Carlos Andrés Betancur Yepes.

**Data Analyst| Jr Data Scientist.**

**Phone: +57 3207712588 | E-mail:** [**ndrsbtncr@gmail.com**](https://mail.google.com/mail/u/0/)

Trained scientist with strong skills for mathematical and computational statistical interpretation of complex industry problems, with relevant complementary training courses such as **SQL, Power BI**, and **Python** focused on library manipulation **(Pandas, JSON, RegEX)** for cleaning and organizing **structured** and **unstructured** data. Likewise, other tools for **Machine Learning** such as **Scikit-Learn**, **Pycaret** **Numpy, Matplotlib** and **Seaborn**.

**EXPERIENCE**

**ManpowerGroup Colombia (January 2023 - Present)**

* Development and Automation of **Power BI** dashboards integrated with tools such as **Python** and **SQL.**
* Consumption of APIs with **Python** for information extraction using **web scraping** techniques and manipulation of **JSON** format documents.
* Optimizing queries by applying advanced **Power Query** and **M languaje** usage for big data projects.
* Building advanced formulas in **DAX language** and designing interactive reports leveraging the various functionalities in **Power BI** visuals for **descriptive** and **predictive analytics.**
* Collaborating on a **natural language processing** project that implemented **cosine similarity** techniques to correlate texts and interacted with the **GPT chat API** for model validation.

**PROJECTS**

**DATA MINING (May 2022- December 2022)**

* Creation of **DataFrames** for **structured** data in **CSV** and **TXT** file formats.
* Extraction of **unstructured** data from web pages in **HTML** format using **Web-Scraping** techniques, **RegEx**, and **JSON** text formats.
* Sentiment analysis using **Natural Language Processing (NLTK, SpaCy, SciPy, TextBlob)**.
* Training models in **Machine Learning** such as **Linear Regression** and **Decision Trees**

**BIG DATA ANALYST****(May 2022-December 2022)**

* Processing and debugging of missing or outlier values ​​in unstructured files using **Natural Language Processing** (NLP) and **Machine Learning** (ML) techniques.
* Provide information and support business decisions.
* Establish financial and operational predictions of future trends.

**BANCOLOMBIA (October 2022 - November 2022)**

* Creation of an automatic news categorization system for corporate clients.
* Unsupervised classification of the data by using algorithms such as **K-means** and **Transformers**.

**MATERIALS SCIENCE RESEARCH GROUP, Universidad de Antioquia (June 2021 – December 2022)**

* Development of new theoretical studies on optical layers of experimentally verified non-linear materials.
* Implementation of news algorithms for the solution of differential equations in Wolfram Mathematica languaje.

**SKILLS.**

* Python, C++, SQL, Power BI, Microsoft Excel, Power Pivot, Power Query, M languaje, Data Mining, Machine Learning, Matlab, Wolfram Mathemática, Mathematical and statistical modeling, basic English, basic German.

**EDUCATION**

**PROFESSION**

**Universidad de Antioquia**

* Physics (2014-2022)

**CERTIFICATIONS**

**Universidad de Antioquia**

* Big Data (April 2022 - November 2022)

**MinTic**

* Fundamentals in data analytics (November 2022 – January 2023)

**Udemy**

* Basic, Intermediate, and Advanced Excel. (December 2022 - December2022)
* SQL Server from beginner to advanced. (January 2023 – February 2023)
* Power BI from beginner to advanced. (January 2023 – April 2023)

**REFERENCES**

Senior Data Scientist Michal Hemmerling

* E-mail: [Michal.hemmerling@udea.edu.co](mailto:Michal.hemmerling@udea.edu.co)
* Phone1: +55 16988511933
* Phone2: +55 16988067251

Physicist Lucas Quiceno

* E-mail: [lucas.quiceno@udea.edu.co](mailto:lucas.quiceno@udea.edu.co)
* Phone: 3024611660

**CONTACT**

Phone: +57 3207712588

E-mail: [ndrsbtncr@gmail.com](https://mail.google.com/mail/u/0/)

Document: 1038337471

Birthplace: Frontino-Antioquia

Birthdate: 26/11/2022

Address: CL 38A # 50A-14

Copacabana, Antioquia