SOUMIK PURKAYASTHA

E-mail: soumikp@umich.edu **Phone:** +1-734-881-5075

EDUCATION

- **University of Michigan, Dept. of Biostatistics** *September 2019 -*
  Graduate student research assistant and Ph.D. student. Current GPA: 4.0+
- **University of Michigan, Dept. of Biostatistics** *September 2019 - April 2021*
  Master of Science in Biostatistics. GPA: 4.0+
- **Indian Statistical Institute** *July 2017 - June 2019.*
  Master of Statistics, First Class. Specialisation: Biostatistics. GPA: 4.0
- **St. Xavier's College (Autonomous), Kolkata** *July 2014 - June 2017.*
  Bachelor of Science (Hons.) in Statistics, First Class. GPA: 4.0

PROFESSIONAL EXPERIENCE

- **Walmart Labs, Bangalore, India** May 2018 - July 2018
  Statistical Analyst
  - Worked on data query and analysis of very large data sets. Built interactive apps using **RShiny**, with special emphasis on data visualisation by using Plotly.
  - Improved existing online grocery **forecasting models**. Built new forecasting models using both **traditional and state space models**
  - Built **ensemble classification models** to identify stores with unusual customer behaviour on basis of cross-sectional data.

- **Apple Inc., Cupertino, USA** May 2021 - August 2021
  Statistical Analyst
  - Developed statistical models to solve challenging analytical problems that help understand user behaviour
  - Build classifiers to predict user search intent, improve data quality for ML training and evaluation.
  - Considering both user behavior and human annotation data, develop appropriate analyses pipelines, iterate on different statistical methods, and incorporate new innovations into the Siri Search product.

COMPUTING SKILLS *Proficiency*

- **Language: R** advanced
- **Language: C++, Python and SQL** intermediate
- **Software: SAS and Minitab** intermediate

PROJECT EXPERIENCES

- **University of Michigan, Ann Arbor, MI.** September, 2019 - present.
  - *Song lab COVID-19 group* and *COV-IND-19 study group:* An interdisciplinary group of researchers who use data and modeling to generate timely reports and recommendations about COVID-19.
  - *MIDAS Data Science Challenge:* Analysis and inference of results from J.D. Power 2019 Auto Insurance Survey. **Best Project Award at the 2019 Michigan Data Science Challenge.**
- **Indian Statistical Institute, India** July 2017-June 2019.
  - *On minimum Bregman divergence estimation:* a robust estimation method developed by making use of Bregman divergences. **Best Masters Project Award at Indian Statistical Institute.**

- *Prediction of monthly Hilsa catch using time series models:* Bayesian structural time series-based method proposed to analyze and predict yield of Hilsa fish in estuarine region of West Bengal, India.
- *Shortest Path Problems: A review of various algorithms and some real world problems.* Application to arbitrage problems.
- *Multivariate Classification of Fraud in Swiss Bank Notes:* Used various classification techniques to classify Swiss bank notes as real or fake. Structural differences identified using Principal Component analysis and Factor analysis.

- **St. Xavier's College (Autonomous), Kolkata, India** July 2015-May 2017.
  - *Newsboy Problem: Optimal order quantity in presence of random supply.* Formulation of inventory management problem in context of random supply and demand. **This project received a Poster Presentation Award at the 103rd Indian Science Congress (2016)**
  - *Estimation of location parameter of symmetric distributions using sample quantiles.* Novel quantile-based estimator for non-symmetric distributions proposed and their performance evaluated using both real life and simulated data. **Best Dissertation in St. Xavier's College, Kolkata.**

PUBLICATIONS

2020

1. Debashree Ray, Maxwell Salvatore, Rupam Bhattacharyya, Lili Wang, Jiacong Du, Shariq Mohammed, **Soumik Purkayastha** et al. *Predictions, role of interventions and effects of a historic national lockdown in Indias response to the COVID-19 pandemic: data science call to arms.* Harv Data Sci Rev. 2020; 2020(Suppl 1): 10.1162/99608f92.60e08ed5

2. **Soumik Purkayastha**, Maxwell Salvatore, and Bhramar Mukherjee. *Are women leaders significantly better at controlling the contagion during the COVID-19 pandemic?.* Journal of health and social sciences 5, no. 2 (2020): 231: 10.1101/2020.06.06.20124487

3. Lu Tang, Yiwang Zhou, Lili Wang, **Soumik Purkayastha**, Leyao Zhang, Jie He, Fei Wang, and Peter XK. Song. *A Review of Multi-Compartment Infectious Disease Models.* International Statistical Review 88, no. 2 (2020): 462-513: 10.1111/insr.12402

4. Yiwang Zhou, Lili Wang, Leyao Zhang, Lan Shi, Kangping Yang, Jie He, Bangyao Zhao, William Overton, **Soumik Purkayastha**, and Peter Song. *A spatiotemporal epidemiological prediction model to inform county-level COVID-19 risk in the United States.* Harvard Data Science Review (2020): 10.1162/99608f92.79e1f45e

5. Sandip Giri, **Soumik Purkayastha**, Sugata Hazra, Abhra Chanda, Isha Das, and Sourav Das. *Prediction of monthly Hilsa (Tenualosa ilisha) catch in the Northern Bay of Bengal using Bayesian structural time series model.* Regional Studies in Marine Science 39 (2020): 101456: 10.1016/j.rsma.2020.101456

6. Ritwik Bhaduri, Ritoban Kundu, **Soumik Purkayastha**, Mike Kleinsasser, Lauren J. Beesley, and Bhramar Mukherjee. *Extending the Susceptible-Exposed-Infected-Removed (SEIR) model to handle the high false negative rate and symptom-based administration of COVID-19 diagnostic tests: SEIR-fansy.* Under review at Statistics in Medicine.

2021

7. Maxwell Salvatore, Deepankar Basu, Debashree Ray, Mike Kleinsasser, **Soumik Purkayastha**, Rupam Bhattacharyya, and Bhramar Mukherjee. *A comprehensive public health evaluation of lockdown as a non-pharmaceutical intervention on COVID-19 spread in India: National trends masking state level variations.* BMJ Open (2020): 10(12): 10.1136/bmjopen-2020-041778

8. **Soumik Purkayastha**, Rupam Bhattacharyya, Ritwik Bhaduri, Ritoban Kundu, Xuelin Gu, Maxwell Salvatore, Swapnil Mishra, and Bhramar Mukherjee. *A comparison of five epidemiological models for*

transmission of SARS-CoV-2 in India. BMC Infect Dis 21, 533 (2021): 10.1186/s12879-021-06077-9

9. **Soumik Pukayastha**, Ritoban Kundu, Ritwik Bhaduri et. al. *Estimating the wave 1 and wave 2 infection fatality rates from SARS-CoV-2 in India.* BMC Res Notes 14, 262 (2021). 10.1186/s13104-021-05652-2

10. Lauren Zimmermann, Subarna Bhattacharya, **Soumik Purkayastha**, Ritoban Kundu, Ritwik Bhaduri, Parikshit Ghosh, and Bhramar Mukherjee. *SARS-CoV-2 Infection Fatality Rates in India: Systematic Review, Meta-Analysis and Model-Based Estimation.* Studies in Microeconomics 9, no. 2, December 2021. 13779. https://doi.org/10.1177/23210222211054324.

11. **Soumik Purkayastha** and Ayanendranath Basu. *On minimum Bregman divergence inference.* arXiv preprint arXiv:2008.06987 (2020). Under review at Metrika.

12. Maxwell Salvatore, Rupam Bhattacharyya, **Soumik Purkayastha**, Lauren Zimmermann, Debashree Ray, Aditi Hazra, Michael Kleinsasser, Thomas Mellan, Charlie Whittaker, Seth Flaxman, Samir Bhatt, Swapnil Mishra, Bhramar Mukherjee. *Resurgence of SARS-CoV-2 in India: Potential role of the B.1.617.2 (Delta) variant and delayed interventions.* Under review at Science.

AWARDS

2020 (a) University of Michigan, Department of Biostatistics - Richard G. Cornell Fellowship for outstanding academic performance.

2019 (a) University of Michigan, Michigan Institute of Data Science - Best project at Michigan Data Science Challenge.

(b) Indian Statistical Institute - Sabyasachi Roy Memorial Gold Medal for the best Master's degree project.

2018 (a) Indian Statistical Institute - Scholarship for good academic performance.

2017 (a) Indian Statistical Institute - Scholarship funded by Government of India.

(a) St. Xavier's College - Best Bachelor's degree dissertation.

2016 (a) Indian Science Congress Association - Poster Presentation Award.