**Project Name** - GCP-Learning-05-PanMergeGrabhX,

**Project ID** - arvindershinh-05

**Set Project**

<https://cloud.google.com/dataproc/docs/guides/setup-project>

**Create bucket from GCP console**

<https://cloud.google.com/dataproc/docs/guides/setup-project>

arvindershinh\_pan-merge-graphx\_bucket05

**Set project in gclud CLI**

gcloud config set project arvindershinh-05

**Create dataproc using below command in gcloud CLI**

<https://cloud.google.com/dataproc/docs/guides/setup-project>

gcloud dataproc clusters create arvindershinh-cluster05 --region=asia-south2

**make directory and add PAN\_Merging\_Graphx.jar file**

mkdir pan\_merge\_graphx

**Add JAR file in bucket**

cd pan\_merge\_graphx

gsutil cp PAN\_Merging\_Graphx.jar gs://arvindershinh\_pan-merge-graphx\_bucket05

**List items in bucket**

gsutil ls gs://arvindershinh\_pan-merge-graphx\_bucket05

**Submit Job in dataproc from GCP console**

Region - asia-south2

Cluster - arvindershinh-cluster05

Main class or jar - gs://arvindershinh\_pan-merge-graphx\_bucket05/PAN\_Merging\_Graphx.jar

Job ID - job-pan-merge-graphx-05