuales debe responder a una pregunta determinada, Para cada perspectiva se definen metas e indicadores de mediciones.

* Desarrollo y Aprendizaje (Learning and Growth): ¿Podemos continuar mejorando y creando valor?
* Interna del Negocio (Internal Business): ¿En qué debemos sobresalir?
* Del cliente (Customer): ¿Cómo nos ven los clientes?
* Financiera (Financial): ¿Cómo nos vemos a los ojos de los accionistas?

**55. Diferencias entre una Cuadro de Marco Integral/ Balance Scorecard (BSC).**



**56. Defina Dinámica de Sistemas**

La dinámica de sistemas es una metodología para analizar, modelar y entender el comportamiento temporal en entornos complejos. Se basa en la identificación de los bucles de realimentación entre los elementos, y también en las demoras en la información y materiales dentro del sistema. De esta manera se puede estructurar a través de modelos matemáticos la dinámica del comportamiento de estos sistemas reales, con el objetivo de mejorar la comprensión de los mismos y rediseñarlos para lograr una mayor eficiencia.

Actualmente se usa en el sector público y privado para el análisis y diseño de políticas, ya que es posible armar modelos que simulen los impactos de dichas decisiones. La base de la metodología es el reconocimiento de que la estructura de cualquier sistema es tan importante como determinar el comportamiento individual de cada uno de sus elementos.

**57. ¿Qué son los Modelos Mentales?**

Un modelo mental es todo aquel que se genera como representación de la realidad que nos rodea. Sólo posee aquellos elementos considerados más relevantes para la representación de esta realidad, por lo que se trata de un modelo incompto de la misma.

Hablando en el contexto de la organización, sería un modelo general que todos sus integrantes poseen de cuales son y cómo interactúan las distintas partes que conforman a la misma y/o a su exterior.

La dinámica de sistemas plantea modelar estos sistemas mentales a través de ciertas herramientas que luego permitan generar modelos formales. Estos últimos son modelos matemáticos que permiten expresar hechos, variables, parámetros, entidades y relaciones entre todos ellos de forma precisa. Mediante la formulación de ecuaciones es posible describir cualquier tipo de interacción que genere efectos a lo largo del tiempo.

**58. Desarrolle Diagrama Causales.**

Es un diagrama que recoge los elementos clave del Sistema y las relaciones entre ellos. Tienen el propósito de mostrar de manera sencilla modelos mentales acerca de las estructuras y estrategias del sistema.

La relación entre una variable A y otra B del sistema se representará mediante una flecha,



leyéndose "A influencia a B".

Sobre la flecha se indicará mediante signo + o - el tipo de relación



siendo positiva cuando las variaciones de A y B son del mismo sentido, y



negativa en caso de variación de sentido contrario.

Los diagramas causales se clasifican según su estructura en:

1. diagramas abiertos, de estructura simple



b) diagramas cerrados, de estructura compleja o bucles de realimentación



**59. ¿Cuál es la Metodología de Aplicación de la Dinámica de Sistemas?**

Según establecen los autores, el proceso de modelización y simulación es de naturaleza iterativa.

1. Identificación del problema. Dicho problema, para poder ser modelado, ha de tratarse de un problema dinámico en el cual pueda detectarse algún tipo de realimentación.
2. Una vez definido el problema, es preciso identificar aquellos factores considerados relevantes que parecen interactuar y ser los causantes del problema de partida.
3. Deben detectarse los circuitos o bucles de realimentación existentes y que enlazan las decisiones con las acciones.
4. Una vez detectadas las realimentaciones existentes, la siguiente fase consistirá en formular las políticas que sirven de guía en la toma de decisiones dentro del sistema analizado.
5. Teniendo en cuenta las variables y decisiones detectadas en las fases tercera y cuarta respectivamente, se elaborará un modelo matemático que recoja las políticas de decisión previamente definidas, así como las fuentes de información existentes.
6. A partir de las ecuaciones que componen el modelo matemático, se generará el comportamiento o evolución del sistema estudiado a lo largo del tiempo.
7. En la medida en que los resultados obtenidos de la simulación con el modelo se aproximen a la evolución real, el modelo será validado.
8. Revisión del modelo hasta tanto sea aceptable como representación del sistema real.
9. Una vez que el modelo ha sido contrastado, podrá ser empleado para la simulación de nuevas políticas frente a las ya existentes con el fin de establecer cuales son los cambios adecuados para mejorar el comportamiento o evolución del sistema en su conjunto.
10. Llevar a la práctica aquellas políticas o decisiones que, como consecuencia de la simulación se han perfilado como las más adecuadas para eliminar las alteraciones o fluctuaciones detectadas en la evolución del sistema.

**60. Desarrolle Sistemas Estables e Inestables y Sistemas con factor externo.**

Un sistema es estable cuando se haya dominado por un bucle negativo. La estructura básica de los sistema estables está formada por un Estado Deseado y por un Estado Real del sistema, estos dos estados se comparan (Diferencia), y en base a este valor el sistema toma una Acción para igualar el Estado Real al Deseado.



Por otro lado, los sistemas inestables son aquellos que están dominados por un bucle de realimentación positivo, lo que produce que “crezcan” ininterrumpidamente. Un caso de sistema inestable es la reproducción de conejos.



Un sistema que basa sus objetivos en la realidad, y únicamente pretende mejorarla está abocado a una tendencia permanente a los bajos resultados. Por otro lado, si obtiene sus objetivos de una base externa al sistema es inmune a este tipo de procesos.



El sistema en sí ya es capaz de desarrollar acciones para lograr igualar el Estado Real al Estado Deseado, reduciendo paulatinamente la Diferencia; ahora con la ayuda externa esta reducción es mucho más rápida y habitualmente es igual a cero al cabo de unos pocos periodos.

**61. Desarrolle Sistemas Hiperestables, Oscilantes y Sigmoidales.**

Cuando un sistema está formado por múltiples bucles negativos, cualquier acción que intenta modificar un elemento no se ve contrarrestado sólo por el bucle en el que se haya dicho elemento, sino por todo el conjunto de bucles negativos que actúan en su apoyo, súper-estabilizando el sistema.



Sistemas oscilantes se trata de sistemas que, en su intento de alcanzar el estado deseado, comienzan a fluctuar alrededor del mismo, sin lograr llegar a un estado de estabilidad.

Esta fluctuación se puede dar por dos factores:

La presencia de efectos secundarios a las acciones realizadas para corregir la Diferencia, los cuales no son tenidos en cuenta y afectan al Estado Real resultante.  


* La presencia de demoras entre las distintas interacciones de las partes del sistema, las cuales generan acumulaciones indeseadas y resultados indeseados.
* 

Los sistemas sigmoidales se refiere a aquellos sistemas que se comportan de manera similar a la descrita por la función sigmoide, exhibiendo un crecimiento acelerado hasta estabilizarse.

Es el caso de sistemas inestables (con un lazo de retroalimentación positiva) donde se agrega un lazo de retroalimentación negativa para regular la tasa de aumento neto.



62. Defina Arquetipos Sistémicos.

Los arquetipos sistémicos son patrones de comportamiento de un sistema, son situaciones que se repiten permanentemente, por costumbre, generando errores en el comportamiento de una organización.

Es la descripción de un proceso, que ayuda a reconocer comportamientos repetitivos, para encontrar sus puntos de apalancamiento, es decir, permite saber cuál es el cambio adecuado para eliminar el límite más importante que sufre el sistema, a través de lo cual ganará dinamismo en una forma más que proporcional.