

Higher Education Data Warehousing Forum



HEDW

Long Beach Community College's Insight Process
Focusing on a Right-Size Solution with Phytorion

Presentation Overview



- Insight to our challenges
- Identifying the right-size strategy
- Choosing a partner
- Implementing the project
- College and Phytorion roles
- Data warehouse demo



Long Beach City College



- Two year California Community College
- Located in Los Angeles metropolitan region
- Over 3,000 classes each term
- 1,000+ courses
- 27,000 students
- 1575 employees
- 2 campuses



What Did We Need?



- A degree of independence for business analysts
- A way around limits imposed by joins and number of tables
- The ability to access data without impacting production
- Flexibility and fleetness of product was very important

Problems before Data Warehouse



- Reports ran against production dragging down performance on transactions



- Reports took too long to run due to complexity of PS data, joins and effective data rows



- Limited experience writing complex reports by Business Analysts

The Good Old Days

- Reports developed by IT
- Reports from disconnected databases
- Less demand for ad hoc data



How Did We Get Data Before?

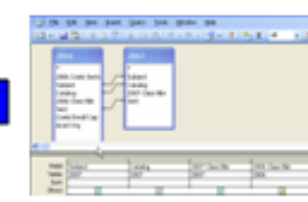


Push it into Excel

	A	B	C	D
1				
2				
3				
4				
5				
6				
7				



	A	B	C	D
1				
2				
3				
4				
5				
6				
7				



	A	B	C	D
1				
2				
3				
4				
5				
6				
7				

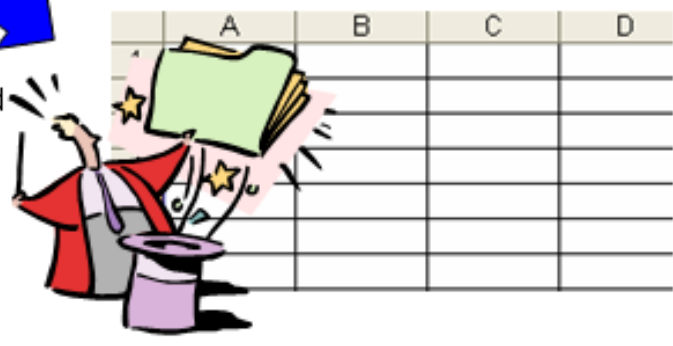
Back to Excel

Back to Access

Back to Excel



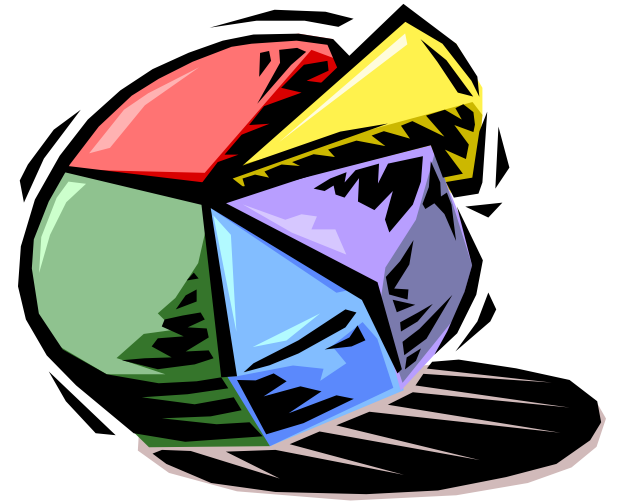
Send out completed Reports via email



Here We Go Again

What? You want to ask another question about that slice of the pie?

Start over.



Why a Data Warehouse?



Transactional systems:

- Optimized for processing
- Thousands of tables
- Effective dated
- Tree structures
- Complex joins to report on needed data
- Live data

What about a Dimensional Model?



- Optimized for reporting
- Trees flattened to allow for quick rollup
- Effective dates transformed to valid date ranges and current flags
- Dimensions allow efficient data slicing
- Lots of derived customized fields
- Near live data

Joining Tables is Easier in a DW



- Surrogate ID's or SID's allow for high-speed efficient joins between fact and dimension tables



- Naming conventions ensure that there's never any doubt about which fields to join



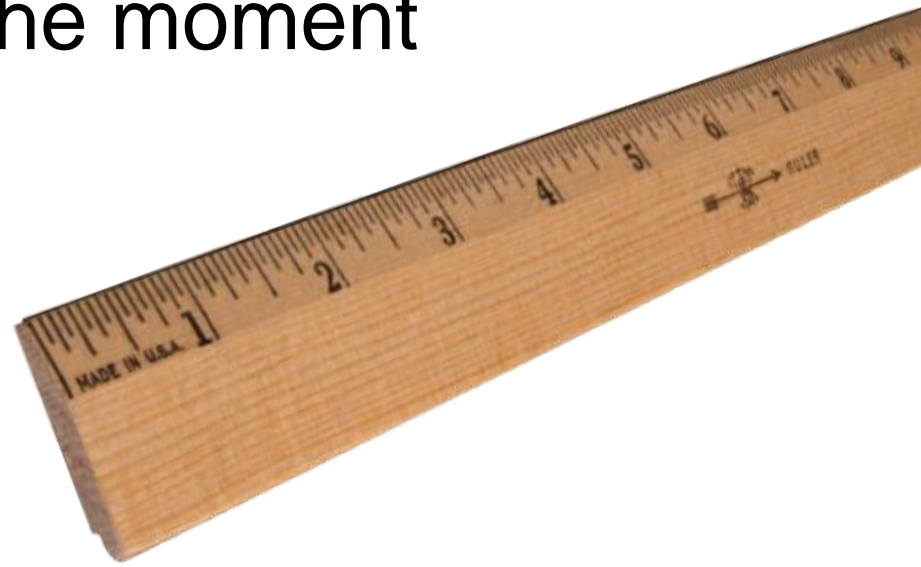
Problems Getting to Your Data?



What is the Right Size?



- A flexible, scalable implementation
- High value: low cost for great results
- Large scale solutions – does one size ever fit?
- In house applications are often designed for the moment



The Right Size Answer for Us



- Build a data warehouse
- Use outside expertise
 - No learning curve
 - A faster time to deployment
- Design 80% jumping off solution
- Expect to grow and expand
- Knowledge transfer

Why a Custom DW with Phytorion?



- In depth expertise and experience
- Time savings
- They knew how to consolidate the data, speed processing
- Reporting sophistication
- Standardized design process
- Ability to interpret our business requirements
- Knowledge transfer

Project Steps



1. What are the business questions?
2. Understand the requirements
3. What PeopleSoft tables are needed
4. Which data is fact and which is dimensional?
5. Decide level of granularity needed
5. Build the data warehouse
6. Test the data warehouse
7. Review the Visio documents and data maps
8. Start writing reports!

Our Story – What's Included



- Counseling Datebook and Assessment
- Financial Aid
- Admissions & Records
- Human Resources
- Academic Services
- Student Services
- Financials

Data Highlights



- Characteristics of Students
- Degrees, Certificates, Honors
- Information about each course
- Information about classes
- Roll ups → Division → School → Dept → Subject
- State reporting and complex derived fields
- Financial information
- Human Resources employee information
- Student GPA, load, status, test scores
- ISIR, PELL, & FA characteristics

What Else? Derived Fields!

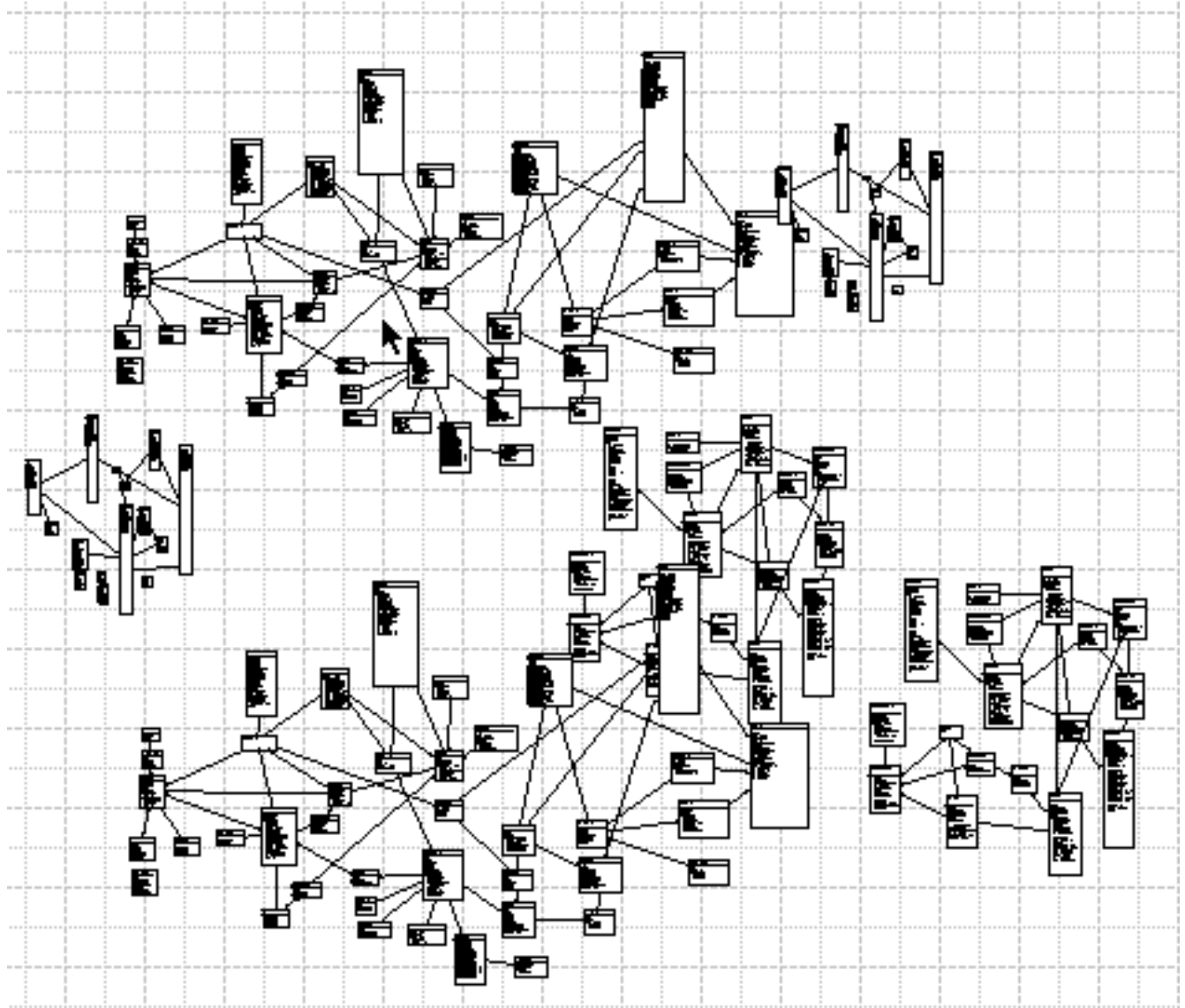


- **Course Key:** HIST 10 (Subject + Catalog Number)
- **FTES** full time equivalent student:
If attendance type = "W" then count residents enrolled *
class hours/meeting * 17.5 / 525
- **Weeks of Instruction:**
actual weeks calculated according to dates
- **Days:**
converted from Y & N's to Monday, Tuesday, etc.
- **Meeting Patterns:** converted to MWF
- Save time and cut down on errors



Anything is possible with a customized approach

Start with this Structure



Transformed it to This!



STUDENT DIMENSION
 STU_PK
 STU_DATE_TIME
 STU_EMPID
 STU_SSN
 STU_NAME
 STU_GENDER
 STU_BIRTHDATE
 STU_CAMPUS_EMAIL_CD
 STU_CAMPUS_EMAIL
 STU_CAMPUS_ADDR_CD
 STU_CAMPUS_ADDR1
 STU_CAMPUS_ADDR2
 STU_CAMPUS_ADDR3
 STU_CAMPUS_CITY
 STU_CAMPUS_STATE
 STU_CAMPUS_ZIP
 STU_CAMPUS_PHONE_CD
 STU_CAMPUS_PHONE
 STU_CAMPUS_PHONE_EXT
 STU_HOME_EMAIL_CD
 STU_HOME_EMAIL
 STU_HOME_ADDR_CD
 STU_HOME_ADDR1
 STU_HOME_ADDR2
 STU_HOME_ADDR3
 STU_HOME_CITY
 STU_HOME_STATE
 STU_HOME_ZIP
 STU_HOME_PHONE_CD
 STU_HOME_PHONE
 STU_HOME_PHONE_EXT
 STU_CITZ_STATUS_CD
 STU_CITZ_SDESC
 STU_CITZ_LDESC
 STU_ETHNIC_CD
 STU_ETHNIC_PRIME
 STU_ETHNIC_PRIME_SDESC
 STU_ETHNIC_PRIME_LDESC
 STU_ADMIS_EXT_PK
 STU_FINAD_EXT_PK
 STU_STUFN_EXT_PK
 STU_CITZ_LDESC
 STU_ETHNIC_CD
 STU_ETHNIC_PRIME
 STU_ETHNIC_PRIME_SDESC
 STU_ETHNIC_PRIME_LDESC
 STU_ADMIS_EXT_PK
 STU_FINAD_EXT_PK
 STU_STUFN_EXT_PK

DATE_TIME DIMENSION
 DTM_PK
 DTM_DATE_TIME
 DTM_EMPID
 DTM_TYPE
 DTM_DESCR
 DTM_ACAD_YEAR
 DTM_WEEKS_OF_INSTRUCT
 DTM_CAL_YEAR
 DTM_CAL_QUARTER
 DTM_CAL_MONTH_NBR
 DTM_CAL_MONTH_SDESC
 DTM_CAL_MONTH_LDESC
 DTM_CAL_WEEK_NBR
 DTM_CAL_WEEK_NBR_YTD
 DTM_CAL_DAY_NBR
 DTM_CAL_DAY_SDESC
 DTM_CAL_DAY_LDESC
 DTM_CAL_DAY_NBR_YTD
 SA_VERSION
 SA_RDBMS_VERSION
 DECISION_STREAM_VERSION
 DS_ODS_VERSION
 Other fields that would help us audit changes to the ODS.

CLASS_SECTION ENROLLMENT FACTS
 ENRL_PK
 ENRL_STU_PK
 ENRL_CLASS_FK
 ENRL_DATE_TIME_FK
 ENRL_HEADCOUNT
 ENRL_INT_TAKEN
 ENRL_GRADE_POINTS
 ENRL_GRADE_OFFER_NBR_PK
 ENRL_STU_FK
 ENRL_CLASS_FK
 ENRL_DATE_TIME_FK
 ENRL_HEADCOUNT
 ENRL_INT_TAKEN
 ENRL_GRADE_POINTS
 ENRL_GRADE_OFF

STUDENT DIMENSION
 STU_PK
 STU_DATE_TIME
 STU_EMPID
 STU_SSN
 STU_NAME
 STU_GENDER
 STU_BIRTHDATE
 STU_CAMPUS_EMAIL_CD
 STU_CAMPUS_EMAIL
 STU_CAMPUS_ADDR_CD
 STU_CAMPUS_ADDR1
 STU_CAMPUS_ADDR2
 STU_CAMPUS_ADDR3
 STU_CAMPUS_CITY
 STU_CAMPUS_STATE
 STU_CAMPUS_ZIP
 STU_CAMPUS_PHONE_CD
 STU_CAMPUS_PHONE
 STU_CAMPUS_PHONE_EXT
 STU_HOME_EMAIL_CD
 STU_HOME_EMAIL
 STU_HOME_ADDR_CD
 STU_HOME_ADDR1
 STU_HOME_ADDR2
 STU_HOME_ADDR3
 STU_HOME_CITY
 STU_HOME_STATE
 STU_HOME_ZIP
 STU_HOME_PHONE_CD
 STU_HOME_PHONE
 STU_HOME_PHONE_EXT
 STU_CITZ_STATUS_CD
 STU_CITZ_SDESC
 STU_CITZ_LDESC
 STU_ETHNIC_CD
 STU_ETHNIC_PRIME
 STU_ETHNIC_PRIME_SDESC
 STU_ETHNIC_PRIME_LDESC
 STU_ADMIS_EXT_PK
 STU_FINAD_EXT_PK
 STU_STUFN_EXT_PK

CAREER_TERM DIMENSION
 CTRM_PK
 CTRM_DATE_TIME
 CTRM_STUD
 CTRM_ACAD_CAREER
 CTRM_ACAD_CAREER_SDESC
 CTRM_ACAD_CAREER_LDESC
 CTRM_ACAD_CAREER_NBR
 CTRM_INSTITUTION
 CTRM_INSTITUTION_SDESC
 CTRM_INSTITUTION_LDESC
 CTRM_STRM_CD
 CTRM_STRM_SDESC
 CTRM_STRM_LDESC
 CTRM_STRM_BEGIN_DT
 CTRM_STRM_END_DT
 CTRM_WEEKS_OF_INSTRUCT
 CTRM_TERM_CATEGORY_CD
 CTRM_TERM_CATEGORY_SDESC
 CTRM_TERM_CATEGORY_LDESC
 CTRM_ACAD_YEAR
 CTRM_WITHD_CD
 CTRM_WITHD_SDESC
 CTRM_WITHD_REASON_CD
 CTRM_WITHD_REASON_SDESC
 CTRM_WITHD_REASON_LDESC
 CTRM_WITHDRAW_DATE
 CTRM_ACAD_PROG_PRIME_CD
 CTRM_ACAD_PROG_PRIME_SDESC
 CTRM_ACAD_PROG_PRIME_LDESC
 CTRM_ACAD_PROG_STATUS_CD
 CTRM_ACAD_PROG_STATUS_SDESC
 CTRM_ACAD_PROG_STATUS_LDESC
 CTRM_ACAD_PROG_ACTION_CD
 CTRM_ACAD_PROG_ACTION_SDESC
 CTRM_ACAD_PROG_ACTION_LDESC
 CTRM_ACAD_PROG_REASON_CD
 CTRM_ACAD_PROG_REASON_SDESC
 CTRM_ACAD_PROG_REASON_LDESC
 CTRM_ACAD_LOAD_CD
 CTRM_ACAD_LOAD_SDESC
 CTRM_ACAD_LOAD_LDESC
 CTRM_ACAD_LEVEL_BOB_CD
 CTRM_ACAD_LEVEL_BOB_SDESC
 CTRM_ACAD_LEVEL_BOB_LDESC
 CTRM_ACAD_LEVEL_EOT_CD
 CTRM_ACAD_LEVEL_EOT_SDESC
 CTRM_ACAD_LEVEL_EOT_LDESC
 CTRM_CLR_GPA_RANGE
 CTRM_CLR_GPA_RANGE

CLASS DIMENSION
 CLS_PK
 CLS_DATE_TIME
 CLS_COURSE_ID
 CLS_COURSE_OFFER_NBR
 CLS_TERM
 CLS_TERM_SDESC
 CLS_TERM_LDESC
 CLS_SESSION_CD
 CLS_SESSION_SDESC
 CLS_SESSION_LDESC
 CLS_CLASS_SECTION
 CLS_INSTITUTION_CD
 CLS_INSTITUTION_SDESC
 CLS_INSTITUTION_LDESC
 CLS_CAMPUS_CD
 CLS_CAMPUS_SDESC
 CLS_CAMPUS_LDESC
 CLS_LOCATION_CD
 CLS_LOCATION_SDESC
 CLS_LOCATION_LDESC
 CLS_ADDRESS1
 CLS_CITY
 CLS_STATE
 CLS_ZIP
 CLS_GROUP_CD
 CLS_GROUP_SDESC
 CLS_GROUP_LDESC
 CLS_ORG_CD
 CLS_ORG_SDESC
 CLS_ORG_LDESC
 CLS_SUBJECT_CD
 CLS_SUBJECT_SDESC
 CLS_SUBJECT_LDESC
 CLS_ORP_CD
 CLS_HIGRS_CD
 CLS_CATALOG_NBR
 CLS_COURSE_LDESC
 CLS_CLASS_NBR
 CLS_CLASS_COMPONENT_CD
 CLS_CLASS_COMPONENT_SDESC
 CLS_CLASS_COMPONENT_LDESC
 CLS_CLASS_TYPE
 CLS_CLASS_TYPE_SDESC
 CLS_CLASS_TYPE_LDESC
 CLS_ASSOCIATED_CLASS_NBR
 CLS_START_DATE
 CLS_END_DATE
 CLS_FACILITY_CODE1
 CLS_FACILITY_SDESC1
 CLS_FACILITY_LDESC1
 CLS_STANDARDS_MTG_PATTERN1
 CLS_MTG_START_TIME1
 CLS_MTG_END_TIME1
 CLS_FACILITY_CODE2
 CLS_FACILITY_SDESC2
 CLS_FACILITY_LDESC2
 CLS_STANDARDS_MTG_PATTERN2
 CLS_MTG_START_TIME2
 CLS_MTG_END_TIME2
 CLS_INSTRUCTOR_EMPID
 CLS_INSTRUCTOR_NAME
 CLS_INSTRUCTOR_PHONE
 CLS_ENROL_TOTAL

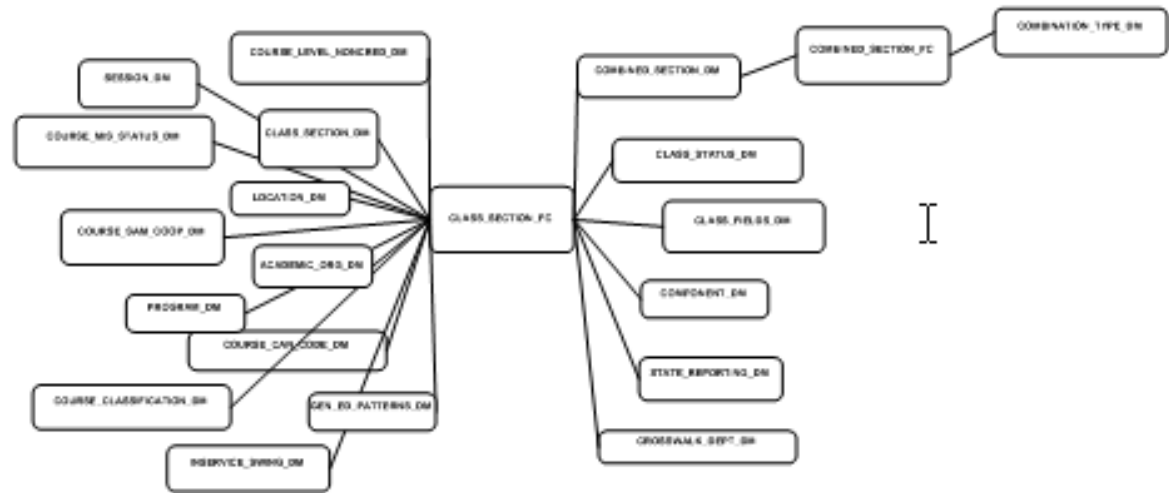


Delivered Documents – Data Maps



CLASS_SECTION_DM		BUILD_NAME=CLASS_SECTION_DM_B9					
CLASS_SECTION_DM	CLASS_SECTION_SID	SID					
CLASS_SECTION_DM	COURSE_ID	M	S-K1		PS_CLASS_TBL	CRSE_ID	
CLASS_SECTION_DM	COURSE_OFFER_NUMBER	M	S-K2		PS_CLASS_TBL	CRSE_OFFER_NBR	
CLASS_SECTION_DM		FUNC				COURSE_OFFERING_DS_CONCAT:CHAR(11);'.CRSE_ID,CRSE_OFFER_NBR	
CLASS_SECTION_DM	COURSE_OFFERING_SID	L	T		STG_COURSE_OFFERING_DM	COURSE_OFFERING	COURSE_OFFERING_DM_KEY,MAX(VALID_FROM_DATE)<=STA
CLASS_SECTION_DM	TERM	M	S-K3		PS_CLASS_TBL	STRM	
CLASS_SECTION_DM	TERM_SID	L	T		TERM_DM	TERM_SID	TERM=STRM
CLASS_SECTION_DM	SESSION_CODE	M	S-K4		PS_CLASS_TBL	SESSION_CODE	
CLASS_SECTION_DM	SESSION_SID	L	T		SESSION_DM	SESSION_SID	TERM=STRM,SESSION_CODE
CLASS_SECTION_DM	CLASS_SECTION	M	S-K5		PS_CLASS_TBL	CLASS_SECTION	
CLASS_SECTION_DM		M	S		PS_CLASS_TBL	START_DT	
CLASS_SECTION_DM		M	S		PS_CLASS_TBL	SUBJECT	
CLASS_SECTION_DM	CATALOG_NUMBER	M	S		PS_CLASS_TBL	CATALOG_NBR	
CLASS_SECTION_DM		FUNC				TRIM_CATALOG_NE	TRIM:CHAR(10),CATALOG_NBR
CLASS_SECTION_DM	COURSE_KEY	FUNC					CONCAT:CHAR(18);',SUBJECT,TRIM_CATALOG_NBR
CLASS_SECTION_DM	CLASS_DESCR	M	S	NA	PS_CLASS_TBL	DESCR	
CLASS_SECTION_DM	CLASS_NUMBER	M	S		PS_CLASS_TBL	CLASS_NBR	
CLASS_SECTION_DM	ASSOCIATED_CLASS	M	S		PS_CLASS_TBL	ASSOCIATED_CLASS	
CLASS_SECTION_DM	PRIMARY_INSTR_SECTION	M	S		PS_CLASS_TBL	PRIM_INSTR_SECT	
CLASS_SECTION_DM		L	T		CLASS_SECTION_DM_B2	PAY_TYPE_DISTINC	CRSE_ID,STRM,SESSION_CODE,CLASS_SECTION
CLASS_SECTION_DM	MIXED_PAY_FLAG	FUNC	N				CASE:CHAR(1),PAY_TYPE_DISTINCT_COUNT>1;'Y',DEFAULT,'N'
CLASS_SECTION_DM		M	S		PS_CLASS_TBL	INSTITUTION	
CLASS_SECTION_DM		LJ	S		PS_SCTN_CMBND	SCTN_COMBINED_I	INSTITUTION,STRM,SESSION_CODE,IGNORE(SCTN_COMBINED_I
CLASS_SECTION_DM		FUNC				SESSION_PLUS_SCT	CONCAT:CHAR(7);',SESSION_CODE,SCTN_COMBINED_ID
CLASS_SECTION_DM	COMBINED_KEY	FUNC					CASE:CHAR(7),SCTN_COMBINED_ID='';',DEFAULT,SESSION_PL

Delivered Documents – User Docs



Fact Table: Class Section FC

Granularity: One row per course per term per session per class section

Primary PK Tables: PK CLASS_IDB

Fact: Enrollment and waitlist capacities and percentages full, FTE percentages for various instructor types

Dimensions:

Class Section Dim: Contains descriptive class information such as course ID, course description, class number, course title, equivalent courses, subject and cable number

Session Dim: Contains information about the academic session (description, start/end dates)

Course SAM Coop Dim: Contains codes and descriptions for attributes SAM Priority (CB09) and Coop Work Experience (CB10)

Course Classification Dim: Contains code and descriptions for course classification (CB11)

Course Can Code Dim: Contains codes and descriptions for can code (CB15) and can sequence code (CB15)

State Reporting Dim: Contains codes and descriptions for accounting method (XB01) and contracted code (XB04)

Delivered Document – ETL Training

DecisionStream 7.1 Training Guide Chapter 2: Builds

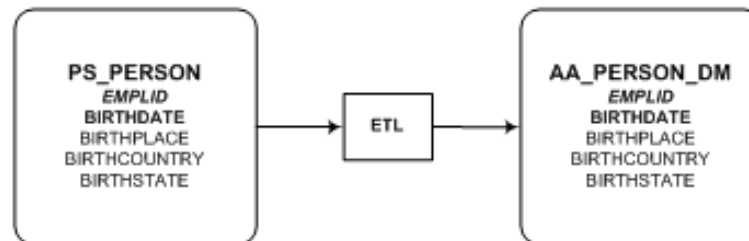


Chapter 2: Builds

This walk through will take you through the steps to create a build. This build will move data from the PS_PERSON in the source database to a warehouse table AA_PERSON_DM. The diagram below details the columns you will use from PS_PERSON to populate AA_PERSON_DM.

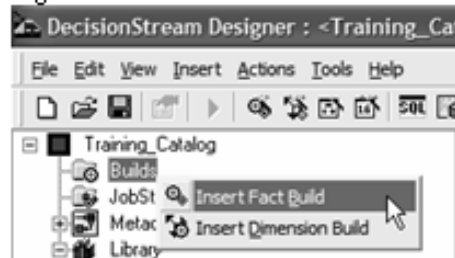
NOTE: This Training Guide is based on PeopleSoft version 8.9. If you are on an earlier version, you will have to make the following changes in the training exercises.

- Substitute PS_PERSONAL_DATA for PS_PERSON.



Creating A New Build

1. Right click Builds and select Insert Fact Build.



Right Size



- Solid base: meeting 80% of needs
- Off and running
 - Usable by business analysts
 - Rolled out to front line workers
- An Insight into LBCC
 - Exec level decisions supported
 - Important new initiatives supported
 - Faster analytics
- Versatile platform for the future

What We Got



- Instance success – Up and running in 4 months!
- Our definitions, our data – every college is truly different



- Concrete results that everyone can see and benefit from
- Knowledge transfer



- Fewer paths to take and fewer joins to make

What did we Learn?

- New data architecture
- Dimensional Modeling Schemas
- ETL build tool
- And that old dogs can learn new tricks
 - Data definitions
 - Dimensional structures
 - Reporting tools



A True Partnership



- Customized, not turn key
- Phytorion guided us to what's most useful
- Designed with us
 - Data elements we wanted
 - Unique derived fields created
 - Our priorities
 - Strategize to solve **our** problems with best practice expertise
- We had naive users and they were patient

What do True Partners Do?



- Don't impose a solution
- Have a true collaborative approach
- Provide thorough documentation and exceptional support



- Best of all.....

No black box

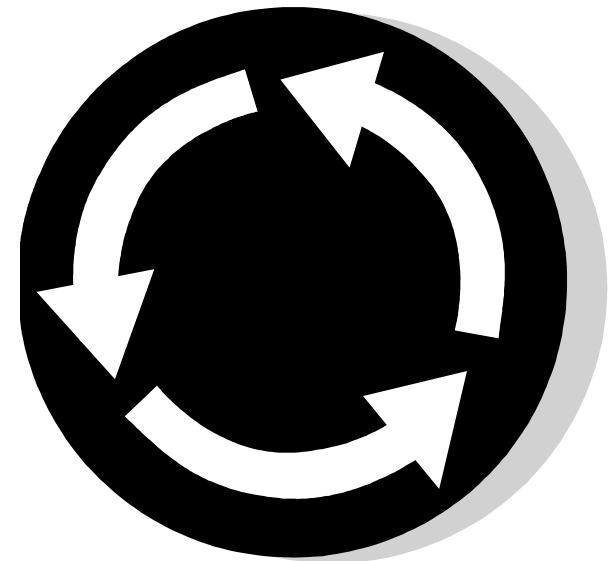
Because researchers like to know where the data is coming from



Project Roles



- Project Manager
 - Project liaison
 - High level needs
 - Set timelines and resources
- Research
 - Functional liaison
 - Build reports and test data
 - Verified output
- Phytorion
 - Business requirements
 - Star schemas
 - Developed ETL code
 - Training
- Information Technology
 - Hardware needs
 - Technical support



College Roles



- Executive sponsorship was critical
- Liaison from Executive Committee
- Project Team
 - Designers of project
 - Business Analysts
 - IT staff

Live Remote from California



What We've Done in the Last Year



- Identified data and built a data warehouse
- Deployed reports
- Deployed a complex enrollment management application
- Automated updates to projections
- Set up security based on role and area
- Building more complex reports and developing a Dashboard



Questions and Answers



Contact Information



- **Long Beach City College**

- **Dr. Linda Umbdenstock**

- Administrative Dean, Planning

- lumbdentock@lbcc.edu

- **Janice Miller**, Research Systems Analyst

- jmiller@lbcc.edu



- **Phytorion**

- **Yiorgos Marathias**, President

- Yiorgos.Marathias@Phytorion.com

- **Ashley Silverburg**, Chief Architect

- Ashley.Silverberg@Phytorion.com

