**Benjamin Cox**

931-675-4534 • Petersburg, TN, 37144 • bjcbencox97@gmail.com • [LinkedIn](https://www.linkedin.com/in/benjammin97) • [GitHub](https://github.com/benjammin97)

**PROFESSIONAL EXPERIENCE**

**UAH** – *Huntsville, AL* Dec 2022- Present

* Suggested process improvements to help meet business goals.
* Gathered and reviewed information for quality and accuracy.

**Frontdoor** – *Memphis, TN (Remote)* June 2022 - Aug 2022

**Process Improvement Analyst Intern**

* Conducted data analysis on quantitative historical data using the DMAIC method and multiple commercial tools to evaluate the potential for standardization for the cash-in-lieu business process.
* Performed complex queries (Joins, With statements) on multiple datasets using Snowflake and Looker. Analyzed those queries to bring meaningful insights in Excel using Pivot Tables and Excel formulas. Developed those insights into clear visualizations using Power BI. Integrated those visualizations into the presentation with PowerPoint.
* Presented the completed analysis to multiple levels of management including C-suite managers. Presented to a technical and a nontechnical audience. The research resulted in the start of a pilot with an estimated $4.8M cost reduction.

**Resolution LLC** – *Huntsville, AL* Mar 2022 – May 2022

**GIS Analyst**

* Managed and leveraged GIS resources (ARCMAP, GAIT, Data Reviewer) to create maps and graphic reports. Collected and handled data, provided mapping services, performed technical research and analysis.

**Academy of Leaders** – *Orlando, FL (Remote)* Oct 2021 – Mar 2022

**Tutor**

* Educated multiple students in STEM courses from 6-12 grade. Above 80% of students maintained a B average or better. One senior went to her first-choice college.

**EDUCATION**

**The University of Alabama in Huntsville (UAH)** – *Huntsville, AL*

MS – Analytics, May 2023

**Coursera** – *Online*

Google Data Analytics Certification, September 2021

**The University of Alabama at Birmingham (UAB)** – *Birmingham, AL*

BS – Mathematics, May 2021

**DATA SKILLS**

* Data Science Skills: Machine Learning (Unsupervised and Supervised), Statistical Analysis, Database Management, Regression, Classification, Data Visualization, Data Preprocessing, Data Balancing, Data Modeling, Data Analysis, Data Sets, Excel Macros, Pivot Tables
* Data Science Tools: Snowflake, Looker, Power BI, Jupyter Notebook, Excel, Oracle, Tableau
* Programming Languages: Python, R, Java, C, SQL

**DATA PROJECTS**

**Potential Customer Prediction using Supervised Machine Learning**

DATA TRAINING IN PYTHON | CLASSIFICATION & REGRESSION IN PYTHON| DATA BALANCING IN PYTHON

* Predicted the number of claims and the cost of claims a potential customer test dataset would have based on the accuracy scores (F1 and mean absolute error used) of a 75/25 independent test split of the training dataset for Machine Learning models (k-nearest neighbors, decision trees) including both Classification and Regression. Balancing techniques such as undersampling, oversampling, and SMOTE were used to improve F1 scores.
* Training dataset had ~60000 rows, but data binning reduced the number of deleted rows by 40%. Data usage was maximized with hot encoding. Test dataset had ~7500 rows.
* Company saved an average of $200 per customer by only choosing customers who are predicted to have 0 or 1 claim.

**Policyholder Identification with Unsupervised Machine Learning**

DATA PREPROCESSING IN PYTHON | CLUSTERING IN PYTHON | KMEANS & DBSCAN IN PYTHON

* Computed the optimal number of clusters to categorize a group of auto insurance policyholders by implementing the k-means and the DBSCAN clustering algorithms in Python. The size was ~60000 rows of data and contained a large number of missing values.
* Developed and presented a PowerPoint presentation to teacher and classmates with the first half aimed at a nontechnical audience and the second half aimed at a technical audience using data visualizations created with Matplotlib and Seaborn packages.
* Profits increased by 30% by implementing different rates for different risk levels.

**Movie Rental Relational Database**

DATA MINING IN SQL| DATABASE MANAGEMENT IN SQL| SCRIPTING WITH PL/SQL

* Built a relational database from scratch using a movie rental company’s records. Drew an entity-relationship diagram to illustrate the functionality of the relational database.
* Performed complex queries on the data including join statements to display members who have not paid their rental fees, to convert date data from a standard m-d-y format to numbers of weeks passed, and creating a PL/SQL block statement to display the top n movies in numbers of sales.
* Several actionable business insights were drawn from the questions answered which had the potential to reduce costs by 15% and increase profits by 10%.

**Cyclistic Bike Share Case Study**

DATA PREPROCESSING IN R | DATA ANALYSIS IN R | DATA VISUALIZATION IN R

* Cleaned, processed, and analyzed over 3.5M rows of data in R, using the tidyverse, janitor, ggplot2, lubridate packages, and basic machine learning. Resulted in clean, accurate data to set up clear, insightful visualizations.
* Generated, in R(ggplot), a line graph comparing the number of riders to the time of day, and generated bar graphs comparing the member riders to casual riders’ behavior based on day of the week as well as behavior based on the length of ride.
* Several recommendations were made to increase rider membership based on the data, such as a weekend membership, a one-hour membership, and a 3PM-6PM membership.