

PROFESSIONAL TRAINING

ACADEMIC YEAR: 2021-2023

DEPARTMENT: MSC BIOINFORMATICS AND DATA SCIENCE

(2021-2023)

COURSE TITLE: PROFESSIONAL TRAINING

SEMESTER: 2



COURSE COORDINATOR

DR.DANIEL ALEX ANAND

ASSOCIATE PROFESSOR

DEPARTMENT OF BIOINFORMATICS

SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY

CHENNAI-600119

SUBMITTED BY

ASWINY AR

MSC. BIOINFORMATICS AND DATASCIENCE-2020-2022

SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY

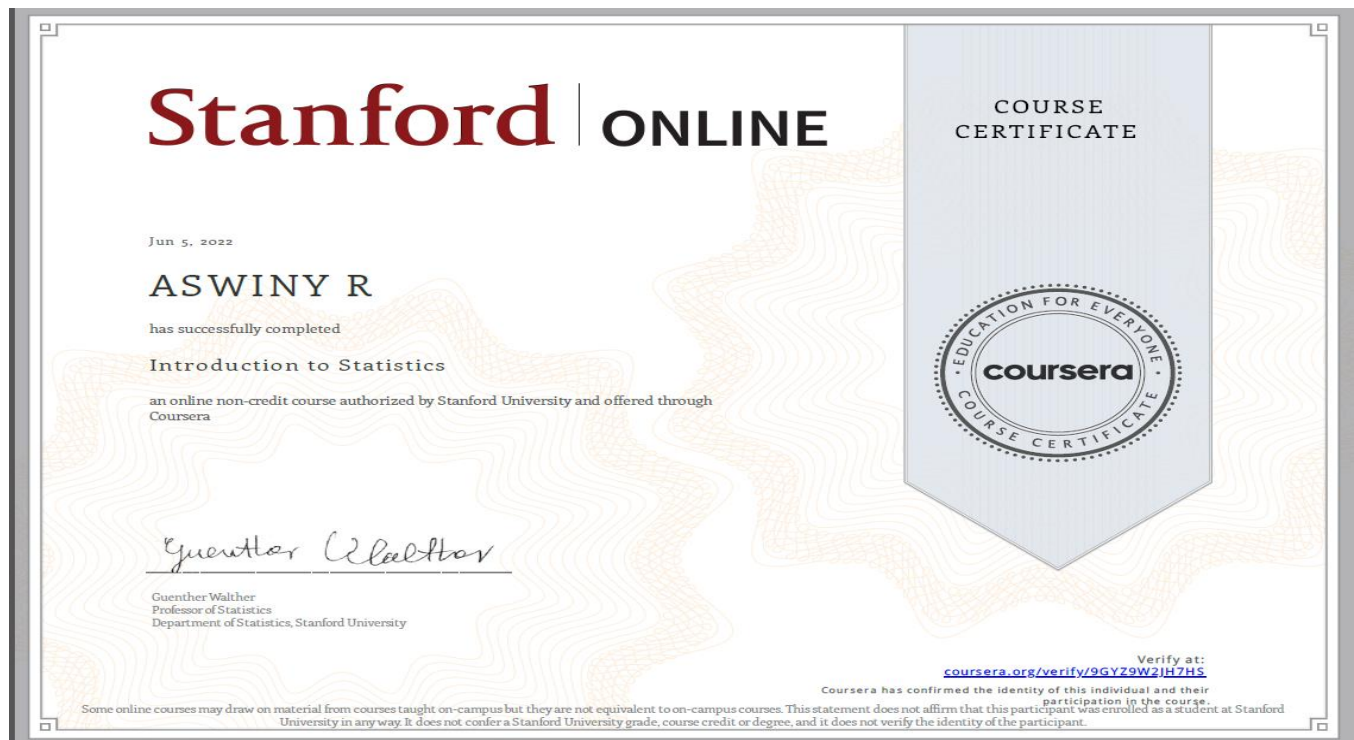
INTRODUCTION TO STATISTICS

OBJECTIVES:

In this course, I am able to perform exploratory data analysis, understand key principles of sampling, and select appropriate tests of significance for multiple contexts. I have gained the foundational skills that prepare me to pursue more advanced topics in statistical thinking and machine learning.

SKILLS

- Descriptive Statistics
- Sampling and Randomized Controlled Experiments
- Probability
- Sampling Distributions and the Central Limit Theorem
- Regression
- Common Tests of Significance
- Resampling, Multiple Comparisons



MACHINE LEARNING WITH PYTHON

OBJECTIVE

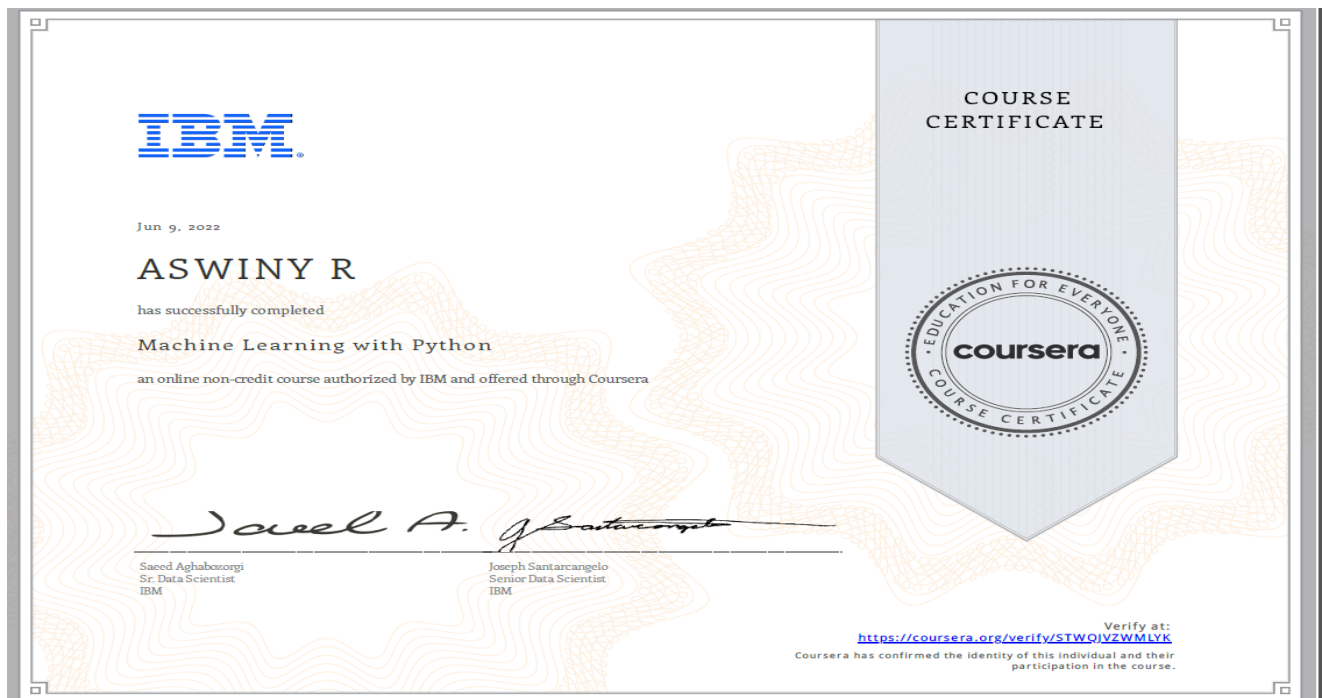
This course dives into the basics of machine learning using an approachable, and well-known programming language, Python. In this course, I have been reviewing two main components: First, I have learnt about the purpose of Machine Learning and where it applies to the real world. Second, I get a general overview of Machine Learning topics such as supervised vs unsupervised learning, model evaluation, and Machine Learning algorithms.

- Give examples of Machine Learning in various industries.
- Outline the steps machine learning uses to solve problems.
- Provide examples of various techniques used in machine learning.
- Describe the Python libraries for Machine Learning

SKILLS

- Python libraries
- Machine learning
- Regression
- Hierarchical clustering

- K-Means clustering



SQL FOR DATA SCIENCE

OBJECTIVE:

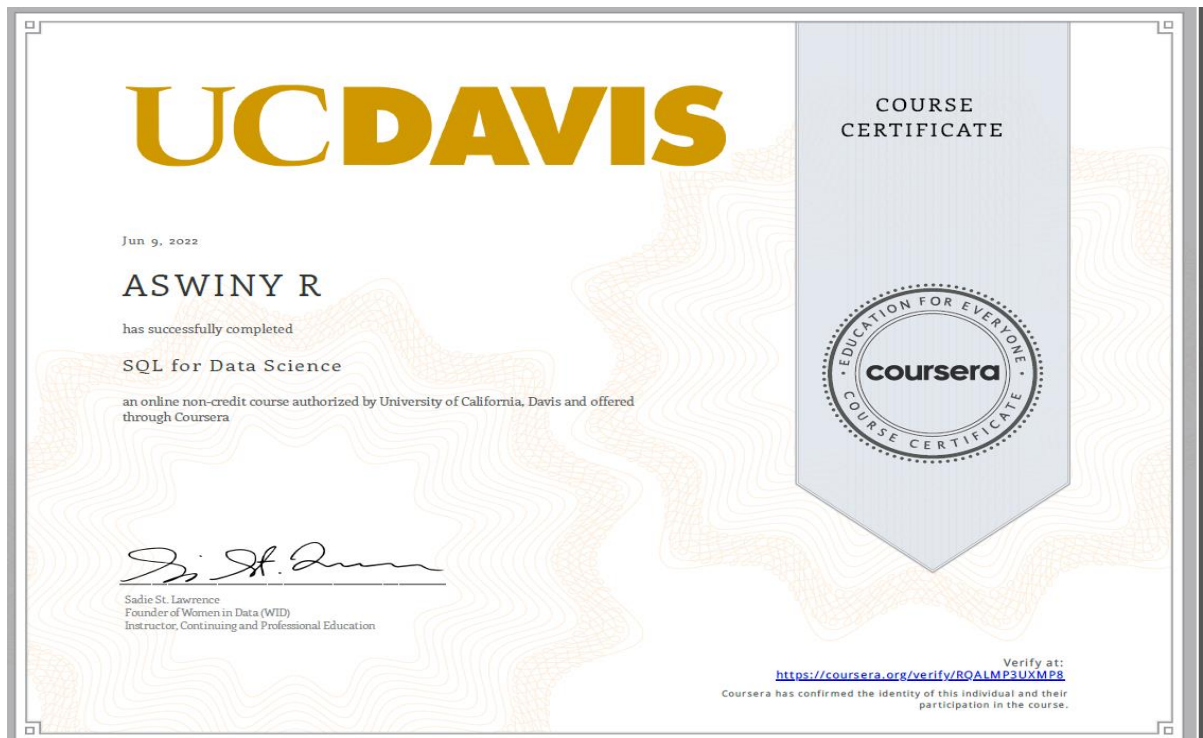
This course is designed to give me a primer in the fundamentals of SQL and working with data so that I can begin analyzing it for data science purposes. This course starts with the basics and assumes you do not have any knowledge or skills in SQL. It will build on that foundation and gradually have write both simple and complex queries to help me select data from tables. I started to work with different types of data like strings and numbers and discuss methods to filter and pare down your results.

- Identify a subset of data needed from a column or set of columns and write a SQL query to limit to those results
- Use SQL commands to filter, sort, and summarize data.
- Create an analysis table from multiple queries using the UNION operator
- Manipulate strings, dates, & numeric data using functions to integrate data from different sources into fields with the correct format for analysis.

SKILLS

- Data science
- Data analysis

- SQL



DRUG DISCOVERY

OBJECTIVE

In this course I learn the drug discovery process up to the filing of an Initial New Drug Application. Each week I learn the steps that a pharmaceutical or biotech company goes through to discover a new therapeutic drug. In this course you will be able to:

- Understand the pharmaceutical and biotechnology market a changing landscape
- Learn the major aspects of the drug discovery process, starting with target selection, to compound screening to designing lead candidates.
- Recognize current modern drug discovery based on the lock-and-key theory, which attempts to use one single compound to hit one target to combat the related disease.
- Increase understanding of the various drug discovery tools and methods that are used for finding, identifying and designing a new drug.
- Define and understand the regulatory responsibilities for drug discovery to file an Investigational New Drug Application

SKILLS:

- Pharmacology
- Drug Development

- Drug Development
- Pharmacokinetics



DATA VISUALIZATION

OBJECTIVES:

Data visualization is the practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from. The main goal of data visualization is to make it easier to identify patterns, trends and outliers in large data sets.

SKILLS

- Line chart
- Bar charts and heatmaps
- Scatter plots
- Distribution
- Choose plot types and custom styles



PYTHON

OBJECTIVE

Python is a computer programming language often used to build websites and software, automate tasks, and conduct data analysis. Python is a general-purpose language, meaning it can be used to create a variety of different programs and isn't specialized for any specific problems.

SKILLS

- Function and getting help
- Booleans and conditionals
- Lists
- Loops and list comprehensions
- Strings and dictionaries
- Working with external libraries



INTRO TO MACHINE LEARNING

OBJECTIVE

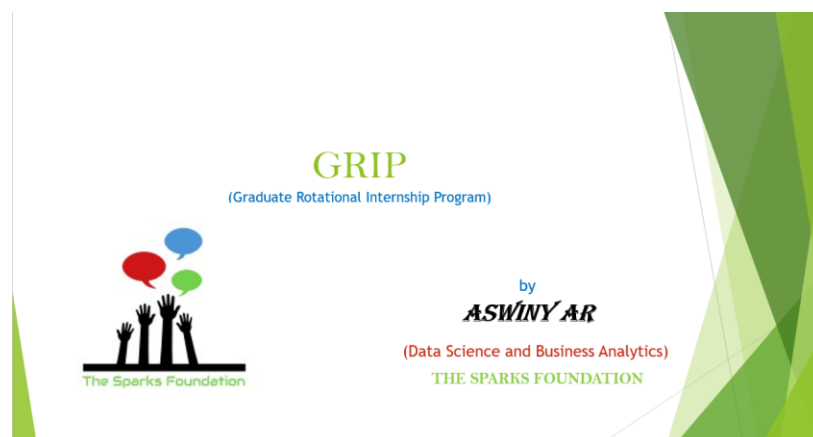
Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy

SKILLS

- How models work
- Basic data exploration
- First machine learning model
- Model validation
- Underfitting and overfitting
- Random forests
- Machine learning competitions



INTERNSHIP



About GRIP at The Sparks Foundation

The Sparks Foundation (TSF) is a non-profit organization registered in India and Singapore. We envision a world of enabled and connected little minds, building the future. We aim to inspire students, help them innovate, and let them integrate to build the next generation of humankind. We help the students to integrate and help each other, learn from each other, and do well together.

The Graduate Rotational Internship Program (GRIP) is the flagship program of TSF in which students, recent graduates, and professionals focus on technical skills development as well as professional profile improvement on LinkedIn. The program gives young aspiring minds a learning platform to connect with students and professionals from varied diversity, background, skills, and countries. During the Covid-19 pandemic, the format of GRIP is a 1-month, unpaid, and virtual internship.

SKILLS

- Python
- Google Collab
- Machine learning
- Exploratory data analysis
- Data Visualization

SELECTION CERTIFICATE



THE SPARKS FOUNDATION



THIS IS PRESENTED TO

ASWINY AR

for successful selection as an intern at The Sparks Foundation for function Data Science & Business Analytics.



PRANAV DUBEY
DIRECTOR

01/25/2022

DATE



CODE : B8U7SVK4AD
Verify at:
<https://truecertificates.com/verification>

COMPLETION CERTIFICATE



I have successfully completed my role in this organization, by completing and submitting the tasks given before the timeline and did peer-review for my fellow peers.

THANK YOU

