

# Steven Cocke

Woodbridge, VA 22193 ♦ Cell: (737)-212-4172 ♦ [sfcocke@gmail.com](mailto:sfcocke@gmail.com)

## Clearance

---

- ♦ Active TS/SCI Clearance w/ CI Polygraph

## Links

---

- ♦ LinkedIn: <https://www.linkedin.com/in/steven-cocke/>

## Education

---

- ♦ Master's Degree: Data Science with Concentration in Machine Learning
  - Southern Methodist University – Dallas, TX
  - Graduated March 2020, 3.9 GPA
  - Curriculum: <https://datascience.smu.edu/academics/courses/>
  - NASA Project: Published paper in a peer-reviewed journal; created the automation of spectroscopic detection of molecules in planetary atmospheres using machine learning techniques
- ♦ Bachelor's Degree: Systems and Information Engineering
  - University of Virginia - Charlottesville, VA
  - Graduated May 2016, 3.2 GPA
  - Curriculum: <https://engineering.virginia.edu/departments/engineering-systems-and-environment/academics/systems-engineering-program/se-curricula>

## Employment History

---

- ♦ Parsons Inc. – Centerville, VA  
**Senior Data Scientist**, October 2021 to Present
  - Description limited due to intelligence classification but some areas can be elaborated on further
  - Contract was bid on and won with myself as the single consultant on-site supporting a newly created data science division within a government agency
  - Implemented an agile work environment within two separate teams with staggering project cycles and coached the division on data science/consulting/branding fundamentals. Provided guidance into data science best practices and successfully brought in new technologies to satisfy unmet organizational needs (model repositories, ETL tools, etc)
  - Created a standard project proposal process with minimum requirements that was implemented in order to better organize the requests that the division receives
  - Guided the implementation of workflow analyses and solution architect workspaces between the division and the internal customers we engaged with. Also facilitated the communication between the division and my direct team back at Parsons.
  - Supported multiple on-going projects at one time from a data science perspective by assisting in writing and debugging code, testing pre-built functionalities, applying natural language processing (NLP) and computer vision (CV) techniques where relevant, and researching methodologies in order to implement and solve new requirements. NLP related tasking included document preprocessing/applying optical character recognition (OCR), topic modeling, and entity tagging. CV related tasking included creation and evaluation of object detection and classification models
  - Assisted in adopting the AWS platform into the team's working environment. Adapted previously created projects to the AWS environment, as well as worked to adapt code from one classification level to another

# Steven Cocke

- Technology and related skills include Python, Pytorch, Huggingface, Spacy, R, AWS Sagemaker/ECR, Docker, GitLab, ArcGIS Pro, and Linux
- ◆ The Knowledge House (TKH) – Remote; Part-time  
**Data Science Instructor**, September 2021 to Present
  - Assist a cohort of 40 students through the year-long Data Science Learning Track (DSLTL). The goal of TKH is to support lower-income young adults by providing a free education in order to place them in internships or jobs that meet a minimum salary requirement by the end of the program. Responsible for leading virtual live classes as well as assisting other instructors in break-out sessions, grading homework and projects, updating course material to keep it relevant, and holding office-hours weekly in order to guarantee student success. Also developed the curriculum and created Phase assessments/mid-terms.
  - The DSLTL aims to take students from little to no programming experience to junior data scientists capable of using a technology-based Python programming language to access, clean, and analyze data, as well as work directly with databases using both Python and SQL. Phase I of the program reviews Python Fundamentals. Phase II of the program reviews core data science concepts such as cleaning/mining data, basic statistics and measures of central tendency, linear regression, and implementing supervised and unsupervised machine learning algorithms. Phase III of the program covers more advanced data science topics such as natural language processing and computer vision, using R as a juxtaposition to Python, and fine-tuning consulting and presentation skills.
- ◆ General Dynamics (GDIT) – Springfield, VA  
**Data Scientist**, July 2020 to November 2021
  - Accountable for creating a deep-learning object detection pipeline that utilizes an HPC/Slurm multi-GPU environment. Full pipeline consists of ingesting satellite imagery, labeling imagery and/or video, preprocessing and augmenting of images, model-training in different deep-learning frameworks, implementing a model repository, creating an automated program for inference across a cluster of 10 nodes, and outputting the results to visualization tools. Technology and related skills include LabelBox, Weights & Biases, Roboflow, Python, Linux, Slurm, Tensorflow, Pytorch, MxNet, OmniSci, Elastic Search, and Kinetica.
  - Coordinated weekly meetings between our team and vendors such as NVIDIA, Roboflow, Weights & Biases, OmniSci in order to troubleshoot and resolve issues or set up workshops in order to advance projects and related knowledge.
  - Responsible for presenting/live demoing the object detection pipeline to GDIT and NGA leadership, as well as incorporate feedback, provide proof of concept, future plans, and project schedules.
- ◆ Driven Brands – Charlotte, NC  
**Senior Data Analyst**, January 2020 to July 2020
  - Reported directly to the VP of Marketing Intelligence. Responsible for defining project scope and timelines, as well as providing updates to the business. Created presentations and delivered the learnings/recommendations to the executive team and Driven Brands business segments. Utilized Google Cloud Storage, Google BigQuery, R, Python, and Alteryx.
  - Lead the effort of analyzing the Digital Marketing Platform (DMP) by working inbetween Salesforce and the Driven Brands Digital Marketing team. Implemented Markov Chain analysis to find optimal digital touchpoint paths in order to optimize digital cost-savings and define device targeting strategy for each audience of customers.
  - Lead the analytical effort of providing fleet health metrics to the Driven Brands Fleet team on a monthly basis, as well as for adhoc requests. Performed clustering analyses of different brand's shops across the U.S. to discover patterns and optimal attribution in order to create baseline targets for 2020 shop strategies. Created processes to ingest third-party data in order to perform fleet penetration analysis. Created Python scripts to automate the process of reverse address lookups in order to append additional business information via Google APIs.

# Steven Cocke

- Combined COVID-19 variables, macro variables, and attributes created in-house to forecast sales for the remainder of 2020 in order to analyze the affect of the virus on businesses.
- ◆ Honeywell – Charlotte, NC  
**Business Intelligence and Datawarehouse Lead**, April 2019 to December 2019
  - Responsible for the entire data management life cycle of a commercial excellence program that reports directly to the CEO of Honeywell. The life cycle includes the aggregation and consolidation of data from various sources – both internal and external, ETL of data via software workflows as well as SQL manipulation, the management of landing tables within a datawarehouse, and the ownership of the sources and security of a data visualization platform that supplies 100+ users
  - Performed initial data quality assessment and exploratory data analysis as a separate deliverable to leadership
  - Created initial machine learning classifier/regresser models for predictive analytics using python and integrated it with ETL workflows. Models were both supervised and unsupervised.
- ◆ Informatica – Austin, TX  
**Professional Services Consultant**, February 2019 to April 2019  
On top of the roles and responsibilities held as an Associate:
  - Execute business and technical roles for Data Quality and Data Integration projects by delivering and implementing scalable and high-performing solutions for major banks, healthcare, and manufacturing companies
  - Lead meetings with customer department managers and directors to guide them with the configuration and the architecture of their solutions on how to best meet their business requirements and data stewardship activities, as well as provide knowledge transfer
  - Prepare and present both detailed and high-level product demonstrations and technical documentations, ranging from installations to product capability demos, for customers from different departments and skill backgrounds
  - Work alongside senior consultants, customers, and business partners both on short-term and long term engagements, to provide high quality and efficient Informatica expertise in accordance with customer expectation
- ◆ Informatica – Austin, TX  
**Professional Services Associate Consultant**, July 2016 to February 2019
  - Implemented Informatica Data Integration, Data Quality, and Master Data Management software, a set of ETL (Extract, Transform, Load) tools, that allows large sets of data to be cleaned, transformed, and warehoused – dramatically cutting overhead costs and improving business processes for customers
  - Developed and unit tested mappings and workflows, which are responsible for sending millions of records from the source table to final base object table, according to functional design specifications – Operating systems: Windows, Linux; Databases: SQL Server, Oracle, AmazonS3; Languages: SQL, Java
  - Worked on an internal project that involved creating an Informatica Cloud Healthcare Accelerator
  - Participated actively on Informatica product enablement, whether it was learning a new product or a new version of a product, to acquire new knowledge or to stay up-to-date with software patches and releases
- ◆ Department of Defense – NSWC Naval Base, Dahlgren, VA  
**Systems Safety Engineer**, June 2014 to May 2016
  - Assisted Navy program's compliance with Military Standards (MIL-STD) and other requirements/directives that defined and guided Program Managers (PMs) in ordnance with system safety that had to be met to field their respective systems

# Steven Cocke

- Sought approval for projects from the Weapon Systems Explosives Safety Review Board (WSESRB) and the Naval Ordnance Safety and Security Activity (NOSSA)

## Certifications

---

- ◆ Certified SAFe 5 Practitioner – Scaled Agile
  - <https://scaledagile.com/>
- ◆ Certified Entry-level Python Programmer
  - <https://pythoninstitute.org/certification/>
- ◆ Certified Associate in Python Programming
  - <https://pythoninstitute.org/certification/>

## Publications

---

- ◆ Automated Spectroscopic Detection And Mapping Using ALMA and Machine Learning Techniques
- ◆ UVA Emergency Department Patient Flow Simulation and Analysis
- ◆ Blockchain Technology – A Survey and Tutorial
- ◆ Methods for Access Control in Relational and Non-Relational Databases

## Relevant Classes

---

- |   |                                |                               |
|---|--------------------------------|-------------------------------|
| ◆ Data and Network Security                 | ◆ Visualization of Information | ◆ Systems Evaluation          |
| ◆ Data Mining                               | ◆ Cloud Computing              | ◆ Systems Engineering         |
| ◆ File Organization and Database Management | ◆ Machine/Deep Learning        | ◆ Linear Statistical Modeling |
|   | ◆ Time Series Analysis         |                               |

## Skills/Technology Summary

- |                                  |   |                                    |
|----------------------------------|---|------------------------------------|
| ◆ Consulting and Data Analytics  | ◆ Java, SQL, R, Python, SAS, XML                      | ◆ Google Cloud Platform            |
| ◆ Data Engineering               | ◆ Google BigQuery, Oracle, SQL Server, MySQL, MongoDB | ◆ Linux, Windows, Amazon S3        |
| ◆ System Analysis and Projection | ◆ Slurm, Multi-GPU Processing, HPC                    | ◆ Microsoft Applications           |
| ◆ Tableau, Alteryx               | ◆ Roboflow, LabelBox, Docker                          | ◆ OmniSci, ElasticSearch, Kinetica |
| ◆ Tensorflow, Pytorch, ONNX      | ◆ AWS Sagemaker/ECR/S3                                | ◆ ARC GIS Pro                      |
| ◆ RabbitMQ, AmazonSQS            |   |                                    |

**Steven Cocke**

*References furnished upon request*