



357 - 99250369



ioakeim.h@gmail.com



Nicosia, Cyprus

IOAKEIM HADJIMPALASIS

PROJECTS

[DataGlitch](#) - Python Toolking for Messy Data

[Github series](#) - Feature Engineering

[Data Science Portfolio](#)

TECH SKILLS

Data Cleaning: Data Accuracy, Missing Values, Outliers, Categorical Data (Cardinality Reduction, Encoding), Feature Scaling and Selection.

Machine Learning: Linear Regression, Logistic Regression, Decision Trees, Random Forest, KNN, k-means, PCA, Association Rule Learning, Causal Impact Analysis.

Automation: Automating data processing tasks and streamlining repetitive processes, including file management and organization.

Other: ETL, Web scraping, Linux CLI

CORE SKILLS

Problem-solving

A self-taught programmer with a proven track record of effectively extracting valuable insights from cluttered information. Proficient at devising plans for complex issues, staying adaptable, and achieving optimal outcomes through evidence-based decision-making.

Communication

Regularly collaborates with colleagues and clients, leveraging a diverse background and unique perspective to understand individual needs and motivations. Facilitates effective communication with both technical and non-technical stakeholders.

Teamwork

Adaptable, open to feedback, and committed to clear, efficient work. Strong creativity and collaboration skills that drive successful project outcomes while valuing and integrating team ideas and perspectives.

PROFILE

ioakeim is a data analytics consultant with experience from reputable organizations in the UK and Cyprus. He is deeply passionate about data cleaning and engineering, and he actively contributes to the community through open-source tools and educational resources. ioakeim is dedicated to upholding best coding practices and adopts a research-oriented approach to deliver efficient, user-friendly software solutions. His diverse background equips him with unique problem-solving abilities, which he applies using Python, R, and SQL.

WORK EXPERIENCE

Data Analytics Consultant - KPMG (CY)

Feb 2022 - Present

Data Quality Project for an Insurance Company

- Used Python to automate the extraction and processing of data for the development of a BI dashboard. This dashboard informed key stakeholders about the state of their data and facilitated the validation of user information.
- Made strategic use of regular expressions to enhance variable quality by extracting information that had been mistakenly stored in other columns.
- Leveraged web sources to collect data from postal services. These data were cross-referenced with the user's information to determine the city for each user.
- Applied fuzzy string matching techniques to identify valid email addresses and standardize their domains.
- Contributed to the development of an NLP algorithm based on brute-force search for translating Greek names to English.
- Involved in the design of a scoring system with the goal of identifying duplicate records. This system assigned scores based on multiple variables that met specific criteria.
- Thoroughly tested and modularized the code to ensure maintainability, achieving clarity through logical structuring.

Churn Prediction for a Communication Service Provider

- Debugging and testing substantial volumes of existing Python code used for pre-processing, model fitting and learning.
- Created a script that leveraged historical data to generate new features, calculating product churn and product switch recursively every three months over a four-year period.
- Developed a rudimentary fbprophet model to assess its potential for forecasting revenue.
- Employed Pandas to construct contingency tables for analyzing product competition among the client and other vendors.
- Evaluated data quality by producing correlation plots using Matplotlib.



357 - 99250369



ioakeim.h@gmail.com



Nicosia, Cyprus

IOAKEIM HADJIMPALASIS

EDUCATION

MSc Data Science and Engineering

2023 - Present

MSc Psychology

2020 - 2021

- **Thesis:** Associative Learning in Humans. Data were analysed using two one-way Mixed ANOVAs and several Pearson Correlations.
- **Student support:** Live sessions for master's students who requested support with statistics. Explained theoretical and practical topics.

BSc Sports and Exercise Science

2017 - 2020

- **Thesis:** Physical Exercise on Mental States. Designed the methodological procedure, recruited and tested participants in a laboratory, analysed the data collected, discussed findings and successfully developed a scientific report.

CerTHE Sports Science and Coaching

2016 -2017

- **Coaching:** Designed and conducted boxing sessions for beginners, exploring various leadership styles to foster an autonomous learning environment.

COURSES

CS50's Introduction to Programming with Python

Functions, Variables | Conditionals | Loops | Exceptions | Libraries | Unit Tests | File I/O | Regular Expressions | Object-Oriented Programming | Et Cetera

Data Science Infinity

SQL | Python (Numpy, Pandas, Matplotlib) | AB Testing | Machine Learning (Data Preparation, Supervised & Unsupervised Models) | Advanced Applications of Scikit-learn

WORK EXPERIENCE

Data Analytics Consultant - KPMG

Feb 2022 - Present

Recalculation for Financial Reporting in the Telco and Forex Sectors

- Comparing system populations with SAP Accounting using Microsoft SQL Server. Importing data from local files, creating tables and performing aggregations.
- Performing data cleaning using regular expressions, creating new features, and conducting table joins with Pandas, while effectively communicating data issues to the client and seeking feedback on the implementation of logic.

Customer Segmentation Tool for an Energy and Petroleum Company

- Managed the ETL pipeline by adjusting the extraction connection from the SQL database and incorporated relative paths for flexible csv file exports.
- Resolved code errors and converted the tool into an executable for enhanced usability.
- Maintained constant communication with the client about project challenges and solutions.

Digital Transformation for a Financial Services Company

- Conducted basic data checks on R-generated data using Microsoft SQL Server.
- Performed data manipulation tasks such as table joins from csv files using Pandas.

Quantitative Social Research Placement - Office for National Statistics (UK)

Jun 2021 - Sep 2021

- Used R to analyse Census data from 2011 and 2021. Decided on suitable variables, cleaned these and produced descriptives that helped senior leadership diagnose potential issues.
- Performed a literature review about the mental health of veterans in England and provided a concise summary of the main findings using tables and references.