DECAL.org
DATABASE
FINAL
EDISON
NICOLE
STEVEN
ALLEN
JESSICA
AIME
KAIA
CLIENT OBJECTIVES

Familiarize students about the Decal course initiation process.

Generate customized reports about decal classes, rankings, and facilitators.

Reduce time and stress in Decal course initiation process.

Provide publicity for Decal in general.

Background | EER | Relational Schema | Normalization | Queries | Forms
DATABASE OBJECTIVES

Convenience

• Make information for Decal on course interest, demand, and rankings accessible in a database.

Efficiency

• Streamline course time-slot scheduling for Decal courses.

Intelligent Record

• Provide advanced functions and reports to:
  • Rank board member performance.
  • Prioritize Classes by "interest factor."
  • Rank categories of classes.

Background | EER | Relational Schema | Normalization | Queries | Forms
BEFORE: SIMPLIFIED EER

Cal_Student

First
Middle
Last
Name
ID
E-mail
Phone

Course

CCN
Title
Site
Semester
App
Class Size
Category
Units
description
Syllabus

Enrolled

Wait_Listed

Facilitates

Facilitator

DeCal_Student

(0,N)
t,o

(0,N)
DeCal_Student

Facilitates

Event

Location

Start_Date

Wait_Listed

Attending

Section

Equipment

Time

Sponsor

Section_ID

Time

Location

EID

Event

Location

Name

Equipment

Start_Date

Wait_Listed

(0,N)

(0,1)

(1,N)

(1,N)

(0,M)

(0,M)

(0,M)

(1,M)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)

(1,1)
RELATIONAL SCHEMA

Strong Entities

1. Course(CID, Cname, DeptID, Category, StartDate, Equipment, Units, Year-Semester, Description, Syllabus, AID, FID, EID)
2. Department(DeptID, Department_name)
3. Faculty_Sponsor(FID, DeptID, Fname, Lname)
4. CalStudent(SID, FName, Mname, Lname, DOB, Gender, Phone, Email)
4a Facilitator(SID)
4b DecalStudent(SID)
4c Volunteer(SID)
4d Board_Member(SID, position, Username, Password, Year-Semester, Number of Semesters Involved)
9. Event(Eid, EName, Location, EDate, StartTime, EndTime, Flyer)
9a DeCal Expo(EID)
9b Fundraiser(EID, Amount_Raised)
9c Outreach(EID)
9d Social(EID)
9e Internal(EID)
10. Sponsor(SpId, Date_added)
   10a Individual(SpId, FName, Lname, email, phone, Address, City, State, ZIP)
   10b Corporate(SpId, Name, Liaison, email, phone, Address, City, State, ZIP)
   10c StudentOrg(SpId, Name, Liaison, email, phone)
RELATIONAL
SCHEMA, CONT’D

11. Timeslot (TID, start_hour, date)
12. TablingShift (TsID, start_hour, date,
13. OfficeHours (OHid, day, Start_Time, End_Time, )
14. Advertising (Aid, start_date,
   14a WebMedia (Aid14, url)
      14aa WebPage (Aid14a, PageViews)
      14ab FacebookPage (Aid14a, PageName, Likes)
      14ac TwitterAcct (Aid14a, AccountName, Followers)
   14b Flyering (Aid14, flyer_attachment)
15. Survey (SuID, date_submitted, CID1,SectionID18)
   15a StudentSurvey (SuID15, how-did-you-hear-about-decal, year, major
   15b CourseEval (SuID15,Cname, facilFname, facilLname, year, major, courserating
   15c FacilitatorEval (SuID15, facilFname, facilLname, year, major, fac_rating
16. Building (BID, Name)
17. Budget (BudID, BudgetName, SID4, Amount)
18. CourseProposal(ProposalID, CID, SID4a, Year-Semester, DeanID19, SID4d, AS_Approval)
19. Dean(DeanID, DeptID2, Fname, Lname)
Weak Entities
20. Section(CID1, SectionID, Start_Time, End_Time, Weekdays, BID16, RoomNum20, SID4A)
21. Donation (SpID10, Donation_Date, Amount)
22. Room (BID16, RoomNum, Capacity, Equipment, Scheduler)
23. Account (BudID17, AccntID, Amount, Purpose)

N-M Relationships
24. F_facilitates_coursesection(SID4A, CID1, SectionID20)
25. DStudent_enrolled_coursesection(SID4B, CID1, SectionID20)
26. DStudent_waitlisted_coursesection(SID4B, CID1, SectionID20)
27. V_Works_Shift(SID4B, Eid9, TID11)
28. CStudent_attends_Event(SID4, EID9)
29. CStudent_takes_Survey(SID4, SuID9)
30. Event_Sponsored_By(EID9, SpID, Description)
31. BMember_Hosts_OH(SID4D, OHid13)
32. BMember_InChargeOf_Ads(SID4D, Aid14)
33. BMember_Attends_Tabling(SID4D, TsID12)
34. Donation_Placed_In_Account(SpID10, Donation_Date19, BudID17, AccntID21, Amount)
35. Building_Assoc_With_Dept(DeptID16, BID16)
36. FacilitatorSubmitsCProposal(ProposalID18, SID4A, Datesubmitted)
NORMALIZATION ANALYSIS

BCNF

F_facilitates_coursesecion(SID^{4A}, CID^{1}, SectionID^{20}, Cname, DeptID, Category, Fname, Mname, Lname)
CID -----> Cname, DeptID, Category (violates BCNF since LHS (CID) is not a superkey)
SID -----> Fname, Mname, Lname (violates BCNF since LHS (SID) is not a superkey)

Decompose into BCNF:
1. To deal with CID -----> Cname, DeptID, Category, decompose into :
   R0(SID^{4A}, CID^{1}, SectionID^{20}, Fname, Mname, Lname, ) and
   R1(CID, Cname, DeptID, Category)

2. To deal with SID -----> Fname, Mname, Lname , decompose into :
   R2( SID^{4A}, CID^{1}, SectionID^{20} )
   R3(SID, Fname, Mname, Lname)

So we end up with R1 , R2 , R3
2\textsuperscript{nd} Example: BCNF
DStudent_enrolled_coursesection\((\text{SID}^4\text{B}, \text{CID}^1, \text{SectionID}^{20}, \text{Cname,DeptID,Category,Start\_Time, End\_Time, Weekdays, BID}^{16}, \text{RoomNum}^{20}, \text{Fname, Mname, Lname})\)

\begin{itemize}
  \item CID, SectionID \rightarrow Start\_Time, End\_Time, Weekdays, BID, RoomNum (violates BCNF since LHS (CID, SectionID) is not a superkey)
  \item SID \rightarrow Fname, Mname, Lname (violates BCNF since LHS (SID) is not a superkey)
\end{itemize}

\textbf{Decompose into BCNF:}

\begin{itemize}
  \item To deal with CID, SectionID \rightarrow Start\_Time, End\_Time, Weekdays, BID, RoomNum
    \begin{align*}
      R_0&((\text{SID}^4\text{B}, \text{CID}^1, \text{SectionID}^{20}, \text{Cname,DeptID,Category,Fname,Mname, Lname})) \\
      R_1&((\text{CID,SectionID,Start\_Time, End\_Time, Weekdays, BID, RoomNum}))
    \end{align*}
  \item To deal with SID \rightarrow Fname, Mname, Lname
    \begin{align*}
      R_2&((\text{SID}^4\text{B}, \text{CID}^1, \text{SectionID}^{20}, \text{Cname,DeptID,Category})) \\
      R_3&((\text{SID,Fname, Mname, Lname}))
    \end{align*}
\end{itemize}

So we end up with \(R_1, R_2, R_3\)
<table>
<thead>
<tr>
<th>No.</th>
<th>Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Students want to see rankings... Classes ranked by “interest factor.”</td>
</tr>
<tr>
<td>2</td>
<td>I want to see the future demands of timeslots for classes in planning... Can you model the timeslot demand for classes?</td>
</tr>
<tr>
<td>3</td>
<td>I want to reward those that contribute the most... Track Board Member Involvement</td>
</tr>
<tr>
<td>4</td>
<td>I want to know reward the best facilitators... Track facilitator performance</td>
</tr>
<tr>
<td>5</td>
<td>Which departments offer the most or famous classes? Forecast interest and demand for categories of classes</td>
</tr>
</tbody>
</table>
No 1

Students want to see rankings to help sign-up for Decals

Classes Ranked by “Interest Factor”
Query 1: Classes Ranked by ‘Interest Factor’

What is the ‘Interest Factor?’

- Score that predicts the demand for a course
- Based on a weighted average of the enrollment, length of waitlist, webpage activity, and course survey responses
- Can be used by both DeCal board and Decal Students
Finding Enrollment Data

-SELECT Course.CID, Count(DStudent_Enrolled_CourseSection.SID) AS CountOfSID
FROM Course LEFT JOIN DStudent_Enrolled_CourseSection ON Course.CID = DStudent_Enrolled_CourseSection.CID
WHERE (((Course.[Year-Semester])=2013))
GROUP BY Course.CID;

Results in:

![Result Tables]

Background | EER | Relational Schema | Normalization | Queries | Forms
Finding Course Webpage Traffic

Run the following query to sum the number of pageviews on each course webpage during the month before registration:

```
SELECT Course.CID, Sum(Webpage.PageViews) AS SumOfPageViews
FROM Course LEFT JOIN Webpage ON Course.AID = Webpage.Aid
WHERE ((Webpage.[Start Date])>#1/1/2013#) AND ((Webpage.[End Date])<#2/2/2013#))
GROUP BY Course.CID;
```

Results in:
Finding the Average Course Rating

-SELECT Course.Cid, Avg(CourseEval.[Course Rating (1-10)]) AS [AvgOfCourse Rating]

FROM Course LEFT JOIN CourseEval ON Course.CID = CourseEval.CID

GROUP BY Course.Cid;

Results in:

![Average Course Ratings Table]

<table>
<thead>
<tr>
<th>Cid</th>
<th>AvgOfCourse Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>2443</td>
<td>10</td>
</tr>
<tr>
<td>2448</td>
<td>8</td>
</tr>
<tr>
<td>2455</td>
<td>9</td>
</tr>
<tr>
<td>2470</td>
<td>7</td>
</tr>
<tr>
<td>2596</td>
<td>8</td>
</tr>
</tbody>
</table>
Calculating the Interest Factor

1. Scale all field values to a 0-100 scale
2. Apply weights to each field
   - 10% Pageviews
   - 35% Number Waitlisted
   - 30% Number Enrolled
   - 25% Avg. Course Rating
3. Calculate the sum for each course!
Calculating the Interest Factor

<table>
<thead>
<tr>
<th>Cname</th>
<th>Page Views</th>
<th>Number Enrolled</th>
<th>Number Waitlisted</th>
<th>Avg Course Rating</th>
<th>Interest Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scientific Academic Presentations</td>
<td>20000</td>
<td>45</td>
<td>5</td>
<td>8</td>
<td>4.425</td>
</tr>
<tr>
<td>Bicycle Mechanics and Repair</td>
<td>35000</td>
<td>50</td>
<td>11</td>
<td>9</td>
<td>6.025</td>
</tr>
<tr>
<td>Radiolab: Science and Philosophy for the Masses</td>
<td>40000</td>
<td>27</td>
<td>0</td>
<td>7</td>
<td>2.96</td>
</tr>
<tr>
<td>Across Cultures and Time</td>
<td>25000</td>
<td>75</td>
<td>8</td>
<td>9</td>
<td>6.15</td>
</tr>
<tr>
<td>BookWorlds- Star Trek: Where No Course Has Gone</td>
<td>60000</td>
<td>30</td>
<td>3</td>
<td>8</td>
<td>4.025</td>
</tr>
</tbody>
</table>

Interest Factor

Interest Factor Breakdown

Background | EER | Relational Schema | Normalization | Queries | Forms
No. 2

I want to ensure decal is offering classes at optimum times...

Timeslot Demand Model
Query 2: Timeslot Demand Model/Optimization

How will this help?

- Common for Cal students to take multiple decal courses

- Selection of a course's timeslot done with little reference to other course timeslots

- Would act as a tool to encourage a wider spread of available class times for Decal as a whole
Part 1: A Weekly Graphical Representation

Given a set of Sections, construct a graph of total classes offered for each day of the week.

The graph's constraints are as follows:

- let the x axis be populated by 7 members, one for each day of the week
- let the y axis represent the number of decal courses offered on that day
Weekly Representation, example

Example data entries:

<table>
<thead>
<tr>
<th>SectionID</th>
<th>CID</th>
<th>Start_Time</th>
<th>End_Time</th>
<th>WeekDays</th>
<th>BID</th>
<th>RoomNumb</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2443</td>
<td>5:00:00 PM</td>
<td>6:30:00 PM</td>
<td>Tu</td>
<td>5</td>
<td>201</td>
</tr>
<tr>
<td>2</td>
<td>2448</td>
<td>4:30:00 PM</td>
<td>6:30:00 PM</td>
<td>Th</td>
<td>6</td>
<td>101</td>
</tr>
<tr>
<td>3</td>
<td>2455</td>
<td>6:30:00 PM</td>
<td>7:00:00 PM</td>
<td>W</td>
<td>9</td>
<td>105</td>
</tr>
<tr>
<td>4</td>
<td>1298</td>
<td>5:00:00 PM</td>
<td>7:00:00 PM</td>
<td>Th</td>
<td>10</td>
<td>1010</td>
</tr>
<tr>
<td>5</td>
<td>5555</td>
<td>6:00:00 PM</td>
<td>8:00:00 PM</td>
<td>F</td>
<td>5</td>
<td>555</td>
</tr>
<tr>
<td>6</td>
<td>6666</td>
<td>4:00:00 PM</td>
<td>6:00:00 PM</td>
<td>W</td>
<td>4</td>
<td>444</td>
</tr>
<tr>
<td>7</td>
<td>7777</td>
<td>7:00:00 PM</td>
<td>9:00:00 PM</td>
<td>M</td>
<td>2</td>
<td>222</td>
</tr>
<tr>
<td>8</td>
<td>8888</td>
<td>8:00:00 PM</td>
<td>9:00:00 PM</td>
<td>Su</td>
<td>1</td>
<td>111</td>
</tr>
<tr>
<td>9</td>
<td>9999</td>
<td>3:00:00 PM</td>
<td>4:30:00 PM</td>
<td>W</td>
<td>4</td>
<td>444</td>
</tr>
<tr>
<td>10</td>
<td>1010</td>
<td>1:00:00 PM</td>
<td>3:30:00 PM</td>
<td>W</td>
<td>2</td>
<td>222</td>
</tr>
<tr>
<td>11</td>
<td>1110</td>
<td>2:00:00 PM</td>
<td>4:00:00 PM</td>
<td>Tu</td>
<td>3</td>
<td>333</td>
</tr>
<tr>
<td>12</td>
<td>1212</td>
<td>3:00:00 PM</td>
<td>5:30:00 PM</td>
<td>Tu</td>
<td>4</td>
<td>444</td>
</tr>
</tbody>
</table>
Run the following query to group and aggregate based on number of classes per day:

-SELECT Weekdays, COUNT(*) FROM Section GROUP BY Weekdays

Results in:

<table>
<thead>
<tr>
<th>Weekdays</th>
<th>Expr1001</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>1</td>
</tr>
<tr>
<td>M</td>
<td>1</td>
</tr>
<tr>
<td>Su</td>
<td>1</td>
</tr>
<tr>
<td>Th</td>
<td>2</td>
</tr>
<tr>
<td>Tu</td>
<td>3</td>
</tr>
<tr>
<td>W</td>
<td>4</td>
</tr>
</tbody>
</table>
Weekly Representation, Graph

# Classes offered/day

Background | EER | Relational Schema | Normalization | Queries | Forms
Part 2: A Day-to-Day Graphical Representation

Given a set of Sections, construct a graph of total classes currently in session for a particular half-hour timeslot interval during a particular day

- let the x axis represent each half hour timeslot in a day i.e. 4:30-5:00, 7:00-7:30

- let the y axis represent the number of Decal courses in session at that time (not necessarily starting at that time)
Run the following query to aggregate the number of sections in session for each timeslot of a particular day:

```
SELECT Weekdays, COUNT(*)
FROM Section WHERE (Start_Time <= {..., 4:00, 4:30, 5:00, 5:30, 6:00, ...} AND End_Time > {..., 4:00, 4:30, 5:00, 5:30, 6:00, ...})
GROUP BY Weekdays;
```
Daily Representation,
Example Graph (Wednesday)
Query 3

I want to reward Board Members...

Track Board Member Involvement
Query 3: Track Board Member Involvement and Contribution

• How will this help?

• By having a numerical value of the contribution of Board members, DeCal will be able to see who has been doing more work than others and who has not been.

• This opens various doors for DeCal to create minimum quota a Board Member must meet or even giving awards to members who have contributed an excellent amount.

• This can also help DeCal board members see where they are lacking effort and contribution, so that the DeCal board can adjust accordingly.
Query 3: SQL Queries

### Semesters Involved

```sql
SELECT Board_MEMBER.[Number of Semesters Involved], Board_MEMBER.SID FROM Board_MEMBER;
```

### Office Hours

```sql
SELECT Count(OfficeHours.OHid) AS CountOHid, BMember_Hosts_OH.SID FROM OfficeHours INNER JOIN BMember_Hosts_OH ON OfficeHours.OHid = BMember_Hosts_OH.OHid;
```

### Advertising Projects

```sql
SELECT Count(BMember_InChargeOf_Ads.Aid) AS CountOfAid, BMember_InChargeOf_Ads.SID FROM BMember_InChargeOf_Ads GROUP BY BMember_InChargeOf_Ads.SID;
```

### Tableing and Tabling Shift

```sql
SELECT Count(TablingShift.TSID) AS CountOfTSID, BMember_Attends_Tabling.SID FROM TablingShift INNER JOIN BMember_Attends_Tabling ON TablingShift.TSID = BMember_Attends_Tabling.TSID GROUP BY BMember_Attends_Tabling.SID;
```

### EER | Relational Schema | Normalization | Queries | Forms
Query 3: Outputs

<table>
<thead>
<tr>
<th>Table</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tabling</strong></td>
<td><em>CountOfTsl</em> 2 RNA SID 22448593</td>
</tr>
<tr>
<td></td>
<td>1 RNA SID 22983505</td>
</tr>
<tr>
<td></td>
<td>3 RNA SID 24557457</td>
</tr>
<tr>
<td></td>
<td>2 RNA SID 25645744</td>
</tr>
<tr>
<td><strong>Semesters Involved</strong></td>
<td><em>Number Of Seme</em> 2 SID 22222222</td>
</tr>
<tr>
<td></td>
<td>3 SID 33323233</td>
</tr>
<tr>
<td></td>
<td>2 SID 34343434</td>
</tr>
<tr>
<td></td>
<td>2 SID 35265356</td>
</tr>
<tr>
<td></td>
<td>1 SID 35454540</td>
</tr>
<tr>
<td><strong>Advertising Projects</strong></td>
<td><em>CountOfAid</em> 1 SID 22345760</td>
</tr>
<tr>
<td></td>
<td>3 SID 22448593</td>
</tr>
<tr>
<td></td>
<td>2 SID 22945899</td>
</tr>
<tr>
<td></td>
<td>5 SID 22983505</td>
</tr>
<tr>
<td></td>
<td>4 SID 24557457</td>
</tr>
<tr>
<td></td>
<td>3 SID 25645744</td>
</tr>
<tr>
<td><strong>Office Hours</strong></td>
<td><em>CountOfOH</em> 2 RNA SID 22448593</td>
</tr>
<tr>
<td></td>
<td>5 RNA SID 22983505</td>
</tr>
<tr>
<td></td>
<td>6 RNA SID 24557457</td>
</tr>
<tr>
<td></td>
<td>2 RNA SID 25645744</td>
</tr>
<tr>
<td><strong>Board Member Meeting</strong></td>
<td><em>CountOfEID</em> 2 SID 22448593</td>
</tr>
<tr>
<td></td>
<td>5 SID 22983505</td>
</tr>
<tr>
<td></td>
<td>6 SID 24557457</td>
</tr>
<tr>
<td></td>
<td>2 SID 25645744</td>
</tr>
<tr>
<td><strong>Total Contribution</strong></td>
<td>SID 22448593 FName Phil LName Mansour</td>
</tr>
<tr>
<td></td>
<td>32.9 SID 24557457 FName Emmelyn LName Hsieh</td>
</tr>
<tr>
<td></td>
<td>24.3 SID 25645744 FName Aditya LName Kaulagi</td>
</tr>
</tbody>
</table>

Background | EER | Relational Schema | Normalization | Queries | Forms
Query 4

I want to know the best facilitators…

Track facilitator performance
Query 4: Track Facilitator Performance

Goal:

Analyze how facilitator/course ratings and facilitator involvement in the DeCal program affect total enrollment in the facilitator’s courses.

Justification:

Determines which factors significantly affect enrollment in DeCal classes so that board members can recommend areas that course facilitators may want to target to increase/maximize enrollment in their classes.
Query 4: Model

• Multiple Linear Regression on several inputs to see how number of sections taught, participation in Decal Expo, Course evaluation, and Facilitator evaluation affect course enrollment.

• Used Excel and the Data Analysis Toolpak to perform the necessary calculation.

• Determined all variables with a p-value < 0.10 to be an area that facilitators and board members should target when looking to increase enrollment.
Query 4: SQL Queries

**Enrollment in Facilitator’s DeCal Classes**

```sql
SELECT Count(DStudent_Enrolled_CourseSection.SID) AS CountOfSID, F_Facilitates_CourseSection.SID
FROM DStudent_Enrolled_CourseSection, F_Facilitates_CourseSection
WHERE ((((DStudent_Enrolled_CourseSection.SID) = F_Facilitates_CourseSection.SID) AND ((DStudent_Enrolled_CourseSection.CID) = F_Facilitates_CourseSection.CID)))
GROUP BY F_Facilitates_CourseSection.SID;
```

**CourseEvalResults**

```sql
SELECT Section.Facilitator, Facilitator.FName, Facilitator.LName, Count(Section.SectionID) AS CountOfSectionID
FROM Section, Facilitator
WHERE (((Section.Facilitator) = [Facilitator], [SID]))
GROUP BY Section.Facilitator, Facilitator.FName, Facilitator.LName;
```

**Participation in Expo**

```sql
SELECT Count(DeCal_Explo.EID) AS CountOfEID, Facilitator.SID
FROM DeCal_Explo, Facilitator, V_Works_Shift
WHERE (((V_Works_Shift.SID) = [Facilitator], [SID]) AND ((DeCal_Explo.EID) = [V_Works_Shift], [EID]))
GROUP BY Facilitator.SID;
```

**FacilitatorEvalResults**

```sql
SELECT Facilitator.SID, Avg(CourseEval.[Course Rating (1-10)]) AS [AvgOfCourse Rating (1-10)]
FROM CourseEval, Facilitator
WHERE (((Facilitator.FName) = CourseEval.FacilitatorFName) AND ((Facilitator.Lname) = CourseEval.FacilitatorLname))
GROUP BY Facilitator.SID;
```

**FacilitatorEvalResults**

```sql
SELECT Avg(FacilitatorEval.[Facilitator Rating (1-10)]) AS [AvgOfFacilitator Rating (1-10)], Facilitator.SID
FROM FacilitatorEval, Facilitator
WHERE Facilitator.FName = FacilitatorEval.FacilitatorFname AND Facilitator.LName = FacilitatorEval.FacilitatorLname
GROUP BY Facilitator.SID;
```
Query 4: Sample Access Implementation
Query 4: Sample Access Implementation

![CourseEvalResults](image)

### CourseEval
- SuID
- CName
- FacilitatorFname
- FacilitatorLname
- Year
- Major

### Facilitator
- SID
- FName
- MName
- LName
- DOB
- Gender

### Query Details
- **Field**: SID, Course Rating (1-10), FName, LName, [CourseEval].[Facilitator]
- **Table**: Facilitator, CourseEval
- **Total**: Group By, Avg
- **Sort**: Where
- **Show**: [ ]
- **Criteria**: [CourseEval].[Facilitator]
Query 4: Sample Access Implementation
Query 4: Sample Access Implementation

<table>
<thead>
<tr>
<th>Field</th>
<th>Table</th>
<th>Total</th>
<th>Sort</th>
<th>Show</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facilitator Rating</td>
<td>FacilitatorEval</td>
<td>Avg</td>
<td>Group By</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SID</td>
<td>Facilitator</td>
<td></td>
<td>Where</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FName</td>
<td>Facilitator</td>
<td></td>
<td>Where</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LName</td>
<td>Facilitator</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Query 4: Sample Access Implementation
Query 4: Query Output

<table>
<thead>
<tr>
<th># of sections taught</th>
<th>Enrollmen in Facilitator's DeCal Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>15839189 Kenneth Wilkins</td>
<td>11 15839189</td>
</tr>
<tr>
<td>21377777 Aime Ngongang</td>
<td>17 21377777</td>
</tr>
<tr>
<td>21555555 Jessica Lazarus</td>
<td>12 21555555</td>
</tr>
<tr>
<td>21733333 Steven Chua</td>
<td>18 21733333</td>
</tr>
<tr>
<td>21822222 Allen Fang</td>
<td>32 21822222</td>
</tr>
<tr>
<td>21911111 Nicole Fronda</td>
<td>14 21911111</td>
</tr>
<tr>
<td>23232323 Joshua Loth</td>
<td>10 23232323</td>
</tr>
<tr>
<td>31904810 Maria Torres</td>
<td>10 31904810</td>
</tr>
<tr>
<td>48592901 Oliver Craig</td>
<td>13 48592901</td>
</tr>
<tr>
<td>61028732 Jennifer Wu</td>
<td>8 61028732</td>
</tr>
<tr>
<td>93718382 Patrick Wells</td>
<td>26 93718382</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th># of sections taught</th>
<th>Participation in Expo</th>
</tr>
</thead>
<tbody>
<tr>
<td>15839189</td>
<td>2 15839189</td>
</tr>
<tr>
<td>21377777</td>
<td>2 21377777</td>
</tr>
<tr>
<td>21555555</td>
<td>2 21555555</td>
</tr>
<tr>
<td>21733333</td>
<td>3 21733333</td>
</tr>
<tr>
<td>21822222</td>
<td>3 21822222</td>
</tr>
<tr>
<td>21911111</td>
<td>1 21911111</td>
</tr>
<tr>
<td>23232323</td>
<td>1 23232323</td>
</tr>
<tr>
<td>31904810</td>
<td>1 31904810</td>
</tr>
<tr>
<td>48592901</td>
<td>1 48592901</td>
</tr>
<tr>
<td>61028732</td>
<td>1 61028732</td>
</tr>
<tr>
<td>93718382</td>
<td>2 93718382</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th># of sections taught</th>
<th>CourseEvalResults</th>
</tr>
</thead>
<tbody>
<tr>
<td>15839189</td>
<td>7 15839189</td>
</tr>
<tr>
<td>21377777</td>
<td>8 21377777</td>
</tr>
<tr>
<td>21555555</td>
<td>7 21555555</td>
</tr>
<tr>
<td>21733333</td>
<td>10 21733333</td>
</tr>
<tr>
<td>21822222</td>
<td>8 21822222</td>
</tr>
<tr>
<td>21911111</td>
<td>9 21911111</td>
</tr>
<tr>
<td>23232323</td>
<td>6 23232323</td>
</tr>
<tr>
<td>31904810</td>
<td>7 31904810</td>
</tr>
<tr>
<td>48592901</td>
<td>9 48592901</td>
</tr>
<tr>
<td>61028732</td>
<td>5 61028732</td>
</tr>
<tr>
<td>93718382</td>
<td>8.5 93718382</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th># of sections taught</th>
<th>FacilitatorEvalResults</th>
</tr>
</thead>
<tbody>
<tr>
<td>15839189</td>
<td>9 15839189</td>
</tr>
<tr>
<td>21377777</td>
<td>7 21377777</td>
</tr>
<tr>
<td>21555555</td>
<td>7 21555555</td>
</tr>
<tr>
<td>21733333</td>
<td>10 21733333</td>
</tr>
<tr>
<td>21822222</td>
<td>8 21822222</td>
</tr>
<tr>
<td>21911111</td>
<td>9 21911111</td>
</tr>
<tr>
<td>23232323</td>
<td>6 23232323</td>
</tr>
<tr>
<td>31904810</td>
<td>7 31904810</td>
</tr>
<tr>
<td>48592901</td>
<td>9 48592901</td>
</tr>
<tr>
<td>61028732</td>
<td>5 61028732</td>
</tr>
<tr>
<td>93718382</td>
<td>8.5 93718382</td>
</tr>
</tbody>
</table>

Background | EER | Relational Schema | Normalization | Queries | Forms
Query 4: Regression in Excel

Enrollment = -17.26 + 16.44(CountofSectionID) + 0.18(CountofEID) + 0.60(AvgofCourseRating) + 1.16(AvgOfFacilitatorRating)
Query 4: Best-fit Regression

Enrollment = -17.42 + 16.37(CountofSectionID) + 0.67(AvgofCourseRating) + 1.16(AvgOfFacilitatorRating)
Query 4: Output and Conclusion

Ran the queries, exported the data to Excel

Performed a multiple linear regression

Found a better fit when the CountofEID variable was omitted and therefore did not include this in our final regression equation

Variables with p-values less than 0.10 have a significant effect on enrollment in DeCal classes.

A smaller p-value suggests a greater impact on enrollment, so the following is an ordered ranking of the factors we analyzed:

- Number of sections
- Facilitator rating
- Course rating
How many sports courses have been taught since Fall 2011?

Query 5

No 5

Which Department offers the Most and In-demand Classes...

Forecast Demand in Categories of Classes.
Query 5: Forecast Demand in categories of classes

Justification:

Predicting enrollment for a category of courses can help DeCal determine what kinds of classes they may want to further advertise or what kinds of classes they should be more inclined to approve in order to boost enrollment and overall student involvement in DeCal.
Query 5: Forecast Demand in categories of classes

Total Enrollment in a class category = B_1 + B_2(Avg facilitator rating in a class category) + B_3(Avg course rating in a class category)

Perform a Multiple Linear Regression on these inputs to see how they affect course enrollment of a course category. Use Excel and StatPlus to perform the necessary calculations to predict enrollment in a certain category.
Query 5: SQL

Course_eval score group by course_category and year
SELECT C.category, CE.year, AVG(CE.course_rating) as AvgofCourse_rating
FROM CourseEval CE, Survey S, Course C
WHERE CE.SUID = S.SUID AND C.CID = S.CID
GROUP BY C.Category, CE.year;

Facilitator_eval score group by course_category and year
SELECT C.category, CE.year, AVG(FE.fac_rating) AS AvgFacilitatorRating
FROM FacilitatorEval FE, Survey S, Course C, CourseEval CE
WHERE FE.SUID = S.SUID AND S.CID = C.CID
GROUP BY C.Category, CE.year;

Enrollment (Spring or Fall) group, course and year
SELECT COUNT(SID) AS Enrollment, C.category, CE.year
FROM Dstudent_enrolled_coursesection SE, Course C, CourseEval CE
WHERE SE.CID = C.CID
GROUP BY C.Category, CE.year;
Query 5: Sample Access Implementation

![Diagram of database query and implementation](image)
Query 5: Access Implementation

<table>
<thead>
<tr>
<th>Field</th>
<th>SE</th>
<th>C</th>
<th>CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrollment: SID</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Table: SE</td>
<td>SID, CID, SectionID</td>
<td>Cname, CID, Category, DeptID, StartDate, Equipment</td>
<td>CID, SuID, Facilitator, StudentYear, Year, Major</td>
</tr>
<tr>
<td>Total: Count</td>
<td>Group By</td>
<td>Group By</td>
<td>Where</td>
</tr>
<tr>
<td>Sort:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Show:</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Criteria:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[C],[CID]</td>
</tr>
</tbody>
</table>
Query 5: Access Implementation
Query 5: OUTPUT

Course_eval score group by course_category and year

<table>
<thead>
<tr>
<th>category</th>
<th>Year</th>
<th>AvgOfcourse_rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture</td>
<td>2007</td>
<td>8.5</td>
</tr>
<tr>
<td>Culture</td>
<td>2008</td>
<td>8</td>
</tr>
<tr>
<td>Culture</td>
<td>2009</td>
<td>7.5</td>
</tr>
<tr>
<td>Culture</td>
<td>2010</td>
<td>7.5</td>
</tr>
<tr>
<td>Geography</td>
<td>2008</td>
<td>7.5</td>
</tr>
<tr>
<td>Geography</td>
<td>2009</td>
<td>7</td>
</tr>
<tr>
<td>Geography</td>
<td>2010</td>
<td>7</td>
</tr>
<tr>
<td>Geography</td>
<td>2012</td>
<td>9</td>
</tr>
<tr>
<td>History</td>
<td>2008</td>
<td>8</td>
</tr>
<tr>
<td>History</td>
<td>2009</td>
<td>7.5</td>
</tr>
<tr>
<td>History</td>
<td>2010</td>
<td>9.5</td>
</tr>
<tr>
<td>History</td>
<td>2011</td>
<td>9</td>
</tr>
<tr>
<td>History</td>
<td>2012</td>
<td>7</td>
</tr>
<tr>
<td>Physics</td>
<td>2008</td>
<td>6</td>
</tr>
<tr>
<td>Physics</td>
<td>2009</td>
<td>4</td>
</tr>
<tr>
<td>Physics</td>
<td>2011</td>
<td>5.3333333333333333</td>
</tr>
<tr>
<td>Physics</td>
<td>2012</td>
<td>9</td>
</tr>
<tr>
<td>Science</td>
<td>2008</td>
<td>6.5</td>
</tr>
<tr>
<td>Science</td>
<td>2009</td>
<td>8</td>
</tr>
<tr>
<td>Science</td>
<td>2011</td>
<td>5.5</td>
</tr>
<tr>
<td>Science</td>
<td>2012</td>
<td>7</td>
</tr>
</tbody>
</table>
Query 5: OUTPUT

Facilitator_eval score group by course_category and year
Query 5: OUTPUT

Enrollment (Spring or Fall) group, course and year
Query 5: OUTPUT

Enrollment (Spring or Fall) group, course and year

Background | EER | Relational Schema | Normalization | Queries | Forms
Query 5: OUTPUT

Using SQL output for the above queries, we imported the following data into Excel and organized it in this manner.

<table>
<thead>
<tr>
<th>Culture</th>
<th>Avg Facilitator Rating (x1)</th>
<th>Avg Course Rating (x2)</th>
<th>Enrollment (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>7</td>
<td>8.5</td>
<td>6</td>
</tr>
<tr>
<td>2008</td>
<td>7</td>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td>2009</td>
<td>7</td>
<td>7.5</td>
<td>24</td>
</tr>
<tr>
<td>2010</td>
<td>7</td>
<td>7.5</td>
<td>18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Geography</th>
<th>Avg Facilitator Rating (x1)</th>
<th>Avg Course Rating (x2)</th>
<th>Enrollment (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>8.428571429</td>
<td>7.5</td>
<td>9</td>
</tr>
<tr>
<td>2009</td>
<td>8.428571429</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>2010</td>
<td>8.428571429</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>2012</td>
<td>8.428571429</td>
<td>9</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Physics</th>
<th>Avg Facilitator Rating (x1)</th>
<th>Avg Course Rating (x2)</th>
<th>Enrollment (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>8.125</td>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>2009</td>
<td>8.125</td>
<td>4</td>
<td>32</td>
</tr>
<tr>
<td>2011</td>
<td>8.125</td>
<td>5.333333333</td>
<td>24</td>
</tr>
<tr>
<td>2012</td>
<td>8.125</td>
<td>9</td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Science</th>
<th>Avg Facilitator Rating (x1)</th>
<th>Avg Course Rating (x2)</th>
<th>Enrollment (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>6.666666667</td>
<td>6.5</td>
<td>18</td>
</tr>
<tr>
<td>2009</td>
<td>6.666666667</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>2011</td>
<td>6.666666667</td>
<td>5.5</td>
<td>12</td>
</tr>
<tr>
<td>2012</td>
<td>6.666666667</td>
<td>7</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>History</th>
<th>Avg Facilitator Rating (x1)</th>
<th>Avg Course Rating (x2)</th>
<th>Enrollment (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>6.571428571</td>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td>2009</td>
<td>6.571428571</td>
<td>7.5</td>
<td>24</td>
</tr>
<tr>
<td>2010</td>
<td>6.571428571</td>
<td>9.5</td>
<td>18</td>
</tr>
<tr>
<td>2011</td>
<td>6.571428571</td>
<td>9</td>
<td>18</td>
</tr>
</tbody>
</table>
Query 5: OUTPUT

We then performed linear estimation in Excel to come up with a Regression Equation and Analysis of Variance (ANOVA) for the Culture Course Category:

\[
\text{Enrollment (Y) = 117.5455 - 12.5455 \times \text{Avg Course Rating} (x2)}
\]

<table>
<thead>
<tr>
<th>ANOVA</th>
<th>d.f.</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>108.20455</td>
<td>108.20455</td>
<td>1.4375</td>
<td>0.35333</td>
</tr>
<tr>
<td>Residual</td>
<td>2</td>
<td>150.54545</td>
<td>75.27273</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>258.75</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As you can see, Average Course Rating does have an effect on Enrollment for courses in a category, at least in the case of courses within the Culture Category.

Note, however that our data output does not show any variance for Average Facilitator Ratings, so right now it appears to not have any effect on Enrollment.

However, if our results for the Culture Category looked like this:
Then we can extrapolate the following Regression Equation and ANOVA:

\[
\text{Enrollment (Y) = 129.0000 - 0.6667 \times \text{Avg Facilitator Rating (x1)} - 13.3333 \times \text{Avg Course Rating (x2)}
\]

<table>
<thead>
<tr>
<th>Culture</th>
<th>Avg Facilitator Rating (x1)</th>
<th>Avg Course Rating (x2)</th>
<th>Enrollment (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>7</td>
<td>8.5</td>
<td>6</td>
</tr>
<tr>
<td>2008</td>
<td>8</td>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td>2009</td>
<td>7.5</td>
<td>7.5</td>
<td>24</td>
</tr>
<tr>
<td>2010</td>
<td>9</td>
<td>7.5</td>
<td>18</td>
</tr>
</tbody>
</table>

Can then use this equation and input values for that year’s average facilitator rating and average course rating in the culture category to predict total enrollment for courses in the culture category.
Future Work

- **Improve Queries**
  - Add more graphics and functionality

- **Add more forms**
  - Add form for new course proposal
  - Add form for new budgets

- **Automation**
  - Provide macros to automate class recording into database.
Special Thanks to:

Emmelyn Hsieh, Decal Board

Prof. Ken Goldberg

Animesh Garg
THANK YOU | Questions?