# Ayaan Omair

480-570-4945 | aomair1@tamu.edu | LinkedIn | Portfolio

## Education

**Texas A&M University** — College Station, TX
*Master of Science in Data Science* — *Expected Graduation: Dec. 2025*
- **GPA:** 3.71/4.00

**Arizona State University** — Tempe, AZ
*Bachelors of Science in Mathematics (Statistics)* — *Graduation: May 2024*
- **GPA:** 3.90/4.00 *(summa cum laude)*

## Skills

**Languages & Tools**: Python, R, MySQL, PostgreSQL, Power BI, DAX, Excel, SAS, JavaScript, HTML/CSS, Linux, Tableau, Alteryx, Jupyter, LaTeX
**Libraries & Machine Learning Models**: Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Tensorflow, Pytorch Decision Trees, Regression (Linear, Logistic, Lasso, Ridge, etc.), K-Means, PCA, KNN, SVM, Neural Networks
**Coursework**: Data Mining and Analysis, Applied Linear Regression, Mathematical Statistics, Linear Algebra, Data Wrangling with SQL, Data Analysis with Python, Exploring Data in R/Python, Probability
**Certifications**: Data Analyst *(DataCamp, 2025)*, Data Analysis with Power BI *(DataCamp, 2025)*, Data Analysis with Python *(DataCamp, 2025)*, SQL Specialist *(Global Tech Experience, 2023)*

## Experience

**Data Analyst Intern** — Phoenix, AZ
*Arizona State University* — *January 2024 – May 2024*
- Designed and implemented an ***ETL*** process to support youth soccer coaches by extracting player data from over **100** Excel files, combining and preprocessing it with ***Python (Pandas)***, and loading results into Excel
- Developed a user-friendly interactive *dashboard* in ***Excel*** to visualize comprehensive player insights and statistics
- Coordinated with coaching staff to identify important metrics and adjust how data was shown to meet their needs
- **Tech Stack:** Python, Pandas, Excel

**SQL & Python Trainee** — Remote
*Global Tech Experience* — *May 2023 – July 2023*
- Devised ***SQL*** queries to efficiently extract, analyze, and manipulate complex datasets for actionable insights
- Analyzed and visualized data findings using ***Jupyter Notebook*** and ***Python***
- Gained insights into global business strategy through EDA, summary statistics, and visualization
- Collaborated with a global team to perform different tasks using ***SQL*** and ***Python***
- **Tech Stack:** Python, Matplotlib, SQL, Tableau

## Projects

**Disease Prediction Model** | *Python, Scikit-Learn, Pandas, NumPy, Matplotlib, Jupyter Notebook*
- Constructed *logistic regression* and *decision tree* models in ***Python*** to predict Heart, Kidney, and Skin Disease
- Achieved up to **75%** accuracy and AUC scores as high as **0.84** across all models
- Identified high recall (**78%**) with low precision due to class imbalance between diseased and undiseased cases
- Compared model performance using F1-score and AUC; verified no overfitting by analyzing training vs. test accuracy

**NBA Player Stats Dashboard** | *Power BI, DAX, Power Query, Excel*
- Built an interactive **Power BI dashboard** analyzing **5,300** NBA players with filters, visuals, and KPI cards
- Enabled player-level insights by implementing search filters to display individual career stats (PPG, APG, RPG) through **cards** and season-by-season performance trends using interactive **line charts**
- Visualized **top 10 players** by points, assists, and rebounds using **line and stacked column charts**
- Implemented navigation buttons to link between dashboard pages and enhance user interactivity
- Structured, cleansed, and transformed data using **Power Query** and **DAX** code