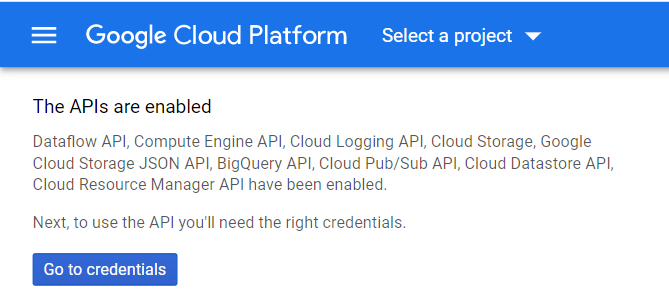
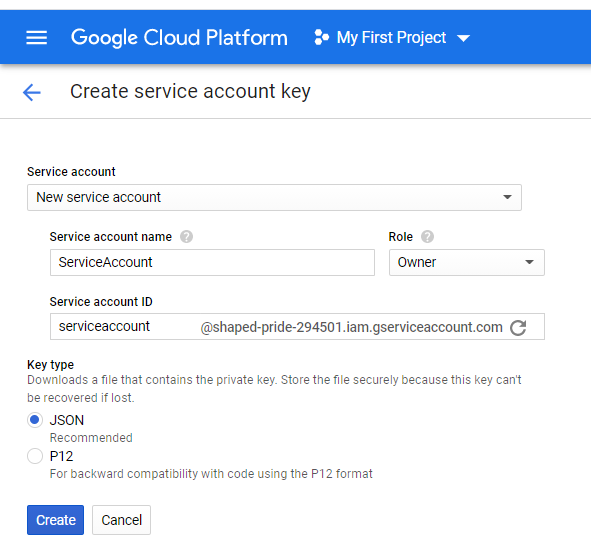
**APACHE BEAM Tutorial**

**Enable the Cloud Dataflow, Compute Engine, Stackdriver Logging, Cloud Storage, Cloud Storage JSON, BigQuery, Cloud Pub/Sub, Cloud Datastore, and Cloud Resource Manager APIs.**

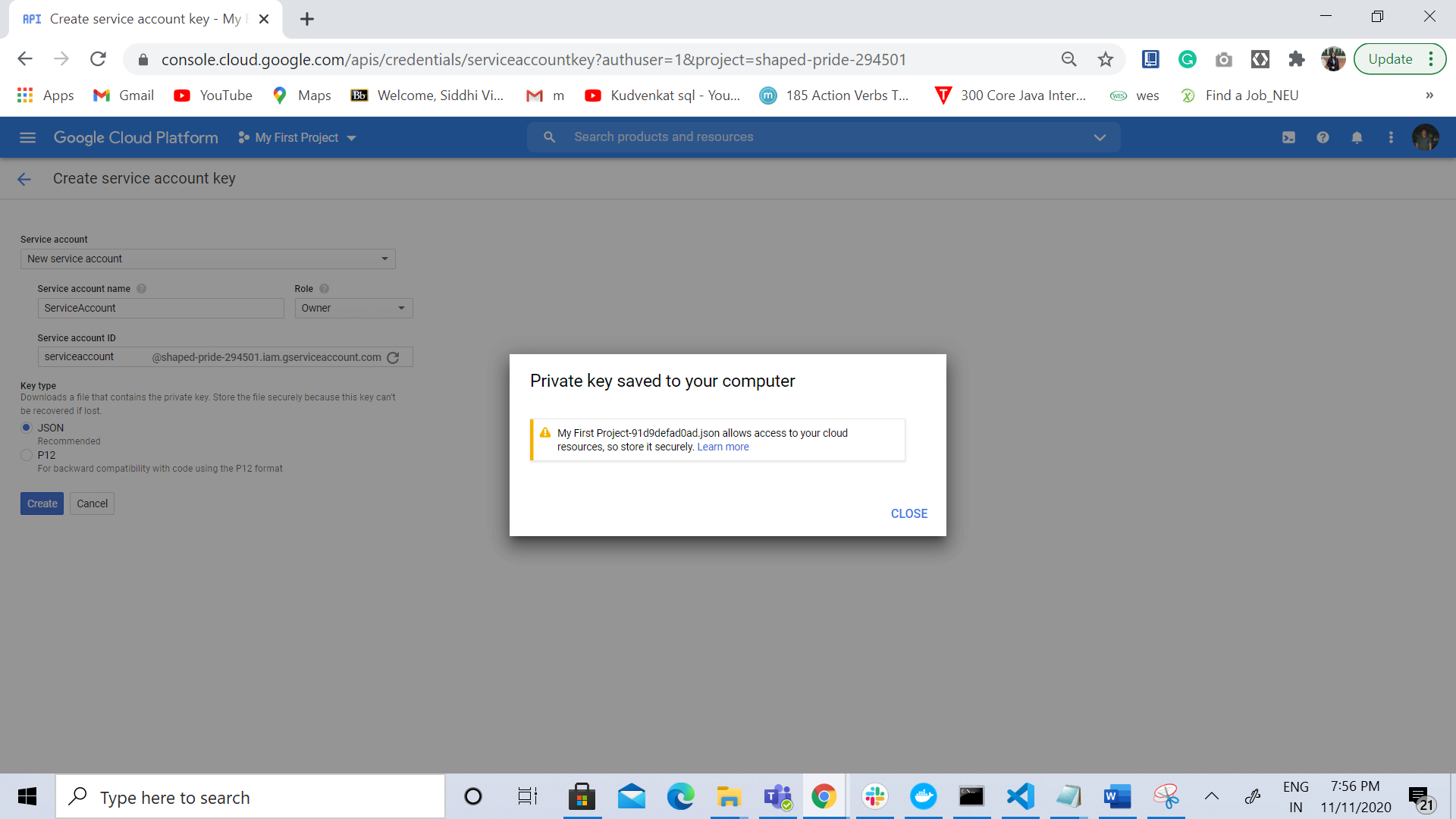
****

**Set up authentication:**

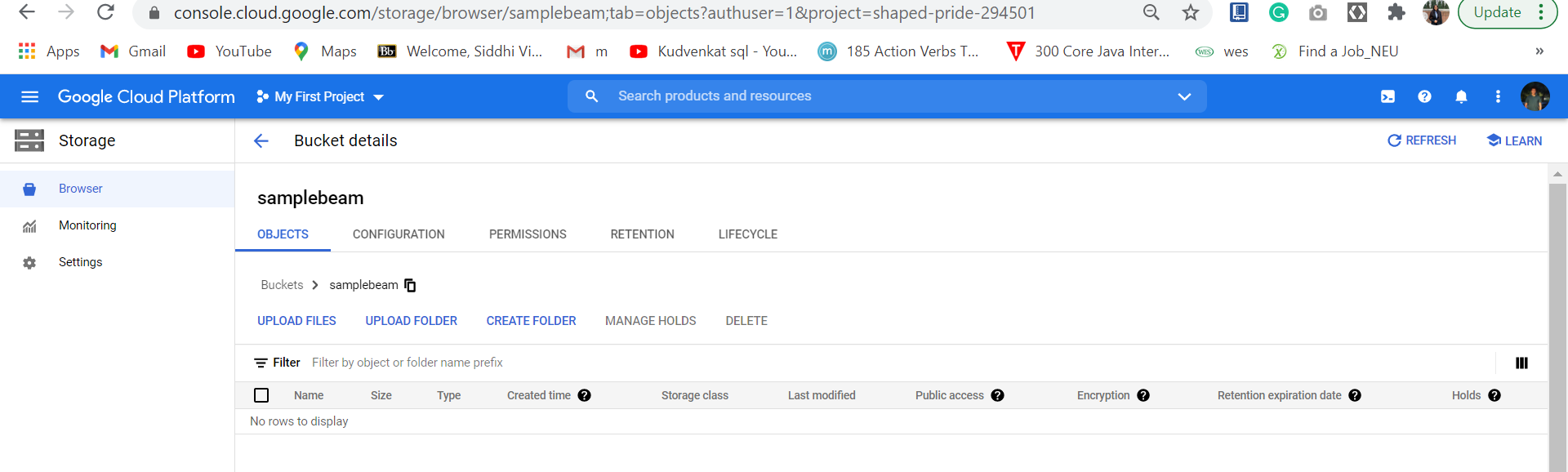
1. **In the Cloud Console, go to the Create service account key page.**

****

**After creating service account key**

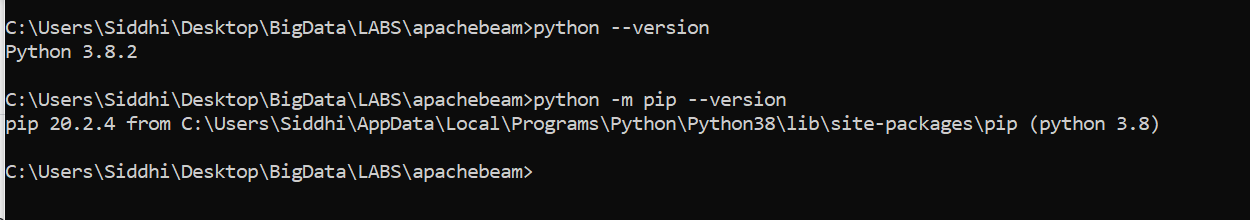
****

**Create bucket**

****

**Set up your environment**

1. **Checking for a working Python and pip installation:**

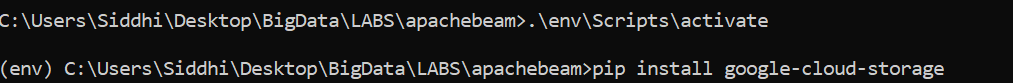
****

1. **Setup and activate a**[**Python virtual environment**](https://cloud.google.com/python/setup#installing_and_using_virtualenv)**.**

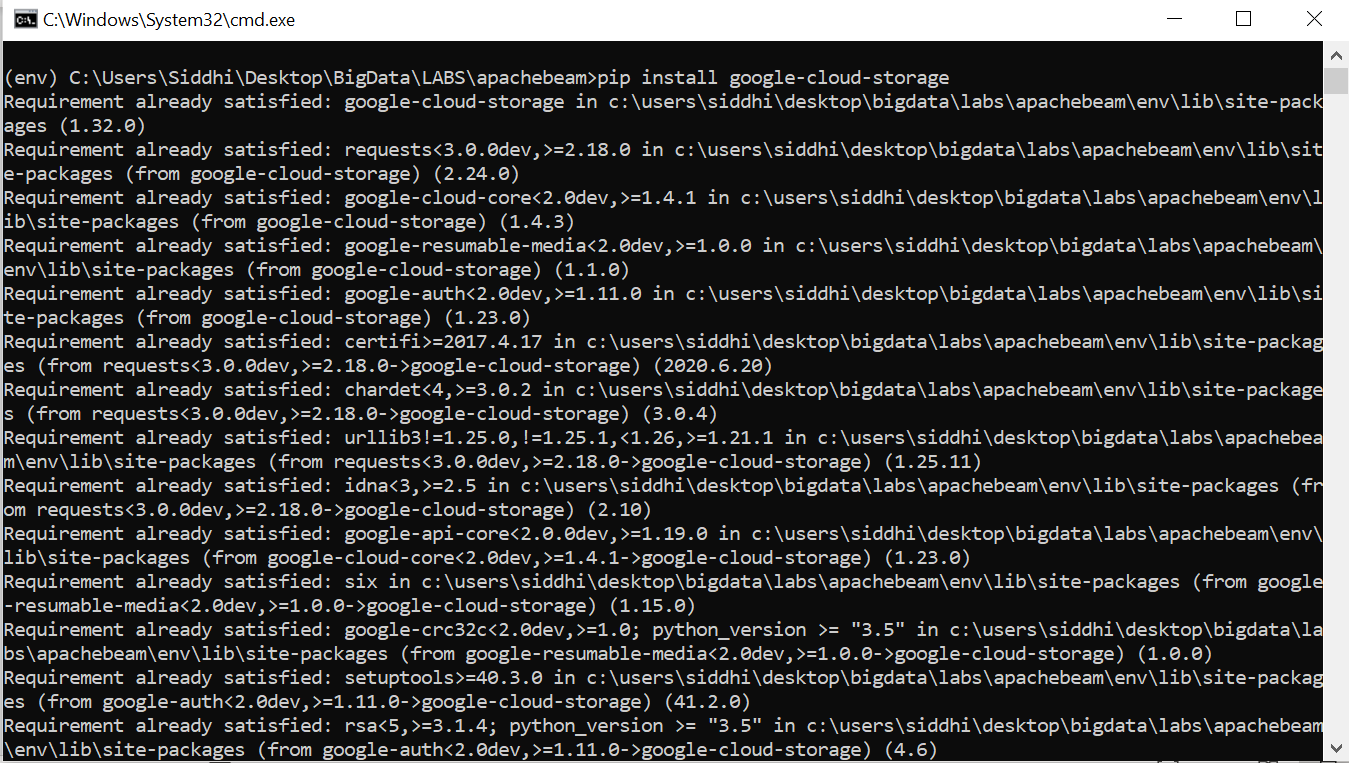
**Using the venv command to create a virtual copy of the entire Python installation.**

****

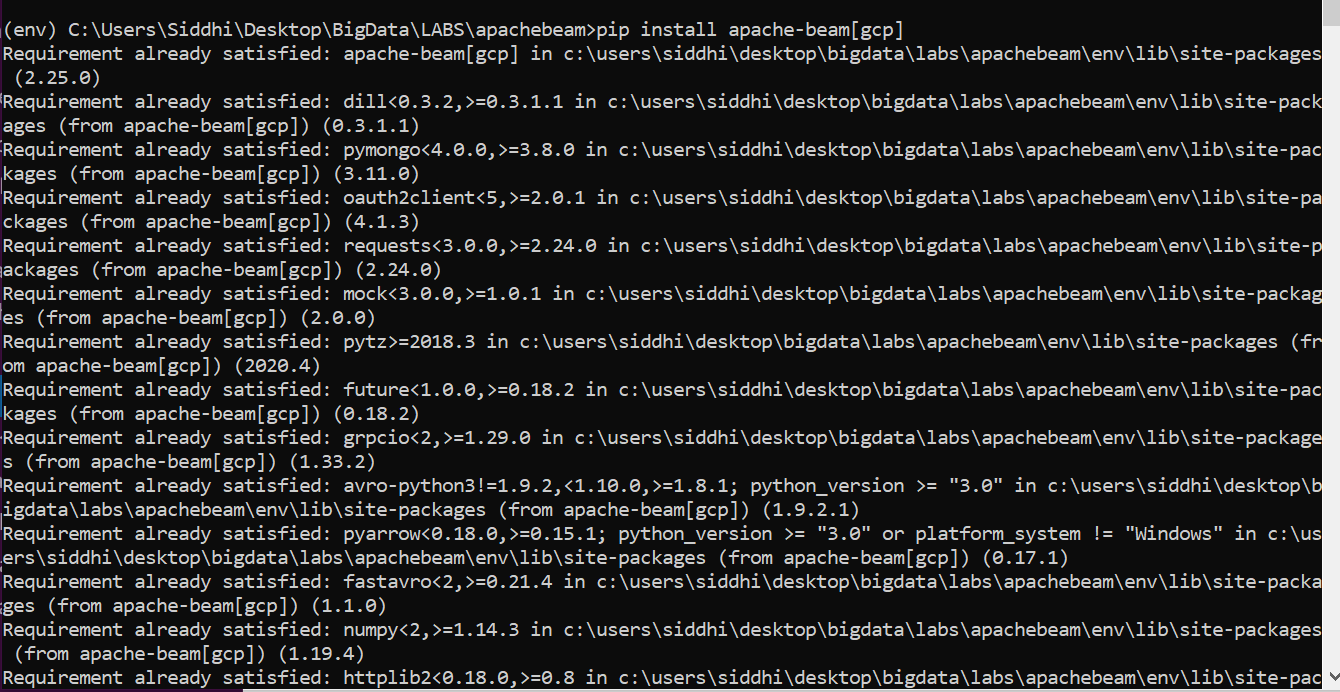
**Setting up shell to use the venv paths for Python by activating the virtual environment.**

****

**Installing packages without affecting other projects or your global Python installation:**

****

**Get the Apache Beam SDK**

****

## Run WordCount locally

## 

## Viewing the output:

## 

## Run WordCount on the Dataflow service

## Defining PROJECT, BUCKET, and REGION variables:

## 

## Running the pipeline:

## 

## Viewing results using GCP

## 

## Viewing the results in the files:

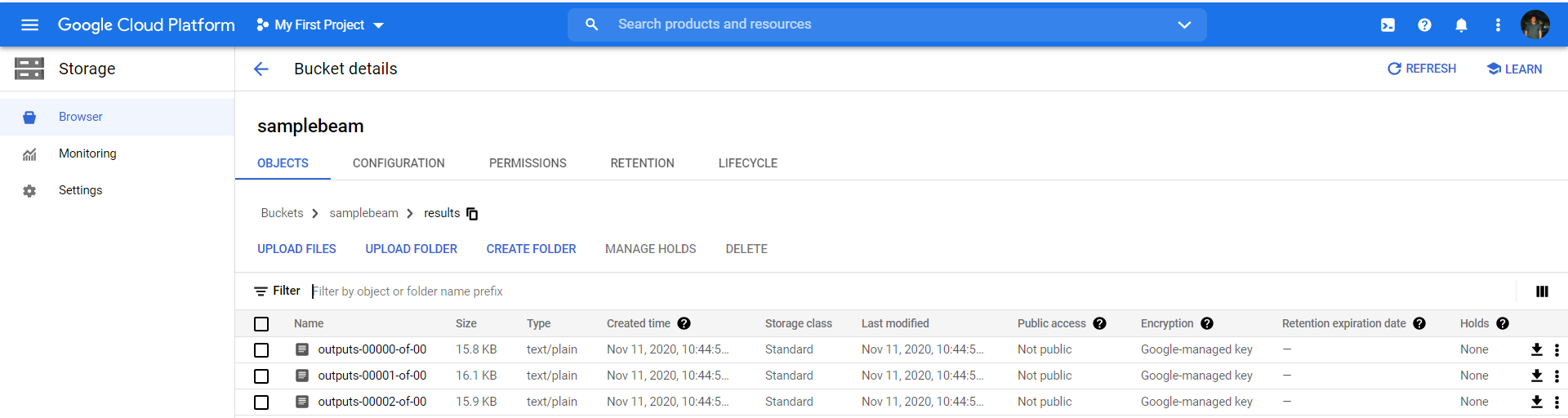
## 

**View Results :-**

1. **Dataflow monitoring UI to view results.**

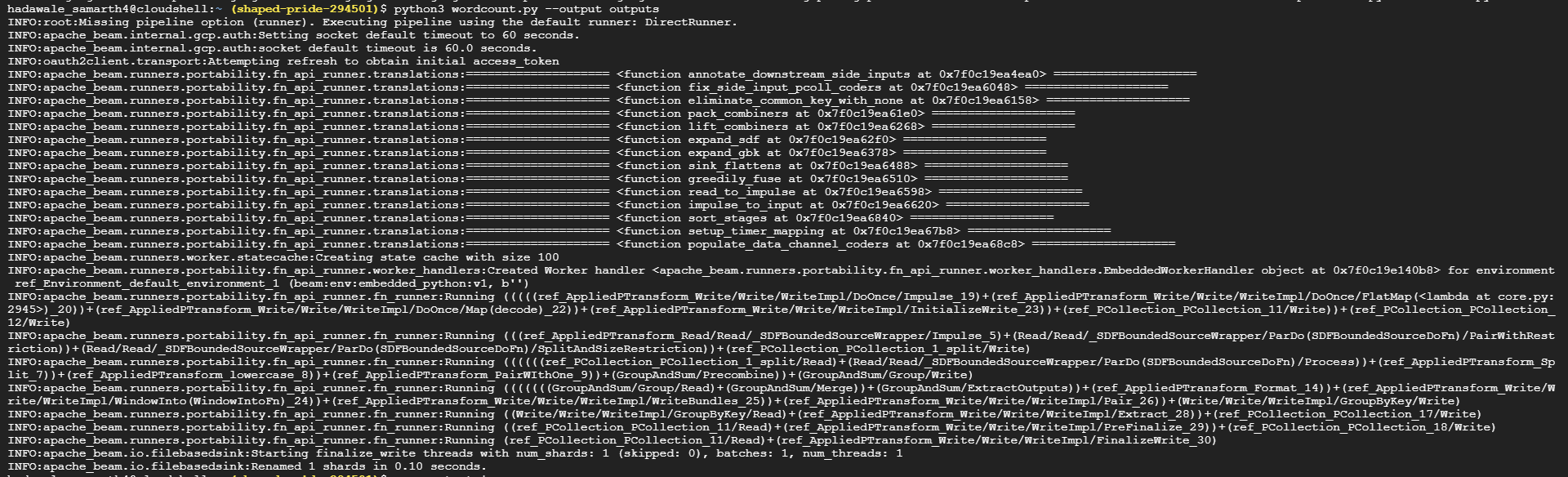
## 

1. **Open the Cloud Storage Browser in the Google Cloud Console**

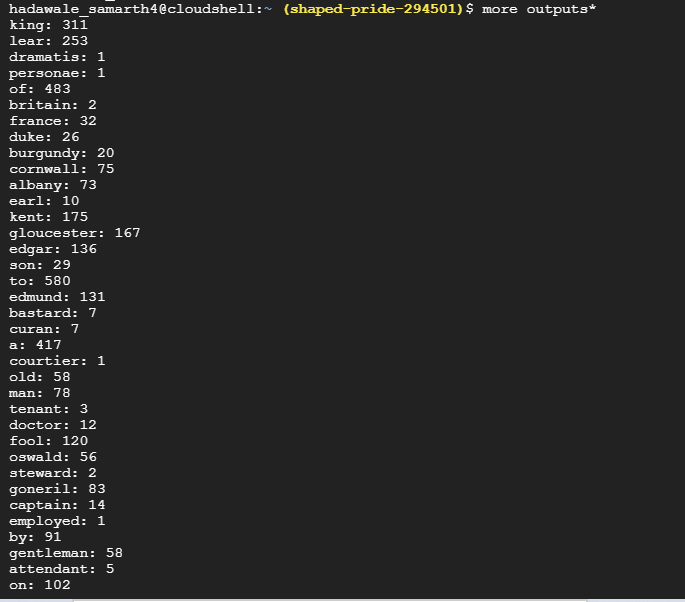


## Modifying pipeline code

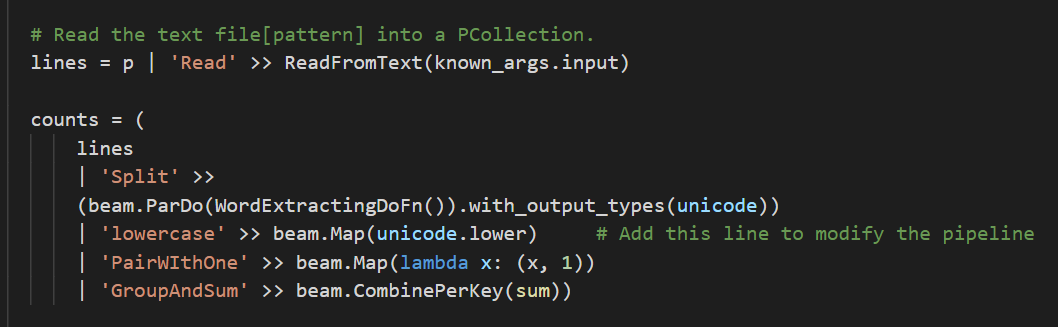
1. **Executing the code :-**



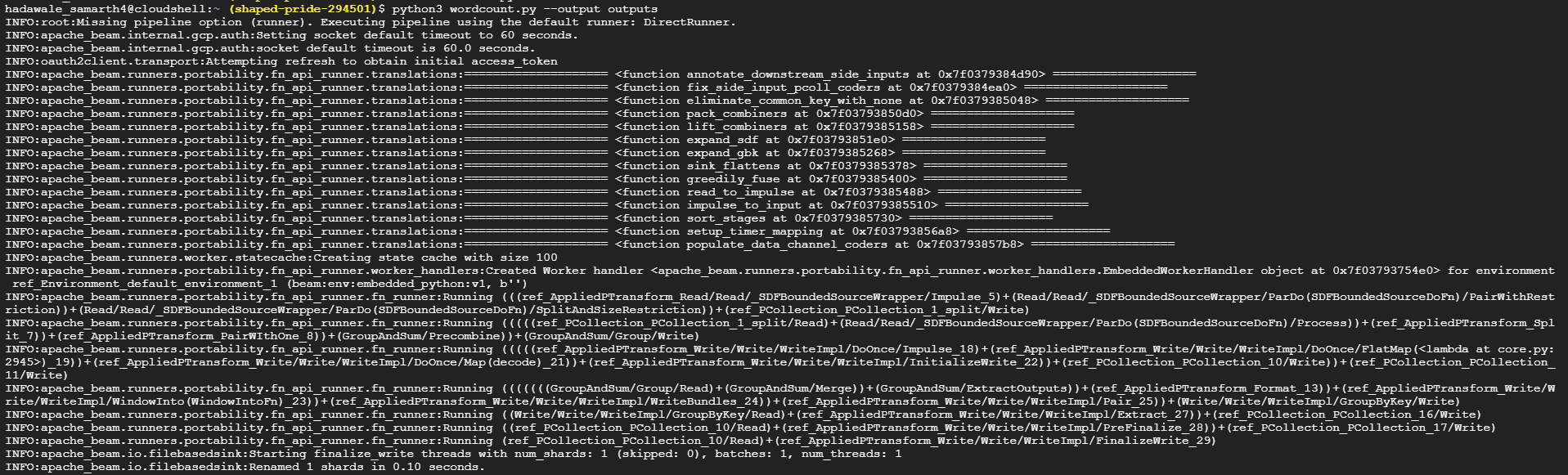
1. **View results of the pipeline code:-**



1. **Modify the pipeline function to lowercase the strings**



1. **Save the file and run the modified WordCount job**



1. **View the results of the modified pipeline code:-**

