Google Cloud Product

Who am I?

Sarah

Lead Data Scientist at zData Inc.

Contracted out to Disney for predictive maintenance

Previously:

- Sr. Statistician at Idaho National Lab
- Statistical Programmer II at The Mayo Clinic

Overview

- Cloud?
- Set up a compute unit
- Run data lab
- Query (big) data
- How to sample subset
- Exploratory Analysis

Working in the Cloud

What is the cloud?

Working in the Cloud

What is the cloud?

A remote server farm that ____ has to maintain, update, and secure.

There are several _____ :

- Azure (Microsoft)
- AWS (Amazon)
- GCP (Google)

Working in the Cloud

Why the cloud?

Working in the Cloud

Why the cloud?

- Scalability
- Ease of access
- Security of Data
- Stability

Compute Engine / Virtual Machine instances

What is CE / VM

Infrastructure as a Service (IaaS) component of Google Cloud Platform (GCP)

Same infrastructure that runs Google’s search engine, Gmail, and YouTube

Compute Engine / Virtual Machine instances

2 ways to create in GCP:

1. Through the UI
2. Through the Google Cloud Shell

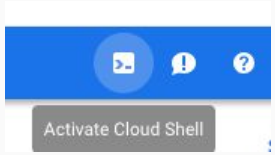
Compute Engine / Virtual Machine instances

2 ways to create in GCP:

1. Through the UI
2. **Through the Google Cloud Shell**

Compute Engine / Virtual Machine instances

Through the Google Cloud Shell

Click on the 

at the top right of the screen

Then a command line should pop up at the bottom of the screen

Type: `gcloud compute zones list`

Why did I do this?

Compute Engine / Virtual Machine instances

In cloudshell type:

```
datalab create <NAME> --zone <ZONE>
```

Replace `<NAME>` with what you want to call your VM
Replace `<ZONE>` with the zone you picked
Pick a zone for your compute engine, preferably near your location

I picked `us-west2-a` since I am in Boise, ID

```
datalab create sarahs-vm --zone us-west2-a
```

Compute Engine / Virtual Machine instances

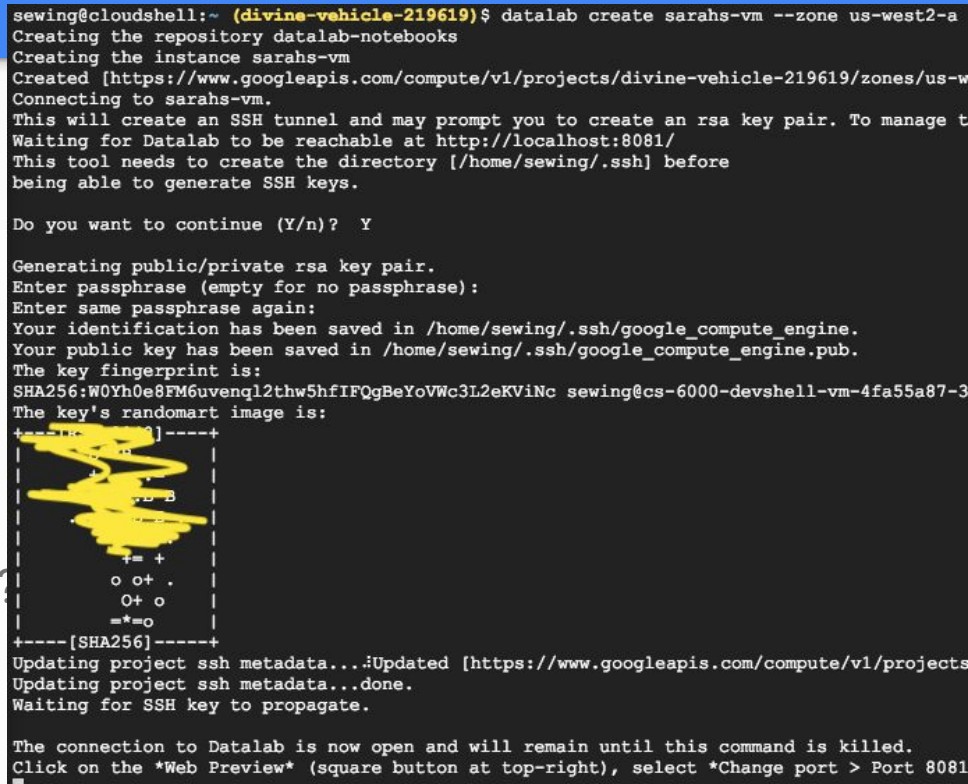
It will take a few minutes to set up

It will ask “Do you want to continue?”

Tell it “Y”

It will provide an output similar to this:

Why did I scribble out part of the image?



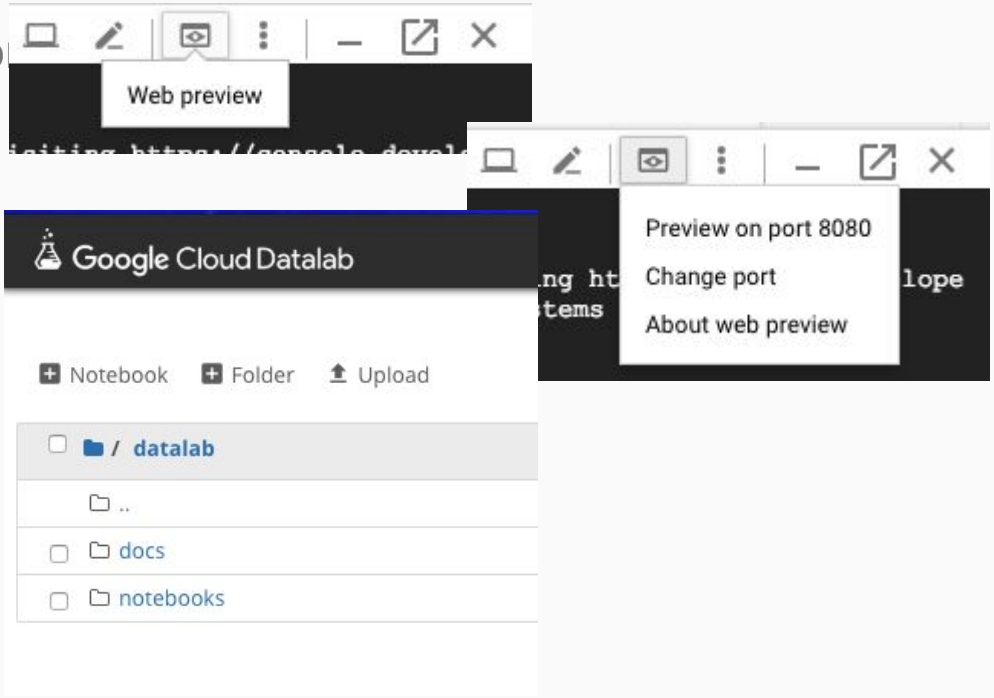
Compute Engine / Virtual Machine instances

At the right top of the cloud shell click o
preview icon.

Pick `change port`

Tell it port `8081`

Then click on `Change and Preview`



Compute Engine / Virtual Machine instances

Other important command, to restart compute engine:

```
datalab connect <instance NAME>
```