[Introduction (Teamwork) 3](#_Toc379653748)

[DAP Design and Implementation (Giridhar K SrinivasaRaghavan) 4](#_Toc1887761586)

[Step 1: Data Analytical Question Formulation 4](#_Toc1635678385)

[Step 2: Data Discovery 4](#_Toc1615259722)

[Step 3: Data Storage Design 4](#_Toc1700180782)

[Step 4: Dataset Preparation 4](#_Toc235666149)

[Steo 5: Data Ingestion 4](#_Toc1529270637)

[Step 6: Data Storage 4](#_Toc848114813)

[Step 7: Data Pipeline Design 4](#_Toc1777204349)

[Step 8: Data Cleaning 4](#_Toc715190994)

[Step 9: Data Structuring 4](#_Toc366715064)

[Step 10: Data Pipeline Implementation 4](#_Toc1211431193)

[Step 11: Data Analysis 4](#_Toc1148269958)

[Step 12: Data Visualization 4](#_Toc322022985)

[Step 13: Data Publishing 4](#_Toc627709591)

[DAP Estimated Cost (Teamwork) 4](#_Toc1750497960)

[Assignment Rubrics 5](#_Toc1981705855)

# **Introduction (Teamwork):**

# **DAP Design and Implementation for 3-1-1 Inquiry Volume – Giridhar K S**

## **Step 1: Data Analytical Question Formulation**

The report's agenda is that The City of Vancouver needs to migrate to AWS. The Below screenshot displays the design flow and various datasets required for the construction and design of the Data Analytics platform for the City of Vancouver to migrate to the Aws cloud platform.

A screenshot of a computer

Description automatically generated

***Figure 1:*** *DAP Design*

The screenshot below displays the essential AWS cloud services required to support and enhance the City of Vancouver during the migration to the AWS cloud platform.

S3: storage service, EC2: Computing service, Athena: Analytical service, AWS Glue: Data Integration Service are a few of the majorly used services during the design of the Data Analytics Platform.

A screenshot of a computer

Description automatically generated

***Figure 2:*** *AWS services used for the design of the AWS cloud platform for the City of Vancouver*

## 

Step 2: Data Discovery

The design implementation requires raw data for the creation of the dataset. We have used the open data website (<https://opendata.vancouver.ca/>) for the discovery and extraction of raw data during this process. The “3-1-1 Inquiry Volume” database is used in this case for the Individual DAP design*.*

A screenshot of a computer

Description automatically generated

***Figure 3:*** *Raw Data Extraction*

A graph on a computer screen

Description automatically generated

*Figure 4: Yearly Volume Flow of Inquiry*

Figure 4 showcases the sum of the total volume flow of inquiries made through various channels during a certain time. The data concentration is targeted only for the years 2023 and 2024 for this particular design implementation.

## **Step 3: Data Storage Design**

A screenshot of a computer

Description automatically generated

*Figure 5: AWS S3 Bucket*

The above screenshot displays the use of AWS cloud service and the creation of the S3 storage bucket for data storage.

## **Step 4: Dataset Preparation:**

After creating the S3 bucket, the next process is the dataset preparation. The screenshot below displays the creation of various folders and primary landing zones where the extracted raw data will be stored.

A screenshot of a computer

Description automatically generated

*Figure 6: Dataset Preparation*

## **Step 5: Data Ingestion**

A screenshot of a computer

Description automatically generated

*Figure 7: Dataset Ingestion*

The above screenshot displays the process of uploading the extracted raw data into the creation landing zones. For this purpose, we navigate through the newly created domain and landing folder under the S3 bucket and add data using the data ingestion process on the AWS cloud service platform.

## **Step 6: Data Storage**

A screenshot of a computer

Description automatically generated

Figure 8: AWS Glue service

To use the raw data in a structured format for analytical purposes, we have used the AWS glue service to create tables and databases.

## **Step 7: Data Pipeline Design**

A diagram of a data flow

Description automatically generated with medium confidence

*Figure 9: Pipeline Design structure*

During the design, we used the AWS Data Pipeline to organize, schedule, and manage regular data processing tasks efficiently and affordably. The service simplified the design of data extract-transform-load (ETL) processes for both structured and unstructured data, whether it's stored on-premises or in the cloud, allowing to tailor these activities to specific business needs.

## **Step 8: Data Cleaning**

A screenshot of a computer

Description automatically generated

## **Step 9: Data Structuring**

A screenshot of a computer

Description automatically generated

In this part of data structuring, we can use the SQL query writer in the AWS services to create properly structured data for analysis.

## **Step 10: Data Pipeline Implementation**

A screenshot of a computer

Description automatically generated

The AWS Data Pipeline is used during the project's design to define data-driven workflows so that tasks can be dependent on the successful completion of previous tasks and a clear process flow is followed.A screenshot of a computer

Description automatically generated

## **Step 11: Data Analysis**

A screenshot of a computer

Description automatically generated

## 

## **Step 12: Data Visualization**

A screenshot of a graph

Description automatically generated

The above screenshot explains the data analysis done under the project clearly using the structured data and completing the job. We can understand the total volume of inquiries made, the channels used, and the time in which these transactions happened.

## Step 13: Data Publishing

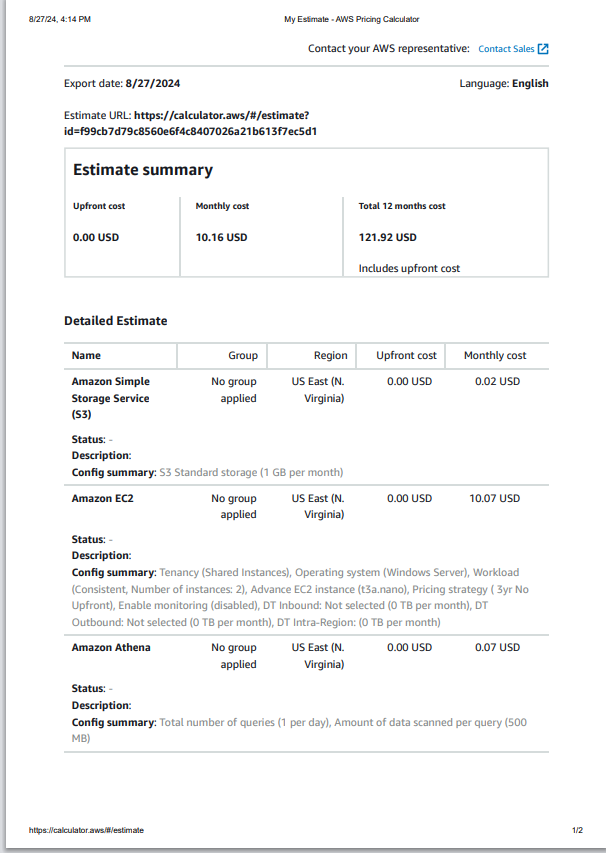
A screenshot of a computer

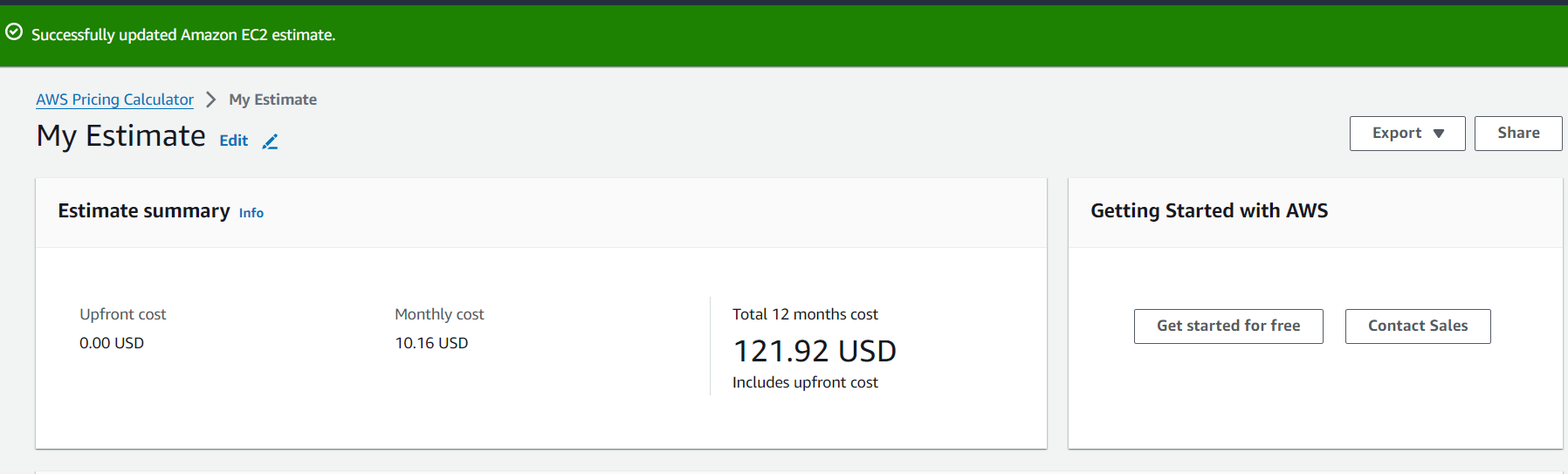
Description automatically generated

The data publishing step clearly showcases the General server and web server instances created to publish the extruded and structed data on the web server and file share on the general server.

# DAP Estimated Cost (Teamwork)

*Giridhar: AWS portal pricing screenshot*





References:

AWS Pricing Calculator. (n.d.). <https://calculator.aws/#/estimate>

Home - city of vancouver open data portal. (n.d.). <https://opendata.vancouver.ca/pages/home/>