

# Gabriela de Queiroz

@gdequeiroz | [linktr.ee/gdq](https://linktr.ee/gdq)

slides and material: [bit.ly/kroz-talks](https://bit.ly/kroz-talks)

**Gabriela de Queiroz** is a *Sr. Engineering & Data Science Manager at IBM*. She manages and leads a team of developers working on Data & AI Open Source projects. She works to democratize AI by building tools and launching new open source projects. She is passionate about making data science available to everybody and is actively involved with several organizations to foster an inclusive community. *She is the founder of AI Inclusive*, a global organization that is helping increase the representation and participation of gender minorities in Artificial Intelligence. She is also the *founder of R-Ladies*, a worldwide organization for promoting diversity in the R community with more than 45 countries. She has worked in several startups and where she built teams, developed statistical models, and employed a variety of techniques to derive insights and drive data-centric decisions. She likes to mentor and share her knowledge through mentorship programs, tutorials, and talks.

December 2, 2020 Wednesday, 3pm

Contact person: Meryem Gurel [mgurel@usfca.edu](mailto:mgurel@usfca.edu) zoom link: <https://usfca.zoom.us/j/5457579593>

WOMEN & DIVERSITY IN ECONOMICS CLUB



# Gabriela de Queiroz

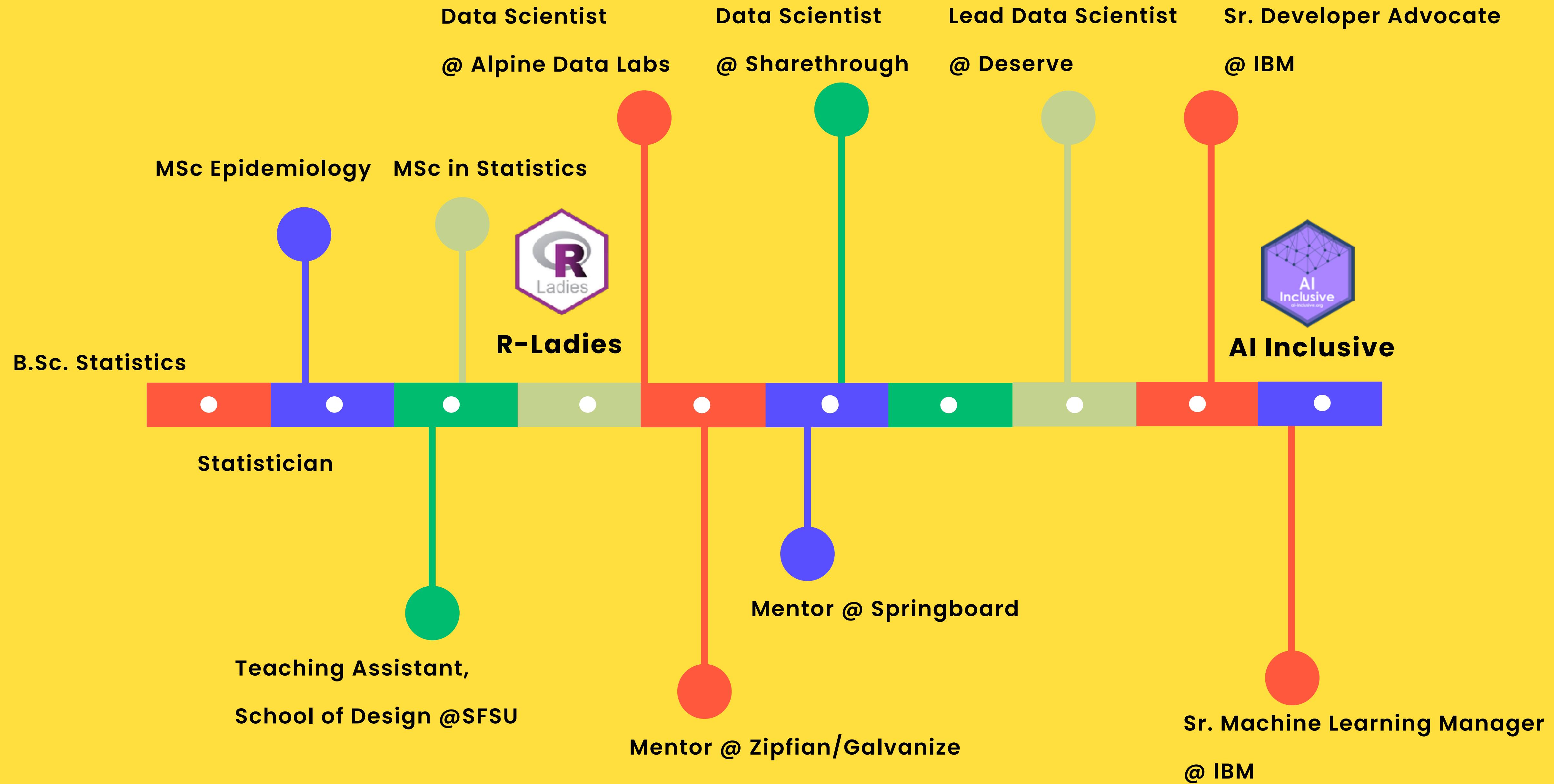
Sr. Machine Learning Manager, IBM

- Founder of **R-Ladies** ([rladies.org](http://rladies.org))
- Founder of **AI Inclusive** ([ai-inclusive.org](http://ai-inclusive.org))
- Member of **R Foundation** ([r-project.org](http://r-project.org))

- B.S. in Statistics
- MSc. in Epidemiology
- MSc. in Statistics



**Data Scientist + Developer Advocate + Open Source Developer + Manager +  
Statistician + Epidemiologist + Community Builder + Mentor + Speaker + Educator**





It was founded in October 2012.

The idea was to give back to the community and create a place where people would feel comfortable, safe and welcome.

A place where people could ask questions, learn together and share.



Wednesday, October 31, 2012  
31 OCT

## Introduction to R (beginners and pre-beginners)



Hosted by  
Gabriela de Queiroz

### Details

Hello R-ladies!

The first meetup will take place on October 31st at the Google office in San Francisco.

For this first meetup, we'll do an introduction to R. We'll go over the following topics:

installing R setting up an R environment (RStudio) basic commands (open files, simple dataset manipulation, simple plots, etc) loading packages the help function and how to read its output

All you need is your laptop and charger.

We look forward to seeing you!

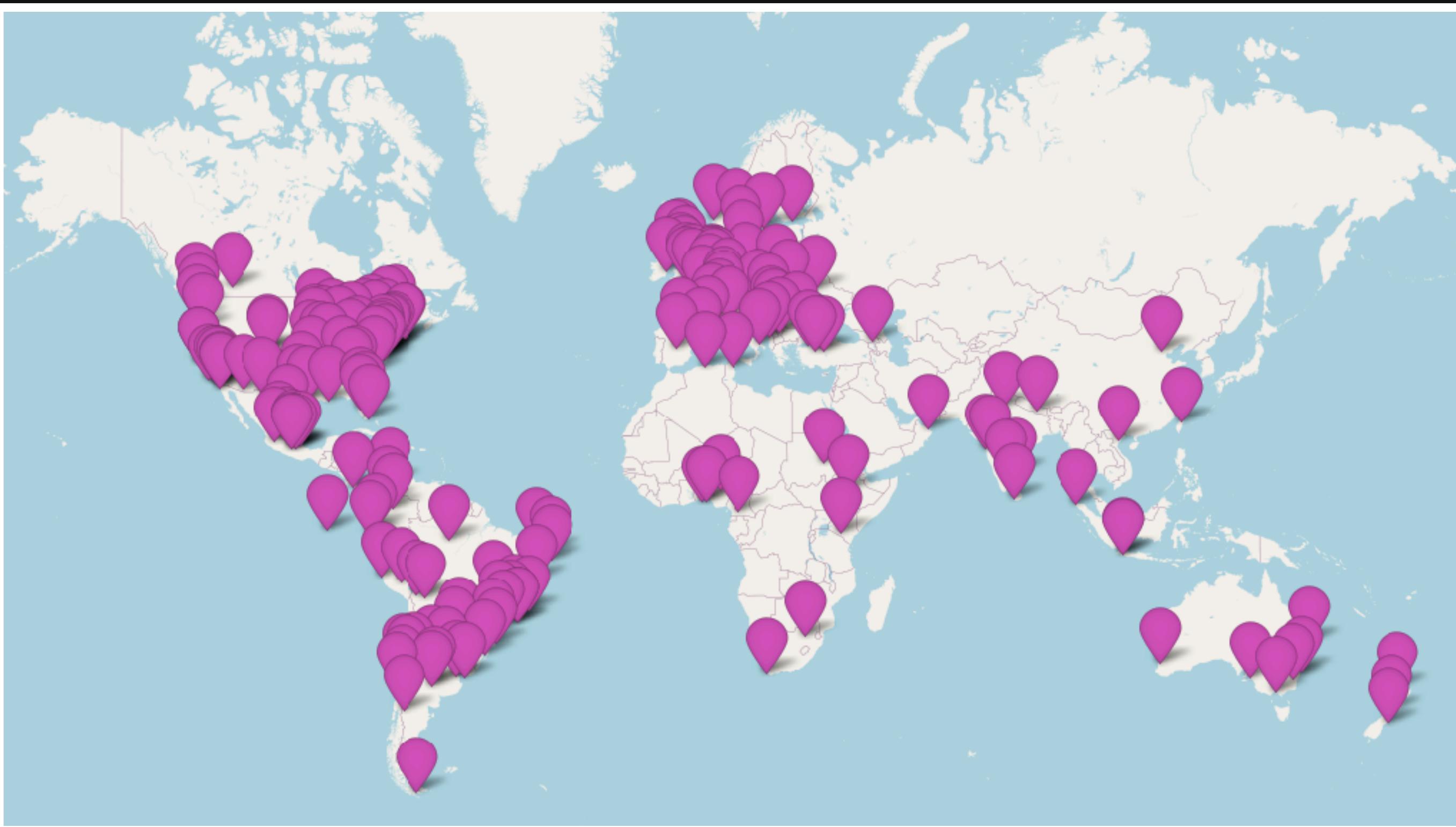


# R-Ladies

[rladies.org](http://rladies.org)



Worldwide organization that promotes diversity in the R community via meetups and mentorship in a friendly and safe environment





# AI Inclusive

**Mission:** Increase the representation and participation of minority groups in Artificial Intelligence

If you want to start a chapter, send us an email: [info@ai-inclusive.org](mailto:info@ai-inclusive.org)

Together, we are building a community to make **AI** more **inclusive** to everyone.

- Website: [ai-inclusive.org](http://ai-inclusive.org)
- Twitter: [bit.ly/ai-inclusive-twitter](https://bit.ly/ai-inclusive-twitter)
- Facebook: [bit.ly/ai-inclusive-facebook](https://bit.ly/ai-inclusive-facebook)
- Instagram: [bit.ly/ai-inclusive-instagram](https://bit.ly/ai-inclusive-instagram)
- Youtube: [bit.ly/ai-inclusive-youtube](https://bit.ly/ai-inclusive-youtube)



# Follow US:

[bit.ly/ai-inclusive-instagram](https://bit.ly/ai-inclusive-instagram)

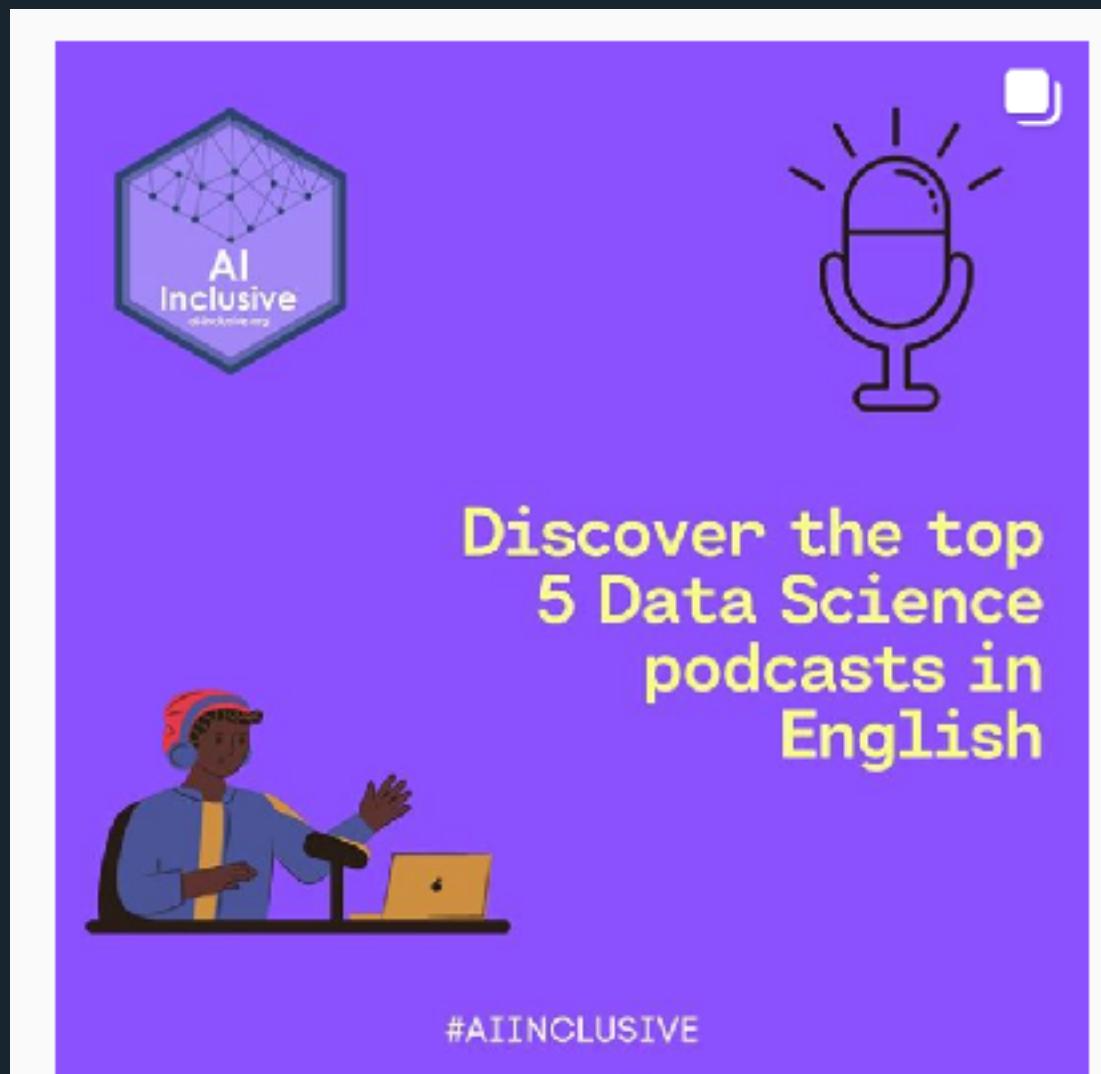
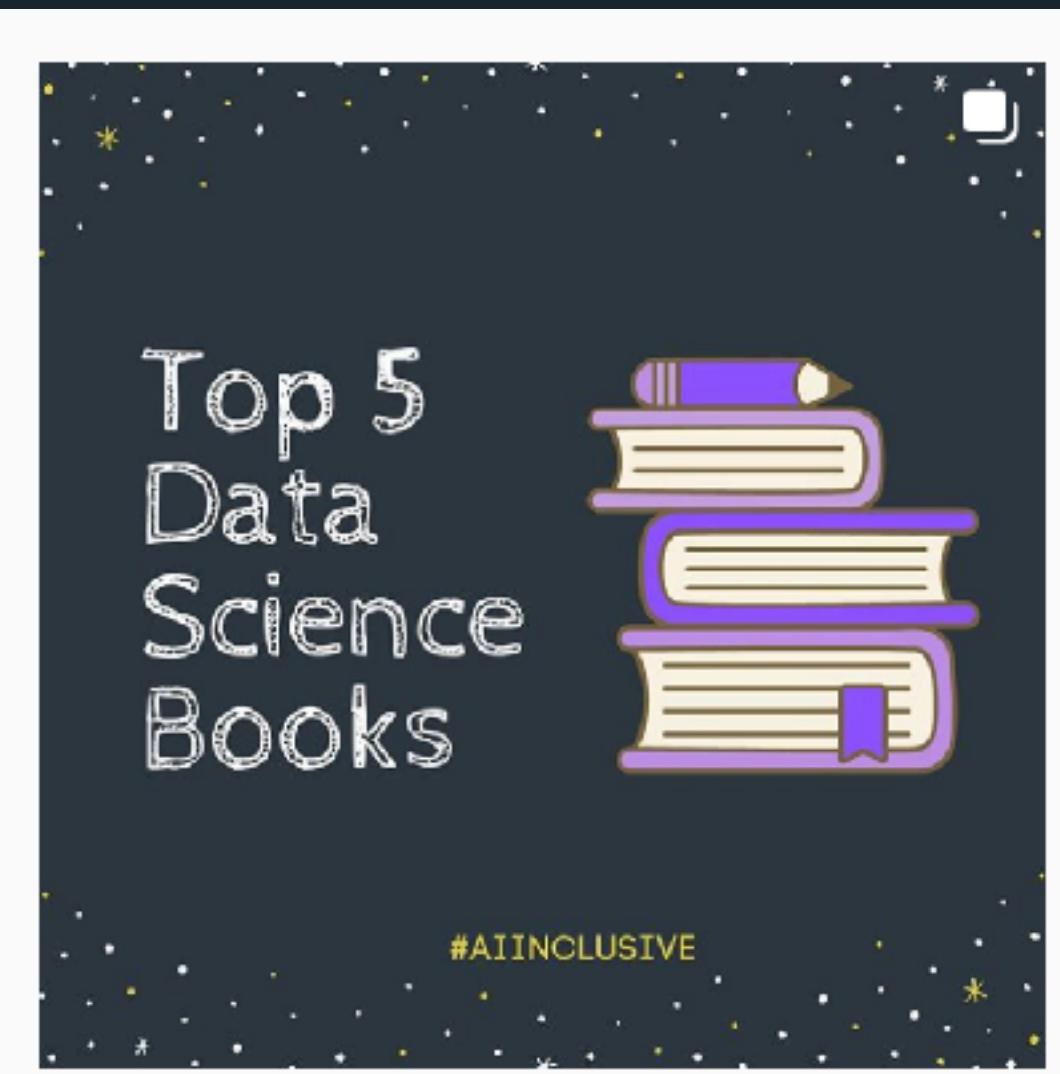
[bit.ly/ai-inclusive-twitter](https://bit.ly/ai-inclusive-twitter)



ai-inclusive.org

Resources on AI, DS, ML

Events, Free Tickets and much more



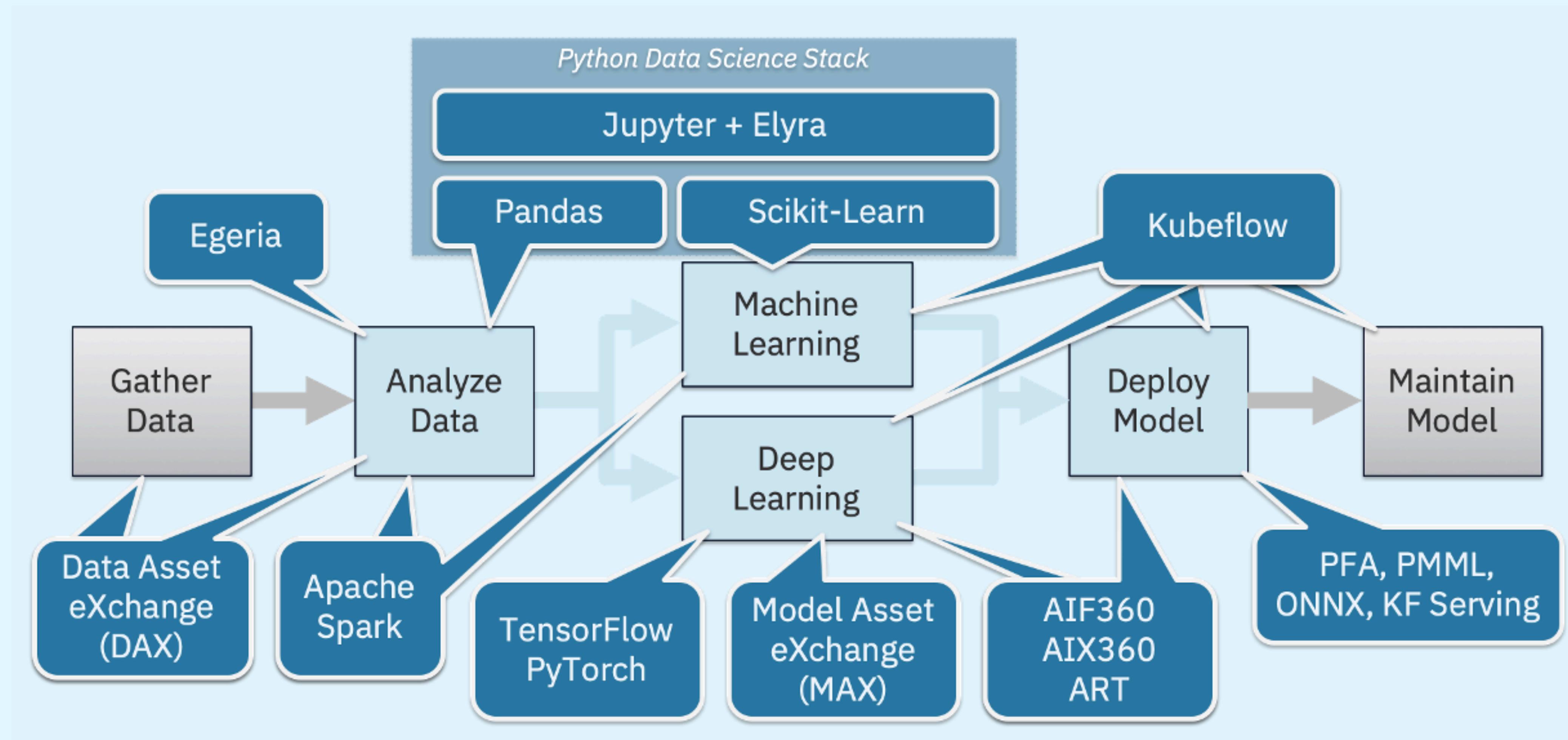
Soon:

Scholarships for Data Science Courses



# We build tools to make AI accessible and available to everybody

([codait.org](http://codait.org))





Open Source @ IBM

# Some Projects

# Model Asset eXchange



[ibm.biz/model-exchange](http://ibm.biz/model-exchange)

## Model Asset eXchange

Free, deployable, and trainable code. A place for developers to find and use free and open source deep learning models.

Try the tutorial



Join the community



[Featured](#) [Deployable](#) [Trainable](#)

Model | Deployable

Toxic Comment Classifier

Detect 6 types of toxicity in user comments

Jun 04, 2019



Model | Deployable, Trainable

Text Sentiment Classifier

Detect the sentiment captured in short pieces of text

Mar 29, 2019



Model | Deployable, Trainable

Image Segmente

Identify objects in an image, additionally assigning each pixel of the image to a particular object.

Sep 21, 2018



Model | Deployable, Trainable

Object Detector

Localize and identify multiple objects in a single image.

Sep 21, 2018



Model | Deployable

Audio Classifier

Identify sounds in short audio clips.

Sep 21, 2018



Model | Deployable

Image Caption Generator

Generate captions that describe the contents of images.

Sep 21, 2018



# Data Asset eXchange



[ibm.biz/data-exchange](http://ibm.biz/data-exchange)

## Data Asset eXchange

Explore useful and relevant data sets for enterprise data science

[Learn More](#)



[What's New](#)



[Get Involved](#)



Dataset | CSV

NOAA Weather Data - JFK Airport

June 30, 2020

Dataset | ICB format

Groningen Meaning Bank - Modified

May 14, 2020

Dataset | CSV

Fashion-MNIST

September 12, 2019



Dataset | JPG, JSON

PubLayNet

October 25, 2019

Dataset | WAV

TensorFlow Speech Commands

March 17, 2020

Dataset | PNG, JSON

PubTabNet

July 20, 2020



# Model Asset eXchange (MAX)

Place for developers/data scientists to find and use  
**free** and **open source** deep learning models



[ibm.biz/model-exchange](http://ibm.biz/model-exchange)

## Model Asset eXchange

Free, deployable, and trainable code. A place for developers to find and use free and open source deep learning models.

[Try the tutorial](#)



[Join the community](#)



[Featured](#) [Deployable](#) [Trainable](#)

Model | Deployable

Toxic Comment Classifier

Detect 6 types of toxicity in user comments

Jun 04, 2019

Model | Deployable, Trainable

Text Sentiment Classifier

Detect the sentiment captured in short pieces of text

Mar 29, 2019

Model | Deployable, Trainable

Image Segmenter

Identify objects in an image, additionally assigning each pixel of the image to a particular object.

Sep 21, 2018

Model | Deployable, Trainable

Object Detector

Localize and identify multiple objects in a single image.

Sep 21, 2018

Model | Deployable

Audio Classifier

Identify sounds in short audio clips.

Sep 21, 2018

Model | Deployable

Image Caption Generator

Generate captions that describe the contents of images.

Sep 21, 2018

[View all models](#)



# MAX Object Detector

Upload an image

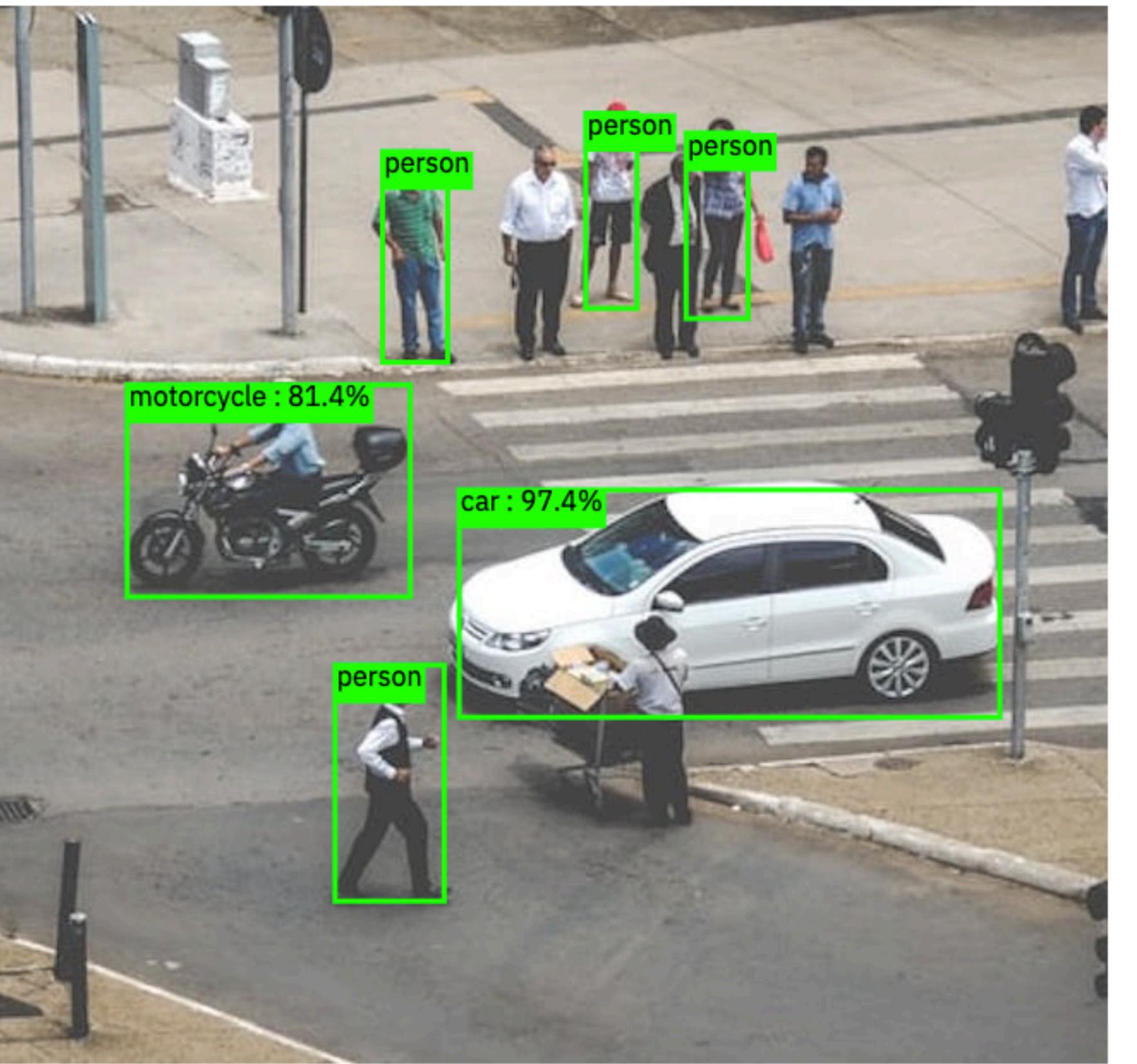
Choose File No file chosen

Submit  Use your webcam

Filter detected objects ⓘ

Probability Threshold: 50%

Labels Found ⓘ



MAX Image  
Caption  
Generator

Upload A New Image ⓘ  
 Choose Files No file chosen  
 Submit

Deselect All  Select All  Delete Uploaded Images

a man holding a snowboard on top of a snow covered slope .

a person riding skis down a snow covered slope .

a person riding a motorcycle on a race track

a tennis racket wth a tennis ball in it

a group of baseball players standing on top of a field .

a crowd of people watching a tennis match .

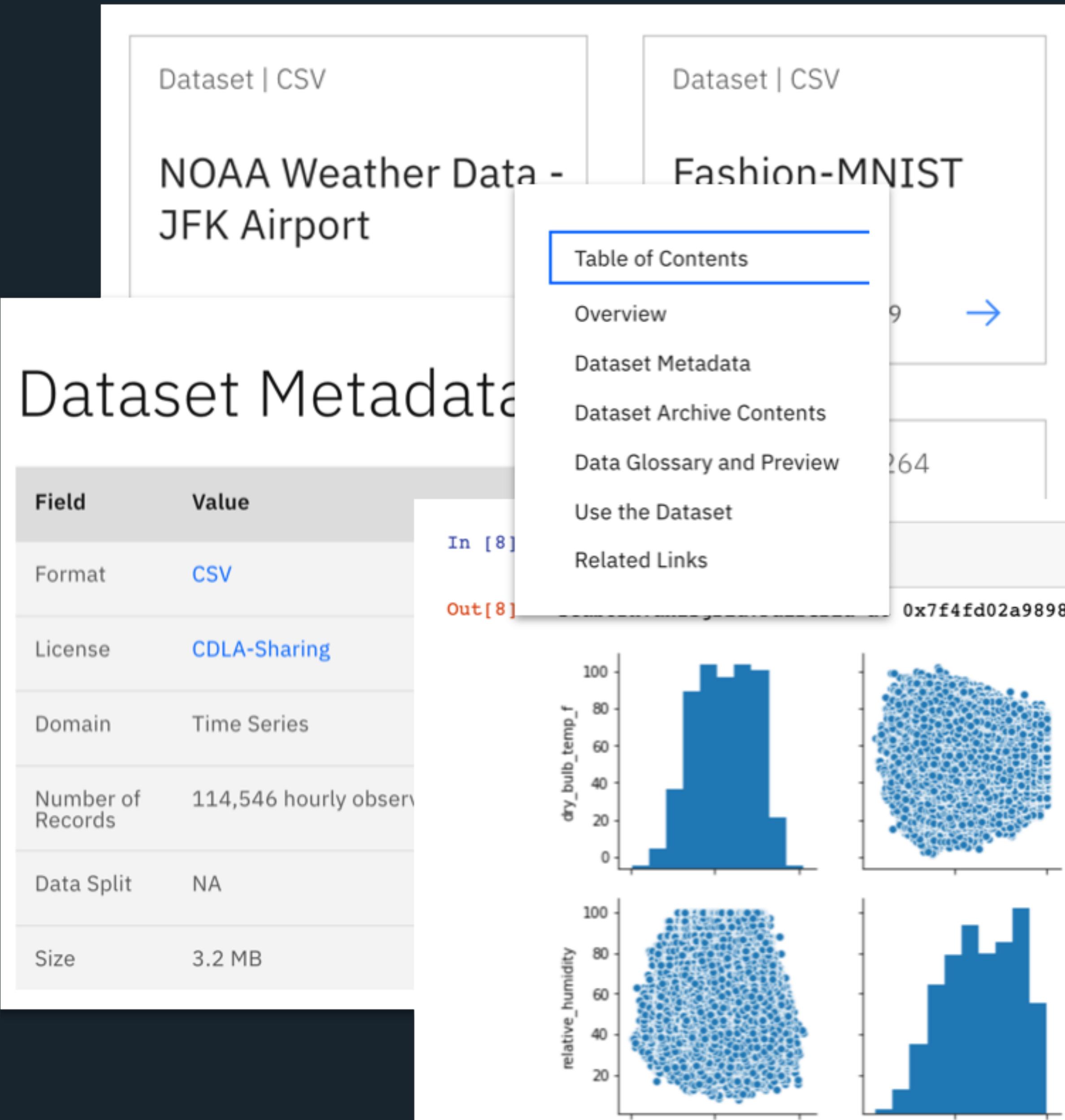
a group of young men playing a game of basketball .

a female tennis player in action on the court .

surfboard  
pitching  
players  
carrying  
playing  
track  
ramp  
man  
close  
people  
tennis  
men  
top  
covered  
holding  
group  
side  
skis  
wave  
action  
young  
race  
swinging  
city  
snowboard  
fire  
Street  
person  
snowbaseball  
crowd  
bikes  
slope  
skateboard  
standing  
match  
dirydrant  
beach

# Data Asset eXchange (DAX)

- Curated repository for **open** datasets from IBM Research and third-parties
- Published under data friendly licenses
- Standardized dataset formats and metadata
- Data sets include starter **notebooks**  
(cleansing, data exploration, analysis)



[ibm.biz/data-exchange](http://ibm.biz/data-exchange)

# NOAA Weather Data – JFK Airport

Local climatological data originally collected at JFK airport.

Save Like

Get this dataset →

Run dataset notebooks →

Preview the data & notebooks →

# NOAA Weather Data – JFK Airport

Part 1 - Data Cleaning

Part 2 - Data Analysis

Part 3 - Time Series Forecasting

```
In [1]: # @hidden_cell
# The project token is an authorization token that is used to access
project resources like data sources, connections, and used by platform
APIs.
from project_lib import Project
project = Project(project_id='...', project_access_token='...')
```

# NOAA Weather Data – JFK Airport

Dataset Metadata

Dataset Preview

Dataset Glossary

Format	<a href="#">CSV</a>
License	<a href="#">CDLA-Sharing</a>
Domain	Time Series
Number of Records	114,546 hourly observations
Data Split	NA
Size	3.2 MB
Data Origin	<a href="#">National Oceanic and Atmospheric Administration (NOAA)</a>
Dataset Version	Version 2 – September 12, 2019 Version 1 – July 16, 2019
Dataset Coverage	Location: New York City Dates: 2010-01-01 through 2018-07-27 Note: To download raw data from NOAA for a different region or date span, follow the steps outlined in the data archive's README.txt.
Agriculture	Detect unseasonal temperature change and alert farmers about potential damage to plants. Energy Regulate solar cell charging hours based on weather type condition and temperature. Regulate wind turbine operation based on wind speed and wind direction. Generate energy demand alerts based on

## Cleaning NOAA Weather Data of JFK Airport (New York)

This notebook relates to the NOAA Weather Dataset - JFK Airport (New York). The dataset contains 114,546 hourly observations of 12 local climatological variables (such as temperature and wind speed) collected at JFK airport. This dataset can be obtained for free from the IBM Developer [Data Asset Exchange](#).

In this notebook, we clean the raw dataset by:

- removing redundant columns and preserving only key numeric columns
- converting and cleaning data where required
- creating a fixed time interval between observations (this aids with later time-series analysis)
- filling missing values
- encoding certain weather features

### Table of Contents:

- [0. Prerequisites](#)
- [1. Read the Raw Data](#)
- [2. Clean the Data](#)
  - [2.1 Select data columns](#)
  - [2.2 Clean up precipitation column](#)
  - [2.3 Convert columns to numerical types](#)
  - [2.4 Reformat and process data](#)
  - [2.5 Create a fixed interval dataset](#)
  - [2.6 Feature encoding](#)
  - [2.7 Rename columns](#)

**ibm.biz/max-tutorial**

Series

# Learning Path: An introduction to the Model Asset Exchange

Learn how to use state-of-the-art deep learning models in your applications or services

[\*\*ibm.biz/dax-tutorial\*\*](http://ibm.biz/dax-tutorial)

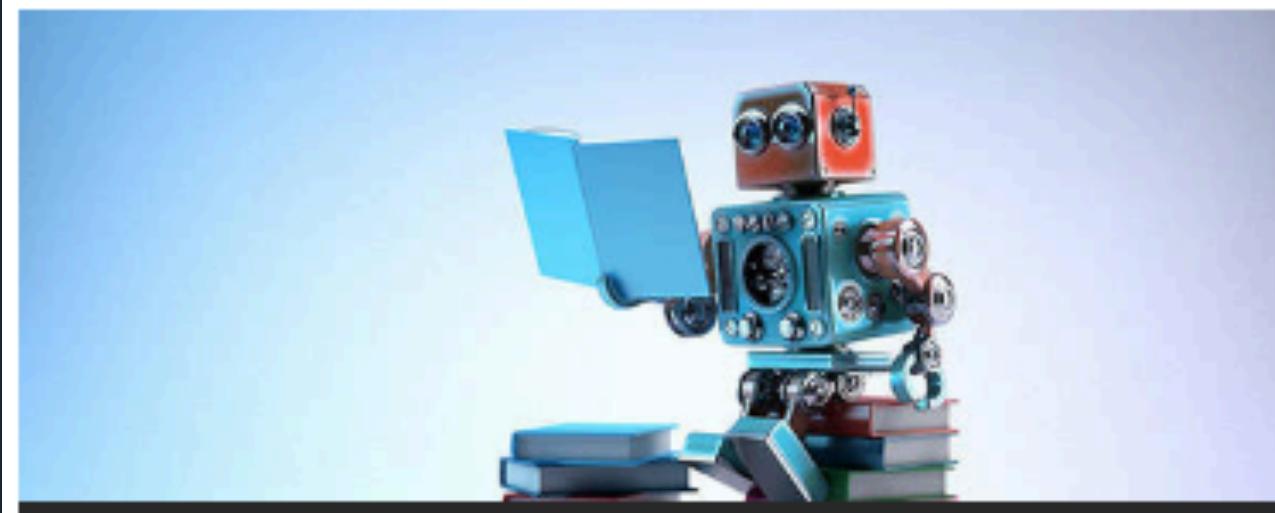
Tutorial

# Get started with the Data Asset eXchange

DAX offers a trusted source for open data sets for AI that are ready to use in enterprise AI applications

# Examples on how to easily consume MAX models

[ibm.biz/max-code-patterns](https://ibm.biz/max-code-patterns)



Code Pattern

Create a machine learning powered web app to answer questions

Nov 05, 2019 →



Code Pattern

Build a web app that recognizes yoga poses using a model from the Model Asset Exchange

Oct 03, 2019 →



Code Pattern

Use your arms to make music

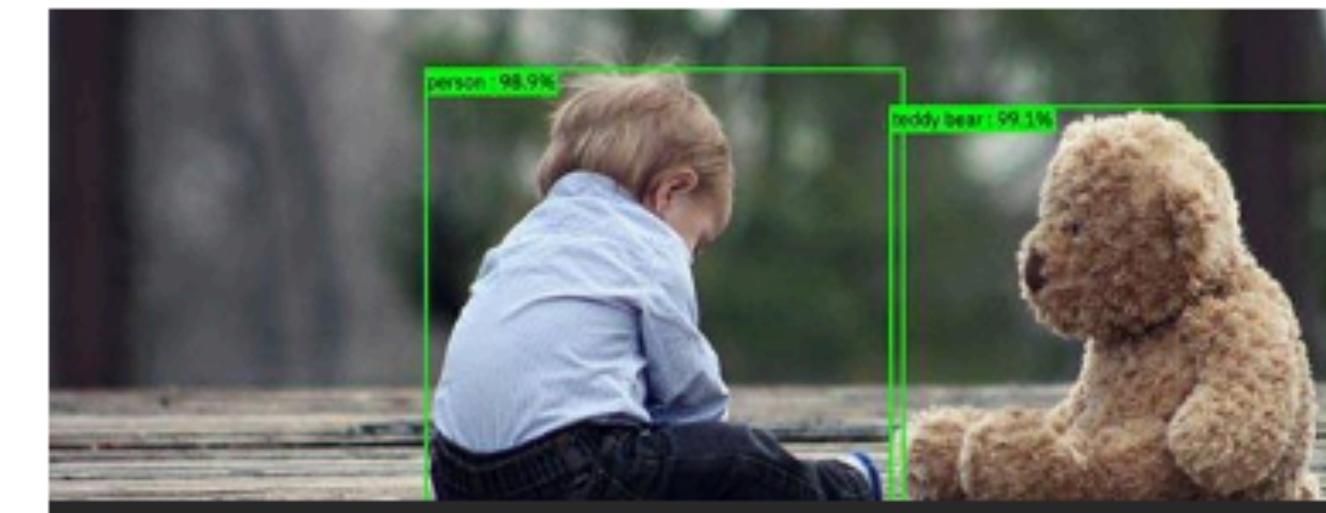
Apr 22, 2019 →



Code Pattern

Create a web app to interact with machine learning generated image captions

Mar 28, 2019 →



Code Pattern

Create a web app to visually interact with objects detected using machine learning

Mar 28, 2019 →



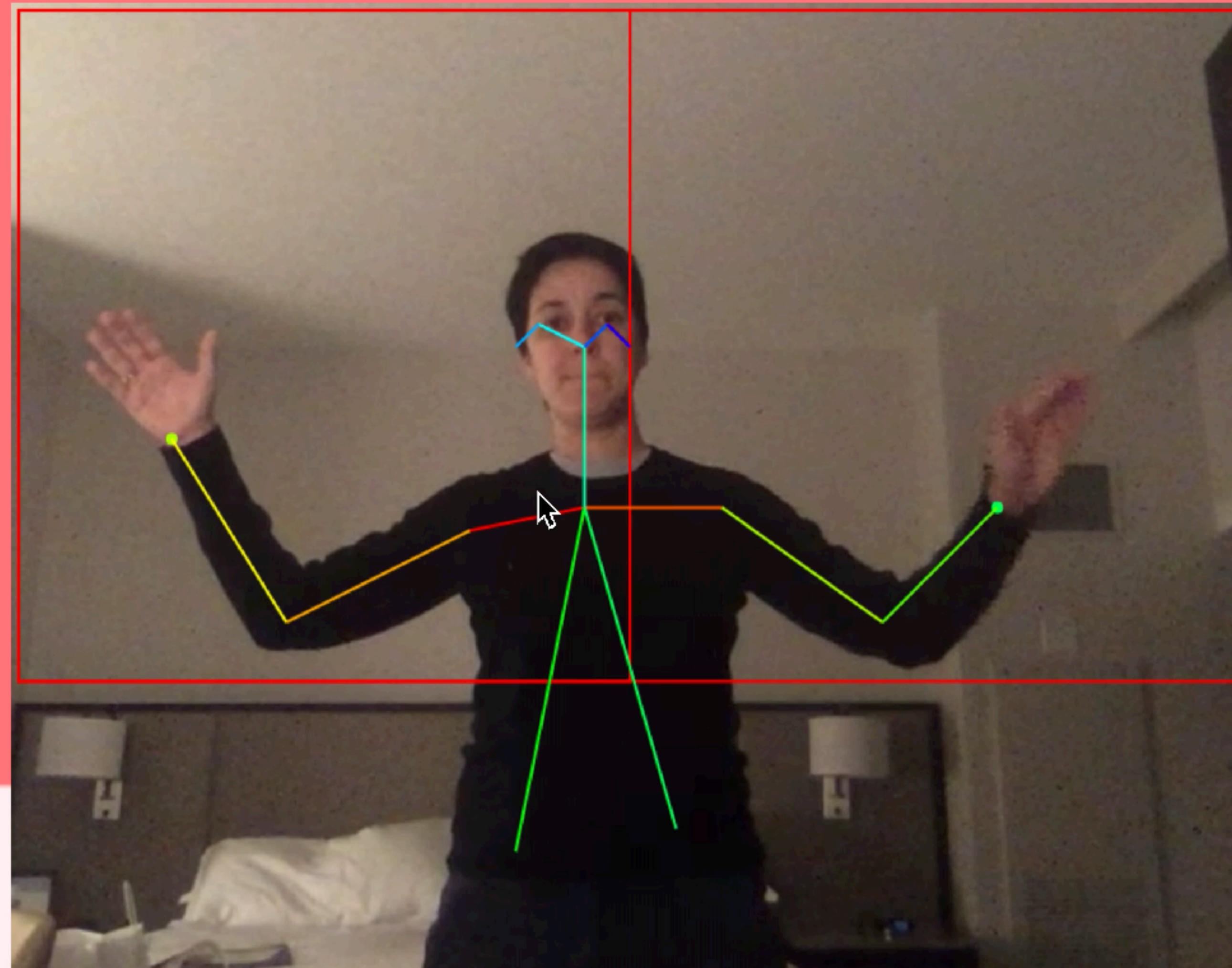
Code Pattern

Deploy a deep learning-powered 'Magic cropping tool'

Mar 28, 2019 →

veremax

a video theremin using OpenPose



# Thank you!

slides & materials: [bit.ly/kroz-talks](https://bit.ly/kroz-talks)



Don't forget to star the repo



ai-inclusive.org



@gdequeiroz | linktr.ee/gdq