Databases

Cap. 1. Introduction. Data Storage. DBMS

Textbook: Ramakrishnan, Gehrke, “Database Management Systems”, McGraw Hill, 2003

2022 UPT

Assoc.Prof.Dr. Dan Pescaru

Course organization

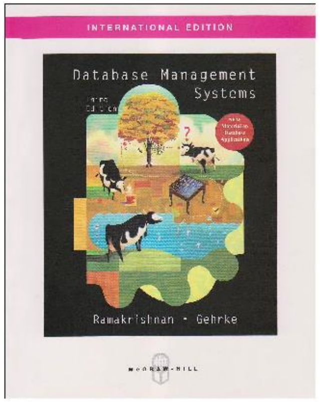
1. Course
   - Dan Pescaru
   - Monday, 12:00-14:00
2. Labs
   - B623
3. Evaluation
   - Lab mark (1/2)
   - Exam (1/2)

Contents

1. Introduction. DBMS. Data Storage.
2. Data modeling. ER conceptual model. Relational data model.
3. Relational algebra. Operators. Normal forms. DB query languages.
4. Indexing. DB optimization. DB Dictionary.
5. SQL. Using DDL. Constraints. Active queries.
6. SQL. Data projection, selection, ordering and joins.
7. SQL Subqueries. Data union and difference.
8. SQL. Data aggregation.
9. SQL Execution plan. Introduction in query optimization.
10. Implementing DB clients in Oracle APEX. Forms and reports.
11. Implementing Web databases using MySQL and PHP.
12. Overview of database administration and security.

Bibliography

- R- Ramakrishnan şi J. Gehrke, "**Database Management Systems**", 3rd edition, ISBN 007-2465-63-8, McGraw-Hill, 2003



- **Oracle® 21c Database SQL Language Reference**, July 2022
  https://docs.oracle.com/en/database/oracle/oracle-database/21/sqlrf/

- **MySQL 8.0 Reference Manual – SQL Syntax**, March 2022
  https://dev.mysql.com/doc/refman/8.0/en/sql-statements.html

Data storage and retrieval

1. Basic storing: the file
2. Physical data organization: directory structure
3. Some advantages:
    1. Simplicity
    2. Linear access (text) or sequential (binary)
4. Some disadvantages:
    1. Searching data in files – complex application specific algorithms
    2. Lack in data protection
    3. Security and access control – at the OS level

DB and DBMS

1. Solution: using a database
2. Database (DB)= a data storage alternative for very large, integrated collection of data. Include support for efficient physical organization on external support, advanced searching and retrieval algorithms, data protection, and security mechanisms
3. Database Management System (DBMS) = a software system designed for database storage and management

Why DBs are important?

1. Wide spread nowadays: almost all people use them every day
   1. Most websites are built on DBs
   2. Telecom systems (mobile phones calls)
   3. Banking systems
   4. Merchandising, etc.
2. High demand on the employment market for DB specialists (DB administrators, analysts and designers, DB App programmers etc.)

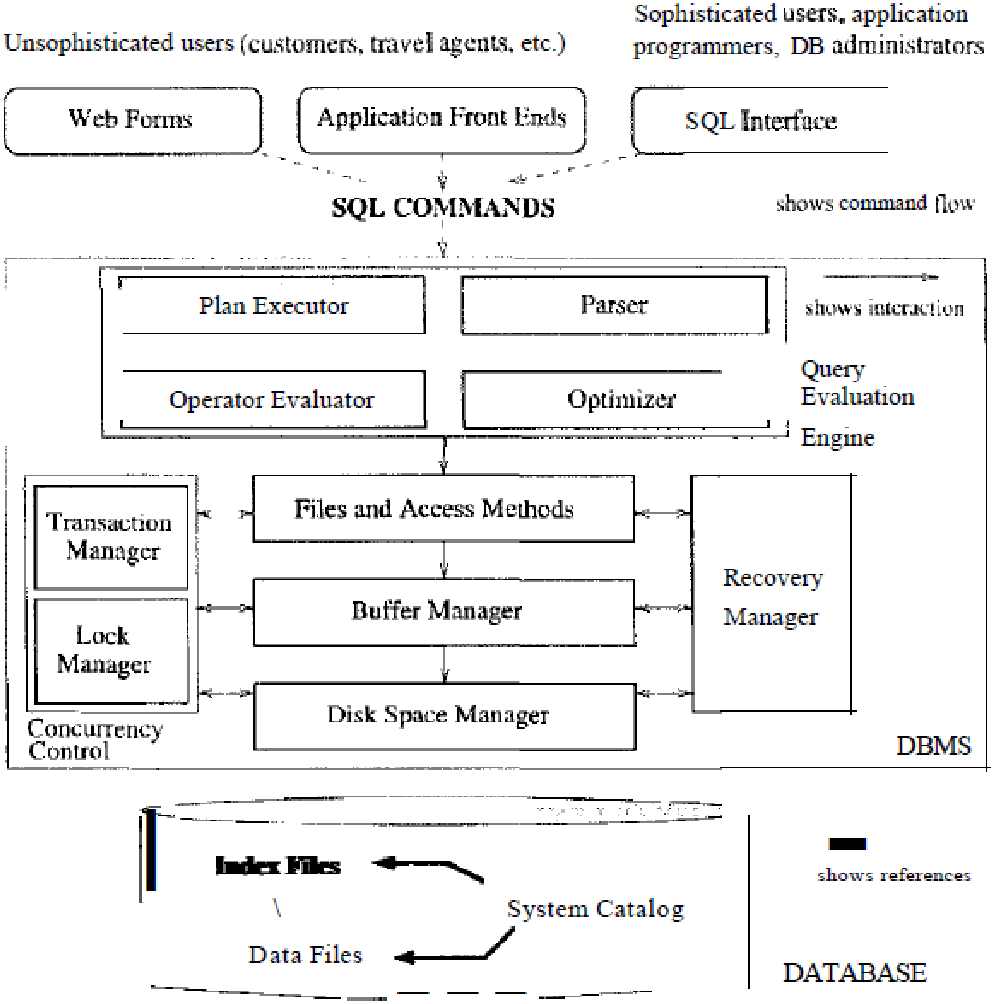
DBMS Advantages

1. Data independency / multiple applications
2. Data access and retrieval efficiency (based on indexes and query optimization)
3. Enable RAD techniques
4. Support for data integrity
5. Support for data security
6. Application domain independency for data management
7. Scalability

Main DBMS Functions

1. Data storage and retrieval
2. Index management
3. Query processing
4. Access control mechanism (users, groups/ roles, privileges)
5. Data integrity control mechanisms (triggers, constraints, concurrency control)
6. Crash recovery mechanisms, replication
7. Transactions

Architecture of a relational DBMS



*Ref: Ramakrishnan, Gehrke, “Database Management Systems”, McGraw Hill, 2003

Specialized DBMS

1. Multimedia databases (video on demand, police fingerprints DB, photo journalism library, etc.)
2. ERP (Enterprise Resource Planning)
   - Substantial layer of application-oriented features on top of a DBMS
   - Support a set of customizable enterprise common tasks (e.g., inventory management, human resources planning, financial analysis)
   - SAP, PeopleSoft, Baan, Siebel

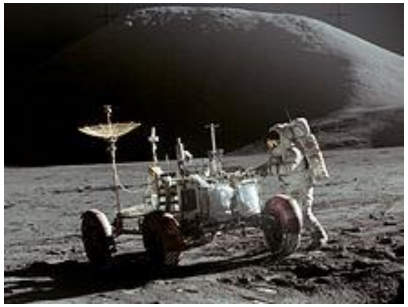
Support for massive research projects

1. The human genome mapping 1987-2003
   - US Department of Energy’s Office of Health and Environmental Research
   - Results: database known as GenBank (http://www.ncbi.nlm.nih.gov/nucleotide)
2. SETI (search for extraterrestrial intelligence)
   - University of California, Berkeley, May 1999
   - Over 290,000 computers with 617 teraFLOPS
3. NASA's Earth Observation System
   - Global observations of the land surface, biosphere, atmosphere, and oceans of the Earth
   - Collects data from over 23 artificial satellites

History. The hierarchical model

1. The hierarchical model – 1960

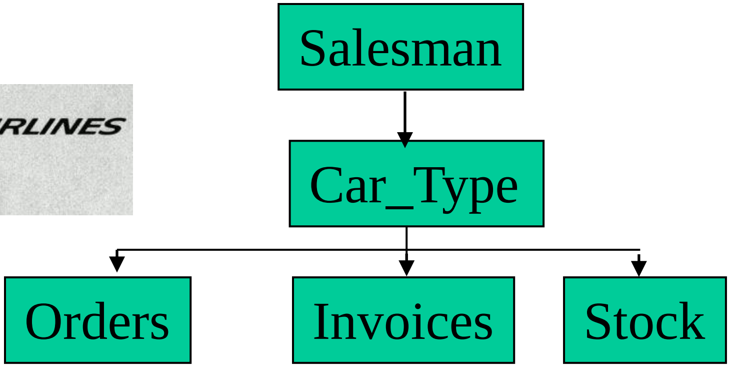
- Three structure (1-to-many relationships)
- IBM IMS – Information Management System – designed for the Apollo program, it was used to inventory the very large bill of materials for Saturn V moon rocket and Apollo space vehicle



NASA: http://www.hq.nasa.gov/

- American Airlines + IBM – SABRE

AA AMERICAN AIRLINES SABRE Reservation System by IBM



Salesman

Car_Type

Orders

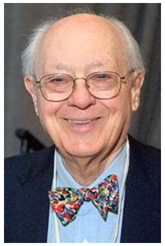
Invoices

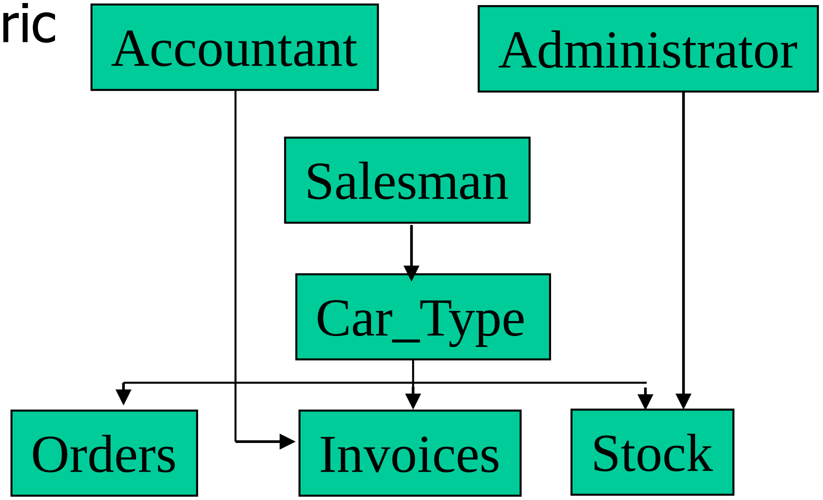
Stock

History. The network model

1. The network model – 1960

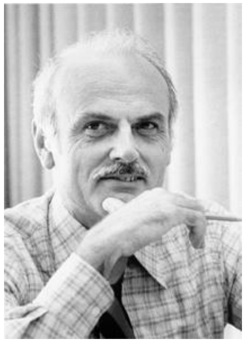
- Generalized graph structure
- Codasyl ("Conference on Data Systems Languages" consortium – COBOL)
- IDS (Integrated Data Store) - designed by Charles Bachman at General Electric
- 1973: Bachman won the ACM's Turing Award (⇔Nobel Prize) for his work in the database area





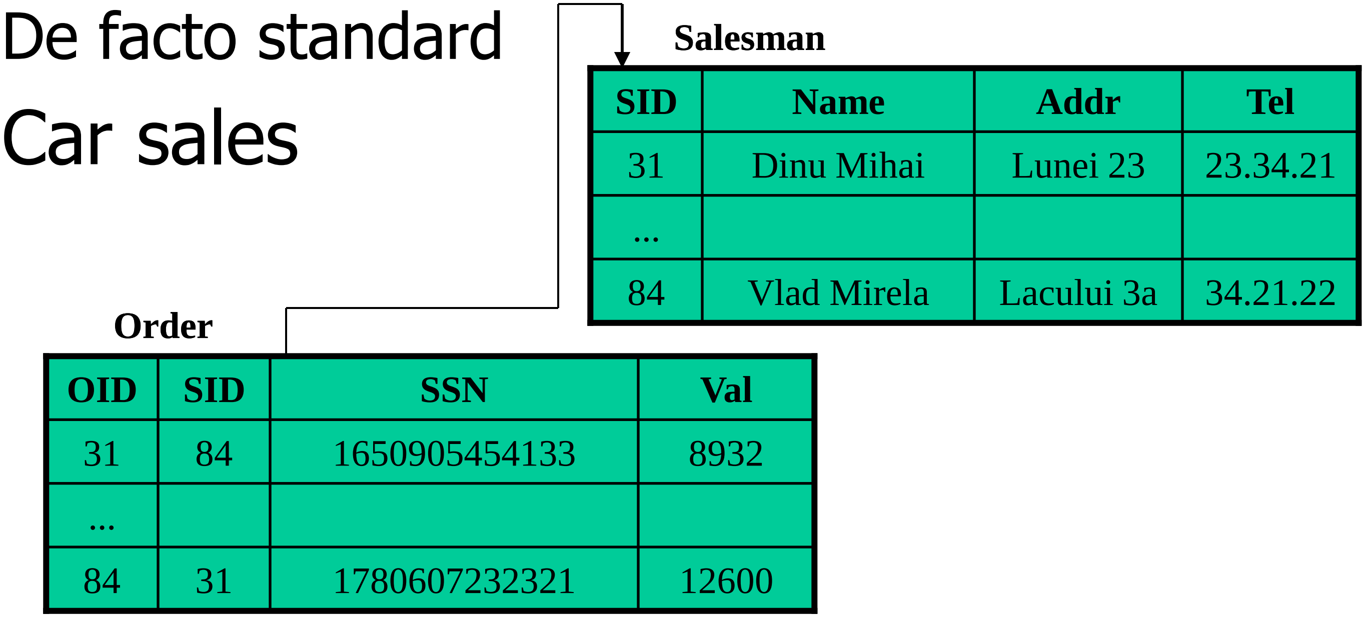
Today. The relational model

1. The relational model – 1970
   - Edgar Codd, at IBM's San Jose Research Laboratory
     - He won the Turing Award in 1981
   - Novelty: a firm mathematical foundation
   - Navigational systems for Desktops
     - xBase: FoxPro, Clipper, Visual dBase
   - Declarative – SQL (IBM's System R project)
     - Oracle, IBM DB2, Ms SQL Server, MySQL



The relational model

1. Main characteristics
   - Describes Tables and Links between them
   - Based on relational algebra
   - Powerful query language (SQL)
   - De facto standard
2. Ex: Car sales



Salesman

SID	Name	Addr	Tel
31	Dinu Mihai	Lunei 23	23.34.21
…			
84	Vlad Mirela	Lacului 3a	34.21.22

Order

OID	SID	SSN	Val
31	84	1650905454133	8932
…			
84	31	1780607232321	12600

Relational: navigation vs. querying

1. xBase
   - Navigational, imperative language
   - Low level operations (allows high optimizations)
   - Requires low resources (ideal for standalone DB old Desktop PCs)
2. SQL
   - Declarative querying language
   - Nowadays standard for medium and large size server based databases
   - Flexible and efficient

Novel DB models

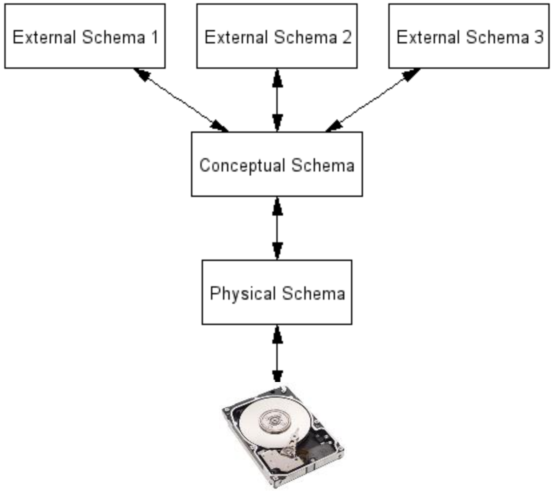
1. The Object Database model – 1990
   - Persistency concepts to OO Languages
   - ODMG - Object Query Language (OQL)
   - ORION (MCC), Jasmine (Fujitsu), O2, POET, ObjectStore, JADE
   - Compromise solution: object-relational model
2. The Post-relational model – 2000
   - fast key-value stores and document oriented databases
   - No-SQL (MongoDB, Redis, Apache Cassandra, etc.)

Data and requirement analyzes

1. First step in DB design
2. Uses a semantic data model
   - Abstract, high-level data model used to describe the data in an enterprise
   - Serves as the starting point for DB modeling and design
3. Semantic data models
   - Entity Relationship (ER) - pictorially denote entities and the relationships among them (Logical modeling)
   - UML – more general than ER (Business, System, Logical, Physical, Hardware Modeling)

Levels of Abstraction

1. The data in a DBMS is described at three levels of abstraction
   - Physical: how the data is stored and where it is stored in database
   - Conceptual: describe the model of data
   - External: simplified domain-specific views



The physical schema

1. Specifies storage details
   - How the relations described in the conceptual schema are actually stored
   - How data is split into files and records sorting strategy
   - What auxiliary data structures are needed (e.g. index files). They are essential for data access efficiency
   - Good design imply a deep understanding of how the data is typically accessed

The conceptual schema

1. Known sometimes as the logical schema
    - Describes the stored data in terms of the data model of the DBMS
    - Contains details about data (e.g. data type)
2. Ex: a relational model for University DB:

    Students(sid: string, name: string, birth: date, year: real)

    FacultyStaff(fid: string, fname: string, sal: real)

    Courses(cid: string, cname: string, credits: integer)

    Rooms(no: integer, address: string, capacity: integer)

    Enrolled (sid: string, cid: string, grade: string)

    Teaches(fid: string, cid: string)

The external schema

1. Allow data access to be customized (and authorized) at the level of individual users or groups of users (or external applications)
2. Consists of a collection of one or more views and relations from the conceptual schema
3. A view is basically a relation, but the records in a view are not stored in the DBMS
4. The external schema design is guided by end user requirements

Peoples who work with databases

1. DBMS implementers (employed by Oracle, Microsoft, IBM etc.)
   - Build DBMS software
2. Database end users – from a large variety of fields
   - Accountants, managers, secretaries, etc.
3. Database application programmers
   - Develop packages that facilitate data access for end users
4. Database administrators
   - Design of the Conceptual and Physical Schemas
   - Manage DB Security and Authorization
   - Are responsible for data availability and crash recovery
   - Perform Database Tuning

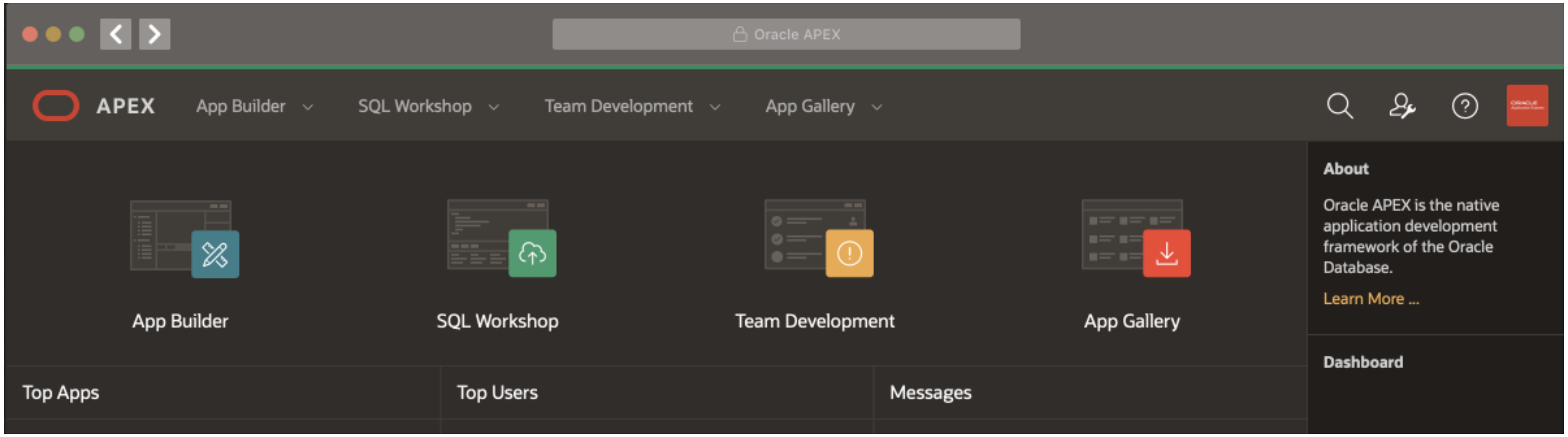
Course examples

1. Course and lab examples uses
   - Oracle XE 11g, APEX
   - MySQL 8.x
   - Visual dBase Plus v.2.7

2. Programming and Querying Languages
   - SQL, xBase
   - Java, Javascript, PHP

Oracle APEX

1. Oracle free solution for DB Web application development. Has four components
   - Application Builder
   - SQL Workshop
   - Team Development
   - Administration



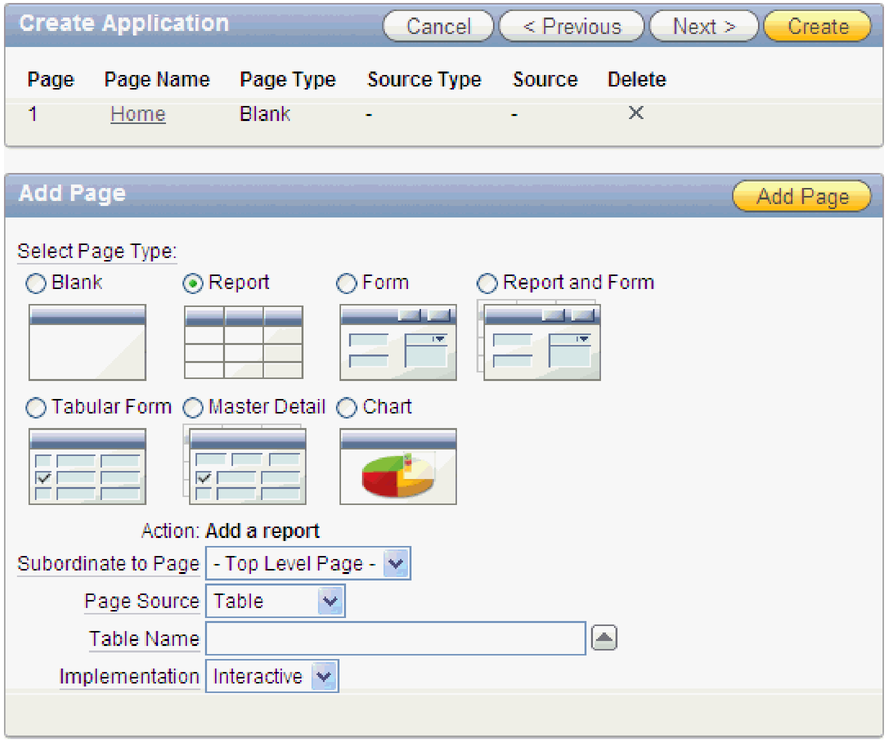
About

Oracle APEX is the native application development framework of the Oracle Database.

Learn More ...

APEX Pages

1. Main types: forms and reports



Page	Page Name	Page Type	Source Type	Source	Delete
1	Home	Blank	-	-	✕

Select Page Type:

○ Blank ● Report ○ Form ○ Report and Form

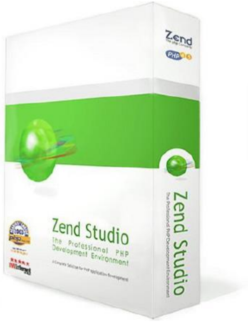
○ Tabular Form ○ Master Detail ○ Chart

Action: Add a report

Subordinate to Page	- Top Level Page -
Page Source	Table
Table Name	
Implementation	Interactive

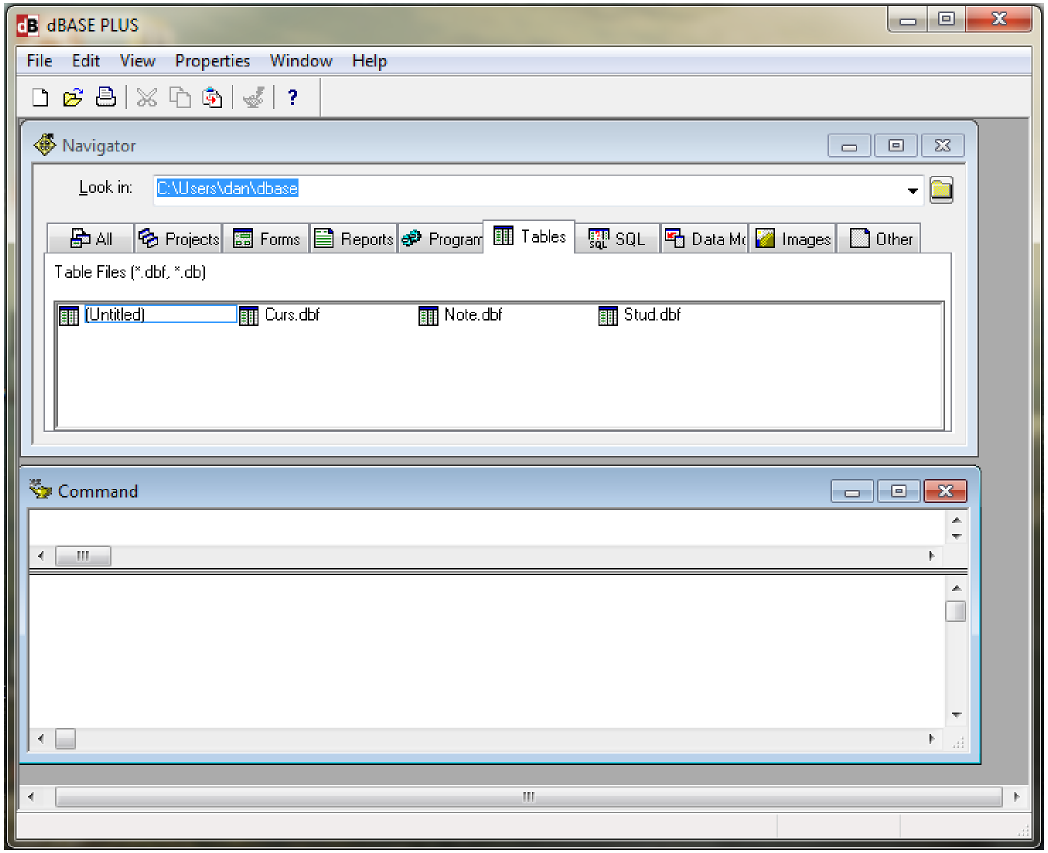
MySQL+PHP

1. MySQL Comunity Server – free at http://dev.mysql.com/
2. Web server – e.g. Apache at http://httpd.apache.org/
3. PHP module – free at http://www.php.net/
4. Alternative: the AMP (Apache, MySQL,PHP) solution stacks
   - WAMP (Windows), LAMP (Linux), MAMP (Mac)
   - XAMPP (Cross-platform)
5. Any text or HTML editor (or more complex programming environments like Eclipse or ZendStudio)



dBase Plus 2.7

1. RAD Environment
2. Navigator – for project management
3. Command – xBase Interpreter interface (commands + results)


dBASE PLUS
File
Edit
View
Properties
Window
Help
Navigator
Look in:
C:\Users\dan\dbase
All
Projects
Forms
Reports
Program
Tables
SQL
Data Mo
Images
Other
Table Files (*.dbf, *.db)
(Untitled)
Curs.dbf
Note.dbf
Stud.dbf
Command