

## Lesson 13:

# BI Tools

- Các công cụ BI
- Thực hành xây dựng báo cáo và biểu đồ

Giảng Viên: Trần Đăng Hòa

Trợ Giảng: Nguyễn Chí Thanh





**Trần Đăng Hòa**

Big Data Architecture & Data Governance

## Chuyên gia. **Trần Đăng Hòa**

---

- Xây dựng, thiết kế Data warehouse
- Thiết kế kiến trúc Data Lake
- Phân tích và khoa học dữ liệu (DA - DS)
- Kiến trúc hạ tầng và mạng lưới hệ thống Big Data
- Quản trị dữ liệu
- Admin 'Cộng đồng Big Data Việt Nam'

# Mục lục

## 1. BI tools Overview

## 2. BI tools

## 3. Chart types

## 4. Hướng dẫn Birt

## 5. Thực hành



# 1. BI tools Overview



# Front room

- Front room là lớp public của warehouse nơi business user nhìn thấy và làm việc hàng ngày.
- Các kiểu ứng dụng BI quan trọng nhất bao gồm:
  - ✓ **Truy vấn truy cập trực tiếp**
  - ✓ **Báo cáo tiêu chuẩn**
  - ✓ **Các ứng dụng phân tích**
  - ✓ **Dashboard và Scorecard**
  - ✓ **Data mining và model**
  - ✓ **Operational BI**
- Những ứng dụng BI này được phân phối tới business user thông qua các interface:
  - ✓ **BI portal và custom front end**: cho phép người dùng dễ dàng truy cập ứng dụng BI thông qua giao diện web để thực hiện các truy vấn phức tạp và
  - ✓ **Giao diện thiết bị cầm tay**
  - ✓ **Instantaneous BI (EII)**: một dạng real time data warehouse architecture với kết nối trực tiếp từ hệ thống source transaction tới user's screen.

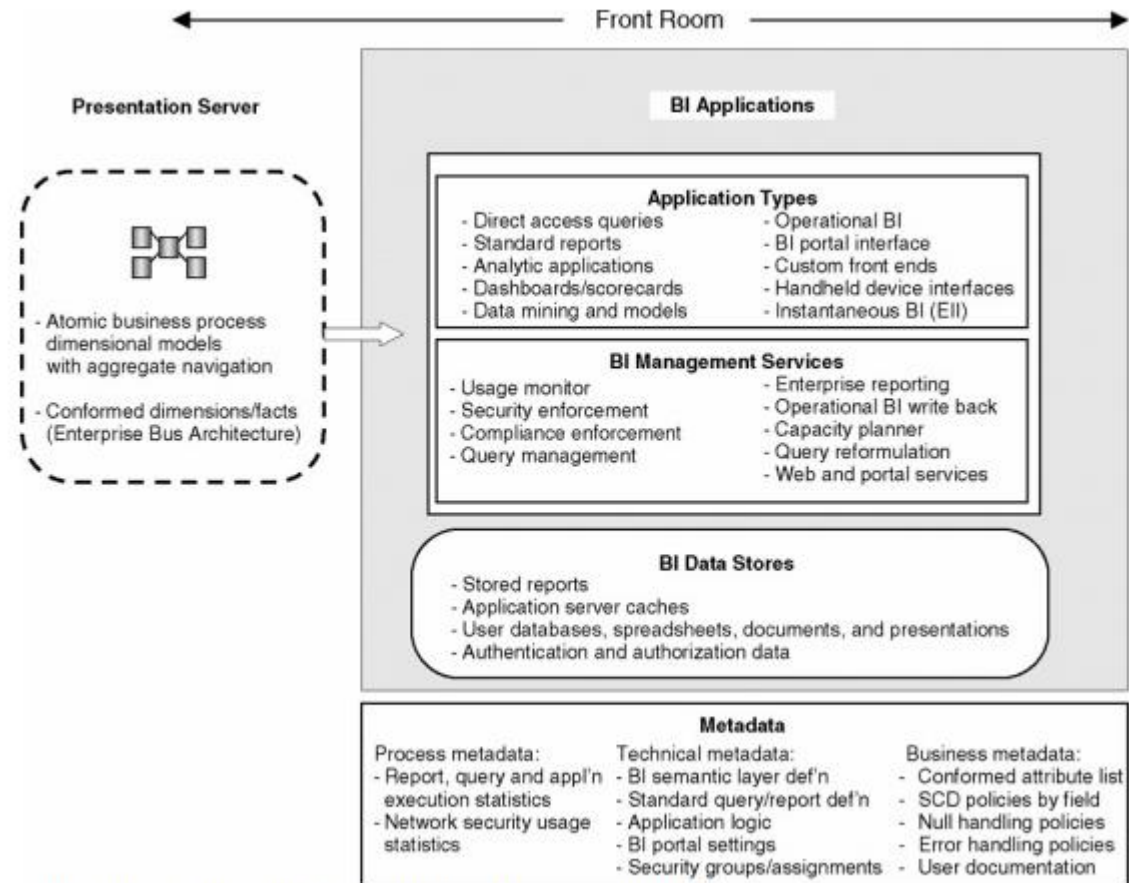


Figure 4-4: Front room technical architecture model

# BI Management Service

Có 3 dạng chính đã mang data access service ra khỏi desktop và vào application tier.

- Đầu tiên là mua sức mạnh của DW/BI sẵn có.
- Thứ 2 là sử dụng các công cụ client nhỏ giúp giảm tải cho desktop client và đưa phần lớn chức năng chia sẻ được tới một application server.
- Thứ 3, tốt nhất cho mọi hệ thống DW/BI hiện tại, đó là chuyển qua service oriented architecture (SOA).



# Enterprise Reporting Services

Các hệ thống báo cáo doanh nghiệp lớn có xu hướng xuất hiện khi hệ thống ERP không thể xử lý khối lượng transaction và reporting. Nên thiết lập một enterprise reporting project hoặc một nhóm chỉ tập trung thực hiện phần này.

Các yêu cầu đối với công cụ báo cáo doanh nghiệp bao gồm:

- **Môi trường phát triển báo cáo.** Bao gồm hầu hết các adhoc tool.
- **Định nghĩa báo cáo linh hoạt.** Gồm các document layout như dashboard với các đồ thị và bảng trên cùng một page. Các tool cần hỗ trợ kiểm soát định dạng cho cả hiển thị và in ấn.
- **Máy chủ thực thi báo cáo.**
- **Khả năng điều chỉnh tham số.**
- **Lập lịch thực hiện báo cáo dựa trên time-based hoặc event-based.**
- **Thực hiện lập hoặc tách kết quả.**
- **Phân phối báo cáo linh động.** Báo cáo có thể cung cấp qua nhiều phương thức như email, web, bản in, fax với nhiều định dạng html, PDF, database table, hoặc spreadsheet.
- **Cung cấp khả năng truy cập dạng publish-subscribe.**
- **Report Linking.** Là phương pháp đơn giản để hiển thị chi tiết, nếu đã có báo cáo chạy trước cho tất cả các phòng ban, bạn có thể bấm vào tên phòng ban đó trong bảng tóm tắt và báo cáo chi tiết hiển thị lên.
- **Khả năng duyệt các thư viện báo cáo.** Đây là một loại metadata tham chiếu đến mô tả của mỗi report gồm nội dung, lần chạy cuối cùng.
- **Phân phối hàng loạt.**
- **Công cụ quản lý môi trường báo cáo.** Nên gồm schedule, monitor, troubleshoot các report problem, khả năng giám sát việc sử dụng và loại bỏ các báo cáo không được sử dụng.
- **User access security.** Dịch vụ báo cáo cần xác thực người dùng và xác định họ có quyền xem báo cáo hay không

# BI Data Stores

Về lý thuyết không nên có bất kỳ data store nào trong front room. Tuy nhiên thực tế tập kết quả trả về có thể được đưa vào 1 local analytic store như OLAP engine, Access, thậm chí Excel hoặc bên trong BI application, hệ thống downstream:

- Stored Reports
- Local User Databases
- Data Store phân tích dùng một lần
- Kết quả từ ứng dụng phân tích
- Hệ thống Downstream
- Data Store security



## 2. BI tools



# Tableau

## Tổng quan cách sử dụng

- Tableau là công cụ trực quan dữ liệu có giao diện thân thiện với người dùng, dễ dàng sử dụng các tính năng cơ bản. Khi thiết kế Dashboard, bạn cần tạo ra một worksheet và kéo nó vào container.
- Dù vậy, đối với người dùng không chuyên có thể tận dụng khả năng phân tích cú pháp mà không cần qua đào tạo. Một ưu điểm nữa của Tableau là phần mềm này có cộng đồng khá đông đảo, người dùng có thể dễ dàng trang bị các kiến thức, kỹ năng và kinh nghiệm từ những người dùng khác

# Tableau

Tính năng BI

Ưu điểm:

- Hỗ trợ nhiều nguồn dữ liệu: Relational database, NoSQL database, Multi-dimensional database, Big Data Platform, File data sources (Excel, csv, txt, Json, pdf, mdb, Tableau)
- Khả năng trực quan dữ liệu mạnh mẽ.
- Sử dụng công nghệ trực quan độc nhất, Tableau có thể nhanh chóng phân tích dữ liệu bằng cách thể hiện kết quả phân tích bằng màu sắc, hình dạng, kích thước. Đội ngũ phát triển của Tableau vẫn đang tiếp tục nghiên cứu các loại biểu đồ mà người dùng ưa chuộng. Vì thế, Tableau có thể hiển thị biểu đồ tốt nhất cho dữ liệu.

Nhược điểm:

- Tính năng BI còn hạn chế: Tableau vẫn thiếu một chút tính năng cho một công cụ phân tích kinh doanh hoàn hảo. Ví dụ như báo cáo quy mô lớn, xây dựng các bảng dữ liệu và giao diện tĩnh (static layouts). Ngoài ra, các định dạng xuất và in cũng bị hạn chế, khá bất tiện khi cần chia sẻ.
- Các vấn đề về tùy chỉnh: Tableau khá khó để có thể tùy chỉnh hoặc nhúng vào bản vẽ kiến trúc của công ty

# Tableau

## Giá thành

Giá của phần mềm sẽ tương ứng với số lượng người dùng. Mức giá này sẽ phù hợp với các doanh nghiệp lớn thay vì các doanh nghiệp vừa và nhỏ.

Every deployment requires at least one Creator		
<b>Tableau Creator</b>	<b>Tableau Explorer</b>	<b>Tableau Viewer</b>
\$70	\$35	\$12
USD/user/month billed annually	USD/user/month billed annually	USD/user/month billed annually   min. 100 Viewers required
Discover insights with a powerful suite of products that support your end-to-end analytics workflow.	Explore trusted data and answer your own questions faster with full self-service analytics.	View and interact with dashboards and visualizations in a secure, easy-to-use platform.
Includes D TABLEAU DESKTOP P TABLEAU PREP And one Creator license of S TABLEAU SERVER	Includes one Explorer license of S TABLEAU SERVER	Includes one Viewer license of S TABLEAU SERVER

# Tableau

<https://www.youtube.com/watch?v=IR5Aiy5umvw>

**LINPACK**  
FOR TABLEAU

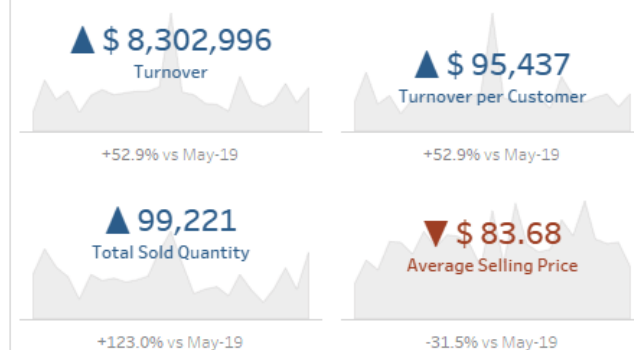
Sales Dashboard  
Cockpit

Current Month: Jun | Current Year: 2019 | Country: (All) | Customer: (All)

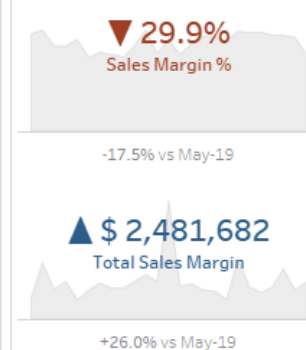
Performance Scope: Current Month vs Previous Month | Product Group: (All) | Filter 4: (All)

Jun-19 vs May-19

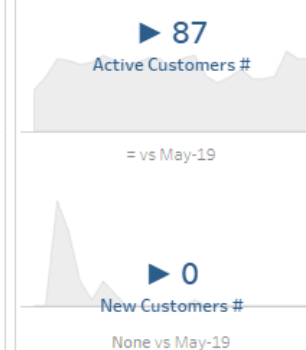
## SALES REVENUES



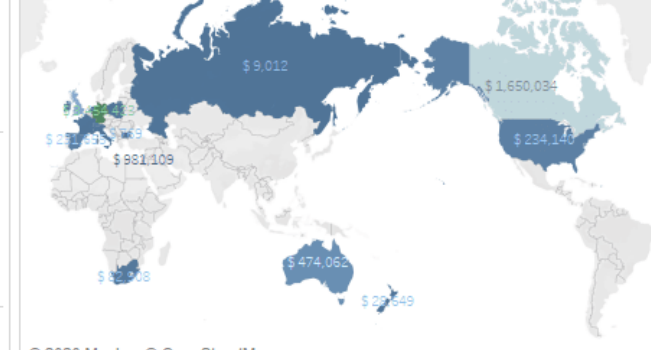
## PROFITABILITY



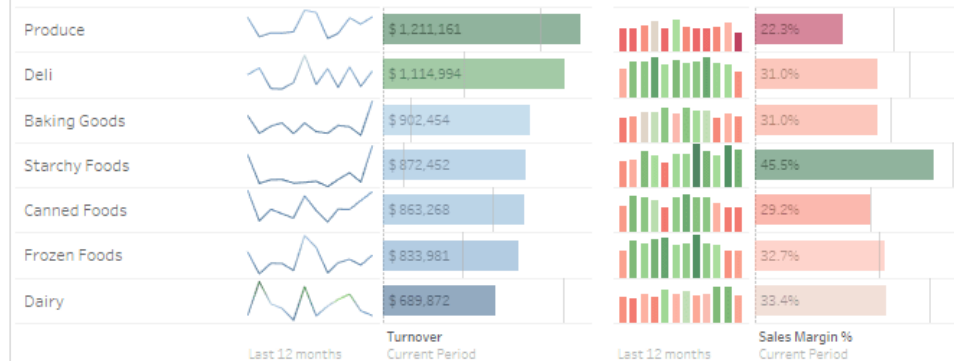
## CUSTOMERS



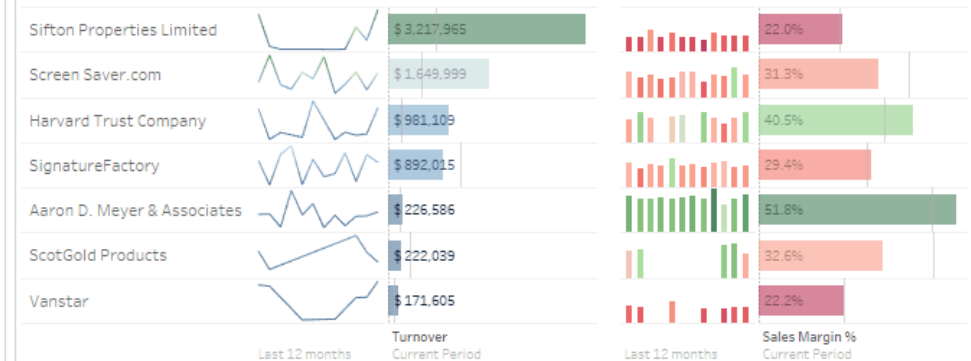
## TURNOVER PER COUNTRY



## PRODUCT GROUP PERFORMANCE



## TOP CUSTOMER



# Power BI

## Tổng quan cách sử dụng

Đến từ Microsoft, nâng các công cụ trực quan dữ liệu của Excel Pivot Table và Excel lên một tầm cao mới. Được cung cấp video và tài liệu học tập khá chi tiết để giúp người dùng nhanh chóng làm quen.

## Tính năng BI

### Ưu điểm:

- Tích hợp Excel: Power BI có thể kết nối với bất kì phần mềm Office nào của Microsoft. Bạn có thể đặt dữ liệu thô vào Excel - nơi bạn có thể xem lại đằng sau các hình ảnh đã trực quan.
- Trực quan hóa tập lệnh R: Là công cụ duy nhất trong TOP 5 hỗ trợ R, sử dụng khả năng phân tích và trực quan phong phú của R để trình bày và phân tích dữ liệu nâng cao như dự báo.

### Nhược điểm:

- Vấn đề về quản lý: Power BI không hỗ trợ phân quyền đa cấp, không thể kiểm soát mức độ chi tiết của cấp độ cột được người dùng xem và không kiểm soát hoàn chỉnh quyền nhóm.
- Khả năng xử lý dữ liệu chưa tốt cho phiên bản miễn phí: Bản miễn phí bị giới hạn lượng dữ liệu có thể xử lý. Khi đã đạt dung lượng 2GB, bạn phải nâng cấp lên bản trả phí để giảm thời gian xử lý.



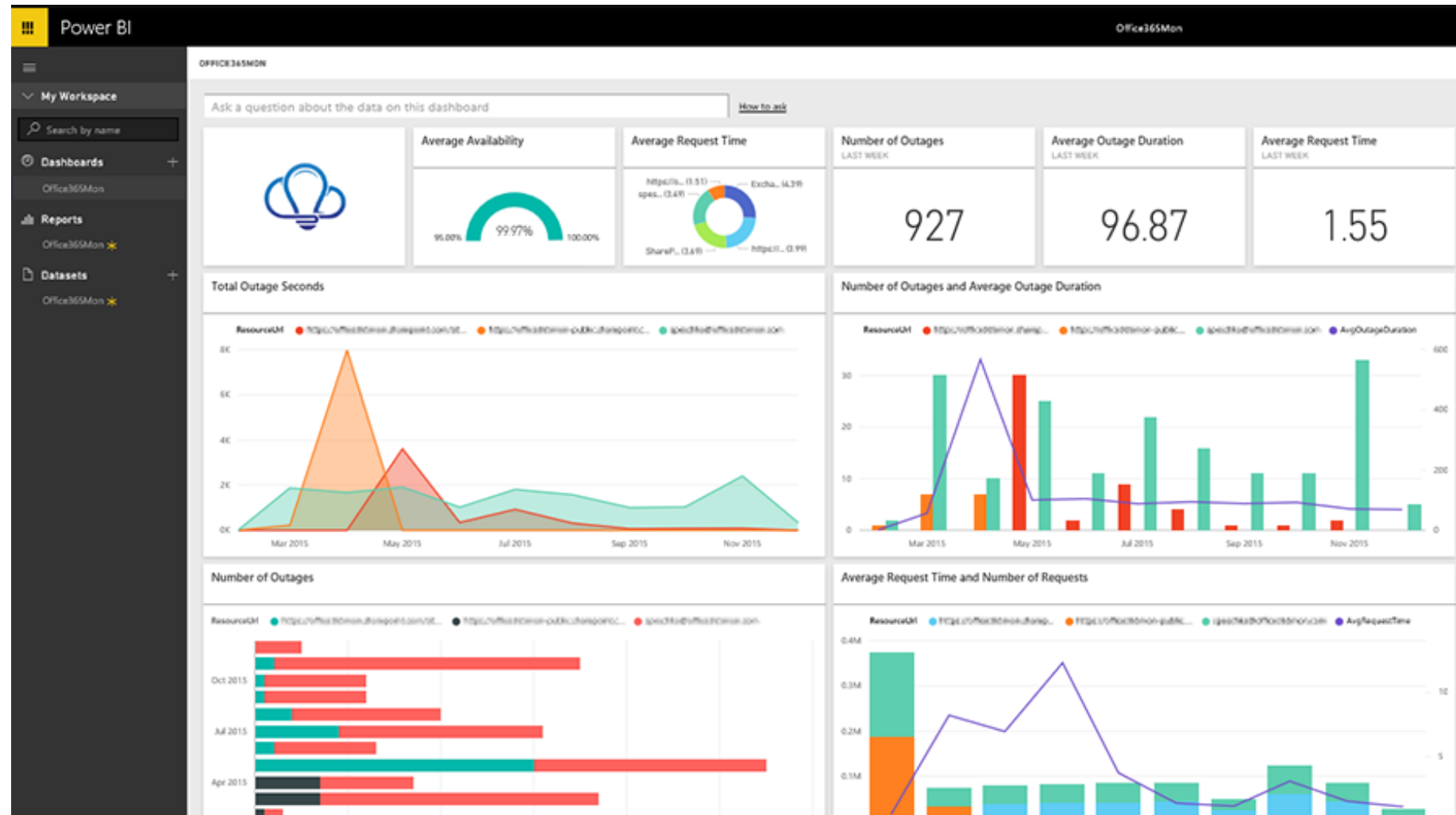
# Power BI

## Giá thành

- Nếu bạn có ngân sách nhỏ và tìm kiếm mức giá phải chăng, Power BI chính là câu trả lời. Bản Pro (Pro Version) hiện chỉ \$9.99/user/tháng, thấp hơn bất kì sản phẩm cùng loại

# Power BI

[https://www.youtube.com/watch?v=yKTSLffVGbk&feature=emb\\_logo](https://www.youtube.com/watch?v=yKTSLffVGbk&feature=emb_logo)



# QlikView

## Tổng quan cách sử dụng

QlikView là một công cụ tự phục vụ hoạt động với công nghệ liên kết của riêng mình và công nghệ bộ nhớ trong. Giao diện trực quan, dễ sử dụng, đi kèm tính năng tìm kiếm thông minh, không cần xây dựng khối, phù hợp cho phân tích ad hoc hơn là các phân tích hằng ngày.

## Tính năng BI

### Ưu điểm:

- Công nghệ bộ nhớ trong: QlikView tính toán dữ liệu trong khi dữ liệu chuyển từ backend sang frontend. Vì không có dữ liệu được tổng hợp hoặc tính toán trước được lưu trữ, nên bộ nhớ hệ thống có thể được lưu và tốc độ truyền dữ liệu nhanh.
- Tìm kiếm dữ liệu trực tiếp và gián tiếp: Bạn có thể tìm dữ liệu cả trực tiếp và gián tiếp. Trực tiếp, bạn có thể nhập dữ liệu và lấy thông tin, gián tiếp, nhập nội dung liên quan đến dữ liệu mong muốn và nhận tất cả dữ liệu liên quan.

### Nhược điểm:

- Không thân thiện với người dùng không có kỹ thuật: Hạn chế này có thể khiến người dùng cuối khó khăn để làm quen.
- Không hỗ trợ map charts: Trong khi biểu đồ bản đồ đang ngày một phát triển, tính năng này lại không hỗ trợ sẵn trong QlikView. Tạo các biểu đồ bản đồ địa lý khá phức tạp, trừ khi có tiện ích bổ sung được mua từ bên thứ ba

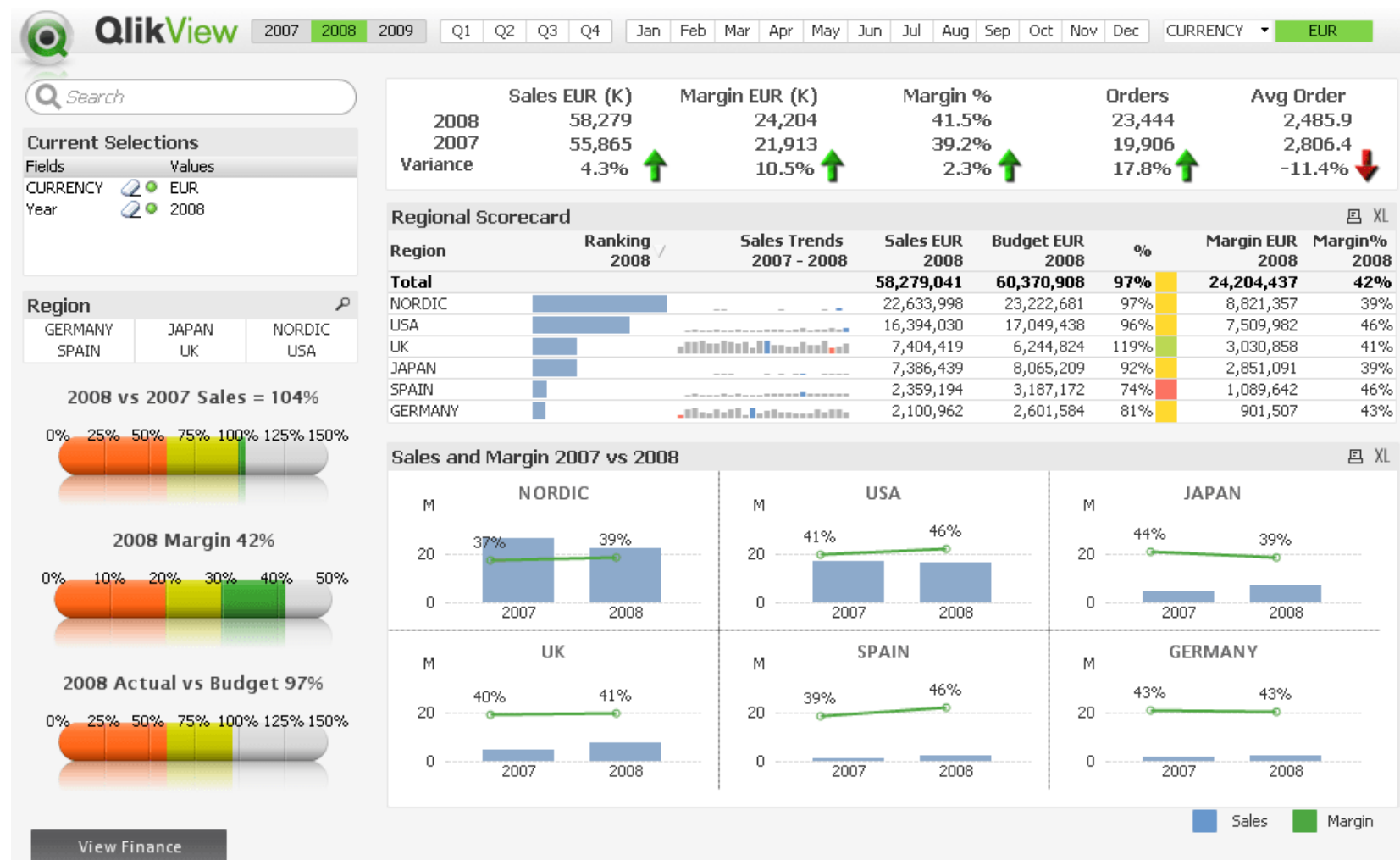
# QlikView

## Giá thành

Mỗi người dùng riêng trên QlikView có giá \$1350 và người dùng đồng thời là \$15000. Giấy phép cho máy chủ là \$35000, các dịch vụ khác có sẵn với chi phí bổ sung.

# QlikView

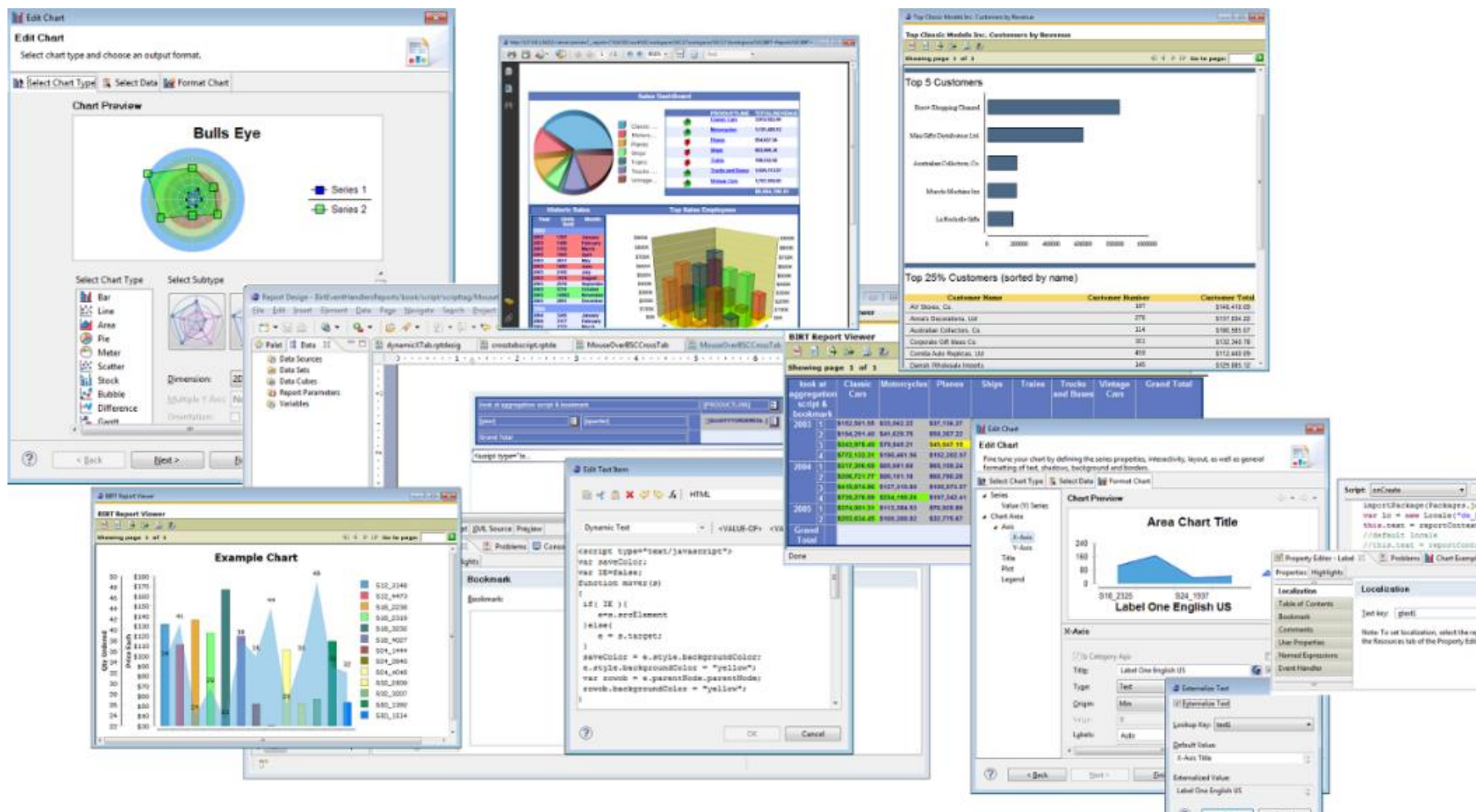
<https://www.youtube.com/watch?v=oSwr50bn9F8>



# Birt

Birt là viết tắt của Business Intelligence Reporting Tools. Birt cung cấp các phương pháp báo cáo, bao gồm danh sách, biểu đồ và các trực quan hóa dữ liệu khác. Birt là một trong những phần mềm mã nguồn mở miễn phí mạnh mẽ với nhiệm vụ chính là tạo các báo cáo có thể nhúng ở mọi nơi.

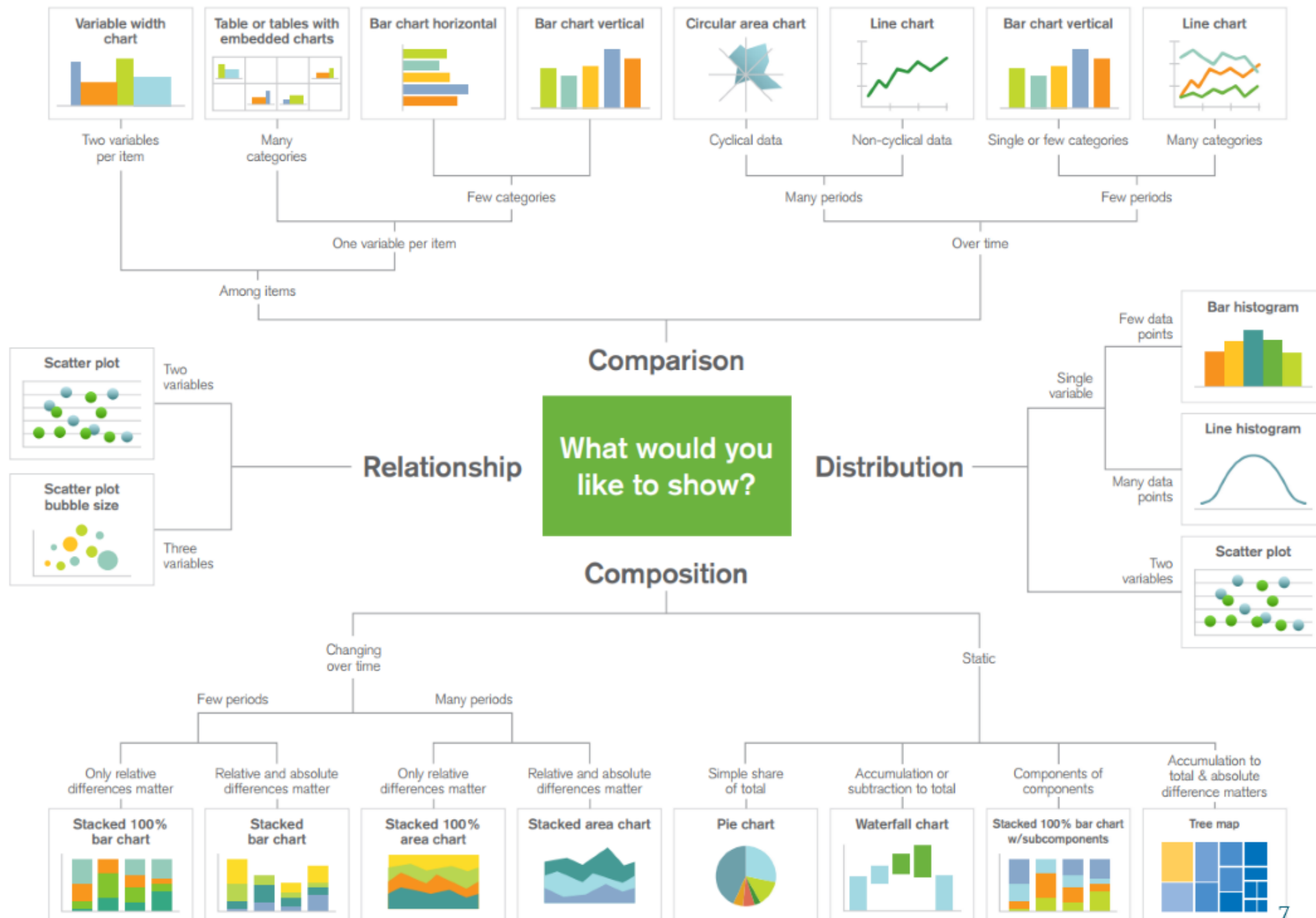
Đây là dự án cấp cao nhất của Eclipse dành cho Business Intelligence dựa trên Java.





### 3. Chart types





## 4. Hướng dẫn Birt



# Thành phần

**Report editor** : Là cửa sổ cho phép người sử dụng thiết kế giao diện báo cáo và preview báo cáo. Report editor có 5 trang, người dùng có thể chọn qua 5 tab tại phía dưới của cửa sổ:

Layout: Trang thiết kế giao diện

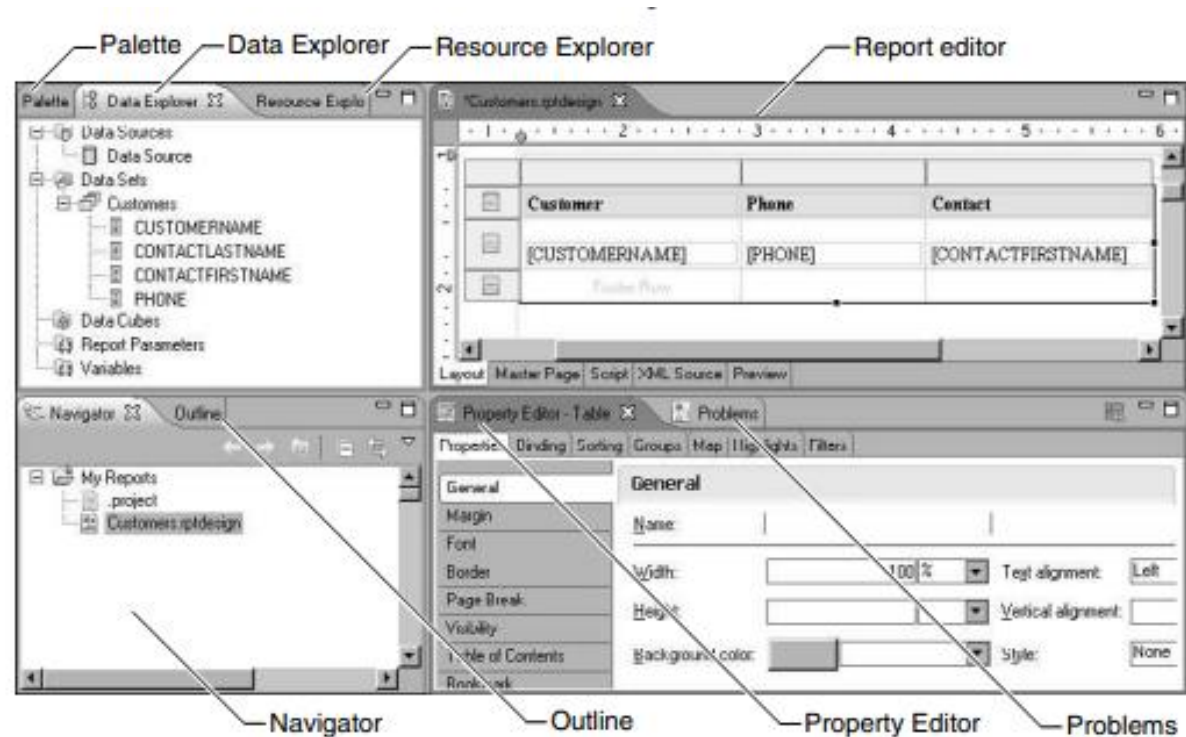
Master page: Hiện thị các mục của một trang báo cáo như: số trang, ngày giờ

Script editor: Nơi để thêm các đoạn code java vào báo cáo

XML source: Hiện thị cấu trúc/ cấu hình của báo cáo theo định dạng XML

Preview: Chạy thử báo cáo và xem kết quả

**Palette**: Là nơi chứa các nguyên tố (Report elements) để thiết kế báo cáo như: tables, grid, label ...Sử dụng bằng cách kéo thả các nguyên tố vào trang Layout của Report editor



# Thành phần

**Data explorer:** Định nghĩa nguồn dữ liệu cho báo cáo.

Data source: Cấu hình kết nối tới DB

Data set: Định nghĩa các dữ liệu trong báo cáo được lấy từ bảng nào, schema nào, trường nào

Data cube: Giống như data set nhưng nếu dữ liệu data set được định nghĩa ở dạng phẳng thì dữ liệu data cube được định nghĩa ở dạng không gian nhiều chiều

Report parameter: khai báo các biến dùng cho data set hoặc data cube

Variable: Khai báo các biến dùng để thiết kế giao diện báo cáo

**Resource explorer:** Nơi chứa các thư viện dùng chung cho các báo cáo. Thư viện này thường được hiểu là các kho phong chữ, bảng biểu đã được format theo một form nhất định ...

**Property editor:** Nơi chứa các thuộc tính của các report elements(ở trong Palette)

**Navigator:** Cây thư mục hiển thị các projects và báo cáo, thư viện thuộc project

**Outline:** Nơi hiện thị cấu trúc của báo cáo được tổ chức theo dạng hình cây

**Problems:** Nơi hiện thị các lỗi khi thiết kế báo cáo

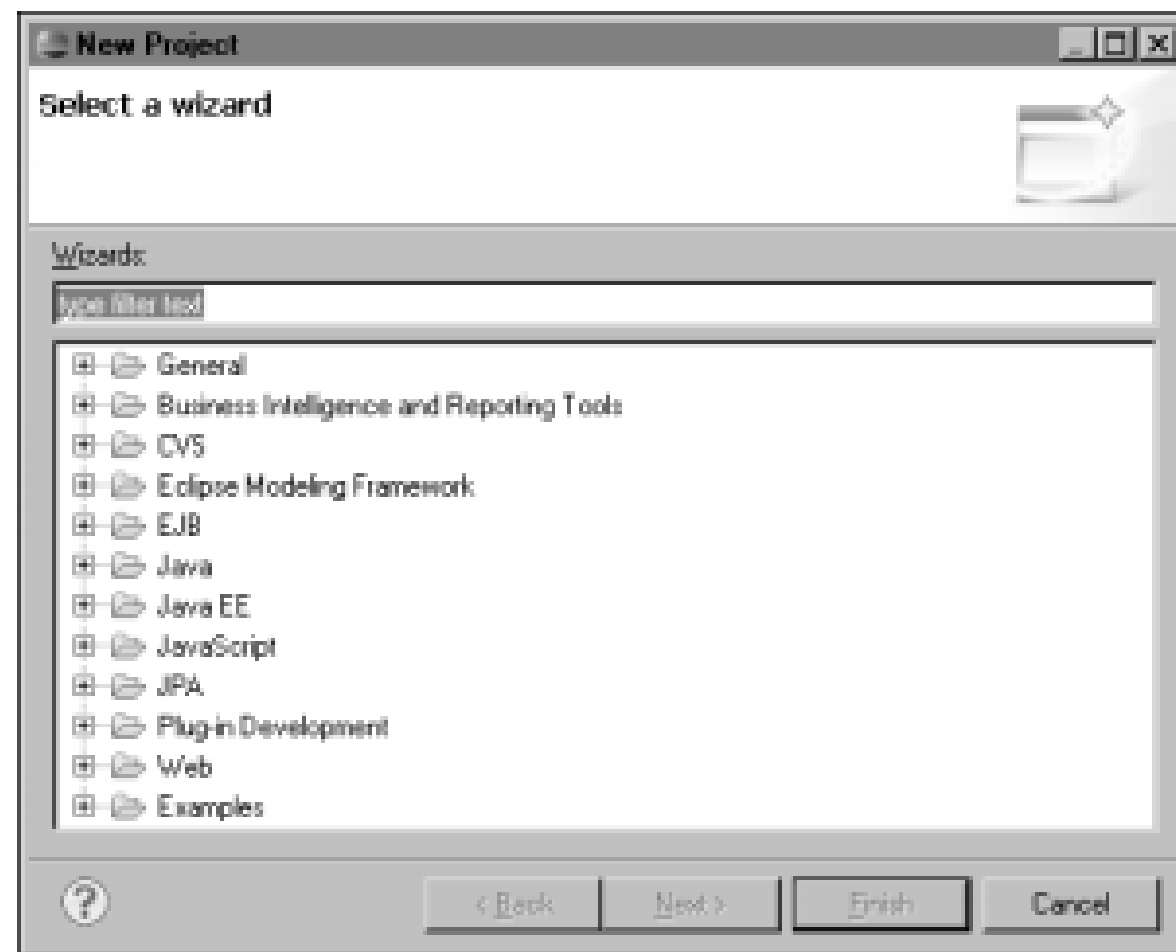
# Dựng báo cáo

Task 1 : Khởi tạo project mới

Chọn File/New/Project/New Project. Cửa sổ hiển thị các loại projects

Mở **Business intelligent and reporting tools** chọn **Report project** rồi ấn **next**

Trong cửa sổ New Report Project đặt tên cho project rồi ấn **next** → **Finish**. Bây giờ bạn có thể thấy project của bạn trong tab Navigator





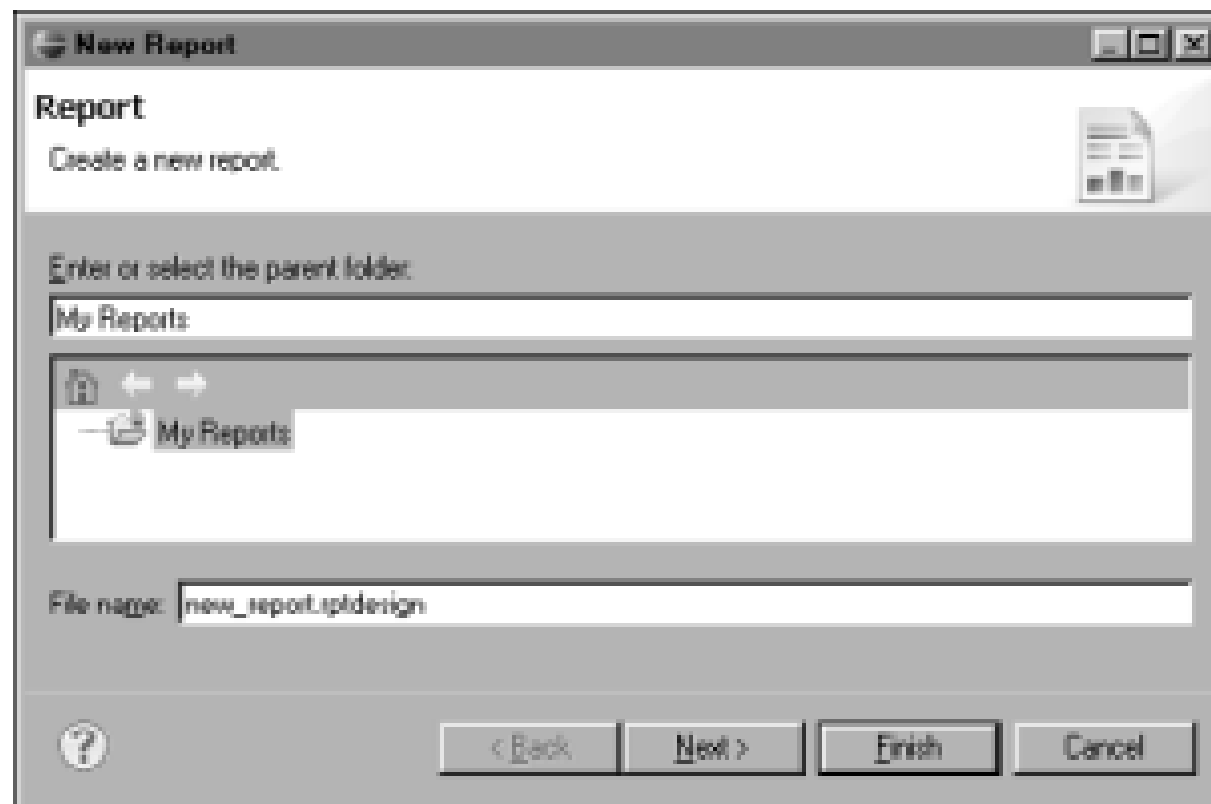
# Dựng báo cáo

## Task 2: Khởi tạo báo cáo

Bạn có thể khởi tạo một báo cáo mới bắt đầu bằng một báo cáo trắng(Blank report) hoặc có thể chọn các template đã có sẵn. Ở đây sẽ hướng dẫn tạo báo cáo bắt đầu bằng blank report

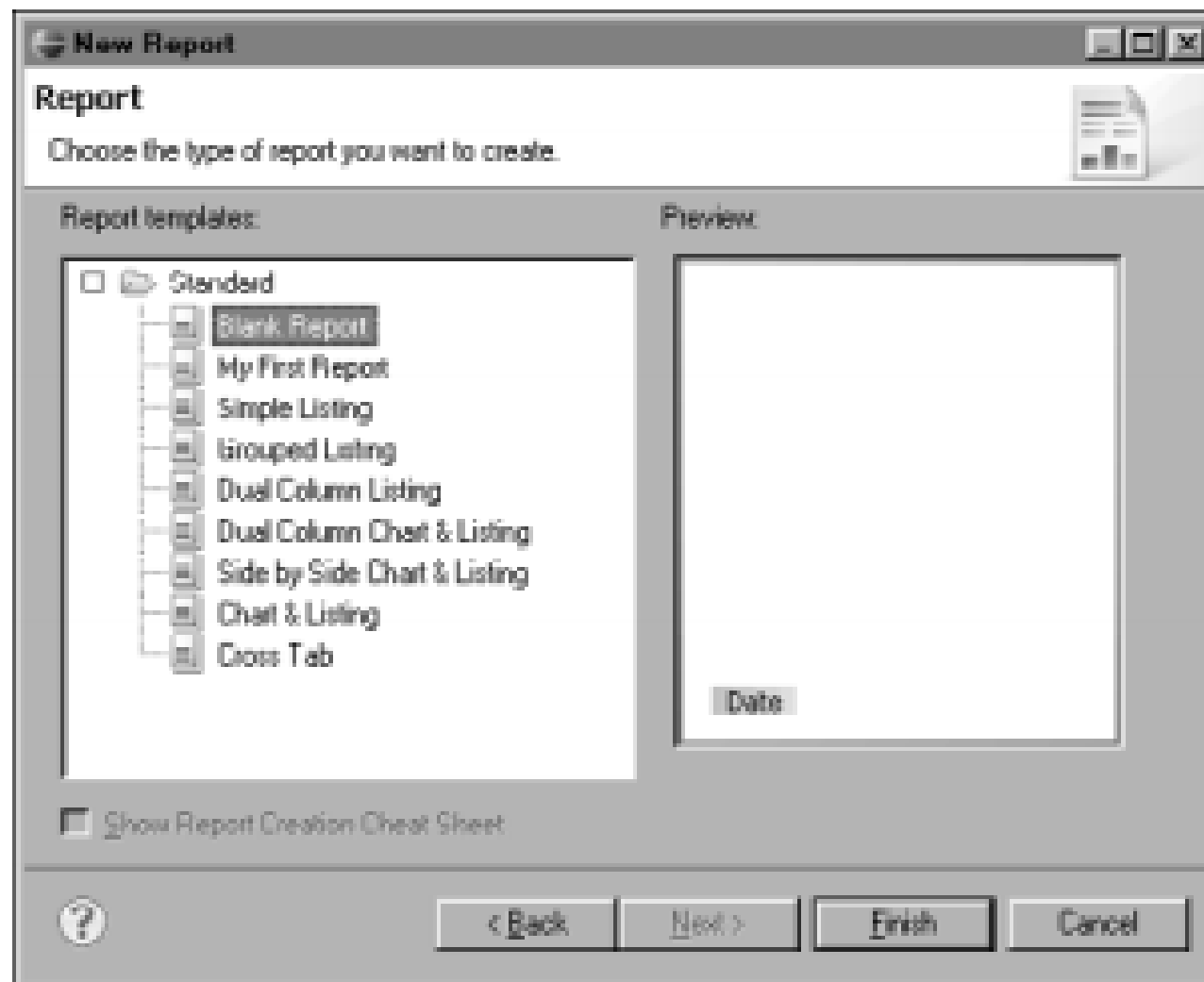
Chọn **File/New/Report. New Report**. Đặt tên cho báo cáo , chọn project cho báo cáo rồi ấn **next**

Chọn project mà bạn muốn để báo cáo vào, đặt tên cho báo cáo và ấn **next**



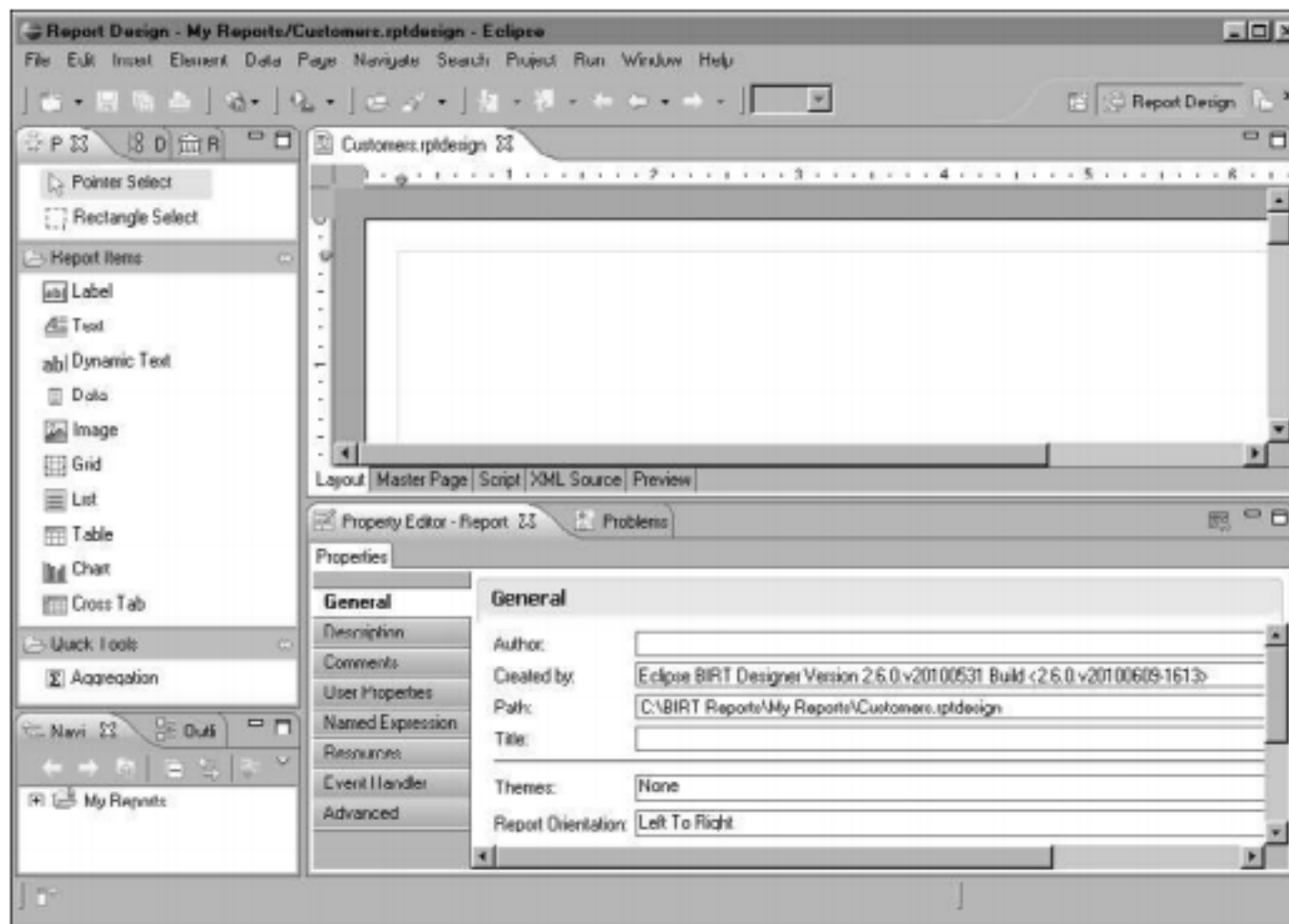
# Dựng báo cáo

Chọn **blank report** rồi ấn **finish**



# Dựng báo cáo

Một blank report sẽ có dạng như sau

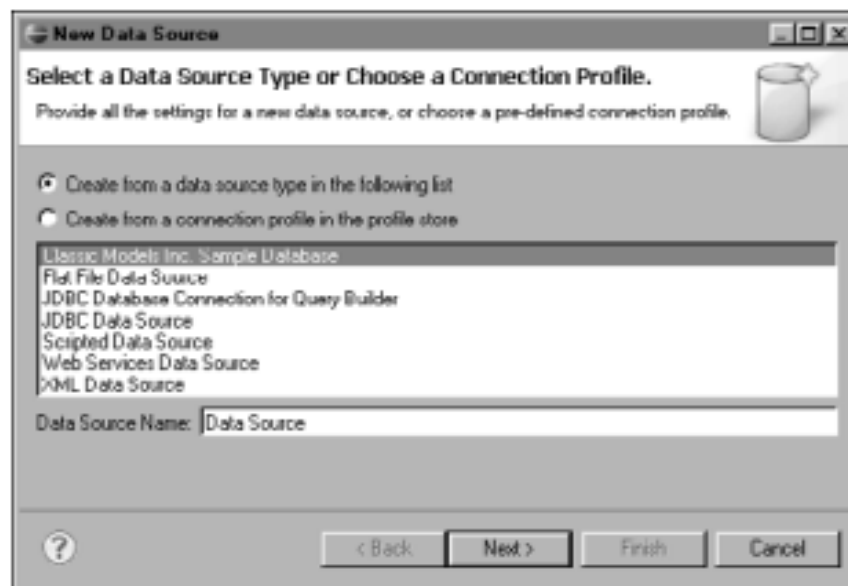
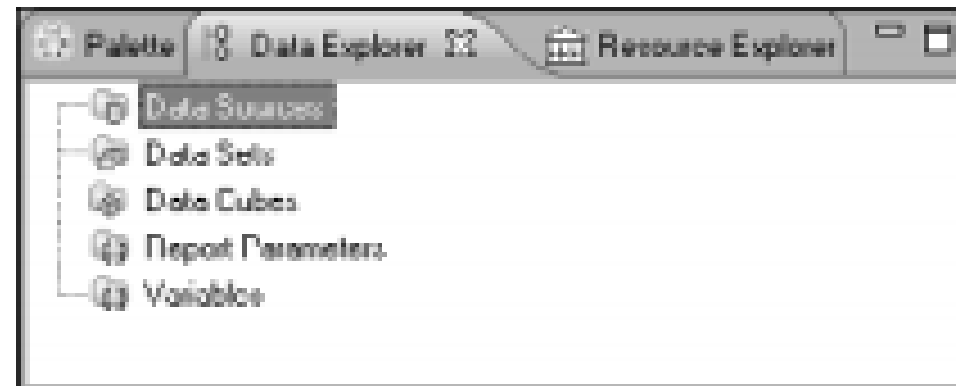


# Dựng báo cáo

## Task 3: Xây dựng data source

Chọn tab data Explorer(Nằm bên trái màn hình), Nếu không thấy có thể chọn **Window/Show/View/Data Explorer**

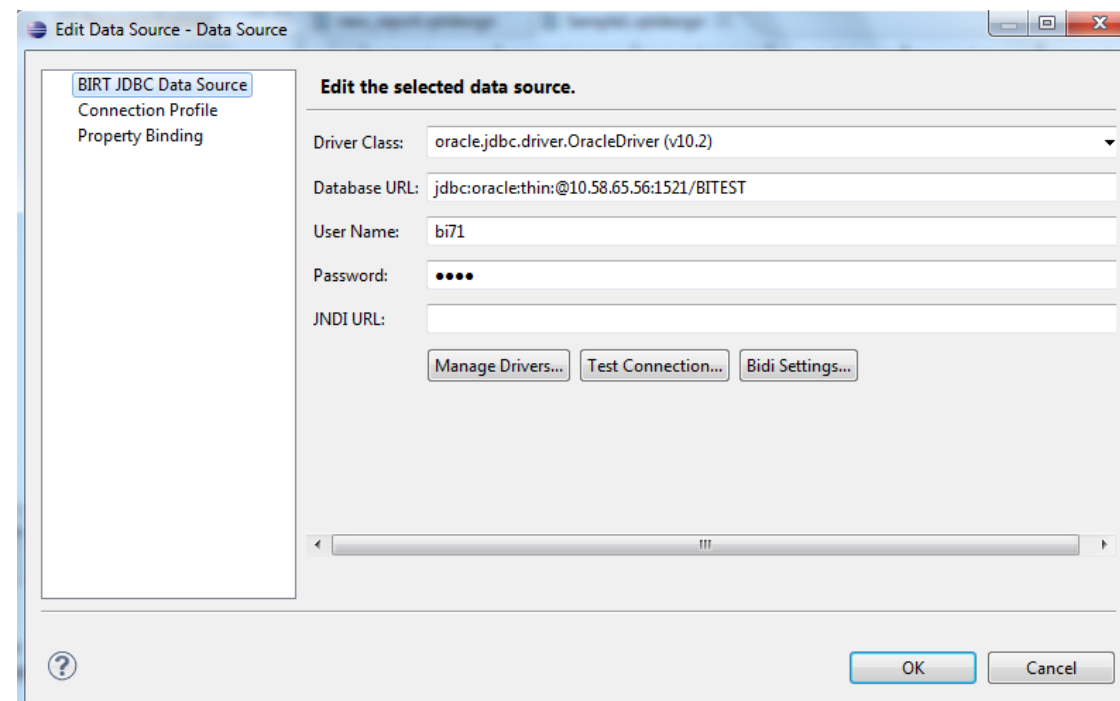
Ấn chuột phải vào **Data source** rồi chọn **New Data Source**.  
Cửa sổ New Data Source sẽ xuất hiện như hình vẽ dưới đây



# Dựng báo cáo

Chọn **JDBC data source**, rồi điền thông tin như hình dưới.  
**Driver class** dùng JDBC driver, **database**: oracle. Chú ý nếu không chọn được JDBC oracle driver thì vào Manage driver và add đường dẫn tới thư mục chứa file JDBC driver

Ấn OK. Sau khi hoàn tất đối tượng sẽ được hiển thị như sau

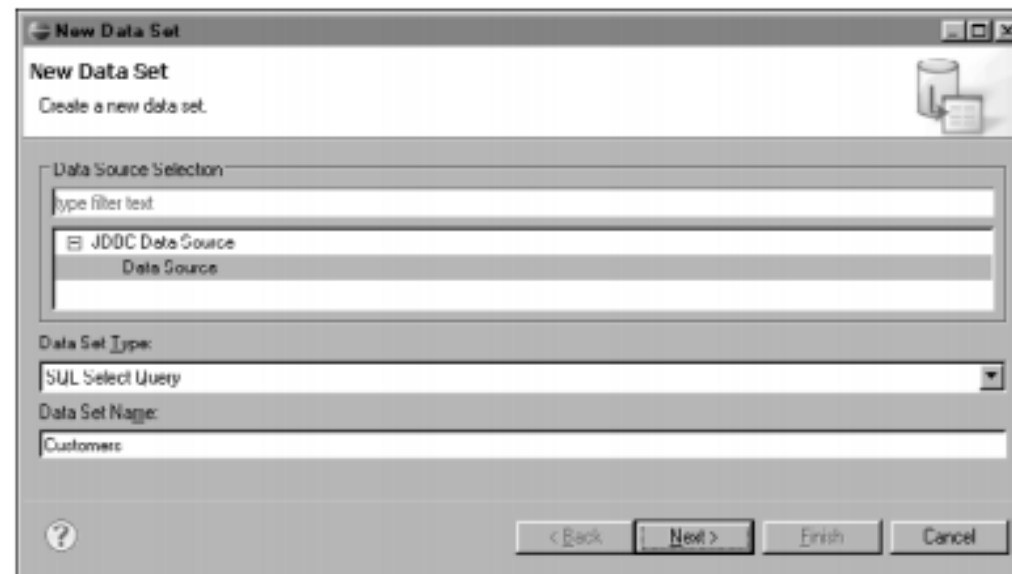


# Dựng báo cáo

## Task 4: Xây dựng DataSet

Trong tab **Data Explorer** ấn chuột phải vào **Data Sets** chọn **New Data Set**

Trong cửa sổ New Data Set chọn Data Source, Data Set Type( chọn Sql Select Query) và đặt tên cho Data Set.



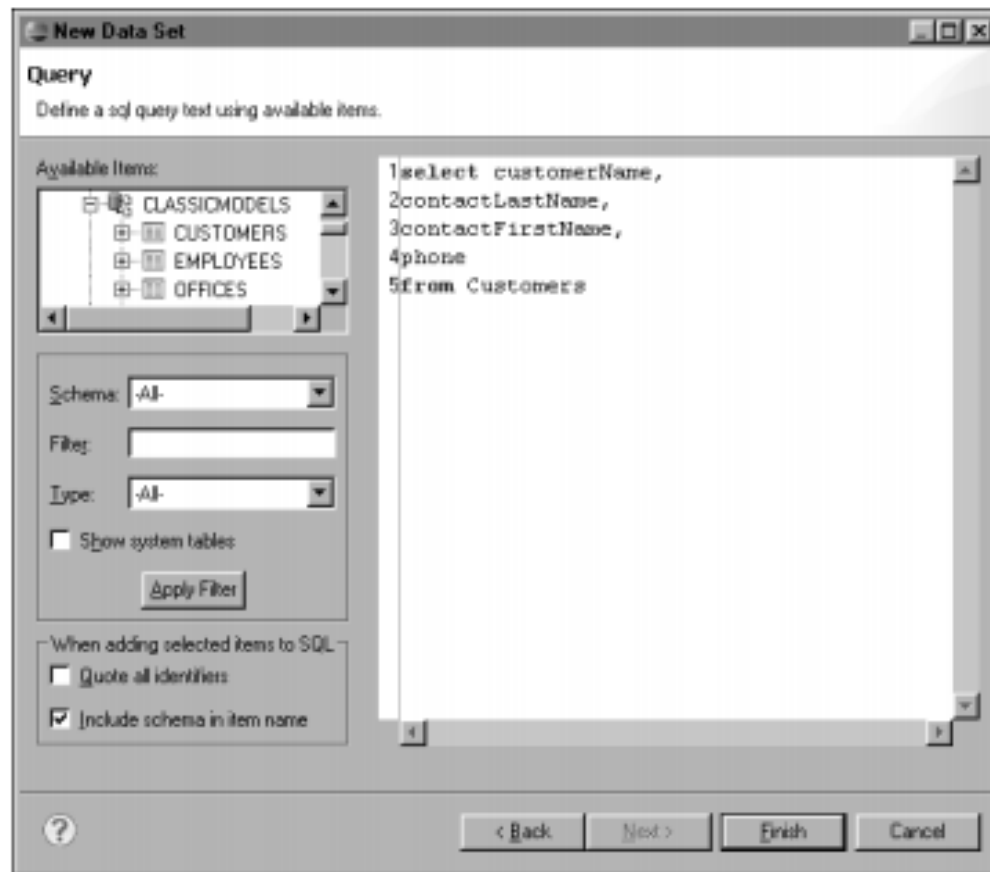
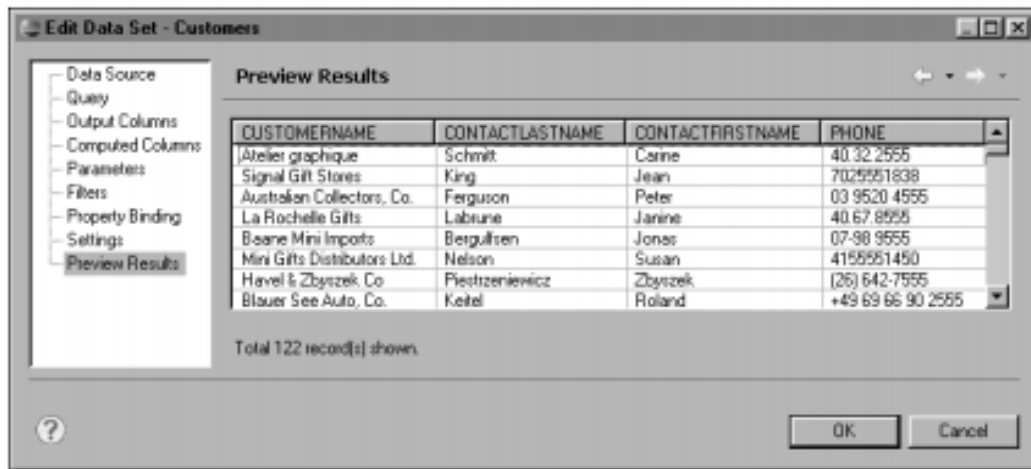


# Dựng báo cáo

Ấn **Next**, cửa sổ Query sẽ hiện lên

Trong phần Text area bạn có thể tùy biến câu lệnh SQL, chọn các tiêu chí, chiều muốn xem trên báo cáo. Ấn **finish** để save Data set

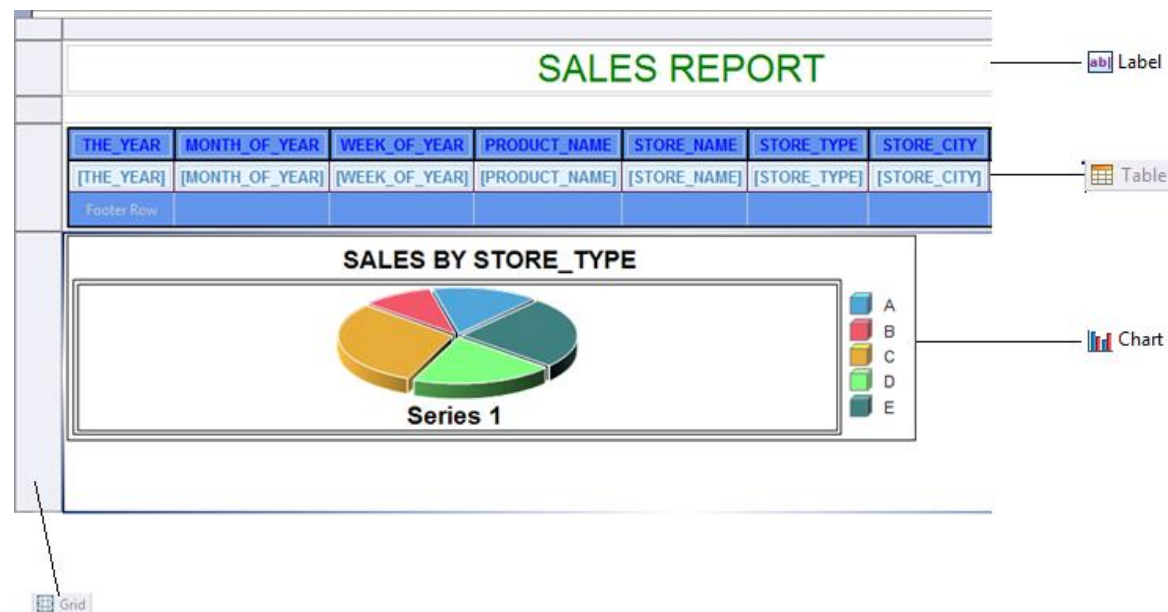
Nếu câu lệnh SQL chuẩn xác thì trong cửa sổ Edit data set (Chuột phải vào **Data sets** bạn vừa tạo và chọn **Edit**)/**Preview Result** sẽ hiện thị kết quả truy suất



# Dựng báo cáo

## Task 5: Thiết kế giao diện:

Chọn tab **Palatte**( Xem 2.1), ở đây chứa các thành phần để xây dựng một layout cho báo cáo. Người dùng có thể kéo thả các thành phần này vào cửa sổ Layout để thiết kế. Một thiết kế cơ bản cho báo cáo sẽ như sau:



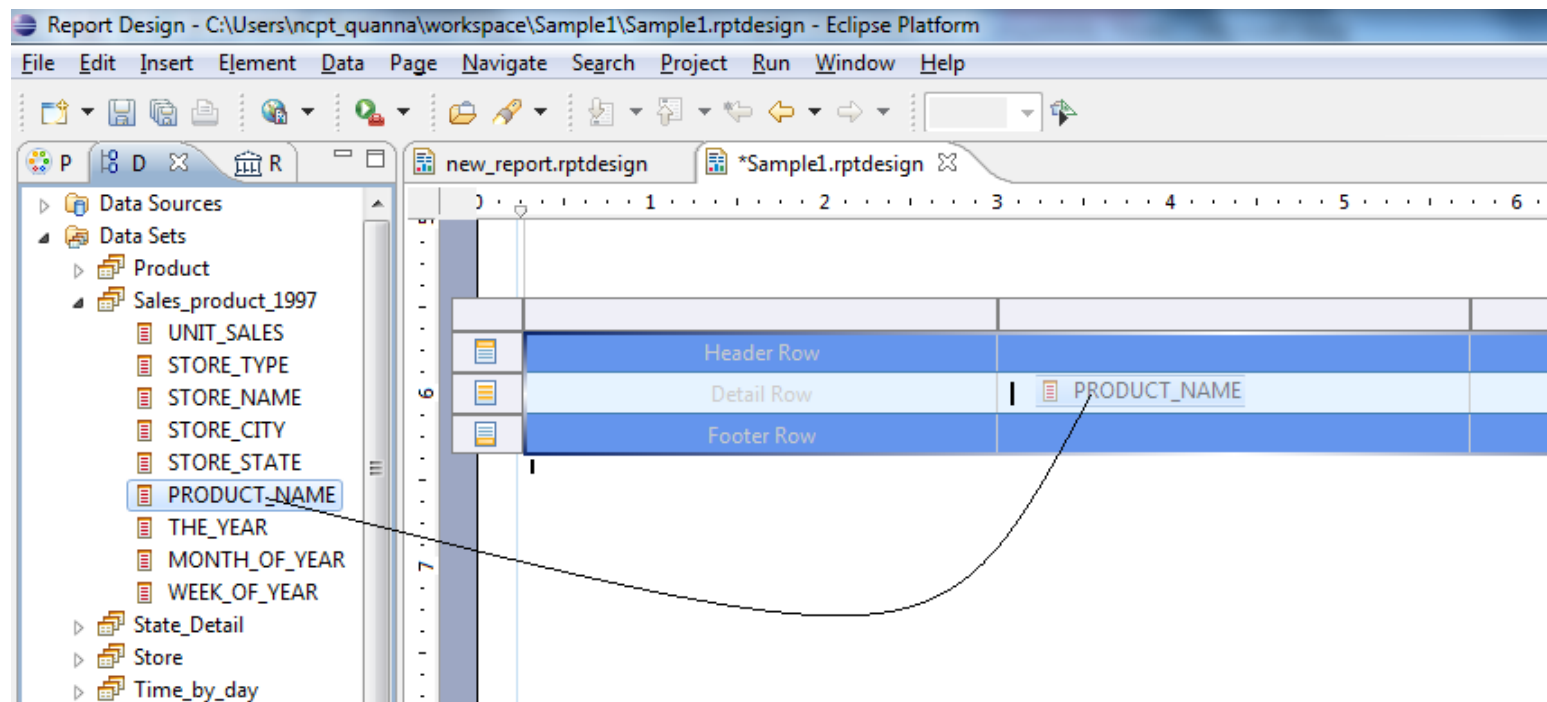
**Grid:** Được hiểu như một tấm “lưới” dùng để xác định vị trí mà người thiết kế muốn đặt báo cáo lên màn hình. Grid có dạng hàng và cột, các hàng và cột này tạo thành các cell, một Grid có thể có 1 hoặc nhiều cell (Tùy chọn theo người dùng). Thông thường với mọi báo cáo Grid luôn được sử dụng như một nền chính, sau đó mới bỏ các thành phần khác vào Grid để tạo giao diện

**Table:** Bảng dùng để hiển thị dữ liệu báo cáo. Thành phần của một bảng bao gồm HEADER/DETAIL/FOOTER. Để định nghĩa cột nào sẽ hiển thị dữ liệu nào thì người dùng sẽ gắp cột dữ liệu được định nghĩa trong **Data Set** để gắp thả vào cột tương ứng. Phần **header row** sẽ tự động lấy tên như tên được đặt trong data set. **Detail row** sẽ hiển thị dữ liệu được truy suất từ database. **Footer row** dùng để thực hiện các hàm như sum, max,min ...

# Dựng báo cáo

**Label:** Nhãn dùng để đặt tiêu đề cho báo cáo

**Chart:** Hiển thị dữ liệu dưới dạng biểu đồ



# Dựng báo cáo

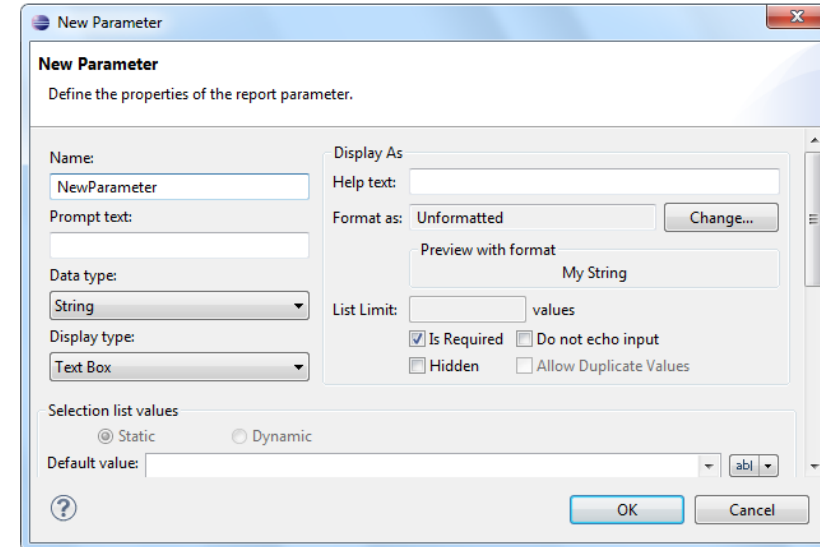
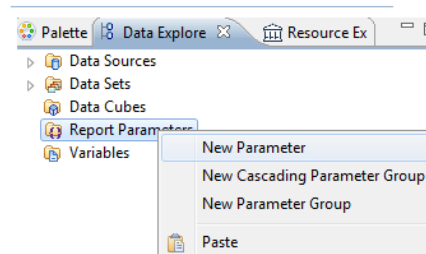
## Task 6: Sử dụng tham số trong báo cáo

Tham số được nói đến ở đây là tham số cho người dùng tùy chọn để xem dữ liệu mà họ mong muốn lên báo cáo. Tham số sẽ được cấu hình trong khi tạo Data set (**Task 4**). Để có thể giúp hình dung một các cụ thể hơn ta sẽ làm qua ví dụ sau:

Giả sử một báo cáo bán hàng đơn giản:

Month	Product	Unit sales
-------	---------	------------

Select Month, product, unit\_sales  
from sale\_fact\_table  
where month = ?  
and product = ?



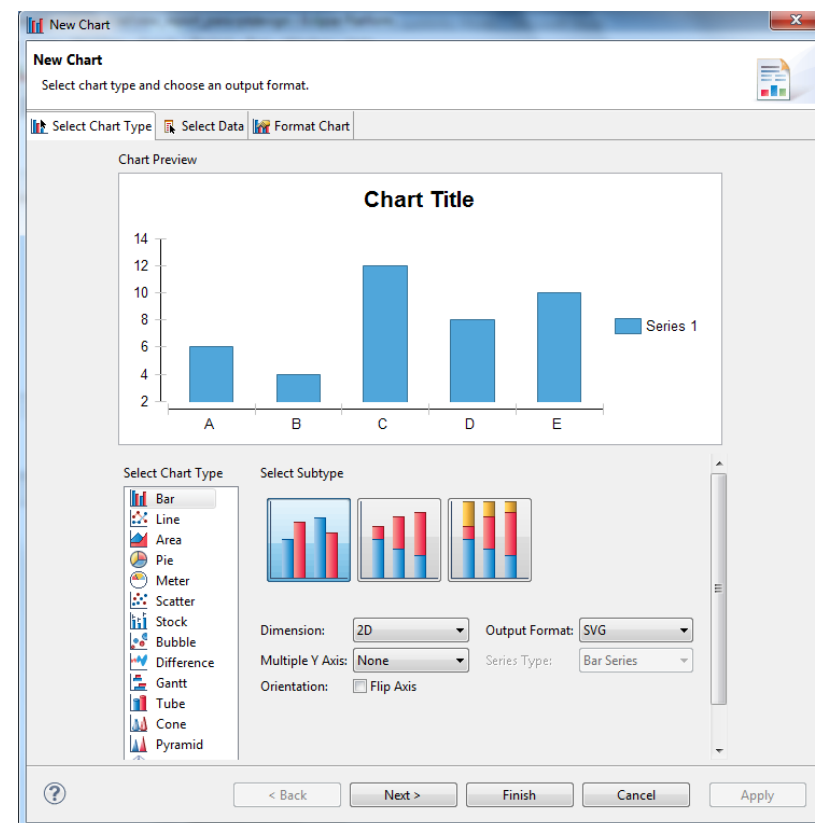
# Dựng báo cáo

## Task 7: Vẽ biểu đồ

Để tạo biểu đồ người dùng sẽ sử dụng thành phần **chart** trong panel **Palette**.

Cửa sổ **New Chart** sẽ xuất hiện khi người dùng kéo thả thành phần **chart** vào tab **Layout** của panel **Report editor**, cửa sổ gồm có 3 tabs: **Select Chart Type**, **Select Data**, **Format Chart**.

**Select Chart Type:** Cho phép người dùng chọn loại biểu đồ để biểu diễn số liệu. Các loại biểu đồ được birt cung cấp đầy đủ và rất đa dạng, từ các biểu đồ cơ bản như bar chart, line chart, pie chart đến các biểu đồ nâng cao như difference chart, grant chart.... Tùy vào từng loại biểu đồ người dùng có thể chọn cách hiển thị 2D, 2D with Depth, hay là 3D, thêm trục tung vào biểu đồ(secondary axis), và chọn định dạng hình ảnh cho biểu đồ : bitmap, jpg, png...



## 5. Thực hành



# Practice

## 1. Dùng Data source sample dựng báo cáo:

- Hiển thị danh sách giao dịch bán hàng
- Vẽ biểu đồ đường doanh số hàng hóa theo đất nước

## 2. Dùng CSDL MySQL xây dựng báo cáo hiển thị:

- Danh sách sinh viên kèm điểm các Môn, cho phép chọn Môn
- Biểu đồ thể hiện phân phối mật độ điểm trong Môn