



# Conference Schedule

View past PyData event schedules here (<https://pydata.org/past-events.html>).

## General Sessions – Monday Nov. 4, 2019

	Central Park West (6501)	Central Park East (6501a)	Winter Garden (5412)	Belasco (6203)
8:00 AM		Breakfast & Registration		
9:00 AM		Opening Remarks		
9:15 AM		Keynote (/nyc2019/schedule/presentation/62/keynote-kelly-jin/) Kelly Jin		
10:05 AM	The Secret Life of Python (/nyc2019/schedule/presentation/secret-life-of-python/) Steve Dower	Time series for scikit-learn people (/nyc2019/schedule/presentation/time-series-for-scikit-learn-people/) Ethan Rosenthal	Bringing mental health data to doctors (/nyc2019/schedule/presentation/mental-health-data-to-doctors/) Bill Lynch	Unconference
10:45 AM		Coffee Break		
10:55 AM	Pandas vs Koalas: The Ultimate Showdown! (/nyc2019/schedule/presentation/pandas-vs-koalas-the-ultimate-showdown/) Amanda Moran	A Crash Course in Applied Linear Algebra (/nyc2019/schedule/presentation/crash-course-in-applied-linear-algebra/) Patrick Landreman	Working with Maps: Extracting Features for Traffic Crash Insights (/nyc2019/schedule/presentation/with-maps-extracting-features-for-traffic-crash-insights/) Jenny Turner-Trauring	Unconference
11:40 AM	Scalable Machine Learning with Dask (/nyc2019/schedule/presentation/machine-learning-with-dask/) Tom Augspurger	Small Big Data: using NumPy and Pandas when your data doesn't fit in memory (/nyc2019/schedule/presentation/big-data-using-numpy-and-pandas-when-your-data-doesnt-fit-in-memory/) Itamar Turner-Trauring	Painting A Picture of Public Data (/nyc2019/schedule/presentation/a-picture-of-public-data/) Kamal Abdelrahman	
12:20 PM		Lunch		

	Central Park West (6501)	Central Park East (6501a)	Winter Garden (5412)	Belasco (6203)
1:20 PM	Launching a new warehouse with SimPy at Rent the Runway (/nyc2019/schedule/presentations-a-new-warehouse-with-simpy-at-rent-the-runway/) Meghan Heintz	Pandas house-keeping and optimization with numba (/nyc2019/schedule/presentations-house-keeping-and-optimization-with-numba/) Diego Torres Quintanilla	Conda-press, or Reinventing the Wheel (/nyc2019/schedule/presentations-press-or-reinventing-the-wheel/) Anthony Scopatz	
2:05 PM	The physics of deep learning using tensor networks (/nyc2019/schedule/presentations-physics-of-deep-learning-using-tensor-networks/) Marianne Hoogeveen	Propensity Score Matching: A Non-experimental Approach to Causal Inference (/nyc2019/schedule/presentations-score-matching-a-non-experimental-approach-to-causal-inference/) Michael Johns	Effective Python and R collaboration (/nyc2019/schedule/presentations-python-and-r-collaboration/) Daniel Rodriguez	Unconference
2:50 PM	Is Spark still relevant? Multi-node CPU and single-node GPU workloads with Spark, Dask and RAPIDS. (/nyc2019/schedule/presentations-spark-still-relevant-multi-node-cpu-and-single-node-gpu-workloads-with-spark-dask-and-rapids/) Eric Dill	Same API, Different Execution (/nyc2019/schedule/presentations-api-different-execution/) Saul Shanabrook	Using Graph Nets (GNNs) to predict molecular properties (/nyc2019/schedule/presentations-graph-nets-gnn-to-predict-molecular-properties/) Chaya D Stern, Yuanding Wang	
3:30 PM	Break			
3:40 PM	Quantifying uncertainty in machine learning models (/nyc2019/schedule/presentations-uncertainty-in-machine-learning-models/) Samuel Rochette	Generating realistic, differentially-private data sets using GANs (/nyc2019/schedule/presentations-realistic-differentially-private-data-sets-using-gans/) Joshua Falk	Simplified Data Quality Monitoring of Dynamic Longitudinal Data: A Functional Programming Approach (/nyc2019/schedule/presentations-data-quality-monitoring-of-dynamic-longitudinal-data-a-functional-programming-approach/) Jacqueline Gutman	
4:25 PM	Spark Backend for Ibis: Seamless Transition Between Pandas and Spark (/nyc2019/schedule/presentations-backend-for-ibis-seamless-transition-between-pandas-and-spark/) Li Jin, Hyonjee Joo	Build an AI-powered Pet Detector in Visual Studio Code (/nyc2019/schedule/presentations-an-ai-powered-pet-detector-in-visual-studio-code/) Katherine Kampf	Geo Experiments and CausalImpact in Incrementality Testing (/nyc2019/schedule/presentations-experiments-and-causalimpact-in-incrementality-testing/) Jessica Tyler	Unconference
5:10 PM	What we learned by running a large custom Bayesian forecasting model in production (/nyc2019/schedule/presentations-we-learned-by-running-a-large-custom-bayesian-forecasting-model-in-production/) Jens Fredrik Skogstrom	Semantic modeling of data science code (/nyc2019/schedule/presentations-modeling-of-data-science-code/) Evan Patterson	Zarr vs. HDF5 (/nyc2019/schedule/presentations-vs-hdf5/) Joe Jevnik	
5:50 PM	Social Reception			
8:30 PM				

# General Sessions – Tuesday Nov. 5, 2019

	Central Park West (6501)	Central Park East (6501a)	Winter Garden (5412)	Music Box (5411)	Ambassador (6202)	Belasco (6203)
8:00 AM				Breakfast & Registration		
9:15 AM				Keynote (/nyc2019/schedule/presentation/63/keynote-bauer-wiggins/) Chris Wiggins		
10:05 AM	Deep Dive into scikit-learn's HistGradientBoostingClassifier and Regressor (/nyc2019/schedule/dive-into-scikit-learns-histgradientboosting-and-regressor/) Thomas J Fan	Julia for Pythonistas (/nyc2019/schedule/for-pythonistas/) Kelly Shen	Improve the efficiency of your Big Data application (/nyc2019/schedule/the-efficiency-of-your-big-data-application/) Francesc Alted, Christian Steiner		Fireside Chat (/nyc2019/schedule/chat-bauer-wiggins/) Chris Wiggins	
10:45 AM				Coffee Break		
10:55 AM	Every ML Model Deserves To Be A Full Micro-service (/nyc2019/schedule/ml-model-deserves-to-be-a-full-micro-service/) Romain Cledat	Dealing With Imbalanced Classes in Machine Learning (/nyc2019/schedule/with-imbalanced-classes-in-machine-learning/) Aditya Lahiri	TBA		Unconference	Introduction to Bayesian Modeling with Stan: No Statistics Background Required (Pt 1) (/nyc2019/schedule/to-bayesian-modeling-with-stan-no-statistics-background-required-pt-1/) Breck Baldwin
11:40 AM	Clean Machine Learning Code: Practical Software Engineering Principles for ML Craftsmanship (/nyc2019/schedule/machine-learning-code-practical-software-engineering-principles-for-ml-craftsmanship/) Moussa Taifi Ph.D.	Discover your latent food graph with this 1 weird trick (/nyc2019/schedule/your-latent-food-graph-with-this-1-weird-trick/) Alex Egg, Emily A Ray, Parin Choganwala	Colorism in High Fashion (featuring: K-Means Clustering) (/nyc2019/schedule/in-high-fashion-featuring-k-means-clustering/) Malaika Handa	High-Performance Data Science at Scale with RAPIDS, Dask, and GPUs (/nyc2019/schedule/performance-data-science-at-scale-with-rapids-dask-and-gpus/) Keith Kraus	Unconference	
12:20 PM				Lunch		
1:20 PM				Keynote (/nyc2019/schedule/presentation/64/keynote-sara-seager/) Sara Seager		

	Central Park West (6501)	Central Park East (6501a)	Winter Garden (5412)	Music Box (5411)	Ambassador (6202)	Belasco (6203)
2:10 PM	Building a maintainable plotting library (/nyc2019/schedule/a-maintainable-plotting-library/) Colin Carroll, Hannah Aizenman, Thomas Caswell	Type-Driven Automated Learning with Lale (/nyc2019/schedule/driven-automated-learning-with-lale/) Martin Hirzel	Data-centric exploration using intake, dask, hvplot, datashader, panel, and binder (/nyc2019/schedule/centric-exploration-using-intake-dask-hvplot-datashader-panel-and-binder/) Julia Signell	Genetic algorithms: Making errors do all the work (/nyc2019/schedule/algorithms-making-errors-do-all-the-work/) Raman Tehlan	Unconference	Introduction to Bayesian Modeling with Stan: No Statistics Background Required (Pt 2) (/nyc2019/schedule/to-bayesian-modeling-with-stan-no-statistics-background-required-pt-2/) Breck Baldwin
2:55 PM	tf-explain: Interpretability for Tensorflow 2.0 (/nyc2019/schedule/explain-interpretability-for-tensorflow-20/) Raphaël Meudec	Implementing Lightweight Random Indexing for Polylingual Text Classification (/nyc2019/schedule/lightweight-random-indexing-for-polylingual-text-classification/) Ian Whalen	Should I develop my own DS library? Maybe. (/nyc2019/schedule/i-develop-my-own-ds-library-maybe/) Piero Ferrante	How and why to put your Jupyter notebooks in Docker containers (/nyc2019/schedule/and-why-to-put-your-jupyter-notebooks-in-docker-containers/) Brian Austin	Unconference	
3:35 PM Break						
3:45 PM	Free Your Esoteric Data Using Apache Arrow and Python (/nyc2019/schedule/your-esoteric-data-using-apache-arrow-and-python/) Maciej Wojton	The Inspection Paradox is Everywhere (/nyc2019/schedule/inspection-paradox-is-everywhere/) Allen Downey	The Echo-Chamber of Your Social Media Feed (/nyc2019/schedule/echo-chamber-of-your-social-media-feed/) Tamar Yastrab	Managing Stakeholders: The Key to a Successful Data Science Project (/nyc2019/schedule/stakeholders-the-key-to-a-successful-data-science-project/) Lauren Oldja	Unconference	Introduction to Bayesian Modeling with Stan: No Statistics Background Required (Pt 3) (/nyc2019/schedule/to-bayesian-modeling-with-stan-no-statistics-background-required-pt-3/) Breck Baldwin
4:30 PM	Reproducibility in ML Systems: A Netflix Original (/nyc2019/schedule/in-ml-systems-a-netflix-original/) Ferras Hamad	TBD (/nyc2019/schedule/James Powell	A How-to guide for migrating legacy data applications (/nyc2019/schedule/how-to-guide-for-migrating-legacy-data-applications/) Marius van Niekerk, Rohit Kapur	Building Software and Communities With Peer Review (/nyc2019/schedule/software-and-communities-with-peer-review/) Noam Ross	Unconference	
5:15 PM	Lightning Talks					
6:00 PM	Closing Remarks					
6:30 PM						

## Tutorial Sessions — Wednesday Nov. 6, 2019

	Winter Garden (5412)	Music Box (5411)	Radio City (6604)	Broadway (5202)	Ambassador (6202)
8:00 AM		Curated Track: Statistics	Curated Track: NLP		
8:00 AM			Breakfast & Registration		
9:00 AM	Introduction to pandas (/nyc2019/schedule/pres-to-pandas/) Marc Garcia, Jeff Reback, Tom Augspurger	An Introduction to Probability and Statistics (/nyc2019/schedule/pres-introduction-to-probability-and-statistics/) Will Kurt	Introduction to NLP (/nyc2019/schedule/pres-to-nlp/) Mariel Frank	Visualizing the 2019 Measles Outbreak in NYC (with Python) (/nyc2019/schedule/pres-the-2019-measles-outbreak-in-nyc-with-python/) Carlos Afonso	HoloViz sprint
10:30 AM			Coffee Break		
10:45 AM	Advanced Software Testing for Data Scientists (/nyc2019/schedule/pres-software-testing-for-data-scientists/) Raoul-Gabriel Urma	How to Prove You're Right: A/B Testing with SciPy (/nyc2019/schedule/pres-to-prove-youre-right-ab-testing-with-scipy/) Hillary Green-Lerman, Michoel Snow	Introduction to Language Modeling (/nyc2019/schedule/pres-to-language-modeling/) Aditi Khullar, Eugene Tang	Swiftly turn Jupyter notebooks into pretty web apps (/nyc2019/schedule/pres-turn-jupyter-notebooks-into-pretty-web-apps/) Michal Mucha	HoloViz sprint
12:15 PM			Lunch		
1:15 PM	Machine learning from scratch using the scientific Python stack (/nyc2019/schedule/pres-learning-from-scratch-using-the-scientific-python-stack/) Lara Kattan	New Trends in Estimation and Inference (/nyc2019/schedule/pres-trends-in-estimation-and-inference/) Cameron Davidson-Pilon	Neural Networks for Natural Language Processing (/nyc2019/schedule/pres-networks-for-natural-language-processing/) Matti Lyra	From Raw Recruit Scripts to Perfect Python (in 90 minutes) (/nyc2019/schedule/pres(raw-recruit-scripts-to-perfect-python-in-90-minutes/) Stanley van der Merwe, Petr Wolf	Pandas sprint
2:45 PM			Break		
3:00 PM	Hacking the Data Science Challenge (/nyc2019/schedule/pres-the-data-science-challenge/) Michoel Snow, Hillary Green-Lerman	Bayesian Inference for Fun and Profit (/nyc2019/schedule/pres-inference-for-fun-and-profit/) Mitzi Morris	Role playing Annotation workshop (/nyc2019/schedule/pres-playing-annotation-workshop/) Agata Sumowska, Bhargav Srinivasa Desikan, Laurence Warner, Lev Konstantinovskiy	A Primer on Gaussian Processes for Regression Analysis (/nyc2019/schedule/pres-primer-on-gaussian-processes-for-regression-analysis/) Chris Fonnesbeck	Pandas sprint
4:30 PM			Wrap Up		
5:00 PM					

Subscribe to Receive PyData Updates

## CODE OF CONDUCT

PyData is dedicated to providing a harassment-free conference experience for everyone, regardless of gender, sexual orientation, gender identity and expression, disability, physical appearance, body size, race, or religion.

We do not tolerate harassment of conference participants in any form. All communication should be appropriate for a professional audience including people of many different backgrounds. Sexual language and imagery is not appropriate for any conference venue, including talks. Be kind to others. Do not insult or put down other attendees. Behave professionally. Remember that harassment and sexist, racist, or exclusionary jokes are not appropriate for PyData. Attendees violating these rules may be asked to leave the conference without a refund at the sole discretion of the conference organizers. Thank you for helping make this a welcoming, friendly event for all.

## CONTACT



## ACCOUNT

[Login \(/nyc2019/accounts/login/\)](#)

[Sign up \(/nyc2019/accounts/signup/\)](#)

PYDATA IS AN EDUCATIONAL PROGRAM OF

# NUMFOCUS

## OPEN CODE = BETTER SCIENCE

(<https://numfocus.org>)

NumFOCUS is committed to protecting your data. In compliance with the new data regulations, we have updated our Privacy Policy (<https://numfocus.org/privacy-policy>) and invite you to review our new practices. We have also updated PyData registration with new data collection procedures.