

# Reporting Informes

Unidad 3. Explotación de almacenes de datos

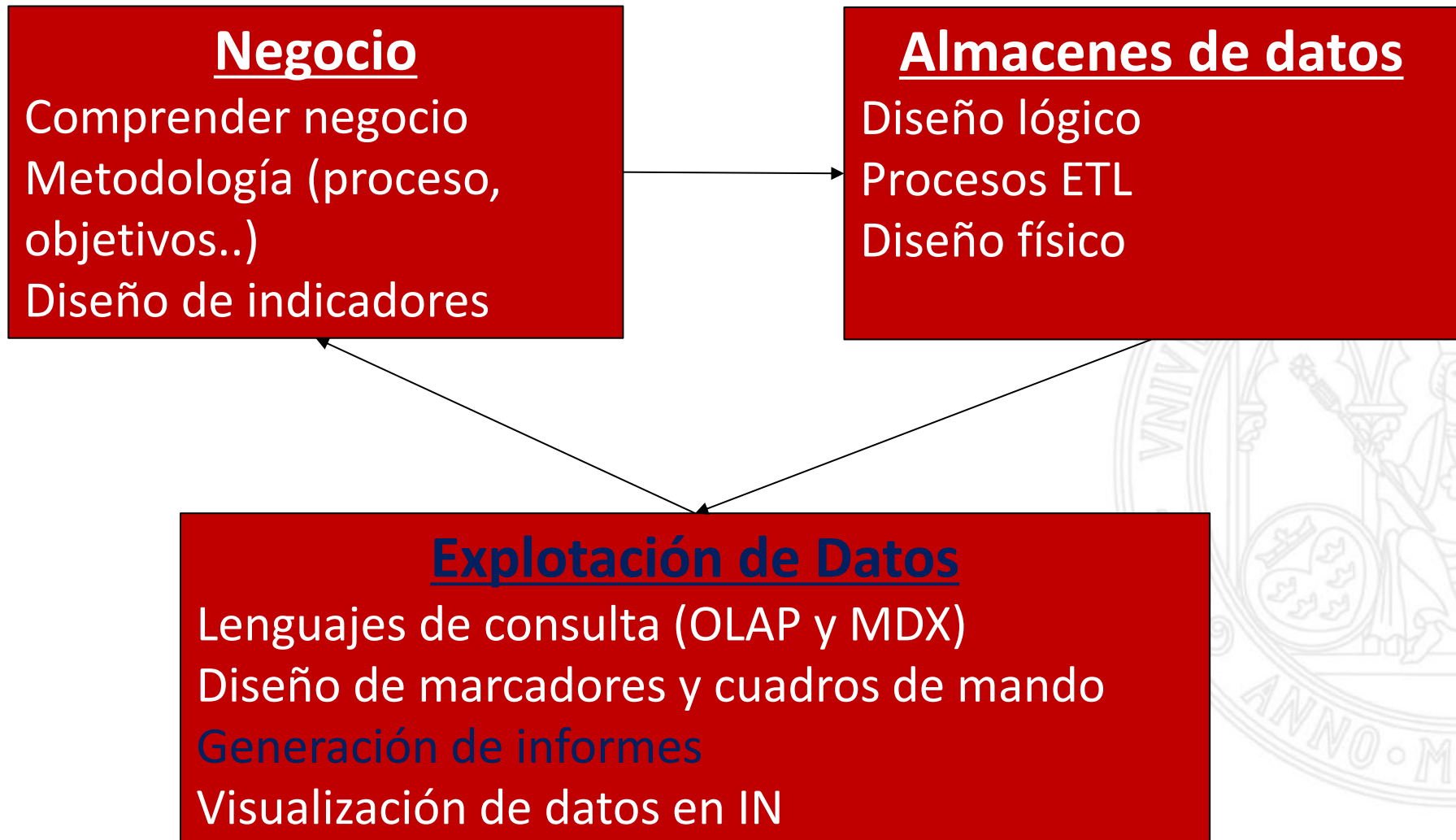
U3.3 - Reporting

UNIVERSIDAD DE  
MURCIA

### Unidad 3.3. – Generación de informes

- Fundamentos de los informes
- Herramientas para generar informes





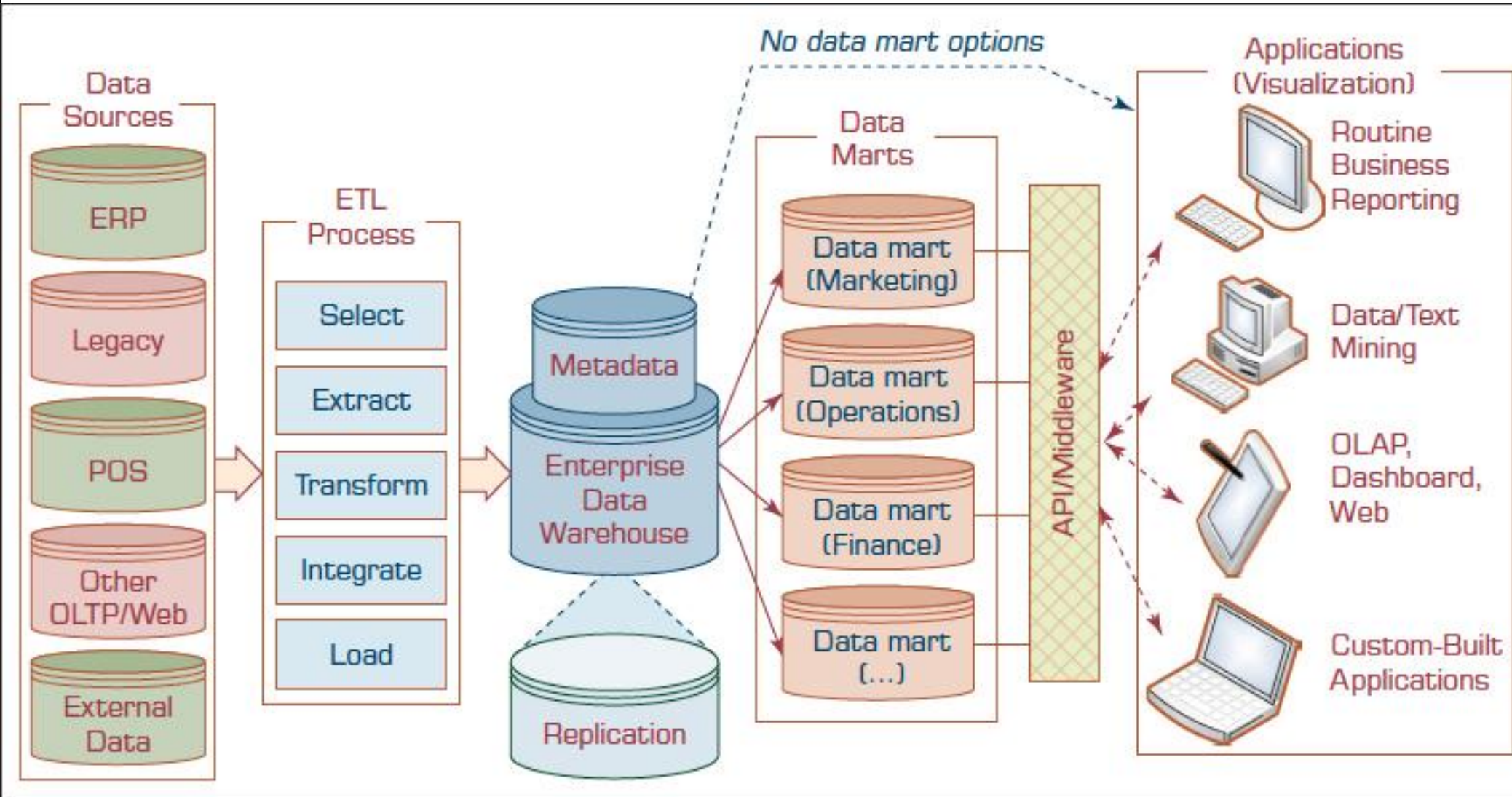
Unidad 3. Explotación de almacenes de datos

# FUNDAMENTO DE LOS INFORMES



# Unidad 3. Explotación de almacenes de datos

## Visión global del proceso de explotación



# Unidad 3. Explotación de almacenes de datos

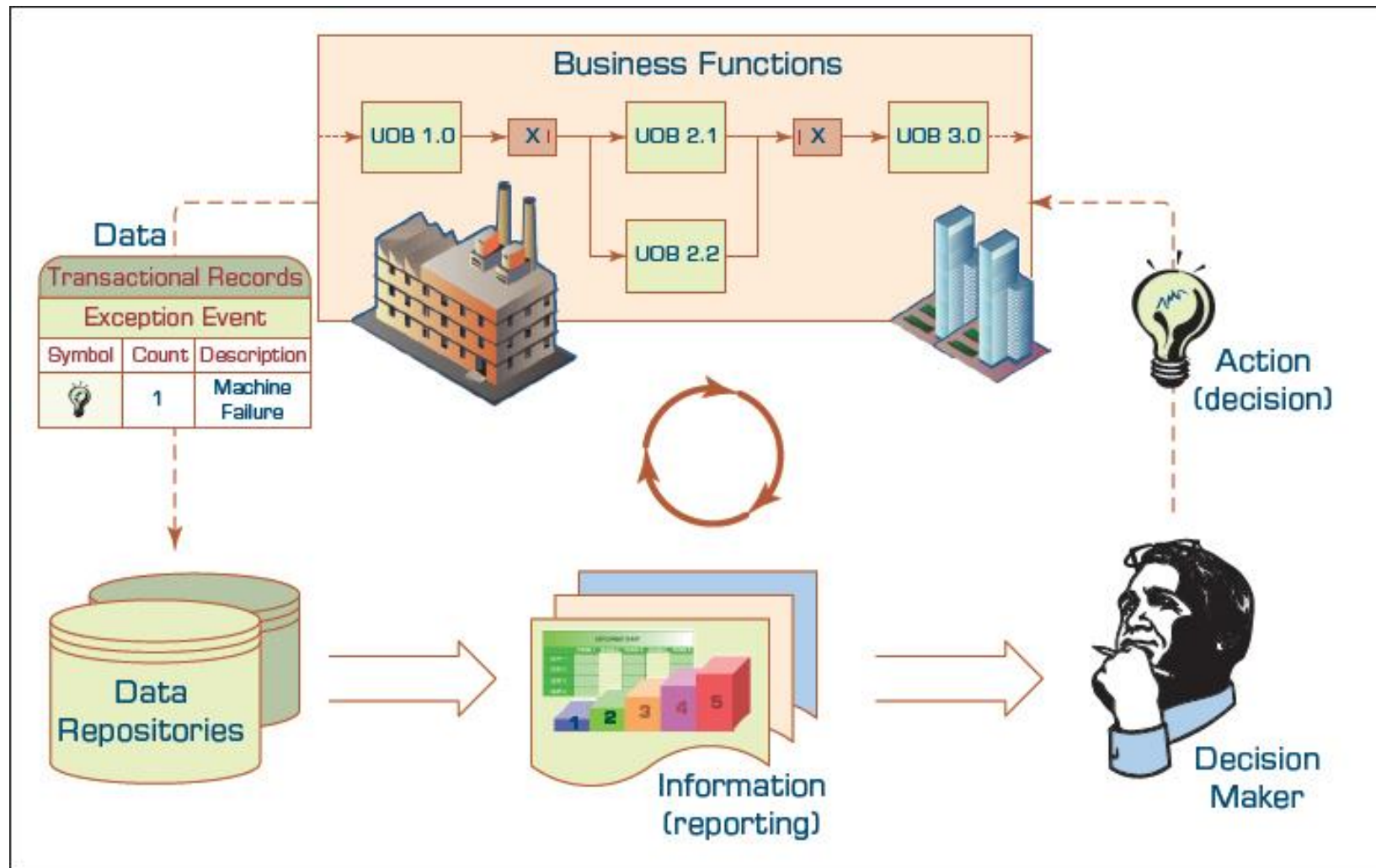
## Visión global de los informes

- Los actores que toman decisiones necesitan tomarlas con exactitud y en el momento adecuado
- La información se suele transmitir en formato escrito (digital o papel), aunque en algunas ocasiones pueda ser oralmente
- *“Un reporte es cualquier artefacto de comunicación preparado con la **intención** específica de transmitir información de una forma **digerible**, para **quien** la necesite, **cuando** y **donde** la necesite”*
  - Periodos de tiempo específicos, eventos, ocurrencias o sujetos (*¿sobre qué?*)
  - Organizado en narrativa, gráficos, y/o formato tabular (*¿cómo?*)
  - Quien preparará y consumirá el reporte (*¿quién?*)
  - Preparado periódicamente (recurrente) o cuando sea necesario (ad-hoc) (*¿cuando?*)
  - Lugar donde el reporte será consumido (*¿dónde?*)



# Unidad 3. Explotación de almacenes de datos

## Diagrama del proceso de informes



# Unidad 3. Explotación de almacenes de datos

## Características clave de los informes

- **Fácil de usar:** Democratización de BI con herramientas amigables (no sólo personal IT) – Funcionalidad de ayuda, diseño GUI, decisiones...
- **Basados en la web:** Acceso del usuario como servicio (sin cliente de software), menos soporte técnico. ¿Seguridad y disponibilidad?
- **Rapidez:** Rapidez de peticiones y manipulación (DW – informe)
- **Interoperabilidad:** Interpreta datos de múltiples fuentes y la salida se puede usar fácilmente
- **Profundización (drill-down):** Posibilidad de quitar capas, profundizar, re-organizar la información, pivotar por una dimensión
- **Versatilidad en comunicaciones:** Intercambio de datos entre varios usuarios y entre departamentos, informes programados estandarizados, alertas basadas en condiciones...



# Unidad 3. Explotación de almacenes de datos

## Tipos de usuarios en informes

- Identificar diferencias y puntos comunes en tipos de usuarios, para especializar servicios e interfaz
- Tres tipos de usuario BI con diferentes habilidades:
  - **Consumidores de información:** Consumen los reportes estándar pre-existentes (o los reciben) para hacer una tarea. No los crean
  - **La clase intermedia de BI:** Pueden hacer consultas e informes mediante una herramienta con GUI (ligeras habilidades técnicas)
  - **Usuario admin:** Incluye los desarrolladores y administraciones, creando reportes y controlando el acceso y permisos de distribución
- El nivel de autoridad no tiene un rol en terminar quién es quién
- La mayoría de los usuarios BI son gestores de nivel medio y por debajo
  - Trabajadores con gran conocimiento de cada departamento

# Unidad 3. Explotación de almacenes de datos

## Tipos de informes

- **Formato:** Ficheros propietarios (Excel, Tableau Workbook), interfaces con tecnologías web, ficheros HTML/PDF...
- **Usuario:** Construir los reportes para otros, auto-generación de informes, o gestión de los informes mediante un portfolio pre-generado
- **Frecuencia:** Periódico o bajo demanda (ad-hoc)
- **Acceso:** Acceso pull-oriented, información crítica push-oriented, alertas automáticas (usuario vs. preferencias de gestores)
- **Canal:** Llevar el reporte al usuario a través de la web, email, suite profesionales (Slack/Microsoft 365/Google Workspace), apps (informes embebidos)

# Unidad 3. Explotación de almacenes de datos

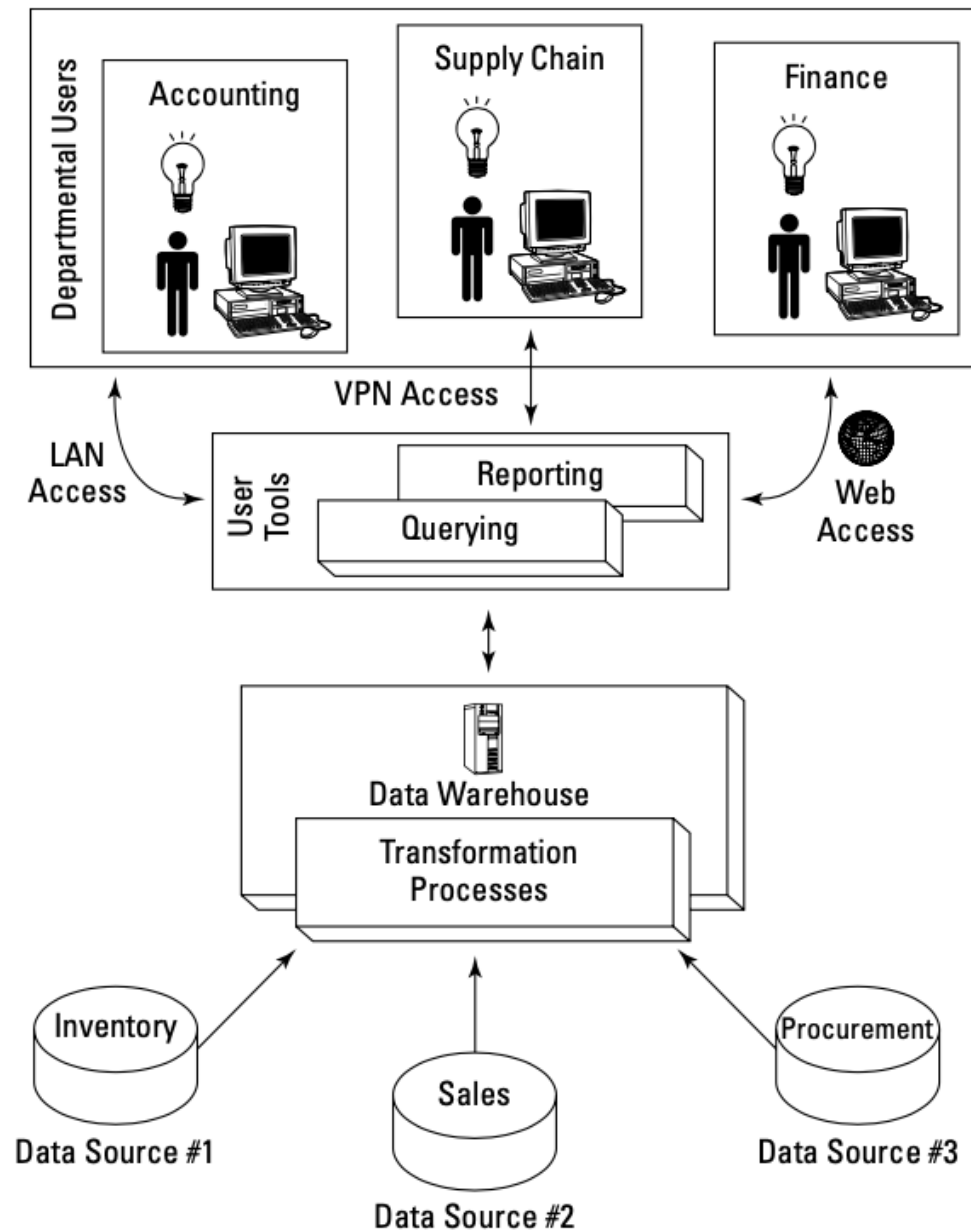
## Características aplicaciones de informes

- Construcción de consultas sin SQL
- Posibilidad de usar drag-and-drop para crear peticiones y reportes
- Características básicas de análisis de datos (por ejemplo, crear un miembro calculado en el reporte)
- Características gráficas que permiten a los usuario construir figuras
- Publicación de información sencilla: email, internet, intranet, donde los compañeros de la empresa puedan accederla
- Distintas posibilidades de exportación
- Calendario y programación automática para organizar informes
- Seguridad para datos sensibles

# Unidad 3. Explotación de almacenes de datos

## El diseño centrado en el usuario en informes

- El diseño centrado en el usuario es clave:
  1. Comprensión clara de los requisitos del usuario y la tarea
  2. Incorporación retroalimentación del usuario
  3. Involucración temprana y activa para evaluar el diseño de un producto de BI
  4. Integración con otras actividades de desarrollo
  5. Iterar en el proceso
- Si tienes que cambiar el diseño más tarde, costará 10 veces más
  - O los usuarios no lo usarán...



# Unidad 3. Explotación de almacenes de datos

## Metodologías de diseño centrado en usuario

- Estándar ISO 13407: Cuatro actividades – toma de requisitos, especificación de requisitos, diseño, y evaluación

Method	Cost	Output	Sample size	When to use
Focus groups	Low	Non-statistical	Low	Requirements gathering
Usability testing	High	Statistical and non-statistical	Low	Design and evaluation
Card sorting	High	Statistical	High	Design
Participatory design	Low	Non-statistical	Low	Design
Questionnaires	Low	Statistical	High	Requirements gathering and evaluation
Interviews	High	Non-statistical	Low	Requirements gathering and evaluation



# Unidad 3. Explotación de almacenes de datos

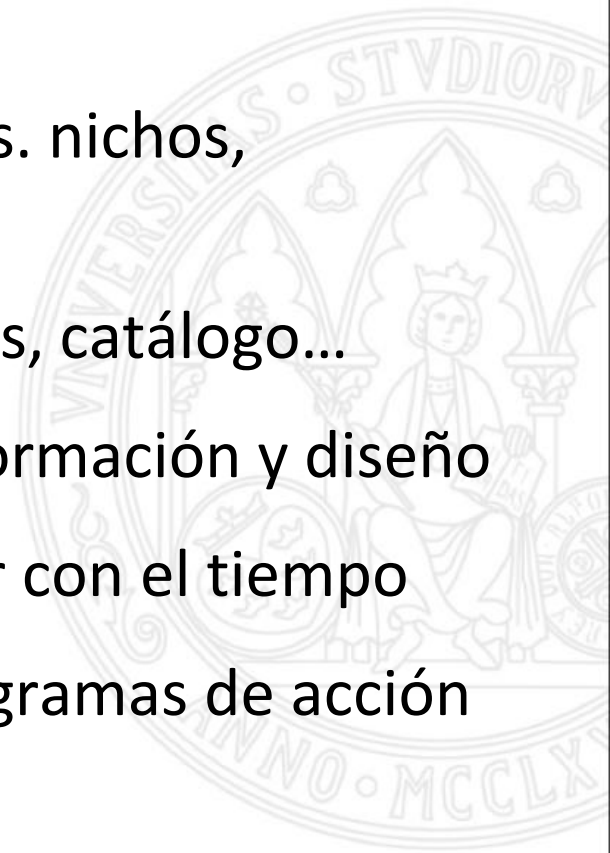
## Comparación con dashboard y scorecard

- ¿Qué conceptos son diferentes de los dashboard y scorecard?
- El cuadro de mando integral (scorecard) es un marco de trabajo para monitorizar las metas de la organización
  - Los cuadros de mando o informes pueden ser parte de las acciones o del programa
- Los cuadros de mando siempre están disponibles, suelen ser basados en herramientas web, y las fuentes de datos se suelen actualizar o están conectadas a la fuente
  - No suelen tener una narrativa escrita
- Un informe suele contener instantáneas, que pueden incluir gráficas/tabla y suelen tener narrativa escrita
  - Pueden ser por web, pero también por ficheros
- Tanto cuadro de mando como reporte se centran en un problema, usuario, dimensión temporal...

# Unidad 3. Explotación de almacenes de datos

## Buenas prácticas de informes

- **Identifica tu audiencia y sus necesidades:** Sigue una metodología centrada en el usuario
- **Preparar la tubería de datos y sus pasos:** Reproducible y automatizado
- **Se flexible y escalable:** Soluciones en la nube vs. nichos, demasiadas herramientas/reportes...
- **Gobierna tus datos:** Maneja el acceso, permisos, catálogo...
- **Comunica tu historia con datos:** Destaca la información y diseño
- **Itera y evoluciona:** Abierto a adaptar y cambiar con el tiempo
- **Cultiva las competencias en datos:** Diseño programas de acción para desarrollar esas habilidades



Unidad 3. Explotación de almacenes de datos

# **HERRAMIENTAS PARA GENERAR INFORMES**



# Unidad 3. Explotación de almacenes de datos

## Visión global de herramientas

- Basadas en Markdown
  - R Markdown
  - Jupyter notebooks
- Ficheros exportados por software propietario
  - Excel (.xlsx)
  - Tableau workbooks (.wtb) y Tableau Reader
- Entornos de programación
  - Librería de informes Jasper
  - Eclipse BIRT
- Herramientas de BI dedicadas
  - Informes Pentaho



# Unidad 3. Explotación de almacenes de datos

## Reportes basados en Markdown

- Nociones básicas de Markdown
  - Lenguaje de marcado para crear texto con formato
  - Cabeceras, listas, imágenes, bloques de citas, reglas horizontales, tablas, pie de página, texto...
  - Markdown: <https://pandoc.org/MANUAL.html#pandocs-markdown>
- Múltiples formatos de salida como HTML, PDF, o Word
  - Para exportar como PDF, necesitas una distribución de LaTeX
    - Windows: <https://miktex.org>
    - Mac: <http://www.tug.org/mactex/>
- También es importante para aproximaciones de ciencia reproducible (datos abiertos, análisis abiertos y artículos interactivos) <http://www.tug.org/mactex/>



# Unidad 3. Explotación de almacenes de datos

## R Markdown (.Rmd)

- Informes con tres partes: Cabecera, escritura en Markdown (narrativa) y bloques de código en R
- Son procesados con un paquete llamado *knitr*
- R Studio hace el proceso muy sencillo y directo
- Tutorial de R Markdown:  
<https://rmarkdown.rstudio.com/lesson-1.html>
- Cheatsheet de R Markdown: <https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf>



# Unidad 3. Explotación de almacenes de datos

## Ejemplo de R Markdown en R Studio

UNIVERSIDAD DE  
MURCIA

RStudio

InvestigacionMetodos.Rmd

```
1 ---
2 title: "Investigacion y Metodos"
3 author: "Jose A. Ruiperez Valiente"
4 date: "4/10/2019"
5 output: html_document
6 ---
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10 ```
11
12 ## R Markdown
13
14 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
15
16 When you click the Knit button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:
17
18 ```{r cars}
19 summary(cars)
20 ```
21
22 ## Including Plots
23
24 You can also embed plots, for example:
25
26 ```{r pressure, echo=FALSE}
```

InvestigacionMetodos.html

Investigacion y Metodos

Jose A. Ruiperez Valiente

4/10/2019

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

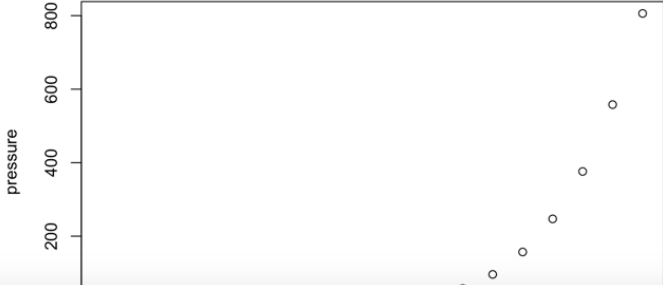
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

##	speed	dist
## Min.	: 4.0	Min. : 2.00
## 1st Qu.:	12.0	1st Qu.: 26.00
## Median :	15.0	Median : 36.00
## Mean :	15.4	Mean : 42.98
## 3rd Qu.:	19.0	3rd Qu.: 56.00
## Max. :	25.0	Max. : 120.00

Including Plots

You can also embed plots, for example:



Cabecera

Bloques código

Bloques narrativos

R Markdown compila a -->

Console

```
~/InvestigacionMetodos.Rmd
> read_csv()
Error in read_csv() : could not find function "read_csv"
> library(readr)
> ?read_csv()
Not all of the characters in ~/InvestigacionMetodos.Rmd could be encoded using ASCII. To save using a different encoding, choose "File | Save with Encoding..." from the main menu.
>
```



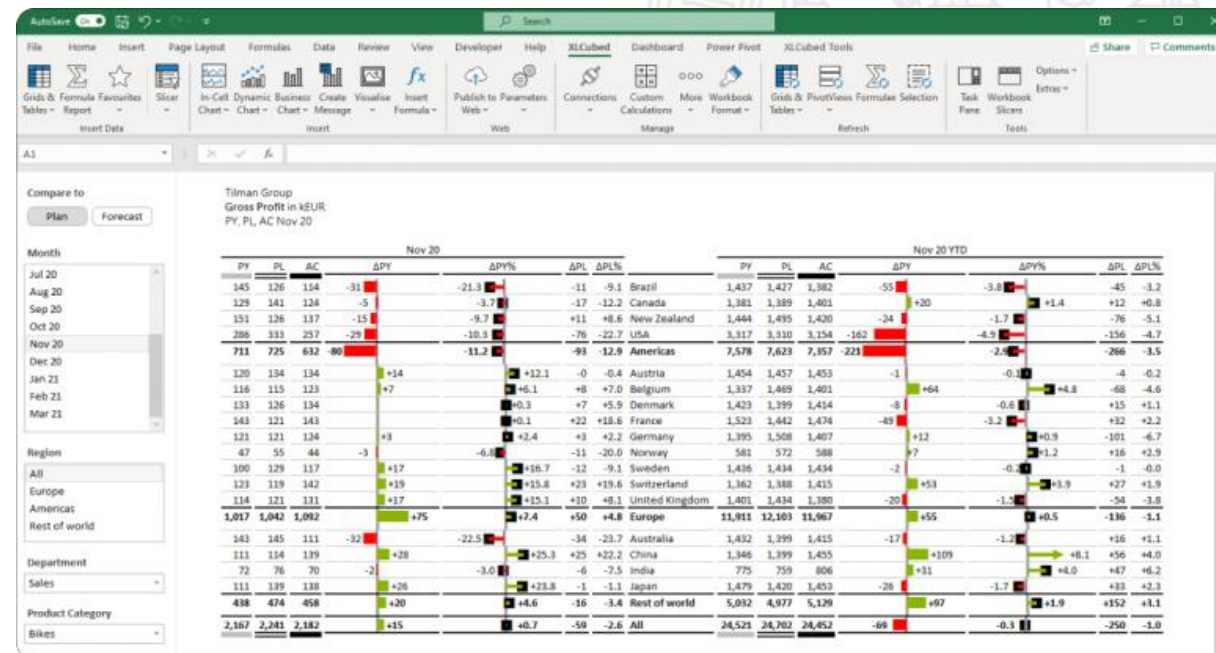
# Unidad 3. Explotación de almacenes de datos

## Jupyter notebook

- Jupyter notebooks (.Nb) ([ejemplo](#))
- Crear y compartir documentos computacionales  
También combina código con Markdown
- Da soporte a más de 40 lenguajes de programación
- Los notebooks pueden ser compartidos a través de varias opciones como email, Dropbox, GitHub y [nbviewer](#)
  - También a través de webs, como JupyterLab y Google Colab
- Salida interactiva rica: HTML, LaTeX... a través de [nbconvert](#)
  - *jupyter nbconvert --to html*
- Integración de herramientas de big data

UNIVERSIDAD DE  
MURCIA

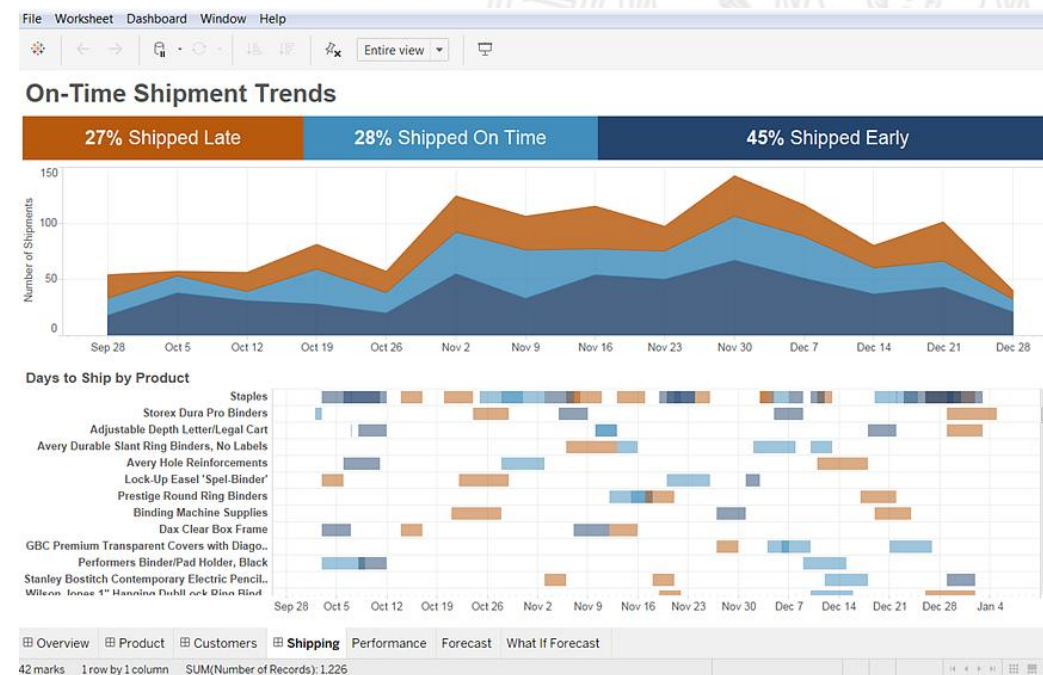
- staña
- 



# Unidad 3. Explotación de almacenes de datos

## Tableau Workbooks (.twbx)

- Tableau: Frecuentemente usado como servicio en la nube o con servidor
  - En este escenario, el pago se usa por licencias de usuario
- En este escenario, se usa un formato propietario de fichero exportando el informe
- El formato se abre gratuitamente con Tableau Reader
- Se crea con Tableau Desktop





# Unidad 3. Explotación de almacenes de datos

## Propiedades de la Jasper Reports Library

UNIVERSIDAD DE  
MURCIA

- Motor de informes muy popular
- Software abierto
- Variedad de formatos de salida: HTML, PDF, Office, Excel, Word
- Aproximación clásica: Edición de XML
- Aproximación moderna: Jasper Studio (IDE Eclipse)



# Unidad 3. Explotación de almacenes de datos

## Jasper Reports Library – Formato XML

- Plantilla de reporte: BasicReport.xml

```
<jasperReport name="BasicReport" >
  <parameter name="Title" class="java.lang.String"/>
  <queryString><![CDATA[select name, cost from product]]>
</queryString>
  <field name="NAME" class="java.lang.String"/>
  <field name="COST" class="java.lang.Double"/>
  <title>
    <band height="50">
      <textField>
        <reportElement x="0" y="0" width="200" height="50" />
        <textFieldExpression class="java.lang.String">$P{Title}</textFieldExpression>
      </textField>
    </band>
  </title>
  <pageHeader>
    <band>
    </band>
  </pageHeader>
  <columnHeader>
    <band height="20">
      <staticText>
        <reportElement x="180" y="0" width="180" height="20" />
        <textElement>
          <font isUnderline="true"/>
        </textElement>
        <text><![CDATA[NAME]]></text>
      </staticText>
```

- title
- pageHeader
- columnHeader
- Detail
- columnFooter
- pageFooter
- summary



# Unidad 3. Explotación de almacenes de datos

## Jasper Reports Library – Formato XML

### Report template: BasicReport.xml

```
<jasperReport name="BasicReport" >
  <parameter name="Title" class="java.lang.String"/>
  <queryString><![CDATA[select name, cost from product]]>
</queryString>
  <field name="NAME" class="java.lang.String"/>
  <field name="COST" class="java.lang.Double"/>
  <title>
    <band height="50">
      <textField>
        <reportElement x="40" y="0" width="100" height="15"/>
        <textElement/>
        <textFieldExpression class="java.lang.Integer"><![
</textFieldExpression>
      </textField>
    </band>
  </title>
  <pageHeader>
    <band>
    </band>
  </pageHeader>
  <columnHeader>
    <band height="20">
      <staticText>
        <reportElement x="180" y="0" width="180" height="20"/>
        <textElement>
          <font isUnderline="true"/>
        </textElement>
        <text><![CDATA[NAME]]></text>
      </staticText>
```

- Parameters \$P{name}
- Fields \$F{name}
- Variable \$V{name}

textField containing a page  
number using a variable  
*PAGE\_NUMBER* defined internally  
By JasperReport

# Unidad 3. Explotación de almacenes de datos

## Jasper Reports Library – Código Java

### Java code

```
// First, load JasperDesign from XML and compile it into
JasperReport
JasperDesign jasperDesign =
JasperManager.loadXmlDesign("BasicReport.xml");
JasperReport jasperReport =
JasperManager.compileReport(jasperDesign);
// Second, create a map of parameters to pass to the report.
Map parameters = new HashMap();
parameters.put("ReportTitle", "Basic JasperReport");
parameters.put("MaxSalary", new Double(25000.00));
// Third, get a database connection
Connection conn = Database.getConnection();
// Fourth, create JasperPrint using fillReport() method
JasperPrint jasperPrint = JasperManager.fillReport(jasperReport,
parameters, conn);
// You can use JasperPrint to create PDF
JasperManager.printReportToPdfFile(jasperPrint,
"BasicReport.pdf");
// Or to view report in the JasperViewer
JasperViewer.viewReport(jasperPrint);
```

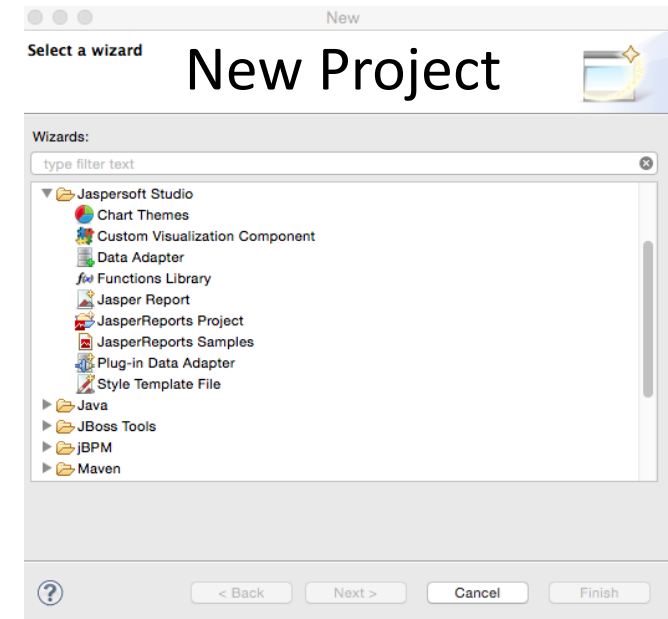
```
<jasperReport name="BasicReport" >
  <parameter name="Title" class="java.lang.String"/>
  <queryString>{CDATA[select name, cost from product]}>
</queryString>
  <field name="NAME" class="java.lang.String"/>
  <field name="COST" class="java.lang.Double"/>
  <title>
    <band height="50">
      <textField>
        <reportElement x="0" y="0" width="200" height="50" />
        <textFieldExpression class="java.lang.String">${P{Title}}</textFieldExpression>
      </textField>
    </band>
  </title>
  <pageHeader>
    <band>
      </band>
    </pageHeader>
    <columnHeader>
      <band height="20">
        <staticText>
          <reportElement x="180" y="0" width="180" height="20" />
          <textElement>
            <font isUnderline="true"/>
            </textElement>
            <text>{CDATA[NAME]}</text>
          </staticText>
        </band>
      </columnHeader>
    </pageHeader>
  </pageHeader>
</jasperReport>
```

# Unidad 3. Explotación de almacenes de datos

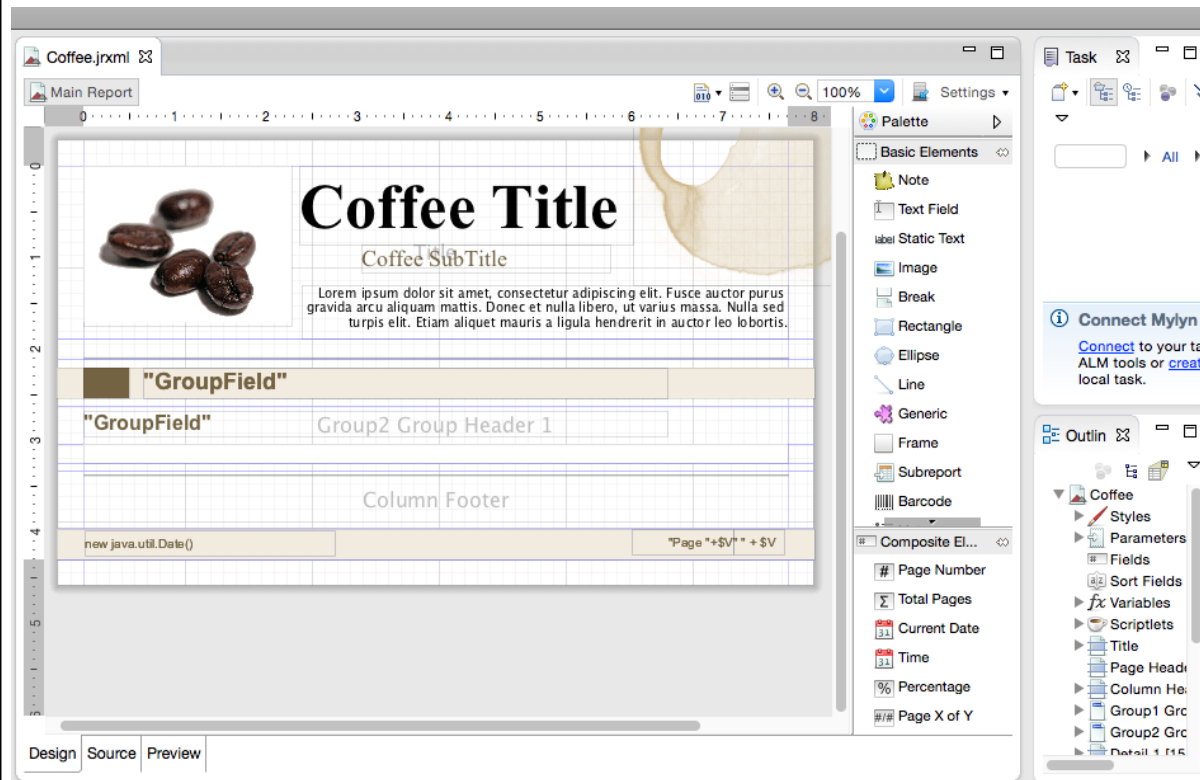
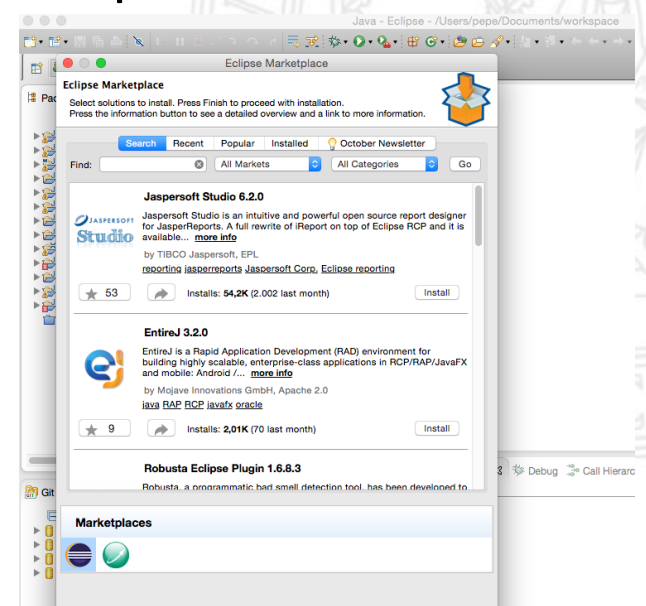
## Jasper Reports Library – Eclipse Studio

UNIVERSIDAD DE  
MURCIA

- Comunidad: Motor de reportes embebido, generación XML, pdf...
- Comercial: gráficas rápidas y widgets



## Eclipse Market Place





# Unidad 3. Explotación de almacenes de datos Eclipse BIRT

UNIVERSIDAD DE  
MURCIA

- Proyecto de Eclipse
- Software abierto
- Similar a Jasper

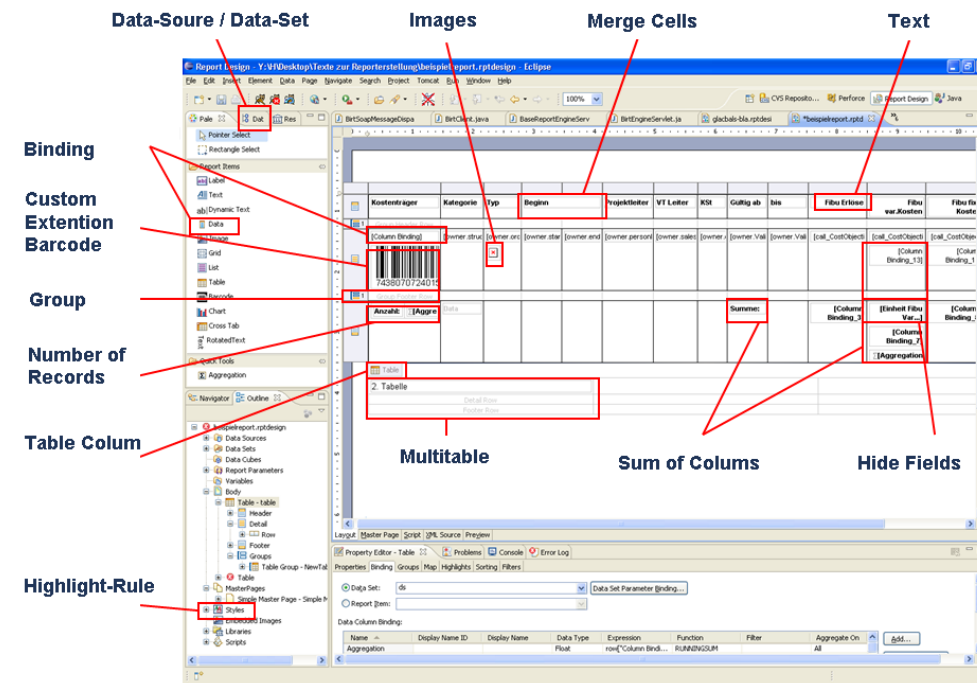
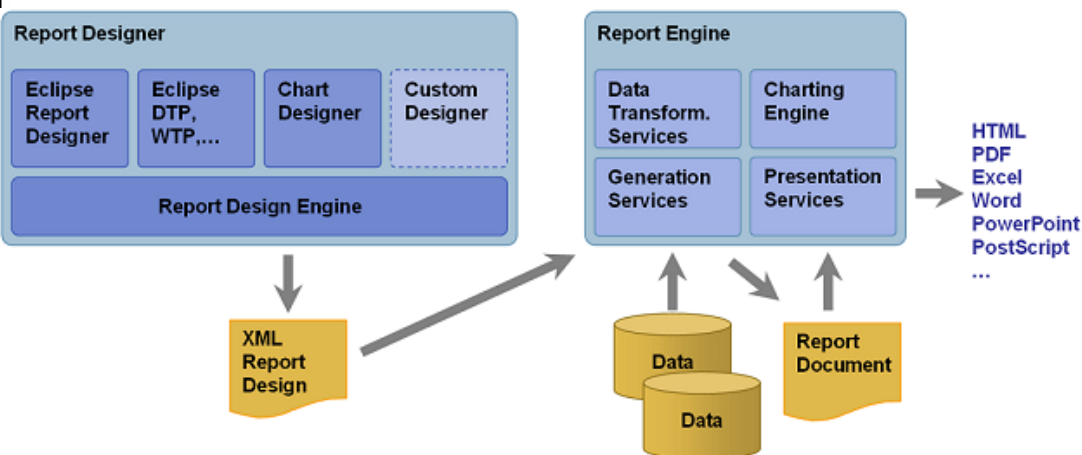


Figure from:  
<http://www.eclipse.org/birt/about/architecture.php>

<https://eclipse.github.io/birt-website/docs/overview-designer/>

# Unidad 3. Explotación de almacenes de datos

## Pentaho Report (jFreeReport)

- Software abierto (Pentaho, adquirida por Hitachi)
  - Librería Java (jFreeReport)
  - Aplicación independiente de escritorio
  - Integración sencilla con el framework de Pentaho
- SDK de informes: Motor clásico + documentación + librerías (Java)
- Motor de informes: Librería java para generar informes
  - Arquitectura cliente/servidor
- Diseñador de informes: herramienta independiente
  - Herramienta visual para informes complejos

Complete comparative at:

<http://www.innoventsolutions.com/comparison-matrix.html>

# Unidad 3. Explotación de almacenes de datos

## Comparación de herramientas de informes

	JASPER	BIRT	PENTAHO
Designer	Eclipse plug-in	Eclipse project	Stand alone
Multicolumn	Y	N	N
Mutiple source data	N	Y	Y
JDBC drivers	24	2	>40
Hadoop support	Y	Y	Y
Pentaho integration	N	N	Y
Output	HTML, PDF EXCEL, PPT,OPENOFFICE, PS,...	HTML, PDF EXCEL, PPT,OPENOFFICE, PS, FLASH,...	HTML, PDF, EXCEL,...

Complete comparative at:  
<http://www.innoventsolutions.com/comparison-matrix.html>