GASP 2020

November 6, 2020, 9:00 am- 6:00 pm Eastern Time

For the online version of this program, see: https://sites.google.com/view/gasp2020

9:00am- 9:05am   Welcome and introductions
Chairperson: Mike Jadoo

9:05am- 9:20am   "Networking effectively in a virtual environment" Robert Brandau- Career Confidence

9:25am-10:05 am   Keynote speaker- Scott Gibbons

Chief Data Officer at Department of Labor

Watch the session recording on YouTube

10:10am-11:30am   Session 1 Regular presentations
Theme: Use cases in Data Visualization and Big data

Chairperson: S. Gwynn Sturdevant

10:10am-10:30am "Visualizing the Impact of COVID-19 on WIC" , *Kelsey Gray* - Insight Policy Research [CODE]

10:30am-10:50am "Modern Data Science Techniques for Visualization", *James Cutler*-
U.S. Army [PRESENTATION]

10:50am-11:10am "Analyze Global COVID-19 data with data visualization", *Mythili Krishnan*
[CODE] [PRESENTATION]

11:10am-11:30am "Profiling statistical programming language performance for largescale data tasks", *Cecile Murray*- US Census [CODE and Material]

Watch the session recording on YouTube

11:33 am -12:33pm Lightning round 1

Chairperson: Mitchell Shuey

11:33am-11:43am "Augmented CPS data on industry and occupation", *Peter Meyer*- Bureau of Labor Statistics, *Kendra Asher*, Bureau of Labor Statistics [PRESENTATION]

11:43am-11:53am "APIs: you're probably not using them and why you probably should", *Josiah Parry*- RStudio [PRESENTATION]

11:53am-12:03pm "Online Winsorization of Construction Price Statistics", *Charles Coleman* -US Census Bureau [CODE] [PRESENTATION]

12:03pm-12:13pm " Libraries and E-Rate Funding: How I Built a Data Pipeline in R from a Scheduled R Script on AWS to a Shiny Dashboard", *Bree Norlander* -University of Washington Technology & Social Change (TASCHA) Group , Chris Jowaisas-University of Washington Technology [CODE and SLIDES]

12:13pm-12:23pm "A picture is worth a thousand labor hours: pinpointing efficiency opportunities in the processing of data from large-scale, complex surveys.", *Lauren Warren* - RTI International

12:23pm-12:33pm "Crop classification and uncertainty assessment at 10 m resolution using Google Earth Engine", *Luca Sartore* -USDA NASS

Watch the session recording on YouTube

12:33pm-1:05pm   Lunch

1:05pm- 3:05pm   Session 2 Regular presentations

Theme: Artificial Intelligence and Natural Language Processing techniques

Chairperson: Peter Meyer

1:05pm-1:25pm "Building Neural Networks in R with External Machine Learning Engines", *Philip Waggoner* - University of Chicago [CODE]

1:25pm-1:45pm "Using Named Entity Recognition to Identify Substances used in SelfMedication of Opioid Withdrawal" , *Peter Baumgartner*- RTI International [PRESENTATION]

1:45pm-2:05pm"From Natural Language Processing and Graph Data Models to Knowledge Graphs", *Chris Marokov*- Office of Management and Budget (OMB), Executive Office of the President [PRESENTATION]

2:05pm-2:25pm"Detecting Pharmaceutical Innovations in News Articles and Financial Filings", *Gizem Korkmaz*- Biocomplexity Institute, University of Virginia , Devika Mahoney-Nair- Biocomplexity Institute & Initiative, University of Virginia , Neil Alexander Kattampallil- Biocomplexity Institute & Initiative, University of Virginia, Gary Anderson -National Center for Science and Engineering Statistics, National Science Foundation [PRESENTATION]

2:25pm-2:45pm"Adopting prior distributions for the variance-covariance matrices to fully specify an area-level bivariate hierarchical Bayes three-fold model for proportions of adult competency",*Andreea Erciulescu*- Westat [PRESENTATION]

2:45pm-3:05pm"Tornado Property Loss Scale: Up to $8 Billion by 2025 Classification, Dependence, and Prediction of Tornado Events in the U.S." ,*Thilini Mahanama*- Texas Tech University [PRESENTATION]

Watch the session recording on YouTube

3:10 - 4:20pm   Lightning round 2

Chairperson: Brandon Kopp

3:10pm-3:20pm "Easily Automating R Scripts on Windows", *James Cutler*- U.S. Army[CODE] [PRESENTATION] [CODE2]

3:20pm-3:30pm "Using Julia for Dissemination of Information About the General Election in Puerto Rico", *José Bayoán Santiago Calderón*- Biocomplexity Institute & Initiative, University of Virginia [CODE] [PRESENTATION]

3:30pm-3:40pm "Toward an Integrated Record Linkage Environment at the Census Bureau", *Yves Thibaudeau*- US Census Bureau [PRESENTATION]

3:40pm-3:50pm"When a Dashboard Won't Do: Automating Static Data Visualization of COVID-19 Data in Python", *Alexander Preiss*- RTI International [PRESENTATION]

3:50pm-4:00pm"Work Stoppages Web Scrape Project", *Drake Gibson*- Bureau of Labor Statistics, David Oh-Bureau of Labor Statistics [PRESENTATION]

4:00pm-4:10pm"Automatic Production of Large Spatial Price Rasters with Python", *Kevin Hunt*- USDA NASS

4:10pm-4:20pm"Using Deep Learning to Classify Open Text Field Comments in a Panel Study", *Kristin Chen*-Westat [PRESENTATION]

Watch the session recording on YouTube

4:25 -5:50pm   Session 3 regular presentations

Theme: Topics in DevOps and Machine Learning

Chairperson: Wendy Martinez

4:25pm-4:45pm"Data Catalog with Machine-Learning based engines for Statistical Programming", *Pragyansmita Nayak*- Hitachi Vantara Federal [PRESENTATION]

4:45pm-5:05pm"Automated Revenue Prediction Modeling", *Thom Miano*- RTI International [PRESENTATION]

5:05pm-5:25pm"Adaptive Computational Reproducibility in a Shifting R Package Landscape", *Kiegan Rice*- NORC at the University of Chicago ,Dr. Heike Hofmann, Iowa State University [PRESENTATION]

5:25pm-5:45pm"Deploying R Models to Implement REST Production Endpoints", *Frank Font*- FINRA, Yan Cheng-FINRA, David Llanos-FINRA [PRESENTATION]

Watch the session recording on YouTube

5:50pm - 6:00pm Closing and Knowledge Harvesting

Facilitator: Mike Jadoo