Business Enhancer

# Project Idea

Connect to any database in the world from a Java Applet and analyse the data.

# Project Concept

Fields: Ip Address or domain name

Database Server access required

Connect to the database and draw graphs defined in statistics.

Calculate the various statistical measures.

Color the graphs, select the scales of the graphs.

The HTML/php code is leveraged to enter the ipaddress/domain name, database name, and login credentials of the database, table names, selecting the X Axis and Y Axis and the various charts.

[Understand how to send messages from php to Java]

The applet only draws the charts and does the calculation of the various measures.

On the website, provide the prototype for piechart, histogram, and regression – scatter plot and one type of probability distribution.

Finally, this product must have implemented the entire Quantitative Statistics for Business Managers.

Design the workspace in the Applet.

Report Generation

# Project Planning

Study Phase – Applet Skeleton, Code for reading parameters from HTML code, database connectivity from Applet leveraging JDBC, draw various graphs leveraging JApplets.

Duration: 7 days

Design Phase – Design the HTML page, Applet.

Duration: 5 days

Implementation Phase – HTML code, Applet Code, JDBC code, JApplet Code – 7 days

Testing Phase: Test cases, testing – 5 days

Debugging Phase: Debug and fix the bugs. – 7 days

Release Phase: Upload the html and applet code online. – 1 day

Number of Resources: TBD.

# Project Task Breakdown

Draw various graphs leveraging JApplet.

Draw an arc

Draw a circle

Divide the circle into arcs based on fields.

Draw a line

Plot points

Draw a bar graph

Plot a curve

## Statistics to be calculated

Mean

Median

Mode

Standard Deviation

Variance

Co-Variance

Coefficient of Variance

Coefficient of Correlation

Coefficient of Determination

## Report Generation

Archiving generated reports to MS-Excel or MS-Word document.

# Project Tracking

|  |  |  |
| --- | --- | --- |
| Study Phase | | |
| Tasks | Status |  |
| Applet Skeleton |  |  |
| Code for reading parameters from HTML code |  |  |
| Database connectivity from Applet leveraging JDBC | Completed | Tested using MS-Access Database |
| Integration of database connectivity and graphics with JApplet |  |  |
| Draw an arc | Completed |  |
| Draw a circle | Completed |  |
| Divide the circle into arcs based on fields. | Deferred |  |
| Draw a line | Completed |  |
| Plot points | Completed |  |
| Draw graph | Completed |  |
| Draw a bar chart | Completed |  |
| Draw a side-by-side chart | Completed |  |
| Draw a frequency histogram | Completed |  |
| Draw linear regression prediction line |  |  |
| Plot a curve |  |  |
| Draw a scatter plot | Completed |  |
| Draw a distribution curve |  |  |
| Draw X-Axes, Y-Axes | Completed |  |
| Number the Axes | Completed |  |
| Label the Axes, Title of the Chart |  |  |
| Design the JPanel – One screen | Completed |  |
| Design the Tabs of the JPanel | Completed |  |
| Convert data to co-ordinates. | Completed |  |
| Design the statistics variables |  |  |
| Generate Linear Equation |  |  |
| Generate Non-Linear Equation |  |  |
| Generate Quadratic Equation |  |  |
| Generate Exponential Equation |  |  |
| Generate trendline equation |  |  |
| Create Tables of Data from data read from database | Completed |  |
| Field Listeners |  |  |
| Save Connections listener |  |  |
| Save As Excel file listener |  |  |
| Save As MS Word file listener |  |  |
| Save As PDF file listener |  |  |
| Disconnect Database listener | Completed |  |
| Display - |  |  |
| View Visual Representation listener | In Progress |  |
| View Selected Fields listener |  |  |
| View Data listener | Completed |  |
| View Analysis listener |  |  |
| View Report listener |  |  |
| Piechart - |  |  |
| Add listener | Completed |  |
| Remove listener | Not Required |  |
| Modify listener |  |  |
| Scatterplot - |  |  |
| Add listener | Completed |  |
| Remove listener | Not Required |  |
| Modify listener |  |  |
| Draw Graph - |  |  |
| Add listener | Completed |  |
| Remove listener | Not Required |  |
| Modify listener |  |  |
| Report - |  |  |
| Generate listener |  |  |
| Edit listener |  |  |
| Help - Business Enhancer Help listener | In Progress |  |
| CATEGORICAL DATA |  |  |
| Summary Tables | Completed |  |
| Descriptive Statistics from Summary Table - means, medians, percentiles, standard deviations, coefficients of variation,  sums, and counts for various combinations of grouping  (break) variables. – Create a variable summary report | In Progress |  |
| Contingency Table | Completed |  |
| Bar Chart | Completed |  |
| Pareto Chart | Completed |  |
| Side-by-side bar chart | Completed |  |
| NUMERICAL DATA |  |  |
| Ordered Array | Completed |  |
| Frequency Distribution | Completed |  |
| Relative Frequency Distribution | Completed |  |
| Percentage Distribution | Completed |  |
| Cumulative Percentage Distribution | Completed |  |
| Stem-and-leaf display | Completed |  |
| Histogram | Completed |  |
| Polygon | Completed |  |
| Cumulative Percentage Polygon | Completed |  |
| BoxPlot | Completed |  |
| Normal Probability Plot | Completed |  |
| Mean | Completed |  |
| Median | Completed |  |
| Mode | Completed |  |
| Quartiles | Completed |  |
| Geometric Mean | Completed |  |
| Range | Completed |  |
| Interquartile Range | Completed |  |
| Standard Deviation | Completed |  |
| Variance | Completed |  |
| Coefficient Of Variation | Completed |  |
| Skewness | Completed |  |
| Kurtosis | Completed |  |
| Index Numbers |  |  |
| TWO NUMERICAL VARIABLES |  |  |
| Scatter plot | Completed |  |
| Time-Series plot | Completed |  |
| CATEGORICAL AND NUMERICAL VARIABLES CONSIDERED TOGETHER |  |  |
| Multidimensional contingency tables | Completed |  |
| Pivot tables | Completed |  |
| Business Analytics |  |  |
| Control Charts | 20-09-2016 |  |

# Terminologies used for fields

Find the dependent variable (Target variable) using the values of the independent variable (Stimulus Variable).

## Regression

Simple Regression – Linear, Non-Linear

Multiple Regressions – Linear, Non-Linear

Y = b0 + B1X + Eetha

Y = b0 + B1X1 + …… + Bk1 + Eetha

Y= b0 + (1/(b1+b2X1)) + X3 ^ b3 + Eetha

Non-Linear Regression – Non-linear wrt coefficients not variables.

Ordinary Least Squares – Assumptions

Diagnostic tests

Parameter estimation

## MLR steps

Start with a hypothesis or belief

Estimate unknown model parameters(b0, b1, b2, b3, b4,…..)

Specify probability distribution of random error term – assumed to be normal distribution

Check assumptions of regression(normality, heteroscedasticity, multi-collinearity)

Evaluate the model.

Use the model for prediction and estimation

Sum of squares method or least square methods to predict the parameters.

Least Squares Method

Belief - >Y=b0+b1x1+b2x2+…………..bkxk + e

Fitted Model -> Y^ = b0^+b1^x1+b2^x2…………….bk^xk

Minimise SSE = ∑(Yi – Yi^)^2 -> lsr method

Estimation of parameters in multiple regression

SSE = ∑i=1,n e^2 = ∑i=1,n(yi – b0 - ∑j=1,n(bjxij)^2 --------------------------- 1)

The least squares estimates must satisfy,

dSSE/db0 = -2 ∑i=1,n(yi-bo^- ∑j=1,k(bjxij)) = 0 ----------> 2)

and

dSSE/dbj = -2 ∑i=1,n(yi-bo^- ∑j=1,kbj^xij)xij =0 ---------3)

k+1 such equations

R^2 ->Explains n% of variation in Y by model

P value must be <= 0.05 for variable to be statistically significant.

Statistically insignificant variables not because of multicollenearity and then drop the variable.

Interpret the coefficients.

Use standardised beta values when all variables are standardised.

Run regression, we get betas.

Standardised B = Sx/Sy \* bi

Sx, Sy -> std deviations of x and y

Bi -> raw beta

Partial correlation coefficient measures the relationship b/w two variables(x and y) when the influence of all other variables (say x2, x3, ……….) on these two variables (y and x1) are removed.

Semi-partial or part correlation coefficient measures the relationship between two variables say x1 and y, when the influence of all other variables say x2, x3, ……xk connected with these two variables y and x1 from one of the variables x1.

Partial Correlation

R12,3 =(R12-R13R23)/√(1-R13^2)(1-R23^2)

R13,2=(R13-R12R32)/√(1-R12^2)(1-R32^2)

R23,1=(R23-R21R31)/√(1-R21^2)(1-R31^2)

Correlation between y1 and x2, when the influence ofx3isremovedfrombothy1andx2.

Correlation between y1 and x3, when the influence ofx2isremoved frombothy1andx3.

Correlation between x2 and x3, when the influence ofy1isremovedfrombothx2andx3.

Null Hypothesis : b1=b2……=bk=0

Alternative Hypothesis : No all are 0

F Statistic:

F = MSR/MSE

Relationship between R^2 and F:

F = ( R2/ k) /(1-R2) /(n-k-1)

Testing for Significance of Individual Parameters

H0 : Bi = 0

HA : Bi not equal to 0

t = Bi(estimated)/Se(Bi(estimated))

T-test:

By rejecting the null hypothesis, we can claim that there is a statiscally signicnt relationship between the response variable Y and explanatory variable Xi.

Testing Model Portions - Partial F Test

---------------------------------------

Full Model

--------------

Y = b0+b1X1+b2X2+..........+bkXk + eeta

Reduced Model (r<k)

---------------------------

Y=b0+b1X1+b2X2+............+brXr+eetha

Test H0: br+1 = bk = 0

Partial F = (SSEreduced - SSEfull)/(k-r) / MSEfull

* *R-square and Adjusted R-square* are used to test the overall model fitness.
* *F test* is used to test the overall model statistical significance.
* *Partial F test* is used to test portions of the model.
* *T test*is used to test the statistical significance of individual explanatory variables.

Adjusted R2 = 1 - [(1-R2)(n-1)]/[n-k-1]

Existence of Multicollinearity –

Method I – Variance Inflation Factor.

VIF associated with introducing a new variable Xj is

given by:

VIF(Xj) = 1/(1-Rj ^ 2)

Rj ^2 is the coefficient of determination for the regression of Xj as dependent variable

The standard error of the corresponding beta is inflated by square root of VIF

* *n-1* dummy variables are created for a categorical variable with n categories.
* Two frequent ways of deriving new variables from the existing explanatory variables are *ratios and interactions*.
* *Multi-collinearity* is nothing but the high correlation between explanatory variables.
* Multi-collinearity can lead to *unstable* regression coefficients.

Variable selection

Forward selection

Backward selection

Stepwise regression

# Design of Business Enhancer

Read the data

Illustrate the data

Solve the data

Explain the data

For a scatter plot, find the trend line

## Pseudo-Algorithm for calculating trendline from data

## Calculating the Slope (m) of the Trendline

Step 1

Consider this data set of three (x,y) points: (1,3) (2, 5) (3,6.5).

Let n = the number of data points, in this case 3.

Step 2

Let a equal n times the summation of all x-values multiplied by their corresponding y-values, like so: a = 3 x {(1 x 3) +( 2 x 5) + (3 x 6.5)} = 97.5

Step 3

Let b equal the sum of all x-values times the sum of all y-values, like so: b = (1 + 2 + 3) x (3 + 5 + 6.5) = 87

Step 4

Let c equal n times the sum of all squared x-values, like so: c = 3 x (1^2 + 2^2 + 3^2) = 42

Step 5

Let d equal the squared sum of all x-values, like so: d = (1 + 2 + 3)^2 = 36

Step 6

Plug the values that you calculated for a, b, c, and d into the following equation to calculate the slope, m, of the regression line: slope = m = (a - b) / (c - d) = (97.5 - 87) / (42 - 36) = 10.5 / 6 = 1.75

## Calculating the y-intercept (b) of the Trendline

Step 1

Consider the same data set. Let e equal the sum of all y-values, like so: e = (3 + 5 + 6.5) = 14.5

Step 2

Let f equal the slope times the sum of all x-values, like so: f = 1.75 x (1 + 2 + 3) = 10.5

Step 3

Plug the values you have calculated for e and f into the following equation for the y-intercept, b, of the trendline: y-intercept = b = (e - f) / n = (14.5 - 10.5) / 3 = 1.3

Step 4

Plug your values for m and b into a linear equation to reveal the final trendline equation: Trendline equation: y = 1.75x + 1.3

## Calculating the Slope (m) of the Trendline

Double slopeOfTrendline()

{

Int n; // the number of data points, in this case 3.

// Let a equal n times the summation of all x-values multiplied by their corresponding y- values

Double a;

Double b;//Let b equal the sum of all x-values times the sum of all y-values

Double c;//Let c equal n times the sum of all squared x-values

Double d;//Let d equal the squared sum of all x-values

//Calculate slope m from a, b, c, and d

Double m;

m = (a - b) / (c - d);

return m;

}

## Calculating the y-intercept (b) of the Trendline

Double slopeOfTrendline()

{

Double n;//Number of Data points

Double e; //Sum of all y-values, like so: e = (3 + 5 + 6.5) = 14.5

Double f;//Slope times the sum of all x-values

//Calculate y-intercept of the trendline from e and f

Double b;// (e - f) / n

Return b;

}

Double[] y getTrendlineCoordinates(Double m, Double b, Double[] x)

{

//Trendline equation

Int n;

N=x.length;

For(I from 0 to (n-1))

{

Y[i] = m \* x[i] + b;

}

Return y;

}

Slope;

*α*=*n*∑(*xy*)−∑*x*∑*yn*∑*x*2−(∑*x*)2α=n∑(xy)−∑x∑yn∑x2−(∑x)2

Offset:

*β*=(∑*y*−*α*∑x)/n

Trendline formula:

*y*=*αx*+*β*

public static AffineTransform getScaleInstance(double sx,

double sy)

public AffineTransform(double m00,

double m10,

double m01,

double m11,

double m02,

double m12)

Constructs a new AffineTransform from 6 double precision values representing the 6 specifiable entries of the 3x3 transformation matrix.

Parameters:

m00 - the X coordinate scaling element of the 3x3 matrix

m10 - the Y coordinate shearing element of the 3x3 matrix

m01 - the X coordinate shearing element of the 3x3 matrix

m11 - the Y coordinate scaling element of the 3x3 matrix

m02 - the X coordinate translation element of the 3x3 matrix

m12 - the Y coordinate translation element of the 3x3 matrix

public static final int TYPE\_TRANSLATION

This flag bit indicates that the transform defined by this object performs a translation in addition to the conversions indicated by other flag bits. A translation moves the coordinates by a constant amount in x and y without changing the length or angle of vectors.

public double getScaleX()

Returns the X coordinate scaling element (m00) of the 3x3 affine transformation matrix.

Returns:

a double value that is the X coordinate of the scaling element of the affine transformation matrix.

public double getScaleY()

Returns the Y coordinate scaling element (m11) of the 3x3 affine transformation matrix.

Returns:

a double value that is the Y coordinate of the scaling element of the affine transformation matrix.

public void setToScale(double sx,

double sy)

Sets this transform to a scaling transformation. The matrix representing this transform becomes:

[ sx 0 0 ]

[ 0 sy 0 ]

[ 0 0 1 ]

Parameters:

sx - the factor by which coordinates are scaled along the X axis direction

sy - the factor by which coordinates are scaled along the Y axis direction

DisplayMetrics metrics = new DisplayMetrics();

getWindowManager().getDefaultDisplay().getMetrics(metrics);

// Convert from dots per inch to dots per centimetre.

int xdpc = metrics.xdpi / 2.54;

int ydpc = metrics.ydpi / 2.54;

1cm = how many pixels

multiple of 10

input - max

round max to next multiple of 10

and leave the last digit of max and assign 0 to all the remaining digits

Divide the resulting number by 10

----------------------------------------------

0 max\_n

number of pixelsperunitdata = x

number of pixels for max\_n = max\_n \* x

1 marking = 5 \* x

value of 1 marking = max\_n / 5

0 10 20 30..............100

for(int i = max\_n;i>0;i

1 mark = x pixels

final mark = x \* max\_n

max\_n = max number of pixels

## Design of HTML page leveraging PHP.

IP Address Type IPV4 or IPV6

Server Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_ or Server IP Address \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_(32bit or 128bit)

Database Name \_\_\_\_\_\_\_\_\_\_\_\_\_

On Selecting server and database names, the tables fields variables gets updated.

## Design of Applet

Different tabs for visual representations and analyses.

Upper portion of JApplet has a menu.

Database File Actions

Database->

Add->Dialog Box appears with IPAddress or domain name of the database server, DatabaseName, login and password, database type, Add Button

Select Field->

Table1->IntField,FloatField,TextField,BooleanField,…..

IntField->Field1, Field2,……

FloatField->Field3,Field4,…..

Table2-> IntField,FloatField,TextField,BooleanField,…..

BooleanField->Field5,Field6,…….

Fieldn -> Select Rows

File -> file operations

Fields -> Select the fields to be leveraged for analysis. The first n fields selected will be leveraged for the selected action.

File -> Save analysis to file, load data from file, Exit from Application

Load Data from File -> updates fields and data stores

Fields Menu Item -> Table -> List of Tables-> For each table item -> Field Types-> Fields for each field type item of the menu.

Actions -> PieChart, ScatterPlot, Graph, Bar Chart, Side-By-Side Chart, Histogram

//Once Action is selected, below widgets appear,

//From Table\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Add Field Type\_\_\_\_\_\_\_\_\_\_\_ Add Field\_\_\_\_\_\_\_\_\_\_\_\_\_(this gets //updated based on Tablename and fieldtype.

//The number of fields that can be added depends on the actions.

PieChart -> Draw, Display % and name in the chart, Add Statistics

ScatterPlot -> Add TrendLine, Linear Regression, Non linear Regression,Logistic Regression, Decision Tree, Neural Networks

Linear Regression – Simple, Multiple

Non Linear Regression – Simple, Multiple

Logistic Regression

Neural Network

# Login Screen Design

Store the database name , type and IP address of the database server leveraged to store the login credentials in a configuration file.

Default database name is called BusinessEnhancer\_Users.

Default database type is MySQL.

Default IP Address is 127.0.0.1

Store login credentials for accessing the application.

If IP address is not entered during login, it takes default IP address from the file.

Once connected to database servers.

Add data from selected fields

Go to Action -> select an action -> select fields -> Draw the visual representation -> Show/Hide Data

Select Data’s fields have to be stored in a List.

Show/Hide Data displays data in a table or does not display anything.

Action -> Select Piechart , etc, and proceed.

PROJECT INITIATION SHEET

What is the Project?

What is the purpose of the project?

What are the specific project objectives?

What(if any) are the major issues of the project?

What will be involved in the project?

What is the background to the project?

What are the requirements of the project?

What are the deliverables of the project?

Who are the key stakeholders?

# Predictive Analytics Process

1. Define Project:

Define the project outcomes, deliverables, scoping of the effort, business objectives, identify the data sets which are going to be used.

1. Data Collection:

Data mining for predictive analytics prepares data from multiple sources for analysis. This provides a complete view of the customer interactions.

1. Data Analysis:

Data Analysis is the process of inspecting, cleaning, transforming, and modelling data with the objective of discovering useful information, arriving at conclusions.

1. Statistics:

Statistical Analysis enables to validate the assumptions, hypotheses and test them using standard statistical models.

1. Modeling:

Predictive modelling provides the ability to automatically create accurate predictive models about future. There are also options to choose the best solution with multi model evaluation.

1. Deployment:

Predictive Model Deployment provides the option to deploy the analytical results in to the everyday decision making process to get results, reports and output by automating the decisions based on the modelling.

# Design Considerations

Create a query for each field and store it in Hash Table.

The hash table has analytic data as key and Result Set as value.

Every time Action is called this hashtable gets cleared and gets filled when Select Data button is clicked.

The Data, Visual Representation and analysis, all of them leverage this Hashtable.

MenuItem of Action

Select Data Button

Hashtable<String, Result Set> resultsets

Analysis

Visual representation

Data

MenuItem of Action does the following,

Clears the resultsets hashtable

Removes all components in Data Panel, Visual Representation Panel, Analysis Panel and Report Panel.

Displays the Configuration Panel.

The Data, Visual Representation and Analysis, all these take data from hashtable resultsets.

The Hashtable resultsets keys are formed by concatenating the analyticdataname, fieldType and fieldname with “;” as delimiter.

How to use resultsets?

Extract tokens analyticdata, fieldtype and fieldname using the key.

The header name = analyticdata

//if conditions based on fieldtype

Use the fieldname as it is to get the data.

Using the key get the resultset.

Get the data and store in ArrayList.

How do you define the scope of the project?

What planning needs to be done?

What resources do I need?

When does it need to be done by?

And, how much will it cost?

Create the WBS.

Identify the project deliverables.

WBS is revisited after identifying the project risks.

SUBTASK

SUBTASK

SUBTASK

PROJECT

SUBTASK

SUBTASK

SUBTASK

TASK 3

TASK 2

TASK 1

5 key elements of Project Schedule

1. Defining the activities
2. Sequencing these activities
3. Estimating the activity resources required
4. Estimating the activity duration to complete
5. Developing the schedule

Scheduling Table

|  |  |  |
| --- | --- | --- |
| Step1 : What do I need to do?(define the activities)  Step2 : Sequencing | Step3 : What resources do I need?(Estimating activity resources) | Step4 : How much time do I need?(Estimating activity duration) |
|  |  |  |
|  |  |  |
|  |  |  |

The initial draft of the schedule table and refine multiple times

Gnatt Chart – Time Schedule Chart

Scope – requirements, activities, scope, WBS

Action plan – what, who, when, where, how

Sequence

Dependencies for every activity

Schedule is represents by Gantt Chart

Timeframe

People, material, equipment

Budget

Milestone Log

Quality Management Processes

# Risk Assessment Template

|  |  |  |
| --- | --- | --- |
| Risk Identified | Likelihood of risk occurring | Risk Response |
|  |  |  |

# Risk Management Plan

# PESTLE Risk Framework

#### Higher risk project

If your project reflects a higher risk profile, it is often good idea to revisit your project idea/concept and explore alternative options by applying some critical and creative thinking techniques. In such a situation, take time to rethink your approach and then re-do the PESTLE profile. Before proceeding with your project, it is advisable that you take sufficient time to engage with your team and broader stakeholders to identify appropriate risk responses to reduce the likelihood and impact of these risks on your project during execution.

Am I delivering to the planned specifications and requirements?

Am I able to secure supply of the necessary resources?

Will I be able to manage and control the project scope?

How will I be able to control quality?

# Team Charter

Goals & Objectives What we do?

Roles & Responsibilities Who does what?

Processes & Procedures How we execute?

Relationships How we work together?

Leadership How our leader leads the team?

External Environment How we impact, or are effected by others?

# Project Kickoff

All stakeholders have shared vision

Culture – use tools

# Effective Communication involves,

Delivering information

Giving clear instructions

Listening to responses

Providing constructive feedback

Successful management of conflict

# Project Closure

Obtaining project acceptance

Finalising and terminating resources

De-Establishing the team

Capturing remaining costs and closing project related accounts

Compiling a final project report

What is the current schedule and cost status for your project?

How much will it cost to complete the project?

When will the project be completed?

If there are schedule slippages and/or cost overruns, what will be the forecasts at completion?

Are there potential problems that need to be addressed?

activity duration times

resource usage

actual costs

contracted services

documentation

warranties

unresolved issue

What is the project?

How will you go about planning for project execution?

How will you engage with your team and other stakeholders throughout the process?

What strategies will you put into place to monitor and review the whole project management life cycle?

# Organization Structure

Based on Functions –

Proprietor

Software Development Team

Developers, Testers,

Design Team

Marketing Team

Roles And Responsibilities Of Proprietor

* HR functions – Recruitment, Interviewing,
* Accounting – Subscribers’ fees, Salary decisions and disbursement

# Organisation Culture

**"**Service, quality and highly motivated workforce depend on the culture of the organization, and when one of the business units fail to get the right balance—it might be profit ahead of service, or it might be productivity ahead of human resources—the image of the organization will deteriorate."

Creation of Culture through the way we :

* Select People
* Induct People
* Reward Behaviour
* Emulate Role Models
* Identify what is important

Organisation Structure and Organisation Culture impact Employee Behaviour.

FIRST TIME MANAGER

* TRUST THE TEAM MEMBERS
* DIFFERENTIATE TREE FROM THE FOREST
* STAKEHOLDERS
* PERFORMANCE REVIEWS
* PERFORMANCE MANAGEMENT
* CONFLICT MANAGEMENT
* UNDERSTAND INFORMAL NETWORKS IN THE ORGANISATION
* INTEGRITY
* GOOD LISTENER
* GOOD OBSERVER
* ABILITY TO MANAGE PEOPLE
* INCLUSIVE RESULTS

WHAT DOES BUSINESS ENHANCER DO?

Business Enhancer is a Predictive Analytics tool based on the DCOVA framework.

# Business Enhancer:

Defines the data leveraged to solve a problem or meet a business objective.

Collects the data from appropriate sources.

Organises the data collected by representing them as tables.

Visualizes the data collected by developing charts.

Analyzes the data collected.

Creates reports from the analysis.

It also publishes to the dashboard webpage.

Business Enhancer can be leveraged to,

uncover unforeseen relationships between the various variables.

develop optimization models that impact an organisation's planning, strategy and operations.

collects and processes data sets of all sizes including Big Data.

# What is Big Data?

Data that are being collected in large volumes, at very fast rates and are of various forms and can be organised or unorganised.

# What are you waiting for?

If you have defined your business objective, then leveraging Business Enhancer must be your strategic goal for making business decisions.

## Definitions

Discrete variables

Continuous variables

## Measurement scales for variables

Nominal, Ordinal scales -> Categorical variables

Interval, ratio scales -> Numerical variable

Nominal ->yes/no, male/female, name

Ordinal -> fair, good, excellent, A/AA/AAA/AAA+

# Data Sources

Any RDBMS,

Excel Sheet,

Text files

## Data Cleaning

Presently, there is no data cleaning done by the tool. Data Cleaning must be done in your database or excel sheet prior to being collected by the tool. Require to get inputs from customer.

## Recoded Variables

This is implemented as user-defined analytics data.

The entire data set that is collected is leveraged for further processing.

# Contingency Table

Create an instance of class ContingencyTable with n number of categorical variables

Find how many combinations of 2 variables you can create, given the number of variables

Let number of categorical variables = n

Int numberofpartialtables = combination (int n)

Int combination (n)

{

Return total number of combinations of 2 variables taken at a time.

}

Create numberofpartialtables number of NXM contingency tables – partial tables

For each combination of 2 variables

{

Create Vector of Vector of uniquevalues to store data of 2 variables.

Each Vector of Vector represents a contingency table.

Create a Vector of Vector of Integer containing the counts for each combination of xij.

}

Store the vectors of each contingency table in a Vector.

For every two variables

One frequency contingency table is created

Three percentage contingency table is created

One percentage of total

One percentage of row total

One percentage of column total

In the view of contingency table, leverage this Vector<Vector<Vector>>>. It may be a Vector<Vector<HashMap<T, Integer>>>

# Multidimensional Contingency Table

Same as Contingency Table, except that you mention the number of variables to be taken at a time.

# Pivot Table

Same as Contingency Table except that we mention what goes in as rows and what variables go in as columns.

|  |  |  |  |
| --- | --- | --- | --- |
| **Organise/Visualize** | **Tests Conducted** | **Conclusions** |  |
| Summary Table |  |  |  |
|  |  |  |  |
| Contingency Table | Pearson’s chi-squared test |  |  |
| Contingency Table | G-test |  |  |
| Contingency Table | Fisher’s exact test |  |  |
| Contingency Table | Barnard’s test |  |  |
| Contingency Table |  | Check if data is dependent |  |
| Contingency Table |  | Check if data is independent |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Combination Formula**

A [formula](http://www.mathwords.com/f/formula.htm) for the number of possible [combinations](http://www.mathwords.com/c/combination.htm) of *r* objects from a [set](http://www.mathwords.com/s/set.htm) of *n* objects. This is written in any of the ways shown below.

http://www.mathwords.com/b/b_assets/binomial%20coefficient%20notation.gif

All forms are read aloud "*n* choose *r*."

|  |  |
| --- | --- |
| Formula: | http://www.mathwords.com/b/b_assets/binomial%20coefficient%20formula.gif |

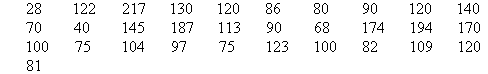
**Permutation Formula**

A [formula](http://www.mathwords.com/f/formula.htm) for the number of possible [permutations](http://www.mathwords.com/p/permutation.htm) of *k* objects from a [set](http://www.mathwords.com/s/set.htm) of *n*. This is usually written *n*P*k* .

|  |  |
| --- | --- |
| Formula: | http://www.mathwords.com/p/p_assets/permutation%20formula.gif |
| Example: | How many ways can 4 students from a group of 15 be lined up for a photograph? |
| Answer: | There are 15P4 possible permutations of 4 students from a group of 15. |

#### Example 6

The number of calls from motorists per day for roadside service was recorded for the month of December 2003.  The results were as follows:

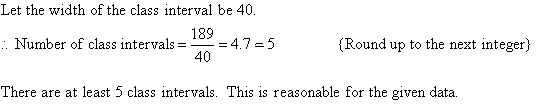


Set up a frequency table for this set of data values.

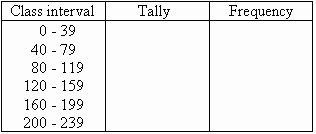
##### Solution:

To construct a frequency table, we proceed as follows:

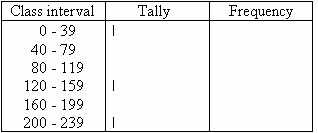




Step 1:  Construct a table with three columns, and then write the data groups or class intervals in the first column.  The size of each group is 40.  So, the groups will start at 0, 40, 80, 120, 160 and 200 to include all of the data.  Note that in fact we need 6 groups (1 more than we first thought).

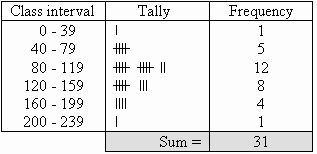
****

Step 2:  Go through the list of data values.  For the first data value in the list, 28, place a tally mark against the group 0-39 in the second column.  For the second data value in the list, 122, place a tally mark against the group 120-159 in the second column.  For the third data value in the list, 217, place a tally mark against the group 200-239 in the second column.

****

We continue this process until all of the data values in the set are tallied.

Step 3:  Count the number of tally marks for each group and write it in the third column.  The finished frequency table is as follows:



# Milestones Achieved

November 2015 – January 15th 2016 – R&D on AWT and developed a toolkit.

January 16th 2016 - Started coding of BusinessEnhancer application

April 16th 2016 – Completed Summary Table creation – It was a day of fulfilment and achievements.

May 5th 2016 – Completed Contingency Table creation

July 29th 2016 – Completed Pivot Table creation – A terrific day – The best program I ever wrote in my career.

August 10th 2016 – Completed Frequency Distribution – 10 days.

January 3rd 2018 – Completed JAXB coding and reading from writing to xml files. All templates in an xml file.

|  |  |  |  |
| --- | --- | --- | --- |
| PRODUCT | BRAND | CITY | RETAILER |
| Soap | Cinthol | Bengaluru | Food World |
| Toothpaste | Closeup | Calcutta | More |
| Shampoo | Sunsilk | Srinagar | Ramesh Stores |
| Talcolm Powder | Ponds | Bengaluru | Food Hall |
| Daycream | Ponds | Bengaluru | Food Hall |
| Soap | Liril | Bengaluru | Food World |
| Toothpaste | Closeup | Calcutta | More |
| Shampoo | Sunsilk | Calcutta | Ramesh Stores |
| Talcolm Powder | Cinthol | Calcutta | Food World |
| Daycream | VLCC | Srinagar | More |
| Soap | Lux | Srinagar | Ramesh Stores |
| Soap | Lux | Srinagar | Ramesh Stores |
| Toothpaste | Colgate | Bengaluru | Food World |
| Shampoo | Sunsilk | Chikmagalur | Food World |
| Talcolm Powder | Liril | Chikmagalur | Food World |
| Daycream | VLCC | Chikmagalur | Food World |
| Soap | Cinthol | Mandya | Ramesh Stores |
| Toothpaste | Colgate | Mandya | Ramesh Stores |
| Shampoo | Head And Shoulders | Calcutta | Ramesh Stores |
| Talcolm Powder | Ponds | Mandya | Ramesh Stores |
| Daycream | Olay | Chikmagalur | Food Hall |
| Soap | Cinthol | Chikmagalur | Food Hall |
| Toothpaste | Pepsodent | Chikmagalur | More |
| Shampoo | Clinic | Bengaluru | More |
| Talcolm Powder | Yarley | Bengaluru | More |
| Daycream | VLCC | Srinagar | More |
| Daycream | VLCC | Srinagar | Food World |
| Daycream | VLCC | Mandya | Food World |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Name | Total Marks | Result | Grade | Social Studies | Maths | Science |
| Suma | 281 | Pass | A | 90 | 95 | 96 |
| Gopalakrishna | 260 | Pass | A | 100 | 100 | 60 |
| Gopala | 291 | Pass | A | 100 | 95 | 96 |
| Krishna | 115 | Pass | C | 40 | 35 | 40 |
| Suseela | 300 | Pass | A | 100 | 100 | 100 |
| Suma Gopalakrishna | 276 | Pass | A | 91 | 92 | 93 |
| Suseela Gopalakrishna | 239 | Pass | B | 100 | 74 | 65 |
| Vicky | 56 | Fail | D | 20 | 10 | 26 |
| Gopala | 288 | Pass | A | 94 | 96 | 98 |
| Suseela Gopalakrishna | 232 | Pass | B | 90 | 72 | 70 |
| Suma G | 279 | Pass | A | 91 | 93 | 95 |
| Rani | 55 | Fail | D | 17 | 18 | 20 |

Design for Pivot Table

Decide the Row Labels

Decide the Column Labels

# Pivot table

From Wikipedia, the free encyclopedia

In [data processing](https://en.wikipedia.org/wiki/Data_processing), a **pivot table** is a data summarization tool found in data visualization programs such as [spreadsheets](https://en.wikipedia.org/wiki/Spreadsheet) or [business intelligence](https://en.wikipedia.org/wiki/Business_intelligence) software. Among other functions, a pivot table can automatically sort, count, total or give the average of the data stored in one table or spreadsheet, displaying the results in a second table showing the summarized data. Pivot tables are also useful for quickly creating unweighted [cross tabulations](https://en.wikipedia.org/wiki/Cross_tabulation). The user sets up and changes the summary's structure by [dragging and dropping](https://en.wikipedia.org/wiki/Drag_and_drop) fields graphically. This "rotation" or pivoting of the summary table gives the concept its name.

The term *pivot table* is a generic phrase used by multiple vendors. In the United States, [Microsoft Corporation](https://en.wikipedia.org/wiki/Microsoft) has [trademarked](https://en.wikipedia.org/wiki/Trademark) the specific [compound word](https://en.wikipedia.org/wiki/CamelCase) form, *PivotTable*.[[1]](https://en.wikipedia.org/wiki/Pivot_table#cite_note-1)

Pivot tables can be seen as a simplification of the more complete and complex [online analytical processing concepts](https://en.wikipedia.org/wiki/Online_analytical_processing) (OLAP).

## Contents

* [1 History](https://en.wikipedia.org/wiki/Pivot_table#History)
* [2 Mechanics](https://en.wikipedia.org/wiki/Pivot_table#Mechanics)
* [3 Implementation](https://en.wikipedia.org/wiki/Pivot_table#Implementation)
  + [3.1 Report filter](https://en.wikipedia.org/wiki/Pivot_table#Report_filter)
  + [3.2 Column labels](https://en.wikipedia.org/wiki/Pivot_table#Column_labels)
  + [3.3 Row labels](https://en.wikipedia.org/wiki/Pivot_table#Row_labels)
  + [3.4 Summation values](https://en.wikipedia.org/wiki/Pivot_table#Summation_values)
* [4 Application support](https://en.wikipedia.org/wiki/Pivot_table#Application_support)
* [5 Online analytical processing](https://en.wikipedia.org/wiki/Pivot_table#Online_analytical_processing)
* [6 See also](https://en.wikipedia.org/wiki/Pivot_table#See_also)
* [7 References](https://en.wikipedia.org/wiki/Pivot_table#References)
* [8 Further reading](https://en.wikipedia.org/wiki/Pivot_table#Further_reading)

## History

In their book *Pivot Table Data Crunching*,[[2]](https://en.wikipedia.org/wiki/Pivot_table#cite_note-2) Bill Jelen and Mike Alexander refer to [Pito Salas](https://en.wikipedia.org/wiki/Pito_Salas) as the "father of pivot tables". While working on a concept for a new program that would eventually become [Lotus Improv](https://en.wikipedia.org/wiki/Lotus_Improv), Salas noted that spreadsheets have patterns of data. A tool that could help the user recognize these patterns would help to build advanced data models quickly. With Improv, users could define and store sets of categories, then change views by dragging category names with the mouse. This core functionality would provide the model for pivot tables.

[Lotus Development](https://en.wikipedia.org/wiki/Lotus_Software) released Improv in 1991 on the [NeXT](https://en.wikipedia.org/wiki/NeXT) platform. A few months after the release of Improv, [Brio Technology](https://en.wikipedia.org/wiki/Brio_Technology) published a standalone [Macintosh](https://en.wikipedia.org/wiki/Macintosh) implementation, called DataPivot (with technology eventually patented in 1999[[3]](https://en.wikipedia.org/wiki/Pivot_table#cite_note-3)). [Borland](https://en.wikipedia.org/wiki/Borland) purchased the DataPivot technology in 1992 and implemented it in their own spreadsheet application, [Quattro Pro](https://en.wikipedia.org/wiki/Quattro_Pro).

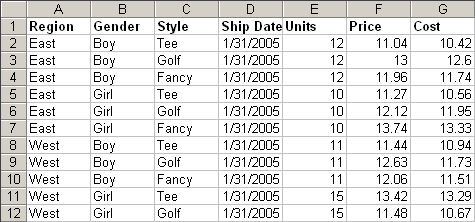
In 1993 the Microsoft Windows version of Improv appeared. Early in 1994 [Microsoft Excel](https://en.wikipedia.org/wiki/Microsoft_Excel) 5 [[4]](https://en.wikipedia.org/wiki/Pivot_table#cite_note-4) brought a new functionality called a "PivotTable" to market. Microsoft further improved this feature in later versions of Excel:

* Excel 97 included a new and improved PivotTable Wizard, the ability to create calculated fields, and new pivot cache objects that allow developers to write [Visual Basic for Applications](https://en.wikipedia.org/wiki/Visual_Basic_for_Applications) macros to create and modify pivot tables
* Excel 2000 introduced "Pivot Charts" to represent pivot-table data graphically

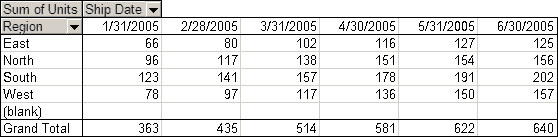
In 2007 Oracle Corporation made PIVOT and UNPIVOT operators available in [Oracle Database](https://en.wikipedia.org/wiki/Oracle_Database) 11g.[[5]](https://en.wikipedia.org/wiki/Pivot_table#cite_note-5)

## Mechanics

For typical data entry and storage, data usually appear in *flat* tables, meaning that they consist of only columns and rows, as in the following example showing data on shirt types:

[](https://en.wikipedia.org/wiki/File:Pivottable-Flatdata.png)

While tables such as these can contain many data items, it can be difficult to get summarized information from them. A pivot table can help quickly summarize the data and highlight the desired information. The usage of a pivot table is extremely broad and depends on the situation. The first question to ask is, "What am I seeking?" In the example here, let us ask, "How many *Units* did we sell in each *Region* for every *Ship Date?*":

[](https://en.wikipedia.org/wiki/File:Pivottable-Pivoted.PNG)

A pivot table usually consists of *row*, *column* and *data* (or *fact*) fields. In this case, the column is *Ship Date*, the row is *Region* and the datum we would like to see is (sum of) *Units*. These fields allow several kinds of [aggregations](https://en.wikipedia.org/wiki/Aggregate_function), including: sum, average, [standard deviation](https://en.wikipedia.org/wiki/Standard_deviation), count, etc. In this case, the total number of units shipped is displayed here using a *sum* aggregation.

## Implementation

Using the example above, software will find all distinct values for *Region*. In this case, they are: *North*, *South*, *East*, *West*. Furthermore, it will find all distinct values for *Ship Date*. Based on the aggregation type, *sum*, it will summarize the fact and display them in a multidimensional chart. In the example above, the first datum is 66. This number was obtained by finding all records where both *Region* was *East* and *Ship Date* was *1/31/2005*, and adding the *Units* of that collection of records (*i.e.*, cells E2 to E7) together to get a final result.

Pivot tables are not created automatically. For example, in Microsoft Excel one must first select the entire data in the original table and then go to the Insert tab and select "Pivot Table" (or "Pivot Chart"). The user then has the option of either inserting the pivot table into an existing sheet or creating a new sheet to house the pivot table.[[6]](https://en.wikipedia.org/wiki/Pivot_table#cite_note-6) A pivot table field list is provided to the user which lists all the column headers present in the data. For instance, if a table represents sales data of a company, it might include Date of sale, Sales person, Item sold, Color of item, Units sold, Per unit price, and Total price. This makes the data more readily accessible.

| **Date of sale** | **Sales person** | **Item sold** | **Color of item** | **Units sold** | **Per unit price** | **Total price** |
| --- | --- | --- | --- | --- | --- | --- |
| 10/01/13 | Jones | Notebook | Black | 8 | 25000 | 200000 |
| 10/02/13 | Prince | Laptop | Red | 4 | 35000 | 140000 |
| 10/03/13 | George | Mouse | Red | 6 | 850 | 5100 |
| 10/04/13 | Larry | Notebook | White | 10 | 27000 | 270000 |
| 10/05/13 | Jones | Mouse | Black | 4 | 700 | 3200 |

The fields that would be created will be visible on the right hand side of the worksheet. By default, the pivot table layout design will appear below this list.

Each of the fields from the list can be dragged on to this layout, which has four options :

1. Report filter
2. Column labels
3. Row labels
4. Summation values

### Report filter

Report filter is used to apply a filter to an entire table. For example, if the "Color of Item" field is dragged to this area, then the table constructed will have a report filter inserted above the table. This report filter will have drop-down options (Black, Red, and White in the example above). When an option is chosen from this [drop-down list](https://en.wikipedia.org/wiki/Drop-down_list) ("Black" in this example), then the table that would be visible will contain only the data from those rows that have the "Color of Item = Black".

### Column labels

Column labels are used to apply a filter to one or more columns that have to be shown in the pivot table. For instance if the "Sales person" field is dragged to this area, then the table constructed will have a values from the column "Sales Person", *i.e.*, one will have number of columns equal to the number of "Sales person". There will also be one added column of Total. In the example above, this instruction will create three columns in the table — Imran, Larry, and Grand Total. There will be a filter above the data — column labels — from which one can select or deselect a particular sales person for the pivot table.

This table will not have any numerical values as no numerical field is selected but when it is selected, the values will automatically get updated in the column of "Grand total".

### Row labels

Row labels are used to apply a filter to one or more rows that have to be shown in the pivot table. For instance if the "Sales person" field is dragged on this area then the other output table constructed will have a values from the column "Sales person", *i.e.*, one will have number of rows equal to the number of "Sales Person". There will also be one added row of "Grand Total". In the example above, this instruction will create three rows in the table — Imran, Larry, and Grand Total. There will be a filter above the data — row labels — from which one can select or deselect a particular sales person for the Pivot table.

This table will not have any numerical values, as no numerical field is selected, but when it is selected, the values will automatically get updated in the Row of "Grand Total".

### Summation values

This usually takes a field that has numerical values that can be used for different types of calculations. However, using text values would also not be wrong; instead of Sum it will give a count. So, in the example above, if the "Units sold" field is dragged to this area along with row label of "Sales person", then the instruction will add a new column, "Sum of units sold", which will have values against each sales person.

|  |  |
| --- | --- |
| **Row labels** | **Sum of units sold** |
| Imran | 14 |
| Larry | 18 |
| Grand total | 32 |

## Application support

Pivot tables are an integral part of a [spreadsheet application](https://en.wikipedia.org/wiki/List_of_spreadsheet_software). In addition to Microsoft Excel, competing software programs such as [Apache OpenOffice Calc](https://en.wikipedia.org/wiki/OpenOffice.org#Components) provide similar functionality; the implementation in Apache OpenOffice and [LibreOffice](https://en.wikipedia.org/wiki/LibreOffice) up to release 3.3 is called [DataPilot](https://en.wikipedia.org/wiki/OpenOffice.org). In version 3.4 of both LibreOffice and OpenOffice, DataPilot is renamed "Pivot Table". Other companies, such as [Quantrix](https://en.wikipedia.org/wiki/Quantrix) and [numberGo](https://en.wikipedia.org/w/index.php?title=NumberGo&action=edit&redlink=1), provide similar implementations.

Pivot functionality can also be found in other data visualization tools, including [business intelligence](https://en.wikipedia.org/wiki/Business_intelligence) packages.

[Google Docs](https://en.wikipedia.org/wiki/Google_Docs) initially allowed the creation of basic pivot tables via the pivot table [gadget](https://en.wikipedia.org/wiki/Google_Gadgets) from [Panorama](https://en.wikipedia.org/wiki/Panorama_Software) called Panorama Analytics, but as of 2011 this gadget provided limited functionality and was extremely slow with large data sets. In May 2011, Google announced the roll-out of a natively hosted pivot table feature in the Google spreadsheets editor.[[7]](https://en.wikipedia.org/wiki/Pivot_table#cite_note-7)

[ZK](https://en.wikipedia.org/wiki/ZK_%28framework%29), an Ajax framework, also allows the embedding of pivot tables in Web applications.

[PostgreSQL](https://en.wikipedia.org/wiki/PostgreSQL), an [object-relational database management system](https://en.wikipedia.org/wiki/Object-relational_database), allows the creation of pivot tables using the *tablefunc* module.[[8]](https://en.wikipedia.org/wiki/Pivot_table#cite_note-8)

[MariaDB](https://en.wikipedia.org/wiki/MariaDB), a MySQL fork, allows pivot tables using the CONNECT storage engine.[[9]](https://en.wikipedia.org/wiki/Pivot_table#cite_note-9)

## Online analytical processing

Excel pivot tables include the feature to directly query an OLAP server for retrieving data instead of getting the data from an Excel spreadsheet. On this configuration a pivot table is a simple client of an OLAP server. Excel's PivotTable not only allows for connecting to Microsoft's Analysis Service, but to any [XML for Analysis](https://en.wikipedia.org/wiki/XML_for_Analysis) (XMLA) OLAP standard-compliant server.

Other OLAP clients are JPivot, Dundas, [IcCube](https://en.wikipedia.org/wiki/IcCube) (Client Library).

Create Visual Control Systems

Improvements in Productivity are data driven, employee centric and continuous and Visual Control Aids are a good mechanism to achieve all these three. Visual Control System is an operational measurement system and is maintained by the operating personnel. Chosen measures can be monitored continuously on appropriate time basis such as daily/shift production. You can create Visual Control Boards in work areas. Your employees collect various types of data and input it into BusinessEnhancer tool that generates visual dashboards that can be displayed in work area and leveraged by team of people who investigate improvement opportunities on a weekly, monthly or quarterly basis and these visuals have a certain impact. Being visual captures one’s mind space and makes stakeholders such as employees think and results in improvement in productivity and waste elimination. It makes employees more involved and encourages them to take ownership of issues they are facing. Then what are you all waiting for, come and become a proud owner of the BusinessEnhancer Tool.

**Visual Control Aids** help us to achieve all three improvements effectively in a single attempt.

BusinessEnhancer tool can be leveraged to build your Quality Assurance System.

Quality is managed statistically leveraging metrics such as Acceptable Quality Level(AQL) and Average Outgoing Quality Level(AOQL).

Aim for Zero Defect

Manufacturing

Defect Rate = Parts Per Million(PPM)

Service

Defect Rate = Defects Per Million Opportunities(DPMO)

These two metrics are used in Six Sigma Quality Process.

The data collected from a project is used to improve the process and achieve zero defect state and for that the measured data is analysed to find the variations and the root cause of a problem and BusinessEnhancer is a tool developed to eliminate these.

This means BusinessEnhancer tool can be used by Six Sigma Coaches. It can be used as tool for Total Quality Management. It can be used to generate Control Charts for highlighting problems and Histograms, Pareto Diagram, Scatter Diagram and graphs for identifying improvement opportunities. Therefore it can be used as a Quality Management Tool.

SPC is a collective set of tools & techniques used to develop a quality assurance system that enables one to make meaningful sense of these variations.

BusinessEnhancer is used as a tool in Statistical Process Control in order to develop a quality assurance system that enables one to give useful meanings to variations in the business processes.

Two types of variations occur in business processes. They are,

* Common causes or random causes
* Non-random causes

Statistics is at the core of modern quality management –Helps operationalize some decisions and keep performance and outcome within limits –Provides basic framework to systematically analyze the quality problem in various business processes –A good mechanism to highlight either an existing quality problem or an impending problem.

SPC is a collective set of tools & techniques used to develop a quality assurance system that enables one to make meaningful sense of these variations.

Variable based method of measuring process

BusinessEnhancer can be leveraged in a Six Sigma Organisation.

|  |  |  |
| --- | --- | --- |
| Organise/Visualise | Planned Date Of Completion | Status |
| Frequency Distribution, Cumulative and Relative Distribution | 03-08-2016 | Completed |
| Bar Chart | 04-08-2016 – 05-08-2016 | Completed |
| Histogram | 06-08-2016 – 07-08-2016 | Completed |
| "Pie Chart” | 08-08-2016 | Completed |
| "Pareto Chart” | 08-08-2016-09-08-2016 | Completed |
| "Side-By-Side Bar Chart” | 09-08-2016 – 10-08-2016 | Completed |
| "Stem-And-Leaf Display” | 11-08-2016-13-08-2016 | Completed |
| "Polygon" | 13-08-2016 | Completed |
| "Cumulative Percentage Polygon” | 14-08-2016 | Completed |
| "Boxplot” | 15-08-2016 | Completed |
| "Normal Probability Plot” | 16-08-2016-18-08-2016 | Completed |
| "Scatterplot" | 19-08-2016 | Completed |
| "Time-Series Plot” | 19-08-2016 | Completed |
| "Graph" | 19-08-2016 – 20-08-2016 | Completed |

Quadrants

qiv qiii qii qi

0 0 0 1 1

0 0 1 0 2

0 0 1 1 3

0 1 0 0 4

0 1 0 1 5

0 1 1 0 6

0 1 1 1 7

1 0 0 0 8

1 0 0 1 9

1 0 1 0 10

1 0 1 1 11

1 1 0 0 12

1 1 0 1 13

1 1 1 0 14

1 1 1 1 15

KNOWN BUGS

Frequency Distribution, Histogram – All tabs are displaying same data.

Can create only one file of each type.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Test Case Number | Test Case | Data Type | Expected Output | Actual Output | Result | Bugs Uncovered |
| BE-T1 | Define Usecases | NA | None of the usecases must be empty in drop down box and selected usecase fields must be displayed in Analytics table |  |  |  |
| BE-T2 | Create new usecase | NA | Add a new usecase name and add fields and save in a file in name value format |  |  |  |
| BE-T3 | Connect to oracle db | ALL | All fields in database must appear in tree format |  |  |  |
| BE-T4 | Connect to excel sheets | ALL | All fields in database must appear in tree format |  |  |  |
| BE-T5 | Connect to access db | ALL | All fields in database must appear in tree format |  |  |  |
| BE-T6 | Connect to mysql db | ALL | All fields in database must appear in tree format |  |  |  |
| BE-T7 | Connect to derby client | ALL | All fields in database must appear in tree format |  |  |  |
| BE-T8 | Connect to derby remote | ALL | All fields in database must appear in tree format |  |  |  |
| BE-T9 | Connect to CSV file | ALL | All fields in database must appear in tree format |  |  |  |
| BE-T10 | Connect to name-value pair text file | ALL | All fields in database must appear in tree format |  |  |  |
| BE-T11 | Check for all data types | ALL | All fields in database must appear in tree format |  |  |  |
| BE-T12 | Link all types of data to usecase fields | ALL | The linked usecase field name must be in green in Analytics table |  |  |  |
| BE-T13 | Unlink linked data field and usecase field | ALL | The unlinked usecase field must turn blue in Analytics table |  |  |  |
| BE-T14 | Connect to multiple databases | ALL |  |  |  |  |
| BE-T15 | Link usecases to fields in different databases | ALL |  |  |  |  |
| BE-T16 | Change the colour of linked usecase field to GREEN color in Analytics Table | ALL |  |  |  |  |
| BE-T17 | Login | NA |  |  |  |  |
| BE-T18 | Logout | NA |  |  |  |  |
| BE-T19 | All panels must be empty if no links are present | ALL |  |  |  |  |
| BE-T20 | Disconnect one database | ALL |  |  |  |  |
| BE-T21 | Disconnect random number of Databases | ALL |  |  |  |  |
| BE-T22 | Disconnect all the databases | ALL |  |  |  |  |
| BE-T23 | Connect new databases after disconnecting all databases | ALL |  |  |  |  |
| BE-T24 | Disconnect a database, connect a new database and reconnect a disconnected database | ALL |  |  |  |  |
| BE-T25 | Reconnect one disconnected database | ALL |  |  |  |  |
| BE-T26 | Reconnect all disconnected databases | ALL |  |  |  |  |
| BE-T27 | Reconnect some disconnected databases | ALL |  |  |  |  |
| BE-T28 | Reconnect some disconnected databases and create new databases | ALL |  |  |  |  |
| BE-T29 |  |  |  |  |  |  |
| BE-T30 |  |  |  |  |  |  |
| BE-T31 |  |  |  |  |  |  |
| BE-T32 |  |  |  |  |  |  |
| BE-T33 |  |  |  |  |  |  |
| BE-T34 |  |  |  |  |  |  |
| BE-T35 |  |  |  |  |  |  |
| BE-T36 |  |  |  |  |  |  |
| BE-T37 |  |  |  |  |  |  |
| BE-T38 |  |  |  |  |  |  |
| BE-T39 |  |  |  |  |  |  |
| BE-T40 |  |  |  |  |  |  |
| BE-T41 |  |  |  |  |  |  |
| BE-T42 |  |  |  |  |  |  |
| BE-T43 |  |  |  |  |  |  |
| BE-T44 |  |  |  |  |  |  |
| BE-T45 |  |  |  |  |  |  |
| BE-T46 |  |  |  |  |  |  |
| BE-T47 |  |  |  |  |  |  |
| BE-T48 |  |  |  |  |  |  |
| BE-T49 |  |  |  |  |  |  |
| BE-T50 |  |  |  |  |  |  |
| BE-T51 |  |  |  |  |  |  |
| BE-T52 |  |  |  |  |  |  |
| BE-T53 |  |  |  |  |  |  |
| BE-T54 |  |  |  |  |  |  |
| BE-T55 |  |  |  |  |  |  |
| BE-T56 |  |  |  |  |  |  |
| BE-T57 |  |  |  |  |  |  |
| BE-T58 |  |  |  |  |  |  |
| BE-T59 |  |  |  |  |  |  |
| BE-T60 |  |  |  |  |  |  |
| BE-T61 |  |  |  |  |  |  |
| BE-T62 |  |  |  |  |  |  |
| BE-T63 |  |  |  |  |  |  |
| BE-T64 |  |  |  |  |  |  |
| BE-T65 |  |  |  |  |  |  |
| BE-T66 |  |  |  |  |  |  |
| BE-T67 |  |  |  |  |  |  |
| BE-T68 |  |  |  |  |  |  |
| BE-T69 |  |  |  |  |  |  |
| BE-T70 |  |  |  |  |  |  |
| BE-T71 |  |  |  |  |  |  |
| BE-T72 |  |  |  |  |  |  |
| BE-T73 |  |  |  |  |  |  |
| BE-T74 |  |  |  |  |  |  |
| BE-T75 |  |  |  |  |  |  |
| BE-T76 |  |  |  |  |  |  |
| BE-T77 |  |  |  |  |  |  |
| BE-T78 |  |  |  |  |  |  |
| BE-T79 |  |  |  |  |  |  |
| BE-T80 |  |  |  |  |  |  |
| BE-T81 |  |  |  |  |  |  |
| BE-T82 |  |  |  |  |  |  |
| BE-T83 |  |  |  |  |  |  |
| BE-T84 |  |  |  |  |  |  |
| BE-T85 |  |  |  |  |  |  |
| BE-T86 |  |  |  |  |  |  |
| BE-T87 |  |  |  |  |  |  |
| BE-T88 |  |  |  |  |  |  |
| BE-T89 |  |  |  |  |  |  |
| BE-T90 |  |  |  |  |  |  |
| BE-T91 |  |  |  |  |  |  |
| BE-T92 |  |  |  |  |  |  |
| BE-T93 |  |  |  |  |  |  |
| BE-T94 |  |  |  |  |  |  |
| BE-T95 |  |  |  |  |  |  |
| BE-T96 |  |  |  |  |  |  |
| BE-T97 |  |  |  |  |  |  |
| BE-T98 |  |  |  |  |  |  |
| BE-T99 |  |  |  |  |  |  |
| BE-T100 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

BUSINESS ENHANCER DATASHEET

BusinessEnhancer product line can be leveraged to increase profitability in any business and generates detailed reports and attractive dashboards.

Business Enhancer Features

Connects and collects data from all types of datasources such as RDBMS databases, CSV format files and Microsoft Excel.   
Organises the raw data into,

* Summary Tables,
* Contingency Tables,
* Multidimensional Contingency Tables,
* Pivot Tables,
* Frequency Distribution Tables.

Provides visualisations creating,

* Piecharts,
* Barcharts,
* Paretocharts,
* Scatterplots,
* Graphs,
* Time-Series Plots,
* Stem-And-Leaf Displays,
* Normal Probability Plots,
* Histograms,
* Control charts - X-Chart, r chart, p chart and ? chart.

Statistical analysis,

**NORMSINV** is an Microsoft **Excel function** that delivers the inverse of the cummulative standarized normal distribution. You enter the “probability that a value Z is up to…” and it returns that value Z (in terms of “sigmas”, because it is the standarized distribution with average 0 and sigma 1).

|  |  |  |  |
| --- | --- | --- | --- |
| #ifndef Pi  #define Pi 3.141592653589793238462643  #endif  double cnd\_manual(double x)  {  double L, K, w ;  /\* constants \*/  double const a1 = 0.31938153, a2 = -0.356563782, a3 = 1.781477937;  double const a4 = -1.821255978, a5 = 1.330274429;  L = fabs(x);  K = 1.0 / (1.0 + 0.2316419 \* L);  w = 1.0 - 1.0 / sqrt(2 \* Pi) \* exp(-L \*L / 2) \* (a1 \* K + a2 \* K \*K + a3 \* pow(K,3) + a4 \* pow(K,4) + a5 \* pow(K,5));  if (x < 0 ){  w= 1.0 - w;  }  return w;  }   |  |  | | --- | --- | | [share](http://stackoverflow.com/a/3512564)[improve this answer](http://stackoverflow.com/posts/3512564/edit) | answered Aug 18 '10 at 13:27  [[https://www.gravatar.com/avatar/c191b7663bac34a35bba3d096712c3c1?s=32&d=identicon&r=PG](http://stackoverflow.com/users/216314/tyler-brock)](http://stackoverflow.com/users/216314/tyler-brock)  [Tyler Brock](http://stackoverflow.com/users/216314/tyler-brock)  15.7k75057 | | |
|  | | |  |  |  |  | | --- | --- | --- | --- | | |  |  | | --- | --- | | 6 |  | | ouch... don't use pow, use Horner's rule. I downvote until this is corrected (please notify me). – [Alexandre C.](http://stackoverflow.com/users/373025/alexandre-c) [Mar 11 '11 at 10:31](http://stackoverflow.com/questions/2328258/cumulative-normal-distribution-function-in-c-c#comment5938898_3512564) | | |  |  | | --- | --- | | 3 |  | | I was going for readability, request denied. – [Tyler Brock](http://stackoverflow.com/users/216314/tyler-brock) [May 27 '11 at 14:14](http://stackoverflow.com/questions/2328258/cumulative-normal-distribution-function-in-c-c#comment7147214_3512564) | | |  |  | | --- | --- | | 5 |  | | this code will lose precision. Horner's rule is stabler (and also faster). – [Alexandre C.](http://stackoverflow.com/users/373025/alexandre-c) [May 27 '11 at 19:04](http://stackoverflow.com/questions/2328258/cumulative-normal-distribution-function-in-c-c#comment7152053_3512564) | | |  |  | | --- | --- | | 1 |  | | why not just use double pK3 = K\*K\*K and so on? – [Daniel Bonetti](http://stackoverflow.com/users/1720085/daniel-bonetti) [Jun 18 '15 at 18:31](http://stackoverflow.com/questions/2328258/cumulative-normal-distribution-function-in-c-c#comment49882968_3512564) | | |  |  | | --- | --- | |  |  | | (As far as I know) But pow is defined as a macro, I think in order to allow good implementations to optimize common powers, such as 2 and 3. So don't give up on pow too soon! – [Aaron McDaid](http://stackoverflow.com/users/146041/aaron-mcdaid) [Dec 4 '15 at 9:09](http://stackoverflow.com/questions/2328258/cumulative-normal-distribution-function-in-c-c#comment55921561_3512564) |   add a comment |
| up vote 5 down vote | The implementations of the normal CDF given here are single precision approximations that have had float replaced with double and hence are only accurate to 7 or 8 significant (decimal) figures. For a VB implementation of Hart's double precision approximation, see figure 2 of West's [Better approximations to cumulative normal functions](http://www.wilmott.com/pdfs/090721_west.pdf).  **Edit**: My translation of West's implementation into C++:  double  phi(double x)  {  static const double RT2PI = sqrt(4.0\*acos(0.0));  static const double SPLIT = 7.07106781186547;  static const double N0 = 220.206867912376;  static const double N1 = 221.213596169931;  static const double N2 = 112.079291497871;  static const double N3 = 33.912866078383;  static const double N4 = 6.37396220353165;  static const double N5 = 0.700383064443688;  static const double N6 = 3.52624965998911e-02;  static const double M0 = 440.413735824752;  static const double M1 = 793.826512519948;  static const double M2 = 637.333633378831;  static const double M3 = 296.564248779674;  static const double M4 = 86.7807322029461;  static const double M5 = 16.064177579207;  static const double M6 = 1.75566716318264;  static const double M7 = 8.83883476483184e-02;  const double z = fabs(x);  double c = 0.0;  if(z<=37.0)  {  const double e = exp(-z\*z/2.0);  if(z<SPLIT)  {  const double n = (((((N6\*z + N5)\*z + N4)\*z + N3)\*z + N2)\*z + N1)\*z + N0;  const double d = ((((((M7\*z + M6)\*z + M5)\*z + M4)\*z + M3)\*z + M2)\*z + M1)\*z + M0;  c = e\*n/d;  }  else  {  const double f = z + 1.0/(z + 2.0/(z + 3.0/(z + 4.0/(z + 13.0/20.0))));  c = e/(RT2PI\*f);  }  }  return x<=0.0 ? c : 1-c;  }  Note that I have rearranged expressions into the more familiar forms for series and continued fraction approximations. The last magic number in West's code is the square root of 2π, which I've deferred to the compiler on the first line by exploiting the identity acos(0) = ½ π. I've triple checked the magic numbers, but there's always the chance that I've mistyped something. If you spot a typo, please comment!  The results for the test data John Cook used in his answer are  x phi Mathematica  -3 1.3498980316301150e-003 0.00134989803163  -1 1.5865525393145702e-001 0.158655253931  0 5.0000000000000000e-001 0.5  0.5 6.9146246127401301e-001 0.691462461274  2.1 9.8213557943718344e-001 0.982135579437  I take some small comfort from the fact that they agree to all of the digits given for the Mathematica results.   |  |  |  | | --- | --- | --- | | [share](http://stackoverflow.com/a/23119456)[improve this answer](http://stackoverflow.com/posts/23119456/edit) | [edited Apr 18 '14 at 7:24](http://stackoverflow.com/posts/23119456/revisions) | answered Apr 16 '14 at 20:26  [[https://i.stack.imgur.com/X7BoO.png?s=32&g=1](http://stackoverflow.com/users/3526684/thus-spake-a-k)](http://stackoverflow.com/users/3526684/thus-spake-a-k)  [thus spake a.k.](http://stackoverflow.com/users/3526684/thus-spake-a-k)  1,08659 | | |
|  | |  |  |  |  | | --- | --- | --- | --- | | |  |  | | --- | --- | |  |  | | How does this compare to erfc ? – [Johan Lundberg](http://stackoverflow.com/users/1149664/johan-lundberg) [May 24 at 19:58](http://stackoverflow.com/questions/2328258/cumulative-normal-distribution-function-in-c-c#comment62350739_23119456) | | |  |  | | --- | --- | |  |  | | That would depend upon the precision guarantees of erfc. There's certainly going to be a slight rounding of the product of the argument and the square root of one half.which may propagate to the final value. Hart's algorithm is claimed to be accurate to double precision for *every* argument, although I've not independantly verified that. In any event both will be much, *much* better than single precision approximations in which float is replaced with double! – [thus spake a.k.](http://stackoverflow.com/users/3526684/thus-spake-a-k) [May 25 at 20:59](http://stackoverflow.com/questions/2328258/cumulative-normal-distribution-function-in-c-c#comment62396979_23119456) |   add a comment | |

|  |  |
| --- | --- |
| up vote 3 down vote | From NVIDIA CUDA samples:  static double CND(double d)  {  const double A1 = 0.31938153;  const double A2 = -0.356563782;  const double A3 = 1.781477937;  const double A4 = -1.821255978;  const double A5 = 1.330274429;  const double RSQRT2PI = 0.39894228040143267793994605993438;  double  K = 1.0 / (1.0 + 0.2316419 \* fabs(d));  double  cnd = RSQRT2PI \* exp(- 0.5 \* d \* d) \*  (K \* (A1 + K \* (A2 + K \* (A3 + K \* (A4 + K \* A5)))));  if (d > 0)  cnd = 1.0 - cnd;  return cnd;  } |

Power = -(1/2)[(X-µ)/rho]2

-(1/2)[(X)]2

*F*(*x*)=∫*x*−∞ ePower

f(X) = 1/(√(2\*∏\*rho))

factor

= 1/(√(2\*∏\*rho))

f(X) = factor \* ePower

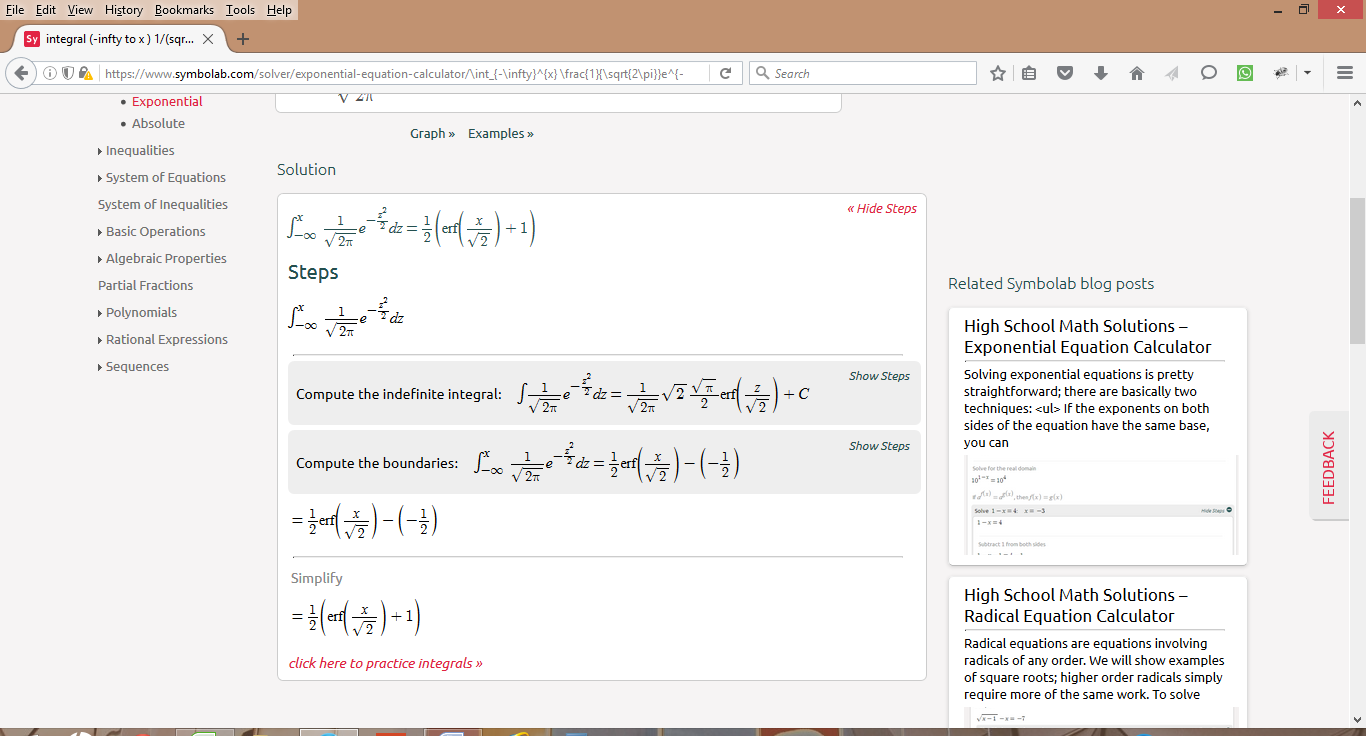
Normal Distribution

The Normal distribution is a continuous probability function that is given by the formula:

Normal Probability Density Function

where μ is the mean of the distribution, σ2 is the variance, and x is the independent variable for which you want to evaluate the function.

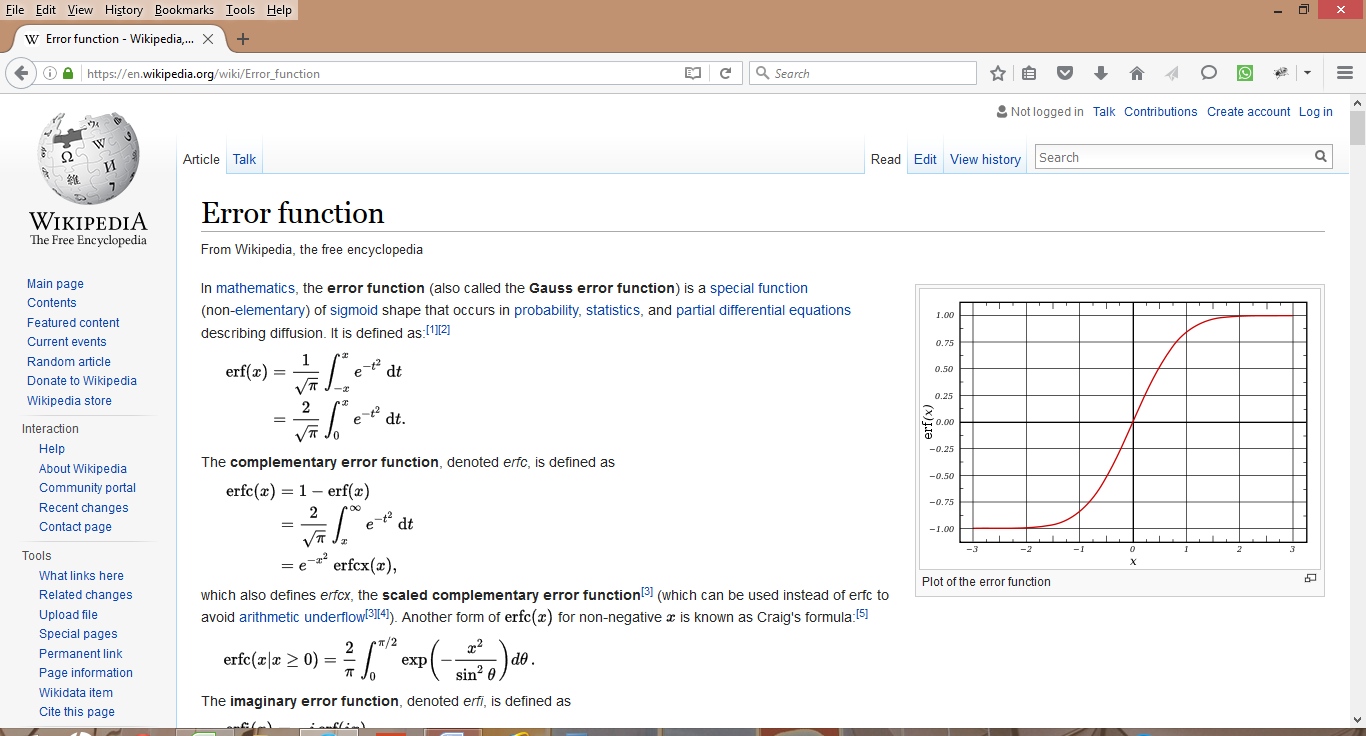
The Cumulative Normal Distribution function is given by the integral, from -∞ to x, of the Normal Probability Density function.



integral from -infinity to x (1/((sqrt(2pi)r) \* e^(-1/2 \* ((x-u)/r)^2) dx

The formula for the cumulative distribution function of the standard normal distribution is

*F*(*x*)=∫*x*−∞ *e*−*x*2/22*π*√



The Excel NORMDIST function calculates the Normal Probability Density Function or the Cumulative Normal Distribution Function for a supplied set of parameters.

The syntax of the function is:

NORMDIST( x, mean, standard\_dev, cumulative )

<https://www.mathsisfun.com/calculus/index.html>

|  |  |
| --- | --- |
| **Taylor Series expansion** | **As** [**Sigma Notation**](https://www.mathsisfun.com/algebra/sigma-notation.html) |
| https://www.mathsisfun.com/algebra/images/taylor-ex.gif | https://www.mathsisfun.com/algebra/images/taylor-ex-sigma.gif |

Normsinv

NORMSINV is an Microsoft Excel function that delivers the inverse of the cummulative standarized normal distribution. You enter the “probability that a value Z is up to…” and it returns that value Z (in terms of “sigmas”, because it is the standarized distribution with average 0 and sigma 1).

Example: NORMSINV(0.5)=0, NORMSINV(0.00135)=-3, NORMSINV(0.9772)=2. NORMSINV(0) and NORMSINV(1) will return error, because they correspond to – infinte sigmas and +infinte sigmas.

Normsinv

Calculate the inverse of the cummulative standarized normal distribution.

You enter the “probability that a value Z is up to…” and it returns that value Z (in terms of “sigmas”, because it is the standarized distribution with average 0 and sigma 1).

Example: NORMSINV(0.5)=0, NORMSINV(0.00135)=-3, NORMSINV(0.9772)=2. NORMSINV(0) and NORMSINV(1) will return error, because they correspond to – infinte sigmas and +infinte sigmas.

Files

Filetype;AbsoluteFilePathName:login;password

Sheetname;fieldtype;fieldname

|---- FileType1

|---- AbsoluteFilePathName

|----- SheetName1

|-----FieldType1

|----FieldName1

|----FieldName2

|-----FieldType2

|----FieldName3

|----FieldName4

|-----SheetName2

|-----FieldType1

|----FieldName1

|----FieldName2

|-----FieldType2

|----FieldName3

|----FieldName4

CUSTOMER QUESTIONAIRE

1. What problems do you face in your daily business?
2. How do you circumvent your problems?
3. How many employees do you have?
4. How many locations do you have?
5. How many departments/divisions are there?
6. How many subsidiary companies are there?
7. How do you manage your suppliers?
8. How do you manage your distributors?
9. How do you manage your warehouses?
10. How do you manage your logistics?
11. Do you face bullwhip effects in your scm?
12. How do you track your products from manufacturing unit up to retailers/dealers?
13. How many manufacturing units do you have?
14. How many customers do you have in each branch?
15. How many customer complaints come from different offices/branches?
16. How do you monitor each employee?
17. How do you manage their customer complaints?
18. How many are attended to?
19. How many are unattended?
20. How many complaints cannot be attended to? Reasons.
21. What are the data you collect daily throughout your organisations?
22. Where do you store your data?
23. What spreadsheet tools do you use?
24. What analytics tool do you use?
25. Are you considering using analytics tool?
26. Do you know what an analytics tool is?
27. What analytics tools are you aware of?
28. When would you like to see the demonstration of Business Enhancer?

str.split(",(?=(?:[^\"]\*\"[^\"]\*\")\*[^\"]\*$)");

This splits the string on , that is followed by an even number of double quotes. In other words, it splits on comma outside the double quotes. This will work provided you have balanced quotes in your string.

Explanation:

, // Split on comma

(?= // Followed by

(?: // Start a non-capture group

[^"]\* // 0 or more non-quote characters

" // 1 quote

[^"]\* // 0 or more non-quote characters

" // 1 quote

)\* // 0 or more repetition of non-capture group (multiple of 2 quotes will be even)

[^"]\* // Finally 0 or more non-quotes

$ // Till the end (This is necessary, else every comma will satisfy the condition)

)

String[] arr = str.split("(?x) " +

", " + // Split on comma

"(?= " + // Followed by

" (?: " + // Start a non-capture group

" [^\"]\* " + // 0 or more non-quote characters

" \" " + // 1 quote

" [^\"]\* " + // 0 or more non-quote characters

" \" " + // 1 quote

" )\* " + // 0 or more repetition of non-capture group (multiple of 2 quotes will be even)

" [^\"]\* " + // Finally 0 or more non-quotes

" $ " + // Till the end (This is necessary, else every comma will satisfy the condition)

") " // End look-ahead

);

**Why Split when you can Match?**

Resurrecting this question because for some reason, the easy solution wasn't mentioned. Here is our beautifully compact regex:

"[^"]\*"|[^,]+

This will match all the desired fragments ([see demo](http://regex101.com/r/dM3wM7)).

**Explanation**

* With "[^"]\*", we match complete "double-quoted strings"
* or |
* we match [^,]+ any characters that are not a comma.

A possible refinement is to improve the string side of the alternation to allow the quoted strings to include escaped quotes.

String[] arr = str.split("(?x) " +

", " + // Split on comma

"(?= " + // Followed by

" (?: " + // Start a non-capture group

" [^\"]\* " + // 0 or more non-quote characters

" \" " + // 1 quote

" [^\"]\* " + // 0 or more non-quote characters

" \" " + // 1 quote

" )\* " + // 0 or more repetition of non-capture group (multiple of 2 quotes will be even)

" [^\"]\* " + // Finally 0 or more non-quotes

" $ " + // Till the end (This is necessary, else every comma will satisfy the condition)

") " // End look-ahead

);

Regex stringDatePattern=new Regex("\*[0-9][0-9][/][0-9][0-9][/][0-9][0-9[0-9][0-9]]\*");

private static String driverName = "org.apache.hive.jdbc.HiveDriver"

jdbc:hive2://localhost:10000/default", "", ""

import java.sql.SQLException;

import java.sql.Connection;

import java.sql.ResultSet;

import java.sql.Statement;

import java.sql.DriverManager;

public class HiveJdbcClient {

private static String driverName = "org.apache.hive.jdbc.HiveDriver";

/\*\*

\* @param args

\* @throws SQLException

\*/

public static void main(String[] args) throws SQLException {

try {

Class.forName(driverName);

} catch (ClassNotFoundException e) {

// TODO Auto-generated catch block

e.printStackTrace();

System.exit(1);

}

//replace "hive" here with the name of the user the queries should run as

Connection con = DriverManager.getConnection("jdbc:hive2://localhost:10000/default", "hive", "");

Statement stmt = con.createStatement();

String tableName = "testHiveDriverTable";

stmt.execute("drop table if exists " + tableName);

stmt.execute("create table " + tableName + " (key int, value string)");

// show tables

String sql = "show tables '" + tableName + "'";

System.out.println("Running: " + sql);

ResultSet res = stmt.executeQuery(sql);

if (res.next()) {

System.out.println(res.getString(1));

}

// describe table

sql = "describe " + tableName;

System.out.println("Running: " + sql);

res = stmt.executeQuery(sql);

while (res.next()) {

System.out.println(res.getString(1) + "\t" + res.getString(2));

}

// load data into table

// NOTE: filepath has to be local to the hive server

// NOTE: /tmp/a.txt is a ctrl-A separated file with two fields per line

String filepath = "/tmp/a.txt";

sql = "load data local inpath '" + filepath + "' into table " + tableName;

System.out.println("Running: " + sql);

stmt.execute(sql);

// select \* query

sql = "select \* from " + tableName;

System.out.println("Running: " + sql);

res = stmt.executeQuery(sql);

while (res.next()) {

System.out.println(String.valueOf(res.getInt(1)) + "\t" + res.getString(2));

}

// regular hive query

sql = "select count(1) from " + tableName;

System.out.println("Running: " + sql);

res = stmt.executeQuery(sql);

while (res.next()) {

System.out.println(res.getString(1));

}

Write a program to construct an optimized DECISION TREE for a given training data and

by using any attribute selection measure

Decision tree learning algorithm

**ij> connect 'jdbc:derby:FSDB;create=true;';**

**ij> show connections**

> ;

CONNECTION0\* - jdbc:derby:FSDB

\* = current connection

**ij> create table balancesheet(companyname varchar(50), capitalwip double, sharecapital double, assets double, liabilities double, equity double, financialyear date, noofshares int, tradepayables real, tradereceivables float);**

0 rows inserted/updated/deleted

**ij> insert into balancesheet values ('B E L', 100002.3341, 217356.0000032, 50132243, 20001.1, 3045221.8265, '2017-03-24', 45323111, 50543, 637726);**

1 row inserted/updated/deleted

ij> insert into balancesheet values ('N A L', 20002.3341, 17356.032, 132243, 4231.1, 353421.8265, '2017-03-23', 4323111, 50543, 637726);

1 row inserted/updated/deleted

ij> insert into balancesheet values ('N A L', 20002.3341, 17456.032, 132243, 4531.1, 323421.8265, '2017-03-24', 323111, 543, 37726);

1 row inserted/updated/deleted

ij> select \* from balancesheet;

COMPANYNAME |CAPITALWIP |SHARECAPITAL |ASSETS |LIABILITIES |EQUITY |FINANCIAL&|NOOFSHARES |T

RADEPAYABLES |TRADERECEIVABLES

B E L |100002.3341 |217356.0000032 |5.0132243E7 |20001.1 |3045221.8265 |2017-03-24|45323111 |5

0543.0 |637726.0

N A L |20002.3341 |17356.032 |132243.0 |4231.1 |353421.8265 |2017-03-23|4323111 |5

0543.0 |637726.0

N A L |20002.3341 |17456.032 |132243.0 |4531.1 |323421.8265 |2017-03-24|323111 |5

43.0 |37726.0

3 rows selected

ij>

A discrete random variable is a whole number that can take any one of the possible values and has to be decided on real-time basis.

For example number of MPs elected to Parliament in 2020 is not random variable.

However number of MPs attending Parliament session today is a random variable and a discrete one.

# Test Cases

## Test Cases for testing creation of Multi-Contingency Tables with one database

|  |  |  |  |
| --- | --- | --- | --- |
| **Test Case ID** | **Input** | **Result(Success/Failure)** | **Bug Description** |
| MCT-1 | Two categorical variables | Success | Must add rowname also along with column name |
| MCT-2 | 10 categorical variables | Failure | Erratic behaviour. One DB – data from two tables |
| MCT-3 | 1 categorical variable and 1 numeric variable | Success |  |
| MCT-4 | 1 categorical variable and 2 numeric variables | Success |  |
| MCT-5 | 1 categorical variable and 10 numeric variables | Failure | 1 cat + 9 num from different tables |
| MCT-6 | 2 categorical variable and 1 numeric variable | Failure |  |
| MCT-7 | 2 categorical variable and 2 numeric variables | Failure |  |
| MCT-8 | 2 categorical variable and 10 numeric variables | Failure |  |
| MCT-9 | 2 categorical variable and 1 numeric variable | Failure |  |
| MCT-10 | 10 categorical variable and 1 numeric variable | Failure |  |
| MCT-11 | 2 categorical variable and 10 numeric variable | Failure |  |
| MCT-12 | 10 categorical variable and 10 numeric variable | Failure |  |
| MCT-13 | Two numeric variables | Success |  |
| MCT-14 | 10 numeric variables | Success | Tested with 9 numeric variables from different tables, panel size is very small |
| MCT-15 | 1 numeric variable and 1 categorical variable | Success |  |
| MCT-16 | 1 numeric variable and 2 categorical variables | Failure |  |
| MCT-17 | 1 numeric variable and 10 categorical variables | Failure |  |
| MCT-18 | 2 numeric variable and 1 categorical variable | Success |  |
| MCT-19 | 2 numeric variable and 2 categorical variables | Failure |  |
| MCT-20 | 2 numeric variable and 10 categorical variables | Failure |  |
| MCT-21 | 2 numeric variable and 1 categorical variable | Failure |  |
| MCT-22 | 10 numeric variable and 1 categorical variable | Failure |  |
| MCT-23 | 2 numeric variable and 10 categorical variable | Failure |  |
| MCT-24 | 10 numeric variable and 10 categorical variable | Failure |  |

## Test Cases for testing creation of Multi-Contingency Tables with multiple databases

|  |  |  |  |
| --- | --- | --- | --- |
| **Test Case ID** | **Input** | **Result(Success/Failure)** | **Bug Description** |
| MCT-1 | Two categorical variables | Failure |  |
| MCT-2 | 10 categorical variables | Failure |  |
| MCT-3 | 1 categorical variable and 1 numeric variable | Failure |  |
| MCT-4 | 1 categorical variable and 2 numeric variables | Failure |  |
| MCT-5 | 1 categorical variable and 10 numeric variables | Failure |  |
| MCT-6 | 2 categorical variable and 1 numeric variable | Failure |  |
| MCT-7 | 2 categorical variable and 2 numeric variables | Failure |  |
| MCT-8 | 2 categorical variable and 10 numeric variables | Failure |  |
| MCT-9 | 2 categorical variable and 1 numeric variable | Failure |  |
| MCT-10 | 10 categorical variable and 1 numeric variable | Failure |  |
| MCT-11 | 2 categorical variable and 10 numeric variable | Failure |  |
| MCT-12 | 10 categorical variable and 10 numeric variable | Failure |  |
| MCT-13 | Two numeric variables | Failure |  |
| MCT-14 | 10 numeric variables | Success | Panel size must be modifiable |
| MCT-15 | 1 numeric variable and 1 categorical variable | Failure |  |
| MCT-16 | 1 numeric variable and 2 categorical variables | Failure |  |
| MCT-17 | 1 numeric variable and 10 categorical variables | Failure |  |
| MCT-18 | 2 numeric variable and 1 categorical variable | Failure |  |
| MCT-19 | 2 numeric variable and 2 categorical variables | Failure |  |
| MCT-20 | 2 numeric variable and 10 categorical variables | Failure |  |
| MCT-21 | 2 numeric variable and 1 categorical variable | Failure |  |
| MCT-22 | 10 numeric variable and 1 categorical variable | Failure |  |
| MCT-23 | 2 numeric variable and 10 categorical variable | Failure |  |
| MCT-24 | 10 numeric variable and 10 categorical variable | Failure |  |

## Test Cases for testing creation of Multi-Contingency Tables with one database and one file

|  |  |  |  |
| --- | --- | --- | --- |
| **Test Case ID** | **Input** | **Result(Success/Failure)** | **Bug Description** |
| MCT-1 | Two categorical variables | Success |  |
| MCT-2 | 10 categorical variables | Failure | Unusal output |
| MCT-3 | 1 categorical variable and 1 numeric variable | Success |  |
| MCT-4 | 1 categorical variable and 2 numeric variables | Failure |  |
| MCT-5 | 1 categorical variable and 10 numeric variables | Failure |  |
| MCT-6 | 2 categorical variable and 1 numeric variable | Failure |  |
| MCT-7 | 2 categorical variable and 2 numeric variables | Failure |  |
| MCT-8 | 2 categorical variable and 10 numeric variables | Failure |  |
| MCT-9 | 2 categorical variable and 1 numeric variable | Failure |  |
| MCT-10 | 10 categorical variable and 1 numeric variable | Failure |  |
| MCT-11 | 2 categorical variable and 10 numeric variable |  |  |
| MCT-12 | 10 categorical variable and 10 numeric variable |  |  |
| MCT-13 | Two numeric variables | Failure |  |
| MCT-14 | 10 numeric variables | Failure |  |
| MCT-15 | 1 numeric variable and 1 categorical variable | Success |  |
| MCT-16 | 1 numeric variable and 2 categorical variables | Failure |  |
| MCT-17 | 1 numeric variable and 10 categorical variables | Failure |  |
| MCT-18 | 2 numeric variable and 1 categorical variable | Failure |  |
| MCT-19 | 2 numeric variable and 2 categorical variables | Failure |  |
| MCT-20 | 2 numeric variable and 10 categorical variables | Failure |  |
| MCT-21 | 2 numeric variable and 1 categorical variable | Failure |  |
| MCT-22 | 10 numeric variable and 1 categorical variable | Failure |  |
| MCT-23 | 2 numeric variable and 10 categorical variable | Failure |  |
| MCT-24 | 10 numeric variable and 10 categorical variable | Failure |  |

## Test Cases for testing creation of Multi-Contingency Tables with one file

|  |  |  |  |
| --- | --- | --- | --- |
| **Test Case ID** | **Input** | **Result(Success/Failure)** | **Bug Description** |
| MCT-1 | Two categorical variables | Success |  |
| MCT-2 | 10 categorical variables | Success |  |
| MCT-3 | 1 categorical variable and 1 numeric variable | Success |  |
| MCT-4 | 1 categorical variable and 2 numeric variables | Success |  |
| MCT-5 | 1 categorical variable and 10 numeric variables | Success |  |
| MCT-6 | 2 categorical variable and 1 numeric variable | Success |  |
| MCT-7 | 2 categorical variable and 2 numeric variables | Success |  |
| MCT-8 | 2 categorical variable and 10 numeric variables | Success |  |
| MCT-9 | 2 categorical variable and 1 numeric variable | Success |  |
| MCT-10 | 10 categorical variable and 1 numeric variable | Success |  |
| MCT-11 | 2 categorical variable and 10 numeric variable | Success |  |
| MCT-12 | 10 categorical variable and 10 numeric variable | Success |  |
| MCT-13 | Two numeric variables | Success |  |
| MCT-14 | 10 numeric variables | Success |  |
| MCT-15 | 1 numeric variable and 1 categorical variable | Success |  |
| MCT-16 | 1 numeric variable and 2 categorical variables | Success |  |
| MCT-17 | 1 numeric variable and 10 categorical variables | Success |  |
| MCT-18 | 2 numeric variable and 1 categorical variable | Success |  |
| MCT-19 | 2 numeric variable and 2 categorical variables | Success |  |
| MCT-20 | 2 numeric variable and 10 categorical variables | Success |  |
| MCT-21 | 2 numeric variable and 1 categorical variable | Success |  |
| MCT-22 | 10 numeric variable and 1 categorical variable | Success |  |
| MCT-23 | 2 numeric variable and 10 categorical variable | Success |  |
| MCT-24 | 10 numeric variable and 10 categorical variable | Success |  |

## Test Cases for testing creation of Multi-Contingency Tables with multiple files

|  |  |  |  |
| --- | --- | --- | --- |
| **Test Case ID** | **Input** | **Result(Success/Failure)** | **Bug Description** |
| MCT-1 | Two categorical variables | Success |  |
| MCT-2 | 10 categorical variables | Success |  |
| MCT-3 | 1 categorical variable and 1 numeric variable | Success |  |
| MCT-4 | 1 categorical variable and 2 numeric variables | Success |  |
| MCT-5 | 1 categorical variable and 10 numeric variables | Success |  |
| MCT-6 | 2 categorical variable and 1 numeric variable | Success |  |
| MCT-7 | 2 categorical variable and 2 numeric variables | Success |  |
| MCT-8 | 2 categorical variable and 10 numeric variables | Success |  |
| MCT-9 | 2 categorical variable and 1 numeric variable | Success |  |
| MCT-10 | 10 categorical variable and 1 numeric variable | Success |  |
| MCT-11 | 2 categorical variable and 10 numeric variable | Success |  |
| MCT-12 | 10 categorical variable and 10 numeric variable | Success |  |
| MCT-13 | Two numeric variables | Success |  |
| MCT-14 | 10 numeric variables | Success |  |
| MCT-15 | 1 numeric variable and 1 categorical variable | Success |  |
| MCT-16 | 1 numeric variable and 2 categorical variables | Success |  |
| MCT-17 | 1 numeric variable and 10 categorical variables | Success |  |
| MCT-18 | 2 numeric variable and 1 categorical variable | Success |  |
| MCT-19 | 2 numeric variable and 2 categorical variables | Success |  |
| MCT-20 | 2 numeric variable and 10 categorical variables | Success |  |
| MCT-21 | 2 numeric variable and 1 categorical variable | Success |  |
| MCT-22 | 10 numeric variable and 1 categorical variable | Success |  |
| MCT-23 | 2 numeric variable and 10 categorical variable | Success |  |
| MCT-24 | 10 numeric variable and 10 categorical variable | Success |  |

# Test Cases for testing creation of Pivot Tables with one database

|  |  |  |  |
| --- | --- | --- | --- |
| **Test Case ID** | **Input** | **Result(Success/Failure)** | **Bug Description** |
| PT-1 | Count of 1 cat |  |  |
| PT-2 | Count of 1 num |  |  |
| PT-3 | Sum of 1 cat |  |  |
| PT-4 | Sum of 1 num |  |  |
| PT-5 | Sum of 10 cat |  |  |
| PT-6 | Sum of 10 num |  |  |
| PT-7 | Avg of 1 cat |  |  |
| PT-8 | Avg of 1 num |  |  |
| PT-9 | Avg of 10 cat |  |  |
| PT-10 | Avg of 10 num |  |  |
| PT-11 | Max of 1 cat |  |  |
| PT-12 | Max of 1 num |  |  |
| PT-13 | Max of 10 cat |  |  |
| PT-14 | Max of 10 num |  |  |
| PT-15 | Min of 1 cat |  |  |
| PT-16 | Min of 1 num |  |  |
| PT-17 | Min of 10 cat |  |  |
| PT-18 | Min of 10 num |  |  |
| PT-19 | Product of 1 cat |  |  |
| PT-20 | Product of 1 num |  |  |
| PT-21 | Product of 10 cat |  |  |
| PT-22 | Product of 10 num |  |  |
| PT-23 |  |  |  |
| PT-24 |  |  |  |
| PT-25 |  |  |  |
| PT-26 |  |  |  |

# Test Cases for testing creation of Multi-Contingency Tables with multiple databases

# Test Cases for testing creation of Multi-Contingency Tables with one database and one file

# Test Cases for testing creation of Multi-Contingency Tables with one file

# Test Cases for testing creation of Multi-Contingency Tables with multiple files

Dashboard Tasks

1. Add tables to dashboard – Sunday – Completed
2. Save the dashboard to files in various formats. - Monday
3. Marshalling and unmarshalling of dashboard vectors(tables and charts) - Tuesday
4. Drag and drop of JFXPanels and JPanels in dashboard - Wednesday
5. Convert the charts and tables into images in bmp, png and jpeg formats. - Monday
6. Clear the dashboard - Sunday

R&D

1. Test dashboard with different layout managers. – GridLayout, Vbox worked

GridLayout selected. - Completed

# CSS Styles

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Class Name** | **Parent Class** | **DEFAULT\_STYLE\_CLASS** | **PseudoClasses** | **StyleableProperties** |
| TableView | Control | table-view | cell-selection  row-selection | -fx-fixed-cell-size |
| TableRow | IndexedCell | table-row-cell |  |  |
| IndexedCell | Cell | indexed-cell | odd even |  |
| Cell | Labeled | cell | selected focused empty filled |  |
| Labeled | Control |  |  | -fx-font -fx-alignment -fx-text-alignment -fx-text-fill -fx-text-overrun -fx-ellipsis-string -fx-wrap-text -fx-graphic -fx-underline -fx-line-spacing -fx-content-display -fx-label-padding -fx-graphic-text-gap |
| Control | Region |  |  | -fx-skin |
| Region | Parent -> Node |  |  | -fx-background-color  -fx-background-insets  -fx-background-radius  -fx-background-image  -fx-background-position  -fx-background-repeat  -fx-background-size  -fx-border-color  -fx-border-insets  -fx-border-radius  -fx-border-style  -fx-border-width  -fx-border-image-source  -fx-border-image-insets  -fx-border-image-repeat  -fx-border-image-slice  -fx-border-image-width  -fx-padding  -fx-position-shape  -fx-scale-shape  -fx-shape  -fx-snap-to-pixel  -fx-background-fills  -fx-background-images  -fx-stroke-borders  -fx-image-borders  -fx-padding -fx-opaque-insets -fx-region-background -fx-region-border -fx-shape -fx-scale-shape -fx-position-shape -fx-cache-shape -fx-snap-to-pixel -fx-min-height -fx-pref-height -fx-max-height -fx-min-width -fx-pref-width -fx-max-width |
| Node |  |  | my-state hover pressed disabled focused show-mnemonics | -fx-cursor -fx-effect -fx-focus-traversable -fx-opacity -fx-blend-mode -fx-rotate -fx-scale-x -fx-scale-y -fx-scale-z -fx-translate-x -fx-translate-y -fx-translate-z visibility |
| TableColumn | TableColumnBase | table-column |  |  |
| TableCell | IndexedCell | table-cell | last-visible |  |
| ScrollPane | Control | scroll-pane | pannable fitToWidth fitToHeight | -fx-hbar-policy -fx-vbar-policy -fx-fit-to-width -fx-fit-to-height -fx-pannable |
| GridPane | Pane->Region |  |  | -fx-grid-lines-visible -fx-hgap -fx-alignment -fx-vgap |
| Group | Parent |  |  |  |

# Appendix

# https://www.w3.org/TR/css3-selectors/#context

A Selector represents a structure. This structure can be used as a condition (e.g. in a CSS rule) that determines which elements a selector matches in the document tree, or as a flat description of the HTML or XML fragment corresponding to that structure.

Selectors may range from simple element names to rich contextual representations.

The following table summarizes the Selector syntax:

| **Pattern** | **Meaning** | **Described in section** | **First defined in CSS level** |
| --- | --- | --- | --- |
| \* | any element | [Universal selector](https://www.w3.org/TR/css3-selectors/#universal-selector) | 2 |
| E | an element of type E | [Type selector](https://www.w3.org/TR/css3-selectors/#type-selectors) | 1 |
| E[foo] | an E element with a "foo" attribute | [Attribute selectors](https://www.w3.org/TR/css3-selectors/#attribute-selectors) | 2 |
| E[foo="bar"] | an E element whose "foo" attribute value is exactly equal to "bar" | [Attribute selectors](https://www.w3.org/TR/css3-selectors/#attribute-selectors) | 2 |
| E[foo~="bar"] | an E element whose "foo" attribute value is a list of whitespace-separated values, one of which is exactly equal to "bar" | [Attribute selectors](https://www.w3.org/TR/css3-selectors/#attribute-selectors) | 2 |
| E[foo^="bar"] | an E element whose "foo" attribute value begins exactly with the string "bar" | [Attribute selectors](https://www.w3.org/TR/css3-selectors/#attribute-selectors) | 3 |
| E[foo$="bar"] | an E element whose "foo" attribute value ends exactly with the string "bar" | [Attribute selectors](https://www.w3.org/TR/css3-selectors/#attribute-selectors) | 3 |
| E[foo\*="bar"] | an E element whose "foo" attribute value contains the substring "bar" | [Attribute selectors](https://www.w3.org/TR/css3-selectors/#attribute-selectors) | 3 |
| E[foo|="en"] | an E element whose "foo" attribute has a hyphen-separated list of values beginning (from the left) with "en" | [Attribute selectors](https://www.w3.org/TR/css3-selectors/#attribute-selectors) | 2 |
| E:root | an E element, root of the document | [Structural pseudo-classes](https://www.w3.org/TR/css3-selectors/#structural-pseudos) | 3 |
| E:nth-child(n) | an E element, the n-th child of its parent | [Structural pseudo-classes](https://www.w3.org/TR/css3-selectors/#structural-pseudos) | 3 |
| E:nth-last-child(n) | an E element, the n-th child of its parent, counting from the last one | [Structural pseudo-classes](https://www.w3.org/TR/css3-selectors/#structural-pseudos) | 3 |
| E:nth-of-type(n) | an E element, the n-th sibling of its type | [Structural pseudo-classes](https://www.w3.org/TR/css3-selectors/#structural-pseudos) | 3 |
| E:nth-last-of-type(n) | an E element, the n-th sibling of its type, counting from the last one | [Structural pseudo-classes](https://www.w3.org/TR/css3-selectors/#structural-pseudos) | 3 |
| E:first-child | an E element, first child of its parent | [Structural pseudo-classes](https://www.w3.org/TR/css3-selectors/#structural-pseudos) | 2 |
| E:last-child | an E element, last child of its parent | [Structural pseudo-classes](https://www.w3.org/TR/css3-selectors/#structural-pseudos) | 3 |
| E:first-of-type | an E element, first sibling of its type | [Structural pseudo-classes](https://www.w3.org/TR/css3-selectors/#structural-pseudos) | 3 |
| E:last-of-type | an E element, last sibling of its type | [Structural pseudo-classes](https://www.w3.org/TR/css3-selectors/#structural-pseudos) | 3 |
| E:only-child | an E element, only child of its parent | [Structural pseudo-classes](https://www.w3.org/TR/css3-selectors/#structural-pseudos) | 3 |
| E:only-of-type | an E element, only sibling of its type | [Structural pseudo-classes](https://www.w3.org/TR/css3-selectors/#structural-pseudos) | 3 |
| E:empty | an E element that has no children (including text nodes) | [Structural pseudo-classes](https://www.w3.org/TR/css3-selectors/#structural-pseudos) | 3 |
| E:link E:visited | an E element being the source anchor of a hyperlink of which the target is not yet visited (:link) or already visited (:visited) | [The link pseudo-classes](https://www.w3.org/TR/css3-selectors/#link) | 1 |
| E:active E:hover E:focus | an E element during certain user actions | [The user action pseudo-classes](https://www.w3.org/TR/css3-selectors/#useraction-pseudos) | 1 and 2 |
| E:target | an E element being the target of the referring URI | [The target pseudo-class](https://www.w3.org/TR/css3-selectors/#target-pseudo) | 3 |
| E:lang(fr) | an element of type E in language "fr" (the document language specifies how language is determined) | [The :lang() pseudo-class](https://www.w3.org/TR/css3-selectors/#lang-pseudo) | 2 |
| E:enabled E:disabled | a user interface element E which is enabled or disabled | [The UI element states pseudo-classes](https://www.w3.org/TR/css3-selectors/#UIstates) | 3 |
| E:checked | a user interface element E which is checked (for instance a radio-button or checkbox) | [The UI element states pseudo-classes](https://www.w3.org/TR/css3-selectors/#UIstates) | 3 |
| E::first-line | the first formatted line of an E element | [The ::first-line pseudo-element](https://www.w3.org/TR/css3-selectors/#first-line) | 1 |
| E::first-letter | the first formatted letter of an E element | [The ::first-letter pseudo-element](https://www.w3.org/TR/css3-selectors/#first-letter) | 1 |
| E::before | generated content before an E element | [The ::before pseudo-element](https://www.w3.org/TR/css3-selectors/#gen-content) | 2 |
| E::after | generated content after an E element | [The ::after pseudo-element](https://www.w3.org/TR/css3-selectors/#gen-content) | 2 |
| E.warning | an E element whose class is "warning" (the document language specifies how class is determined). | [Class selectors](https://www.w3.org/TR/css3-selectors/#class-html) | 1 |
| E#myid | an E element with ID equal to "myid". | [ID selectors](https://www.w3.org/TR/css3-selectors/#id-selectors) | 1 |
| E:not(s) | an E element that does not match simple selector s | [Negation pseudo-class](https://www.w3.org/TR/css3-selectors/#negation) | 3 |
| E F | an F element descendant of an E element | [Descendant combinator](https://www.w3.org/TR/css3-selectors/#descendant-combinators) | 1 |
| E > F | an F element child of an E element | [Child combinator](https://www.w3.org/TR/css3-selectors/#child-combinators) | 2 |
| E + F | an F element immediately preceded by an E element | [Adjacent sibling combinator](https://www.w3.org/TR/css3-selectors/#adjacent-sibling-combinators) | 2 |
| E ~ F | an F element preceded by an E element | [General sibling combinator](https://www.w3.org/TR/css3-selectors/#general-sibling-combinators) |  |

# Project Planning for Completion

|  |  |  |
| --- | --- | --- |
| Dates | Task | Status |
| July 5 – 6 | createStackedAndUnstackedViewScene()  orderedArrayViewScene() | Testing |
| July 7 | frequencyDistributionViewScene()  relativeFrequencyDistributionViewScene() | Testing |
| July 8 | percentageDistributionViewScene()  cumulativePercentageDistributionViewScene() | Testing |
| July 9 | Testing of MDCT | Completed |
| July 10 | Testing of PVT |  |
| July 11 | Bug Fixing |  |
| July 12 | multiDimensionalContingencyTableViewScene() |  |
| July 13 | multiDimensionalContingencyTableViewScene() |  |
| July 14 | pivotTableViewScene() |  |
| July 15 | pivotTableViewScene() |  |
| July 16 | pivotTableViewScene() |  |
| July 17 | Analysis |  |
| July 18 | Analysis |  |
| July 19 | Analysis |  |
| July 20 | Analysis |  |
| July 21 | Analysis |  |
| July 22 | Analysis |  |
| July23 | Analysis |  |
| July24 | Analysis |  |
| July 25 | Analysis |  |
| July 26 | Analysis |  |
| July 27 | Analysis |  |
| July 28 | Analysis |  |
| July 29 | Analysis |  |
| July 30 | Analysis |  |
| July 31 | Analysis |  |
| August 1 | Analysis |  |
| August 2 | Analysis |  |
| August 3 | Analysis |  |
| August 4 | Analysis |  |
| August 5 | Analysis |  |
| August 6 | Analysis |  |
| August 7 | Complete Analysis |  |
| August 8 | Begin Report, Saving Dashboard to files |  |
| August 9 | Report, Saving Connections to file |  |
| August 10 | Report, Saving Report to file |  |
| August 11 | Report |  |
| August 12 | Report |  |
| August 13 | Report |  |
| August 14 | Report |  |
| August 15 | Report |  |
| August 16 | Bug Fixing |  |
| August 17 | Bug Fixing |  |
| August 18 | Bug Fixing |  |
| August 19 | Bug Fixing |  |
| August 20 | Bug Fixing |  |
| August 21 | Final Testing |  |
| August 22 | Final Testing |  |
| August 23 | Final Testing |  |
| August 24 | Final Testing |  |
| August 25 | Final Testing |  |
| August 26 | Final Testing |  |
| August 27 | Bug Fixing |  |
| August 28 | Bug Fixing |  |
| August 29 | Bug Fixing |  |
| August 30 | Bug Fixing |  |

# Tasks Remaining

Getting data from Internet – Web Scrapping

Testing for Big Data

Performance Testing

Copyright

IP

Obfuscation

License – Floating, fixed, demo/trial version – 1 user, 5 users, 10 users, 25 users, 50 users

Designing of License server and client

Certificates

Discrete Variables

Number of interruptions per day

Number of cards

Population of a country

Number of tagged forms in a sample

Eg., Tagged and Untagged forms – Event of interest is Tagged

Sample size 4

Population size 10

How many samples of size 4 are possible in population of 10. This type of data can be created in the tool. For that specify the sample size and select a categorical variable and its unique values are filled in a list box and then select whatever unique values you want and create your raw data.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S1** | **S2** | **S3** | **S4** | **Number of Tagged Forms in the order of 4 forms** |
| U | U | U | U | 0 |
| U | U | U | T | 1 |
| U | U | T | T | 2 |
| U | T | U | U | 1 |
| U | T | U | T | 2 |
| U | T | T | U | 2 |
| U | T | T | T | 3 |
| T | U | U | U | 1 |
| T | U | U | T | 2 |
| T | U | T | U | 2 |
| T | U | T | T | 3 |
| T | T | U | U | 2 |
| T | T | U | T | 3 |
| T | T | T | U | 3 |
| T | T | T | T | 4 |

Analysis - > Categorical - > Organised, Visualised - > Summary Table - > Categorical Distribution - > Table, Graph - > View, Add To Dashboard

Analysis - > Numeric - > Discrete, Continuous - > Organised, Visualised - > TypeOfTables -> NameofAnalysis - > Table, Graph, etc -> View, Add To Dashboard,

-> etc.,

Organised - > Ordered Array, Frequency Distribution, Relative Frequency Distribution, Percentage Distribution, Cumulative Percentage Distribution,

FINAL DEADLINE – SEPTEMBER 2017

# Tasks Identified

1. Complete Pivot Table view and addtodashboard
2. Complete Multicontingency table view and addtodashboard using javafx
3. Complete swapping in chart and table dashboards
4. After swapping in saved chart, it is not getting saved, saved chart is empty
5. Add a dashboard to contain both tables and charts
6. Save connections as xml file
7. Save dashboard to pdf
8. Save dashboard to excelsheet
9. Save dashboard to msword
10. Save dashboard to html
11. Save dashboard to image file
12. Save report to pdf
13. Save report to msword
14. Save report to html
15. Analysis of all 21 chapters – including study of each
16. Fix bugs in adding data files
17. Add web scrapping
18. Add getting data from xml file
19. Obfuscate the code
20. Add licenses to the product
21. Create an installer

# Timeframe

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Task to be completed | Time Required  In days | Allotted days | Remarks | Status |
| Complete Pivot Table view and addtodashboard | 2 |  | Study to add columns within columns and rows within rows in javafx table  Add class Pivottable results into javafx table |  |
| Complete Multicontingency table view and addtodashboard using javafx | 2 |  | Convert multicontingency table to javafx table |  |
| Complete swapping in chart and table dashboards | 1 |  |  | Completed |
| After swapping in saved chart, it is not getting saved, saved chart is empty | 1 |  |  |  |
| Add a dashboard to contain both tables and charts | 1 |  |  |  |
| Save connections as xml file | 2 |  |  |  |
| Save dashboard to pdf | 1 |  |  |  |
| Save dashboard to excelsheet | 1 |  |  |  |
| Save dashboard to msword | 1 |  |  |  |
| Save dashboard to html | 1 |  |  |  |
| Save dashboard to image file | 1 |  |  |  |
| Save report to pdf | 1 |  |  |  |
| Save report to msword | 1 |  |  |  |
| Save report to html | 1 |  |  |  |
| Analysis of all 21 chapters – including study of each | 10 | 4-09-2017 – 15-09-2017 |  |  |
| Fix bugs in adding data files | 1 | 10-11-2017 |  | Completed |
| Add web scrapping | 10 |  |  |  |
| Add getting data from xml file | 5 |  |  |  |
| Storing the usecase definitions in xml file | 2 |  |  |  |
| Converting usecases as namevalue pairs in text file to xml | 2 |  |  |  |
| Obfuscate the code | 5 |  |  |  |
| Add licenses to the product | 15 |  |  |  |
| Create an installer | 7 |  |  |  |
| Total | 73 days |  |  |  |

Theoretically, 73 days, 2 ½ months required for entire product to be completed.

Many of the tasks can be completed on same day

Aiming to release the product on November 2nd 2017, dedicated to Rajyothsava Day.

Next deadline is 31 December 2017.

Project Schedule : 10/02/2018 – 15/02/2018

|  |  |  |  |
| --- | --- | --- | --- |
| Save connections as xml file | 1 |  | 11/02/2018 |
| Save dashboard to pdf | 2 |  | 12-13 |
| Save dashboard to excelsheet | 2 |  |  |
| Save dashboard to msword | 2 |  |  |
| Save dashboard to html | 2 |  |  |
| Save dashboard to image file | 2 |  | 14-15 |
| Storing the usecase definitions in xml file | 2 | Debugging in progress | 10/02/2018 |
| Pivot Table | 5 | In progress | 10-15 |

Remaining tasks

|  |  |  |  |
| --- | --- | --- | --- |
| Remaining data types to be debugged and corrected |  |  |  |
| Analysis |  |  | Begin from 16/02/2018 |
| Correct the disconnect of database and file |  |  |  |
| Exchange of tables in dashboard |  |  |  |
| Creating tables and charts on same dashboard |  |  |  |
| Control Charts |  |  |  |
| Trend Line for graph |  |  |  |

Other tasks

|  |  |  |  |
| --- | --- | --- | --- |
| Copyright |  |  |  |
| IP |  |  |  |
| Licenses |  |  |  |
| Obfuscation |  |  |  |

# Design of xml files

1. Configuration file
   1. Database configuration
      1. This is parsed and tree is formed
      2. The tree is saved into xml file
   2. Usecase configuration
      1. Various templates are Stored in this file
      2. Create new templates and store in this file
      3. Contains default templates
2. Data file
   1. Read data from the xml file and create a tree
   2. This data gets added to Data Files tree.
   3. Save the displayed data that is collected after linking into an xml file and excel file

analytics

{

usecases = {"","industry","services",

"financialInstitution",

"governmentDepartments",

"financialStatement"

{

balancesheet = {

"Capital work-in-progress", "Inventories", "Loose tools",

"Share capital", "Term loans from other parties",

"Total Liabilities", "Accounts payable and Accrued Expenses",

"Accrued Income Tax",

"Application money rcvd for allotment of securities & due for refund and interest accrued thereon",

"Asset Turnover", "ASSETS", "Balances with banks",

"Bank deposits with more than twelve months maturity",

"Bonds/debentures", "Brands /trademarks", "Buildings",

"Capital Advances", "Capital Redemption Reserve", "Capital Reserves",

"Cash And Bank Balances", "Cash and cash equivalents", "Cash on hand",

"Cheques", "Company Name", "Computer software", "Copyrights",

"Current Assets", "Current investments", "Current liabilities",

"Current maturities of finance lease obligations",

"Current maturities of long-term debt",

"Current maturity of long-term debt", "Current Ratio",

"Debenture Redemption Reserve", "Debt Service Cover",

"Debt-to-Equity Ratio", "Deferred Income Taxes",

"Deferred payment liabilities", "Deferred tax assets (net)",

"Deferred tax liabilities (Net)", "Deposits",

"Earmarked balances with banks", "EQUITY AND LIABILITIES", "EQUITY",

"Financial-Year", "Finished goods",

"Fixed assets - Accumulated Depreciation",

"Fixed assets - Net Book Value", "Fixed assets - Original Cost",

"Fixed assets - Total", "Furniture and Fixtures", "Goodwill",

"Income received in advance", "Intangible assets under development",

"Intangible assets", "Interest accrued and due on borrowings",

"Interest accrued but not due on borrowings",

"Inventories", "Investment in Preference Shares", "Investment property",

"Investments in debentures or bonds",

"Investments in Equity Instruments",

"Investments in Government or trust securities",

"Investments in Mutual Funds",

"Investments in partnership firms", "Investments in preference shares",

"Investments", "Land", "Licences and franchise",

"Loans and advances from related parties",

"Loans and advances from related parties", "Loans and Notes Payable",

"Loans repayable on demand from banks",

"Loans repayable on demand from other parties",

"Long term maturities of finance lease obligations",

"Long-term borrowings", "Long-term debt",

"Long-term loans and advances", "Long-term provisions",

"Long-term Trade Receivables", "Marketable Securities",

"Mastheads and publishing titles", "Mining rights",

"Money received against share warrants", "Non-current assets ",

"Non-current investments", "Non-current liabilities",

"Office equipment", "Other Assets", "Other current assets",

"Other current liabilities", "Other intangible Assets",

"Other investments", "Other liabilities", "Other loans and advances",

"Other Long term liabilities", "Other non-current assets",

"Other non-current investments", "Other payables", "Other Reserves",

"Plant and Equipment", "Prepaid Expenses And Other Assets", "Property",

"Provision for employee benefits", "Raw materials", "Recipes",

"Repatriation restrictions", "Reserves and surplus", "Return On Assets",

"Revaluation Reserve", "Secured Borrowings",

"Securities Premium Reserve", "Security Deposits",

"Services and operating rights",

"Share application money pending allotment",

"Share Options Outstanding Account", "Shareholders funds",

"Short-term borrowings", "Short-term loans and advances",

"Short-term provisions", "Stock-in-trade", "Stores and spares",

"Sundry Debtors", "Surplus", "Tangible assets", "Term loans from banks",

"Total Assets", "Total Current Liabilities", "TOTAL",

"Total Liabilities and Shareowners's Equity", "Trade payables",

"Trade receivables", "Trademarks", "Unpaid dividends",

"Unpaid matured debentures and interest accrued thereon",

"Unpaid matured deposits and interest accrued thereon",

"Unsecured Borrowings", "Vehicles", "Work-in-progress",

"Patents and other intellectual property rights", "Formulae", "Models",

"Designs and prototypes", "Drafts on hand",

"Plant And Equipment - Net"

} ,

"scm"

{

{"supplier”

"manufacturingplant

"distributor

{"Name","Product", "Model", "Brand",

"City", "Retailer", "Unit-cost", "Number Of Units Sold",

"Number Of Defect Products", "Distributor Name", "Transport Company"}

"logistics",

"retailer"};

}

};

}

Usecase

Analytics

Variables

/

Parameters

Template Preview

Save Template To File

Preview

# Final lapse work breakdown tasks

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Slno. | Task | Scheduled date | Status | Time required(days) |
|  | Save connections - whatever datafile is connected, its configuration has to be stored in xml. | 18/02 |  | 1/4 |
|  | Save connections - whatever datasource is connected, its configuration is stored in xml | 18/02 |  | 1/4 |
|  | Save Report in MSWORD, PDF, HTML, Image file |  |  | 1/2 |
|  | Save Dashboard in MSWORD, PDF, HTML, Image file | 18/02 | In progress | 1 |
|  | Database disconnect has to be redesigned and implemented. | 18/02 |  | 1/2 |
|  | Datafile disconnect has to be redesigned and implemented. | 18/02 |  | 1/2 |
|  | Data clear has to be tested and debugged. |  |  | 1 |
|  | Data configure has to be implemented. |  |  | 2 |
|  | Complete pivot table |  | Completed | 7 days |
|  | Pivot table – display all the cell value rows - Bug | 24/06-27/06 |  | 2 hours |
|  | Pivot table – two row labels + two column labels bug | 24/06-27/06 |  | 2 hours |
|  | Pivot table – complete Summarise as and show values as | 24/06-27/06 |  | 4 hours |
|  | Pivot table – report filter | 24/06-27/06 |  | 1 day |
|  | Pivot table - redesign to add cat and num variables in one vector. | 24/06-27/06 |  | 1 day |
|  | Draw trendline | 28/06 |  | 1/2 |
|  | Draw 4 control charts. | 28/06-29/06 |  | 1 |
|  | Implement analysis |  |  | 15 |
|  | Implement report |  |  | 10 |
|  | Exchange of tables |  |  | 1/2 |
|  | Implement tables and visuals in one dashboard | 18/02 | In progress | 1 |
|  | Create the user help |  |  | 1 |
|  | Creation of decision tree |  |  | 7 |
|  | Test and debug |  |  | 7 |
|  | Implement for all datatypes |  | This is possible only after all functionalities have been implemented. | 3 |
|  | Add scaling to all charts and tables | 19/02 |  | 0.25 |

Time required to complete : 57 days

11+31 = 42 days

Fit entire scene in visible region.

Append to MSWORD

Add title in pdf

Create HTML files

Design of HTML files

Create an html file that has following skeletal form.

HTML File Title

Dashboard1 | Clicking the appropriate dashboardname, displays the appropriate image file

Dashboard2 | here

... Display the Dashboard name

Dashboardn | Display the image file here