

Corey W. Sykes

Data Scientist | 757.710.0769 | corey@coreysykes.com

SUMMARY

I am an experienced data scientist working within the Advanced Analytics & Strategy division of Cisco's Digital Marketing. I have an expertise in web data and customer behaviors, working on research and development of new ideas with the use of big data, focusing mainly on the petabytes of web data we store. I work on high profile projects that are shifting all of Cisco's Marketing to digital-first and am very capable of providing actionable insights and driving profitable business outcomes with the use of big data. Most recently, I worked on the research, development, and implementation of a Customer Engagements KPI to support a marketing objective of developing deep and rich customer relationships and experiences that drive increased business for Cisco.

TECHNICAL SKILLS SUMMARY

- Solid working understanding of:
 - Processing structured and unstructured data sets within Hadoop's HDFS for complex value added processing and enrichments using Hive, MapReduce, Pig, and User Defined Functions (UDFs) with Java.
 - Designing, architecting, and implementing enterprise-level big data solutions for a variety of analytics works using the Hadoop ecosystem and related tools to develop scalable production code.
 - Data Warehousing concepts with an emphasis on ETL and Life Cycle Development – including requirements analysis, design, development, testing, and implementation.
- Extensive experience in:
 - Apache Hive, Spark, Drill, HBase to query and analyze large data sets stored in HDFS.
 - Software Development Lifecycle Tools like JIRA, Rally, and Git
 - Using Tableau, DOMO, and open-source JavaScript libraries (D3) to create custom visualizations and applications.
 - Creating data pipelines with the use of Shell, Hive, and Sqoop
- Deep knowledge of:
 - Using R and Python's open source libraries to statistically analyze large data sets and create data models that drive profitable business outcomes.
 - Hive dynamic partitioning, bucketing, compression techniques, columnar format files, and YARN architecture.
 - Optimization of Hive queries – map-side joins, reducer-side joins, combiner functions, shuffle & sort.
 - Data story-telling to provide actionable insights to the business.

Languages	PHP, Python, R, HTML/CSS, JavaScript
Big Data	Hive, Spark, Drill, HiveServer2, Pig, Sqoop, Oozie, MapReduce, HDFS, HUE, Druid, ZooKeeper, Oozie, Tidal
Scripting	Bash, PowerShell, Crontab, SQL
Visualization Tools	Tableau, Domo, QlikView, JavaScript (D3), Platfora
Algorithms	Bayesian Networks, Optimization, Decision Trees, Regression, Attribution, Forecasting
Databases	Teradata, Oracle, SQL Server, MySQL, Cassandra, Graphs, MongoDB, DB2
Web Data	Tealium, Adobe SiteCatalyst, Google Analytics, Radian6, DynamicSignal, Graphs, Cypher

Corey W. Sykes

Data Scientist | 757.710.0769 | corey@coreysykes.com

PROFESSIONAL EXPERIENCE

Data Scientist, Cisco Systems, Research Triangle Park, NC

March 2015 – Present

- Work with large, complex, raw server web logs; solve difficult, non-routine analysis problems, applying advanced analytic methods as needed. Conduct end-to-end analysis that includes data gathering and requirements specification, processing, analysis, ongoing deliverables and presentations.
- Build prototype analysis pipelines iteratively to provide insights at scale for stakeholders. A/B & multivariate test. Write scalable code for production data pipelines. Create production dashboards with Tableau and Domo.
- Make business recommendations (forecasting, KPI creation, recommendation engines, attribution analysis) with effective presentations of findings at multiple levels of leadership through visual displays of quantitative information.

Projects

○ **Defining Customer Engagement and Content Performance on Cisco.com**

- Worked within a small team to develop a new metric, coined 'Engagements' for Cisco – one of the KPIs all Cisco Marketing organizations are now goaled and measured against.
- Developed a customer engagement and content performance model by doing deep exploratory data analysis on content across Cisco.com domains, using random forest modeling to find important factors that are indicative of engagement to content and Bayesian networks to calculate metric weightings.
- Continued further to architect and engineer a massive production dashboard with another Marketing team to track performance for all regions and countries

○ **Cisco.com Real-time Offer Personalization Engine**

- Helped deliver insights through data modeling efforts with a group from a Cisco Data Science program graduating class out of University of Washington/ NC State. Contributed with three others to develop a real-time personalized engine (in R, then ShinyR for application use) for recommending offer types (analyst reports, whitepapers, webcasts, etc.) down to the cookie level on Cisco.com by various factor analyses, through extensive exploratory data analysis, followed by GBM modeling.

○ **Using Graph Databases for Path Analysis through Cisco.com**

- Developed a front-end web-based application with various languages (HTML, CSS, JavaScript (D3), jQuery, Python and NetworkX) that allowed us to use both out-of-the box algorithms and our own to analyze popular paths through Cisco.com.

Pricing Analyst, Ferguson Enterprises, Radford, VA

June 2013 – March 2015

- Gained extensive knowledge of Oracle PL/SQL, IBM Cognos, Access, Excel/VBA, SharePoint, and in-house applications by using them to conduct intensive market research on both Ferguson's wholesale pricing structures and close competitors to enhance margin opportunities.

Projects

○ **Ferguson Customized Reporting Systems**

- Created an office LAN Website hosted on Apache, using PHP, AJAX, HTML, CSS, PL/SQL, and jQuery to allow employees to develop customized sales reports by various criteria. Also allows users to run stored procedures via mixing PHP and Cron scripting. The process helped to automate and optimize several weekly business processes, as well as drastically cut costs and time on reporting.

CSykes Web Design, Freelance, Raleigh, NC

October 2013 – Present

- Self-driven learning of many languages, including HTML, CSS, JavaScript, jQuery, MySQL, PHP, and AJAX to create responsive websites. Currently used as a sandbox for learning new languages and visualization techniques.
- Personal Website: <http://www.coreysykes.com/>

Corey W. Sykes

Data Scientist | 757.710.0769 | corey@coreysykes.com

SKILLS & CERTIFICATIONS & AWARDS

Computer Skills

- See technical summary above: very comfortable picking up new technologies and languages on the fly.
- Microsoft Office Specialist Certified in Office 2003, 2007 (Word, Excel, PowerPoint, Access)
- MTA, MSCA for Microsoft SQL

Presentations/Speaking Events

- INFORMS Ops Research & Analytics Conference 2016 – Orlando, FL
 - Presented in the poster session on how Bayesian Networks are used at Cisco.
- 4th Annual Bayesia Conference 2017 – Nashville, TN
 - Presented on recent work involving defining a customer engagement KPI at Cisco
 - Presentation Deck: [Here](#)

EDUCATION

Virginia Polytechnic Institute & State University, Blacksburg, VA

Fall 2009 – Spring 2013

- Bachelor of Science, Economics

Awards

- Pamplin Leadership Award, 2009-2013
- H.C. & C.W. Watson Scholarship, 2009-2013
- Hallie Whealton Trust Fund Scholarship, 2009-2013
- Kiwanis Club Scholarship, 2009
- Masonic Lodge Scholarship, 2009

Involvement

- Virginia Tech Billiards Club President, Treasurer
- Virginia Tech Club Golf Team Member