

# Programming and Database Fundamentals for Data Scientists

## Introduction

Varun Chandola

School of Engineering and Applied Sciences  
State University of New York at Buffalo  
Buffalo, NY, USA  
chandola@buffalo.edu



Introduction

Need for Computer Science in Data Science

# Starting from the beginning

- ▶ What is a *computer*?

# Starting from the beginning

- ▶ What is a *computer*?
  - ▶ Machine that **stores** and **manipulates** **information** under the control of a **changeable program**.
- ▶ What is a *computer program*?

# Starting from the beginning

- ▶ What is a *computer*?
  - ▶ Machine that **stores** and **manipulates information** under the control of a **changeable program**.
- ▶ What is a *computer program*?
  - ▶ A **detailed, step-by-step** set of instructions telling the computer **exactly** what to do.

# What is Computer Science?

- ▶ Computer Science  $\neq$  Study of Computers!!!

# What is Computer Science?

- ▶ Computer Science  $\neq$  Study of Computers!!!
- ▶ **How can we make computers to do what we want.**
- ▶ **What can be computed?**

# The Holy Trinity of Computer Science

## Design

- ▶ Write algorithms to solve problems

## Analyze

- ▶ Is a problem *solvable* (efficiently)?

## Experiment

- ▶ Implement algorithms and assess performance



# Why CS in Data Science?

- ▶ Store and access data
- ▶ Manipulate data
- ▶ Get insights from data
- ▶ Report/visualize results

# References