**SEAN PILI**

**Phone:** 703-835-7263 **Email:** seanpili@gwu.edu **Address:** 131 Moncure Drive, Alexandria, VA 22314

# EDUCATION M.S. Data Science, Anticipated December 2020

|  |  |
| --- | --- |
|  | George Washington University, Washington, District of Columbia  **B.S. Statistics,** Awarded December 2017 |
|  | **Minors:** Actuarial Science and Communication |
|  | Virginia Tech, Blacksburg, VA |
|  | In Major G.P.A: 3.5 Overall G.P.A: 3.2 |
| **HONORS/**  **AWARDS:** | **1st Place, Fall Data Competition at Virginia Tech, *Blacksburg, VA*** Oct 30th-Nov 6th 2017   * Competed against 15 teams of 2-3 undergraduates * Aggregated health data from 30 files to build predictive models that accurately predicted patients’ cholesterol levels without blood-work information (for private, inexpensive screening purposes) * Used Elastic Net regression to deal with multicollinearity between the features of the final models.   **Mu Sigma Rho Honor Society,** Member, September 2017-Present |

# RELATED

# EXPERIENCE Reporting Analyst, *Merchant Link, Silver Spring, MD,* Jun 2018-Present

* Leveraged a summer internship into a part time position
* As an intern: generated and improved the quality of periodic KPI reports in Excel, updated pricing models, conducted market research to identify new business partners and performed a customer segmentation analysis to help the company better understand and service their customers.
* Currently: automating KPI reports in Python and performing a more in-depth customer segmentation analysis.

# Undergraduate Research, *Statistics Department, Virginia Tech*, Aug-Dec 2017

* Worked to develop hypothesis test to determine if a network contains core-periphery structure

**Undergraduate Research,** ***Virginia Tech Biocomplexity Institute,* *Arlington, VA,*** May-Aug 2017

* Participated in their Data Science for the Public Good program (DSPG) for honors credit and worked on three social science focused research projects, the most involved of which entailed finding the best method(s) to create a perturbed or synthetic dataset for sensitive, public policy data to be implemented by SCHEV.
* Used R to web-scrape, query, clean, aggregate, visualize and analyze data for all projects and more specifically used R’s synthpop package to create synthetic datasets.
* Used supervised and unsupervised Machine Learning Methods (in addition to frequency tables and summary statistics) to evaluate the usability of synthetic datasets.

# Product Classification Intern: *USPS, District of Columbia*, May-Aug-2016

* Updated their 2013-14 First-Class Mail (FCM) Product Plan for FY 2017
* Conducted and presented a feasibility analysis on the introduction of a Standard Mail Card Product
* Partly responsible for the re-branding of Standard Mail as Marketing Mail

**COMPUTER**

**SKILLS** R, Python, MySQL, Studio3T, Git Bash, Base SAS, Tableau, Minitab, JMP, and Microsoft Office