UNIT1: INTRODUCTION TO DATA AND DATABASES

CUNY LAGUARDIA COMMUNITY COLLEGE

INTRODUCTIONS

- This course will introduce you to a few main topics:
  - SQL
  - Fundamentals of Python Programming
  - Data Analytics in Python
  - Data visualization and communication.

INTRODUCTIONS 20 MINS

- Introduce yourself to your partners/neighbors.
- Guiding questions to speak about with your partner:
  - What does data mean to you?
  - How can data help your organization?
  - Current roles

OBJECTIVES

- By the end of this course you should be able to:
  - Query data from structured databases using SQL.
  - Load data into Python from multiple data sources, including CSV files, Excel files, and relational databases.
  - Perform data cleaning and transformations in python using various libraries, including: NumPy and pandas
  - Use Python to learn key statistical methods for analyzing data - including: Exploratory data analysis and regression analysis.
  - Visualize, present, and share their insights to their classmates, instructor, and employers, with a capstone project.

UNIT 1 - RELATIONAL DATABASES AND DATA EXTRACTION

- Students will gain an overview of databases and SQL using PostgreSQL. The skills learned during this unit are applicable to any other major SQL database, such as Microsoft SQL Server, MySQL, Oracle and others. SQL is one of the most in demand tech skills – being able to extract data from your organization allows you to manipulate, transform, visualize and answer business related questions.
- Some topics this unit will cover:
  - What is data? How does it relate to business?
  - SQL Statement fundamentals
  - Aggregate functions
  - JOINS
  - Sub Queries
  - Math Functions
  - Creating Databases and Tables

UNIT 1 - RELATIONAL DATABASES AND DATA EXTRACTION

- The point:
  - Introduction to databases.
  - Unit 1 will introduce you to the language of databases – SQL.
  - SQL is used to gather and extract data from databases.

UNIT 2: PYTHON FOR DATA ANALYTICS

- Students will begin to explore Python 3. Python is high level, general purpose programming language – used in application and web development, data analysis, artificial intelligence, and more. This until will expose students to the fundamentals of using Python – setting the stage for more advanced data analytics topics. Some topics this unit will cover:
- Variable assignments
  - Lists
  - Dictionaries.
  - Looping
  - Functions
  - Built-in Functions

UNIT 2: PYTHON FOR DATA ANALYTICS

- Python is one of the most up-to-date, popular, and used programming languages used today.
- It's especially useful for data analytics and data science.

UNIT 3: DATA ANALYTICS & DESCRIPTIVE STATISTICS IN PYTHON

- Building on the topics learned in the previous units, students will dive deeper into data analysis and statistics using Python. Students will be importing data from multiples sources, including SQL databases and excel and csv files. Students will explore real-world business cases using popular and relevant data analytics libraries.
- Some topics this unit will cover:
  - NumPy
  - Pandas
  - Matplotlib and Seaborn
  - Hypothesis Testing
  - Regression

UNIT 3: DATA ANALYTICS & DESCRIPTIVE STATISTICS IN PYTHON

- Statistics allows for organizations to make meaningful and impactful decisions.
- Organizations have tons of data – that data needs to be broken down into information that administration and other users can benefit from.

UNIT 4: DATA VISUALIZATION & COMMUNICATION

- Organizations use dashboards to keep track of their day-to-day, and long-term business objectives – from keeping track of patient admissions in an emergency department, to tracking spend vs. budgets. Students will learn how to create visualizations that tell meaningful stories that viewers and users can clearly identify and understand to help keep track of business objectives.
- Some topics this unit will cover:
  - Data-Ink Ratio
  - ChartJunk
  - Small Multiples
  - Multifunctioning Graph Elements
  - Aesthetics and technique
  - Tools tips
  - Exploratory Data Analysis in Tableau

UNIT 4: DATA VISUALIZATION & COMMUNICATION

- After analysis has been made on data, that data needs to be communicated to others.
- Communication can be in the form of:
  - Reports
  - PowerPoints
  - Dashboards
  - Visualizations

ARE YOU FAMILIAR WITH ANY OF THIS INFORMATION?

- With your table/partner:
- Discuss what information you're familiar with, or have used in the past.
- What are some topics you're interested in learning about?

WHAT IS DATA USED FOR?

In todays technology driven age, data is all around us.

Data is used to recommend products to users (amazon, ads on websites, etc.)

Tailor your experiences on streaming websites (Netflix, hulu, etc.)

Track your location and offer suggestions on travel, restaurants, points of interest.

Forecast spending/savings for an organization.

Predict flu outbreaks before they happen.

Determine how many staff are needed to fill a shift

And way more…

WHY DO WE CARE ABOUT ANALYTICS?

Generate reports for management, administration, physicians, etc.

Gather hidden insights – find patterns to your customers habits, etc.

Improve business practices.

Guide strategy and planning; short and long term goals.

HOW IS THIS COURSE RELEVANT?

- The ability to gather, prepare, and analyze data is one of the most important, and most sought-after skills in the information technology and analytics marketplace. Skills and projects completed during the course will relate directly to actual every-day job functions in the data analytics field and are designed around employer needs. The skills learned during the course will allow students to be better prepared for roles such as:
  - Data Analyst
  - Business Analyst
  - Data Scientist
  - Systems Analyst

WHERE ARE DATA ANALYTICS PROFESSIONALS FOUND?

- Data analytics positions can be found in every sector – such as: healthcare, finance, retail, and even gaming!
  - Data analytics professionals can be found working at companies such as: Facebook, Twitter, Google, ConEdison, Northwell Health, NYC Agencies, and more.