**CODE 3 CAMP**

**Data science** is a rapidly developing field that involves computational principles, methods, and systems for extracting and structuring insights from data.

On a daily basis, large datasets are generated by activities in the sciences, finance society, and commerce. Data scientists are constantly seeking patterns and predicting outcomes from these vast collections of data.

This enables them to extract insights from business data to facilitate business decision-making. In this boot camp mathematical content is combined with principles of statistics, computer science, and software engineering to give all the finalists a steppingstone in a world-class career in data science.

We also believe that the program is designed with sufficient technical depth to equip finalists with the ability to develop novel analytical tools for new scientific applications and industry problems that will emerge in the future.

**Weeks 1: Program Orientation and Data Science Foundation.**

* **Orientation:** Introduction to the Program, whole program expectations, and rules and regulations.
* **Introduction to data science,** overview of data science, its applications, possible data science career path, and the data science lifecycle.
* **Setting up environments** for tools and technologies which we will be using throughout the program.
  + For the Python we will use VS Code, Anaconda, Jupyter Notebook, and Kaggle Notebook.
  + We will also set up mysql workbench and SQL Server for SQl MongoDB and Mongo Atlast for NoSQL.
  + Also we can set a SQL cleint like dbeaver.

**Week 2: Mathematics, Probability, and Statistics Part 1.**

* **Mathematics, Probability, and Statistics for data science.** This will include probability and sampling, mathematical analysis, A/B testing, Increasing test sensitivity, linear algebra and differential calculus.

**Week 3: Mathematics, Probability, and Statistics Part 2.**

* **Mathematics, Probability, and Statistics for data science.** This will include probability and sampling, mathematical analysis, A/B testing, Increasing test sensitivity, linear algebra and differential calculus.

**Weeks 4: Python for Data Science**

* **Introduction to Python programming** for data science, including data types, variables, basic operations, data structures (lists, dictionaries, etc.), control flow, conditional statements, and functions.
* **Advanced Python concepts** including OOPS and how to build Python packages, and modules.
* **Python Libraries** for data science: Pandas, NumPy, Matplotlib, seaborn and Scikit-Learn

**Week 5: SQL for Data Science**

* **Master the Essentials:** Build your SQL foundation with data structures, basic queries, filtering techniques, and functions to unlock your data's potential.
* **Unleash Data Power:** Learn how to join and manipulate data from multiple tables, group it for deeper insights, and unleash the hidden relationships within your information.
* **Solve Real-World Challenges:** Put your SQL skills to the test through case studies and a hands-on project, applying your knowledge to practical data science scenarios and gaining valuable experience.

**Week 6: Data Manipulation and Cleaning**

* Explore data science and analysis python packages, and how to use them on different data science processes (pandas, numpy, matplotlib, seaborn, beautiful soup, selenium, and MySQL-python.
* Data Collection techniques, data cleaning, handling missing values, data preprocessing and analysis.

**Week 7: Exploratory Data Analysis (EDA)**

* Techniques for data visualization using libraries like Matplotlib and Seaborn.
* Statistical analysis, including measures of central tendency and data distribution.

**Week 8: Data Wrangling and Feature Engineering**

* Data wrangling methods, including filtering, merging, and reshaping data.
* Feature engineering: creating meaningful features from raw data.

**Week 9: Machine Learning Fundamentals**

* Introduction to supervised and unsupervised learning.
* Linear regression, logistic regression, and model evaluation metrics.

**Weeks 10: Advanced Machine Learning**

* Decision trees, random forests, and Gradient Boosting Decision Trees
* Clustering techniques (K-means, hierarchical clustering) and dimensionality reduction (PCA).

**Weeks 11 - 13: Deep Learning and Neural Networks**

* Introduction to neural networks and deep learning.
* Building and training neural networks with Pytorch.
* Introduction to special topics such as natural language processing and computer vision using Fastai

**Weeks 14 - 15: Introduction to Generative AI**

* Introduction to LLMS: their use cases and limitations
* Introduction to Retrieval Augmented Generation

**Capstone Project Kickoff**