

AHMED L. RASHED

Columbus, OH 43204 | 781-309-7560 | [LinkedIn Profile](#) | [GitHub Portfolio](#) | arashed07@gmail.com

Professional Summary

Highly analytical and **process-oriented** data cruncher with over 10 years of experience in **data mining, statistical analysis, business intelligence gathering, trending, and benchmarking**. Created **ML models** on device telemetry to **pre-emptively classify CTQ** device failures before Final Test, *cutting rework times by 25%* and *increasing throughput by 10%*. Collaborated with test engineers to develop **ETL processes** to facilitate the **movement of data** between systems, **identifying patterns and insights** in the data, *reducing failure rates from 35% to 2%*.

Technical Skills

Data Analytics: SQL, R Programming, Python, Tableau, Power BI (PowerQuery, DAX), Statistics, Hypothesis Testing.

Machine Learning: Linear/Logistic Regression, GridSearchCV, Random-Forest, XGBoost, Pickle, Decision-Trees.

Database & Project Management: EDA, ETL, SQL Server, Azure Data Studio, Google BigQuery, Agile, Scrum, JIRA, Git.

Business Intelligence: Collaboration, Continuous Improvement, Data Cleaning, Data Visualization, Data Presentation, Requirement Gathering, Problem Solving, Root-Cause Analysis, Turning Questions into Queries and Data into Insights.

Education & Certification

[Google Advanced Data Analytics Certificate](#)

Master of Arts – Physics - Bryn Mawr College

[Google Data Analytics Certificate](#)

Bachelor of Science - Applied Physics - UMBC

Work Experience

- | | | | |
|---|---------------------------|-------------|-------------------|
| <i>Data Analytics Consultant</i> | HENNY PENNY | REMOTE- OH | 11/2022 – 08/2023 |
| Implemented data hooks in production test stand software to identify key metrics and transform raw data into meaningful, actionable information through automated extraction and standardized reporting . | | | |
| <ul style="list-style-type: none">Created various Python scripts (numpy, pandas) using SQL (complex joins) to extract and analyze factory data, enhancing build quality and <i>saving Mfg engineers over 15 hours each week</i> in manual data manipulation.Collaborated with test engineers to develop ETL processes to facilitate the movement of data between systems, identifying patterns and insights in the data, <i>reducing failure rates from 35% to 2%</i>. | | | |
| Key Project: Created python utility to extract and transform data from TestStand SQL database schema into flat CSV files. | | | |
| <ul style="list-style-type: none">Tools Used: SQL Server Management Studio, T-SQL, Python, PYODBC, Pandas.Results: Reduced script to single-file, used <i>auto-py-to-exe</i> to build a stand-alone executable file for users to download. | | | |
| <i>Quality Engineer Consultant</i> | GRANVILLE PHILLIPS | REMOTE- CO | 09/2022 - 01/2023 |
| Developed and integrated data-driven metrics into existing TestStand system to optimize Mfg process. | | | |
| <ul style="list-style-type: none">Implemented Python predictive analytics (Scikit-learn) on test-unit data streams to pre-emptively classify CTQ device faults before Final Test, <i>cutting rework times by 25%</i> and <i>increasing throughput by 10%</i>.Deployed Mfg Statistical-Process-Control dashboards, <i>improving first-pass-yield rates from 75% to 95%</i>. | | | |
| <i>Principal Test Engineer</i> | MKS INSTRUMENTS | Methuen, MA | 03/2021 - 06/2022 |
| Managed test station software projects, ensuring quality and reliability of software through comprehensive testing and analysis . Picked up Python and SQL on-the-job to drive enhanced data analysis and visualization for Tech Ops and Business Unit teams, creating compelling data-driven reports to upper management. | | | |
| <ul style="list-style-type: none">Extracted and extrapolated Mfg data using Python, SQL, and Excel for monthly quality reports, <i>preventing over \$50,000 in lost annual revenue</i> by catching product quality defects BEFORE they leave the factory floor.Designed and built statistical analysis models on large data sets using Azure Data Studio that helped <i>reduce product rework and post-processing test-time by 15%</i>. | | | |

Incorporated extensive upgrading and overhauling legacy product lines. Rewrote test software architecture into **robust, data-driven, user-friendly applications**, *improving quality, reliability, and efficiency*.

- Connected **Tableau dashboards** to centralized **data-pipeline** for real-time production **data tracking**, *improving average time-to-bug-fixes by 20% and highlighting time-sinks* in factory workflows.
- Created **visually impactful dashboards** and **executive data visualizations** in Tableau to track production yields and failure modes, *saving test engineers over 10 hours weekly* in diagnostics and manual reporting.

Data Projects

TikTok Video Classification:

[TVC GitHub Link](#)

12/2023

Exploratory data analysis to identify key metrics and trends to inform and drive appropriate marketing strategy.

- **Tools Used:** Jupyter Notebook, Pandas, Pathlib, Matplotlib, Seaborn, Scikit-learn, Statsmodels.
- **Approach:** Developed machine-learning algorithm to predict whether videos presented claims or opinions to improve triaging process of videos for human review.
- **Results:** Both **RandomForest** and **XGBoost** architectures provided near-perfect models, but **RandomForest** had slightly higher accuracy and less processing, so it is the champion model. The most predictive features all were related to engagement levels generated by the video.

Sales Dashboard Report:

[SDR GitHub Link](#)

11/2023

Demo dashboard mockup with downloaded SQL tables to show regional sales in Power BI Desktop.

- **Tools Used:** Azure Data Studio, T-SQL, **Power BI**, Interactive Data Visualization.
- **Approach:** Cleaned and sliced regional sales dataset with various **SQL queries**. Imported data to **PowerBI** with **PowerQuery** preprocessing to display **Key Process Indicators**. Created various **Slicers** for interactive filter views.

Housing Prediction Model:

[HPM GitHub Link](#)

10/2023

Built and optimized a predictive regression model of housing prices with historical CA housing data.

- **Tools Used:** Jupyter Notebook, Pandas, Pathlib, Matplotlib, Seaborn, Scikit-learn.
- **Approach:** Identified and removed **outliers** from the dataset. Created and modified the attributes of the data (**feature engineering**) to input into the model that is most relevant to **predicting median-house-value**. Tweaked and tuned Linear, GridSearchCV, and RandomForest Regressions to achieve 83% accurate predictions.
- **Results:** Evaluated **RandomForest Regression** as the best overall model to predict median-house-value.

COVID-19 Retrospective:

[Cov19 GitHub Link](#)

09/2023

Explored and visualized COVID's most affected countries, highest death rates, and fastest vaccination rates.

- **Tools Used:** Azure Data Studio, T-SQL, **Tableau**, Interactive Data Visualization.
- **Approach:** Explored and sliced the dataset with various **SQL queries**. Exported tables to CSV files for later importing to **Tableau Desktop**. After exporting and visualizing, we **discovered an error in the calculations**. *New_Vaccinations* included both initial AND booster shots, resulting in *Running_Percent_Vaccination* going over 100%. Had to go back and **reformat our query** to use *people_vaccinated* instead of new or total vaccinations.
- **Results:** Published visualization to **Tableau Public** [LINK](#).

Bike Sharing Analysis:

[BSA GitHub Link](#)

07/2023

Exploratory data analysis to identify key metrics and trends to inform and drive appropriate marketing strategy.

- **Tools Used:** R, **RStudio**, Data Extraction, **Data Cleaning**, Data Transforming, Data Interpretation.
- **Approach:** Scraped the public data sets from the **AWS S3 bucket**. Loaded the CSV files with **fread()** and use **rbindlist()** to merge the individual data sets together. **Transformed data** by adding calculated columns. Generated **aggregate descriptive statistics** of categorical and numerical comparisons to identify A vs B populations.
- **Results:** Minor **trends** in weekday vs weekend usage. Exported data for further time-series analysis.