

Repository and Mining of Temporal Data

Milestone 1 Evaluation

Team Members:

Jessica Nguy jnguy2014@my.fit.edu

Siomara Nieves snieves2014@my.fit.edu

Faculty Sponsor and Client:

Philip Chan pkc@fit.edu

Meetings With the Client:

Sept 5; Sept 18; Sept 29

Progress of Current Milestone:

Task	Completion %	Jessica	Siomara	To-Do
Investigate/Select Tools	100%	50%	50%	none
Investigate additional plugins/sites	60%	20%	40%	Plotting software and web-to-app plugins
Create .CSV program, "Hello World" demos	100%	50%	50%	none
Requirements Document	100%	75%	25%	none
Design Document	100%	50%	50%	none
Test Plan	100%	60%	40%	none

Discussion of Each Accomplished Task:

Task 1: Investigated and selected collaboration tools, programming languages, and algorithms to use for analysis. Currently are using a combination of Slack, Gmail, and Google Docs for collaboration and GitHub to host code. We are using Python for our main language and SQL for our database.

Task 2: Plugins were researched; currently looking at using Django to implement a Python program and host it on the web at the same time. Algorithms that we are using for the project is the calculations for z-score, standard deviation, Pearson Correlation, cross correlation,

and linear regression. Problems arose when team members did not understand how to use Pearson Correlation, cross correlation, and linear regression. Plans have been made to remedy the issue within the next milestone.

Task 3: a .CSV reader program was created as a demo to test the file input from the user and check if the file has the correct formatting with the data. The program simply takes the input file and prints it out on the console.

Task 4: Requirements Document was created using Google Docs. Issues that arose during the writing of this document was the need to be specific for certain requirements. Concepts that the project needed were also brought up as requirements, added to the document, and edited as necessary.

Task 5: Design Document was created using Google Docs. Includes the program's high level functional diagram and sketches for the GUI.

Task 6: Test Plan was created using Google Docs. The document uses requirements from the Requirements Document and elaborated on its function and expected outputs to be considered as correct. There were no issues that appeared during writing this document.

Discussion of Contribution for Each Task:

Jessica: Investigated plugins, R programming language, writing requirements/design/test plan documents. Research into suggested algorithms proposed by faculty sponsor. Still working on how to create a web application using existing back-end Java code.

Siomara: .CSV reader code, plug-ins for web, writing requirements/design/test plan documents, programming languages for project, data-mining in SQL analysis. Discovered Django and switched code to Python.

Plan for the Next Milestone:

Task	Jessica	Siomara
1.) CSV User Input catch cases	Write and test catch cases	Write and test catch cases
2.) Q1	Coding math-equations, visualizations	Client input for target variable, export csv file
3.) Website	Host web site somewhere, 3 pages to start	Try to break website
4.) Data Processing	Set Target Variable, Search	Narrow Data, Find data to use
5.) Database setup for Meta-data, Meta-data inputs	Learn SQL, Set database up	SQL, Set database up

Discussion of Each Planned Task:

.CSV User Input catch cases: Need to add a ‘catcher’ code to indicate if users uploaded incorrect formats for csv - csv files need to have timestamps going across the top row with variables in the first column, or timestamps going down in the first column with variables names in the first row with subsequent data in columns going down.

Q1: Client requires that Question 1 be completed by the next Milestone. Target Variable needs to be set. Other variables need to be narrowed down using another algorithm. Question 1 needs to be able to visualize the data and create a csv export of results.

Website: Needs at least 3 pages. A working main page, a page for data providers, and a page for data consumers. Buttons need to be working.

Data Processing: Set Target Variable, Search, and Narrow Data functions need to be programmed.

Database setup for meta-data: database creation using SQL to save the description of the target variable, its tags, initial and final timestamps, and if data is of public or private use.

Sponsor Feedback:

Task 1:

Task 2:

Task 3:

Task 4:

Task 5:

Task 6:

Sponsor Signature: _____ Date: _____

Sponsor Evaluation

- Sponsor: detach and return this page to Dr. Chan (HC 322)
- Score (0-10) for each member: circle a score (or circle two adjacent scores for .25 or write down a real number between 0 and 10)

Jessica	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10
Siomara	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10

Sponsor Signature: _____ Date: _____