**Phase 1**

**The summarized requirements from Prof:**

propose and understand an appropriate project topic, identify its main data management needs, explore and motivate its potential for interesting queries, and analyze the needed user functionality.

1. What are the main functions that the web-based user interface should provide?
2. How do the different functions work together? Sometimes there are dependencies between different functions.
3. Which real-world data are needed to support the functions identified before?
4. Can such real-world data be found in the Internet?
5. What (colloquial) queries are important for the application?
6. Which public domain and/or proprietary software is needed to perform the task? (The database system used must be CISE Oracle.)

(PDF file) that presents a clear and structured description and motivation of the project topic and requirements that the software solution should later fulfil.

The focus of this project is supposed to be on the database part and not so much on the application part. focus on database queries that evaluate large volumes of stored data. Of course, the application part must be highly functional, and the different user functions must cooperate nicely together. However, a fancy layout design of the user interface is not required but appreciated.

It is important that each group demonstrates in their deliverable that their application would really benefit from database support and that new information (such as trends) can be derived from the stored data. A simple retrieval of data from the database (that is, search) or the pure connection of different tables (that is, joins) are not sufficient. As an example, let us assume that a group selects a sales application as their project topic and stores many daily sales numbers in their database. Of course, one can search for sales data of interest in the database and display them in the user interface. But searching only identifies an interesting subset of all data stored in the database. DBMS are specialized for search tasks, and the respective SQL queries are relatively simply structured. This project aims at more interesting queries that, first, derive new information which is not explicitly stored in the database but can be derived from the data in the database by computations and, second, represent trends (see Section 3).

Arunabh-

Since we are required to do trend analysis and compute the changes in data over a period, based on the datasets, I could think of the following:

For point 5, some of the colloquial queries could be:

* What was the average number of books read during a certain period, ordered by genre?
* Which genres had the highest ratings during a certain period?
* Which genres received more reviews during a certain period?
* What was the percent of shelved books by genre?
* What was the ratio of read and shelved books by genre?