DANA ROCHA

dbridgette.rocha@gmail.com • (858) 353 – 4055 • linkedin.com/in/rochadana • github.com/dana-rocha

EDUCATION:

Northeastern University – Boston, MA
Master of Science in Bioinformatics, Graduate Certificate in Data Analytics (GPA: 3.6/4.0) Expected Dec. 2020
Relevant Courses: Algorithms, Computational Statistics, Statistics for Bioinformatics
Northeastern University – Boston, MA
Bachelor of Science in Biochemistry Aug. 2018

TECHNICAL SKILLS:

Languages: Python 3, R, SQL, HTML5, CSS
Operating Systems: OSX, Windows
Tools: PyCharm, Git, RStudio, Spotfire, Pipeline Pilot, DBeaver, Oracle SQL Developer

TECHNICAL EXPERIENCE:

Merck & Co. – Boston, MA Jan. 2020 – Aug. 2020
Modeling and Informatics Co-op – MRL Computational and Structural Chemistry

- Trained machine learning models and built workflows using Pipeline Pilot to predict new relationships between compounds, genes, and disease
- Queried public and internal databases using SQL to retrieve data for workflows
- Designed interactive dashboards in Spotfire for data analysis on over 150,000 compounds and targets
- Collaborated with scientists within Chemical Biology and Informatics and presented data insights
- Delivered three knowledge-based sets of compounds for phenotypic screening programs in Chemical Biology

RESEARCH EXPERIENCE:

Orig3n Inc. – Boston, MA Jul. 2017 – Dec. 2017
Research Assistant Co-op

- Maintained hiPSC, MSC, and RPE cell cultures for assays and animal studies
- Characterized cell cultures using fluorescence microscopy and flow cytometry
- Represented Orig3n as a brand ambassador at events in New York City, Baltimore, and Lake Tahoe

Brigham and Women's Hospital - **CCI Specimen Processing Lab** – Boston, MA Jun. 2016 – Dec 2016
Lab Technician Co-op

- Processed specimens for clinical studies conducted by the CCI Units and Emergency Department
- Completed all documentation required for individual specimens according to research protocol requirements
- Maintained laboratory instrumentation to performance standards

PROJECTS:

Mapping Manhattan by Distance to Nearest Subway Stations Jun. 2020 – Current

- Curated publicly available dataset based on coordinate data
- Created Python script to compute distance using the Pandas library and Numpy package
- Visualized data with an interactive choropleth map using the Plotly library in Python

College Student Demographics Over Time: Data Visualization Nov. 2019 – Dec. 2019

- Accessed the College ScoreCard API to retrieve demographic data using R
- Visualized patterns of demographic enrollment using the ggplot2 package in R
- Utilized SQL to store and query data in relational databases