COMPETENCIES

- Multivariate analysis
- Regression: Linear, Logistic, Multivariate
- Survey sampling & analysis techniques
- Agile Methodology & SDLC
- Exploratory and confirmatory factor analysis
- Trend analysis
- Predictive Modeling
- Experimental Design
- Categorical data analysis
- ETL on AWS EMR
- Structural equation modeling
- Frequent Pattern Mining
- Sentiment Analysis
- Natural Language Processing
- Topic Modeling (LDA)
- Robotic Process Automation
- Database Quality Assurance
- Jira

SUMMARY OF SKILLS, TOOLS, AND TECHNOLOGIES:

Programming Languages:

- SQL (5 yrs)
- Python (2 yrs)
- R (3 mo)
- SAS (3 yrs)
- PySpark (2 yrs)
- STATA (1 yr)

Software:

- SPSS (2 yrs)
- Tableau (2 yrs)
- SAS Enterprise Miner (1 yr)
- Apache Hue (1 yr)
- UiPath (5 mo)

Operating Systems:

- Unix (5 yrs)
- Linux (2 yrs)
- Windows (15 yrs)

EDUCATION

M.S., *Data Science*, **Oxford University**
B.S., *Business Management*, **University of New South Wales**

Rebecca Kumir

Data Scientist

Kumir_Rebecca@gmail.com	888 17th St, Nowhere, NY
+1 (888) 888-8888	Clearance: Public Trust

EXPERIENCE SUMMARY

Rebecca has developed extensive hands-on experience analyzing large data sets and developing automated analytics to assist key front line and executive staff in making critical business decisions, balanced with a graduate degree in quantitative analytics earned at a major research university. Her skills are most predominantly in predictive modeling, SDLC, natural language processing, trend analysis, machine-learning, and heuristic techniques to create descriptive, predictive, and prescriptive analytics, both object-oriented and non-object-oriented languages like SAS, SQL and Python, distributed computing in AWS EMR, database ETL management and other core analytic processes.

She also has a proven ability to effectively work across boundaries on a variety of teams developing actionable analytic solutions and implementing strategic organizational objectives using both Waterfall and Agile methodologies in a fast-paced commercial style consulting environment. She most recently led the analytics effort at the DOL launching a custom data analytics application designed to provide analysts with the tools to infer actionable insights from their data, automate processes, leverage distributed computing capabilities, and scalable database storage.

EXPERIENCE DETAILS

Outstand Company, Inc., *Data Scientist* 2/2016 – Present

Served as a reporting analyst creating a stream of interactive live data visualizations for the client's team whose overarching objective was to automate and streamline workflow processes and improve overall efficiency, quality, and consistency of decision making for several internal programs. The scope of the team's requirement is to implement a modern case management solution (CMS) built on top of the Appian-based Case Management Platform, conduct business analysis and process reengineering activities, integrate the solution with other systems, conduct all test planning & execution, train stakeholders, deploy the fully functioning/operational solution to the production environment, and then operate and maintain the solution afterwards. Tableau was the primary reporting tool used which was integrated with an Oracle database to create live reports cataloged in JIRA for multiple stakeholders.

Impressive Company, Inc., *Data Scientist* 2/2016 – Present

Analyzed historical audit data to identify trends and predictors of non-compliance, such as industry type, entity size, etc. Managed project goals, timelines, and client relationships; led weekly meetings informing senior clients of progress; coordinated with regression model that was built collaboratively

with project subcontractors to ensure project goals were achieved in unison. Developed predictive model to identify non-compliance in the audit population. The main deliverable was a logistic regression model that was developed with a training dataset of 66,000 data points of historical cases across 5 databases. Decision trees, ensemble methods and SMOTE analysis were also explored during the modeling process. Identified key risks in development and deployment of analytical model, and collaborated with team members to develop risk mitigation plans to ensure client concerns were addressed while maintaining project timeline. Developed model field testing deployment plan and automated case monitoring tools in preparation for full hand-off to Client personnel. Received performance award from Project Manager for exemplary performance.

Impressive University 12/2014 – 1/2016

- *Data Reporting Analyst | Office of Development:* Data reporting analyst for the campus-wide development/fund raising unit with approximately one million worldwide donors. Applied SAS Software to manage, extract, aggregate, and transform data from multiple internal sources to support all data-driven fundraising activities, including donor targeting, campaign results, optimal geographic location for university alumni/development events, and ad hoc reporting. Developed reports requested by senior university executives using SAS tools for reporting and results export to Excel-readable files. Implemented models predicting donation activity by alumni and others. Explained model results and applicability to development officers and executive level staff. Aggregated data from multiple sources to create and manage analytics and reports to support frontline fundraisers. Collaborated with database manager to report data updates and inconsistencies. Identified opportunities to improve internal processes and develop analytical solutions.
- *Market Segmentation Analysis Project Lead:* Developed a regression model from a database of just under one million cases across 395 available variables, designed to predict donor likelihood using nine comparative models. The objective of the project was to deliver a numerical rank for each constituent in the data base designed to provide a more strategic approach to major donor assignment management. Three different regression models were used across three varying population segments designed to maximize accuracy. The models will then be evaluated throughout the next fiscal year to determine its applicability and success. Summary and descriptive statistics were also generated to provide general insight into ideal donor profiles to target for outreach.