# Cem Kazan

Tempe,AZ | kzncem@gmail.com | 480-410-8963 | linkedin.com/in/cem-kazan-data/ | samkazan.github.io/Portfolio/

As a Data Scientist with 2+ years of experience, I have a track record of enhancing sales forecasting accuracy by 20% and reducing data processing by 15%. I am seeking a position that allows me to utilize my skills in Python, SQL, and machine learning to drive business growth and efficiency.

## PROFESSIONAL EXPERIENCE

**HustleHawks**, New Orleans, LA
**Data Science Intern**,                                                                                                     01/2024 -  Present
- Automated report generation using Python & Pandas, reducing manual data processing by 20%
- Enhanced team productivity by 30% and minimized human error in data handling.
- Conducted geospatial analysis in Python, identifying new market regions, leading to a projected 15% increase in sales.
- Addressed critical transaction issues in the company's database, supporting a 20% increase in transaction volume.
- Collaborated with cross-functional teams to analyze business needs, leveraging Python & Tableau public to devise data-driven solutions that bolstered customer retention by 15%.
- Implemented machine learning models utilizing Logistic Regression, Random Forest, & clustering techniques, enhancing sales forecasting accuracy by 20%
- Led a project to clean and organize the company's large-scale data, resulting in a 15% reduction in data processing.

**Arizona State University**, Tempe, AZ
**Research Scientist/Data Analytics**,                                                                                        05/2023 -  Present
- As part of a team of 2 researchers, reported to the Head of the Department, and was responsible for developing a large-scale protein dynamic signature database. Website (https://dynamicsdb.org/)
- Developed a protein dynamic signature database using Python and SQL, providing researchers with ready access to critical data and attracting a 5% monthly increase in user engagement.
- Utilized clustering with RMSD and sequence based metrics. Applied dynamics analyses, and identified dynamically and evolutionarily conserved residues important for protein function and could be a guide for protein engineering.
- Innovated a protein clustering protocol using Python and K-means clustering, achieving an 87% success rate in protein grouping, thereby enhancing the accuracy of protein dynamic studies.

**Four Seasons Resort**, Scottsdale, AZ
**Pastry Chef,**                                                                                                             06/2021 - 07/2022
- Designed training programs for new chefs, leading to a 15% increase in kitchen productivity and a 10% decrease in culinary errors. Oversaw inventory of baking supplies, reducing wastage by 20%.

## PROJECTS

- Predictive Modeling for Yield Excursions in Semiconductor Manufacturing using Machine Learning
- Predicting Academic Dropout or Success in Higher Education,
- Personalized AI, RAG application using Mistral LLM
- Text Classification and Recommendation Systems with PySpark ML and GCP

## EDUCATION

**Bachelor Of Science, Data Science**, *Arizona State University*, Tempe, AZ,                                                  07/2024
- Focus: Business Analytics and Database Management, Awards : Dean's List, GPA: 4.0

**Associate in Applied Science, Audio Engineering and Culinary Arts,** *Scottsdale Community College,*                         06/2021

**Bachelor Of Science, Industrial Engineering**, *Kadir Has University*, Istanbul/Turkey,                                      05/2016

## SKILLS

**Technical Skills**: Programming languages (R, SQL, Python(Sklearn,Pandas,Scikit), PySpark, Spark SQL),Data Visualization (Tableau, Power BI , Matplotlib,Seaborn),Machine Learning & AI (TensorFlow, Keras, PySpark ML), Microsoft Excel, Github, Docker,Flask,Streamlit,Cloud Platforms (AWS,GCP), Statistics (Statistical modeling, hypothesis testing,logistic regression, classification algorithms,predictive modeling, time-series forecasting,decision tree,LLM)