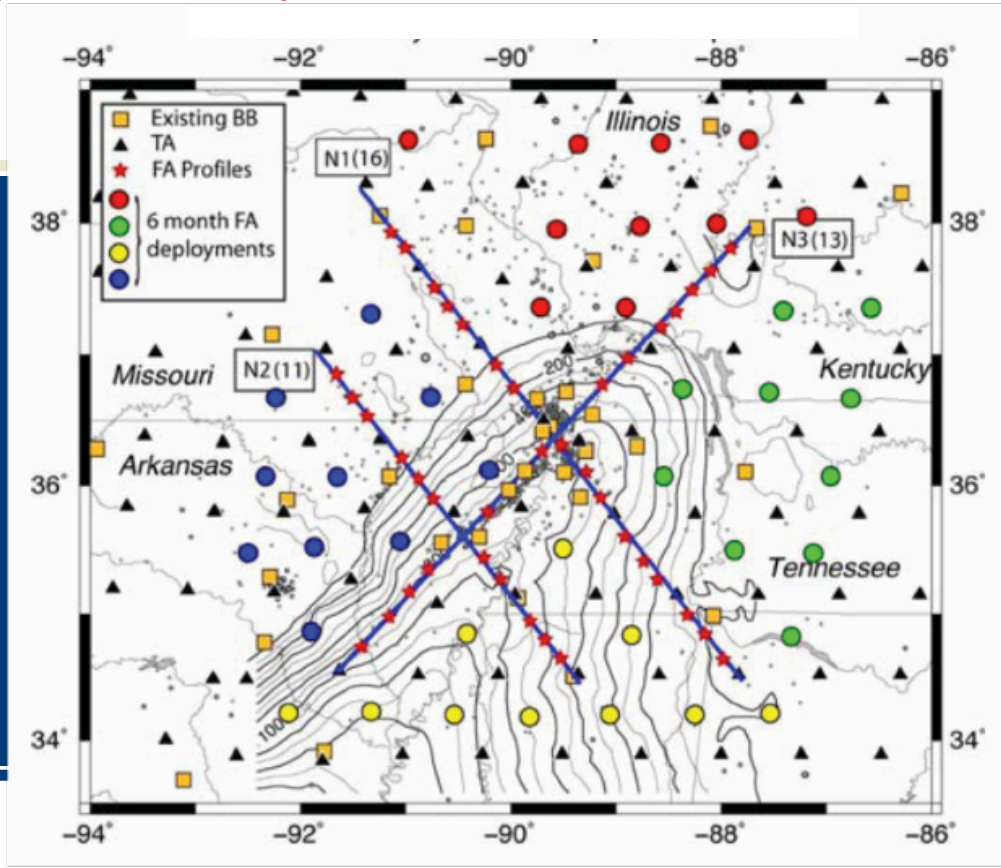
Relational Database Management Systems

A practical introduction using Antelope, the Northern Lithosphere Embayment Experiment (NELE), & the Cooperative New Madrid Seismic Network (CNMSN)



Heather DeShon
Southern Methodist University

IRIS USArray Short-course 2016
Northwestern University
August 3, 2016

NELE

- Earthscope Flexible Array experiment conducted between 2013-2017
- Major Goals Addressed:
  - the primary differences in lithospheric structure between the embayment and the surrounding region
  - the nature of early Cambrian rifting and relationship to pre-existing structure
  - the dynamic processes responsible for basin subsidence, and
  - the relationship of the NMSZ to large scale lithospheric structure.
- Collaborative effort between the CERI @ the University of Memphis (C. Langston, C. Powell, S. Horton, H. DeShon), Penn State (C. Ammon), St. Louis University (B. Hermann)
- More information: http://www.memphis.edu/ceri/nele/abstract.php

Why do seismologists need to know about relational databases (RDB)?

- Increasing complexity of seismology data acquisition requires efficient means to organize and access data
- Only want to type information once and keep it in a central location! A piece of information is only found in one place
- Standard RDB tools to create, edit, combine, and destroy tables already exist
- Entire discipline of study in computer science

How do seismologists use RDBs?

- Processing, analysis and storage of data and associated metadata
- RDBs are at the heart of every seismic/infrasound network operation
- Most RDBs are embedded in combinations of commercial and open source software packages

AQMS

Earthworm

*ITSI

Antelope

Datascope

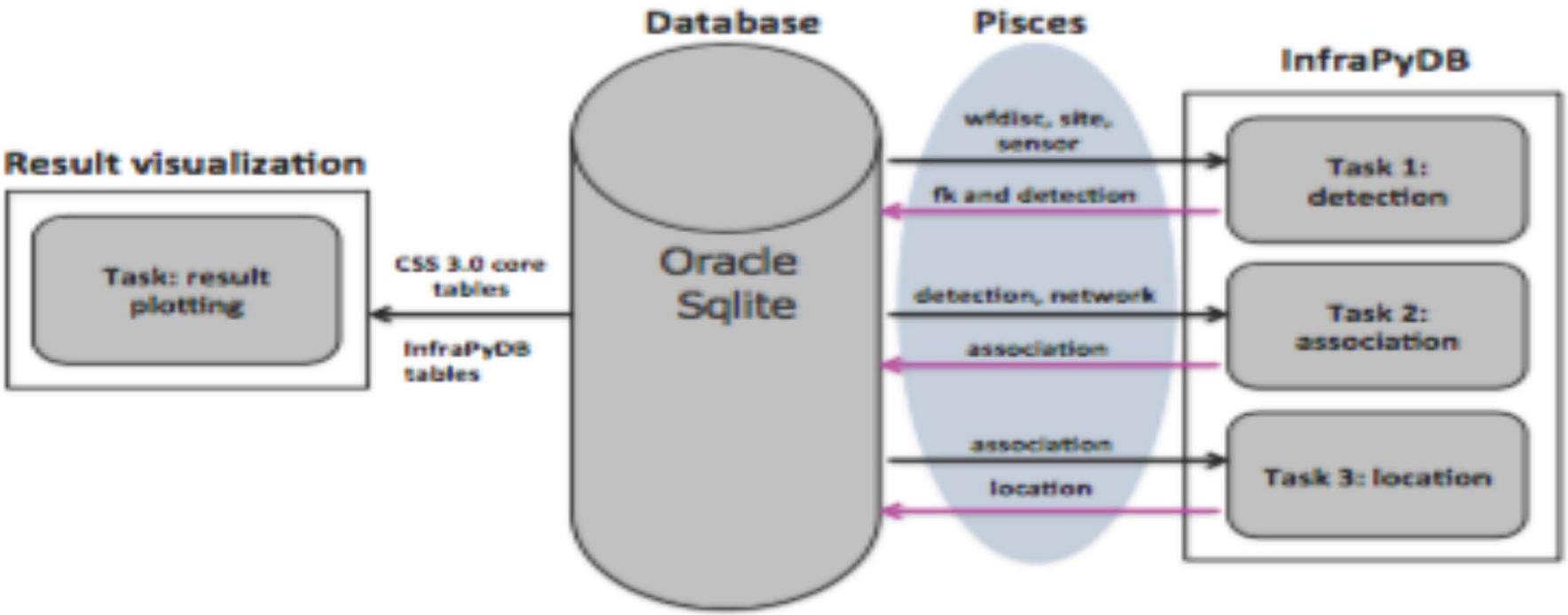
*Boulder RealTime Technologies/ Kinemetrics

SeisComP3

MySQL, SQLite, postgresSQL

*gempa GmbH

The Wrappers



Example from Blom et al. of a set of Python modules wrapped around an Oracle SQLite RDB to analyze infrasound data

What is Antelope?

- Software platform designed to acquire, distribute, archive, and process environmental data

- Provides a standard substrate (middleware) for interconnecting diverse data sources with diverse data processing
  - Commercial and public-domain components
  - Closed and open source components
  - Can handle ANY packetized data
  - Encapsulates diverse specific data formats and uses standard protocols for communication

more about Antelope

- Built upon the Datascope Relational Database
- Handles real-time and offline batch mode data
- Easy integration of user developed modules
- Property of Boulder Real Time Technologies, Inc (BRTT)

What is Datascope?

- Relational database component ➔ ASCII fixed format tables which are related or linked to each other so that every record (row) of the table is unique
- Originally developed at UC-Boulder and distributed through IRIS for use with PASSCAL deployments
- Academic usage was primarily offline batch mode processing for creating seed volumes, phase picking, locations, etc.
- Used to create DMC SEED volumes for archiving

- Seismology-specific processing functions using derivatives of the CSS3.0 schema
  - Data and metadata archiving (waveforms, site and instr char, etc.)
  - Information archiving (picks, hypocenters, etc.)
  - Import/export (SEED, SAC, SEGY, etc.)
  - “routine” network processing (detection, association, location, etc.)
  - Ground motion estimation (similar to Shakemap)
  - Interactive analysis review
  - Archive management functions (backup, restore, copy, excerpt, etc.)
  - Interface with other research tools (SAC, MATLAB, etc.)

- Seismology-specific processing functions using derivatives of the CSS3.0 schema

  - Data and metadata archiving (waveforms, site and instr char, etc.)
  - Information archiving (picks, hypocenters, etc.)
  - Import/export (SEED, SAC, SEGY, etc.)
  - “routine” network processing (detection, association, location, etc.)
  - Ground motion estimation (similar to Shakemap)
  - Interactive analysis review
  - Archive management functions (backup, restore, copy, excerpt, etc.)
  - Interface with other research tools (SAC, MATLAB, etc.)

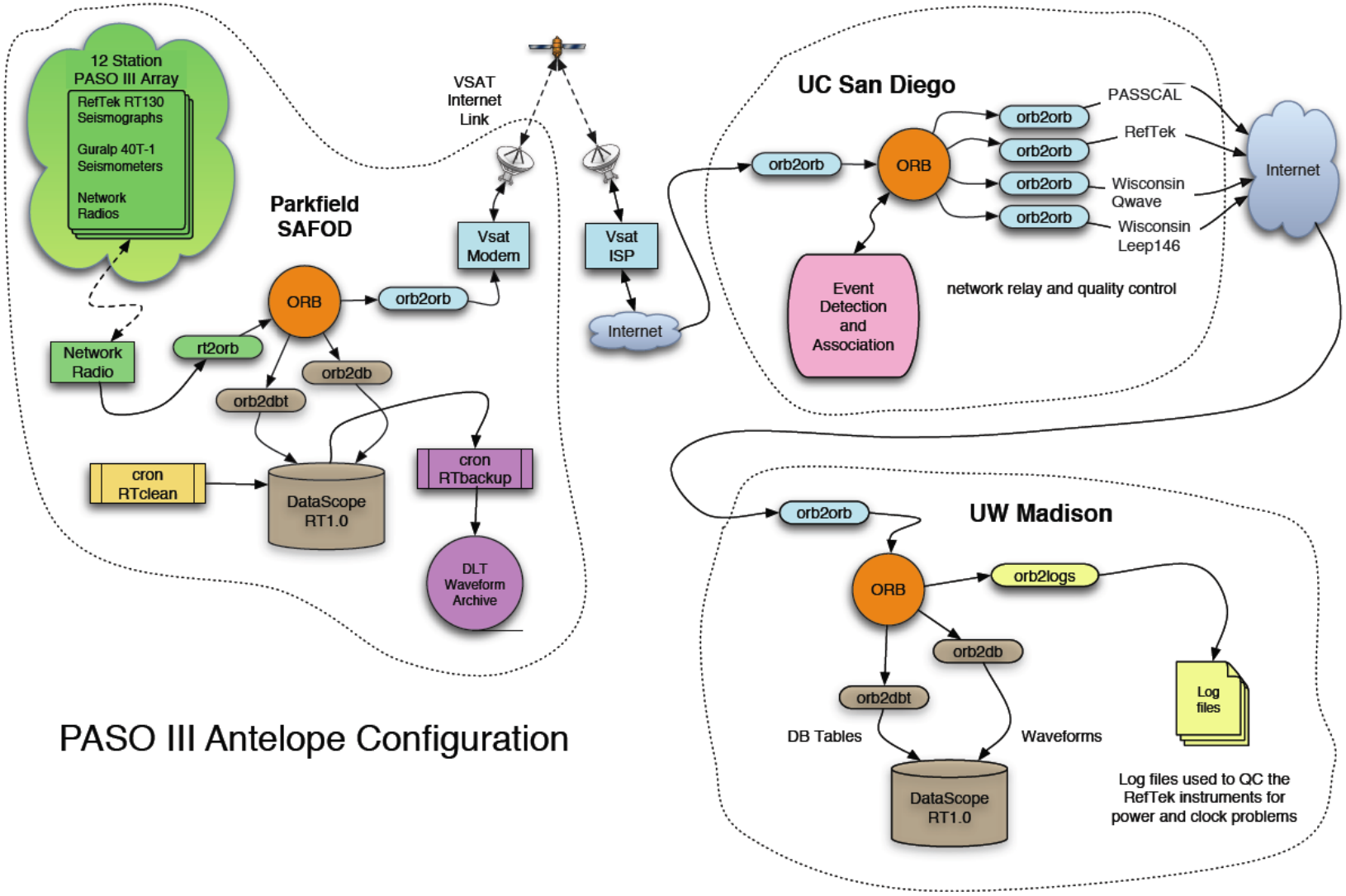
Why use BRTTs Antelope?

- It was originally designed to deal with seismic data
- Contains numerous tools for processing and analyzing seismic data of many formats
- Flexible enough so that users can create their own schema or integrate home-grown location algorithms, magnitude calculators, etc.
- It's FREE for research use at US educational institutions

Real-time or ORB system

- The real time system is built around a large, flexible, non-volatile ring buffer.

- Ring Buffer: a method of using memory within a computer program. “Ring” alludes to the rotation through the buffer of the positions where the next data will be read and written. When moving through the buffer, the writer moves forward one step each time it writes, and when it passes the end of the buffer it starts again at the beginning. The reader moves through the buffer in the same way.



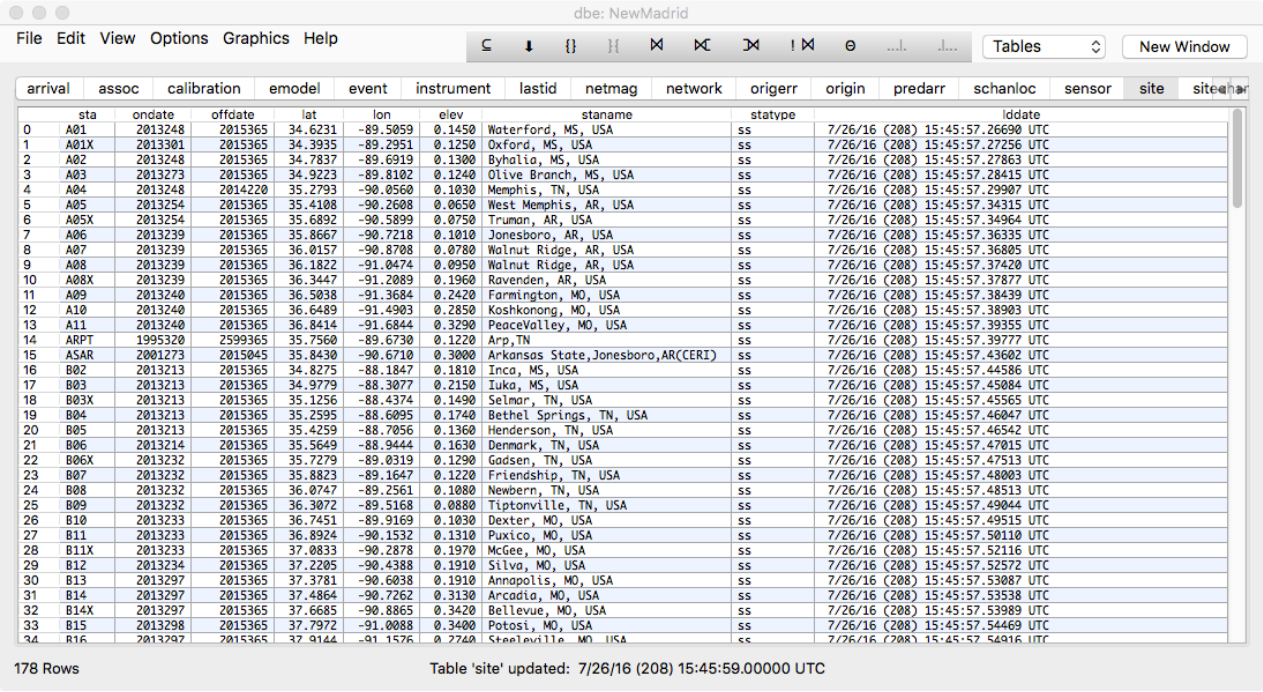
PASO III Antelope Configuration

Offline processing or the DB system

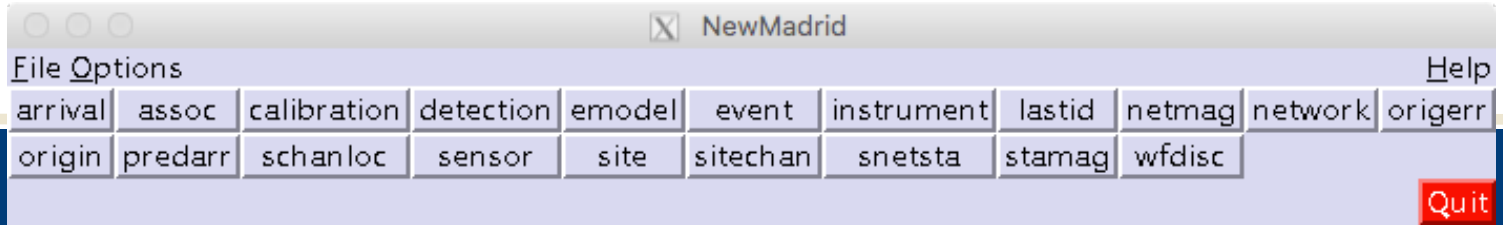
- Offline batch mode processing built upon algorithms originally developed with Datascope
- The common way in which data for temporary or non-telemetered PASSCAL arrays are processed and seed files for DMC archiving are created
- DB: acronym for database which marks algorithms that work with offline data (ie. dbassoc)

dbe - The database editor

- New version with Antelope 5.6

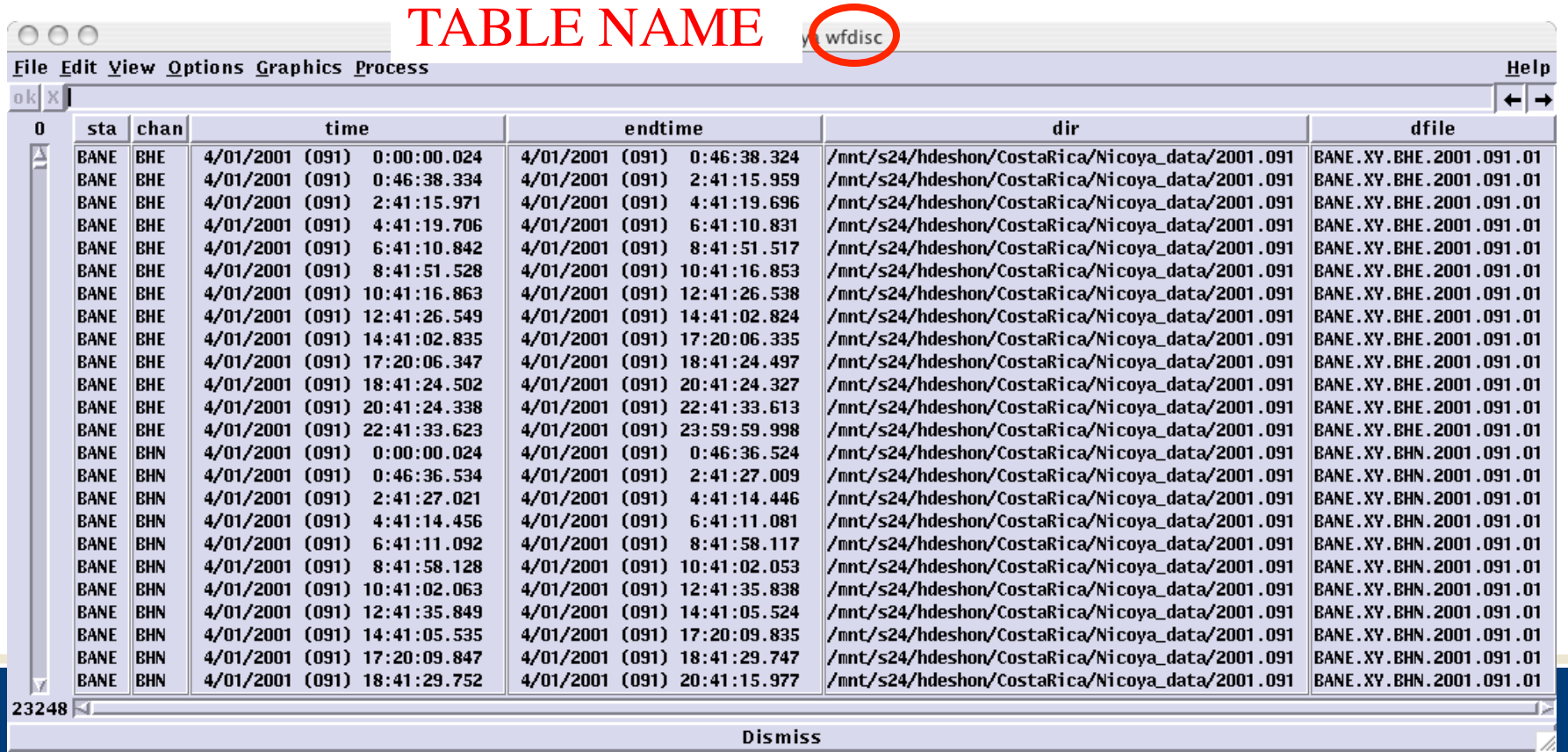


- Slides are going to show the old version, dbe_dep



Relational databases – Some definitions

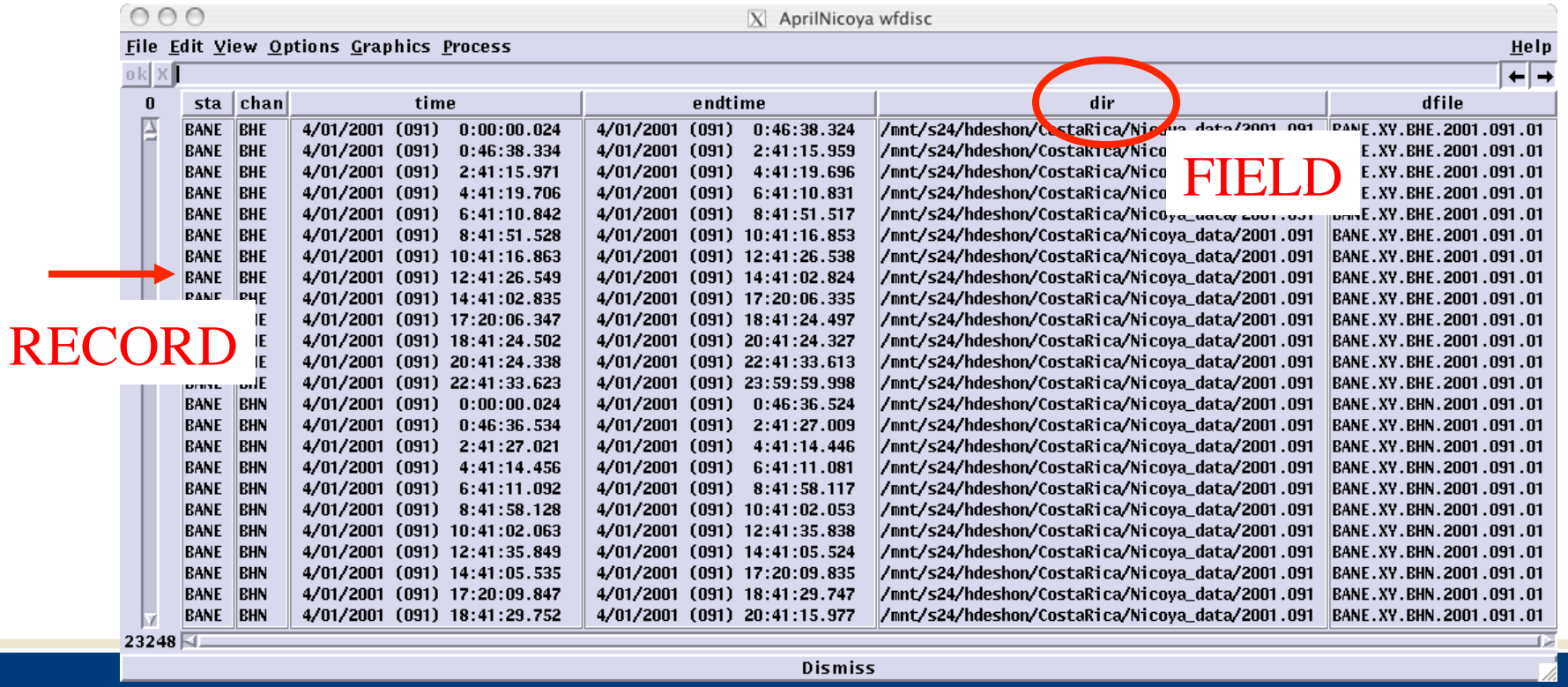
- Table: fix format ASCII text files made up of records (rows) and fields (columns)



sta	chan	time	endtime	dir	dfile
BANE	BHE	4/01/2001 (091) 0:00:00.024	4/01/2001 (091) 0:46:38.324	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHE	4/01/2001 (091) 0:46:38.334	4/01/2001 (091) 2:41:15.959	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHE	4/01/2001 (091) 2:41:15.971	4/01/2001 (091) 4:41:19.696	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHE	4/01/2001 (091) 4:41:19.706	4/01/2001 (091) 6:41:10.831	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHE	4/01/2001 (091) 6:41:10.842	4/01/2001 (091) 8:41:51.517	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHE	4/01/2001 (091) 8:41:51.528	4/01/2001 (091) 10:41:16.853	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHE	4/01/2001 (091) 10:41:16.863	4/01/2001 (091) 12:41:26.538	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHE	4/01/2001 (091) 12:41:26.549	4/01/2001 (091) 14:41:02.824	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHE	4/01/2001 (091) 14:41:02.835	4/01/2001 (091) 17:20:06.335	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHE	4/01/2001 (091) 17:20:06.347	4/01/2001 (091) 18:41:24.497	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHE	4/01/2001 (091) 18:41:24.502	4/01/2001 (091) 20:41:24.327	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHE	4/01/2001 (091) 20:41:24.338	4/01/2001 (091) 22:41:33.613	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHE	4/01/2001 (091) 22:41:33.623	4/01/2001 (091) 23:59:59.998	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
BANE	BHN	4/01/2001 (091) 0:00:00.024	4/01/2001 (091) 0:46:36.524	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
BANE	BHN	4/01/2001 (091) 0:46:36.534	4/01/2001 (091) 2:41:27.009	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
BANE	BHN	4/01/2001 (091) 2:41:27.021	4/01/2001 (091) 4:41:14.446	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
BANE	BHN	4/01/2001 (091) 4:41:14.456	4/01/2001 (091) 6:41:11.081	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
BANE	BHN	4/01/2001 (091) 6:41:11.092	4/01/2001 (091) 8:41:58.117	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
BANE	BHN	4/01/2001 (091) 8:41:58.128	4/01/2001 (091) 10:41:02.053	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
BANE	BHN	4/01/2001 (091) 10:41:02.063	4/01/2001 (091) 12:41:35.838	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
BANE	BHN	4/01/2001 (091) 12:41:35.849	4/01/2001 (091) 14:41:05.524	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
BANE	BHN	4/01/2001 (091) 14:41:05.535	4/01/2001 (091) 17:20:09.835	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
BANE	BHN	4/01/2001 (091) 17:20:09.847	4/01/2001 (091) 18:41:29.747	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
BANE	BHN	4/01/2001 (091) 18:41:29.752	4/01/2001 (091) 20:41:15.977	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01

- Field: has the same attributes (size, type, format) in every table in which it appears

- Record: unique combination of information



0	sta	chan	time	endtime	dir	dfile
	BANE	BHE	4/01/2001 (091) 0:00:00.024	4/01/2001 (091) 0:46:38.324	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHE	4/01/2001 (091) 0:46:38.334	4/01/2001 (091) 2:41:15.959	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHE	4/01/2001 (091) 2:41:15.971	4/01/2001 (091) 4:41:19.696	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHE	4/01/2001 (091) 4:41:19.706	4/01/2001 (091) 6:41:10.831	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHE	4/01/2001 (091) 6:41:10.842	4/01/2001 (091) 8:41:51.517	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHE	4/01/2001 (091) 8:41:51.528	4/01/2001 (091) 10:41:16.853	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHE	4/01/2001 (091) 10:41:16.863	4/01/2001 (091) 12:41:26.538	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHE	4/01/2001 (091) 12:41:26.549	4/01/2001 (091) 14:41:02.824	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHE	4/01/2001 (091) 14:41:02.835	4/01/2001 (091) 17:20:06.335	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHE	4/01/2001 (091) 17:20:06.347	4/01/2001 (091) 18:41:24.497	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHE	4/01/2001 (091) 18:41:24.502	4/01/2001 (091) 20:41:24.327	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHE	4/01/2001 (091) 20:41:24.338	4/01/2001 (091) 22:41:33.613	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHE	4/01/2001 (091) 22:41:33.623	4/01/2001 (091) 23:59:59.998	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHE.2001.091.01
	BANE	BHN	4/01/2001 (091) 0:00:00.024	4/01/2001 (091) 0:46:36.524	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
	BANE	BHN	4/01/2001 (091) 0:46:36.534	4/01/2001 (091) 2:41:27.009	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
	BANE	BHN	4/01/2001 (091) 2:41:27.021	4/01/2001 (091) 4:41:14.446	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
	BANE	BHN	4/01/2001 (091) 4:41:14.456	4/01/2001 (091) 6:41:11.081	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
	BANE	BHN	4/01/2001 (091) 6:41:11.092	4/01/2001 (091) 8:41:58.117	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
	BANE	BHN	4/01/2001 (091) 8:41:58.128	4/01/2001 (091) 10:41:02.053	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
	BANE	BHN	4/01/2001 (091) 10:41:02.063	4/01/2001 (091) 12:41:35.838	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
	BANE	BHN	4/01/2001 (091) 12:41:35.849	4/01/2001 (091) 14:41:05.524	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
	BANE	BHN	4/01/2001 (091) 14:41:05.535	4/01/2001 (091) 17:20:09.835	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
	BANE	BHN	4/01/2001 (091) 17:20:09.847	4/01/2001 (091) 18:41:29.747	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01
	BANE	BHN	4/01/2001 (091) 18:41:29.752	4/01/2001 (091) 20:41:15.977	/mnt/s24/hdeshon/CostaRica/Nicoya_data/2001.091	BANE.XY.BHN.2001.091.01

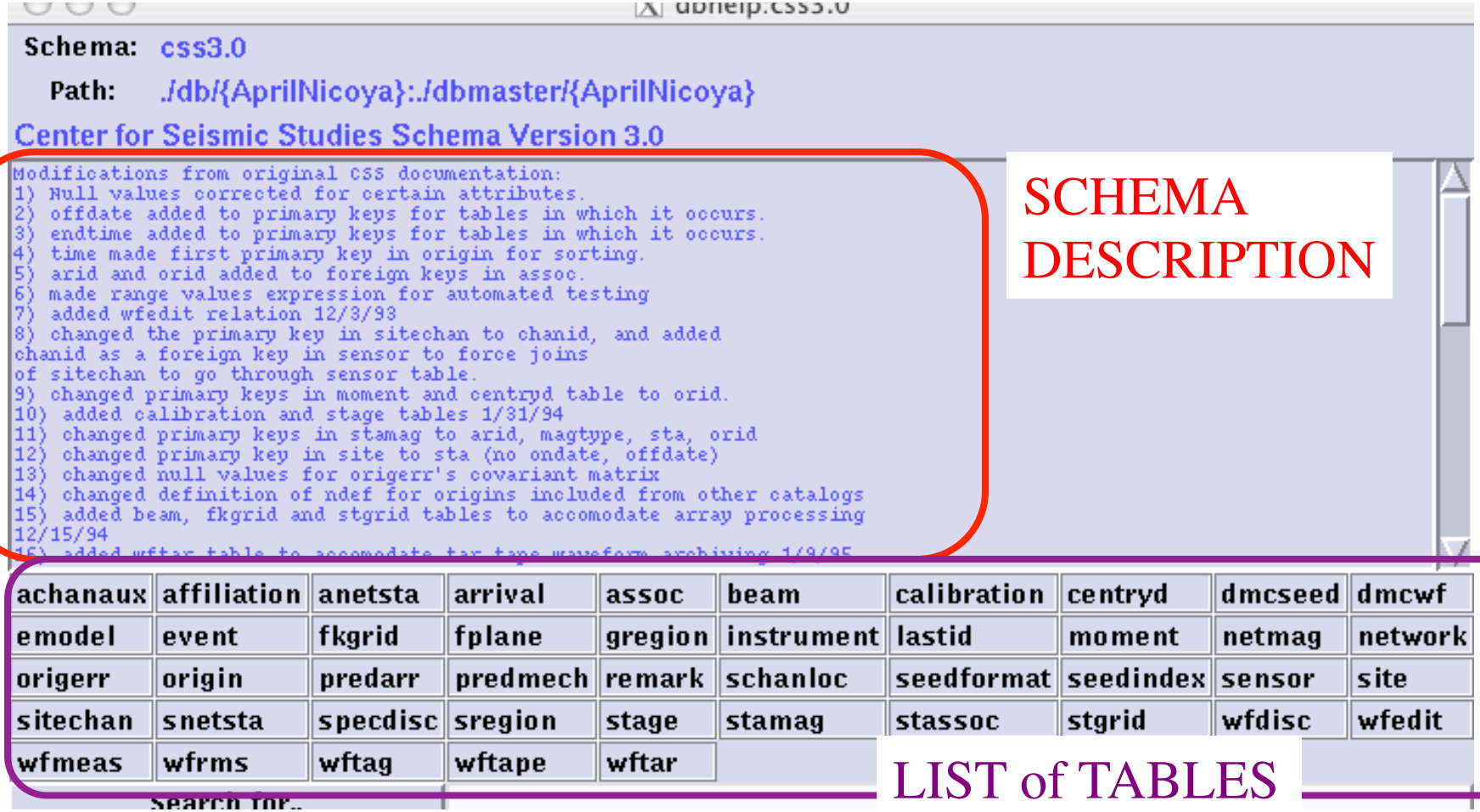
23248

Schema

- The attributes and relations that define the relational database

- Defined by single text file that contains
  - schema description
  - attribute statements for every field
  - relation statements for every table which describe how the tables are linked/related

Example Schema

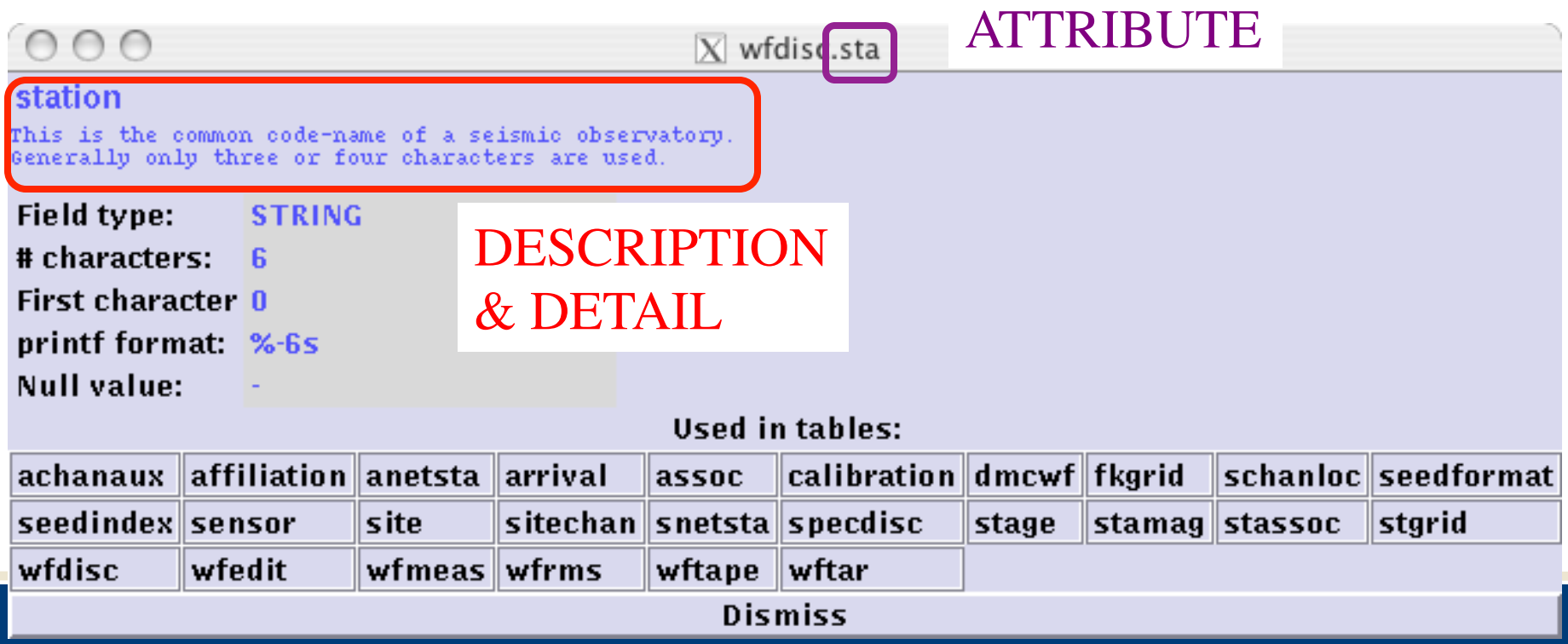


/opt/antelope/4.9/data/schemas

```
2%Schema css3.0
        Description ( "Center for Seismic Studies Schema Version 3.0 " )
        Detail {
                Modifications from original CSS documentation:
                0) units of calib vary according to the instrument, with wfdisc.segtype
                   and instrument.rsptype indicating both sensor type and units
                1) Null values corrected for certain attributes.
                2) offdate added to primary keys for tables in which it occurs.
                3) endtime added to primary keys for tables in which it occurs.
                4) time made first primary key in origin for sorting.
                5) arid and orid added to foreign keys in assoc.
                6) made range values expression for automated testing
                7) added wfedit relation 12/3/93
                8) changed the primary key in sitechan to chanid, and added
                chanid as a foreign key in sensor to force joins
                of sitechan to go through sensor table.
                9) changed primary keys in moment and centryd table to orid.
                10) added calibration and stage tables 1/31/94
                11) changed primary keys in stamag to arid, magtype, sta, orid
                12) changed primary key in site to sta (no ondate, offdate)
                13) changed null values for origerr's covariant matrix
                14) changed definition of ndef for origins included from other catalogs
                15) added beam, fkgrid and stgrid tables to accomodate array processing
                12/15/94
                16) added wftar table to accomodate tar tape waveform archiving 1/9/95
                17) changed all NONULL null values to reasonable values
                18) added wfrms table
                19) added wfmeas table for holding generic waveform measurements
                20) segment origin and stassoc etype field into two
                fields, etype and review, so that analyst review
                status can be kept in origin table
                21) add snetsta, anetsta, schanloc and achanaux tables
                to translate between foreign volumes of SEED or
                autoDRM into local databases.
                22) add specdisc table and associated attributes to support
                spectral estimation processing
                23) add rsprm to specdisc table
                24) added tables dmcseed and dmcwf to database to
                support all the DMC requirements for building DMC seed
                volumes.  Changed default value of fileno to -1.
                25) changed format of chksum from Integer %15d to Real %12.0f
                to make sure table is written properly.  Previous format would
                wrap to negative numbers in some cases which would corrupt the
                database.
                26) added fields calib, calper, samprate, timever to table dmcwf and
                field totbytes to table dmcseed.
                27) changed primary keys of stamag from arid magtype sta and orid
                to magid magtype sta and orid.
                28) changed fields dmcseedfile to dfile and jdate to yearday in tables dmcwf and
                dmcseed.
                29) added focal mechanism calculation related tables and emodel table
                30) extensive changes to wfrms table to make compatible with orbwfrms
                40) made instrument table alternate key all fields except inid and lddate
                41) introduced new calibration table, with new fields samprate, segtype, dlsta,
                dlchan, and lead.
        }
        Timedate lddate
2%
```

Attribute Statement

- Type
- Format
- Null value
- Units
- Range

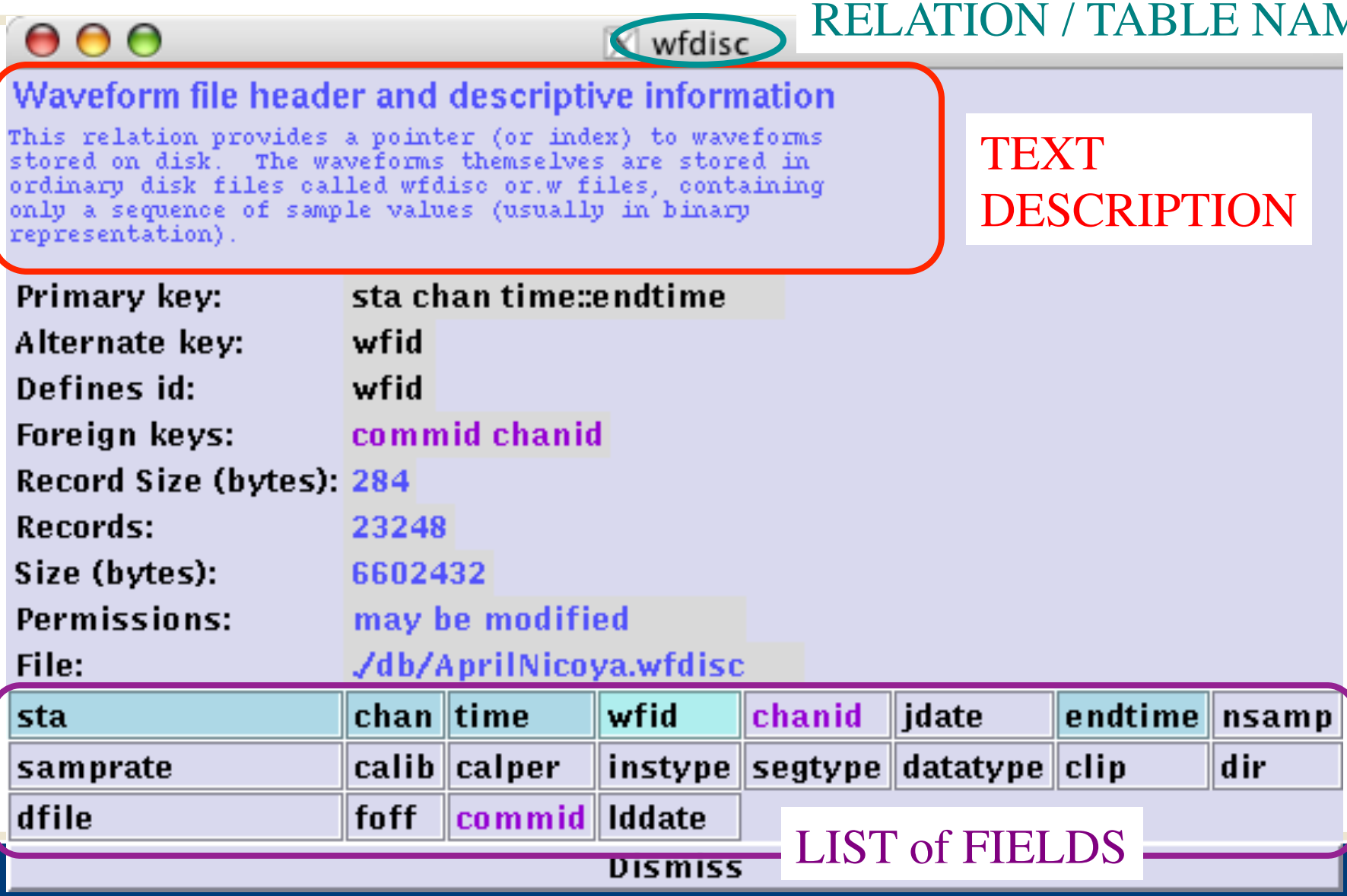


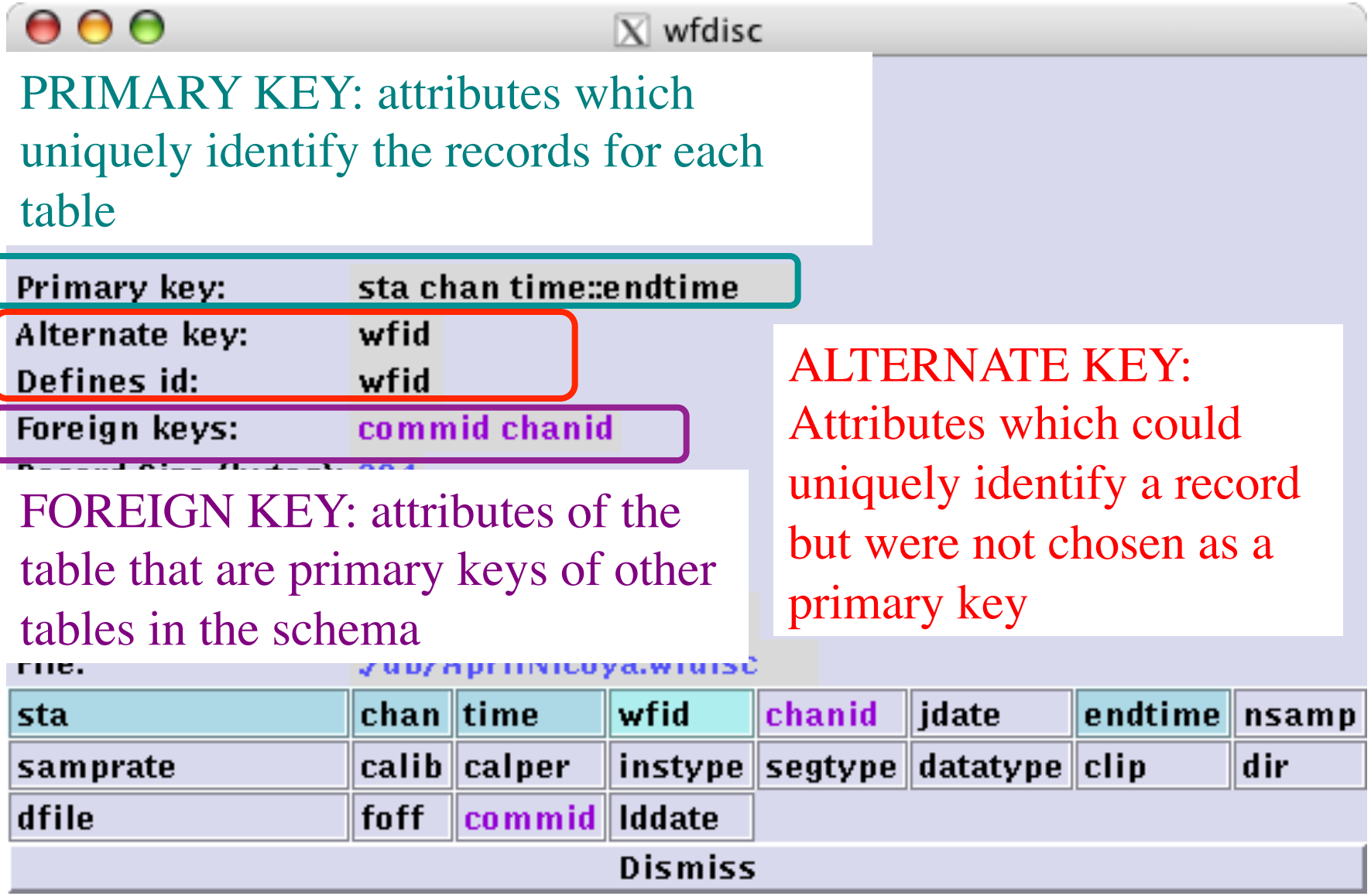
/opt/antelope/4.9/data/schemas

```
                revision level, or other identifier.
        }
        ;

Attribute sta
        String (6)
        Format ( "%-6s" )
        Null ( "-" )
        Description ( "station" )
        Detail {
                This is the common code-name of a seismic observatory.
                Generally only three or four characters are used.
        }
        ;

Attribute stageid
        Integer (8)
        Format ( "%8d" )
        Null ( "-1" )
        Range ( " 0 < stageid " )
        Description ( "stage number in the calibration response" )
        Detail {
                The ordered stage number of this discrete stage in the
                calibration response.  Each individual stage corresponds
                to a sensor, analog filter, A/D converter, or FIR filter.
                The numbering scheme for a seismic system will generally
                assign stageid=1 for the sensor, stageid=2 for the
                anti-alias filter, stageid=3 for the analog-to-digital
                converter, stageid=4 for the first FIR filter, ...
        }
        ;

Attribute staname
        String (50)
        Format ( "%-50s" )
        Null ( "-" )
        Description ( "station description" )
        Detail {
                This is the full name of the station whose code-name is in
                sta.  As an example, one record in the site relation
                connects sta = ANMO to staname = ALBUQUERQUE, NEW MEXICO
                (SRO).
        }
        ;

Attribute stassid
        Integer (8)
        Format ( "%8d" )
        Null ( "-1" )
        Range ( "stassid > 0" )
        Description ( "stassoc id" )
        Detail {
                The wavetrain from a single event may be made up of a
                number of arrivals.  A unique stassid joins those arrivals
                believed to have come from a common event as measured at a
                single station.  Stassid is also the key to the stassoc
                relation, which contains additional signal measurements
                not contained within the arrival relation, such as station
```

sta ATTRIBUTE

Relation Statement




wfdisc
PRIMARY KEY: attributes which uniquely identify the records for each table
Primary key: sta chan time::endtime
Alternate key: wfid
Defines id: wfid
Foreign keys: commid chanid
FOREIGN KEY: attributes of the table that are primary keys of other tables in the schema
ALTERNATE KEY: Attributes which could uniquely identify a record but were not chosen as a primary key
sta
chan
time
wfid
chanid
jdate
endtime
nsamp
samprate
calib
calper
instype
segtype
datatype
clip
dir
dfile
foff
commid
lddate
Dismiss

/opt/antelope/4.9/data/schemas

```
98%Relation wfdisc
        Fields ( sta chan time wfid chanid jdate endtime nsamp samprate calib calper instype segtype datatype clip dir dfile foff commid ldda
        Primary ( sta chan time::endtime )
        Alternate ( wfid )
        Foreign ( commid chanid )
        Defines wfid
        Description ( "Waveform file header and descriptive information" )
        Detail {
                This relation provides a pointer (or index) to waveforms
                stored on disk.  The waveforms themselves are stored in
                ordinary disk files called wfdisc or.w files, containing
                only a sequence of sample values (usually in binary
                representation).
        }
        ;

Relation wfedit
        Fields ( sta chan edid time endtime probtype edittype auth commid lddate )
        Primary ( sta chan time::endtime )
        Alternate ( edid )
        Foreign ( commid )
        Defines edid
        Description ( "Describes a waveform edit" )
        Detail {
                Specifies a time slice for which a specified waveform has
                a problem, identified by the probtype field.
        }
        ;

Relation wfmeas
        Fields ( sta chan meastype filter time endtime tmeas twin val1 val2 units1 units2 arid auth lddate )
        Primary ( sta chan meastype filter time endtime )
        Foreign ( arid )
        Description ( "waveform measurements" )
        Detail {
                This relation provides a general way to store measurements made
                on segments of waveform data. The time::endtime fields give
                the time window of the data for which the measurement is unique.
                tmeas and twin specify the beginning of the measurement time
                for discrete measurements, and the time-span for extended measurements.
                The contents of val1 and val2, described by units1 and units2, depend
                on the type of measurement made.
        }
        ;

98%
```

wfdisc RELATION

Tables are linked together by the primary, alternate, and foreign keys

How links via keys work

Site

File Edit View Options Graphics Help

0	sta	ondate	offdate	lat	lon	elev	staname
	BANE	1999340	2001365	9.9292	-84.9564	0.0490	Bajo de Negros
	BONG	1999341	2001365	9.7535	-85.2078	0.0260	Rio Bongo
	CORO	1999340	2001365	9.9729	-85.1711	0.0500	Corozal
	COYO	1999340	2001365	10.3614	-85.6493	0.0670	Coyolito
	GRAN	1999340	2001365	10.3404	-85.8458	0.0280	Playa Grande
	GUAI	1999339	2001365	10.2702	-85.5105	0.0490	Guaitil
	GUIO	1999342	2001365	9.9228	-85.6584	0.0280	Playa Guiones
	HATA	1999341	2001365	10.4343	-85.2848	0.0440	Hacienda Tamarindo
	HOJA	2000020	2001365	10.0498	-85.4230	0.3560	Hojancha
	INDI	1999351	2001365	9.8647	-85.5022	0.0800	Punto Indio
	JUDI	2000134	2001365	10.1659	-85.5387	0.7070	Juan Diaz
	MARB	1999341	2001365	10.0630	-85.7544	0.0480	Marbella
	PAPA	2000257	2001365	10.5902	-85.6760	0.0380	Golfo Papagayo (Near Playa de Coco)
	PARG	1999340	2001365	10.1983	-85.8230	0.0200	Punto Pargo
	PNCB	1999339	2001365	9.5895	-85.0917	0.0180	Parque Nacional Cabo Blanco
	PNUE	1999341	2001365	9.8411	-85.3352	0.0210	Pueblo Nuevo
	SAJE	2000091	2001365	10.6086	-85.4500	0.0200	San Jeronimo
	SARO	2000029	2001365	10.8390	-85.6094	0.0300	Santa Rosa
	TFER	1999340	2001365	10.2082	-85.2706	0.0690	Tempisque Ferry
	VAIN	1999339	2001365	9.7768	-85.0102	0.0600	Vainilla

20

Dismiss

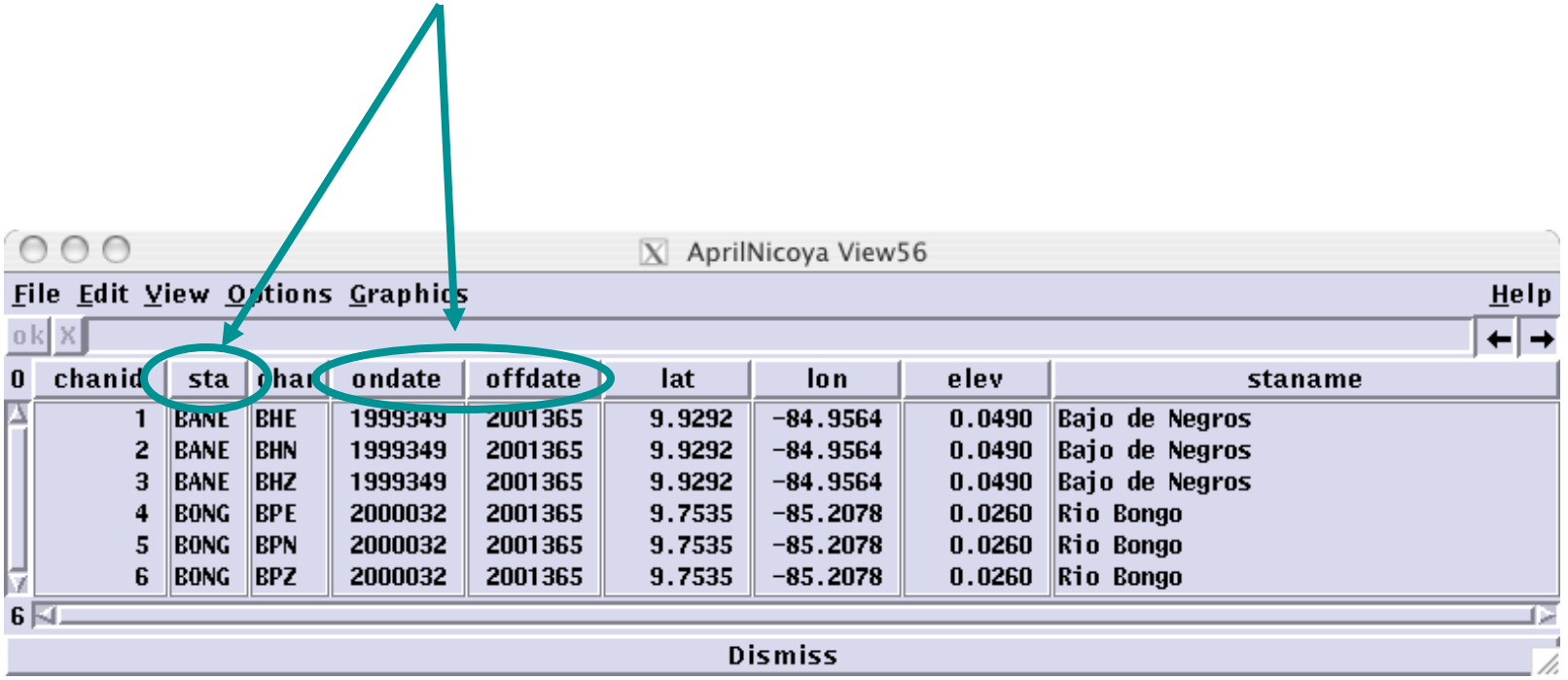
- Site Table Primary Keys
- Sta
- Ondate::Enddate

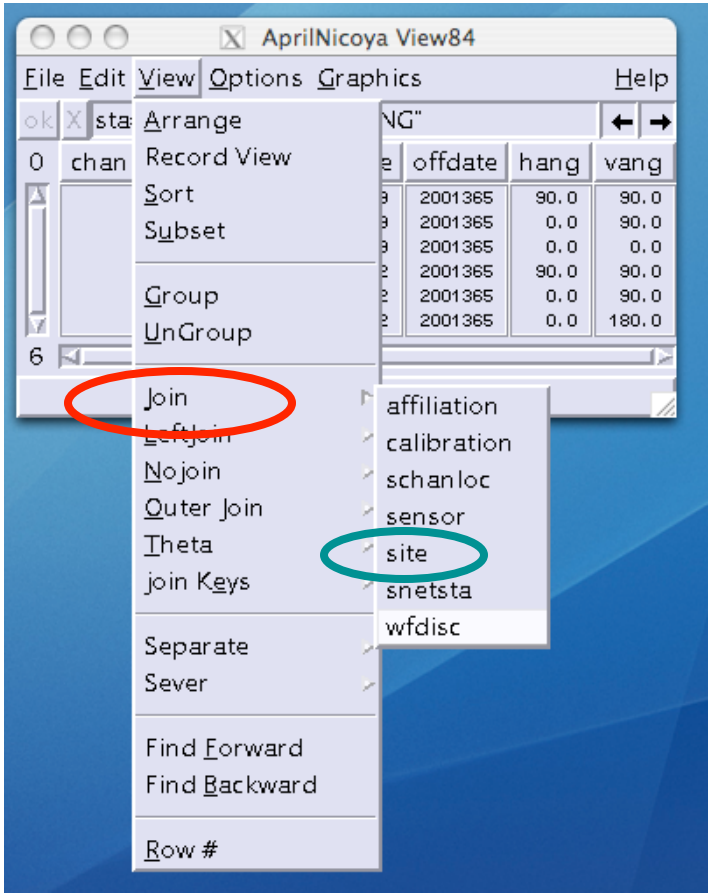
- SiteChan Table Primary Key
  - Chanid
- Alternate Key
  - Sta, Chan
  - Ondate::Enddate

SiteChan Subset

Result of joining the sitechan and site tables

Common primary keys used for the join

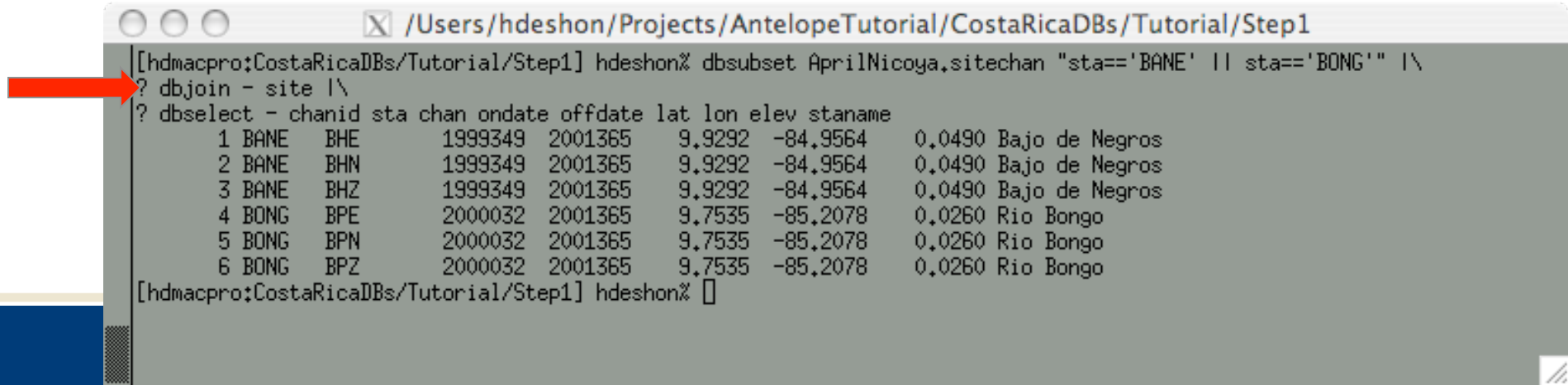




Join using the GUI dbe

or

Join via command line



Relational Algebra

- **Union:** If A and B are sets, then the union of A and B is the set R that contains all elements of A and all elements of B, but no other elements.

- **Difference:** If A and B are sets, then the difference of B and A, is the set R which contains all of elements in B, but not in A.

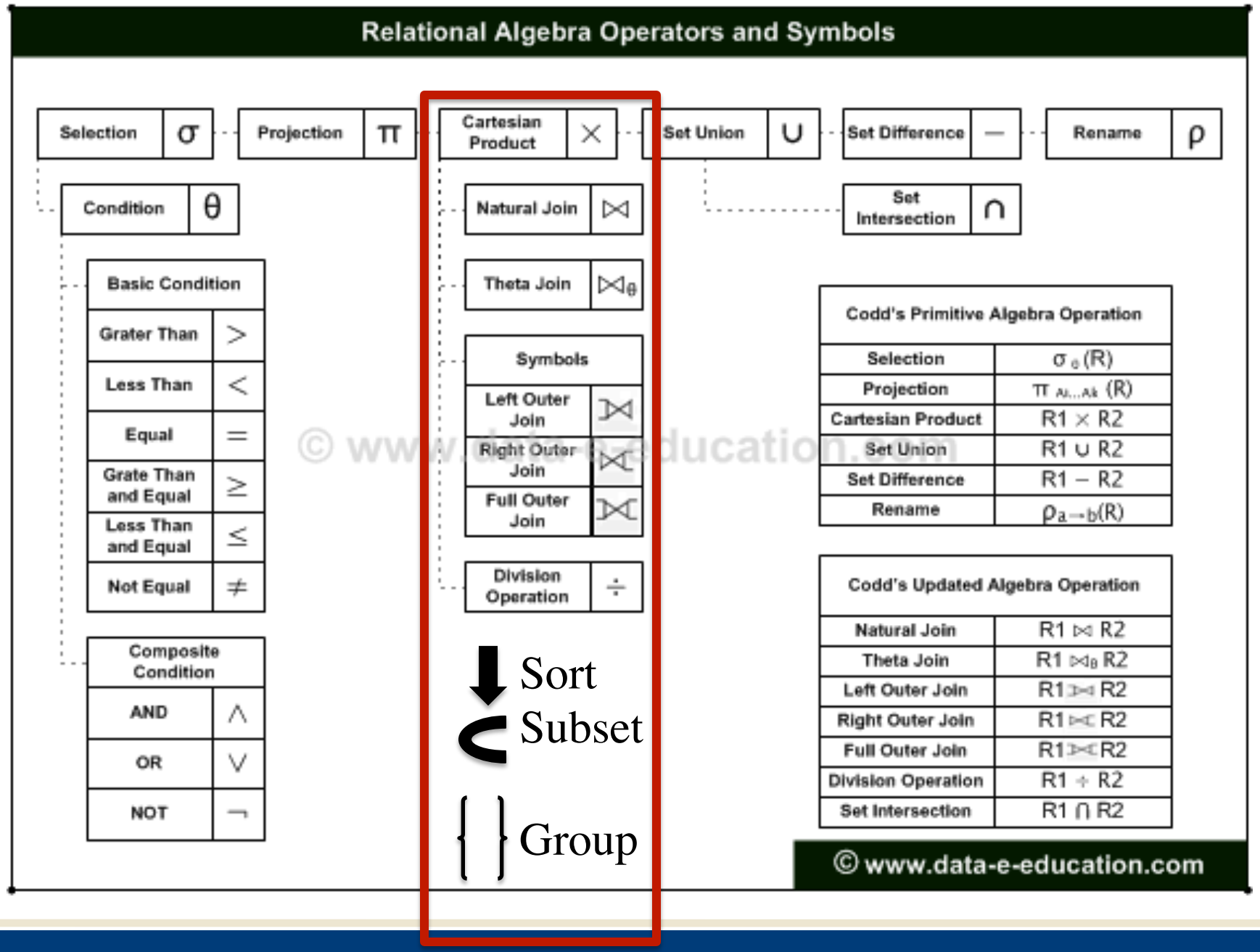
- **Rename:** change the name of a field

Relational Algebra

- **Selection:** selecting records by applying a test using binary operators (<,≤,=≥,>) based on the table fields

- **Projection:** if a1,...,an is a set of field names, the result of projection is defined as the set R that is obtained when all records in R are restricted to the set {a1,...,an}

- **Cartesian Product:** If A and B are sets, the Cartesian product is the set R containing all possible ordered pairs whose first component is a member of A and whose second component is a member of B


Relational Algebra Operators and Symbols
Selection σ
Projection π
Cartesian Product ×
Set Union ∪
Set Difference −
Rename ρ
Condition θ
Natural Join ⋈
Set Intersection ∩
Theta Join ⋈θ
Basic Condition
Grater Than >
Less Than <
Equal =
Grate Than and Equal ≥
Less Than and Equal ≤
Not Equal ≠
Symbols
Left Outer Join
Right Outer Join
Full Outer Join
Division Operation ÷
Composite Condition
AND ∧
OR ∨
NOT ¬
Sort
Subset
Group
Codd's Primitive Algebra Operation
Selection σθ(R)
Projection π Ai,..,Ak (R)
Cartesian Product R1 × R2
Set Union R1 ∪ R2
Set Difference R1 − R2
Rename ρa→b(R)
Codd's Updated Algebra Operation
Natural Join R1 ⋈ R2
Theta Join R1 ⋈θ R2
Left Outer Join R1 ⟕ R2
Right Outer Join R1 ⟖ R2
Full Outer Join R1 ⟗ R2
Division Operation R1 ÷ R2
Set Intersection R1 ∩ R2
© www.data-e-education.com

Relational operations

- Subset

  $dbsubset dbsite ‘lat > 45’

- Sort

  $dbsort dbsite sta

- Join

  $dbjoin db.arrival site

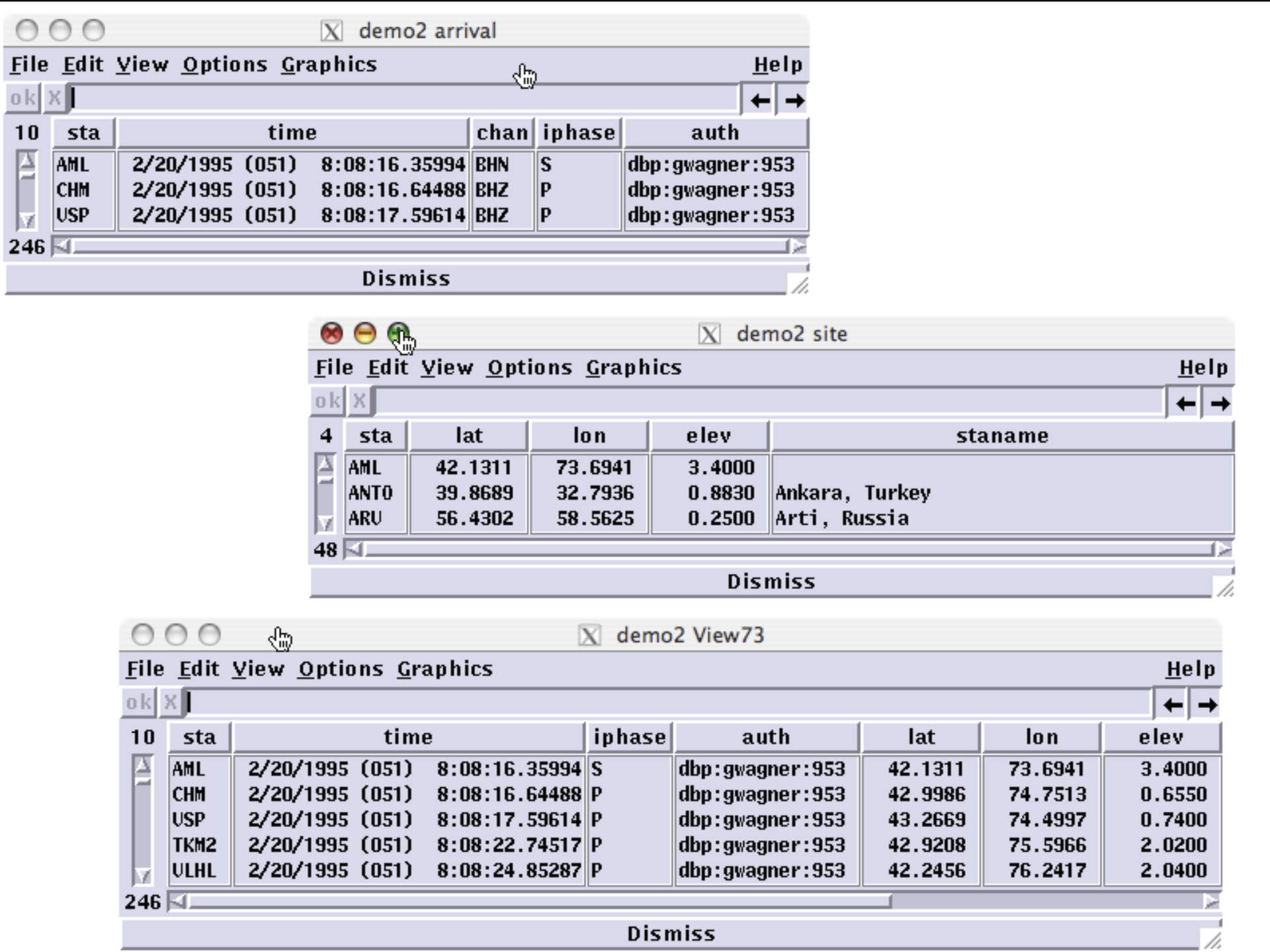
Select/Arrange

%dbselect - sta lat lon elev staname

```
% dbsubset AprilNicoya.site 'lat < 11 && elev > .5' |\
 dbsort - sta |\
 dbjoin - arrival |\
 dbselect - sta lat lon arid arrival.time chan
JUDI     10.1659  -85.5387    30026   986357438.70800 BHZ
JUDI     10.1659  -85.5387    31083   988433330.21200 BHZ
JUDI     10.1659  -85.5387    31093   988479025.88600 BHZ
JUDI     10.1659  -85.5387    31099   988518786.98200 BHZ
JUDI     10.1659  -85.5387    31108   988529215.11200 BHZ
JUDI     10.1659  -85.5387    31114   988558341.28100 BHZ
JUDI     10.1659  -85.5387    31120   988565702.96100 BHZ
JUDI     10.1659  -85.5387    31128   988606800.21600 BHZ
```

What is a join?

- For example, given an arrival pick, you need to know the location of the recording station
- ie, given station name and arrival time, lookup the location of the station in the site table.
- Consider new virtual table which combines arrival and site table; each record has 1 arrival plus matching record from site table


demo2 arrival
File Edit View Options Graphics Help
ok X
10
sta time chan iphase auth
AML 2/20/1995 (051) 8:08:16.35994 BHN S dbp:gwagner:953
CHM 2/20/1995 (051) 8:08:16.64488 BHZ P dbp:gwagner:953
USP 2/20/1995 (051) 8:08:17.59614 BHZ P dbp:gwagner:953
246
Dismiss
demo2 site
File Edit View Options Graphics Help
ok X
4
sta lat lon elev staname
AML 42.1311 73.6941 3.4000
ANTO 39.8689 32.7936 0.8830 Ankara, Turkey
ARU 56.4302 58.5625 0.2500 Arti, Russia
48
Dismiss
demo2 View73
File Edit View Options Graphics Help
ok X
10
sta time iphase auth lat lon elev
AML 2/20/1995 (051) 8:08:16.35994 S dbp:gwagner:953 42.1311 73.6941 3.4000
CHM 2/20/1995 (051) 8:08:16.64488 P dbp:gwagner:953 42.9986 74.7513 0.6550
USP 2/20/1995 (051) 8:08:17.59614 P dbp:gwagner:953 43.2669 74.4997 0.7400
TKM2 2/20/1995 (051) 8:08:22.74517 P dbp:gwagner:953 42.9208 75.5966 2.0200
ULHL 2/20/1995 (051) 8:08:24.85287 P dbp:gwagner:953 42.2456 76.2417 2.0400
246
Dismiss

Relational Operators

< : test for less than

<= : test for less than or equal to

> : test for greater than

>= : test for greater than or equal to

== : test for equal to

~= : test for not equal

Logical Operators

&& : logical AND ; tests that both expressions are true

|| : logical OR ; tests that one or both of the expressions are true

! : logical NOT; tests that expression is true