



## VIII SEMINARIO DE INFORMÁTICA EN LA UC

# DATA WAREHOUSE





# Esquema de la presentación

- Parte I – Una reflexión
- Parte II – ¿Cómo se implementa?
- Parte III – Un ejemplo práctico



# Parte I - Una reflexión



# Data Warehouse

(Almacén de datos)



# Business Intelligence (BI)

(Inteligencia de Negocios (BI))



# ¿Qué es esto de “Inteligencia de Negocios”?

¿Cómo puede ser que un negocio llegue a ser “inteligente”?



# Tal vez debiéramos preguntarnos antes

¿Qué entendemos por Inteligencia?



# Para tratar de contestar esa pregunta, permítanme una reflexión.

Estamos claros que los seres humanos no son como las organizaciones, pero si estamos hablando de inteligencia no podemos dejar de lado algún tipo de analogía con la inteligencia humana.



# ¿Cuándo de decimos que un ser humano es inteligente?

Sin detenernos a pensar demasiado, podríamos decir que:





# Se entendió ya la idea ¿No?

¿Y eso es la inteligencia?



Un hombre inteligente debiera utilizar su capacidad de análisis para conocerse mejor y de esa manera alcanzar su plena madurez.



# Y la madurez definitivamente no es algo simple de alcanzar.

De hecho, (desde nuestra dimensión “más humana”) todos tenemos que seguir nuestro propio camino de madurez.



# Hay caminos fáciles a la madurez.





# Otros no parecen tan fáciles





Finalmente, un ser humano maduro ha sabido procesar su experiencia y ha aprendido a conocerse a sí mismo.

Y con ese auto-conocimiento tiene claras sus potencialidades y ha aprendido a convivir con sus debilidades. Y se analiza constantemente con el propósito de mejorar de manera continua.



Y además es maduro como darse cuenta  
que con su inteligencia y habilidades,  
realiza un aporte a la sociedad.

O dicho de otra manera, se ha  
convencido de que tiene una  
misión en la vida.



# ¿Y si pudiéramos dotar a un negocio de una inteligencia similar?

¿Qué haría la UC por ejemplo si pudiera desarrollar aún más su inteligencia?



Me gustaría pensar que la UC usaría todo lo que sabe hacer –y sabe hacer muchas cosas – para tratar de ser una institución más madura.



Saber, fundamentada en datos objetivos, qué cosas hace bien y en qué hay que mejorar. Saber como puede mejorar para aportar –aún más– a la sociedad.



La UC conociéndose de mejor manera con el objeto de saber finalmente como *servir* mejor a su misión.



Definición (Personal):

Un negocio inteligente es aquel implementa procesos que le permiten analizar la información que dispone, con el fin de ayudar al cumplimiento de su plan estratégico.



## Parte II - ¿Cómo se implementa?



# En términos simples, ahora hay que realizar dos labores

1. Generar una base de datos de análisis.
2. Sobre esos datos, montar una herramienta que facilite el análisis y presentación de los datos.



# Primero, debemos entender el concepto de bases de datos operacionales

- Son las bases de datos que almacenan los datos de los sistemas operacionales.
- Orientados a la operación.
- Orientados a la transacción.
- Usualmente están en diferentes plataformas.
- Están estructurados orientados a la unicidad e integridad de los datos.

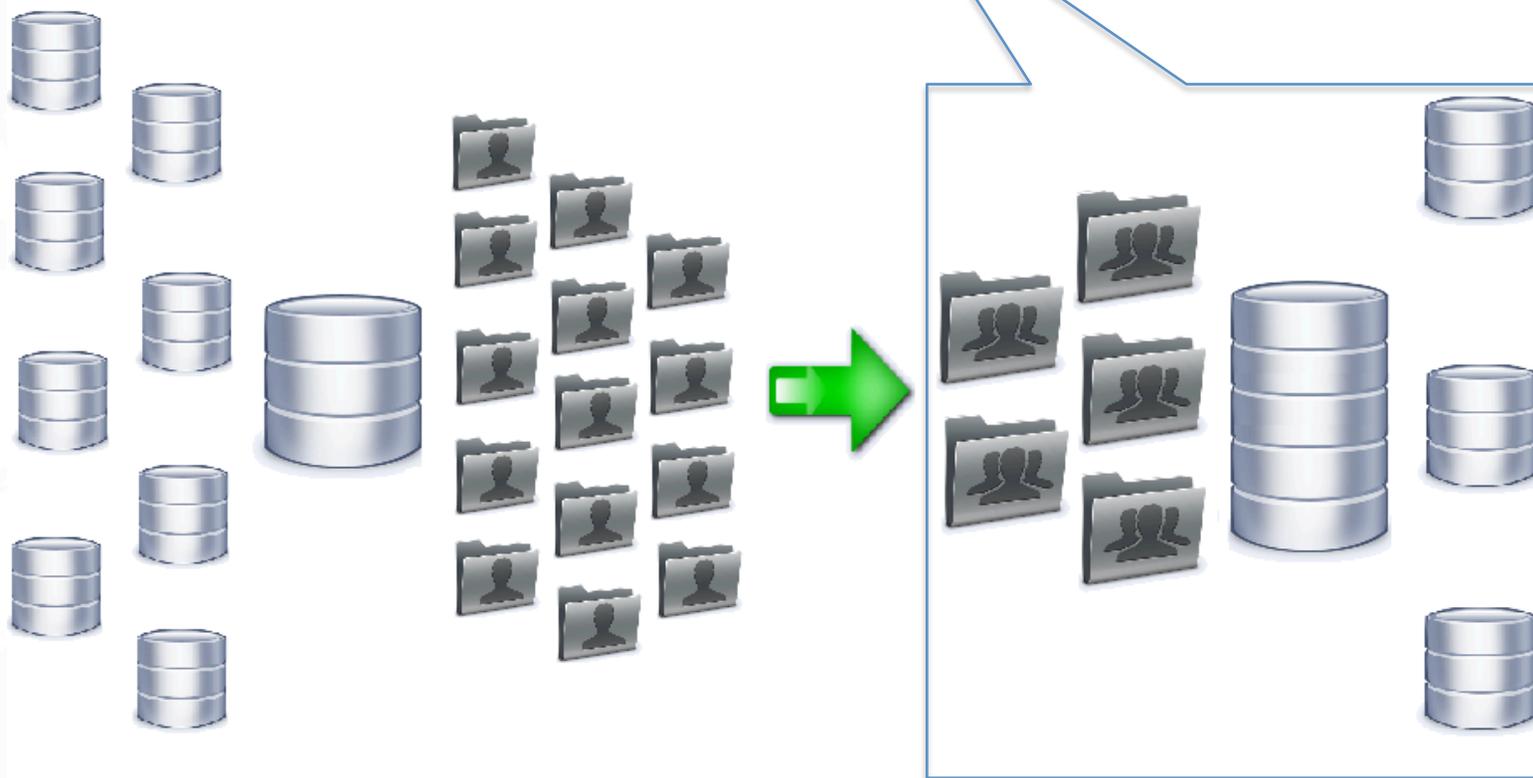


# Sistemas operacionales en la UC





# Se genera el Data Warehouse



Base de datos operacional  
(OLTP: On-Line Transactional Processing)

Base de datos de análisis  
(OLAP: On-Line Analytical Processing)



# Este Data Warehouse tiene las siguientes características

- Estar separada de las bases de datos operacionales.
- Debe contener un resumen de la información de los sistemas operacionales.
- Debe contener información consolidada desde las diferentes fuentes.
- Están estructurados para cumplir con los requerimientos de información definidos.



Bases de datos operacional (OLTP)	Bases de datos de análisis (OLAP)
Orientada a aplicaciones	Orientada a datos
Detallada	Resumida y refinada
Actualizada en línea	Histórica
Uso administrativo	Uso directivo
Datos se corrigen constantemente	Datos no se corrigen
Rendimiento es crítico	Rendimiento no es crítico
No redundante	Puede ser redundante



# Por ejemplo:

## Tenemos la información de un alumno (detalle)



RUT:

Nombre:

Año Admisión:

Carrera:

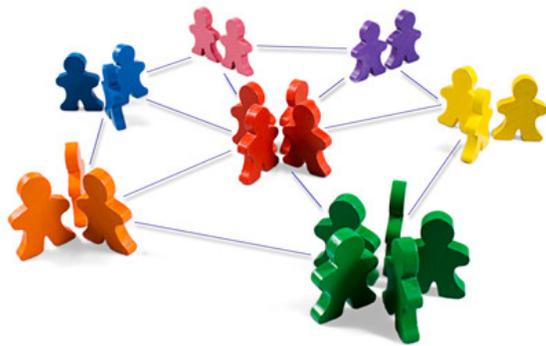
Puntajes PSU:

NEM: <input type="text" value="805.00"/>	Lenguaje y Comunicación: <input type="text" value="842.00"/>	Matemáticas: <input type="text" value="850.00"/>
Historia y Cs. Sociales: <input type="text" value="744.00"/>	Ciencias: <input type="text" value="839.00"/>	Promedio PSU: <input type="text" value="846.00"/>

“Datos”



Desde una “perspectiva académica”,  
puedo querer analizar esa información  
“agrupada” por:



Carrera

Puntajes PSU:

NEM:

Historia y Cs. Sociales:

Lenguaje y Comunicación:

Ciencias:

Matemáticas:

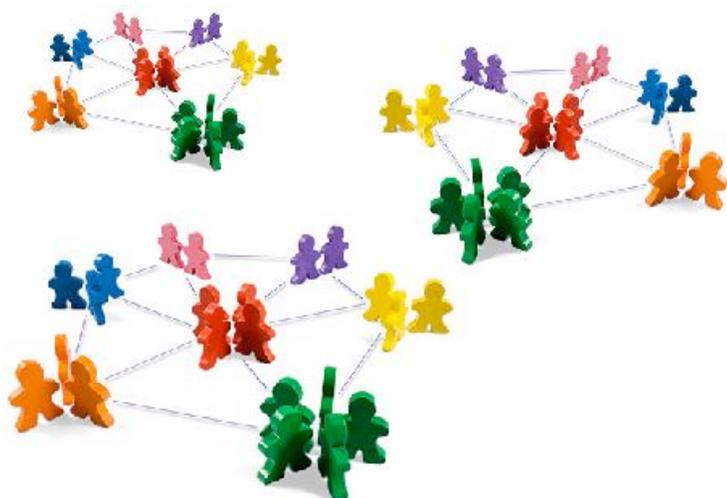
Promedio PSU:

$f(x)$

- Máximo
- Mínimo
- Suma
- Promedio
- Conteo



# O “agrupada” a más alto nivel por:



Unidad Académica

Puntajes PSU:

NEM:

Historia y Cs. Sociales:

Lenguaje y Comunicación:

Ciencias:

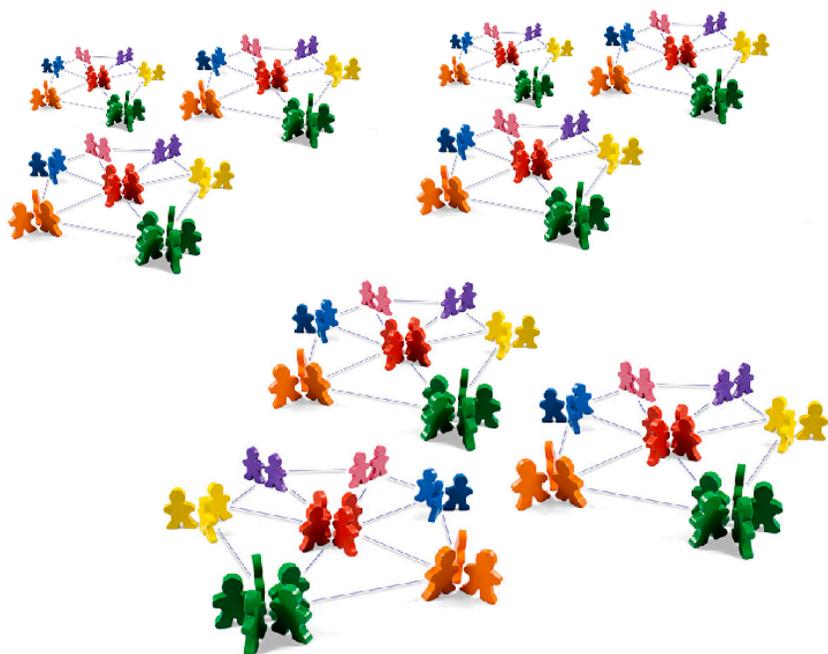
Matemáticas:

Promedio PSU:

**f(x)**



# 0 aún a más alto nivel por:



Facultad

Puntajes PSU:

NEM:

Historia y Cs. Sociales:

Lenguaje y Comunicación:

Ciencias:

Matemáticas:

Promedio PSU:

**f(x)**



# O desde una perspectiva “geográfica”, agrupada por:



Comuna

Puntajes PSU:

NEM:

Historia y Cs. Sociales:

Lenguaje y Comunicación:

Ciencias:

Matemáticas:

Promedio PSU:

**f(x)**



# O desde una perspectiva “geográfica”, agrupada por:



Puntajes PSU:

NEM:

Historia y Cs. Sociales:

Lenguaje y Comunicación:

Ciencias:

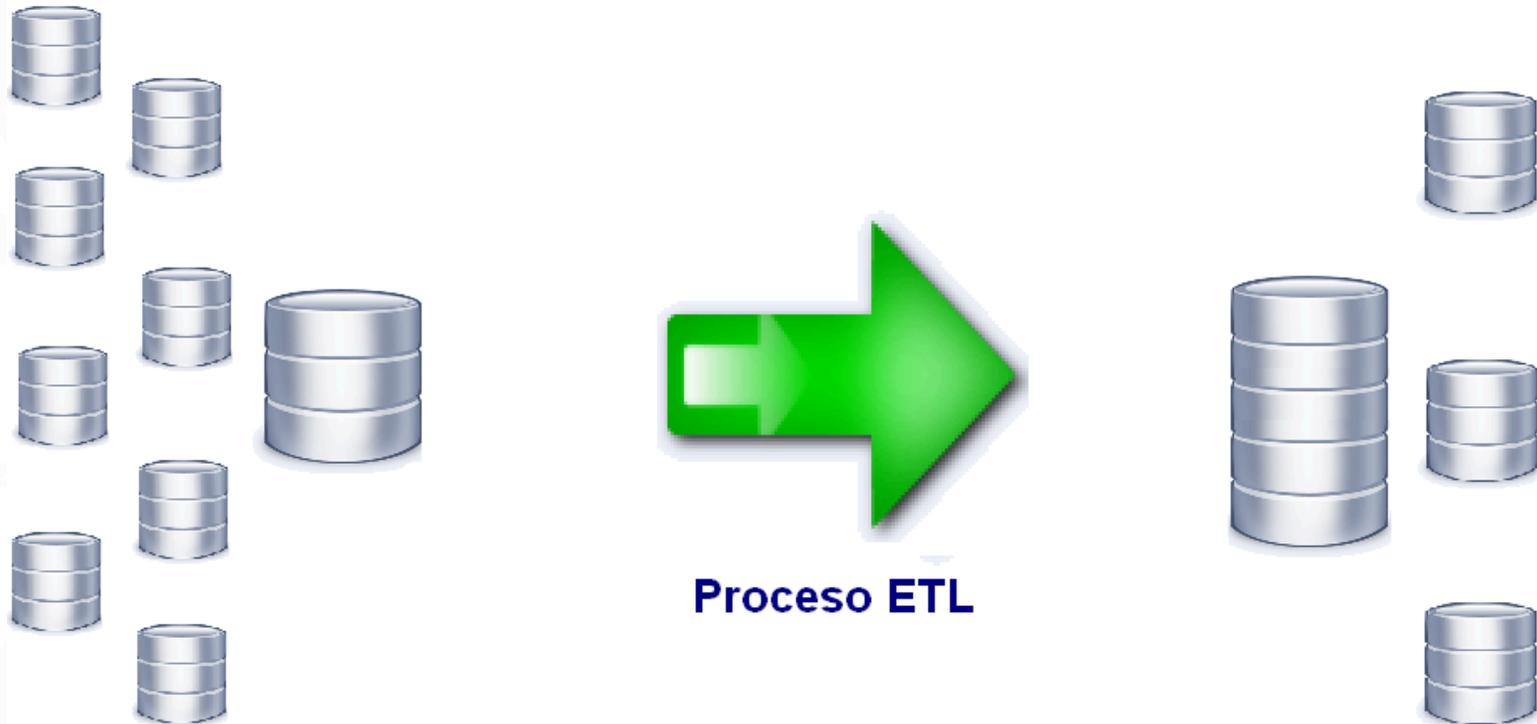
Matemáticas:

Promedio PSU:

**f(x)**



## ¿Cómo poblar el Data Warehouse?



Base de datos operacional  
(OLTP: On-Line Transactional Processing)

Base de datos de análisis  
(OLAP: On-Line Analytical Processing)

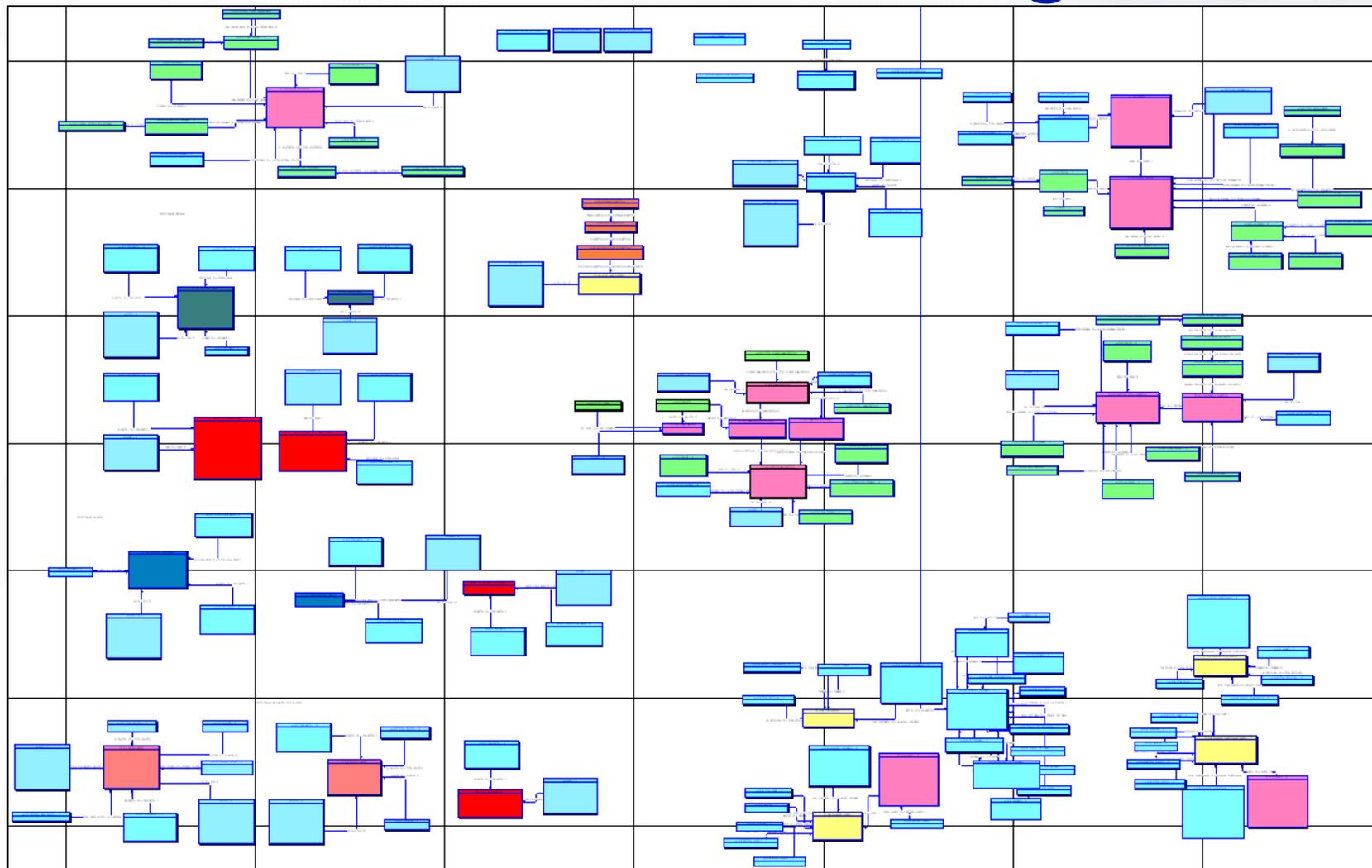


## El proceso ETL lo que hace es:

- Extrae los datos desde las BDO. (E)
- Los transforma bajo criterios comunes. (Ejemplo: sexo, estado civil, agrupaciones, correcciones eventuales de datos). (T)
- Los carga al Data Warehouse. (L)
- Además, estos datos deben ser validados. (V)



# Finalmente obtenemos algo como:





## Y para explotar esa base de datos de análisis necesitamos:

- Conocer o capacitarse en herramientas de consulta a dicha base de datos de manera de hacer consultas ad-hoc a cada pregunta que se requiera.
- Solicitar a informática que realice dichos consultas.



# ¿Todo este trabajo para llegar a algo similar?

Es decir, consultas y reportes estáticos y con una fuerte dependencia de Informática.



# Afortunadamente, la tecnología ha puesto a nuestra disposición excelentes herramientas de análisis de estas bases de datos

Con este tipo de herramientas, los usuarios finales pueden “sacar provecho” del Data Warehouse. Herramientas de BI.



# Y estas herramientas en general lo que entregan es:

- Un modelo de objetos que representan el negocio y que permiten manejar los datos a un nivel abstracto (metadata).
- Capacidad de manejar grandes volúmenes de datos.
- Grandes capacidades para generar salidas como reportes, dashboards.
- Poder cruzar datos y realizar drill's.



- Herramientas para publicar dichas salidas.
- Exportar dichas salidas a variados formatos.
- Distribuir la información a través de diferentes mecanismos como servidores de archivos, sistemas móviles, e-mail.



Ahora, desde el punto de vista de los usuarios, lo realmente relevante es:

El modelo de negocio



# Un ejemplo práctico

## Caso de la UC



# Data Warehouse de Gestión UC



Gestión  
Financiera



Gestión  
Estudiantil



Gestión  
Académica



# Plataforma de BI en la UC





# Los 5 estilos de BI





## ¿Se acuerdan del ejemplo anterior?

Bueno, luego de un proceso de modelamiento de la herramienta...donde de todas maneras hay que solicitar ayuda del área Informática...se obtiene finalmente:

- Atributos
- Jerarquías
- Métricas



## Parte III - Un ejemplo práctico



# Jerarquía Académica



**Facultad**



**Unidad Académica**



**Carrera**



# Un atributo solo (o Jerarquía Pruebas)



**Pruebas de Selección**



# Y la infaltable Jerarquía Tiempo



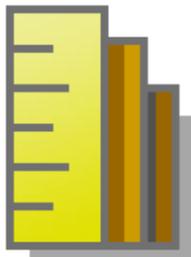
**Año Admisión**



**Período Admisión**



# Unas cuantas métricas



**Mínimo  
Puntaje  
Prueba**



**Máximo  
Puntaje  
Prueba**



**Promedio  
Puntaje  
Prueba**