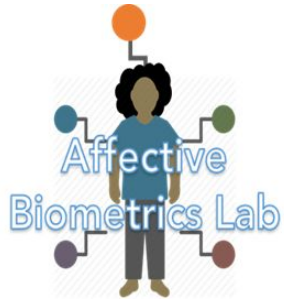
Leveraging the Cloud for Teaching Applied Data Science at HBCUs

Gloria Washington, PhD
Assistant Professor
Howard University
Director Affective Biometrics Lab
gloria.washington@howard.edu

HOWARD UNIVERSITY
1867



About Me: **Gloria Washington**

Howard University
Assistant Professor, Computer Science
Director of Affective Biometrics Lab
BS | Lincoln University
MS & PhD | George Washington University

Areas of expertise

- Human-Centered Computing
- Computer Science Education
- Biometrics

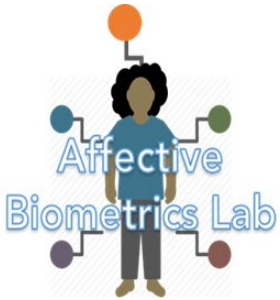
Personal

- I want to help Black Women get PhDs in CS
- Love undergraduate research ...is that a crime?

Affective Biometrics Lab Mission & Research

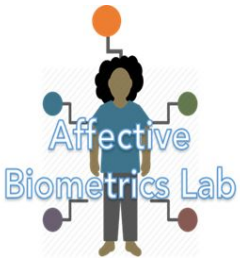
Give voices to individuals or communities that feel marginalized through AI that leverages human physical, physiological, or behavioral characteristics for identity or emotion recognition

1. *Who is it that?*
2. *What do they feel?*
3. *How are they going to react or behave?*
4. *What can I build or develop to positively impact humans' feelings?*



Howard University's Computer Science Course on Applied Data Science

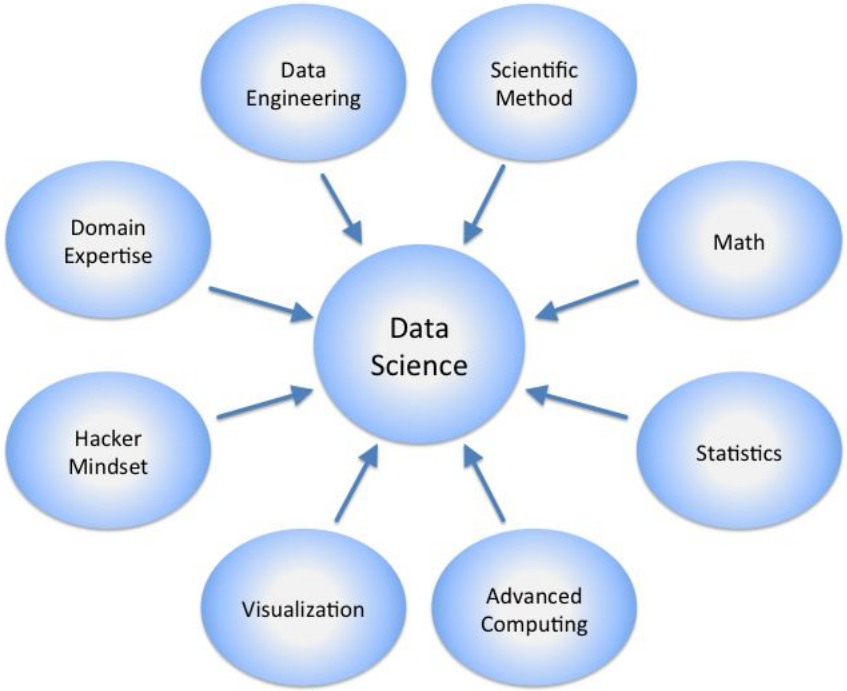
Hands-on course that mixes statistics, analytics, and data visualization techniques for students to obtain entry-level positions in data science



HOWARD UNIVERSITY

Major Topics Covered

1. How to Be A Good Data Scientist
2. Ethics, Fairness, Accountability and Transparency in Data Science
3. Python Programming Using Jupyterlab Notebooks
4. Foundations of Probability & Statistics
5. Regression and Prediction
6. Classification



The Littlest Jupyterhub

Cloud-based server that lets students login and use Jupyter Notebooks instead of downloading complicated software onto their laptops

HOWARD UNIVERSITY

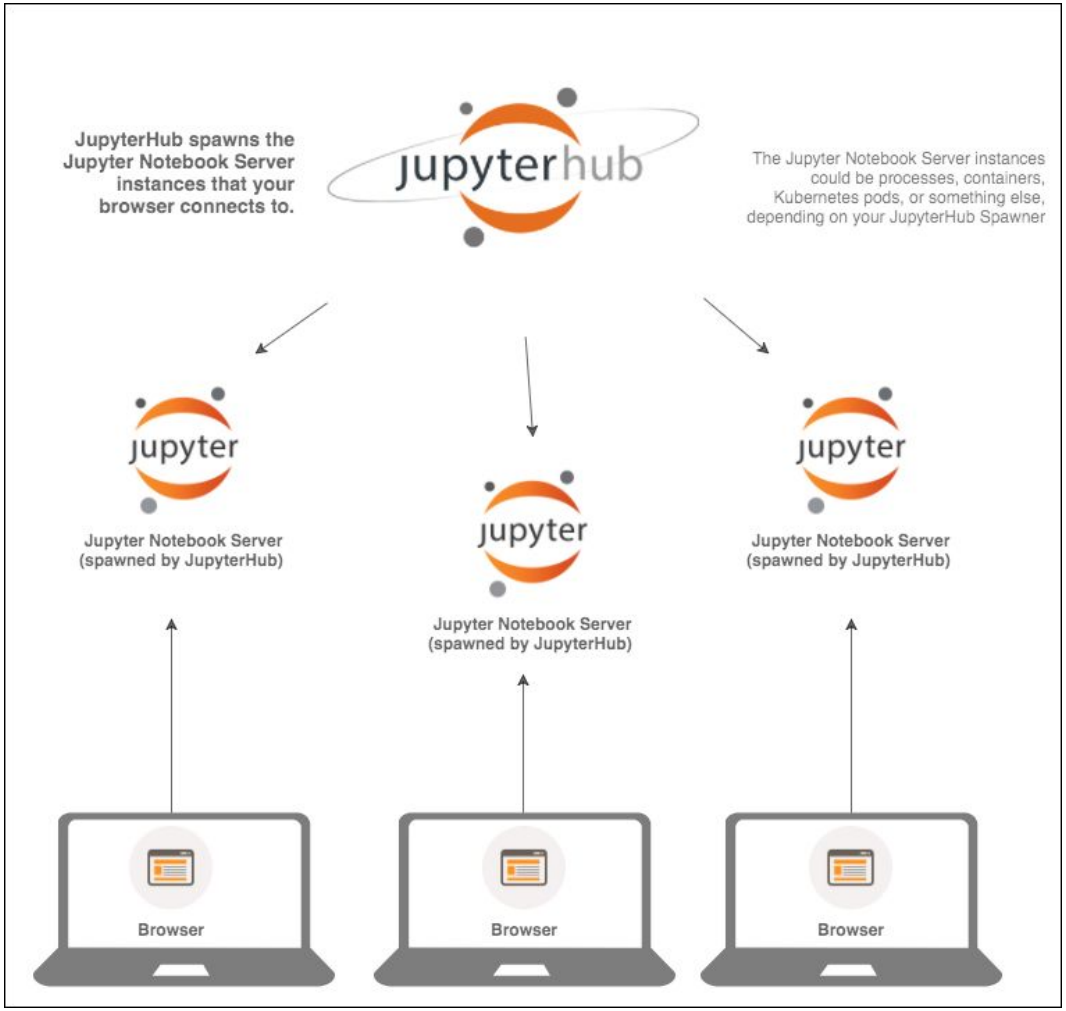
Affective Biometrics Lab

The Littlest Jupyterhub

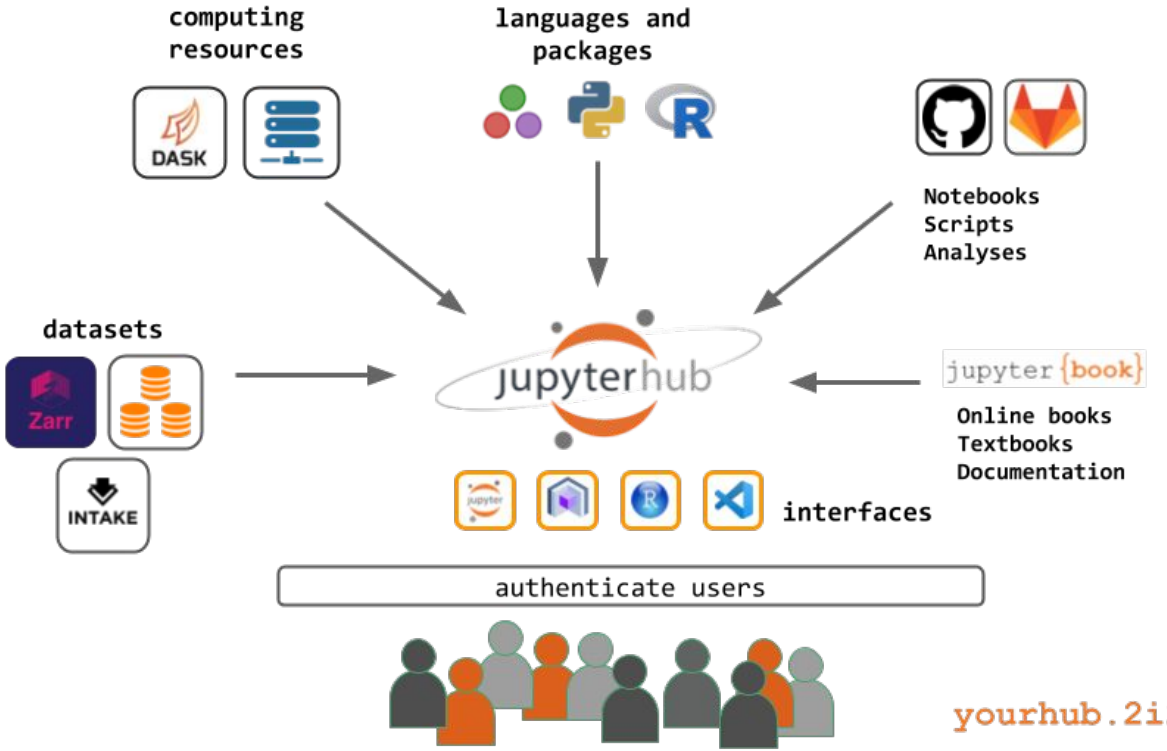
Cloud-based server that lets students login and use Jupyter Notebooks instead of downloading complicated software onto their laptops

Uses Microsoft Azure back-end

Sponsored by Microsoft Data Science Fellows Grant



Partnership with 2i2c Hub



yourhub.2i2c.cloud

DATA 8

CloudBank

Affective Biometrics Lab

HOWARD UNIVERSITY

Python programming in Jupyter Notebooks

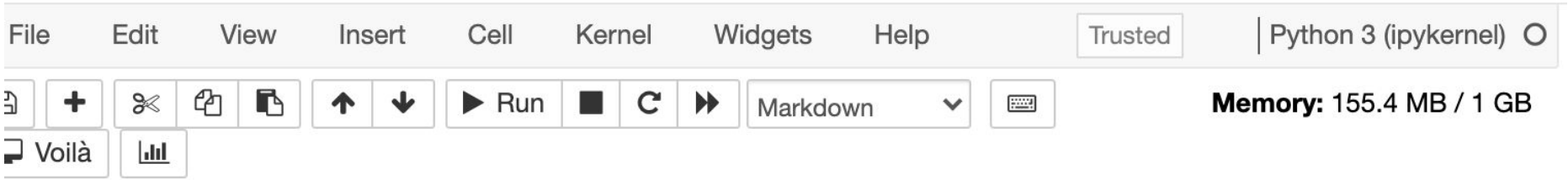
Dataset Manipulation

Table Operations

Math & Stat Operations

Demo:
https://howard.cloudbank.2i2c.cloud/

Back-up for Demo


File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Trusted
Python 3 (ipykernel)
Run
Markdown
Memory: 155.4 MB / 1 GB
Voilà

Lab 2: Pandas Overview

To receive credit for a lab, answer all questions correctly and submit before the deadline.

This lab is due Monday, September 13th at 11:59 PM.

Pandas is one of the most widely used Python libraries in data science. In this lab, you will learn commonly used data wrangling operations/tools in Pandas. We aim to give you familiarity with:

- Creating dataframes
- Slicing data frames (i.e. selecting rows and columns)
- Filtering data (using boolean arrays)

In this lab you are going to use several pandas methods, such as `drop` and `loc`. You may press `shift+tab` on the method parameters to see the documentation for that method. If you are familar with the `datascience` library used in Data 8, this conversion notebook may serve as a useful guide.

Note: The Pandas interface is notoriously confusing, and the documentation is not consistently great. Throughout the semester, you will have to search through Pandas documentation and experiment, but remember it is part of the learning experience and will help shape you as a data scientist!

In [1]:
```
1 import numpy as np
```

Questions?

Gloria Washington, Ph.D.
Assistant Professor of Computer Science
gloria.washington@howard.edu
https://sites.google.com/view/affective-biometrics-lab-resea

HOWARD UNIVERSITY

Affective Biometrics Lab