UNIVERSITY OF DAYTON

Masters in Computer Science Department

Course: Software Project Management (CPS 622 Z1)

PROJECT TITLE: MODERN WAY TO STORE THE DATA

MIDTERM PROJECT REPORT

By

DINESH KUMAR KALA

Student ID: 1017453540

TABLE OF CONTENT: _____ Page No:

1.INTRODUCTION

1.1 DATA STORAGES

Data Storage places a crucial role in Every firms. Now-a-days Every firm is planning to Migrate to Cloud Storages mainly because of their Scalability, Security and Backup of files. Especially big MNC's Like Banks are facing the Critical task to Store Customer Data because of increase of Customers. A Cloud Based Customer Data Platform (CDP) would be the Best Solution to Store Data because less cost (Pay-as-go), Cost increases based on the usage of Storage and best flexibility and Performance.

1.2 PRESENT DATA STORAGE TECHNIQUES:

Big Banks like CHASE Bank, Key Bank, Wells Fargo Provides consumers and businesses with a wide range of financial products and Services. Chase Bank and Key Bank have been attracting the customers because of their best Services and features like, Products and Services, Marketing, Branch Network, online and mobile banking, easy to use Applications and Securities and more branches. Both banks offer similar type of products and services such as Checking Account, Saving accounts, investments, loans, Credit Cards and Purchase Discounts. They are using Different Marketing Techniques to attract the customers and to retain present Customers. When Compare to Branch network, Chase bank is bigger than the KeyBank, which is easier for customers to do the business with them. Both Banks are offering Online and mobile banking to their customers to access their accounts and for transactions from anyplace.

. These Banks are using Different types of Data Storage Techniques like Centralized Database, Distributed Database and Cloud Storage. Both Banks are using this Databases Based on the Data and the information to be Stored.

Chase Bank's Centralized Database is called as Customer Information system (CIS). Where they use to Store information of Customers Account, Transaction history, and Contact information. CIS is protected by a different type of security. Distributed database is used to store the less sensitive Customer data, such as Customer reviews and Marketing Information. In order to increase the performance and Security, these databases are Distributed between multiple servers. Coming to Cloud Storages, Chase Bank Uses Amazon Web Services (AWS) to store the Not sensitive data. AWS is a Secure and safe platform for cloud computing that provides a variety of storage services.

Key Bank also uses the three databases, Key Bank's Centralized database is known as Customer Data Repository (CDR). CDR stores Sensitive Data like Customers information, Accounts, Transaction History. It is located in a secure data center and secured by different types of securities. Same as the Chase bank, Key Bank also uses the Distributed Database to store fewer sensitive data like Marketing Information and customers feedbacks. It is also stored in Multiple Servers for better performance and Security. In Cloud Storages Key Bank Uses Microsoft Azure to store Non-Sensitive Data. Azure is also another cloud provider such as AWS with similar features and services.

Every MNC's Like Chase Bank, KeyBank are planning to migrate all the customer data, Accounts to the Cloud Storages because it is cost-effective way to store data and it is easy to add and remove storage capacity as needed. This is vital for Banks that experience up and downs in data volumes. Security level is high compared to other Storages and New features and services can easily added.

1.3 CLOUD SERVICE PROVIDERS:

There are many Cloud Service Providers like Amazon Web Services (AWS), Microsoft AZURE, Google Cloud, IBM Cloud, Oracle Cloud, Alibaba Cloud, Salesforce Cloud, VMware Cloud, Cisco Cloud. Etc.

1.4 MAJOR COMPETITORS:

Major Cloud Providers are Amazon web services (AWS), Microsoft AZURE, Google Cloud

AMAZON WEB SERVICES TECHNIQUES (AWS):

There are many ways to store the data in AWS, Amazon S3 (Simple, Storage, Services) is a highly Scalable object Storage Services. It is used to store the large amount of Unstructured data, images, videos, files. Amazon EBS (Elastic book Store) it is a traditional technique to store the Vm's Disks. Amazon EFS (Elastic File System) is used to Store the shared files and Documents. Amazon DynamoDB it is a No SQL Database used for storing large amount of Structured Data. Amazon Glacier it is used for rarely accessed data like history. **AMAZON CUSTOMER 360** is the one of the best solutions for banks to migrate all Accounts because it offers wide range to features and Functionalities including Data Integration, Data Collection, Data Cleansing, Data Analysis and Data Activation. It collects data from websites, databases, and it integrates all the sources to the single view and removes Duplicates, Correcting Errors. One of the Best feature is It Analysis the data Using Machine Learning and Artificial Intelligence (AI).

MICROSOFT AZURE:

Azure Providers Many options to store the data, Azure Blob Storage have the best Scalability object Storage Services and it stores large amount of unstructured data, images, videos. AZURE Files uses Server Message Block (SMB) protocol to share files, it can access simultaneously from different VM's. Azure Disk Storage provides Durable Block Storages for Azure Virtual machine. **AZURE PURVIEW** is the best choice to store all bank data because, compliance, Security, Integration, Scalability, Cost-Effectiveness. It is certified as a variety of regulations like HIPAA. It is Highly Scalable and fully Encrypted. One of the best reason is it is integrate with wide software like, salesforces, which make easy to connect with existing System and it is Cost-Effectiveness.

GOOGLE CLOUD:

Same as Azure and AWS, GCP have a various technique in storing data. Cloud Storage has the High Scalability and used for storing large amount of Unstructured data, images, Videos. Persistent Disk is used to Store the Virtual machine Disks. **GCP Customer Match** would be the good option to choose because Advanced Customer Engagement, Targeted Advertising, Data Privacy and Security, Cross Channel Marketing. Data Privacy and Security is the high priority and it is easy to communicate with customers and advertising. It provides the analytical report based on the data.

1.5 CONCLUSION:

Based on the Budget, Scope and Quality, Azure Purview is the best Solution Because of Cost-Effectiveness and Security and Easy to Compatible and It is easy to Connect with Other Software's, Other than Azure Purview, Amazon Customer 360, GCP Customer Match are the best Solutions.

2. Present Data Storage Technique VS Proposal Data Storage technique

2.1 Present Data Storage Technique:

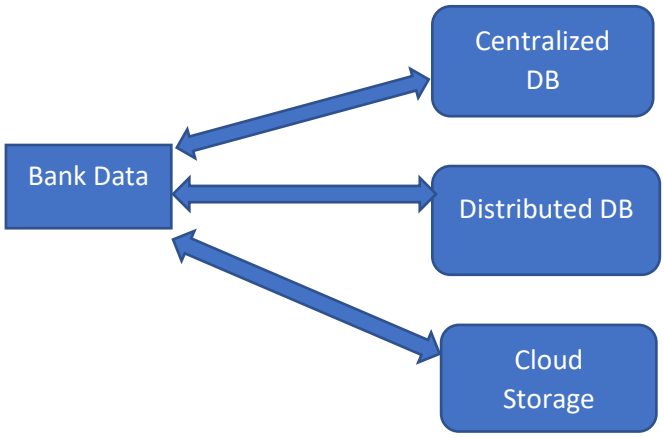


Fig 2.1.1 This figure shows the Present data Storage technique of the bank

Present data Technique of the bank is, it uses the 3 types of Data Storages based on the Confidential Information and Non-Confidential Information.

2.2 Proposal Technique:

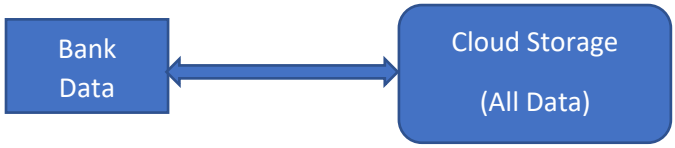
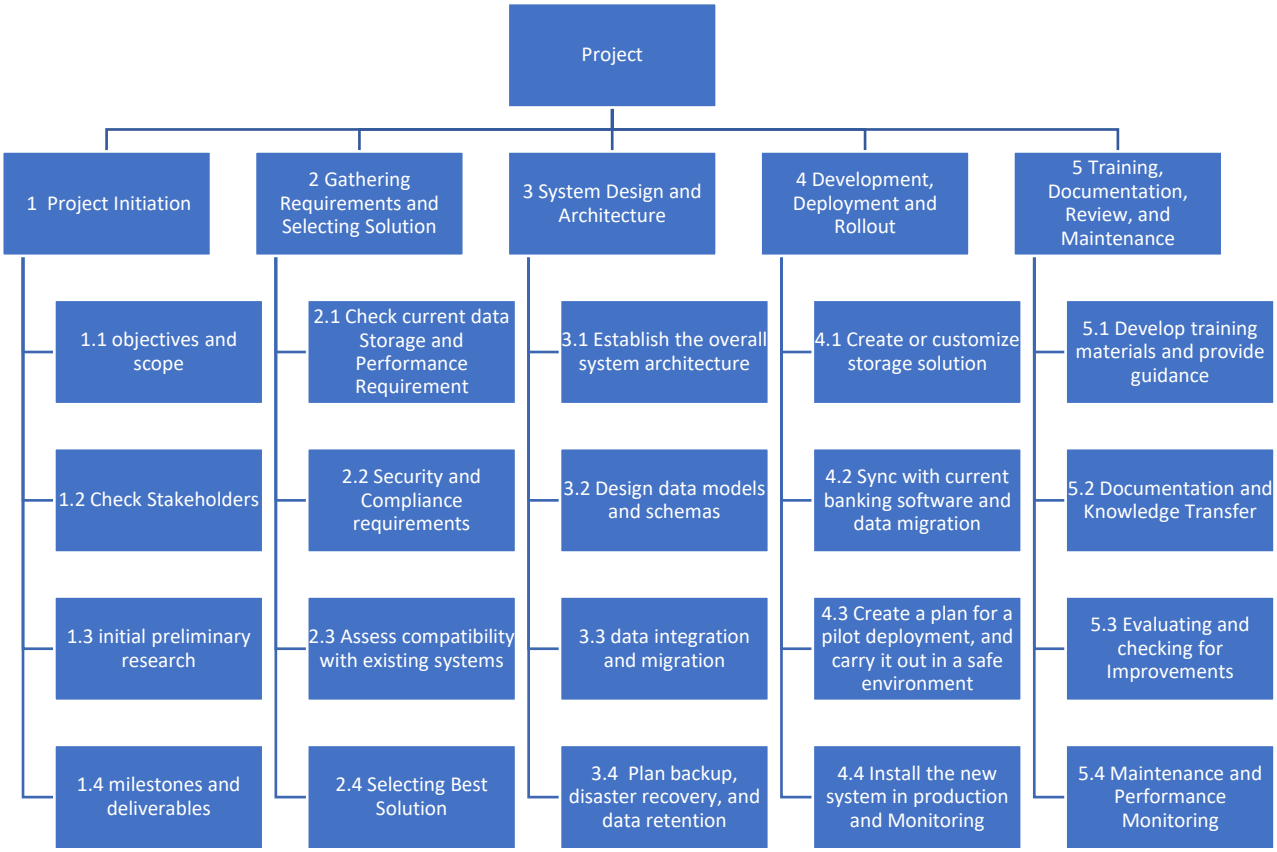


Fig 2.2.1 This figure Shows the proposal Data Storage Technique

In the Proposal Technique bank is Using the Cloud Storage Technique, where all data stored at one place. Cloud is more Preferable because, it provides advanced security and latest techniques and it is Scalable.

3. WORK BREAKDOWN STRUCTURE



3.1BRIEF DESCRIPTION:

To Do this Project, WBS is the Best Approach. This Project is Divided into 5 Tasks and 4 Sub-Tasks on each. Project starts with

1.Project Initiation

This has 4 Subtasks Objectives and Scope, Check Stakeholders, Initial Preliminary research, Milestones and Deliverables. Objectives and Scopes are important for this project. Search for the Stakeholders and doing Initial research and schedule the timeline. Make deliverables and Milestones.

2. Gathering Requirements and Selecting Solutions:

This Task have total of 4 Sub-Tasks, Gathering the All Requirements and Check with the Current Data Storage if the Bank and Perform the requirements, Check with the Security and Compliance Requirements and Assess the Compatibility with the Existing Systems and Choose the best Solution.

3. System Design & Architecture:

Establish the All-System Architecture and design the Data Models and Data Schemas for the Selected Database. Do the Data Integration and Data Migration and Plan the Backup and Disaster Recovery and Data Retention Process. This Tasks and Sub-Tasks are very important for the Project.

4. Development, Deployment Rollout:

This task is important for the Project, where we create a new Storage Solution in Cloud and Merge with the Exists Database Storage and sync with the current banking Software and do the Data Migration. After Creating Data Storage Successfully, Create a plan for a Pilot Deployment and Carry-out in a Safe Environment. Run the All test-Cases and install the new System in Production and Do Monitoring.

5. Training, Documentation, Review and Maintenance:

After the Development and Deployment, The Final Step Are Develop Training materials and Provide Guidance. Do the Documentation and Do the Knowledge Transfer (KT Session's) to the client. Do the Evaluating and Check for Improvements and do Maintenance and Performance Monitoring.

4. SCOPE OF TRIANGLE:

	High	Medium	Low
Scope		✓	
Quality	✓		
Time		✓	
Cost			✓
Resource Availability	✓		

Reasons & Justifications:

Scope: Large Bank already have an Existing Database and migrating to Cloud Storage Which required Numerous Processes and More Integration. So, Scope is Rated as a MEDUIM.

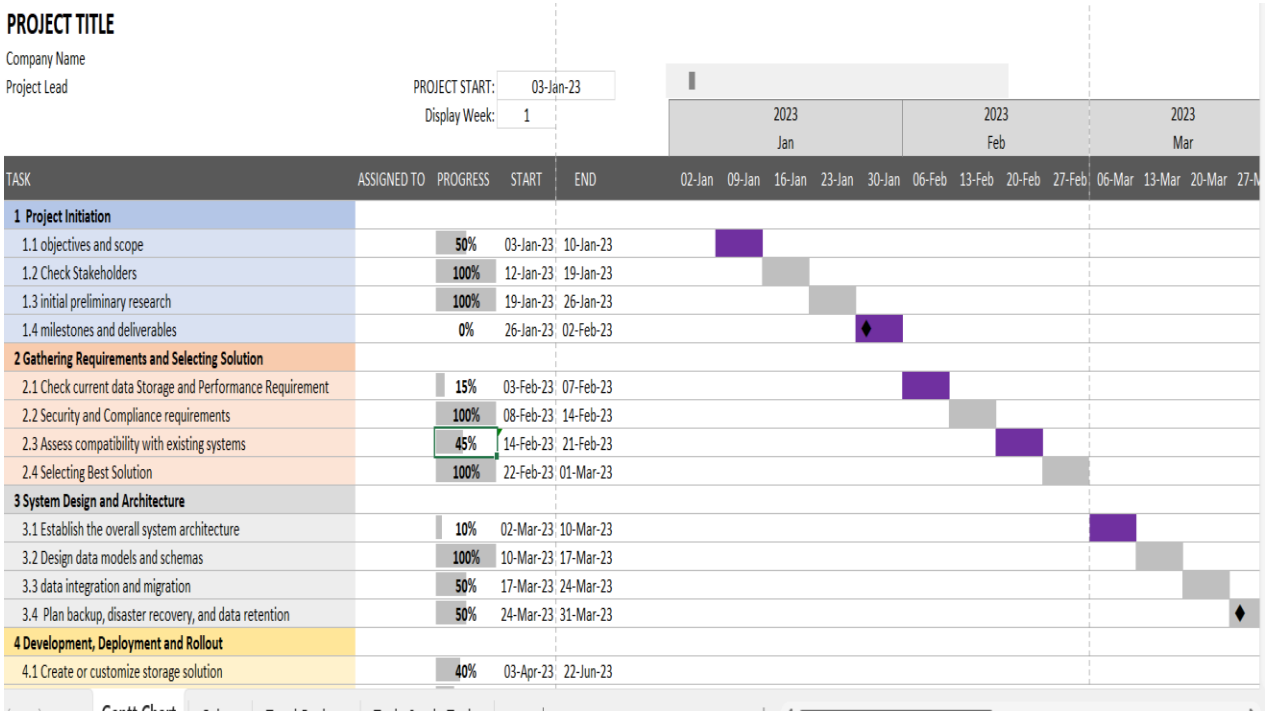
Quality: Quality is Rated as a HIGH Priority because the Big MNC's are mainly Concentrated on Quality of the Project Requirements and they look for Security and Confidentiality.

Time: As they are Migrating all the Accounts to the Cloud Storages, it is important to Complete the Project with in the Given Time Period. So, it is Rated as a MEDIUM.

Cost: Cost is rated as a LOW Priority because large banks Mainly Focus on Scalability, Security and Confidentiality of the Accounts.

Resources Availability: To Design this Project we required a high Quality Team Members and SME because It is a Large Application Where Customers Use this frequently and it monitored Carefully.

5. Gantt Chart



PROJECT TITLE

Company Name

Project Lead

PROJECT START: 03-Jan-23

Display Week: 1

2023 Jan | 2023 Feb | 2023 Mar

02-Jan 09-Jan 16-Jan 23-Jan 30-Jan 06-Feb 13-Feb 20-Feb 27-Feb 06-Mar 13-Mar 20-Mar 27-M

TASK	ASSIGNED TO	PROGRESS	START	END
1 Project Initiation				
1.1 objectives and scope		50%	03-Jan-23	10-Jan-23
1.2 Check Stakeholders		100%	12-Jan-23	19-Jan-23
1.3 initial preliminary research		100%	19-Jan-23	26-Jan-23
1.4 milestones and deliverables		0%	26-Jan-23	02-Feb-23
2 Gathering Requirements and Selecting Solution				
2.1 Check current data Storage and Performance Requirement		15%	03-Feb-23	07-Feb-23
2.2 Security and Compliance requirements		100%	08-Feb-23	14-Feb-23
2.3 Assess compatibility with existing systems		45%	14-Feb-23	21-Feb-23
2.4 Selecting Best Solution		100%	22-Feb-23	01-Mar-23
3 System Design and Architecture				
3.1 Establish the overall system architecture		10%	02-Mar-23	10-Mar-23
3.2 Design data models and schemas		100%	10-Mar-23	17-Mar-23
3.3 data integration and migration		50%	17-Mar-23	24-Mar-23
3.4 Plan backup, disaster recovery, and data retention		50%	24-Mar-23	31-Mar-23
4 Development, Deployment and Rollout				
4.1 Create or customize storage solution		40%	03-Apr-23	22-Jun-23

This the Image of a Gantt Chart, Showing the Tasks and Sub-tasks and along with their start and End Dates and Progress of the work. This Project Is Divided into 5 Tasks With 4 Sub-Tasks on Each and Every Task have a Milestones. This Project is taking the 10 months of time period to Complete the Project. This Chart has a **Display week** cell, we can directly search the any week progress and it is also included with the **Scroll Bar**.

5.1 LEGENDS USED IN GANTT CHARTS:

- Work to be Done

- Completed Work

- This Diamond is Represented as a Milestone

6. All tasks and Major Tasks with Milestones:

ID	TASK NAME	DUE DATE mm-dd-yyyy	MILESTONES	MILESTONE DUE DATE
1	Project Initiation		Approved Project Charter	02-02-2023
1.1	objectives and scope	01-10-2023		
1.2	Check Stakeholders	01-19-2023		
1.3	initial preliminary research	01-26-2023		
1.4	milestones and deliverables	02-02-2023		
2	Gathering Requirements and Selecting Solution	03-01-2023		
2.1	Check current data Storage and Performance Requirement	02-07-2023	Selecting Best Solution	03-01-2023
2.2	Security and Compliance requirements	02-14-2023		
2.3	Assess compatibility with existing systems	02-21-2023		
2.4	Selecting Best Solution	03-01-2023		
3	System Design and Architecture	03-31-2023		
3.1	Establish the overall system architecture	03-10-2023	completion of architecture and design	03-31-2023
3.2	Design data models and schemas	03-17-2023		
3.3	data integration and migration	03-24-2023		
3.4	Plan backup, disaster recovery, and data retention	03-31-2023		
4	Development, Deployment and Rollout	07-27-2023		
4.1	Create or customize storage solution	06-22-2023	Completion of Deployment	07-27-2023
4.2	Sync with current banking software and data migration techniques	07-13-2023		
4.3	Create a plan for a pilot deployment, and carry it out in a safe environment	07-21-2023		
4.4	Install the new system in production and Monitoring	07-27-2023		
5	Training, Documentation, Review, and Maintenance	10-31-2023	Project Completion	10-31-2023
5.1	Develop training materials and provide guidance	08-21-2023		
5.2	Documentation and Knowledge Transfer	08-28-2023		
5.3	Evaluating and checking for Improvements	10-05-2023		
5.4	Maintenance and Performance Monitoring	10-31-2023		

Fig 6.1: Figure Shows All tasks and Major sub-tasks required to do the Project

At the End of Every task, there is a Milestone with a Suggested due dates. This all are major Tasks and Sub-tasks we are going to use to build the Project. Time taking to complete this project is of 10 Months and Every Dates assigned are Reasonable.

7. Team members and Salaries:

Role	Total Hours	Hourly Rate	Total Cost
Project Manager	$250	$75	18,750
Senior Software Engineer	$500	$90	45,000
Junior Developer 1	$1,000	$40	40,000
Junior Developer 2	$1,000	$40	40,000
Junior Developer 3	$500	$40	20,000
Subject-Matter Expert	$100	$200	20,000
Total	**3350**		**1,83,750**

Fig 7.1: Figure Shows the Team members and their Salaries

To do this Project we required 1 Project Manager, One Senior Software Engineer, 3 Junior Developers and mainly Subject-Matter Expert. The Subject-Matter Expert (SME) is a Professional, suggests a best technology and Software tools required for the Project. Project Manager who handle multiple projects take care of the work progress and Senior Software Engineer and 3 Junior Developers Used to develop and test the Project.

8. TOTAL MONTLY BUDGET:

Project Budget				
Months	Cumulative cos	Planned	Actual	Total Budge
January	$15,000	$15,000	$15,000	$4,90,000
February	$55,000	$40,000	$30,000	$4,90,000
March	$95,000	$40,000	$35,000	$4,90,000
April	$1,45,000	$50,000	$45,000	$4,90,000
May	$1,95,000	$50,000	$45,000	$4,90,000
June	$2,55,000	$60,000	$60,000	$4,90,000
July	$3,15,000	$60,000	$60,000	$4,90,000
August	$3,75,000	$60,000	$50,000	$4,90,000
September	$4,25,000	$50,000	$30,000	$4,90,000
October	$4,50,000	$25,000	$20,000	$4,90,000
November	$4,60,000	$10,000	$10,000	$4,90,000
December	$4,70,000	$10,000	$10,000	$4,90,000
total		470000	4,10,000	

Fig 8.1: Figure Shows Total budget and Monthly Budgets

The Figure Shows the Total budget Required for one year, with Planned Budget and Actual Budget and Cumulative Cost is the Addition of Every month of Planned Cost. Which is Equal to the Planned Budget. The Project Completes in the 10th Month (OCTOBER). Remaining two months are for Monitoring the Project.

Graph:

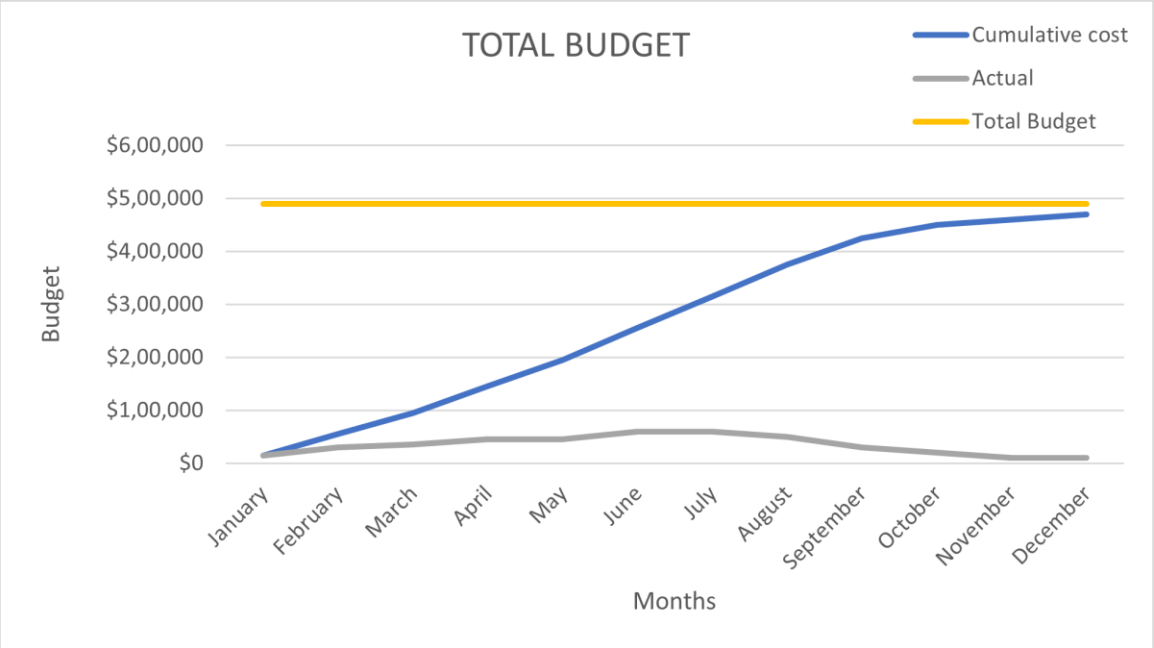


Fig 8.2: This Figure shows Graphical representation of Total Budget and Monthly Budget

This the graph for One year plan with Months as a Non-Dependable variable and Cumulative cost is the Dependent variable and the total Budget for Every Month Which is Straight line.