Project

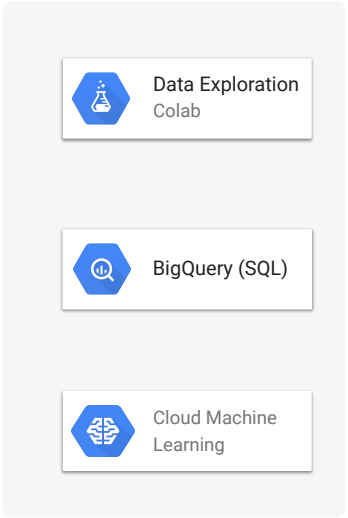
Goal #1

Full data cycle on cloud

1 Run queries on public datasets

2 Explore/Visualize public datasets

3 Predict using Machine Learning



Notes

- Concepts are the key. Easy to go ‘cloud hopping’
- Key goal: Do (a) full data cycle, (b) for 10s-100s of GBs of real data

Summary

When you finish Project . . .

- Full Query-Visualize-Learn data cycle on cloud stack
  - Work on 10s-100s of GBs of data on 100s of machines in cluster
  - See examples of multiple real datasets (COVID, ncaa, github, WorldBank, …)
  - Work with popular SQL/Visualization/Learning libraries
  - Gain intuition of scale, performance, pain and costs in running 100s of queries

- Hop-Skip-Jump to any cloud stack or application based on same concepts

1. Public datasets: [q = “BigQuery public dataset”]

2. Explore public datasets (COVID, NCAA, Weather, Hurricanes, Blockchain…)

3. Let’s pick
   a. [COVID]
      - [Schema]

Sample projects from 2018 and 2019

- When is best time to get away with murder?
- Which cities are at risk for rising pollution?
- Election Ads in 2016
- What factors influence how much a patient pays for their treatment/drug requirements?

⇒ Last year, many strong projects. Students said project helped their internship/jobs experience.

Signup for GCP

Dear Students,

Here is the URL you will need to access in order to request a Google Cloud Platform coupon. You will be asked to provide your school email address and name. An email will be sent to you to confirm these details before a coupon is sent to you.

...

- You will be asked for a name and email address, which needs to match the domain. A confirmation email will be sent to you with a coupon code.
- You can request a coupon from the URL and redeem it until: 1/23/2021
- Coupon valid through: 9/23/2021
- You can only request ONE code per unique email address.

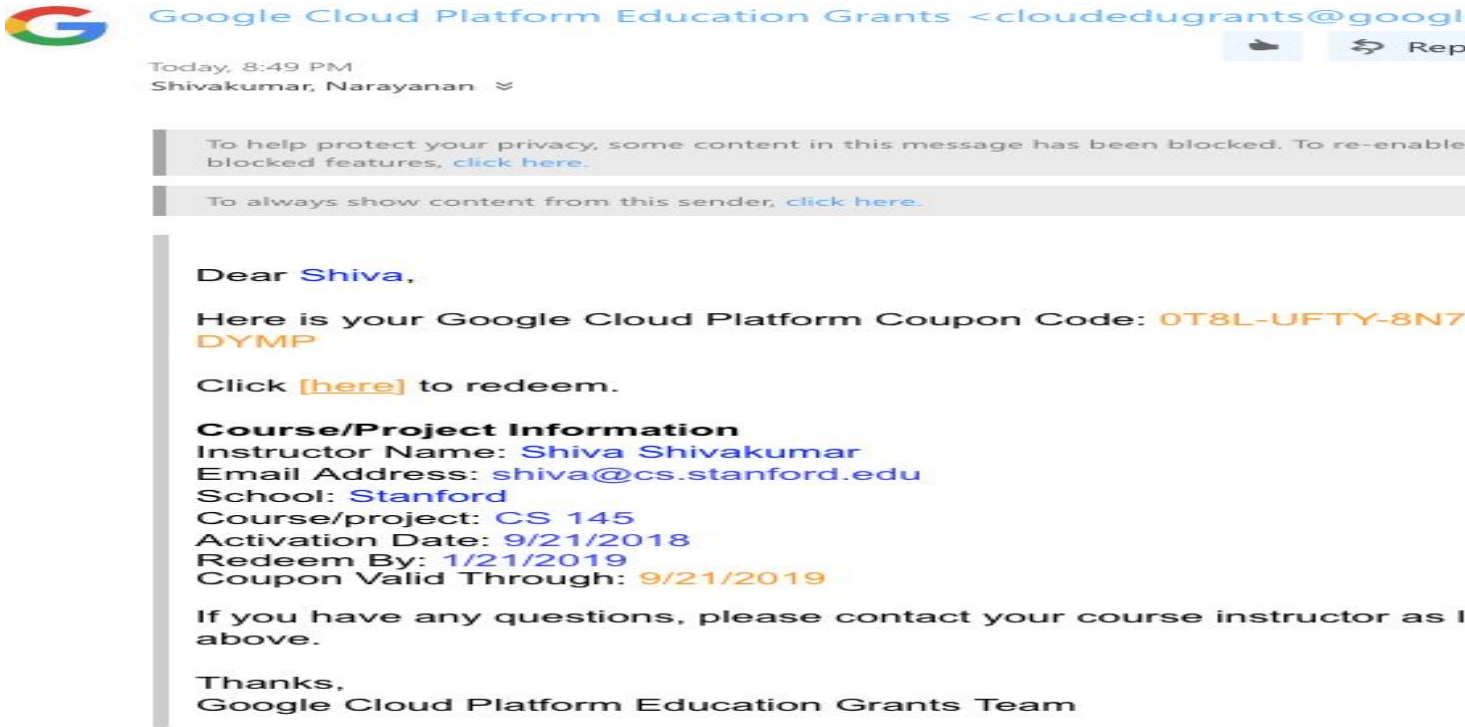
Please contact me if you have any questions or issues.

1. Signup with stanford.edu email
2. You'll get a 50$ coupon in 1-2 minutes
3. Apply to your GMAIL account (signup if you don't have one), **NOT** Stanford account
4. RECOMMEND: separate browser profile

Signup for GCP

Read Project P1

Instructions



Google Cloud Platform Education Grants <cloudedugrants@googl

Today, 8:49 PM
Shivakumar, Narayanan

To help protect your privacy, some content in this message has been blocked. To re-enable blocked features, click here.

To always show content from this sender, click here.

Dear Shiva,

Here is your Google Cloud Platform Coupon Code: 0T8L-UFTY-8N7
DYMP

Click [here] to redeem.

Course/Project Information
Instructor Name: Shiva Shivakumar
Email Address: shiva@cs.stanford.edu
School: Stanford
Course/project: CS 145
Activation Date: 9/21/2018
Redeem By: 1/21/2019
Coupon Valid Through: 9/21/2019

If you have any questions, please contact your course instructor as
above.

Thanks,
Google Cloud Platform Education Grants Team