

Seminario

Representación gráfica



Pedro A. Castillo Valdivieso

Depto Arquitectura y Tecnología de Computadores
Universidad de Granada

pacv@ugr.es

Índice



[1. Introducción]

2. Recomendaciones

3. Errores comunes

4. Manipulación de gráficos

5. Presentación de resultados



Introducción

- La representación gráfica es un paso fundamental
 - “*Una imagen vale más que 1000 palabras*”
 - Ayuda a sacar conclusiones de los estudios
 - No sacar conclusiones == no haber hecho el estudio
 - Ahorra tiempo y facilita la comprensión al lector
 - Permite enfatizar puntos interesantes
- Su diseño es responsabilidad del analista



Tipos de variables

- A la hora de realizar un gráfico, debemos considerar el tipo de variables que vamos a representar:
 - **Cualitativas** (categóricas)
 - Son un conjunto definido de valores
 - Normalmente se expresan con palabras (podremos ordenarlas alfabéticamente)
 - p.ej.: tipo de carga: científica, propósito general, videojuegos
 - **Cuantitativas:**
 - Se expresan mediante números (intervalos continuos o discretos)
 - p.ej. 2,3.5,300

Índice



1. Introducción
- [2. Recomendaciones]**
3. Errores comunes
4. Manipulación de gráficos
5. Presentación de resultados

Recomendaciones

- Para representar gráficamente el resultado de un análisis tenemos las siguientes recomendaciones (no son reglas):
- Minimización del esfuerzo por parte del lector: cuanto más fácil sea leer un gráfico, mejor.

P. ej., poner una caja con la clave del gráfico (leyenda) o etiquetar las curvas directamente

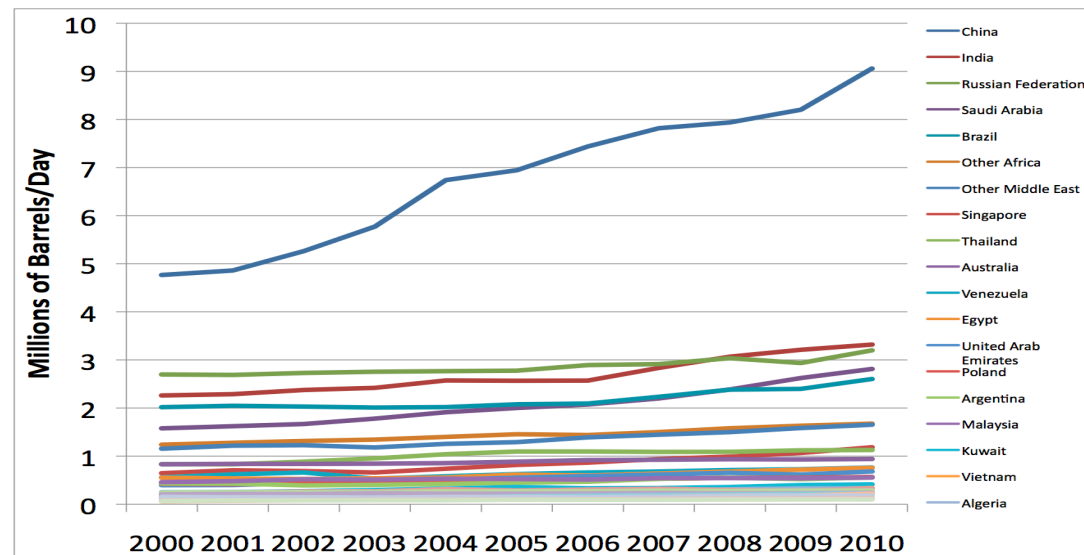


Recomendaciones

- Maximizar información: debe de haber información en el gráfico para hacerlo autosuficiente:
 - Usar palabras o nombres de variables en vez de símbolos
 - Etiquetar adecuadamente los ejes (ej. Tiempo CPU diario).
 - Añadir unidades en las etiquetas (Tiempo CPU (segundos))

- Minimizar la tinta: presentar el máximo de información con el mínimo de tinta.

“Demasiada” información hace el gráfico poco interesante; p. ej., quitar líneas de la rejilla si no son necesarias; ajustar la escala y las variables para número mínimo de cifras y rayas (ej. disponibilidad vs. indisponibilidad).





Recomendaciones

- Usar prácticas comúnmente aceptadas: como el eje X para la variable independiente, el y para la dependiente, el origen en el 0,0, las abscisas y ordenadas en orden creciente.

Si no se hace así, el lector necesitará un esfuerzo extra.

- Evitar la ambigüedad: mostrar los ejes, las subdivisiones, las medidas, escalas, identificar cada una de las barras, no mostrar varias variables



Recomendaciones (resumen)


- Minimización del esfuerzo por parte del lector
- Maximizar información
- Minimizar la tinta
- Usar prácticas comúnmente aceptadas (los ejes, el origen 0,0)
- Evitar la ambigüedad (mostrar los ejes, las subdivisiones, las medidas, escalas...)
- No poner demasiada información ni demasiada poca.

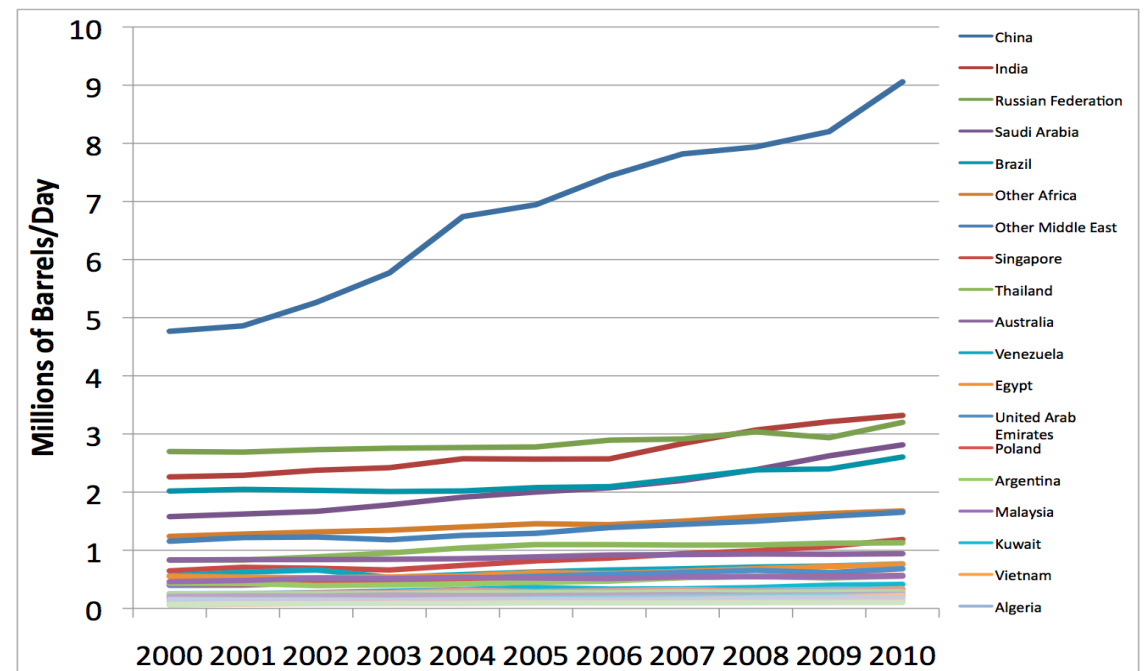
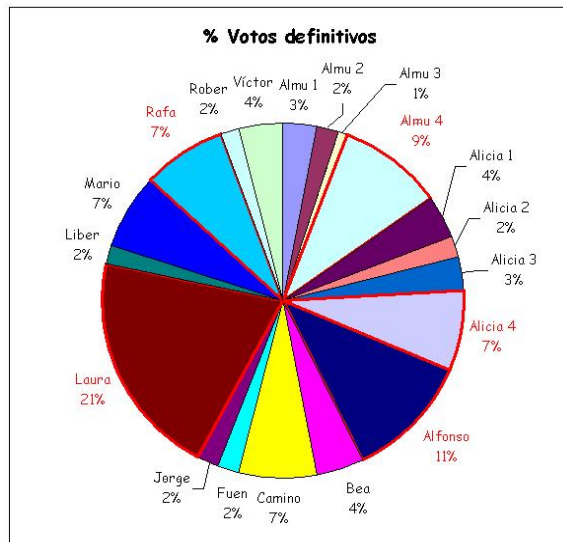
Índice



1. Introducción
2. Recomendaciones
- [3. Errores comunes]**
4. Manipulación de gráficos
5. Presentación de resultados

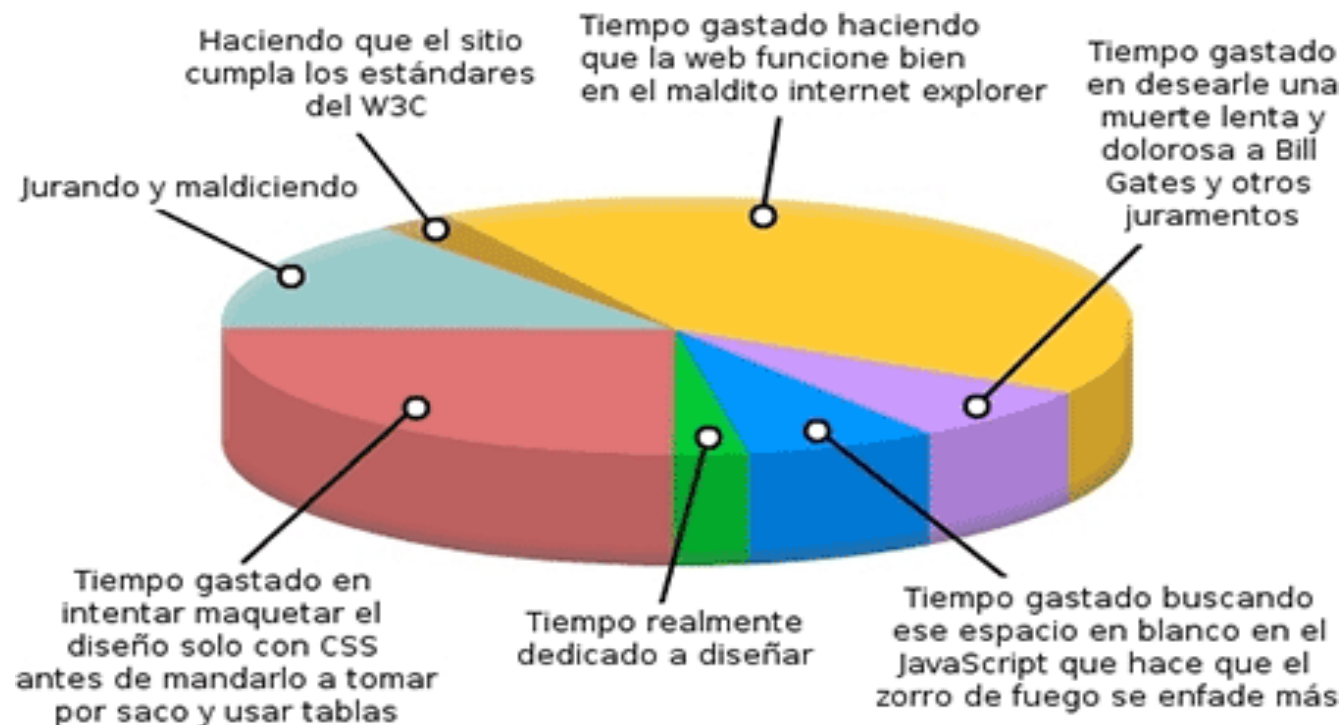
Errores comunes

- Algunos errores que se suelen cometer son los siguientes:
 - Presentar demasiadas alternativas en un solo gráfico: no se deben de meter más de cinco o seis curvas o 10 barras u 8 componentes de tarta en una sola gráfica.
 - En un histograma, cada "cubo" debería tener al menos 5 puntos de datos.
- 
- | Year | China (Billions) | India (Billions) |
|------|------------------|------------------|
| 1950 | 0.5 | 0.4 |
| 1960 | 0.6 | 0.5 |
| 1970 | 0.7 | 0.6 |
| 1980 | 0.8 | 0.7 |
| 1990 | 1.0 | 0.9 |
| 2000 | 1.2 | 1.1 |
| 2010 | 1.4 | 1.3 |
| 2020 | 1.6 | 1.5 |
| 2030 | 1.8 | 1.7 |
| 2040 | 2.0 | 1.9 |
| 2050 | 2.2 | 2.1 |



Errores comunes

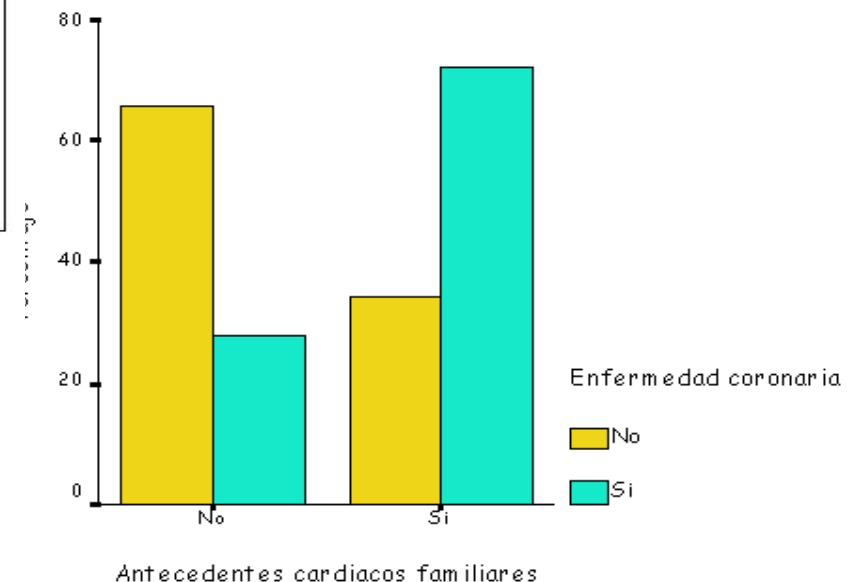
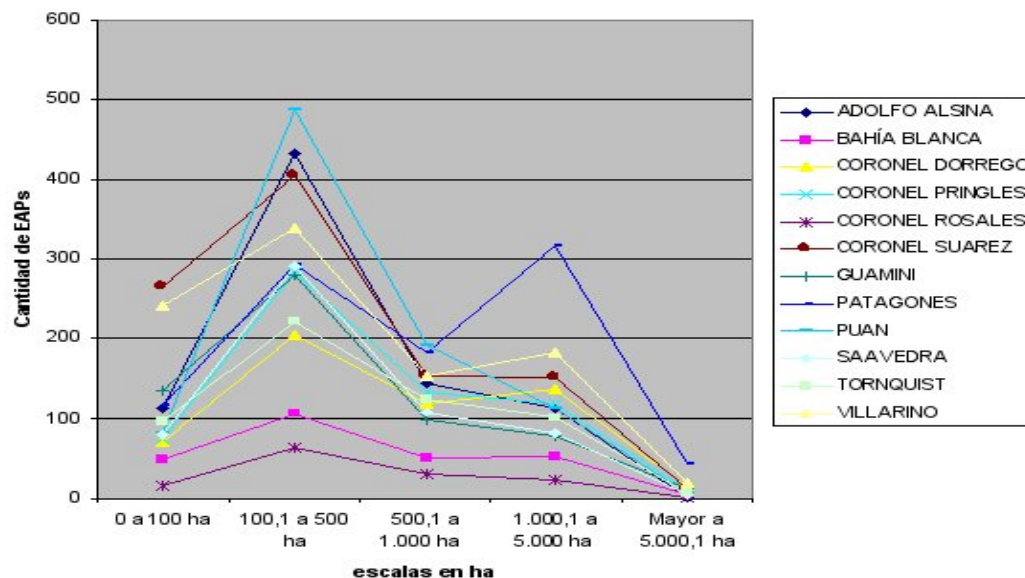
- Usar símbolos en lugar de texto: en vez del nombre dado a una variable, debería usarse lo que contiene una variable; por ejemplo, en vez de %CPU, se debería usar Porcentaje de CPU dedicado al usuario, en un tipo de letra más pequeño.



Errores comunes

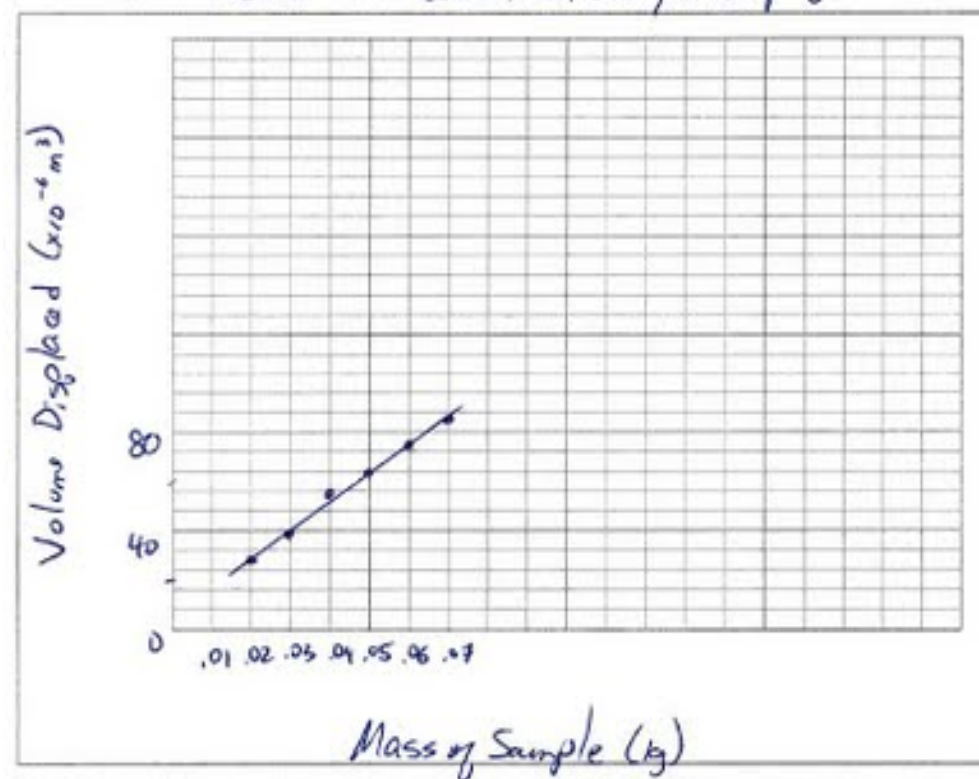
- Añadir información extra “confusa”, p.ej. usar un gráfico de líneas en vez de uno de barras en variables categóricas (no se puede interpolar).

EAPs: cantidad y escalas de extensión



Errores comunes

- Seleccionar rangos de escala de manera incorrecta:
si los mínimos y máximos no permiten ver con claridad los resultados es necesario reajustarlos





Errores comunes (resumen)

- Presentar demasiadas alternativas en un solo gráfico
- Presentar muchas variables y (abscisas) en un solo gráfico
- Usar símbolos en lugar de texto
- Seleccionar mal las escalas
- Usar un gráfico de líneas en vez de uno de barra (variables categóricas => gráfico de barras)

Índice



1. Introducción
2. Recomendaciones
3. Errores comunes
- [4. Manipulación de gráficos]**
5. Presentación de resultados



Manipulaciones

- La representación utilizada nos permite manipular la información para mostrar “medias-verdades”

Manipulaciones

- Usar orígenes no nulos para hacer énfasis de la diferencia: cuando la diferencia es pequeña, se usa un origen no nulo para que parezca más grande.



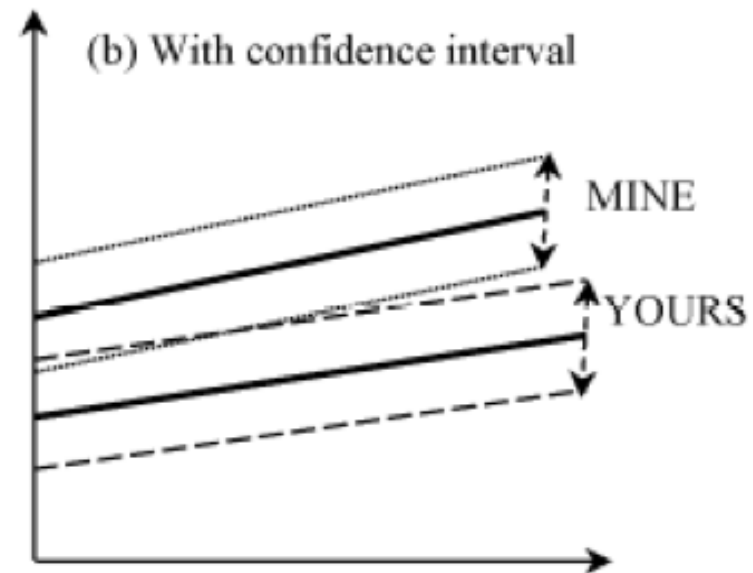
Manipulaciones

- Usar orígenes no nulos para hacer énfasis de la diferencia: cuando la diferencia es pequeña, se usa un origen no nulo para que parezca más grande.



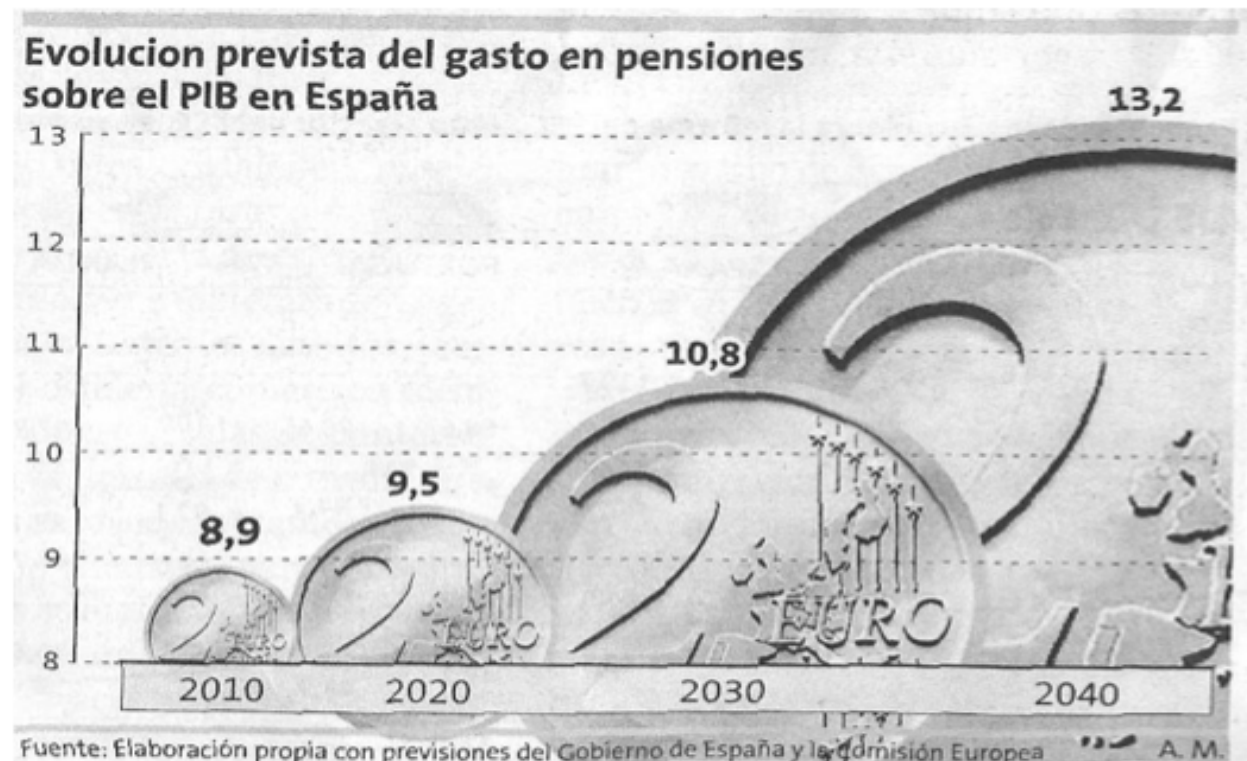
Manipulaciones

- Trazar cantidades aleatorias sin los intervalos de confianza: la media transmite poca información si no se acompaña de la desviación típica, o rango o algún momento.



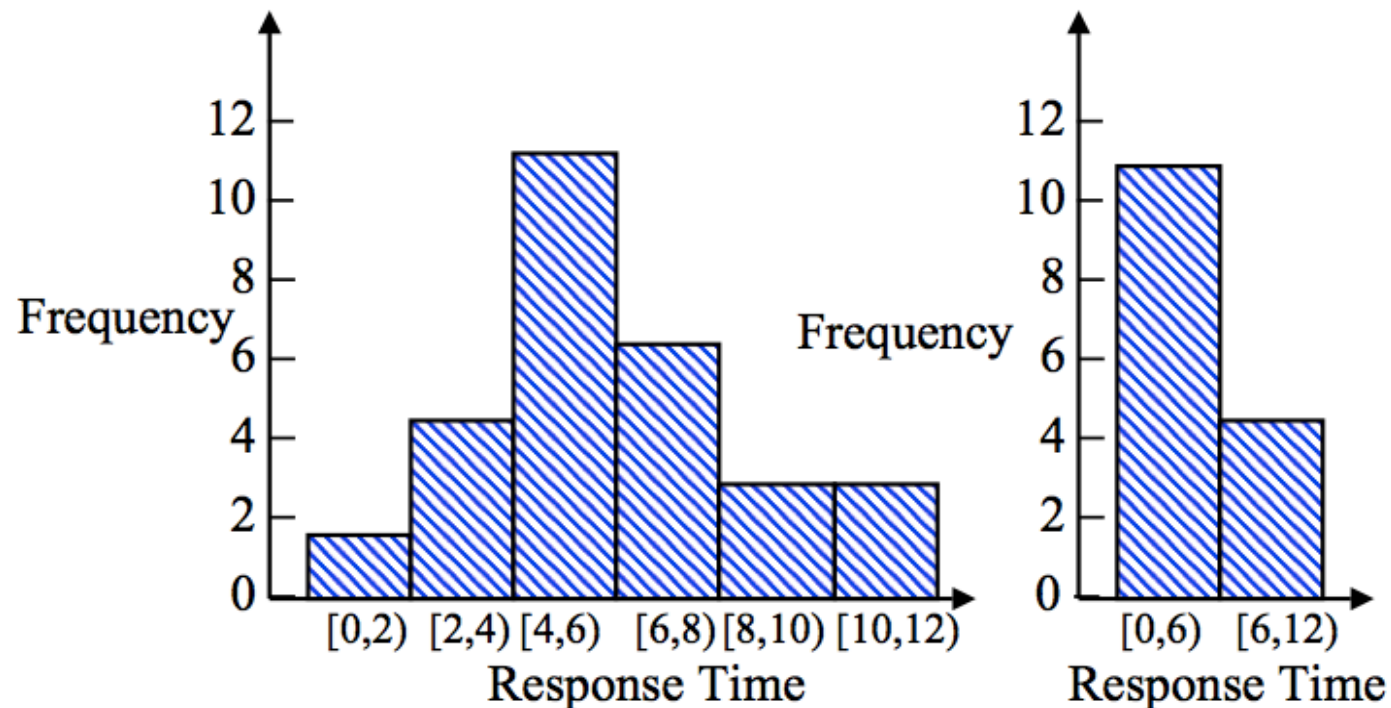
Manipulaciones

- Usar pictogramas escalados por altitud: a veces se usan barquitos, o cochecitos, o similares para representar una serie de datos. Siempre hay que tener en cuenta que es la altura, y no la superficie, lo que representa la cantidad; si se mira solo la superficie, puede resultar que una cantidad parezca n^2 veces mejor que otra, cuando solo es n veces mejor.



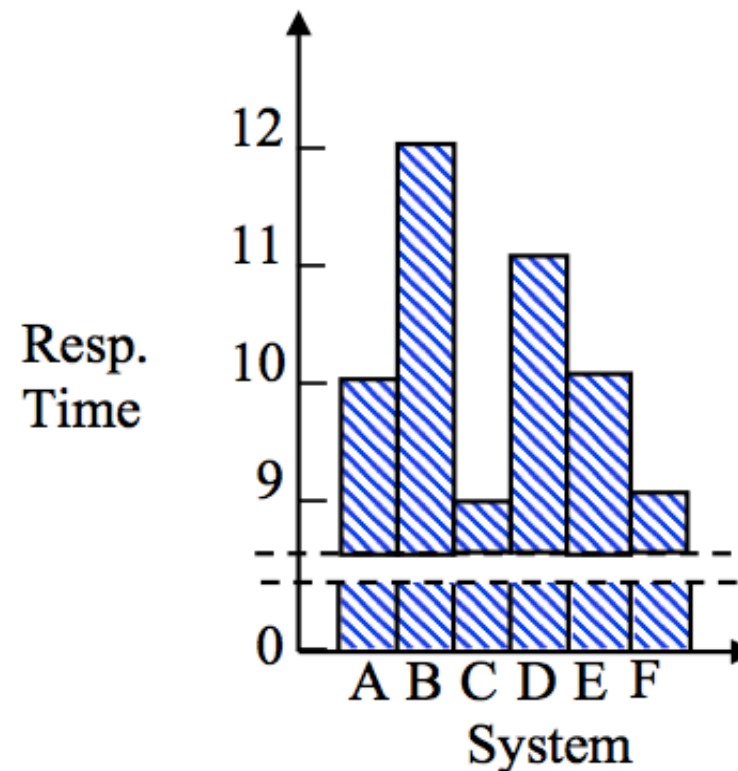
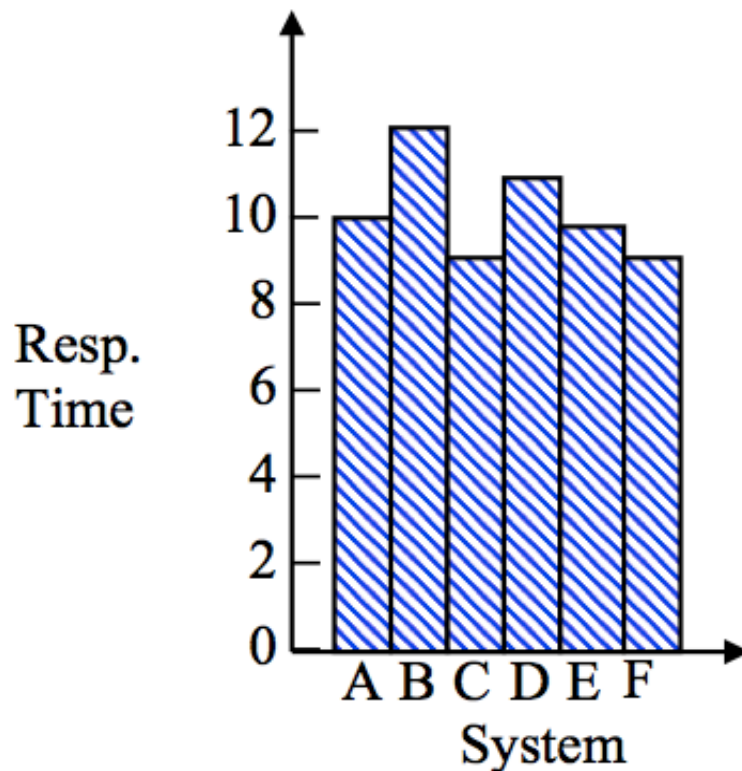
Manipulaciones

- Usar tamaños de celda no adecuados en histogramas
 - hay que aplicar tests estadísticos para comprobar que se haya hecho adecuadamente



Manipulaciones

- Usar escalas partidas en gráficos de columnas.
 - equivalente a cambiar el inicio de los ejes.



Índice



1. Introducción
2. Recomendaciones
3. Errores comunes
4. Manipulación de gráficos
- [5. Presentación de resultados]**



Presentación de resultados. Introducción

- La visualización es esencial en general, pero en evaluación de prestaciones sirve para:
 - **apreciar de un vistazo cuál es el estado del sistema**
 - si hay algún problema, **cómo identificarlo y resolverlo.**
 - visualizar la figura de mérito
 - es una cuantificación para medir el comportamiento del sistema

Presentación de resultados. Introducción

- Para presentar y analizar los resultados obtenidos de la ejecución de un monitor sobre un sistema o una comparativa entre varios sistemas normalmente se usan gráficos de Gantt, o gráficos de Kiviat.

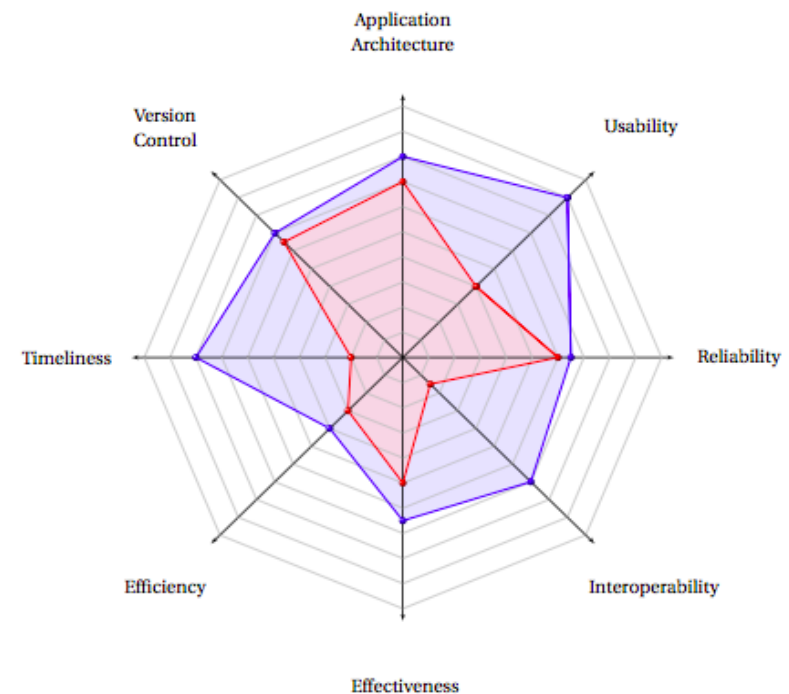
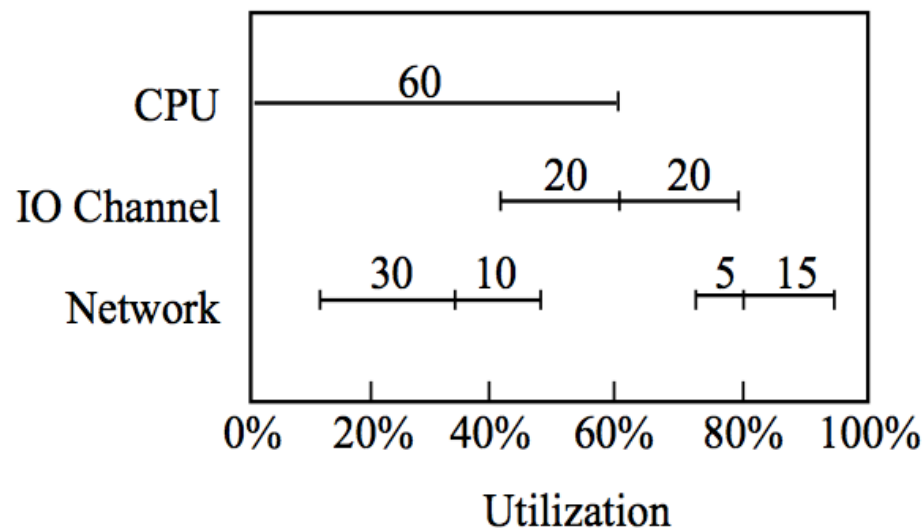


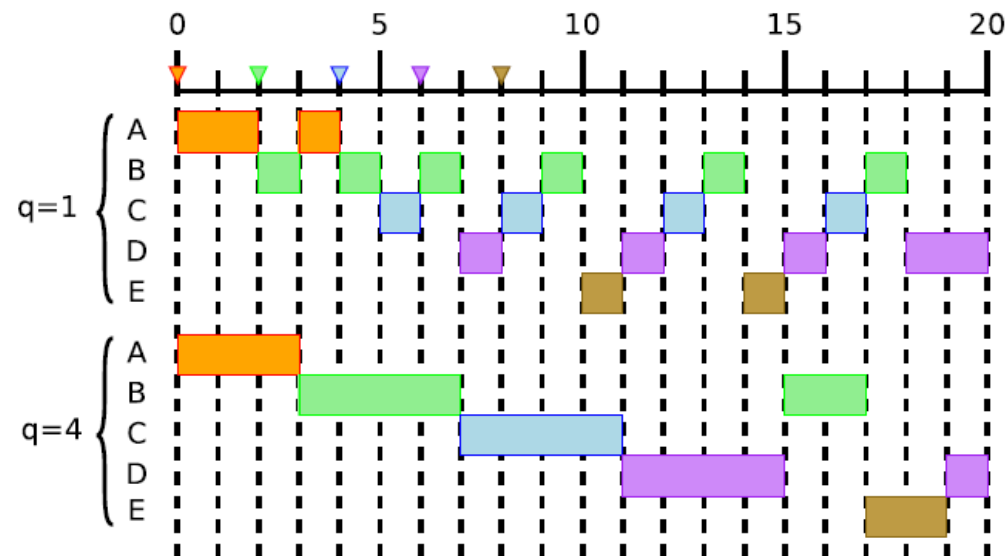


Gráfico de Gantt

- En un gráfico de Gantt, se representa en abscisas el tiempo, y en ordenadas una línea que representa los instantes durante los cuales un recurso ha estado ocupado.
- Permiten ver el grado de solapamiento de utilización de ciertos recursos
 - Identificará los cuellos de botella
- Se suelen usar gráficos similares en ingeniería del software, y, en general, en planificación de proyectos
 - Análisis PERT (Program Evaluation and Review Technique)

Diagramas de Gantt

- Muestran una condición booleana a lo largo del tiempo, p.ej. “*CPU idle*” true/false
- Cada condición se traza con una línea en horizontal





Gráficos de Kiviat

- Representan los porcentajes de uso y solapamiento de diferentes componentes del sistema en los radios de un círculo.
- Teóricamente, es posible ver rápidamente el problema que tiene un sistema.

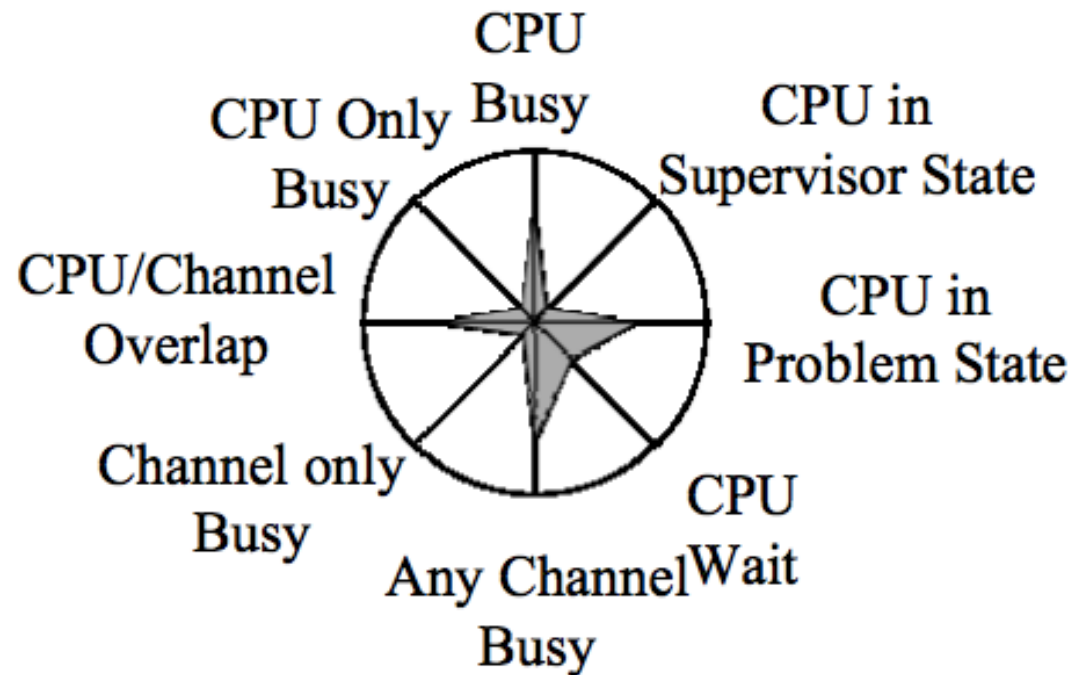
http://en.wikipedia.org/wiki/Radar_chart

http://www.orcacomputer.com/eeHelp/Kiviat_Graphs.htm

<http://it.toolbox.com/blogs/enterprise-solutions/better-kiviat-diagrams-19868>

Gráficos de Kiviat

- Habitualmente está dividido en 8 sectores donde se suele colocar intercaladas magnitudes del tipo “más alto es mejor” con magnitudes del tipo “bajo es mejor”.



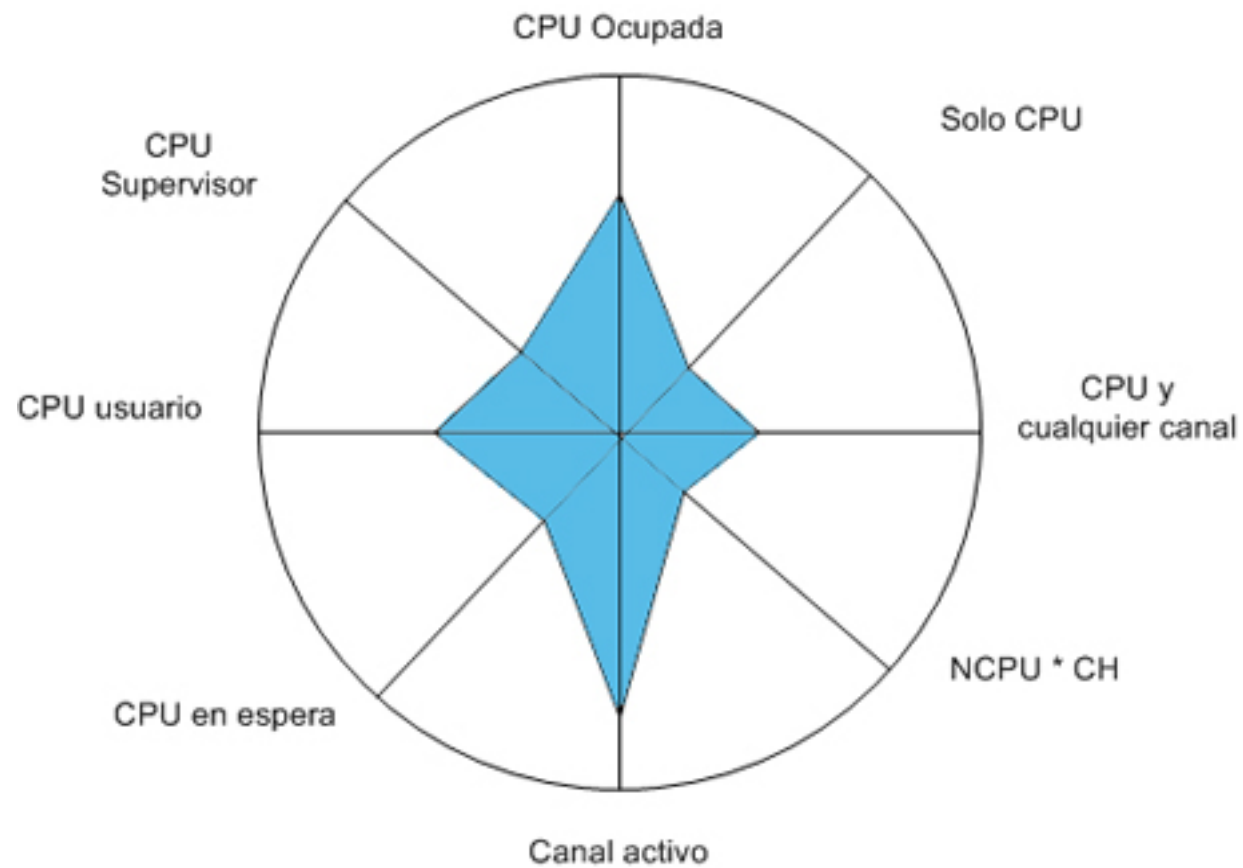


Gráficos de Kiviat

- Se selecciona un número par de variables que hay que estudiar, frecuentemente ocho, la mitad de ellas deben ser buenos índices de prestaciones y la otra mitad malos.
- Se divide el círculo en tantos sectores como variables hay que representar.
- Se enumeran los semiejes secuencialmente, normalmente en sentido horario, comenzando por el semieje vertical superior.
- Se asocian los buenos índices de prestaciones a los semiejes impares y los malos a los pares.

Gráficos de Kiviat

Obtendremos:



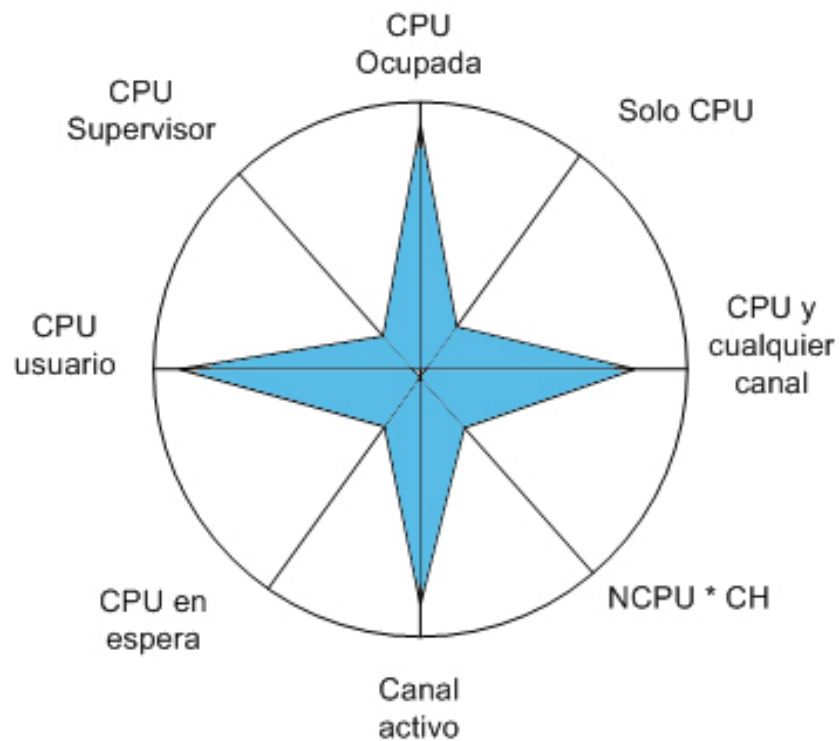


Gráficos de Kiviat

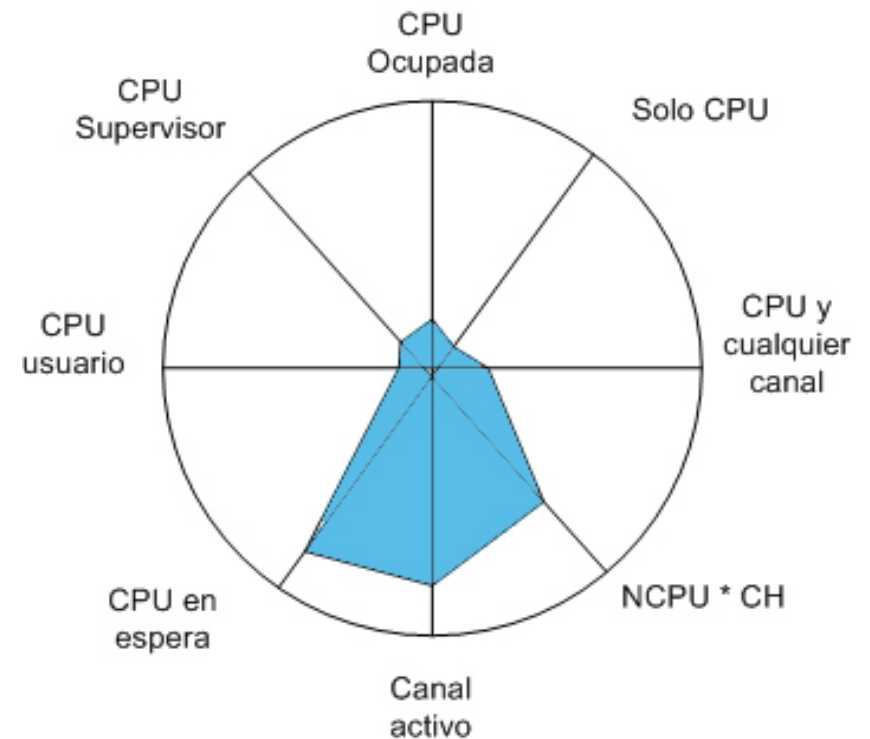
- Diferentes formas resultantes indican diferentes "tipos" de sistemas, el sistema perfecto sería el denominado estrella de Kiviat, una estrella de 3 o 4 puntas.
- Hay que mirar siempre que los indicadores a las 12, 3, 6 y 9 sean altos, y los demás bajos.
- Valores altos de los indicadores impares indican un sistema que no funciona como debiera
- Si los valores pares tienen valores muy diferentes, eso suele indicar un sistema desequilibrado.

Gráficos de Kiviat

Situación ideal:

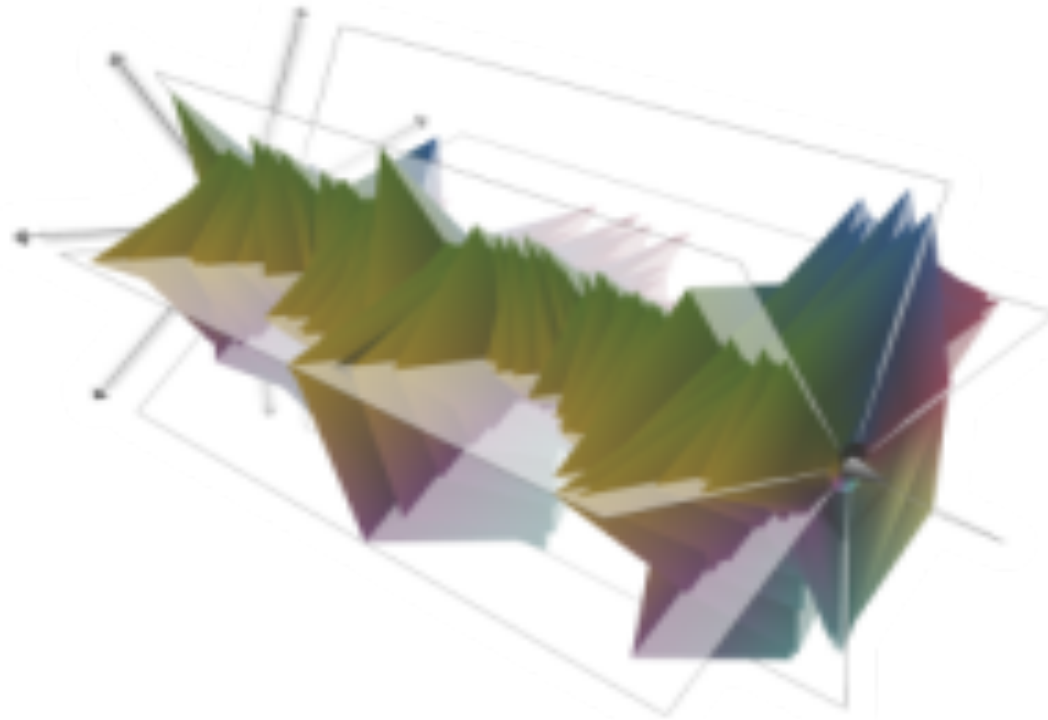


Poco uso de CPU
y mucho de E/S:



Gráficos de Kiviat

- Al gráfico de Kiviat se le puede añadir una tercera dimensión temporal, formando un tubo de Kiviat, que indica la progresión del gráfico de Kiviat a través del tiempo





Gráficos de Kiviat. *FOM*

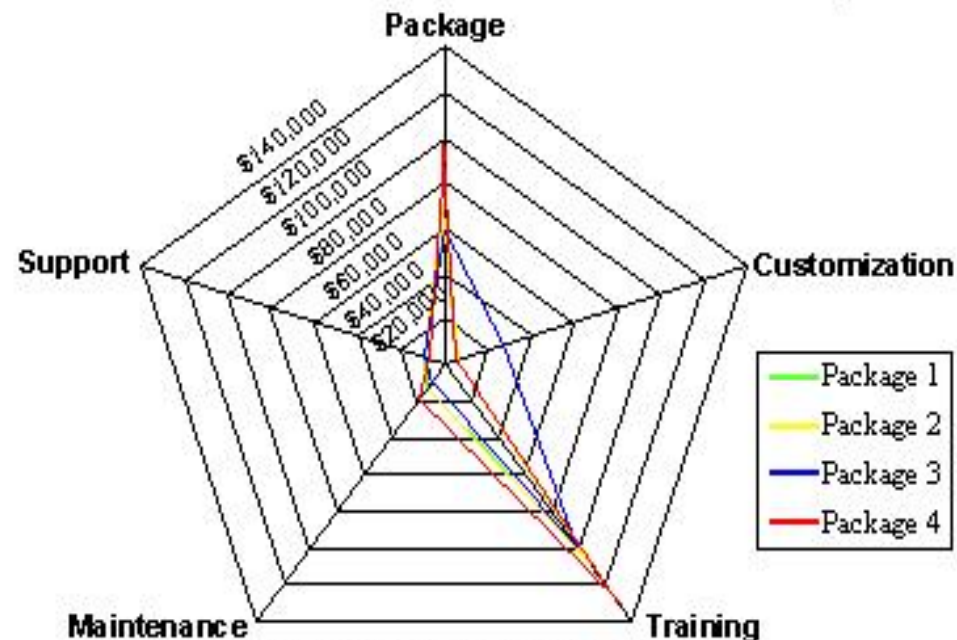
- Figura de mérito (*figure of merit, FOM*).
- Es una cuantificación para medir el comportamiento del sistema.
- La idea es calcular el área del polígono formado y así comparar los sistemas.
- Puede resultar engañoso (ejes con diferentes escalas; valores límites no tienen por qué ser buenos; sistemas con igual FOM no tienen por qué ser igual de buenos).

¿Cómo hacer gráficos Kiviat?

<http://it.toolbox.com/blogs/enterprise-solutions/better-kiviat-diagrams-19868>

secuencia de pasos
detallados para
construir un gráfico
de este tipo

Kiviat Diagram Sample
Evaluation of the costs of different software packages



The smallest footprint is best.

	Package	Customization	Training	Maintenance	Support
Package 1	\$65,000	\$5,000	\$110,000	\$12,000	\$5,000
Package 2	\$70,000	\$6,000	\$110,000	\$12,600	\$5,000
Package 3	\$80,000	\$25,000	\$100,000	\$10,800	\$6,000
Package 4	\$98,000	\$5,000	\$130,000	\$19,600	\$4,000

¿Cómo hacer gráficos Kiviat?

<http://tex.stackexchange.com/questions/10240/problem-drawing-kiviat-diagram>
en LaTeX:

```
\begin{tikzpicture}
\tkzKiviatDiagramFromFile[
  scale=.5,
  label distance=.5cm,
  gap = 1,label space=3,
  lattice = 10]{tableae.dat}
\tkzKiviatLineFromFile[thick,
  color = blue,
  mark = ball,
  ball color = blue,
  mark size = 4pt,
  fill = blue!20]{tableae.dat}{2}
\tkzKiviatLineFromFile[thick,
  color = red,
  mark = ball,
  ball color = red,
  mark size = 4pt,
  fill = red!20]{tableae.dat}{1}
\end{tikzpicture}
```

