Marina Drus

Data Scientist/Data Analyst/Researcher - Independent Work Study

- Email me on Indeed: indeed.com/r/Marina-Drus/76dc8a32f9eabcfc

WORK EXPERIENCE

Data Scientist/Data Analyst/Researcher

Independent Work Study - New York, NY - July 2016 to Present

Analyze, clean, and interpret data using Python; prepare data visualization; build predictive models; consult other
data scientists on predictive modeling; peer review scientific journals.
Selected Science Projects:
• Explored and visualized the Billboard Hot 100 data to discover trends and relationships within various music genres
and songs using Matplotlib, Seaborn, and Tableau.
• Built profitability models for a new liquor storefront in the Iowa state given the liquor sales dataset from Iowa.gov.
• Created and trained machine learning models to predict survival rate among passengers of the legendary Titanic; used
complex imputation techniques and a synthetic minority oversampling technique to obtain precision/recall accuracy of
99%.
• Implemented clustering techniques such as PCA, k-means clusters, hierarchical clustering, and DBSCAN to identify
meaningful clusters of flight delays for FAA regulated airports in the US.
• Explored and executed various feature engineering techniques, synthetic minority oversampling techniques, and multiple machine learning algorithms to predict love at first sight in Speed Dating Kaggle Dataset; achieved validation accuracy of 97% and precision/recall accuracy of 100%.

Data Scientist/Data Analyst/Researcher

University of Kansas - Lawrence, KS - August 2011 to July 2016

Designed and built complex statistical models to predict scientific outcomes using R; analyzed and interpreted large
datasets with thousands of features;
• Originated and generated hypothesis testing to bring innovative findings to the field of social and cognitive science.
• Designed quantitative and qualitative surveys and experimental studies to support theories and hypotheses; conducted
focus groups, in-depth interviews, and observations; collected data; analyzed and interpreted data to provide answers to social science questions.
• Managed, prioritized, and balanced multiple research projects; researched and solved complex theoretical and
statistical problems; supervised, trained, and mentored a team of research assistants.
• Translated, explained, visualized, and presented statistical concepts to a wide range of audiences.
Selected Science Projects with Impact on Social Science:
• Applied deep learning and machine learning concepts to predict cognitive traits of highly creative people; was

featured in a scientific blog; the findings are used for educational purposes in a Nervous System and Behavior course.
• Built a statistical model of prejudice to predict significant changes in intergroup relations in the U.S. over the last 40
years; was honored by the National Science Foundation for the potential to benefit society.
• Established and identified links between paying taxes and happiness; was recognized by the Personality and Social
Psychology Society for contributing to understanding socio-economic behavior.

Data Analyst/Researcher
CUNY, Brooklyn College - Brooklyn, NY - September 2009 to August 2011

Financial/Accounting Specialist
Milbank, Tweed, Hadley, McCloy LLP - New York, NY - February 2003 to August 2011

EDUCATION

Ph.D. in Social Psychology
The University of Kansas

B.A. in Finance & Investment
Baruch College

B.S. in Economics
Petersburg State Transport University

SKILLS

clustering (1 year), data visualization (1 year), machine learning (6 years), predictive modeling (1 year), testing (4 years)

LINKS

https://dmarinav.github.io

https://github.com/dmarinav

ADDITIONAL INFORMATION

TECHNICAL SKILLS

Machine Learning: predictive modeling, classification, regression, clustering, feature engineering, re-sampling techniques, complex imputation techniques
Statistical Methods: inferential modeling, general linear models, time series, hierarchical models, hypothesis testing and
confidence intervals, dimensionality reduction, multivariate analyses, experimental research, A/B testing
Computer: Linux, Git, Python, R, SAS, Hadoop, Apache Spark, Map Reduce, AWS
Data Visualization: Tableau, GoogleVis, Shiny, Flask
Databases: MySQL, PostgreSQL