**Allan Waweru**

(470) 209-3566  [linkedin.com/AllanWaweru](https://www.linkedin.com/in/allan-waweru/)

Kennesaw, GA 30144 [allan.n.waweru@gmail.com](mailto:allan.n.waweru@gmail.com)  [github.com/AllanWaweru](https://github.com/Allanwaweru)  c cc

**EDUCATION**

**The University of Georgia**, Athens, GA Expected Graduation: May 2023

* **Major**: Data Science
* **Relevant Coursework**: Software Development, Database Management, Data Structures, Statistical Programming, Applied Linear Algebra, Applied Regression Analysis, Machine Learning (IP), Data Mining (IP)

**SKILLS**

* Programming Languages: Python (NumPy, SciPy, Pandas, TensorFlow, Keras), Git, R, Tableau, SQL (intermediate)
* Machine Learning: Classification, Regression, Feature Engineering, Data Scraping (see projects)
* Statistics: Hypothesis Testing. Regression, Confidence Intervals, Basic Probability Theory
* Soft Skills: Effective Communication, Self-Starter, Adaptability, Intellectual Inquisitiveness, Business Acumen

**PROJECTS**

**NBA Salary Prediction**

* Built predictive models to calculate the ideal NBA player’s salary based on their game statistics.
* Used Feature Engineering to identify the variables that stood to correlate the most with salary i.e., Points, Assists, Rebounds, Minutes, Age.
* To understand the efficiency of the models, the R-Squared values for each of the models were:
  + Random Forest Regressor: 61.4% | K-Nearest Neighbor: 61.8% | Linear Regression: 64.4%

**Formula 1 Exploratory Data Analysis**

* Exploratory data analysis comparing performances of Lewis Hamilton and Max Verstappen at Abu Dhabi Grand Prix.
* Used the web-scraper Beautiful Soup to parse the HTML data of the F1 webpage into a more readable format.
* Plotted lots of visuals from the discovered race insights with seaborn, matplotlib, and Plotly.

**PROFESSIONAL EXPERIENCE**

**Federal Reserve Bank of Atlanta,** *Data Science Intern* May 2022 – August 2022

* Explored the viability of using NLP to support the legal agreement administration between the reserve bank and various depository institutions.
* Used python(OCR) to extract contact information from the legal documents and automatically input it into the related databases, saving analysts 14 hours of manual work weekly.
* Wrote power query scripts that organized various banks’ ABA number based on merger dates
* Created an interactive cohort analysis in Tableau for the Discount Window team that revealed previously uncovered trends in the districts’ collateral lending.

**UGA EITS Externship**, *Data Warehouse and Analytics*May 2021 – August 2021

* Wrote Domo scripts for an interactive query system that provided housing for students affected by COVID.
* Worked on Amazon Redshift Database with exposure to cloud computing infrastructure solution such as EC2, S3.
* Gained experience with dimensional modeling using star schema and snowflake models.

**Kennesaw State University Bookstore**,*Student Assistant* December 2019 - August 2020

* 1 of 4 students responsible for implementing *Foxtrot RPA* to streamline unstructured employee data from payroll.
* Prepared attendance reports and aided in processing payroll for bookstore staff.
* Increased the annual productivity of the payroll department by 15%, saving an additional $7,500.

**RELEVANT EXPERIENCE & INVOLVEMENT**

**Association for Computing Machinery** October 2020 – Present

* Accepted as 1 of 25 to UGA’s organization that prepares students for careers in computer science related fields.
* Discuss various topics such as big-data analytics, bioinformatics, A.I at weekly meetings, enhancing debate skills.
* Volunteer at various high schools educating the youth of the many opportunities available in Computer Science.

**CERTIFICATIONS**

KPMG Data Analytics Consulting Virtual Internship, Google Data Analysis Certification (IP)

**INTERESTS**

Basketball, Formula 1, Working out, Sneakers