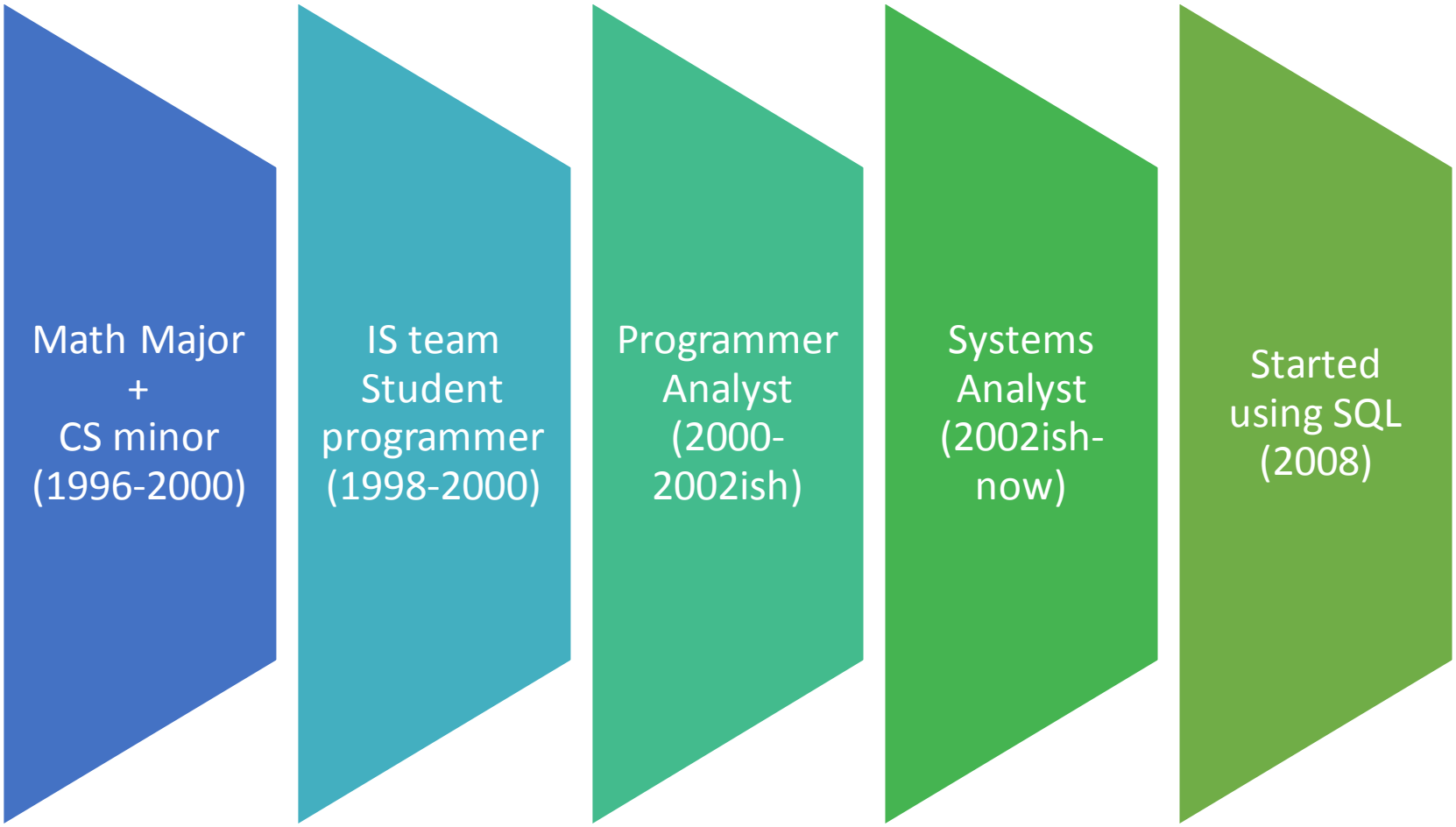
Databases at Calvin College

Monday May 1, 2017

17/SP CS-342

Sarah Greenfield

About me


Math Major
+
CS minor
(1996-2000)
IS team
Student
programmer
(1998-2000)
Programmer
Analyst
(2000-
2002ish)
Systems
Analyst
(2002ish-
now)
Started
using SQL
(2008)

UniData Database

Environment

- Colleague Application runs on a RedHat server “abacus”
- Native DB languages: UniQuery, UniBASIC

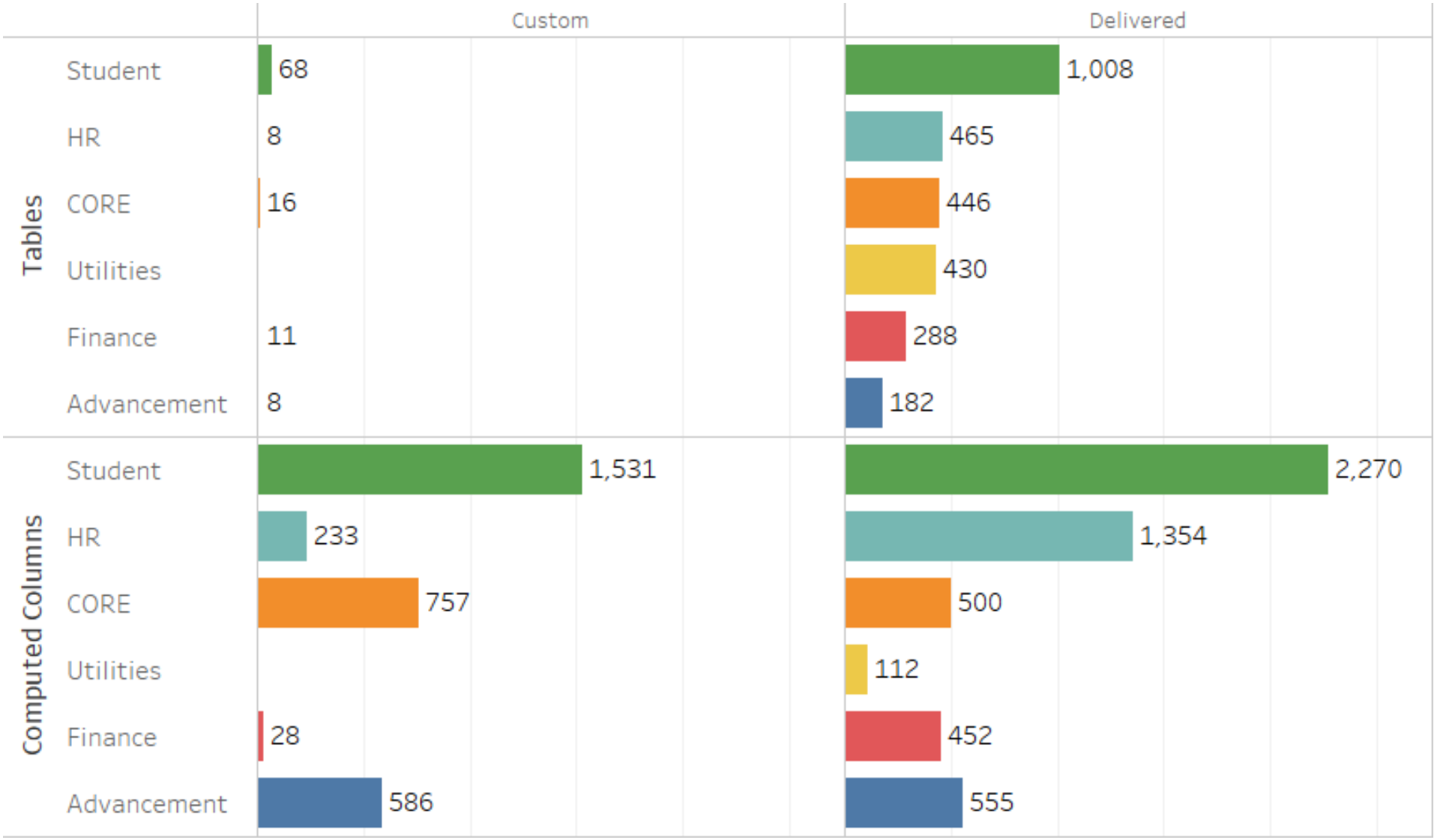
Total size: 91 GB

- Data stored on an Equilogic SAN
- ~32,600 compiled programs to update/report out that data

Runs the administrative business of the college

- Student, HR, Finance, Financial Aid, Registrar, Advancement
- Database vendor: Rocket Software
- Application Vendor: Ellucian

What's in Colleague?

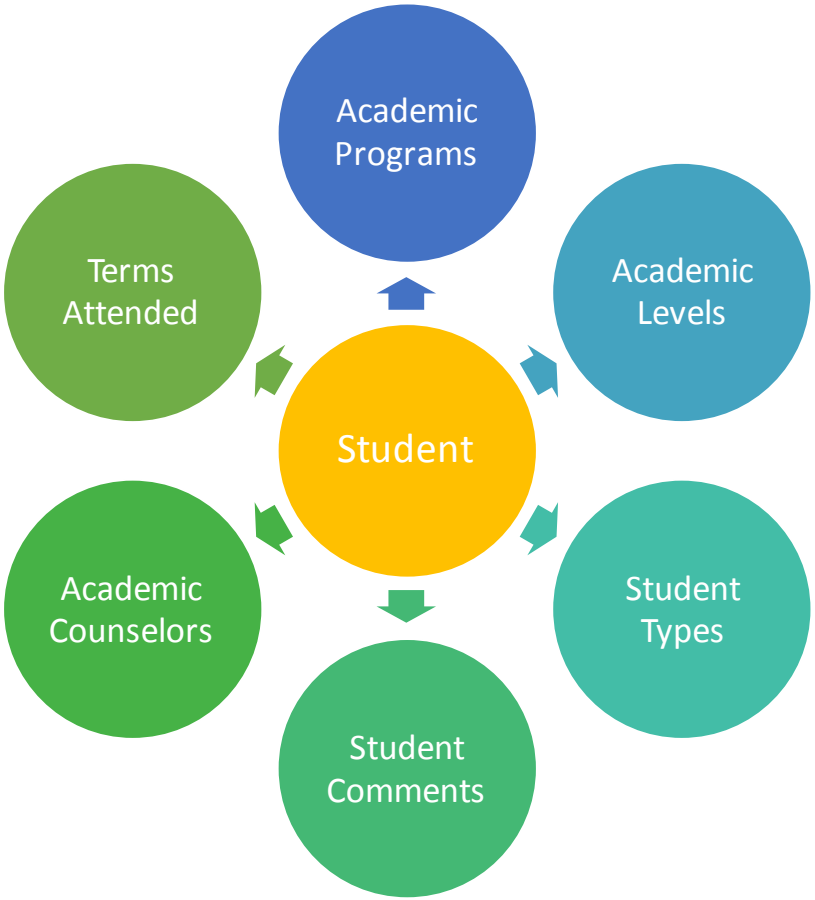


How is UniData different?
Terminology

SQL	UniData
• Table • Row • Column • Key • Foreign Key • Case Insensitive	• File • Record • Field • Record ID • Pointer • Case sensitive (CAPS)

How is UniData different?

Relational database



Multi-valued relational database

Student

Academic Programs	Student Types	Terms Attended
• BS.MATH	• RG - 9/8/1997	• 96/SU
• BA.MATH	• RG - 3/11/1997	• 96/FA
• PB.NONDGR	• RG - 1/28/1997	• 97/IN
	• RG - 9/9/1996	• 97/SP
		• 97/FA
		• 98/IN
		• 98/SP
		• 98/FA
		• 99/IN

How is UniData different?

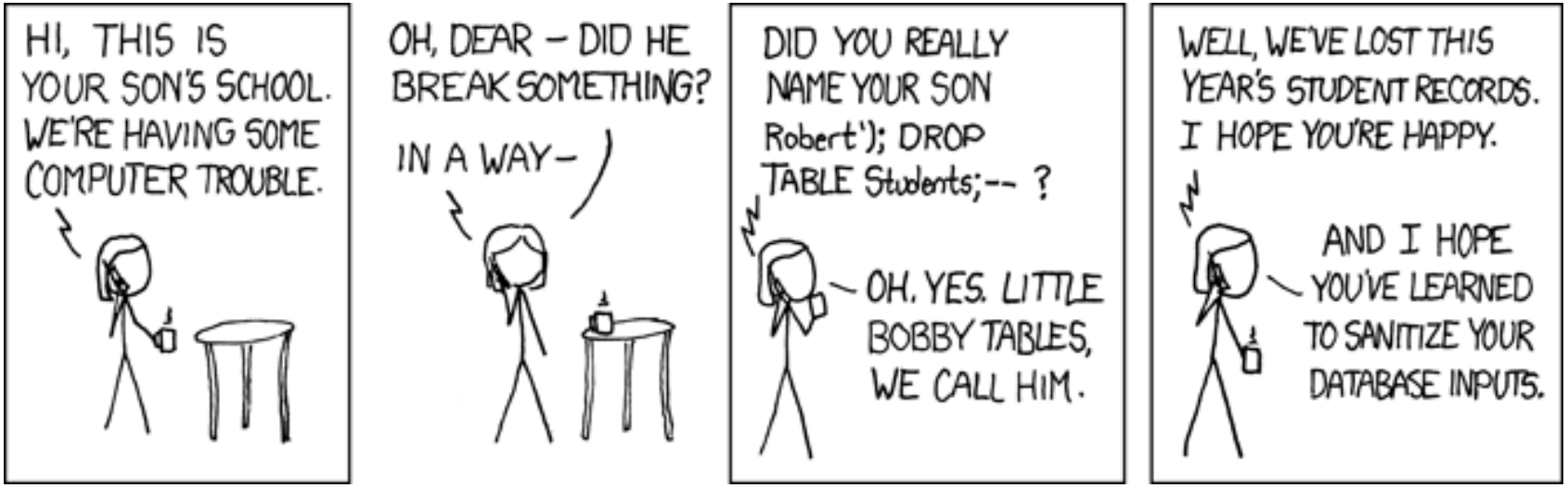
No “join” syntax; no sub-queries

- Need to be able to store record sets for re-use: SAVEDLISTS
- Computed columns are needed in most tables to allow access to data from related tables
- Custom programs written to do those “joins”

Database fields are (very) weakly typed

- 18019 = is this integer, double, string, date?
- Validation must happen on the front end

Data validation



https://imgs.xkcd.com/comics/exploits_of_a_mom.png

Why use UniData?

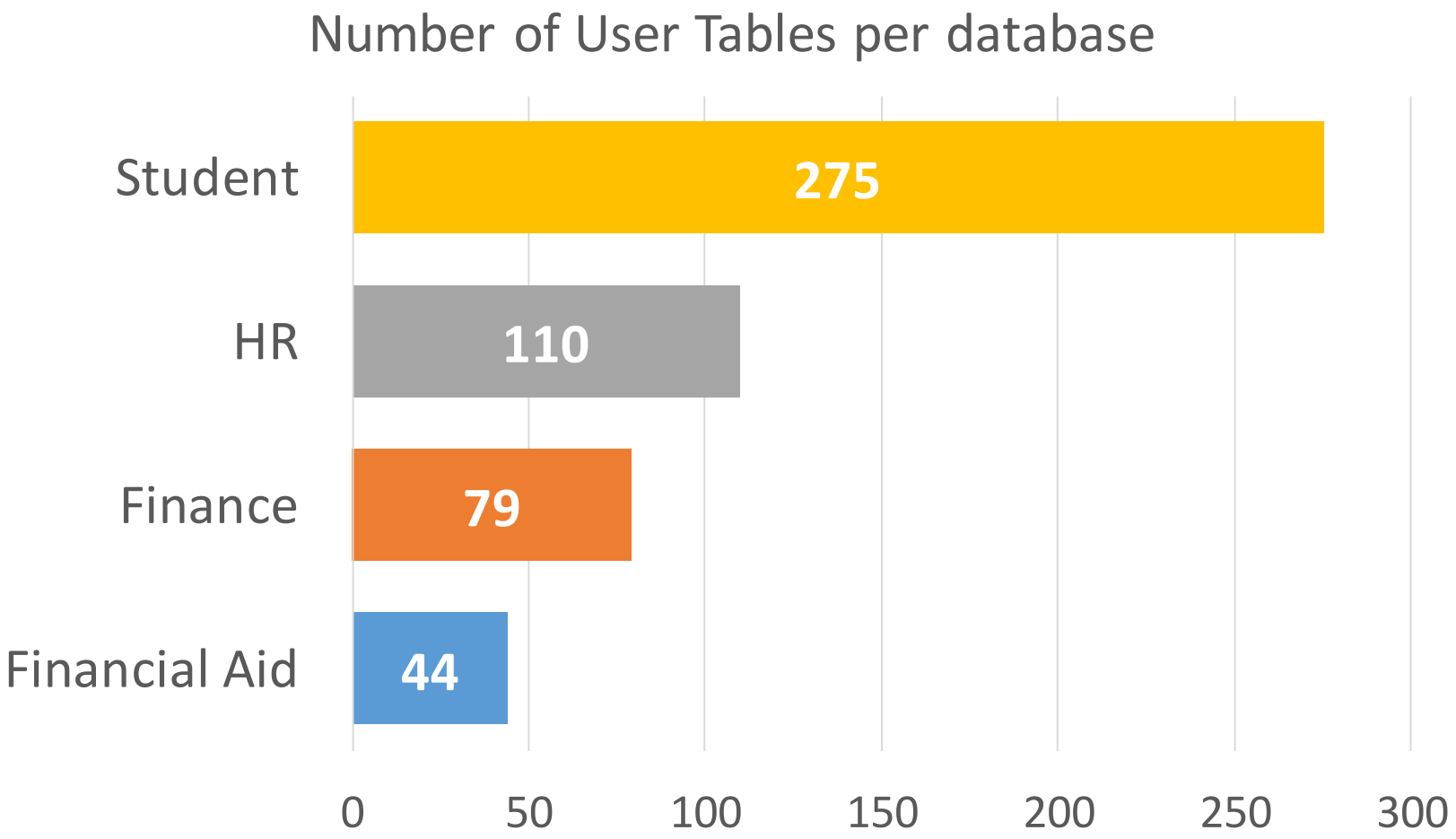
27 years of use

- No other back end database was available for Colleague until 2007
- The application layer STILL runs on UniData
- Identifying what needs to be migrated

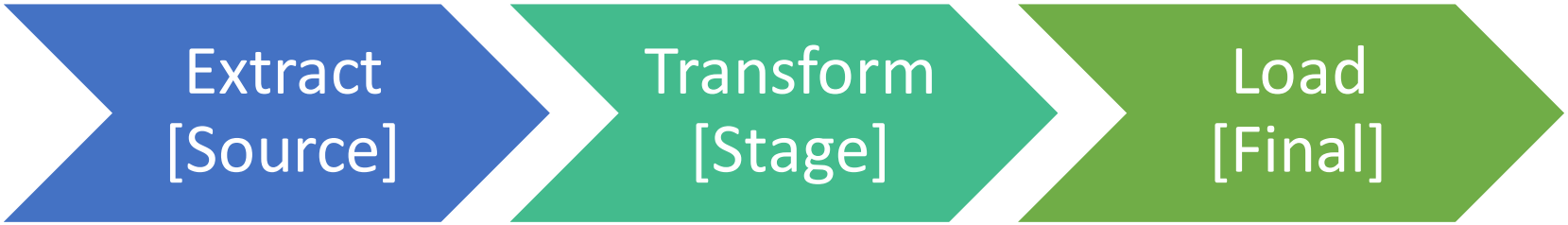
The cost of moving

- Re-writing 1000+ custom programs written at the database level
- Providing a different way to run ~3000 saved user queries
- Migrating 80 GB of data to a more strictly typed database
- Migrating historical Fall and Spring enrollment snapshots for terms prior to 2007
- Training IT staff and end users

SQL reporting data warehouse



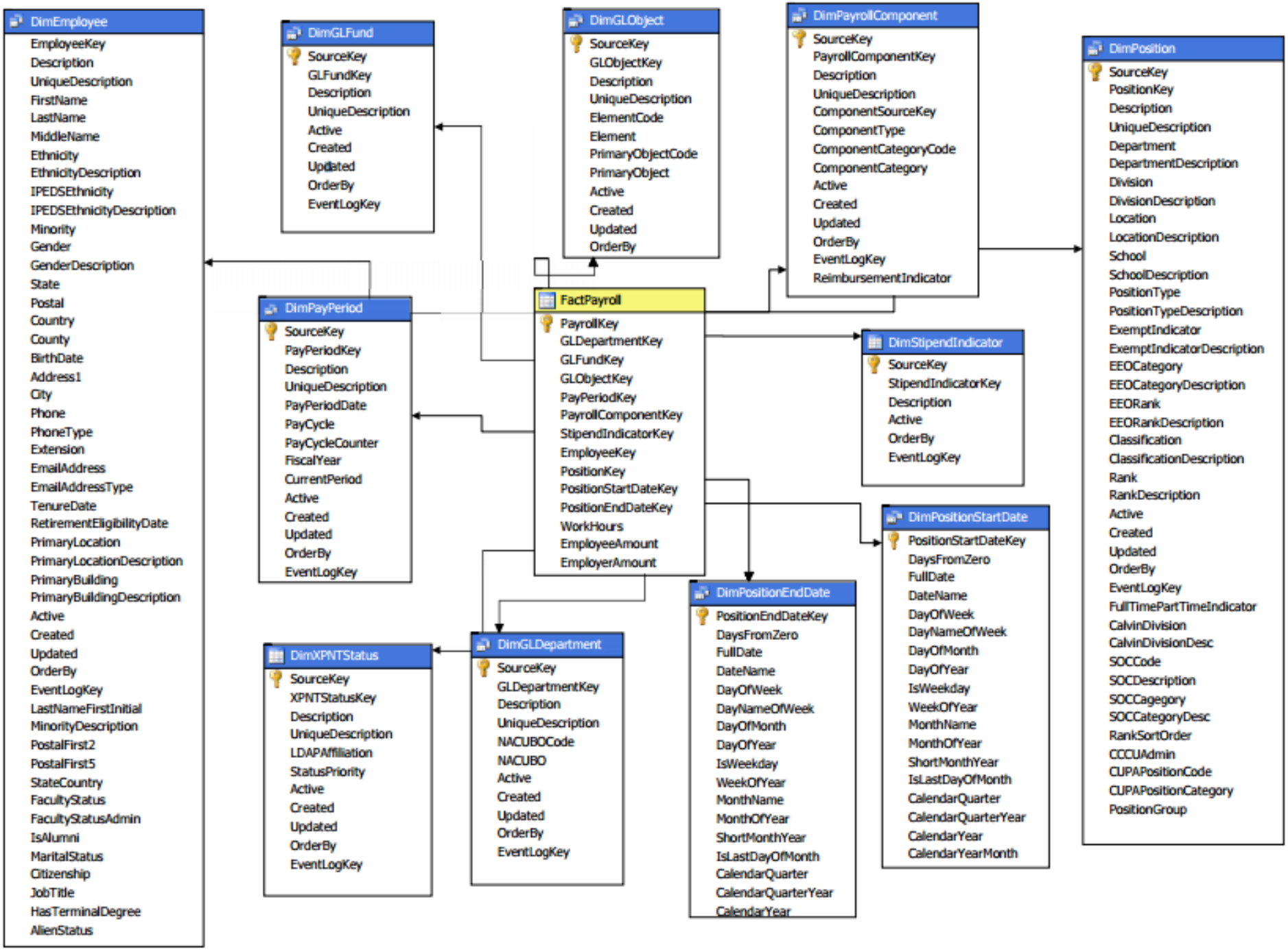
Student ETL Processing

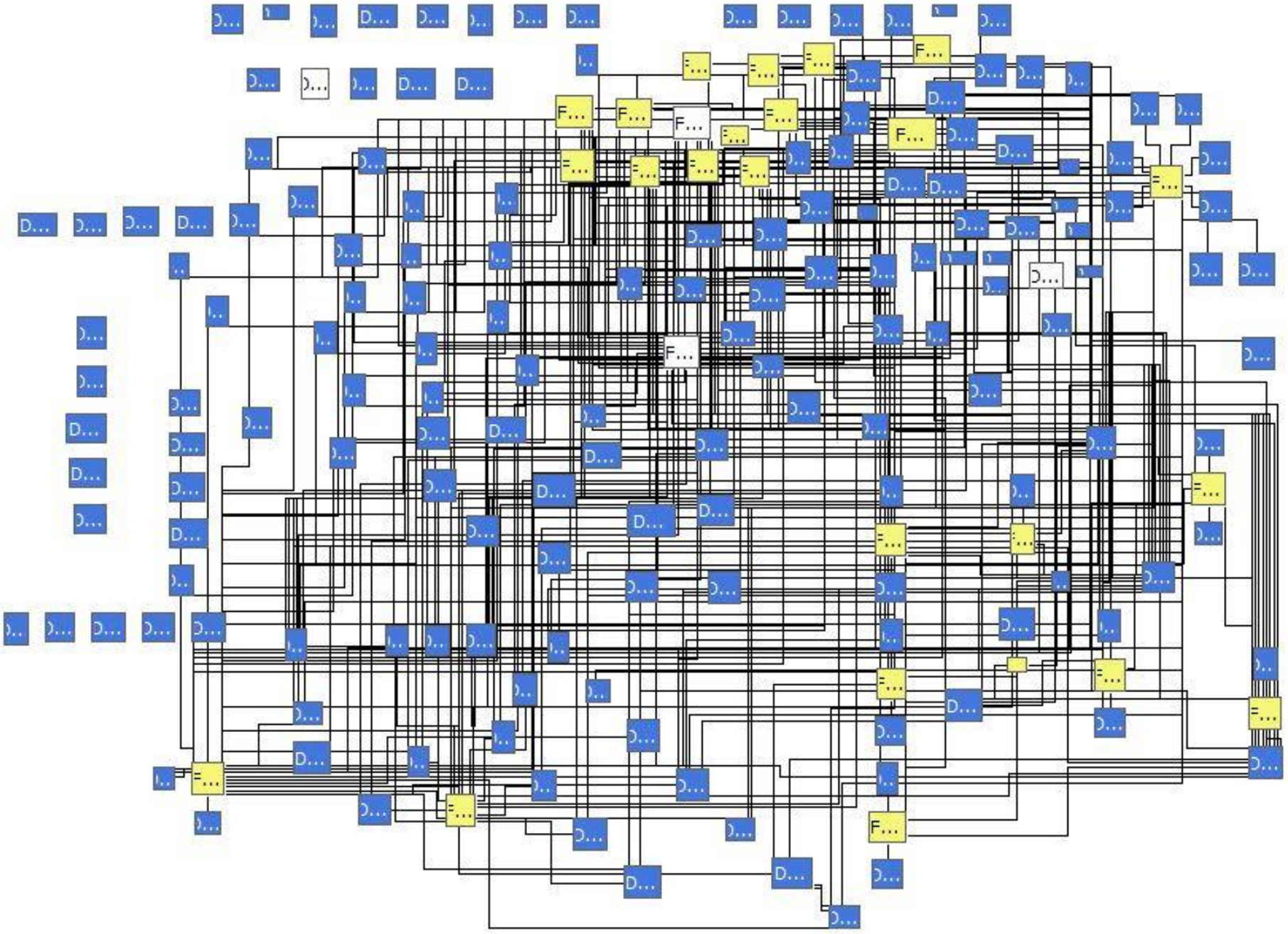


- Export of Colleague data to SQL
- ~20 minutes nightly
- Incremental updates

- Change into star schema: fact and dimension tables
- Predefined relationships

- Data quality firewall
- Transform and load takes ~1 hour


DimEmployee
EmployeeKey
Description
UniqueDescription
FirstName
LastName
MiddleName
Ethnicity
EthnicityDescription
IPEDSEthnicity
IPEDSEthnicityDescription
Minority
Gender
GenderDescription
State
Postal
Country
County
BirthDate
Address1
City
Phone
PhoneType
Extension
EmailAddress
EmailAddressType
TenureDate
RetirementEligibilityDate
PrimaryLocation
PrimaryLocationDescription
PrimaryBuilding
PrimaryBuildingDescription
Active
Created
Updated
OrderBy
EventLogKey
LastNameFirstInitial
MinorityDescription
PostalFirst2
PostalFirst5
StateCountry
FacultyStatus
FacultyStatusAdmin
IsAlumni
MaritalStatus
Citizenship
JobTitle
HasTerminalDegree
AlienStatus
DimGLFund
SourceKey
GLFundKey
Description
UniqueDescription
Active
Created
Updated
OrderBy
EventLogKey
DimGLObject
SourceKey
GLObjectKey
Description
UniqueDescription
ElementCode
Element
PrimaryObjectCode
PrimaryObject
Active
Created
Updated
OrderBy
DimPayrollComponent
SourceKey
PayrollComponentKey
Description
UniqueDescription
ComponentSourceKey
ComponentType
ComponentCategoryCode
ComponentCategory
Active
Created
Updated
OrderBy
EventLogKey
ReimbursementIndicator
DimPosition
SourceKey
PositionKey
Description
UniqueDescription
Department
DepartmentDescription
Division
DivisionDescription
Location
LocationDescription
School
SchoolDescription
PositionType
PositionTypeDescription
ExemptIndicator
ExemptIndicatorDescription
EEOCategory
EEOCategoryDescription
EEORank
EEORankDescription
Classification
ClassificationDescription
Rank
RankDescription
Active
Created
Updated
OrderBy
EventLogKey
FullTimePartTimeIndicator
CalvinDivision
CalvinDivisionDesc
SOCCode
SOCDescription
SOCCagegory
SOCCategoryDesc
RankSortOrder
CCCUAdmin
CUPAPositionCode
CUPAPositionCategory
PositionGroup
DimPayPeriod
SourceKey
PayPeriodKey
Description
UniqueDescription
PayPeriodDate
PayCycle
PayCycleCounter
FiscalYear
CurrentPeriod
Active
Created
Updated
OrderBy
EventLogKey
FactPayroll
PayrollKey
GLDepartmentKey
GLFundKey
GLObjectKey
PayPeriodKey
PayrollComponentKey
StipendIndicatorKey
EmployeeKey
PositionKey
PositionStartDateKey
PositionEndDateKey
WorkHours
EmployeeAmount
EmployerAmount
DimStipendIndicator
SourceKey
StipendIndicatorKey
Description
Active
OrderBy
EventLogKey
DimPositionStartDate
PositionStartDateKey
DaysFromZero
FullDate
DateName
DayOfWeek
DayNameOfWeek
DayOfMonth
DayOfYear
IsWeekday
WeekOfYear
MonthName
MonthOfYear
ShortMonthYear
IsLastDayOfMonth
CalendarQuarter
CalendarQuarterYear
CalendarYear
CalendarYearMonth
DimPositionEndDate
PositionEndDateKey
DaysFromZero
FullDate
DateName
DayOfWeek
DayNameOfWeek
DayOfMonth
DayOfYear
IsWeekday
WeekOfYear
MonthName
MonthOfYear
ShortMonthYear
IsLastDayOfMonth
CalendarQuarter
CalendarQuarterYear
CalendarYear
DimXPNTStatus
SourceKey
XPNTStatusKey
Description
UniqueDescription
LDAPAffiliation
StatusPriority
Active
Created
Updated
OrderBy
EventLogKey
DimGLDepartment
SourceKey
GLDepartmentKey
Description
UniqueDescription
NACUBOCode
NACUBO
Active
Created
Updated
OrderBy
EventLogKey



What do you do with a database?

Get data out of it

Exporting data to use elsewhere

Reporting tools

- Tableau
- Excel
- Portal services

Data for other software systems

- Slate (enrollment CRM)
- Medicat (health services)
- CBORD (meal plans, card swipes)

Etc, Etc, Etc…

- Mailing lists
- Active Directory Groups
- **Ad hoc Queries**

Course Sections taught by Prof. Vander Linden - UniData query

```
* Get the ID number for this user
LIST XNETID 'KVLINDEN' XALL
SELECT COURSE.SECTIONS WITH CSF.FACULTY = 0662493 WITH SEC.CURRENT.STATUS = 'A'
            264 records selected to list 0.
SAVE.LIST KVL.SEC
* What fields should I list?
LIST DICT COURSE.SECTIONS
GET.LIST KVL.SEC
LIST COURSE.SECTIONS SEC.TERM SEC.TERM.START.DATE SEC.SUBJECT SEC.NAME SEC.SHORT.TITLE
GET.LIST KVL.SEC
LIST COURSE.SECTIONS BY.DSND SEC.TERM.START.DATE BREAK.ON SEC.TERM TOTAL EVAL '1' DET.SUP
```

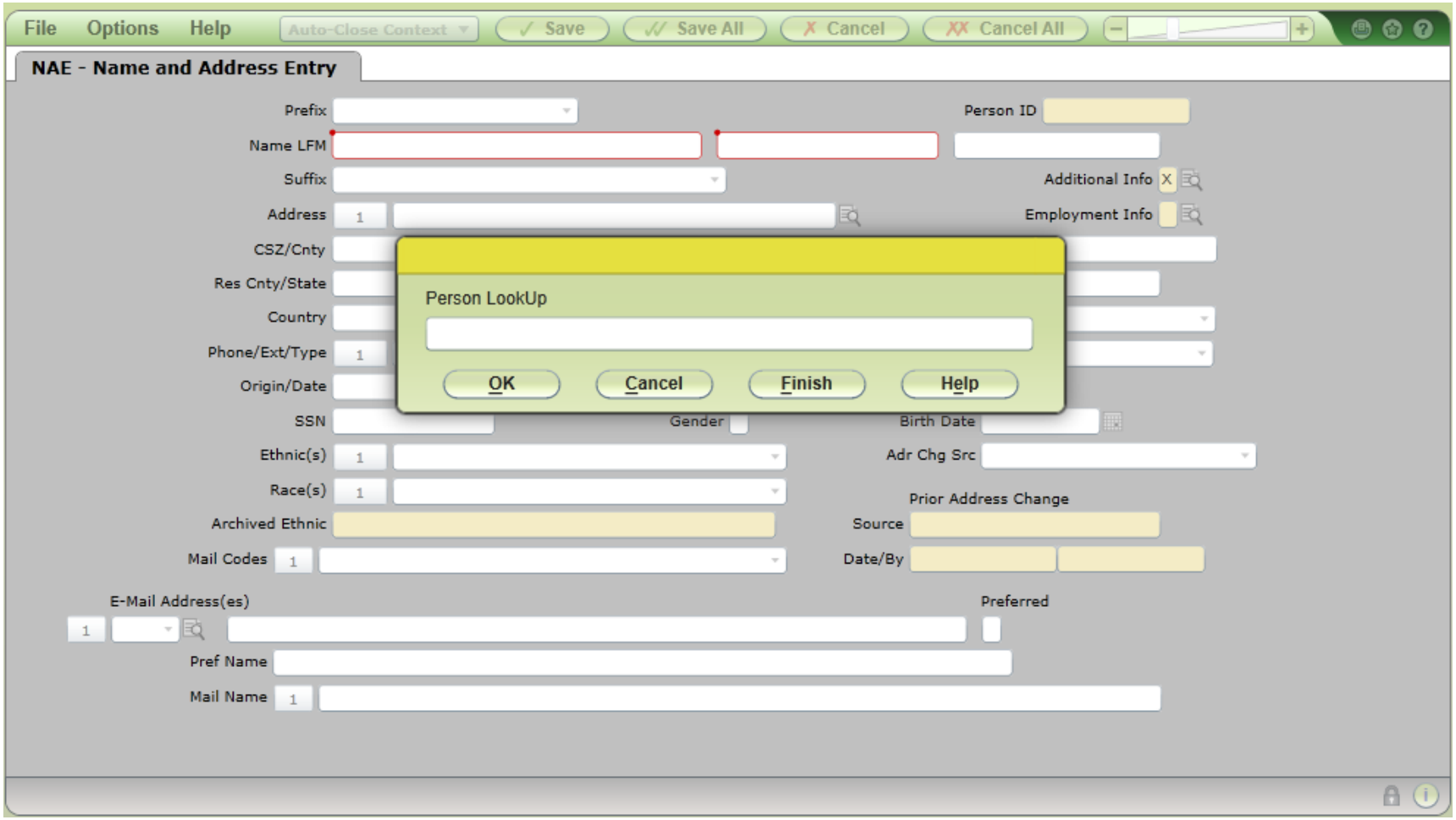
Course Sections taught by Prof. Vander Linden - SQL query

```
select t.SourceKey as Term, t.StartDate, crs.SubjectCode,
        crs.SectionName, crs.Name, crs.Description
from Final.FactClassSchedule c
       join Final.DimTerm t
              on c.TermKey = t.TermKey
       join Final.DimCourse crs
              on c.CourseKey = crs.CourseKey
       join Final.DimFaculty f
              on c.FacultyKey = f.FacultyKey
       join Final.DimSectionStatus ss
              on c.SectionStatusKey = ss.SectionStatusKey
where f.LastName = 'vander linden' and f.FirstName = 'keith'
       and ss.SourceKey = 'a'
       and c.VersionKey = 1
```

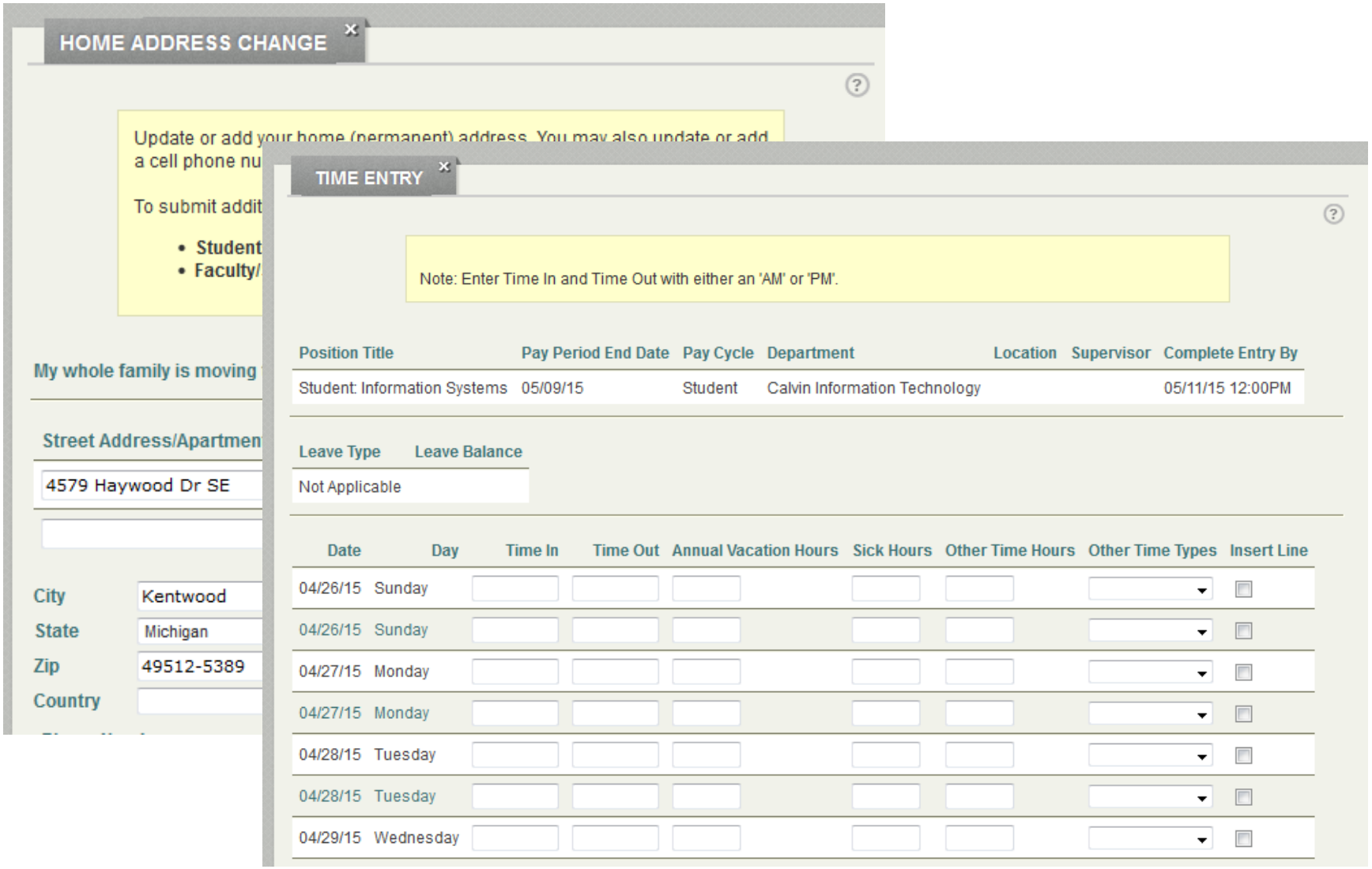
What do you do with a database?

Put data into it

Colleague UI



Web forms


HOME ADDRESS CHANGE
Update or add your home (permanent) address. You may also update or add a cell phone nu
To submit addit
Student
Faculty/
My whole family is moving
Street Address/Apartmen
4579 Haywood Dr SE
City
Kentwood
State
Michigan
Zip
49512-5389
Country
TIME ENTRY
Note: Enter Time In and Time Out with either an 'AM' or 'PM'.
Position Title
Pay Period End Date
Pay Cycle
Department
Location
Supervisor
Complete Entry By
Student: Information Systems
05/09/15
Student
Calvin Information Technology
05/11/15 12:00PM
Leave Type
Leave Balance
Not Applicable
Date
Day
Time In
Time Out
Annual Vacation Hours
Sick Hours
Other Time Hours
Other Time Types
Insert Line
04/26/15 Sunday
04/26/15 Sunday
04/27/15 Monday
04/27/15 Monday
04/28/15 Tuesday
04/28/15 Tuesday
04/29/15 Wednesday

Common data errors

Typos	Data that is "valid" but doesn't make sense	Not following data entry standards
• 123 Main St vs 123 Maine SE • 616-526-60000 • mymail@gmail.co	• Communication date of 4/17/1115 • SAT score of 2600 • Mr. Sarah Greenfield	• Person records with only first and last name • When entering parents and children, siblings are not related to each other • Entering a country on US addresses

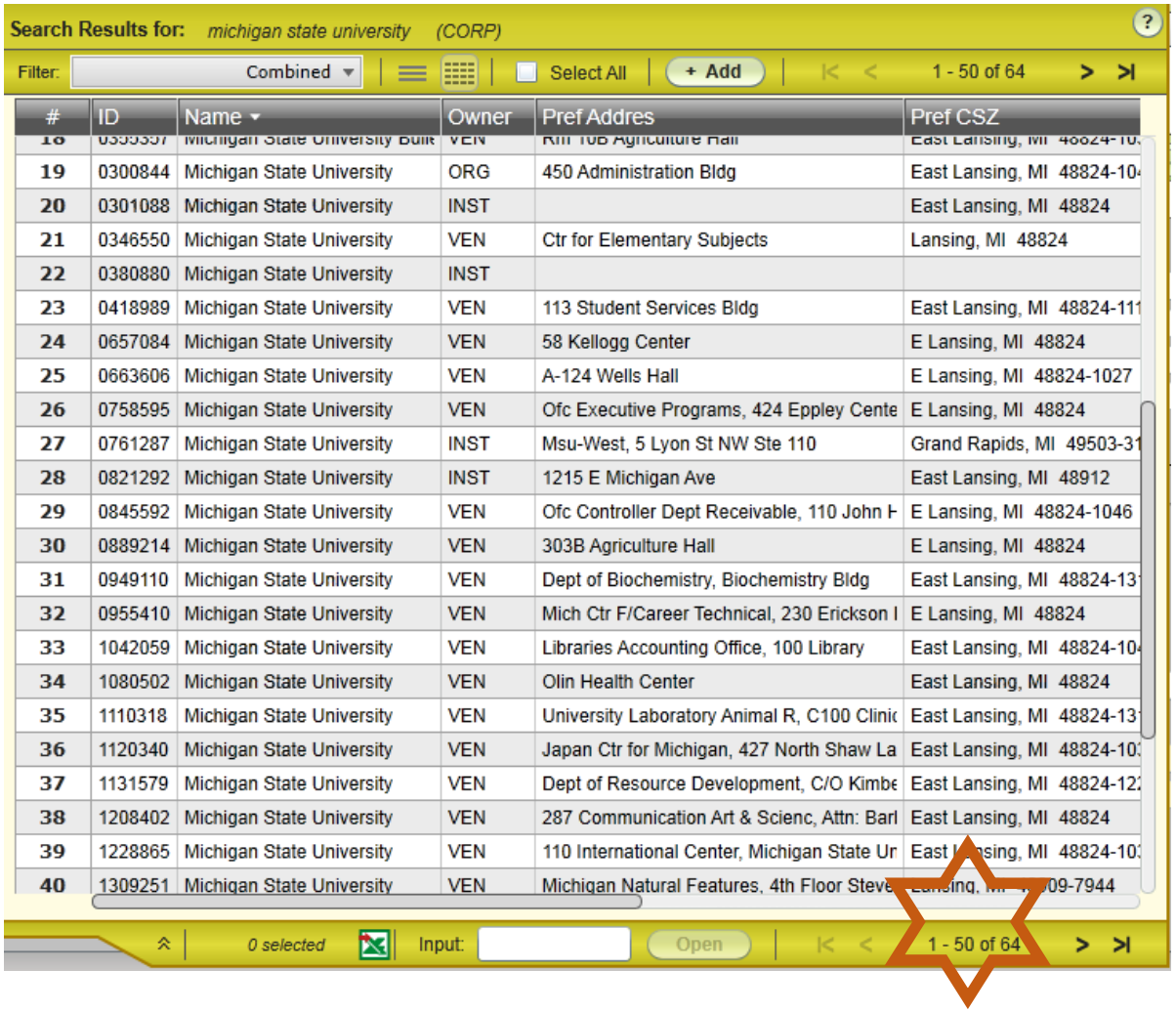
Duplicates

Scope

- 150-200 identified each year
- Can take 1 hour to clean up one record

Cost

- Upset people
- Duplicate mailings
- Incorrect information
- Rework
- Cleanup



#	ID	Name	Owner	Pref Addres	Pref CSZ
19	0300844	Michigan State University	ORG	450 Administration Bldg	East Lansing, MI 48824-10
20	0301088	Michigan State University	INST		East Lansing, MI 48824
21	0346550	Michigan State University	VEN	Ctr for Elementary Subjects	Lansing, MI 48824
22	0380880	Michigan State University	INST		
23	0418989	Michigan State University	VEN	113 Student Services Bldg	East Lansing, MI 48824-111
24	0657084	Michigan State University	VEN	58 Kellogg Center	E Lansing, MI 48824
25	0663606	Michigan State University	VEN	A-124 Wells Hall	E Lansing, MI 48824-1027
26	0758595	Michigan State University	VEN	Ofc Executive Programs, 424 Eppley Cente	E Lansing, MI 48824
27	0761287	Michigan State University	INST	Msu-West, 5 Lyon St NW Ste 110	Grand Rapids, MI 49503-31
28	0821292	Michigan State University	INST	1215 E Michigan Ave	East Lansing, MI 48912
29	0845592	Michigan State University	VEN	Ofc Controller Dept Receivable, 110 John H	E Lansing, MI 48824-1046
30	0889214	Michigan State University	VEN	303B Agriculture Hall	E Lansing, MI 48824
31	0949110	Michigan State University	VEN	Dept of Biochemistry, Biochemistry Bldg	East Lansing, MI 48824-13
32	0955410	Michigan State University	VEN	Mich Ctr F/Career Technical, 230 Erickson I	E Lansing, MI 48824
33	1042059	Michigan State University	VEN	Libraries Accounting Office, 100 Library	East Lansing, MI 48824-10
34	1080502	Michigan State University	VEN	Olin Health Center	East Lansing, MI 48824
35	1110318	Michigan State University	VEN	University Laboratory Animal R, C100 Clinic	East Lansing, MI 48824-13
36	1120340	Michigan State University	VEN	Japan Ctr for Michigan, 427 North Shaw La	East Lansing, MI 48824-10
37	1131579	Michigan State University	VEN	Dept of Resource Development, C/O Kimbe	East Lansing, MI 48824-12
38	1208402	Michigan State University	VEN	287 Communication Art & Scienc, Attn: Barl	East Lansing, MI 48824
39	1228865	Michigan State University	VEN	110 International Center, Michigan State Ur	East Lansing, MI 48824-10
40	1309251	Michigan State University	VEN	Michigan Natural Features, 4th Floor Steve	Lansing, MI 09-7944

Questions?