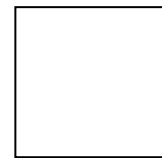




LAB 10

Subject 308-471 Data Warehousing and Data Mining

Topic Building a data warehouse



ชื่อ – นามสกุล (Name and surname)	รหัสนักศึกษา (Student ID)
Koranat Leelertpanyakul	5810210014
Napa Theppalak	5810210209
Pongsakorn Kaewnopparat	5810210297
Wipanee Sornloy	5810210912

1. วัตถุประสงค์ของปฏิบัติการ

1.1. เพื่อให้นักศึกษาเข้าใจการสร้างคลังข้อมูลและเหมืองข้อมูล

Objective: To understand the Data Warehousing and Data Mining.

2. คำสั่งสำหรับปฏิบัติการ

2.1 ให้นักศึกษาปฏิบัติการ ตามหัวข้อปฏิบัติการเป็น รายบุคคล

Lab instruction: Please read the exercise carefully and do it together to your group.

3. คำสั่งสำหรับจัดทำรายงานหรือเอกสาร (Submitting report instruction)

3.1. Please send electronic document by write your name, surname, and student ID in the document and submit to LMS2 of 308-471.

3.2. Please rename filename to 308-471_L10_ and then following by the last 3 digits of your student ID. For example, 308-471_L10_999_998_997_996.

3.3. Only accept the Microsoft Word file.

3.4. The deadline for this exercise is on Mar 21, 2017, 23:59.

Exercise 1:

1. Please create a group of four people.
2. Select one of your senior Projects that fits for doing a Data Warehouse Model.
3. Imagine that you are the people in a company of Project extending 10 years later and want to do Data warehouse. Please analyse as follows.
 - a. Brief your project no more than 100 words or 10 lines.

Predict the drop out of student Science faculty PSU

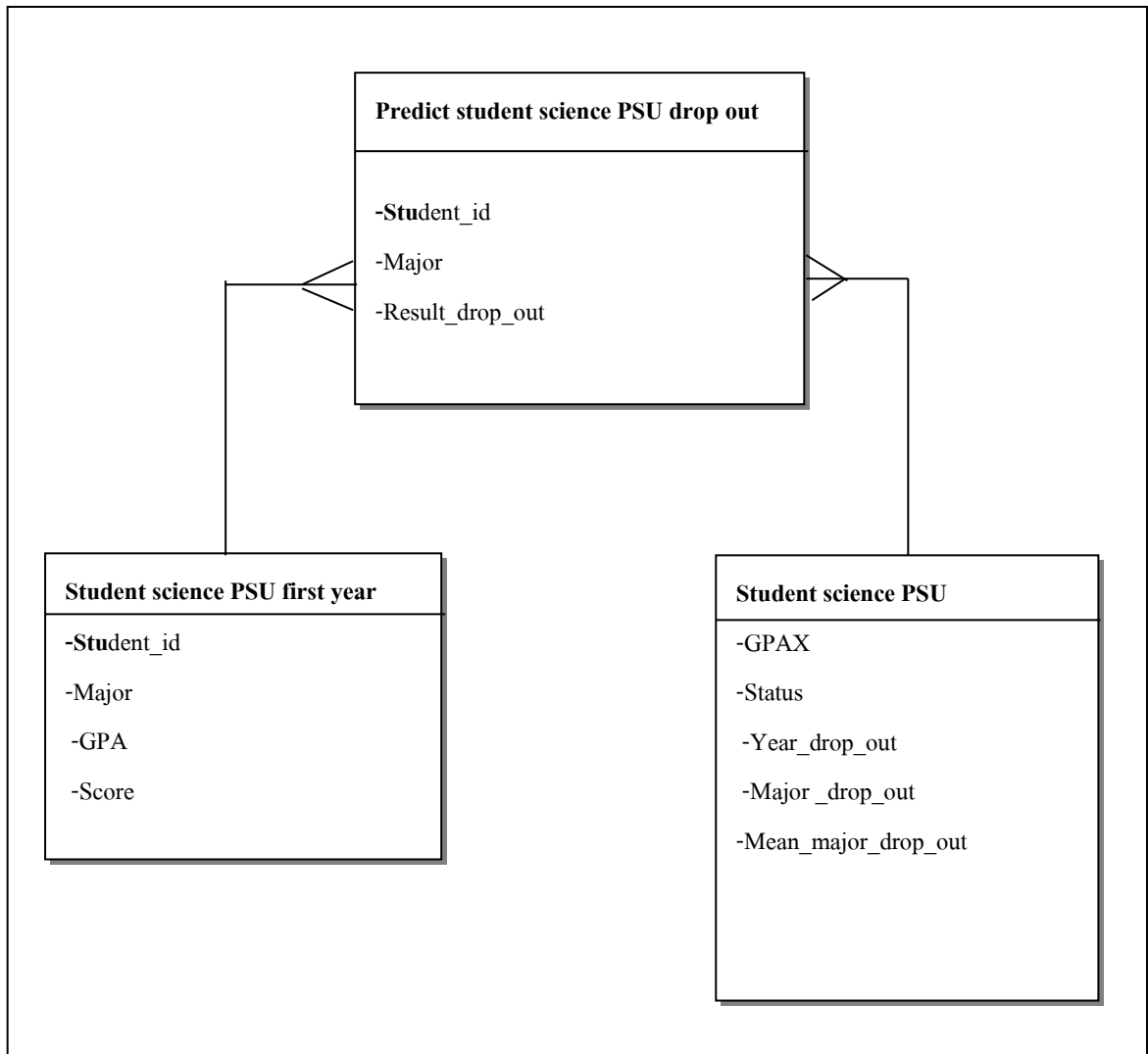
Towards, Science faculty, Prince of Songkhla University is the large faculty. Which have 13 majors and every years have many students drop out. The most students are drop out after they are select majors. So we bring education data mining apply to create model to predict the drop out of student science faculty Prince of songkla university and show predictive effect in the form of scatter plot graphs for develop education and find solutions to these problems.

- b. Details of your information that are suitable for data warehouse feature. (Subject Oriented, Integrated, Time Variant, and Non-volatile.)

- Student ID,
- Major,
- GPAX of 5 main courses
- Score each 5-main course
- GPAX of student
- Education Status,
- Academic year to drop out
- Amount of student that drop out each major
- Mean value of each major that drop out

We bring these information storages to data were house by use information since 2015 – 2017 (3 years).

- c. Analyse the multi-dimensional data model that you can choose between Star schema, Snowflake schema, and Constellation schema. Explain by entity-relationship diagram.



- a. Please assume that you are a group of manager. What are the top 3 subjects of Data Warehouse reports (along 10 years of your business) that you want to see?

- Can bring data mining technique forecast student academic achievement
- Can bring a data mining system that assists students in selecting the appropriate major according to their profile and in their course registering.
- Predict career for student that graduate.