IOAKEIM HADJIMPALASIS

https://ioakeim-h.github.io/
ioakeim.h@gmail.com
357 - 99250369

PROFILE

Ioakeim is a self-taught programmer with professional experience from reputable organisations in the UK and Cyprus. He is particularly skilled in data transformation and has a robust toolbox that can be applied in various scenarios. His statistical expertise comes from the social sciences where he consistently used linear models to answer research questions. Ioakeim is a highly motivated individual seeking a challenging role in the data domain.

EXPERIENCE

Advisor II - Digital Innovation Team, KPMG (CY) — Feb 2022 - Present

Recalculation of Revenue for BDSwiss

- Data cleaning using regular expressions, feature creation and table joins with Pandas.
- Communicating data issues to client in writing and through online meetings.

Data Engineering and Analysis for Cyta Revenue Prediction

- Debugging and testing substantial volumes of 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.

Data Validation and Manipulation for Bank of Cyprus Digital Transformation

- 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.

Professional Placement - Leadership Team, ONS (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.

PROJECTS

DataGlitch

- A Python package currently under development catering to the needs of data engineers and analysts. Its primary capabilities include identifying and rectifying mixed data types in columns, detecting non-ASCII values, and introducing the LittleMCAR statistical test into Python. Additionally, it offers the ability to convert columns to snake_case and implements Tukey's IQR method to address outlier-related concerns.

Miniseries: Advanced Data Engineering for Machine Learning with Python

Missing Data

- Using regular expressions to identify and resolve inconsistencies.
- Using the missigno package for data visualization and rpy2 to extract the LittleMCAR function from R, to quantify missingness and determine its mechanism.
- Testing and implementing DataWig – a deep learning library for the imputation of missing categorical data.

Categorical Data

- Quantifying and reducing cardinality.
- Popular and evidence-based encoding techniques and potential pitfalls. Methods for both low and high cardinality features.

IOAKEIM HADJIMPALASIS

https://ioakeim-h.github.io/
ioakeim.h@gmail.com
357 - 99250369

TECHNICAL SKILLS

Programming: An experienced Python programmer working with Pandas and Numpy on a daily basis. Good knowledge of scikit-learn and proficiency in other programming languages such as R.

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

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

Relational Databases: Integrating PostgreSQL with Python to perform CRUD operations and extract data for analysis. Basic knowledge of theoretical concepts such as normalisation and denormalisation.

CORE SKILLS

Communication: Collaborating with colleagues and clients on a daily basis to resolve complex technical issues and develop plans of action. My diverse background has given me a unique perspective on problem-solving, especially in understanding the needs and motivations of individuals. I believe this skillset is highly advantageous in a technical role, allowing me to communicate effectively with both technical and non-technical stakeholders.

Teamwork: I'm adaptable, open to feedback, and prioritize producing clear and efficient work. My creativity and collaboration skills have led to successful outcomes on every project. I value and incorporate the ideas and perspectives of my team.

Problem-solving: As a self-taught programmer, I possess a proven track record of effectively extracting valuable insights from cluttered information. I excel at developing actionable plans to address poorly defined problems, and remain patient and adaptable in the face of unexpected events. My decision-making is always grounded in both evidence and experience, allowing me to consistently deliver optimal results.

EDUCATION

MSc Psychology, Cardiff University — 2020-2021

- Developed a thesis that focused on associative learning in humans. Data were analysed using two one-way Mixed ANOVAs and several Pearson Correlations. Results were compared against previous literature and were discussed in detail.
- Provided live sessions for master's students who requested support with statistics. Explained theoretical and practical topics and shaped the learning experience according to individual needs and capabilities.
- Critically evaluated interview data using qualitative analysis techniques. Hypotheses were formed and areas of further investigation were identified.

BSc Sports and Exercise Science, University of Essex — 2017-2020

- Designed the methodological procedure of a study, recruited and tested participants in a laboratory, analysed the data collected, discussed findings and successfully developed a scientific report.
- Presented evidence-based arguments in front of an audience in a debate with a fellow student. Explored ways to “sell” my ideas and focused on the limitations of the opposition.
- Explored different approaches to interviewing clients with sensitive health issues. Open-ended questioning, use of affirmations, empathetic and reflective listening were applied with a hypothetical client.

COURSES & CERTS

CS50's Introduction to Programming with Python, edX

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

Feature Engineering for Machine Learning, TrainInData

Syllabus: Missing Data Imputation | Categorical Variables (Cardinality Reduction, Encoding) | Outliers | Feature Scaling

DATA SCIENCE INFINITY, Andrew Jones

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