**Autumn Group**  
17/07/2017

PROJECT PROPOSAL

Central Data Warehouse

central data warehouse is a collection of all databases from multiple pharmacies and clinics within the autumn group. the aim of this project is to integrate all databases into this central data warehouse. it will help changing completely the ways of analising business data, and preparing a foundation for advanced analytics in the near future



*authors*: VINH PHAN & nHU NGUYEN

*VERSION*: 3.0

CONTENT

[**I.** **Summary** 3](#_Toc487574022)

[**II.** **Introduction** 4](#_Toc487574023)

[**III.** **Needs/Problems** 5](#_Toc487574024)

[**IV.** **Goals/Objectives** 6](#_Toc487574025)

[**V.** **The business values** 7](#_Toc487574026)

[**VI.** **Scope of Work** 8](#_Toc487574027)

[**VII.** **Timetable** 8](#_Toc487574028)

[**VIII.** **Budget** 9](#_Toc487574029)

[**IX.** **Key Personnel** 11](#_Toc487574030)

[**X.** **Evaluation** 11](#_Toc487574031)

[**XI.** **Risks** 12](#_Toc487574032)

# **Summary**

In this document, the following tasks are done:

* The limitations of the current analytic method.
* The proposed model and the pros and cons of this model.
* The schedule and the time estimation for the proposed project. The benefits and potential values of this project will also be discussed.

Based on the above examination, there are 2 proposed options will be shown:

* Basic option
* Advanced option

Basic option aims to fully utilize the company hardware and IT resources to get the tasks done. It costs less and still fulfills the objectives of the project.

Advanced option aims to use cloud computing technologies, and adapt industry standards in database design and data analyzing, however it costs more and requires expensive maintenance costs.

In conclusion, the Basic option is chosen because it costs less to develop and maintain. Any future upgrading to cloud computing if required, will be done with ease.

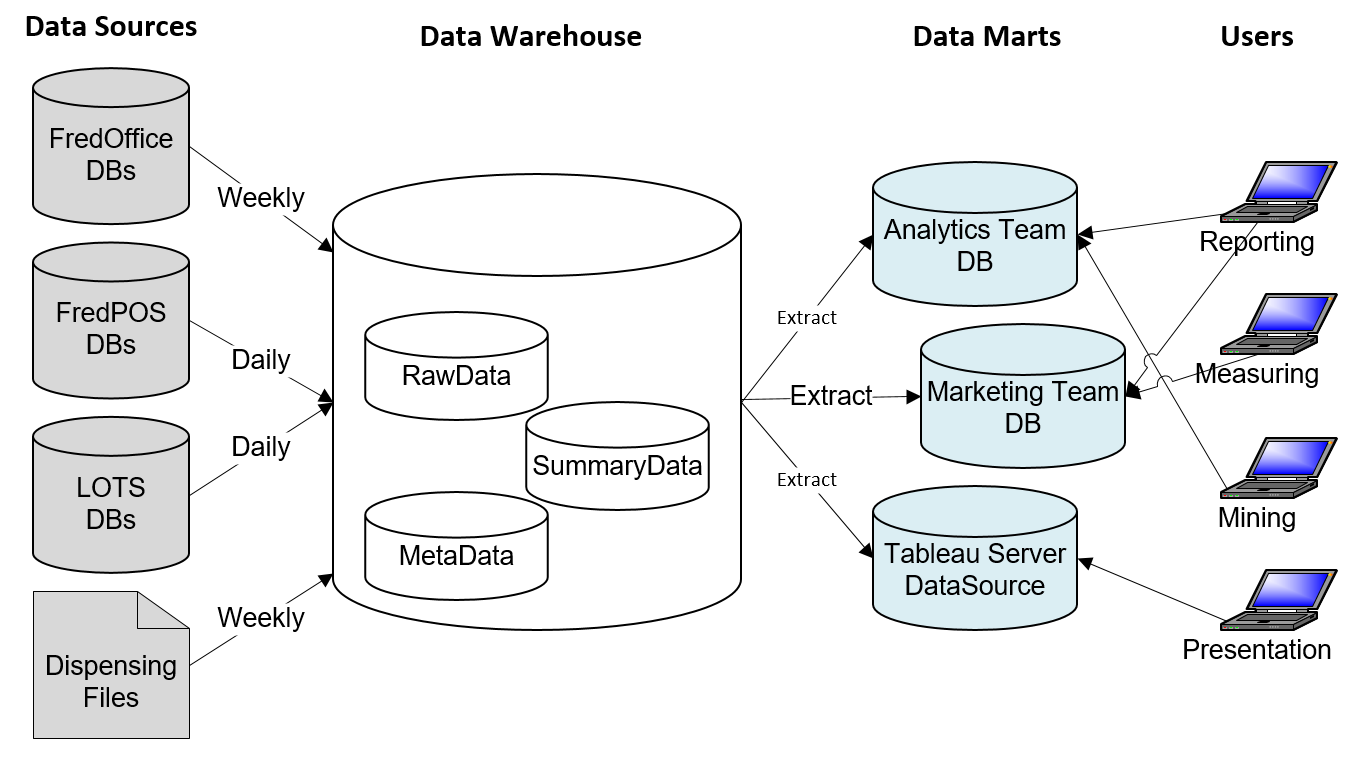
# **Introduction**

*The current situation*

AutumnGroup is a growing company which owns 8 pharmacies and 2 clinics in different locations. Each pharmacy has its own data management system. The managers usually request sales and stock reports every month. To do this, some staffs are needed to collect data from different sources and formats, and then to compile and normalize these data into one agreed format. This process is time consuming and errors are hard to avoid. Furthermore, mangers currently cannot manage stocks in daily activities using these data.

A centralized data warehouse will help to eliminate the above problems and give management extra and powerful tools to analyze business data.

The proposed project will have all pharmacies data stored in one central location. In this way, the managers can have a centralized view of multiple pharmacies and clinics, and see how each site performs compared to others.

*The overview of the project can be visualized as follows*:

* The proposed project includes 4 components that are data sources, data warehouse, data marts, and end users
  + **Data sources** are the databases or flat files from different stores. In particular, the sources come from Fred Dispense, Fred POS, Lots, Fred Office, and Best Practice software provided by vendors.
  + **Data warehouse** keeps historical data from all stores and is queried for business intelligence or other analytic activities. It is typically updated in batches hourly or daily.
  + **Data mart** is a smaller database which only contains data used by a particular department / event

# **Needs/Problems**

|  |  |  |
| --- | --- | --- |
| **#** |  | **Description** |
| *The limitations of the current process* | | |
| 1 | Situation | To analyze data, analytic team has to collect store’s data one-by-one, and uses TeamViewer software to access dispensing computers during working hours |
| Limitations | * Interrupt dispensing process * Slow down the network * Risk of exposing TeamViewer username/pass to outsiders |
| 2 | Situation | The need to analyze accurately and to drill down to the detailed level |
| Limitations | * The data sources both come from Fred and Lots’s reports, they are tailored for only common needs. For example, some detailed Lots’s reports have not been produced ever due to the data extraction limit. The company may have to pay extra to get more detailed reports. * The format of the reports are not consistent. For example, the unique id transactions at Knox differs from Autumn ones. * This is not easy to cleanse data, or to convert it to the same format |
| 3 | Situation | The directors require a combined report which can describe how each store performs  Or a report compares between stores’ sales over the years |
| Limitations | * It may take over a week to combine all backup data for such reports, and there has no way to shorten the time with the current situation. * Cannot drill down to a detailed level of a sale report, the current process cannot handle such a report. |
| 4 | Situation | The sales, profit and inventory reports at each store are required weekly, monthly or even daily. |
|  | Limitations | * Time-consuming to create these reports manually * Cannot have daily reports |
|  | | |

# **Goals/Objectives**

|  |  |
| --- | --- |
| **#** | **Objectives and Description** |
| 1 | *Centralizing data – Data warehouse*   * Store historical data of pharmacies and clinics over the years for analyzing purposes * Address the accuracy problems for different vendor’s software * Have centralized view of multiple stores/clinics * Have valuable data for marketing and sales campaigns * Predict buying trends and customer behavior by analyzing historic data |
| 2 | *Sales, Inventory and Customer databases*   * Address the problems and limitations of the current data analytic method * Gathering data sources for a certain activity, such as Sales, Inventory, and Customer reports * The first step going towards the membership system |
| 3 | *Management & Visualization*   * By using Tableau Server(\*), management team can access reports every time everywhere * Cooperation management between staff in the team * Centralize reports, and schedule regular reports automatically |
| 4 | *Collecting data for advanced analytics later on*   * In the future, these data will be used to support the decision making process by adapting predictive analytics and advanced analysis techniques |

# **The business values**

In this section, the valuable gains for AutumnGroup will be presented. A few scenarios will be used to clarify the huge potentials of this project:

1. The membership system (an application built on top of each POS’s and data will be sent back to the centralized data warehouse):

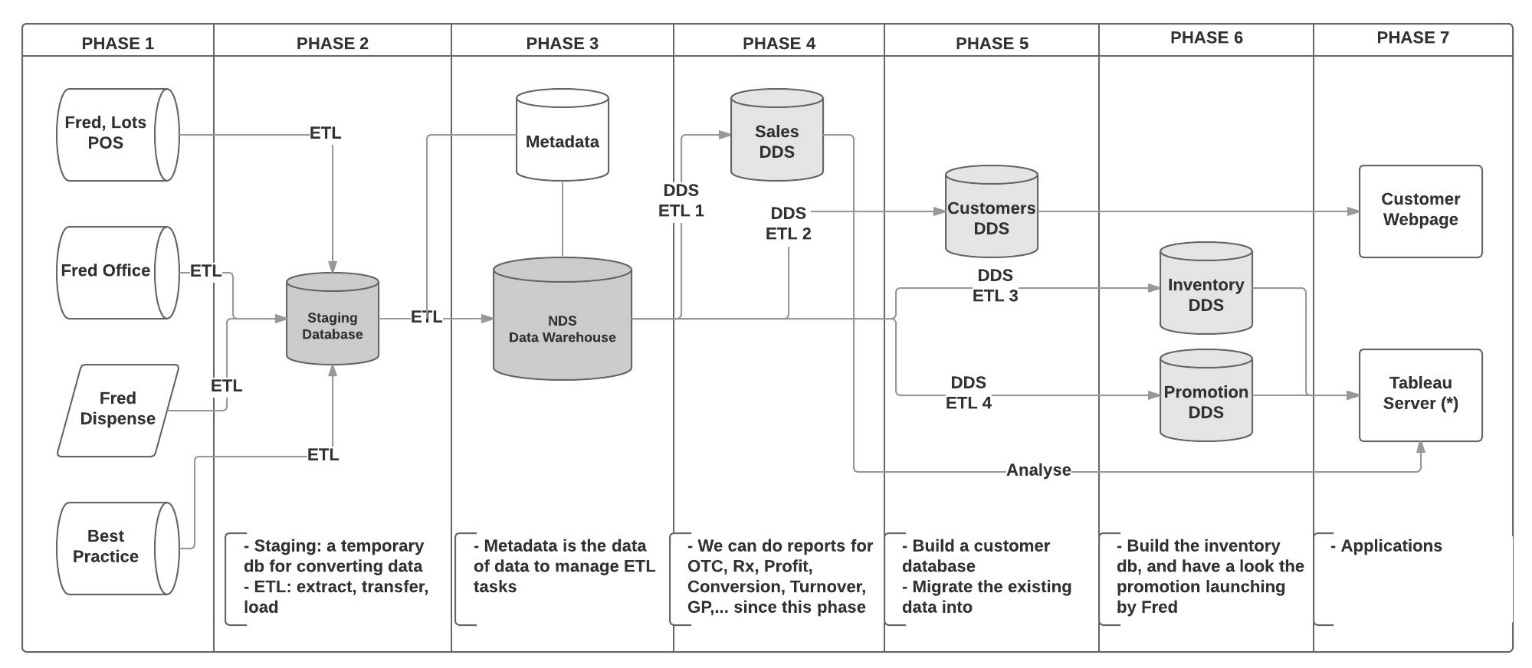
* Customers can be served better by understanding their behaviors
* Promotions and services can be created based on the data from Membership system
* Email marketing, SMS marketing can be used to market company’s products / services

1. The comparison of items’ costs among the stores

* The centralized inventory system can provide a complete view/report of the costs, quantity, stocks …etc… for all stores quickly (**delay only 1 day at most**).
* Provide a useful tool to assist company’s buyers to get the best deal from the suppliers.

1. The centralized database has the flexibility to integrate new store’s data in the future with minimal costs and efforts.
2. This internal, flexible Central Data Warehouse system can help the management making right decisions, serving customers better by understanding their behavior, their needs and their buying patterns.
3. Monthly reports will be created automatically and shared via a link on Tableau among Pharmacist managers quickly.
4. Furthermore, the business intelligence tools can be added in on top of this system whenever there is a need without building it up from the scratch

# **Scope of Work**

****The project can be divided into 7 phases as follows

# **Project Time Frame**

The project can be broken into 7 phases with detailed description of works. Stage 1 will consist of Phase 1 to phase 4, stage 2 will have phase 5 to 7.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Phase | Description of Work | Duration (\*) |
| *Stage 1* | 1 | * Build local staging database prototypes for each store * Build staging database prototype (central location) * Build main database prototype (central location) * Setup static IP for each staging box * Write scripts to transfer data from each store to local staging database and from local to central staging and from central staging to main database * This stage will create a working prototype for this project | 80 hours |
| 2 | * Design staging database for each store * Design main database, including mapping different data sources into a united source for the whole group * Design ETL to cleanse, transfer and load data to databases | 120 hours |
| 3 | * Deploy the staging boxes in each stores (8 pharmacies and 2 clinics) * Deploy staging and main databases | 120 hours |
| 4 | * Extract historic data from all stores into Central Data Warehouse * Run tests and measure system performance * Start the system and monitor the process * Develop user manual | 120 hours |
| *Stage 2* | 5 | * Design DDS Customer * Migrate the existing data into it | 40 hours |
| 6 | * Design DDS Inventory * Do research Fred promotions / events | 40 hours |
| 7 | * Do research for synchronization * Launch a website for registration * Review the performance of the project * Discuss the Tableau server | 40 hours |
|  | **Total** | | 560 hours (14 weeks for one staff) |

During the time frame of this project, weekly catch up will be arranged among project manager, sponsors, consultants, directors.

# **Budget**

The budget for this project consists of the initial capital cost and the ongoing maintenance cost

* The initial capital cost includes the hardware and licenses costs, and the labor costs

|  |  |  |
| --- | --- | --- |
| **#** | Description of Work | Costs |
| 1 | Computer with i7, Memory 16gb, and HDD 2TB running Windows 10 64x *or* Window Server 2012 Essentials Edition | 1500 – 2000AUD |
| 2 | Window SQL Server 2012 Standard (permanent license) | 1000AUD |
| 3 | Tableau Server (1 publisher, 5 users) / year (\*) | Publisher: $70 \* 12  Users: $35 \* 5 \* 12  Total: $2940 / year |
| 4 | Static IP cost | $10/month \* 10 =  100AUD/month |
| 5 | Computers at each store (using as staging servers) **can use existing hardware if it is available** | 800 \* 10  Total: 8000AUD |
| 5 | Labor cost | 560 hours (= 14 weeks/ one working staff) |
|  | Total |  |

At this moment, it is not possible to break down the cost for each pharmacies because different pharmacies have different database systems and they require different efforts to extract, analyse and store into the central database.

* The ongoing maintenance costs

|  |  |  |
| --- | --- | --- |
| # | Description of Work | Costs |
| 1 | Tableau Server Renewal (\*) | $2940 / year |
| 2 | Maintenance Cost | 10 hours/week |
|  | Total |  |

# **Key Personnel**

List the key staff who will be responsible for completion of the project, as well as other personnel involved in the project.

Project Sponsors: Alex Fung & Hailey

Consultant: Hailey

Hardware Manager: Huu Nguyen

Project Manager: Nhu Nguyen

Project Member: Vinh Phan

The project team will have weekly meeting with the sponsors to report the progress of the project, to discuss about any potential problems…etc…

# **Evaluation**

The problem can be evaluated throughout the process. Based on the “scope of work”, this project can be evaluated at then end of Phase 2, Phase 4, Phase 5, and Phase 7

|  |  |
| --- | --- |
| **Phase** | **Evaluation** |
| 2 | To evaluate the risk and possibility of the project  How big the project can be up to |
| 4 | To evaluate the results between traditional and new approach. How different are they? Is there a way to automatic |
| 5 | To evaluate the value of existing customer data  How possible the approach will be, or do we need another solution |
| 7 | A general evaluation |

# **Risks**

This proposed project relies heavily on the infrastructure bases of the pharmacies and the internet connection from each store to the central location. Therefore it can raise the following risks:

* Slow internet connection: data will not be sent to Central as expected (default expectation: only one day delayed)
* Some pharmacies are located interstates therefore the process of installation the hardware and software can be long and it may delay the progress of the project.
* Some extra software such as paid SQL server (standard version) may be needed for each store if the proposed technical method does not work as expected, it leads to budget blowout or extra research time to find a better technical solution. This is a low risk scenario anyway.
* IT staff consists only Nhu Nguyen and Vinh Phan, if one of these staff are sick or busy to do some other higher priority tasks, the dateline will not be met.
* Part of this project requires hardware installation and network setup, this is not calculated in the project time line and may cause a delay as well.
* Most of the current reports will need to be modified and it may require extra time from Hailey and IT staff (Unforeseen time delay)