

Dana Dahrouj

Junior Data Scientist

✉ dahroujdana6@gmail.com ☎ +961 76 068 403

Education

Master of Science in Statistics, <i>Lebanese University</i>	10/2023 – 8/2025
Bachelor of Science in Statistics, <i>Lebanese University</i>	10/2020 – 8/2023

Professional Experience

Data Science Internship, <i>Maaloomatiia</i>	7/2024 – 9/2024
---	-----------------

- Gained hands-on experience with Dataiku for data science.
- Worked on real-world use cases, from data collection to model deployment, ensuring the project met the objectives.
- Conducted extensive data preprocessing, including handling missing values, feature engineering, and data scaling, to prepare datasets for the machine learning model.
- Developed and fine-tuned machine learning models to predict outcomes.
- Communicated findings, and presented results.
- Generated insights and reports from analyzed data, providing valuable information for decision-making processes.

Data Science Internship, <i>Salem Group</i>	8/2023 – 10/2023
--	------------------

- Scraped job openings from various websites.
- Engaged in data cleaning tasks using both Excel and Python, ensuring that datasets were free from inconsistencies, missing values, and errors.
- Conducted data analysis tasks using Python, leveraging libraries such as Pandas, NumPy, Matplotlib, and Seaborn for in-depth exploration and visualization of datasets.

Projects

British Airways Data Science Program

- Performed web scraping to gain company insights.
- Predicted customer buying behavior.

Fashion Recommendation System

- Built a fashion recommendation system by extracting image features using Convolutional Neural Networks (CNNs).
- Leveraged similarity metrics to suggest outfits based on visual content and user preferences.

User Profiling and Segmentation

- Developed user profiles and segments by analyzing demographic and behavioral data to personalize marketing strategies.
- Used hierarchical clustering to group users into distinct categories for more effective engagement.

Credit Card Clustering

- Applied clustering techniques to segment credit card customers based on spending behavior and transaction patterns.
- Used K-means to identify distinct customer groups for targeted marketing.

Instagram Reach Analysis and Prediction

- Developed a model to predict Instagram reach based on user engagement metrics.
- Analyzed historical post data to forecast audience interaction and optimize content strategies.

Topic Modeling

- Performed topic modeling using Python to identify key themes in text datasets.
- Utilized techniques like Latent Dirichlet Allocation (LDA) to uncover hidden patterns and insights.

Air Quality Index Analysis Using Python

- Conducted air quality index analysis using Python.
- Visualized data and used Machine Learning for pollution prediction.

Movies Gain Prediction

- Managed comprehensive movie dataset cleaning.
- Predicted future film accuracy with advanced statistical methods and machine learning algorithms.

Certifications

- Computation and Programming Using Python, *AUB*
- Advanced Designer Certificate, *Dataiku*
- Core Designer Certificate, *Dataiku*
- ML Practitioner Certificate, *Dataiku*
- Intermediate Python, *DataCamp*
- Data Manipulation with pandas, *DataCamp*
- Python data science toolbox, *DataCamp*
- Introduction to SQL server, *DataCamp*
- SQL for joining data, *DataCamp*

Skills

Programming Languages and Software: Python, R, MySQL, SAS, SPSS, Tableau, Power BI, Excel.

Other Technical Skills: Data Analysis, Data Collection, Web Scraping, Data Cleaning, Machine Learning, Statistics.

Soft Skills: Communication, Problem Solving, Presentation, Time Management.

Languages

English: Professional Working Proficient

French: Professional Working Proficient

Arabic: Native