

CLINICAL REVIEW

HbA1c Range Determination for Diabetes Treatment in Google Cloud Compute Engine and SerpAPI

Ramamurthy Valavandan*, Hanumantha R, Kumaraswamy Reddy, Prasanth Parayatham, Vikram Sharma, Surasa Mukherjee, Vijay Yadav, Pragathi S, Vijay R, Nitin Ambekar, Dinesh Sai Teja Neeli, Santosh Baran, Vijender Singh, Praveen B and Musheer Ahmed N

Kyndryl Solutions Private Limited, GCP Guild Project, Karnataka, India

Correspondence should be addressed to Ramamurthy Valavandan, Kyndryl Solutions Private Limited, GCP Guild Project, 6th Floor, D-1, Manyata Tech Park, Bangalore 560024, Karnataka, India

Received: 28 June 2022; Accepted: 17 July 2022; Published: 24 July 2022

Copyright © Ramamurthy Valavandan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

In this paper, we have taken a use case in the healthcare domain for the determination of factors for an accurate range of HbA1c in diabetics' treatment. Google search engine application program interface (API), SerpAPI@ is evaluated for the execution of compute engine in google cloud. As a service in Infrastructure as a Service (IaaS), the compute engine gives the accurate value in HbA1c in the work orchestration in airflow (astronomer) of Apache and cloud composer. We have also developed the application layer for every work stream in python and application output is stored in google cloud storage. The accuracy of diabetic range determination is calculated in python as a measure for the healthcare practitioner to provide the right line of treatment for diabetic patients. A python program is developed to connect to the google search with a wrapper of SerpAPI in the google cloud scheduler and create the JSON file schema for parameterization of search engine API. The results are shared with customers in the healthcare domain. We have also addressed the challenges of diabetic treatment during long-term COVID-19.

KEYWORDS

Google cloud; Diabetics; HbA1c; COVID-19, Application modernization; Infrastructure as a service; Compute engine; Google search engine; Python; Cloud storage; Cloud composer

INTRODUCTION

Diabetes is a global endemic with rapidly increasing prevalence in both developing and developed countries. The American Diabetes Association has recommended glycated hemoglobin (HbA1c) as a possible substitute for

fasting blood glucose for the diagnosis of diabetes. HbA1c is an important indicator of long-term glycemic control with the ability to reflect the cumulative glycemic history of the preceding two to three months.

HbA1c not only provides a reliable measure of chronic hyperglycemia but also correlates well with the risk of long-term diabetes complications.

Elevated HbA1c has also been regarded as an independent risk factor for coronary heart disease and stroke in subjects with or without diabetes.

The valuable information provided by a single HbA1c test has rendered it a reliable biomarker for the diagnosis and prognosis of diabetes.

HbA1c is sugar bound to hemoglobin in red blood cells. It can help diagnose diabetes. HbA1c is an indicator of glucose levels in long term.

HbA1C is a lab report of blood that is used in the diagnosis of diabetes procedure and line of treatment. In a simple the lower the HbA1c within the normal range is the better for the diabetics.

As the HbA1c number raises, the higher the risk of diabetes and related health issues. Determination of HbA1c is very essential for medical terms [1] and getting the accuracy in HbA1c is challenging considering the factors of test report sampling, long-term, and nonavailability of applications for factoring HbA1c.

Given the gravity of health contribution, we have created the lab environment in Google Cloud with application deployment of SerpAPI for Google Search Engine connect, parsing of JSON, and finally arriving at the accurate range of HbA1c.

For workload for the computation, we've used the Google Cloud Engine service for the determination of the long-term parameter of HbA1c in search Application Program Interface (SerpAPI) and Compute Engine for calculations. We also followed the healthcare standards of the World Health Organization on diabetics' treatments as an objective.

As per the recommendation of WHO, HbA1c can be used as a diagnostic test for diabetes providing that stringent quality assurance tests are in place and assays are standardized to criteria aligned to the international reference values, and there are no conditions present which preclude its accurate measurement.

An HbA1c of 6.5% is recommended as the cut point for diagnosing diabetes. A value of less than 6.5% does not exclude diabetes diagnosed using glucose tests.

Quality of evidence assessed by GRADE: Moderate.

Strength of recommendation based on GRADE criteria: Conditional.

DIABETES - A SILENT KILLER

More than 37 million children and adults have diabetes in the United States. The cost of diabetes.

The national cost of diabetes in the U.S. in 2017 was more than \$327 billion, up from \$245 billion in 2012 (Figure 1 and Figure 2).

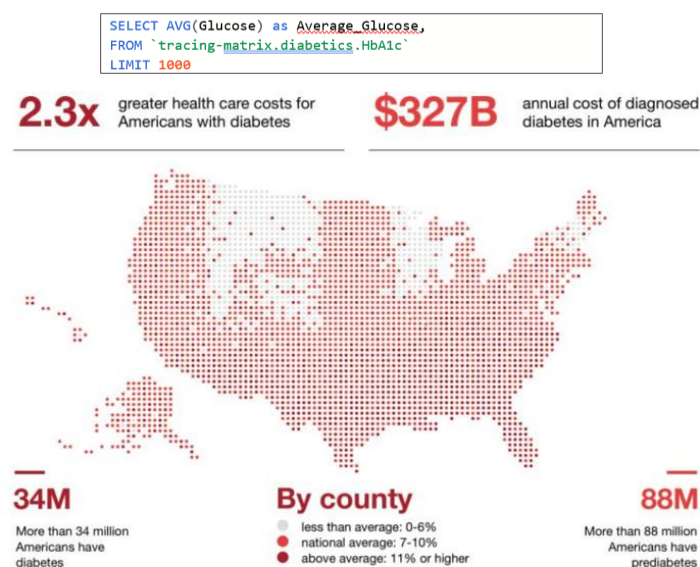


Figure 1: Diabetes average in US.

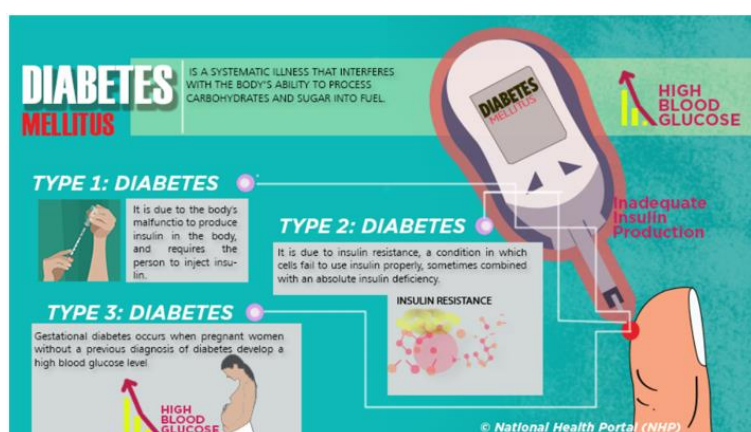


Figure 2: Reference in diabetes mellitus | National health portal of India (nhp.gov.in).

WHAT IS HbA1c?

Glycated hemoglobin (HbA1c) is an important indicator of long-term glucose levels and has been recommended for use in the diagnosis of diabetes mellitus (DM) by the American Diabetes Association (ADA) and the World Health Organization (WHO) [2]. HbA1c is created when glucose binds to hemoglobin, which is found in red blood cells [3]. Glycated hemoglobin (HbA1c) was initially identified as “unusual” hemoglobin in patients with diabetes over 40 years ago.

The lifespan of a red blood cell varies from person to person but averages out to approximately 117 days in men and 106 days in women. Therefore, the HbA1c is an index of average blood sugar levels over the preceding 2 to 3 months. Of these 2 months - 3 months period, the immediately preceding 30 days contribute 50% to HbA1c [4].

CHALLENGES IN HbA1c

Glycated hemoglobin (HbA1c) is an important indicator of glycemic control in diabetes mellitus, based on which important diagnostic and therapeutic decisions are routinely made. However, there are several situations in which

the level of HbA1c may not faithfully reflect the glycemic control in a given patient. Important among these is the use of certain non-diabetic medications, which can affect the HbA1c levels in different ways.

Normal adult hemoglobin consists predominantly of HbA1, of which HbA1c comprises approximately 5% [4]. However, HbA1c does not always correlate with blood glucose levels. It is possible to have high blood glucose but low HbA1c due to the presence of hemoglobin variants, inflammation, or other factors that increase red blood cell turnover.

After that discovery, numerous small studies were conducted correlating it to glucose measurements resulting in the idea that HbA1c could be used as an objective measure of glycaemic control.

The A1C-Derived Average Glucose (ADAG) study included 643 participants representing a range of A1C levels. It established a validated relationship between A1C and average glucose across a range of diabetes types and patient populations [5-13]. HbA1c was introduced into clinical use in the 1980s and subsequently has become a cornerstone of clinical practice [7].

HbA1c reflects average plasma glucose over the previous eight to 12 weeks. It can be performed at any time of the day and does not require any special preparation such as fasting. These properties have made it the preferred test for assessing glycaemic control in people with diabetes. More recently, there has been substantial interest in using it as a diagnostic test for diabetes and as a screening test for persons at high risk of diabetes.

Owing in large part to the inconvenience of measuring fasting plasma glucose levels or performing an OGTT, and day-to-day variability in glucose, an alternative to glucose measurements for the diagnosis of diabetes has long been sought. HbA1c has now been recommended by an International Committee and by the ADA as a means to diagnose diabetes. Although it gives equal or almost equal sensitivity and specificity to a fasting or post-load glucose measurement as a predictor of prevalent retinopathy, it is not available in many parts of the more recently in the DETECT-2 analysis Overall, the performance of HbA1c has been similar to that of fasting or 2 hours plasma glucose.

For all three measures of glycaemia, the value above which the prevalence of retinopathy begins to rise rapidly has differed to some extent between studies. Although HbA1c gives equal or almost equal sensitivity and specificity to glucose measurement as a predictor of prevalent retinopathy, it is not available in many parts of the world, and in general, it is not known which is better for predicting microvascular complications. World. Also, many people identified as having diabetes based on HbA1c will not have diabetes by direct glucose measurement and vice versa.

The relationship between HbA1c and prevalent retinopathy is similar to that of plasma glucose, whether glucose and HbA1c are plotted in deciles in vigintiles or as continuous variables. This relationship was originally reported in the Pima Indians and has also been observed in several other populations including Egyptians, the NHANES study in the USA in Japanese.

HOW HIGH BLOOD SUGAR CREATES HbA1c

Red blood cells (RBCs) are exposed to glucose in the blood. The higher the blood glucose level gets the more HbA1c forms throughout the lifespan of a red blood cell [3].

The lifespan of a red blood cell varies from person to person but averages out to approximately 117 days in men and 106 days in women. Therefore, the HbA1c is an index of average blood sugar levels over the preceding 2 to 3 months. Of this 2 to 3-month period, the immediately preceding 30 days contribute 50% to HbA1c [3].

Normal adult hemoglobin consists predominantly of HbA1, of which HbA1c comprises approximately 5% [4].

However, HbA1c does not always correlate with blood glucose levels. It is possible to have high blood glucose but low HbA1c due to the presence of hemoglobin variants, inflammation, or other factors that increase red blood cell turnover. Glycated hemoglobin (HbA1c) was first described by Rahbar et al. in 1969 [1]. Subsequent studies showed that the level of HbA1c correlated well with glycemic control over a period of 2 months to 3 months, leading to the gradual incorporation of the test into clinical practice in the 1980s [2].

With the publication of the diabetes control and complications trial [3] and the United Kingdom prospective diabetes study, [4] both of which correlated the HbA1c levels to the development of diabetes complications, HbA1c estimation has become established as a cornerstone of diabetes management.

Hemoglobin (Hb) is a tetramer formed of two alpha and two beta-globin chains. On exposure to high levels of blood glucose, hemoglobin gets non-enzymatically glycated at different sites in the molecule. HbA1c is formed when glucose gets added to the N-terminal valine residue of the beta chain of Hb [5].

The levels of HbA1c in the blood reflect the glucose levels to which the erythrocyte has been exposed during its lifespan (approximately 117 days in men and 106 days in women). Therefore, the HbA1c is an index of the level of glycemic control over the preceding 2 to 3 months. Of this period, the immediately preceding 30 days contribute 50% to HbA1c [5].

A number of methods are in use for measuring HbA1c. The most widely used assay utilizes high-performance liquid chromatography (HPLC). Other methods include boronate affinity assay and immunoassays. Efforts are underway to standardize the assay worldwide [6].

The main use of HbA1c is in the monitoring of patients with diabetes. HbA1c targets of less than 6.5% to 7% have been proposed by various organizations to define good control of diabetes [7,8]. The efficacy of anti-diabetic medications is also assessed based on their HbA1c lowering capacity. Many algorithms for the management of diabetes utilize HbA1c as the basis for major therapeutic decisions. Recently, HbA1c has also been recommended as a tool for the diagnosis of diabetes [9].

HbA1c ASSAY INTERFERENCES

- 1) HbA1c Methods: Effects of hemoglobin variants (HbC, HbS, HbE, and HbD traits) and elevated fetal hemoglobin (HbF).

HbA1c, also called A1C, is a measure of the amount of glucose attached to the hemoglobin (Hb) in red blood cells. The higher the glucose levels over the previous 2 months - 3 months, the higher the A1C. The A1C test is

used to monitor the glucose levels of patients who have been diagnosed with diabetes. In people who have hemoglobin variants such as HbS (sickle cell trait), some A1C tests give falsely high or low readings that can lead to the over-treatment or under-treatment of diabetes.

Laboratories use many different methods for measuring A1C, but some of these methods can give inaccurate results when the patient has a hemoglobin variant such as sickle cell trait or if there is an elevated level of fetal hemoglobin (HbF). Doctors or patients interested in getting information about the accuracy of a particular A1C method for patients with hemoglobin variants should first find out which method your laboratory is using.

The following table lists the 20 methods most often used to measure A1C and whether the method is affected by HbC, HbS, HbE or HbD trait or by elevated HbF. Methods are listed in alphabetical order by manufacturer. The criteria used to determine whether or not a method shows interference that is clinically significant (indicated by “Yes”) is $\geq \pm 6\%$ at 6 and/or 9% A1C. If your diabetes patient has a hemoglobin variant, your lab should use a method that does not show interference from that variant in order to produce an accurate A1C result (Table 1).

Table 1: 20 methods most often used to measure A1C and whether the method is affected by HbC, HbS, HbE or HbD trait or by elevated HbF.

Method	Interference from HbC	Interference from HbS	Interference from HbE	Interference from HbD	Interference from Elevated HbF
Abbott Architect c Enzymatic	No	No	No	No	-
Alere Afinion	No	No	No	No	\$
Arkray ADAMS A1c HA-8180V (Menarini)	No	No	HbA1c not quantified (no for ver. EU 1.41)	HbA1c not quantified (no for ver. EU 1.41)	No <30%
Beckman HbA1c Advanced B00389 Manual Application on DxC 700 AU AU system	No	No	No	No	\$
Beckman HbA1c Advanced B93009 Online Application on DxC 700 AU	No	No	No	No	\$
Beckman Synchron System Unicel DxC	No	No	No	No	\$
Bio-Rad D-100 (A1c program)	No	No	No	No	-
Bio-Rad Variant II Turbo 2.0	No	No	No	No	No <25% HbF
Ortho-Clinical Vitros	No	No	No	No	\$
Roche Cobas c513	No	No	No	No	\$
Sebia Capillarys 2 Flex Piercing	No	No	No	No	No <15% HbF
Siemens DCA Vantage	No	Yes†/ No*	Yes†/ No*	No	No <10% HbF
Siemens Atellica	No	No	No	No	\$
Siemens Dimension	No	No	No	No	\$
Tosoh G8 ver. 5.24, 5.28	No	No	No	No	No <30% HbF
Trinity HPLC	No	No	No	No	No <15% HbF

UNIT CONVERSION

HbA1c is usually measured in percent or mmol/mol (mili-mols per mol).

Measuring HbA1c percentages are considered to be the “old way” and most labs nowadays report its values as mmol/mol.

If you need to convert between the two, the Australian Diabetes Society recommends the following formula [11]:

Old (%) = 0.0915 New (mmol/mol) + 2.15%.

New (mmol/mol) = 10.93 Old (%) - 23.5 mmol/mol.

As an example, using this formula, 5% HbA1c equals 31 mmol/mol (54.7-23.5).

GOOGLE CLOUD COMPUTE

Compute Engine for SerpAPI

Google cloud engine for connecting the SerpAPI.

SerpAPI Project

The project scope is to connect to the Google Cloud Search Engine with sponsored project of SerAPI LLC with acknowledgement of 40,000 credits and 10,000 throughputs in Google Cloud as application modernization.

Dataset used

Google cloud uses the diabetics dataset of Indian population.

Google cloud services used in the projects are Google compute engine, Google storage, and Data analytics service. 'Big Query'

Application development

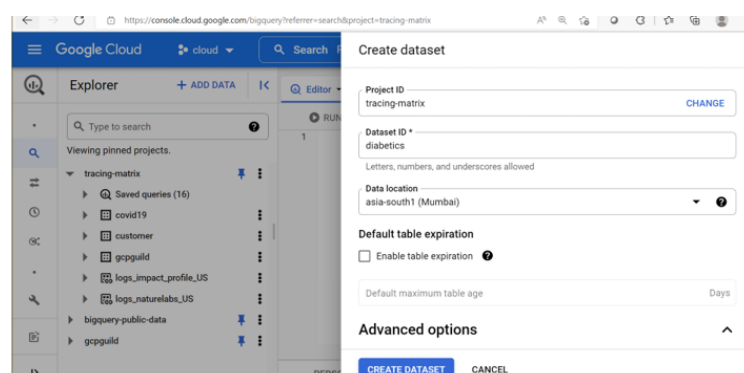
Python and dataset are accessible on-demand in google cloud.

Google Cloud - Big Query

Google cloud data analytics service, big query is considered for processing the diabetics treatment and identification of range of HbA1c.

Big Query External Table

The Big Query SQL is created for an external table dataset in the 'tracking-matrix'. The syntax for an external table for the project (Figure 3 and Figure 4).



← → ↻ https://console.cloud.google.com/bigquery/referrer=search&project=tracing-matrix Aⁿ 🔍 🏠 🔄 📄

Create table

Source

Create table from

Drive

Select Drive URI *

https://docs.google.com/spreadsheets/d/1EHDpKXb4WjNZfB0dYGh4amTWhXtwo0Ce191UBgYZgpw/edit#gid=14199329

File format

CSV

Destination

Project *

tracing-matrix

BROWSE

Dataset *

diabetics

CREATE TABLE

CANCEL

```
CREATE EXTERNAL TABLE `tracing-
matrix.diabetics.india_predictor`
(
  Pregnancies INT64,
  Glucose INT64,
  BloodPressure INT64,
  SkinThickness INT64,
  Insulin INT64,
  BMI FLOAT64,
  DiabetesPedigreeFunction FLOAT64,
  Age INT64,
  Outcome INT64
)

OPTIONS(
  skip_leading_rows=1,
  format="CSV",
  uris=["https://docs.google.com/spreadsheets/d/1EHDpKXb4WjNZfBOd
YGh4amTWhtXtwo0Ce191UBgYZgpw/edit?gid=14199329"]
);
```

*diabetics

×

+

DISABLE EDITOR TABS

▶ RUN

⋮

✔ This query will process 0 B when run.

```
1 CREATE EXTERNAL TABLE tracing-matrix.diabetics.india_predictor
2 (
3     Pregnancies INT64,
4     Glucose INT64,
5     BloodPressure INT64,
6     SkinThickness INT64,
7     Insulin INT64,
8     BMI FLOAT64,
9     DiabetesPedigreeFunction FLOAT64,
10    Age INT64,
11    Outcome INT64
12 ) |
13 OPTIONS(
14     skip_leading_rows=1,
15     format=CSV,
16     uris=["https://drive.google.com/file/d/1loVPGR5fDYitT12u45nIUMp_V6i0Me-i/view?usp=sharing"]
17 );
```

The screenshot shows the DBeaver interface for the 'india_predictor' schema. The 'Filter' section is active, displaying a table with columns: Field name, Type, Mode, Collation, Policy Tags, and Description. The table lists five fields: Pregnancies, Glucose, BloodPressure, SkinThickness, and another field (partially visible). All fields are of type INTEGER and are NULLABLE. The 'EDIT SCHEMA' button is visible at the bottom left.

Field name	Type	Mode	Collation	Policy Tags	Description
Pregnancies	INTEGER	NULLABLE			
Glucose	INTEGER	NULLABLE			
BloodPressure	INTEGER	NULLABLE			
SkinThickness	INTEGER	NULLABLE			

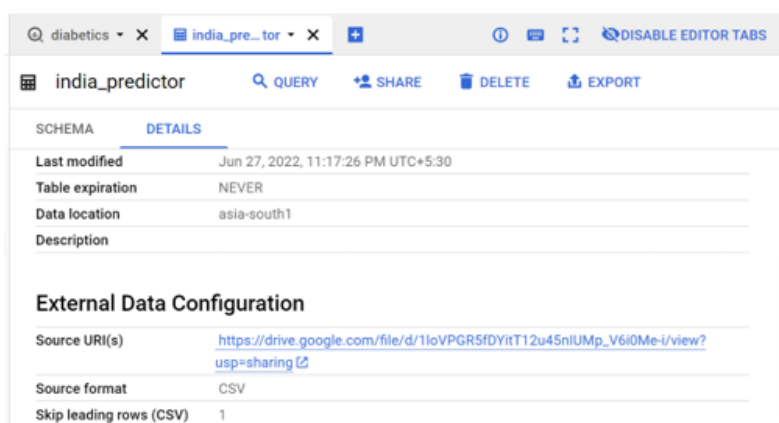


Figure 3: Big query external table.

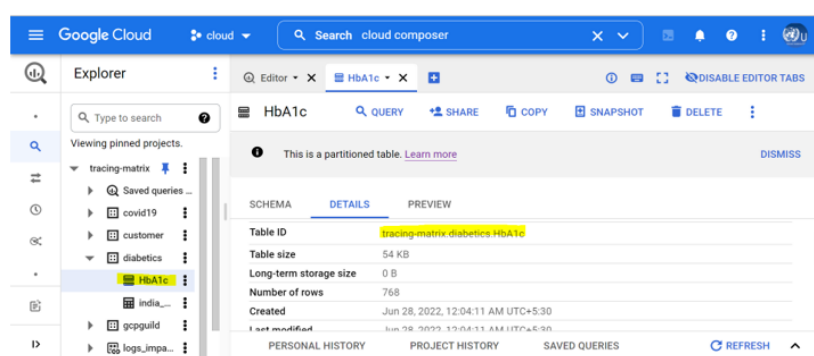


Figure 4: tracing-matrix.diabetics.india_predictor is created.

GCloud CLI (Gcloud Commands)

The Gcloud command line is your gateway to manage and interact with the cloud and offers a variety of options to automatically parse and format the results. Here, we have used in this use case demonstrates gcloud commands together with python to extract the data and create the external table with the utility to auto-generate the formats for Google Cloud-Native Services.

gcloud auth activate-service-account

--key-file "nature-labs-key.json"

GOOGLE GDK CLI

Gcloud Utilities in the Project

In order to set the tracing-matrix as a project, in CLI the following command is performed (Figure 5).

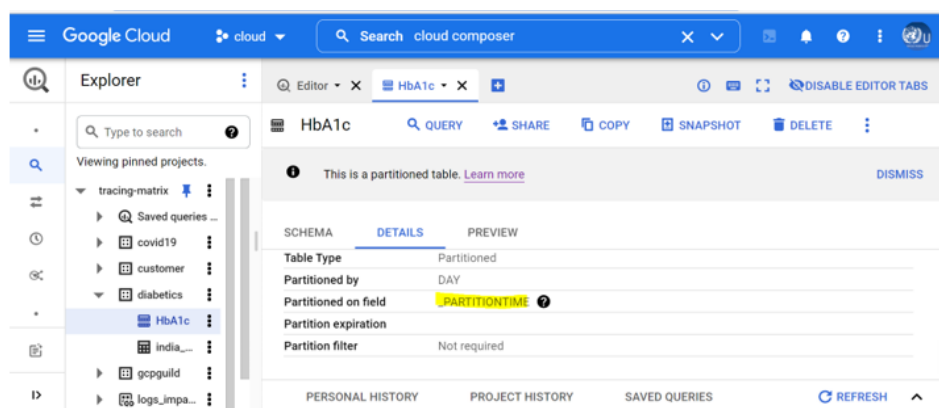


Figure 5: tracing-matrix.diabetics.india_predictor.

```
gcloud config set project tracing-matrix
```

then, to view the config settings in the GCP

```
gcloud config list
```

```
[accessibility]
```

```
screen_reader = False
```

```
[compute]
```

```
region = us-central1
```

```
[core]
```

```
account = nature-labs@tracing-matrix.iam.gserviceaccount.com
```

ABOUT DATASET

Context

This dataset is originally from the National Institute of diabetes and digestive and kidney diseases. The objective of the dataset is to diagnostically predict whether or not a patient has diabetes, based on certain diagnostic measurements included in the dataset. Several constraints were placed on the selection of these instances from a larger database. In particular, all patients here are females at least 21 years old of Pima Indian heritage.

Content

The datasets consist of several medical predictor variables and one target variable, Outcome. Predictor variables includes the number of pregnancies the patient has had, their BMI, insulin level, age, and so on.

AUTOMATION OF GOOGLE CLOUD PROVISIONING

Cloud Composer is a fully managed workflow orchestration service, that create, schedule, monitor, and manage workflows that span across clouds and on-premises data centers. Cloud Composer is built on the popular Apache Airflow open source project. Operates using the Python programming language.

The benefit of Cloud Composer in Apache Airflow, Cloud Composer work orchestration in Airflow, there is no need for installation or management overhead. Cloud Composer helps you create Airflow environments quickly

Use Airflow-native tools, such as the powerful Airflow web interface and command-line tools, so you can focus on your workflows and not your infrastructure.

AUTOMATING INFRASTRUCTURE WITH CLOUD COMPOSER

The use case of a Google Search Engine API in a real-world scenario for diabetic's project on Big Query for processing the Diabetics treatment and identification of the range of HbA1c.

Hundreds of VMs belong to different parts of the projects in diabetics patents and compute the HbA1c, and connect to pub/sub, web interface and search engine of Google engine and each workflow is requiring different backup schedules.

Scaling applies not only to our example with Compute Engine VMs but to any infrastructure component for which a scheduled process needs to be run

Cloud Composer excels at these complex scenarios because it's a full-fledged workflow engine based on Apache Airflow hosted in the cloud and not just an alternative to Cloud Scheduler or cron.

Airflow DAGs, which are flexible representations of a workflow, adapt to real-world needs while still running from a single codebase.

Create custom Airflow operators and a sensor for Compute Engine.

Create a Cloud Composer workflow using the Airflow operators and a sensor. Schedule the workflow to back up a Compute Engine instance at regular intervals.

Google Services and Airflow Listing

Compute Engine

GKE

Cloud Storage

Cloud Composer Environment

Key Features of Cloud Composer

Hybrid and multi-cloud

Ease your transition to the cloud or maintain a hybrid data environment by orchestrating workflows that cross between on-premises and the public cloud. Create workflows that connect data, processing, and services across clouds to give you a unified data environment.

Fully managed workflow orchestration

Cloud composer's managed nature and Apache airflow compatibility allows you to focus on authoring, scheduling, and monitoring your workflows as opposed to provisioning resources.

Integrates with other google cloud products

End-to-end integration with google cloud products including big query, dataflow, Dataproc, datastore, cloud storage, pub/sub, and AI Platform gives users the freedom to fully orchestrate their pipeline.

Supports hybrid and multi-cloud

Author, schedule, and monitor your workflows through a single orchestration tool-whether your pipeline lives on-premises, in multiple clouds, or fully within Google Cloud.

Open source

Cloud Composer is built upon [Apache Airflow](#), giving users freedom from lock-in and portability. This open source project, which Google is contributing back into, provides freedom from lock-in for customers as well as integration with a broad number of platforms, which will only expand as the Airflow community grows.

Easy orchestration

Cloud Composer pipelines are configured as directed acyclic graphs (DAGs) using Python, making it easy for any user. One-click deployment yields instant access to a rich library of connectors and multiple graphical representations of your workflow in action, making troubleshooting easy. Automatic synchronization of your directed acyclic graphs ensures your jobs stay on schedule.

Contribution of Google Cloud Guild

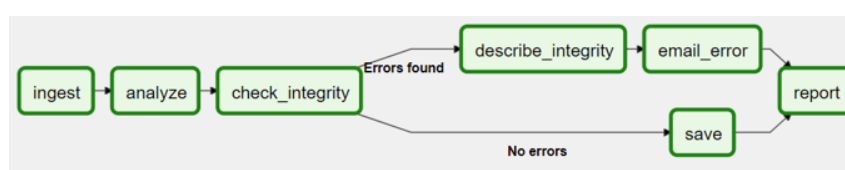
Kyndryl's services would cover the following activities:

The project is supported by more than 200 hours of work by Google Cloud Guild members of Kyndryl.

- Setting up of the Google Cloud account
- Provisioning automation of Google Cloud services, projects, Airflow, and Astronomer.
- Proving the automation in Automating infrastructure with Cloud Composer
- Airflow & DAG - control

Architecture Overview

- A DAG specifies the dependencies between Tasks, and the order in which to execute them and run retries; the Tasks themselves describe what to do, be it fetching data, running analysis, triggering other systems, or more.
- An Airflow installation generally consists of the following components:
- A scheduler, which handles both triggering scheduled workflows, and submitting Tasks to the executor to run.
- An executor, which handles running tasks. In the default Airflow installation, this runs everything inside the scheduler, but most production-suitable executors actually push task execution out to workers.
- A web server, which presents a handy user interface to inspect, trigger and debug the behavior of DAGs and tasks.
- A folder of DAG files, read by the scheduler and executor (and any workers the executor has)
- A metadata database, used by the scheduler, executor, and webserver to store state.
- Airflow is a platform that lets you build and run workflows.
- A workflow is represented as a DAG (a Directed Acyclic Graph), and contains individual pieces of work called Tasks, arranged with dependencies and data flows taken into account (Figure 6).



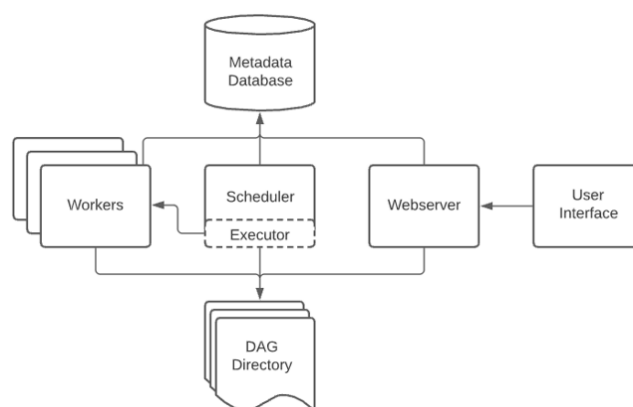


Figure 6: Architecture overview.

DAG Workloads

A DAG runs through a series of Tasks, and there are three common types of tasks you will see:

Operators, predefined tasks that you can string together quickly to build most parts of your DAGs.

Sensors, a special subclass of Operators are entirely about waiting for an external event to happen.

A TaskFlow-decorated `@task`, which is a custom Python function packaged up as a Task.

Internally, these are all actually subclasses of Airflow's BaseOperator, and the concepts of Task and Operator are somewhat interchangeable, but it's useful to think of them as separate concepts - essentially, Operators and sensors are templates, and when you call one in a DAG file, you're making a task (Figure 7 and Figure 8).

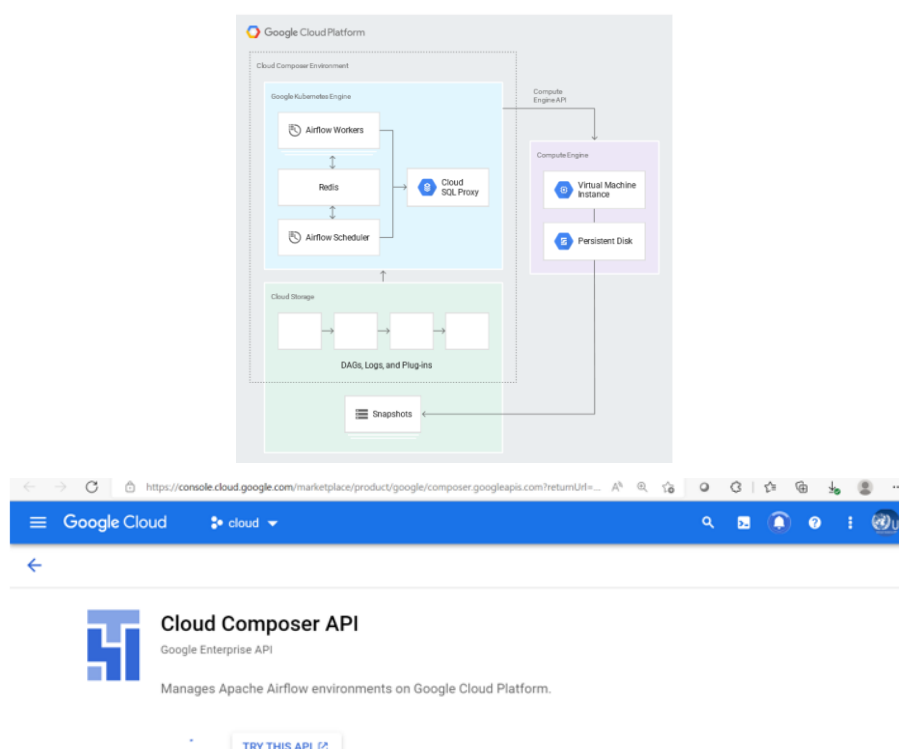


Figure 7: Google cloud platform.

The figure consists of three vertically stacked screenshots of the Google Cloud Composer 'Create environment' page. Each screenshot shows a different section of the form.

- Top Screenshot:** Shows the 'Location' dropdown menu set to 'asia-south1'. Below it, the 'Node configuration' section is partially visible, showing 'Node count' set to 3 and 'Zone' set to 'asia-south1-c'.
- Middle Screenshot:** Shows the 'Node configuration' section in more detail. 'Node count' is 3, 'Zone' is 'asia-south1-c', and 'Disk size (GB)' is set to 100. Below this, the 'OAuth Scopes' and 'Service account' sections are visible, with the service account set to 'nature-labs@tracing-matrix.iam.gserviceaccount.com'.
- Bottom Screenshot:** Shows the 'Cloud SQL configuration' section. The 'Cloud SQL machine type' dropdown is set to 'db-n1-standard-2 (2 vCPU, 7.5 GB memory)'. Below this, the 'Network configuration' section is partially visible.

Figure 8: Creating the environment.

In Airflow, workflows are created using DAGs, or "Directed Acyclic Graphs".

A DAG is a collection of tasks that you want to schedule and run, organized in a way that reflects their relationships and dependencies. DAGs are created in Python scripts, which define the DAG structure (tasks and their dependencies) using code.

Each task in a DAG can represent almost anything—for example, one task might perform any of the following functions:

- Preparing data for ingestion
- Monitoring an API
- Sending an email
- Running a pipeline

A DAG should not be concerned with the function of each constituent task—its purpose is to ensure that each task is executed at the right time, in the right order, or with the right issue handling.

To run workflows, we have created an environment for Google Search API. Airflow depends on many micro-services to run, so Cloud Composer provisions Google Cloud components to run your workflows. These components are collectively known as a Cloud Composer environment.

Environments are self-contained Airflow deployments based on Google Kubernetes Engine. This project work with other Google Cloud services using connectors built into Airflow. We have also created ore environments in a tracing-matrix Google Cloud project. Cloud Composer environments in Asia-south-a, Mumbai region.

Cloud Composer environment uses multiple Google Cloud products, but you don't have to configure each one individually.

Cloud storage

The airflow DAG and logs are stored in a cloud storage bucket.

Google kubernetes engine

The airflow platform is based on a micro-service architecture and is suitable to run in GKE.

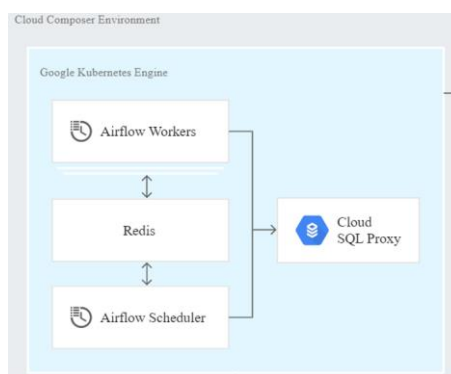
- Airflow workers load workflow definitions from Cloud Storage and run each task, using the Compute Engine API.
- The Airflow scheduler makes sure that backups are executed in the configured cadence, and with the proper task order.
- Redis is used as a message broker between Airflow components.
- Cloud SQL Proxy is used to communicate with the metadata repository.
- Cloud Composer also uses a Cloud SQL instance for metadata and an App Engine Flex app that serves the Airflow UI.
- These resources are not pictured in the diagram because they live in a separate Google-managed project.

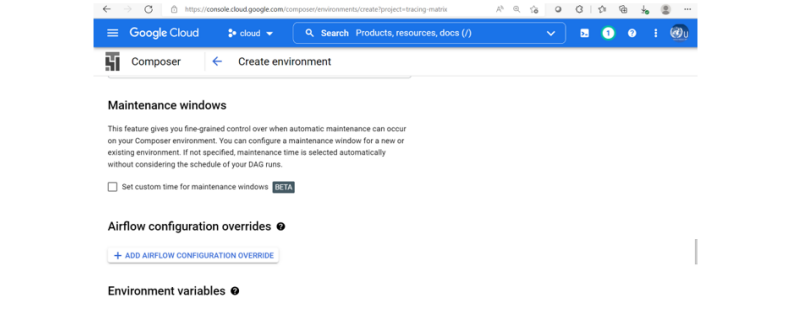
Google Cloud Project Process

Table 2 and Table 3; Figure 9 and Figure 10.

Automation of Google Cloud Provisioning Process	
Organization	Enable cloud composer API
Validation	Set up airflow - astronomers
Airflow	Provision of services, schedulers, projects
Mapping	Map to google cloud
Roles	Assign administrative access
Access	Ensure admin users responsible for setup have the right level of access
Billing	Set up billing
Accounts	Select primary billing account for initial setup
Configuration	Configure resource hierarchy: Configure initial folders and project's structure
Policy	Assign colleagues the right level of access
Support	Choose a support model

Table 2: Google cloud project process.





Maintenance windows

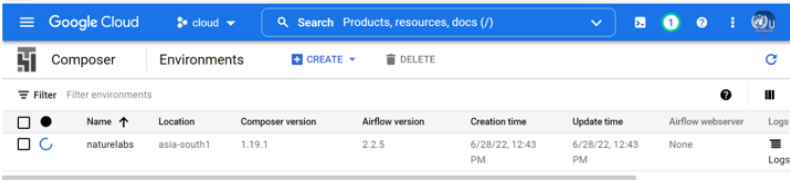
This feature gives you fine-grained control over when automatic maintenance can occur on your Composer environment. You can configure a maintenance window for a new or existing environment. If not specified, maintenance time is selected automatically without considering the schedule of your DAG runs.

☐ Set custom time for maintenance windows **BETA**

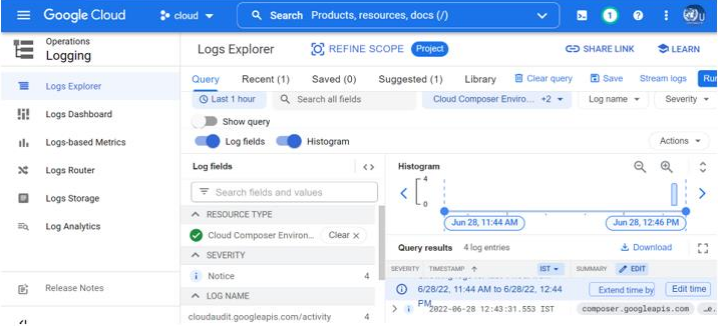
Airflow configuration overrides

[+ ADD AIRFLOW CONFIGURATION OVERRIDE](#)

Environment variables



Name	Location	Composer version	Airflow version	Creation time	Update time	Airflow webserver
naturelabs	asia-south1	1.19.1	2.2.5	6/28/22, 12:43 PM	6/28/22, 12:43 PM	None



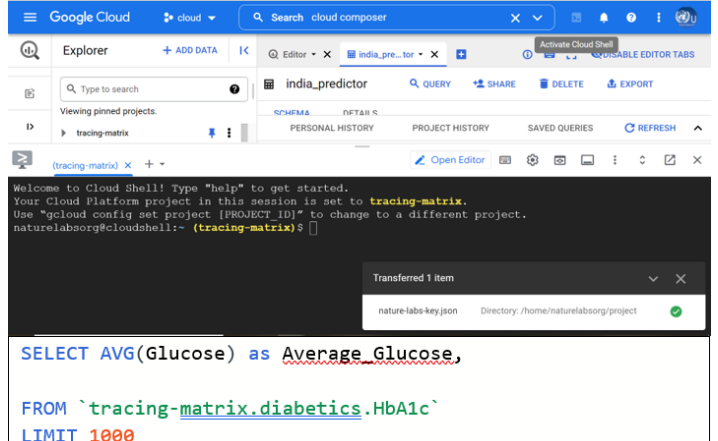
Logs Explorer

Query: Recent (1) Saved (0) Suggested (1) Library

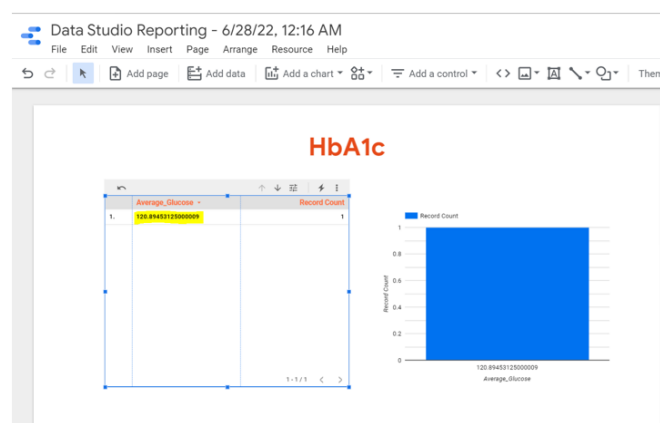
Log fields: Notice (4), Log Name (4)

Query results: 4 log entries

SEVERITY	TIMESTAMP	Log Name
Notice	6/28/22, 11:44 AM to 6/28/22, 12:44 PM	cloudaudit.googleapis.com/activity



```
SELECT AVG(Glucose) as Average_Glucose,
FROM `tracing-matrix.diabetics.HbA1c`
LIMIT 1000
```



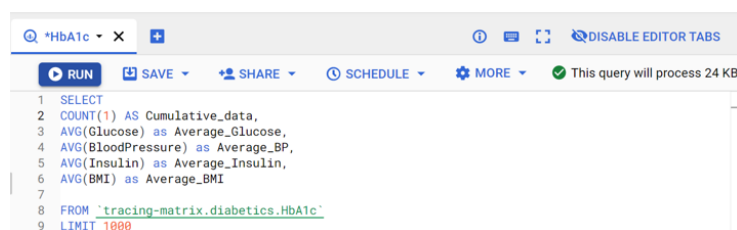


Figure 9: Climb composer environment.

Table 3: Average ratio in patients.

Numbers of Patients	Average Glucose	Average BP	Average Insulin	Average BMI
768	120.89	69.10	79.79	31.99

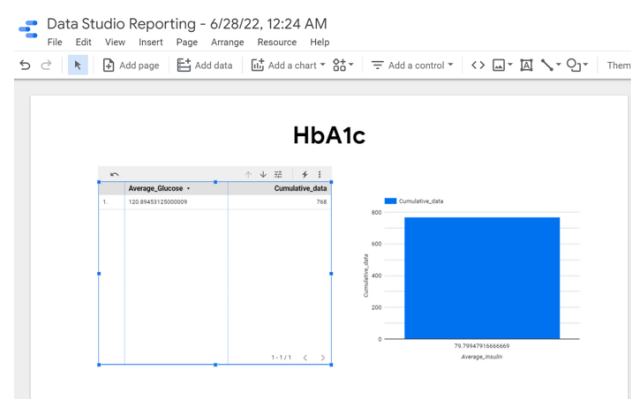


Figure 10: Data studio reporting.

SERPAPI SPONSORSHIP

Google Engine @ Google Cloud Nature Labs is sponsored by SerpApi. SerApi has provisioned with the API Key for scrapes Google and other search engines.

On behalf of Google Cloud Nature Labs, researchers, we express our gratitude to SerpAPI LLC, for provisioning their sponsorship SerpAPI's sponsorship has helped us make our research and social work contribution to speaking out greater audience.

With the advent of SerpAPI, Nature Labs has addressed our research work on Temples in India with the experience of a blazingly fast, super easy to use, and data-rich API in Google Cloud Platform Search Engine on Big Query for Research in Google Cloud Engine. With SerpAPI, Nature Labs will be helping the student community on projects.

ABOUT SERPAPI

SERP API is a real-time API to access Google search results. It solves the issues of having to rent proxies, solving captchas, and JSON parsing.

REFERENCES

1. World Health Organization (2011) Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus: Abbreviated report of a WHO consultation.
2. Kang SH, Jung DJ, Choi EW et al. (2015) HbA1c levels are associated with chronic kidney disease in a non-diabetic adult population: A Nationwide survey (KNHANES 2011–2013). PLoS One 10(12): e0145827.

3. Schwab S, Zierer A, Heier M et al. (2015) Intake of vitamin and mineral supplements and longitudinal association with HbA1c levels in the general non-diabetic population-results from the MONICA/KORA S3/F3 study. *PloS One* 10(10): e0139244.
4. Rahbar S, Blumenfeld O, Ranney HM (1969) Studies of an unusual hemoglobin in patients with diabetes mellitus. *Biochemical and biophysical research communications* 36(5): 838-843.
5. Unnikrishnan R, Anjana RM, Mohan V (2012) Drugs affecting HbA1c levels. *Indian Journal of Endocrinology and Metabolism* 16(4): 528.
6. Sherwani SI, Khan HA, Ekhzaimy A et al. (2016) Significance of HbA1c test in diagnosis and prognosis of diabetic patients. *Biomarker Insights* 11: BMI-S38440.
7. Massi-Benedetti M (2006) Changing targets in the treatment of type 2 diabetes. *Current Medical Research and Opinion* 22(Suppl 2): S5-S13.
8. Little RR, Rohlfing CR, Hanson S et al. (2008) Effects of hemoglobin E and D traits on glycated hemoglobin (HbA1c) measurements by twenty-three methods. *Clinical Chemistry* 54(8): 1277-1282.
9. Little RR, Rohlfing CL, Hanson SE et al. (2012) The effect of increased fetal hemoglobin on 7 common Hb A1c assay methods. *Clinical Chemistry* 58(5): 945-947.
10. Rohlfing CL, Connolly SM, England JD et al. (2008) The effect of elevated fetal hemoglobin on hemoglobin A1c results: five common hemoglobin A1c methods compared with the IFCC reference method. *American Journal of Clinical Pathology* 129(5): 811-814.
11. <https://diabetessociety.com.au/>
12. Smith JW, Everhart JE, Dickson WC et al. (1988) Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. In *Proceedings of the annual symposium on computer application in medical care*. American Medical Informatics Association: 261.
13. <https://github.com/gcpguild/diabetics/blob/main/googlengine.py>