

What I have learnt?

- **Mastered Practical Skills and Knowledge:**

Acquired a comprehensive understanding of the most up-to-date practical skills and knowledge essential for data scientists in their daily roles

- **Proficiency in Tools, Languages, and Libraries:**

Developed expertise in using the tools, languages, and libraries commonly employed by professional data scientists, with a specific focus on Python and SQL.

- **Data Importation and Cleaning:**

Learned the techniques and best practices for importing and cleaning diverse data sets, ensuring the ability to work with a wide range of data sources.

- **Data Analysis and Visualization:**

Gained the skills to effectively analyze and visualize data, employing various methods to extract meaningful insights from complex datasets.

MODULES

- 1.What is Data Science?
- 2.Tools for Data Science
- 3.Data Science Methodology
- 4.Python for Data Science, AI and Development
- 5.Python Project for Data Science
- 6.Databases and SQL for Data Science
- 7.Data Analysis with Python
- 8.Data Visualization with Python
- 9.Machine Learning with Python
- 10.Applied Data Science Capstone



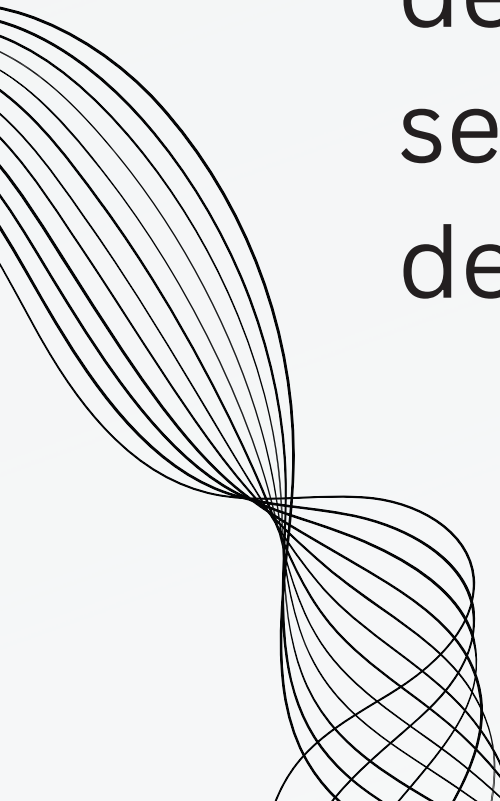
What is Data Science?



Data science is the key to unlocking insights from the vast volumes of data generated in our digital age. It involves collecting, cleaning, analyzing, and interpreting data using advanced algorithms and machine learning models.

Importance of Data Scientists:

Data scientists are pivotal in harnessing the power of data for decision-making, predictive analytics, innovation, personalization, and security. Their expertise ensures organizations make informed decisions, anticipate trends, and maintain the integrity of systems.



Tools for Data Science



1. Programming Languages:

- Python
- R

2. Data Manipulation and Analysis:

- Pandas
- NumPy
- SQL

3. Data Visualization:

- Matplotlib
- Seaborn
- Plotly

4. Statistical Analysis:

- RStudio

5. Machine Learning:

- Scikit-Learn
- TensorFlow
- PyTorch

6. Big Data Processing:

- Apache Spark
- Hadoop

7. Version Control:

- Git

8. Data Storage:

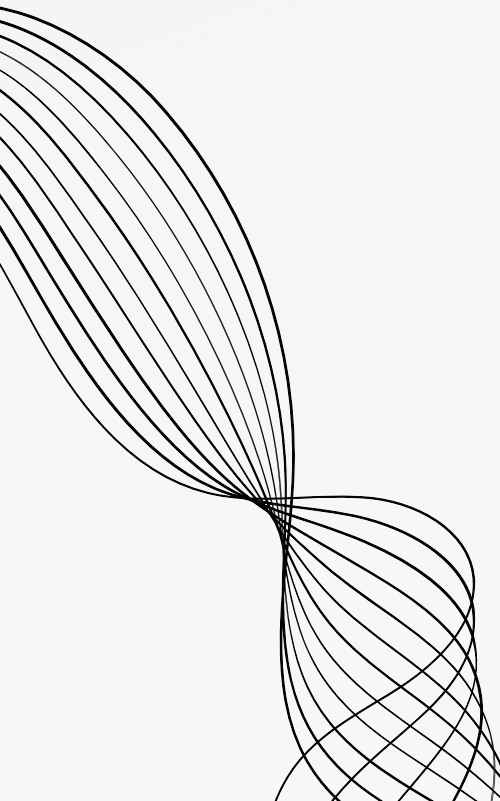
- MySQL
- PostgreSQL
- MongoDB

Data Science Methodology



I've gained a solid grasp of data science methodology, understanding its vital role in guiding analytical processes. Through hands-on application of the CRISP-DM methodology, I navigated six stages with a focus on real-world case studies. This experience enhanced my technical skills and instilled a disciplined problem-solving approach.

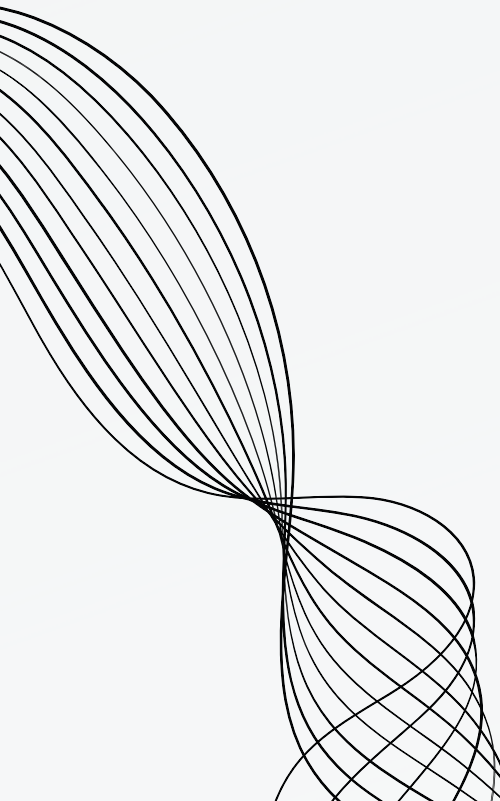
I now possess the ability to judiciously evaluate and select analytic models, be it predictive, descriptive, or classification, aligning them with project goals. Additionally, I've honed my skill in identifying and leveraging suitable data sources, crucial for impactful data science projects.



Python for Data Science, AI & Development



I've mastered Python fundamentals, including data types, expressions, variables, and structures. Proficient in Python logic, I've applied it to key concepts like branching, loops, functions, and object-oriented principles. My expertise extends to vital libraries—Pandas for data manipulation, Numpy for numerical operations, and BeautifulSoup for web scraping. I've honed skills in accessing web data through APIs and scraping in Python within Jupyter Notebooks. This course provides a solid Python foundation, hands-on experience with critical libraries, and the ability to confidently apply these skills to real-world data tasks.

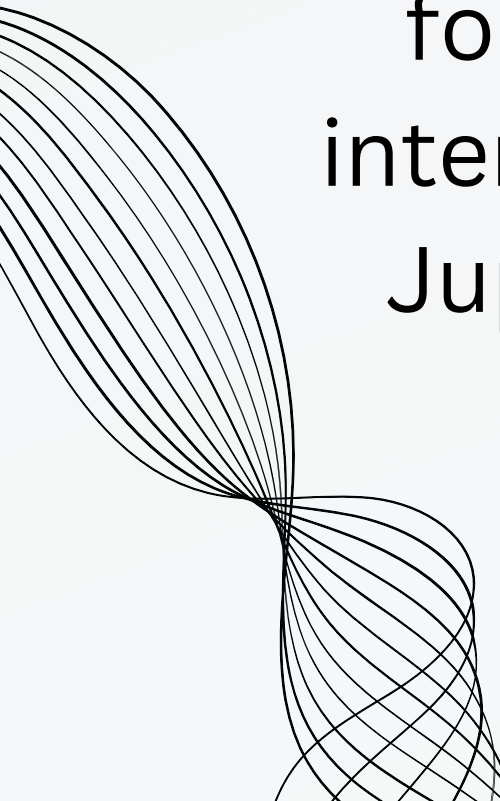


Project for Data Science



I've assumed the role of a Data Scientist/Data Analyst engaged in a hands-on, real-world project. My proficiency in Python, the language synonymous with Data Science and Analysis, has significantly evolved. I've successfully applied Python fundamentals, mastered essential data structures, and honed my skills in handling data with Python.

A key highlight of my learning journey includes the construction of a dynamic dashboard using Python, employing crucial libraries such as Pandas for data manipulation, BeautifulSoup for web scraping, and Plotly for interactive visualizations. This project-based approach, executed within a Jupyter notebook, has not only solidified my Python skills but has also equipped me to create impactful data-driven dashboards.



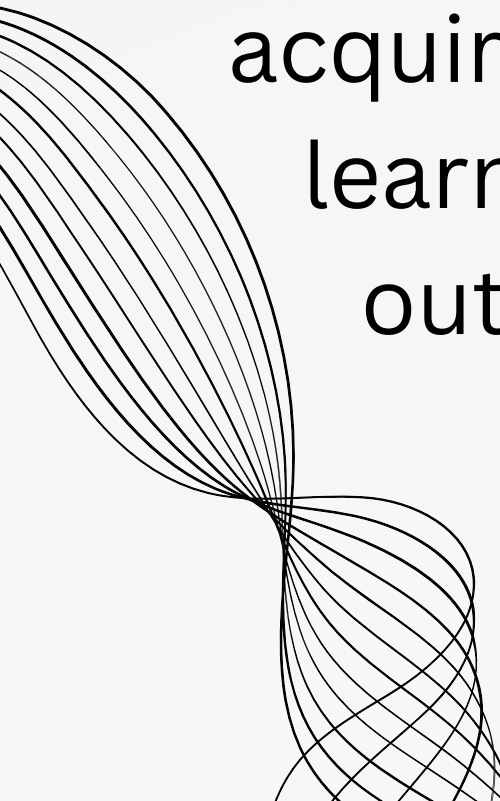
Databases and SQL for Data Science

I've acquired a comprehensive skill set in data analysis by seamlessly integrating SQL and Python. I can proficiently create and manage relational databases, utilizing DDL commands to work with multiple tables. My ability to construct SQL queries has progressed from basic to intermediate levels, encompassing DML commands. Importantly, I've gained mastery in advanced SQL techniques, including the creation of views, implementation of transactions, writing stored procedures, and executing complex joins. This proficiency enables me to wield powerful tools for extracting valuable insights and conducting sophisticated data analysis tasks.

Data Analysis



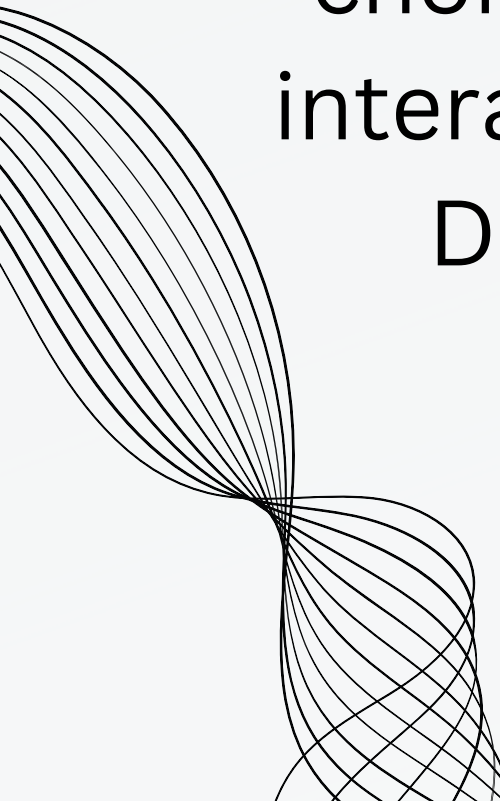
I've honed my Python coding skills, specifically in the context of cleaning and preparing data for analysis. This includes adeptly handling missing values, formatting, normalizing, and creating data bins to ensure data quality. Utilizing libraries such as Pandas, Numpy, and Scipy, I've seamlessly performed exploratory data analysis on real-world datasets. I've become proficient in manipulating data through dataframes, summarizing key insights, understanding data distributions, and implementing correlation analyses. Furthermore, I've acquired the ability to construct and evaluate regression models using the scikit-learn library in machine learning. These skills empower me to not only predict outcomes but also make informed decisions based on comprehensive data analysis.



Data Visualization



I've acquired a robust skill set in implementing compelling visualizations using Python libraries like Matplotlib, Seaborn, and Folium. I can effectively tell data-driven stories through diverse chart types, including line, area, histograms, bar, pie, box, scatter, and bubble plots. Additionally, I've expanded my capabilities to create advanced visualizations, such as waffle charts, word clouds, regression plots, and intricate maps with markers and choropleth representations. The course has also equipped me to construct interactive dashboards, integrating various chart types seamlessly, using the Dash framework and the Plotly library. This proficiency enables me to present data insights in a visually engaging and impactful manner.



Machine Learning

I've developed a comprehensive understanding of various machine learning algorithms and learned to discern their optimal applications. Specifically, I can now effectively compare and contrast linear classification methods such as multiclass prediction, support vector machines, and logistic regression. My proficiency extends to practical implementation, as I can write Python code to apply diverse classification techniques, including K-Nearest Neighbors (KNN), decision trees, and regression trees.

Furthermore, I've gained the ability to evaluate the results of linear, non-linear, and multiple regression on datasets, employing relevant evaluation metrics. This course has not only equipped me with essential skills in SciPy and scikit-learn but also enriched my expertise in machine learning, regression, and classification, providing a solid foundation for data-driven decision-making.