# Instructions for Bellevue Big Data Class (Google Cloud)

## Signing up and Getting Free Credits on Google Cloud

1. Go to the [Google Cloud](https://cloud.google.com) website.
2. Click Get Started for Free.
3. Sign in with your Google account or create one.
4. Follow the prompts to create your new Google Cloud account. You’ll need to provide your credit card details for verification purposes, but you won’t be charged unless you upgrade your account.
5. After setting up, you should have $300 in free credits.

**Remember, the Google Cloud free tier credits expire after 90 days or when they are all used. Always monitor your usage to avoid unexpected charges. Be sure to stop your instance when not in use to conserve your credits.**

## Deploy Your VM

Watch the video tutorial to learn how to deploy your VM in Google Cloud. Separate videos are available for Mac and PC users, so please choose the one that applies to you. The video links below are the SAME videos on blackboard.

* **PC Users:** [Watch the walkthrough here](https://youtu.be/B67pJBVTsLI)
* **Mac Users:** [Watch the walkthrough here](https://youtu.be/qkVhO9gtQ8o)

Starting this week, all work will be done on the Google VM, with your PC or Mac serving only as a client to connect to the VM.

**PC Users:** In future assignment walkthroughs, we’ll be using the Mac terminal while logged into the Google VM. You'll connect to the VM using PuTTY, but the commands will be identical for both systems. The only difference is that PuTTY handles the SSH tunnel for you, so you won't need to open it manually from the command line. Other than that, both PC and Mac users will operate as clients, with all work being completed on the Google VM.

**Note:** Over time, Ubuntu releases will change. Always start with the latest minimal install. Some minimal images do not include Git by default, so you’ll need to install it manually with:

sudo apt update

sudo apt install git -y

## Accessing User Interfaces

Open your tunnels:

* **PC Users:** Use the saved tunnel session for your tunnels. You are required to download [Putty](https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html) and [PuttyGen](https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html) from the official website
* **Mac Users:** Use the following command:

ssh -L 9870:localhost:9870 -L 8088:localhost:8088 -L 8080:localhost:8080 -L 18080:localhost:18080 -L 16010:localhost:16010 -L 8983:localhost:8983 -L 8443:localhost:8443 **USERNAME@EXTERNAL\_IP**

Replace **USERNAME** with your SSH user as shown in the video.  
Replace **EXTERNAL\_IP** with your Google VM external IP address.

* HDFS: <http://localhost:9870>
* YARN: <http://localhost:8088>
* Spark Master: <http://localhost:8080>
* Spark History: <http://localhost:18080>
* HBase: <http://localhost:16010>
* Solr: <http://localhost:8983>
* NiFi: <https://localhost:8443/nifi>

## Software Port List (For Putty SSH Tunnels)

9870, 8088, 8080, 18080, 16010, 8983, 8443

## Key Commands for Copy + Paste

* Clone the Repository:  
  git clone <https://github.com/bellevue-university/dsc650-infra.git>
* Run the Setup Script:  
  sudo ./setup.sh
* Start the Docker Containers:  
  docker-compose up -d
* Stop the Docker Containers:  
  docker-compose down
* Start NiFi:  
  ./nifi.sh start
* Stop NiFi:  
  ./nifi.sh stop

**Remember, the Google Cloud free tier credits expire after 90 days or when they are all used. Always monitor your usage to avoid unexpected charges. Be sure to stop your instance when not in use to conserve your credits.**