Generic Data Analyzer for Stock Data Analysis

Software Requirements Specification

Version 1.1

Team Guide: Prof. Lydia Jane
Members: Raj Biswas (12BCE0567)
Kunal Deb (12BCE0018)
Namratha Pratipatti (12BCE0050)
College Name: VIT University
Department: B.Tech CSE

State: Tamil Nadu

Table of Contents

Description **Page**

1. INTRODUCTION

1.1 Purpose

To build a web based application that will take huge amounts of stock data and give the current stock data and trend analysis of stock data to the user. The application should have the facility to specify the format if input structures of the data and adapt accordingly. Provide functionalities to access distinct elements/top hitters or frequent occurrences/occurrences in a specified range which can be dates or values for example. The analyser must be able to understand new queries (in the predefined format), be scalable and be able to access data spread across machines. It should use distributed data stores like Hadoop and use advanced techniques like Map Reduce. Should have the ability to tag unique data sets. Response time should be reasonable and provide detail analysis of different task executions done.

1.2 Scope

- There are 3 basic users – Registered User, Broker, Admin.
- Any user can register for the stock monitoring process. Each registered user will be given his/her own customized space.
- Depending on space, a registered user can select stocks to monitor and the time duration for monitoring them.
- During the duration specified the stocks selected by the user will be stored and analyzed. Important updates will be automatically shown as alerts to the user.
- Messages can be sent and received by both registered members and brokers.
- Admin oversees functioning of all activities.
- Admin has authority to reject or cancel user monitoring requests in case of any anomalous behavior of the system or if the user has requested for it.
- Admin can updates the bulletin that is read by all members for any notifications.
- Anything not related to the functionalities mentioned above is not considered to lie within the scope of this document.

1.3 Process Model

The process model used in the given project is a fusion of **RAD** (Rapid Application Development) and **Incremental build model**.

Justification

The different modules of the project are divided in such a way that some of them are independent of each other and thus can be implemented simultaneously (i.e. RAD model). In this model each module is designed, implemented, tested and then integrated to form the final result. While there are some modules which are dependent on other modules and thus cannot be implemented till the prior is implemented and tested (i.e. Incremental model). Thus, we have chosen a fusion of the above two models for our project.

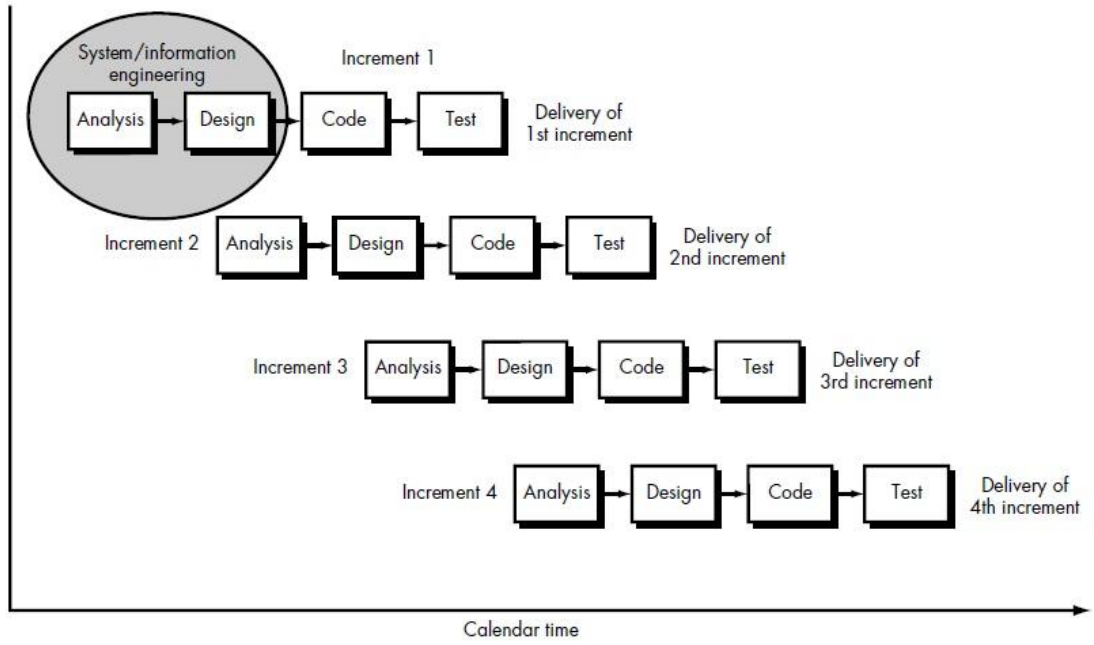


Fig 1.Diagrammatic Representation of Incremental model

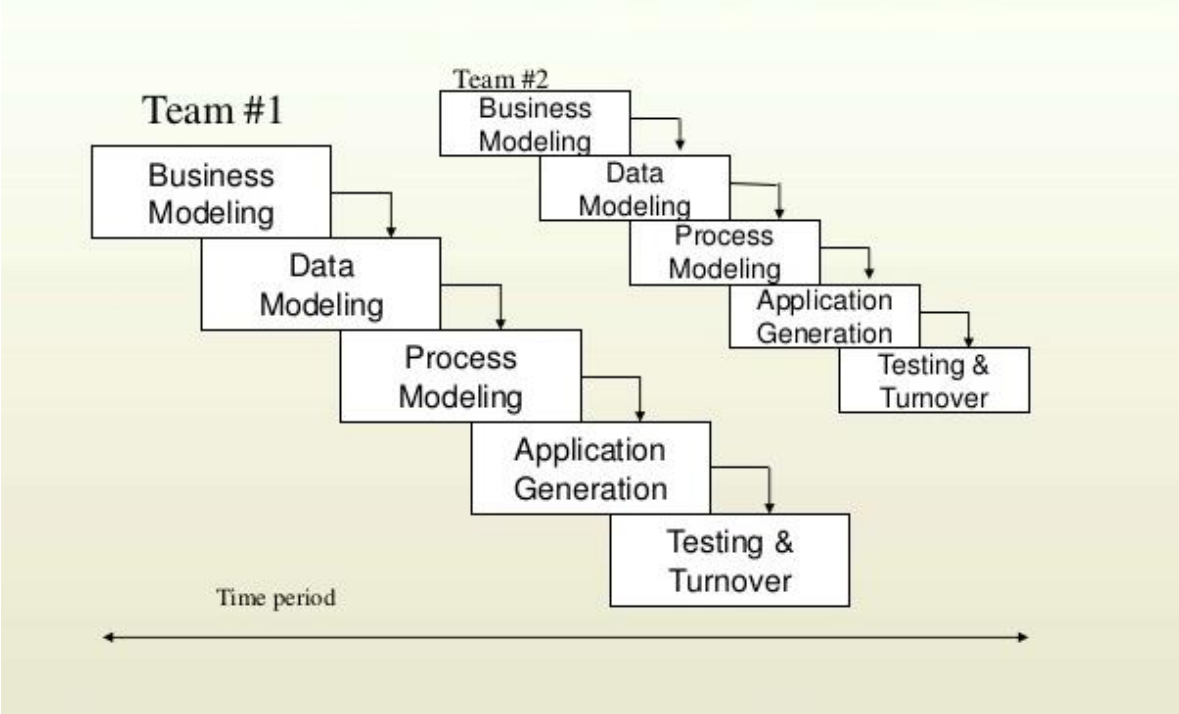


Fig2. Diagrammatical Representation of RAD model

1.4 Definitions, Acronyms, Abbreviations

GSDA

Generic Stock Data Anlayser – It is a web application that will monitor stock data and analyze it.

Admin

Administrator – Has the authority to oversee functionings of all activities and reject request of stock monitoring of any user.

RU

Registered User- Can register to monitor any stock data available in the stock market depending on limit on space provided to user.

BROKER

Can interact with the application similar to the users with the added functionality of maintaining separate records for separate clients.

WAMP

Windows Apache MySQL PHP – It is an application server platform

HTML5

HyperText Markup Language – It will be used to create web pages

CSS

Cascading Style Sheets – It will be used to style web pages

JSP

Java Server Pages- It is used to create dynamic web content.

PHP

PHP: Hypertext Preprocessor- It is a server-side scripting language

UML

Unified Modeling Language-is a standard language for writing software blueprints. The UML may be used to visualize, specify, construct and document

HTTP

Hypertext Transfer Protocol-It's a service protocol.

HDFS

Hadoop Distributed File System- It is the distributed file system in which the stock data will be stored.

MR

Map Reduce- It is the technique used to run algorithms in a parallel distribution manner in the HDFS.

1.5 TOOLS USED

NETBEANS 8.0.2

NetBeans is an integrated development environment (IDE) for developing primarily with Java, but also with other languages, in particular PHP, C/C++, and HTML5. It is also an application platform framework for Java desktop applications and others.
The NetBeans IDE is written in Java and can run on Windows, OS X, Linux, Solaris and other platforms supporting a compatible JVM. It allows applications to be developed from a set of modular software components called *modules*.

WAMP

WampServer is a Windows web development environment. It allows us to create web applications with Apache2, PHP and a MySQL database. PhpMyAdmin allows you to manage easily our databases.

WEB DEVELOPMENT TOOLS

To create and style webpages, HTML5 and CSS will be used. To script pages PHP and Javascript will be used.

1.6 REFERENCES

- Software Engineering, Seventh Edition, Ian Sommerville.
- SWEBOK V3.0, Guide to Software Engineering Body of Knowledge
- Wikipedia: http://en.wikipedia.org
- Netbeans: https://netbeans.org
- Database Management Systems- Navathe

1.7 TECHNOLOGIES TO BE USED

- Web Designing languages: HTML5, CSS
- Scripting Languages: JS, PHP, Python
- PHP Myadmin
- WAMP
- Rational Rose Software Modeler
- Microsoft Visio
- Hadoop
- Map Reduce

2. OVERALL DESCRIPTION

2.1 Product Perspective

The product aims at implementing a secure access of confidential data (user's details) for which SSL can be used. Better component design to get better performance at peak time. Flexible service based architecture will be highly desirable for future.

The product should be very easy to operate and understand. It will also include an alerting system (email or dashboard alerts) if any abnormal behavior detected in the data analyzed.

2.2 Software Interface

Software interfaces provide access to computer resources (such as memory, CPU, storage, etc.) of the underlying computer system.

- **Client on Internet-**Web Browser, Operating System (any)
- **Client on Intranet-** Web Browser, Operating System (any)
- **Web Server-** Apache, Operating System (any)
- **Data Base Server-** MySQL, Operating System (any)
- **Development End-** Web Designing Tools (PHP, Java, JavaScript, HTML, CSS, XML, AJAX, MySQL), OS (Windows),Apache (Web Server).

2.3 Hardware Interface

Minimum requirements:

CLIENT SIDE			
	Processor	RAM	Disk Space
Internet Explorer-6	Intel Pentium III or AMD 800 MHz	128 MB	100 MB

SERVER SIDE			
	Processor	RAM	Disk Space
Web Designing Tools	Intel Pentium III or AMD 800 MHz	1 GB	3.5 GB
MySQL		256 MB	500 MB(Excluding Data Size)

Recommended Requirements:

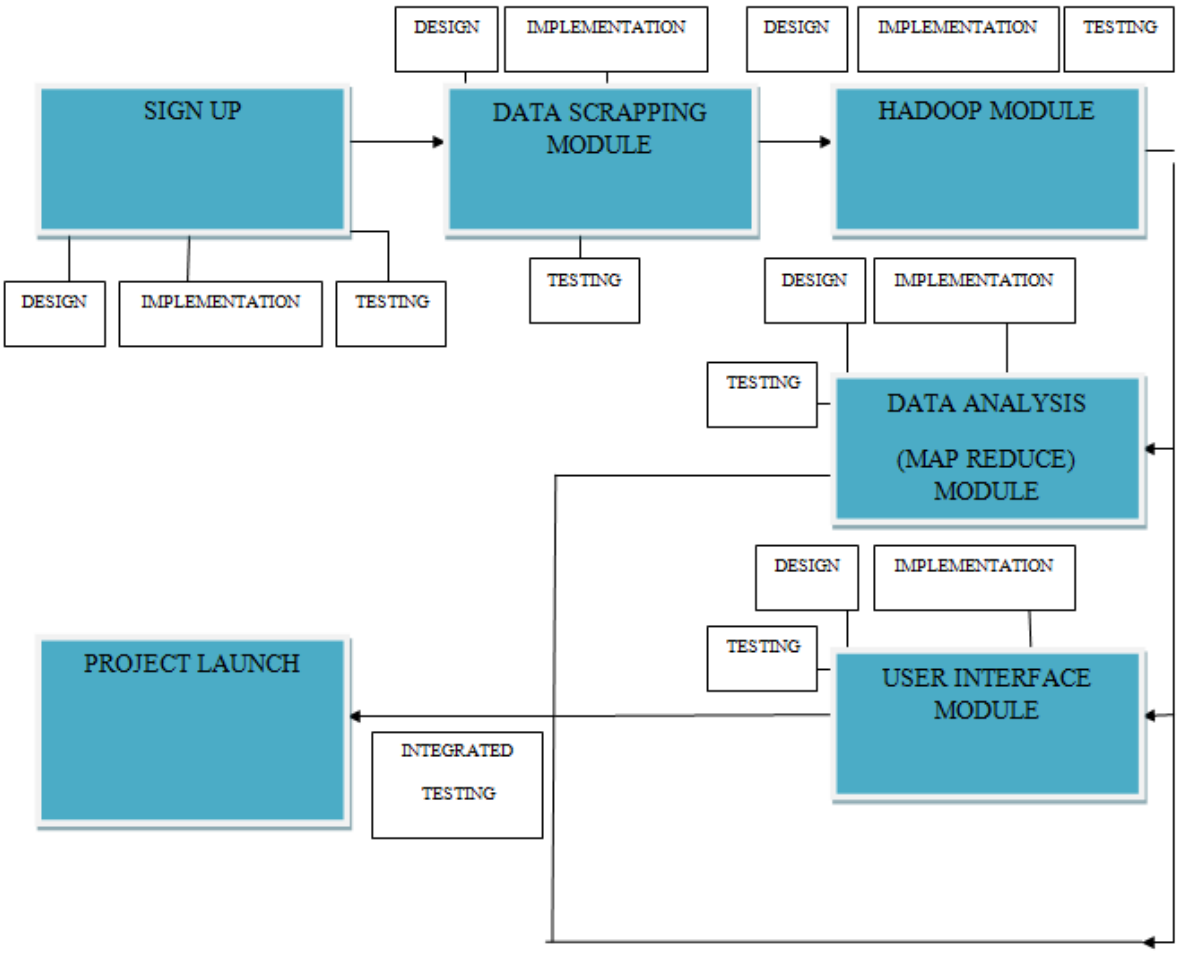
CLIENT SIDE			
	Processor	RAM	Disk Space
Internet Explorer-6	All Intel or AMD - 1 GHZ	256 MB	100 MB

SERVER SIDE			
	Processor	RAM	Disk Space
Web Designing Tools	All Intel or AMD - 2 GHZ	2 GB	3.5 GB
Internet Explorer-6		512 MB	500 MB(Excluding Data Size)

2.4 Constraints

- GUI is only in English.
- Login and password is used for the identification of users.
- This system is working for single server.
- Since the stock markets are not open 24 x 7, the system will not be able to provide data analysis and alerts throughout the day.
- The stock data is available for only those markets which provide an API for directly accessing the current stock data.
- Since the amount of storage given to a user is limited, the number of stocks that can be monitored at a time is limited.

2.5 Process Model Diagram



3.SPECIFIC REQUIREMENTS

3.1 Use Case Report:

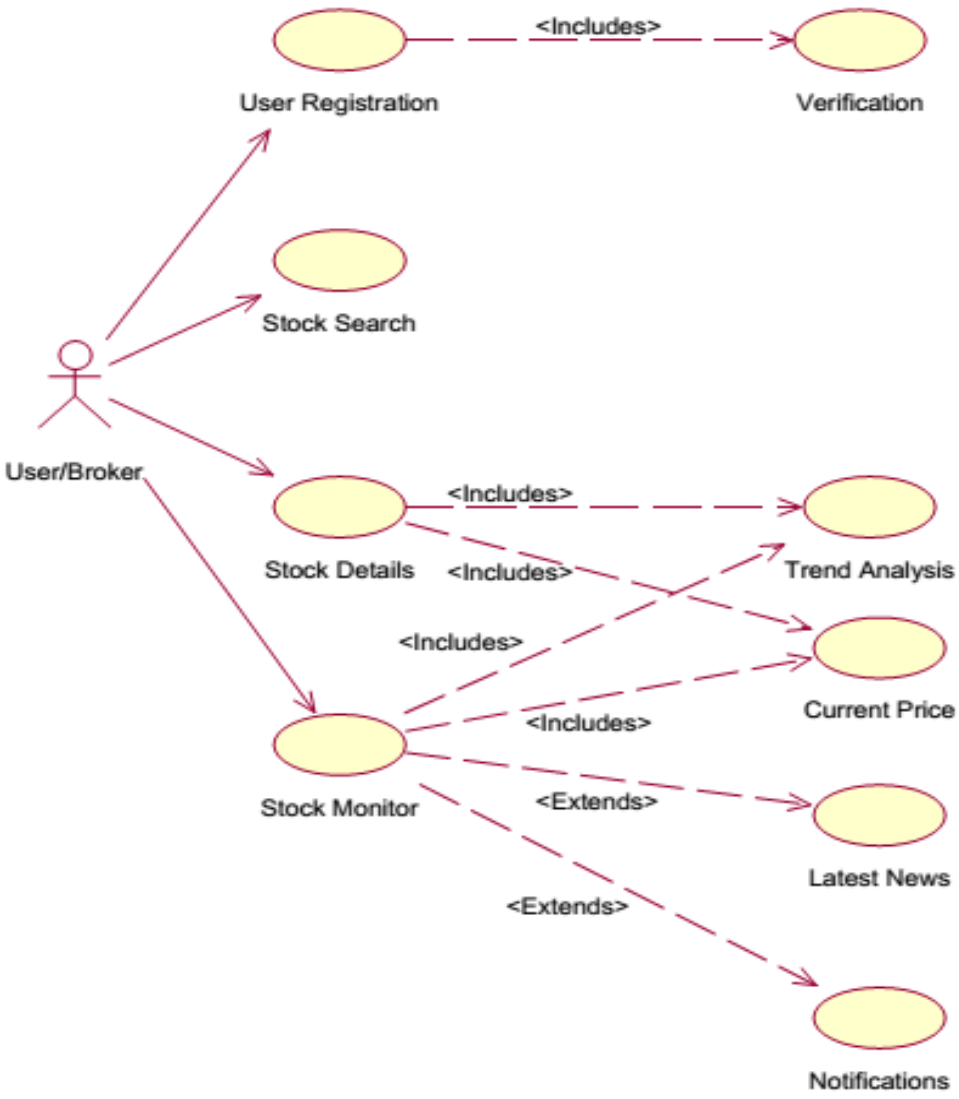


Fig3. Use case Diagram

USE CASE	DESCRIPTION
Registration	The member can register in order to avail the facilities.
Log in (Verification)	Every registered member can access his/her own profile containing personal details.
Stock Search	The registered users can search for specific stocks.
Stock Details	The registered users can view the details of the stock.
Stock Monitor	The users can add stocks for monitoring.
Trend Analysis	The selected stocks are analyzed for trends and viewed.
News Bulletin	The user can view the news bulletin of the stocks added for monitoring.
Borrow item	Any user can take the equipments if its available.
Current Price	Users can check the current prices of the stocks selected for monitoring.
Notifications	The user is given alert notifications if any abnormal trend is observed in the data in the dashboard.
Manage activity	Trainers have full control over the activities.
File complain against member	Trainers can file complain against members registered under him.

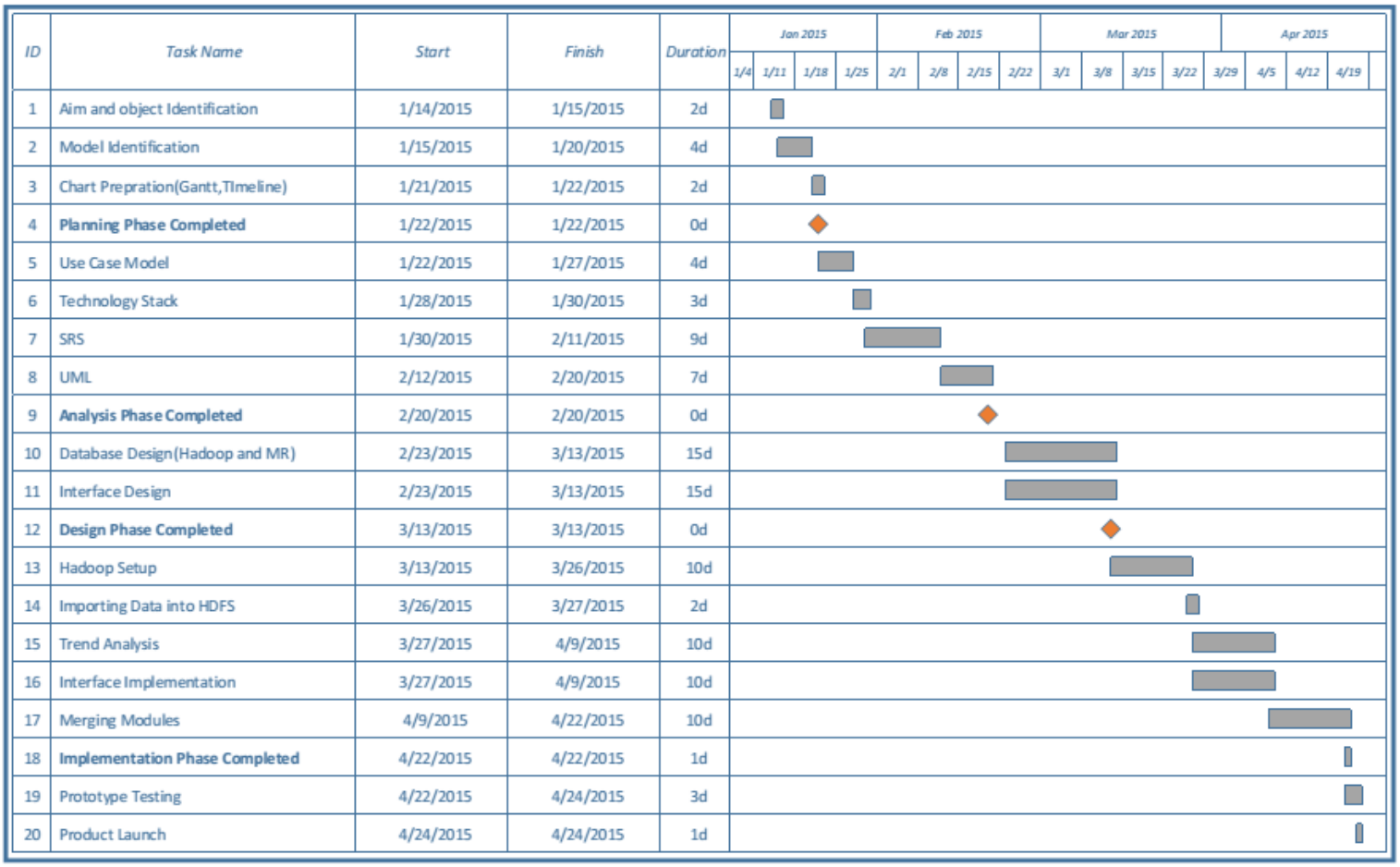
3.2 Schedules and Charts

Major Project Milestones

Activity	Dependency	Duration in Days	Estimated Start Date	Estimated End Date
Aim and Objective Identification (A)	-	2	14-Jan-2015	15-Jan-2015
Project Planning (B)	A	4	15-Jan-2015	18-Jan-2015
Process Model Identification and Scheduling (C)	B	2	18-Jan-2015	20-Jan-2015
Stakeholders and Requirements Identification (D)	B	5	22-Jan-2015	27-Jan-2015
System and Module Design (E)	B,C,D	30	28-Jan-2015	13-Mar-2015
Data Scrapping Module (F)	E	3	13-Mar-2015	16-Mar-2015
HADOOP Module (G)	F	11	16-Mar-2015	26-Mar-2015
Data Analysis Module (H)	G	8	27-Mar-2015	04-Apr-2015
User Interface Module (I)	E	8	27-Mar-2015	04-Apr-2015
Merging Modules				

(J)	E,F,G,H,I	4	05-Apr-2015	08-April-2015
Unit testing (K)	D,F,F,H,I,J	3	09-Apr-2015	11-Apr-2015
Knowledge Transfer and Walkthrough (L)	E,F,G,H,I,J	3	12-Apr-2015	14-Apr-2015
System Integration Testing (M)	K	3	15 -Apr-2015	18-Apr-2015
Performance and Compatibility Testing (N)	D,M	2	19-Apr-2015	20-Apr-2015
Regression Testing (O)	N	2	21-Apr-2015	22-Apr-2015
UA Testing (P)	O	2	23-Apr-2015	24-Apr-2015
Production Deployment (Q)	P	2	24-Apr-2015	25-Apr-2015
Total Hours		135 days*2 hrs	270 hours	

Gantt Chart



ID	Task Name	Start	Finish	Duration
1	Aim and object Identification	1/14/2015	1/15/2015	2d
2	Model Identification	1/15/2015	1/20/2015	4d
3	Chart Prepration(Gantt,TImeline)	1/21/2015	1/22/2015	2d
4	**Planning Phase Completed**	1/22/2015	1/22/2015	0d
5	Use Case Model	1/22/2015	1/27/2015	4d
6	Technology Stack	1/28/2015	1/30/2015	3d
7	SRS	1/30/2015	2/11/2015	9d
8	UML	2/12/2015	2/20/2015	7d
9	**Analysis Phase Completed**	2/20/2015	2/20/2015	0d
10	Database Design(Hadoop and MR)	2/23/2015	3/13/2015	15d
11	Interface Design	2/23/2015	3/13/2015	15d
12	**Design Phase Completed**	3/13/2015	3/13/2015	0d
13	Hadoop Setup	3/13/2015	3/26/2015	10d
14	Importing Data into HDFS	3/26/2015	3/27/2015	2d
15	Trend Analysis	3/27/2015	4/9/2015	10d
16	Interface Implementation	3/27/2015	4/9/2015	10d
17	Merging Modules	4/9/2015	4/22/2015	10d
18	**Implementation Phase Completed**	4/22/2015	4/22/2015	1d
19	Prototype Testing	4/22/2015	4/24/2015	3d
20	Product Launch	4/24/2015	4/24/2015	1d

Fig4. Gantt Chart

Pert Chart

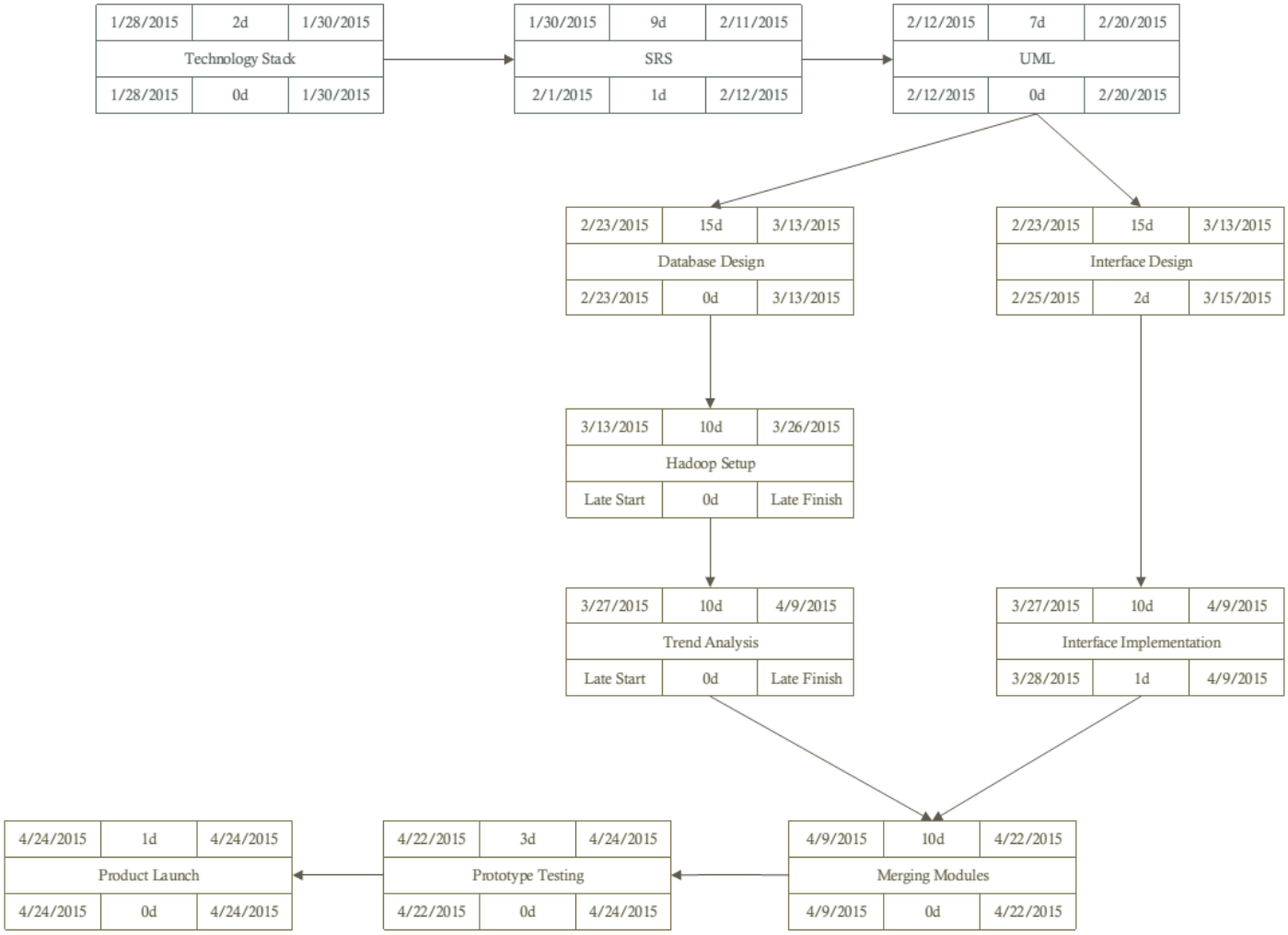


Fig5. Pert Chart