

# IBM Developer Series

Join our webinar series to expand your development skills

Cloud Native Thursdays

Data Science Fridays

Red Hat OpenShift Saturdays

AI Sundays

[meetup.com/IBM-Cloud-MEA/](https://www.meetup.com/IBM-Cloud-MEA/)

# Build and deploy your Machine Learning models effortlessly

---

**Sidra Ahmed**

Developer Advocate, UAE

**Worood Dabbas**

Client Technical Specialist, UAE

**IBM Developer**

Build and deploy your Machine Learning models effortlessly November 20th, 2020 / © 2020 IBM Corporation



# Agenda

- **What is Data Science?**
- **Types of Data Analytics**
- **What is Classification?**
- **Different Classification Algorithms**
- **What is Watson Studio?**
- **Introduction to Jupyter Notebook**
- **Hands On**
- **Wrap Up**
- **Helpful Resources**

# Let's get started

- Sign up/Log in to your IBM Cloud Account  
<http://ibm.biz/JupyterML>

- Follow along for the hands on:  
<https://github.com/ozzael-codes/CustomerChurnAnalysis>



# What is Data Science?

Data Science is the discipline of extracting knowledge and insights from data and sharing discoveries in all this data.

*"Over 2.5 quintillion (18 zeros) bytes of data are created every single day" – (Forbes, 2018)*



Image Source: [http://cyborganthropology.com/Slow\\_Data\\_Movement](http://cyborganthropology.com/Slow_Data_Movement)

# Types of Data Analytics

**Descriptive**

**What Has happened?**

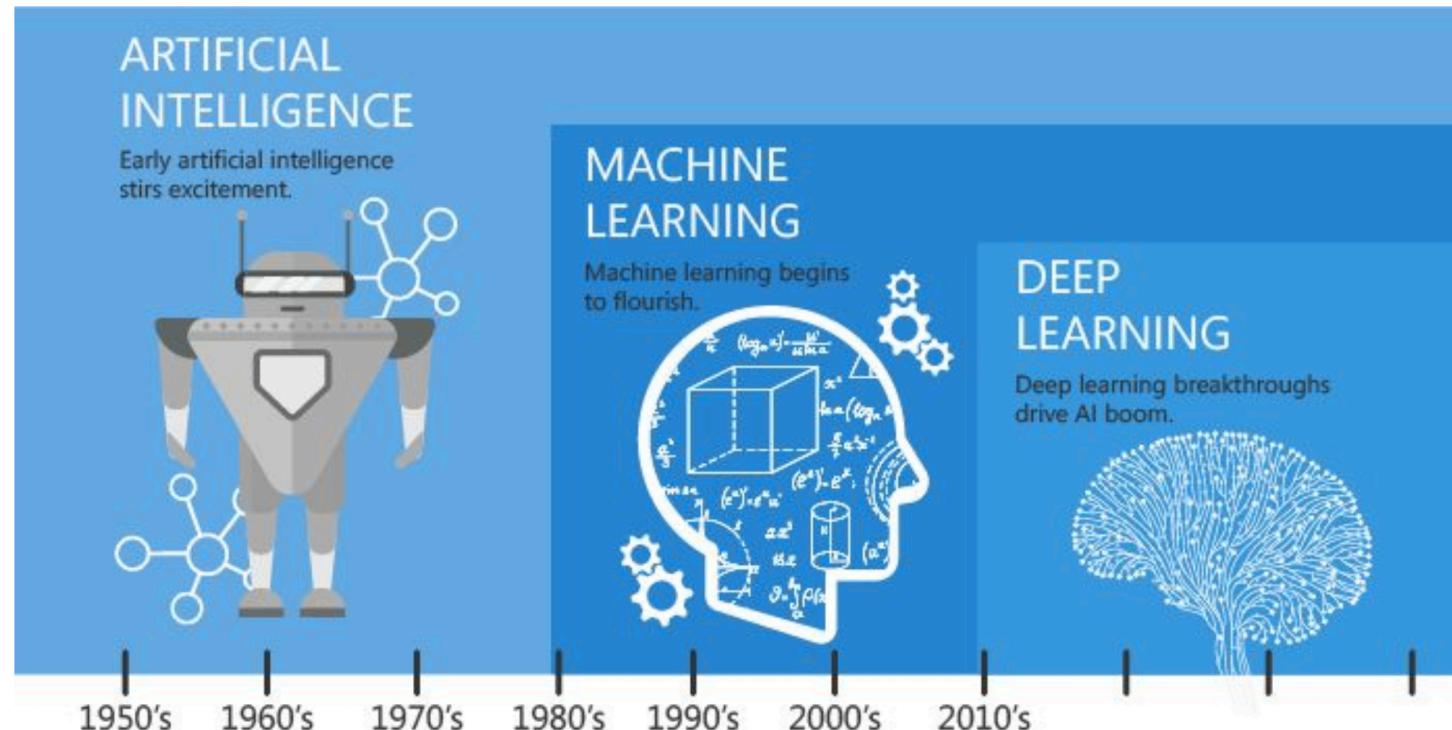
**Predictive**

**What Could Happen?**

**Prescriptive**

**What Should happen?**

# What is Machine Learning?



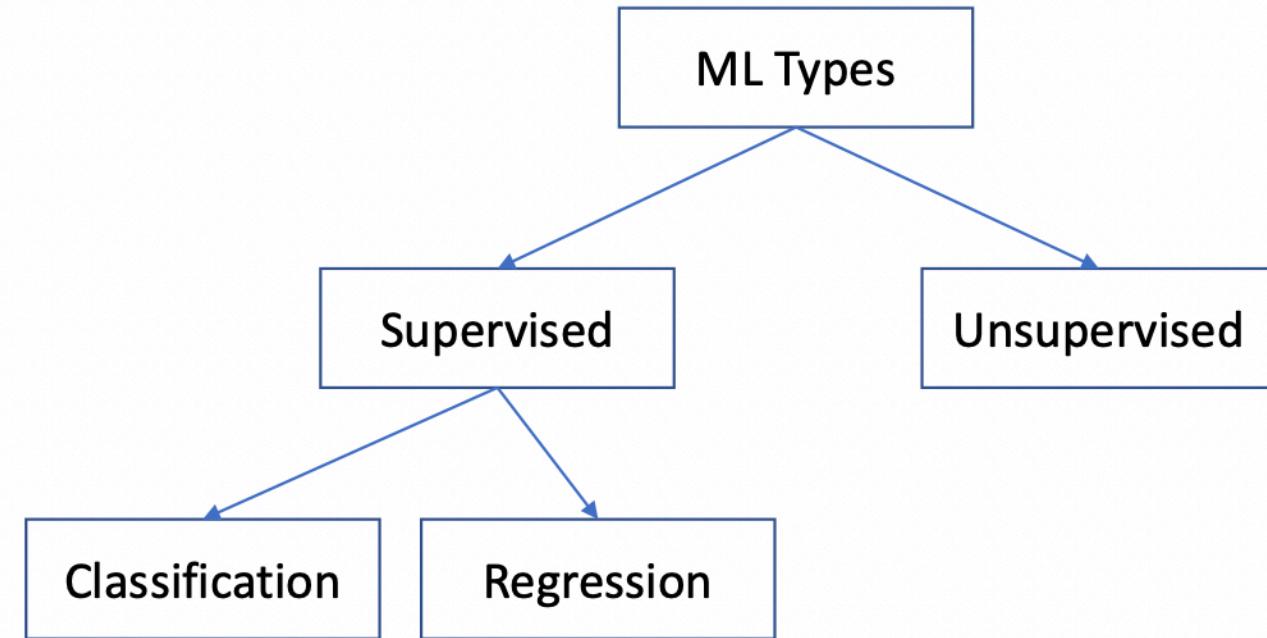
Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.

Image source: <https://towardsdatascience.com/artificial-intelligence-vs-machine-learning-vs-deep-learning-2210ba8cc4ac>

# Machine Learning

- 3 – 9
- 5 – 25
- 7 – 49
- 9 – ?

How did we end up with 81??



# What is Classification?

- Supervised learning approach
- Learning from the input data
- Uses the learning to classify new observations
- Predicts categorical class labels
- Examples: Speech recognition, Face recognition, Car type recognition

# k-nearest neighbors (KNN)

- “Birds of a feather flock together.”
- Classifies data points based on the points that are most similar to it.
- Non-parametric, lazy learning
- Classified by majority vote of K nearest neighbors
- Assigned to class most common among K neighbors
- Examples: Pattern recognition, Recommender systems, image classification

Get started at: <https://ibm.biz/WebinarClassification>

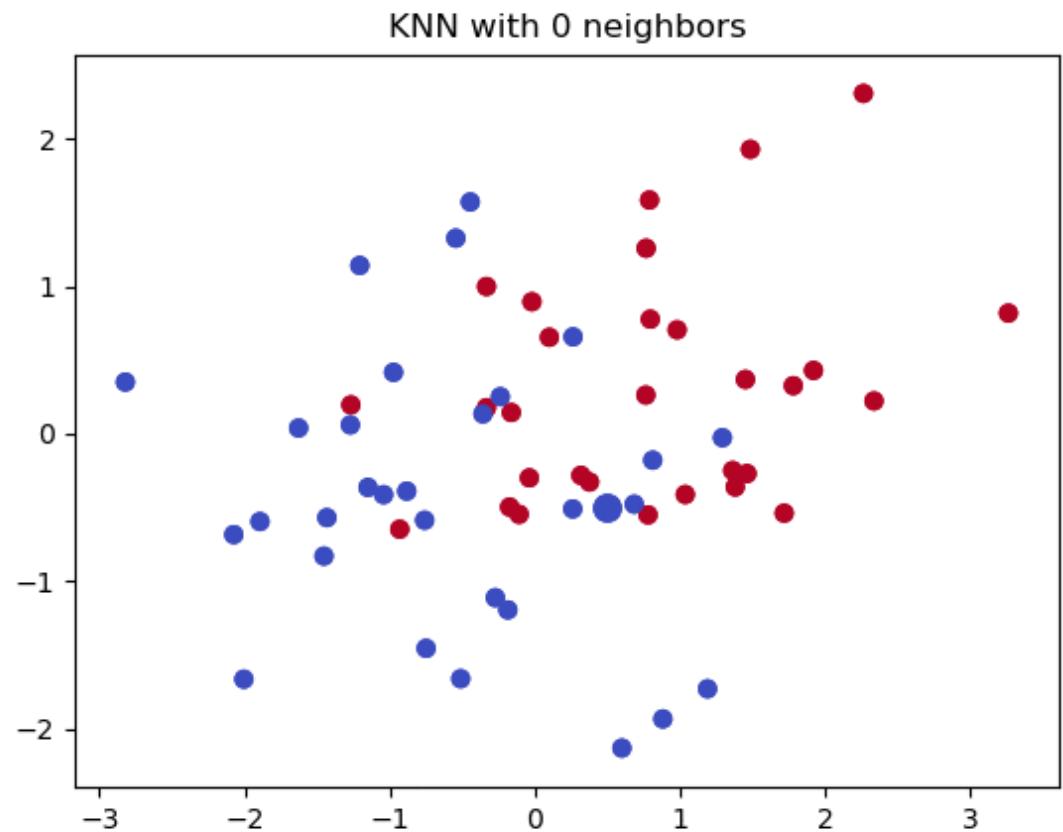


Image source: <https://importq.wordpress.com/2017/11/24/mnist-analysis-using-knn/>

# Support Vector Machines (SVM)

- Produces significant accuracy with less computational power
- Hyperplane in an N-dimensional space
- Objective is to find a plane that has the maximum margin
- Maximizing margin distance provides some reinforcement for future data

Get started at: <https://ibm.biz/WebinarClassification>

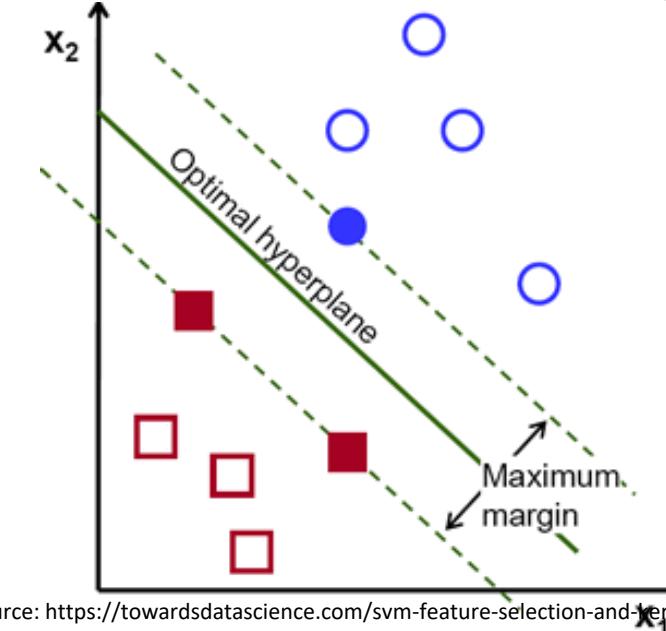
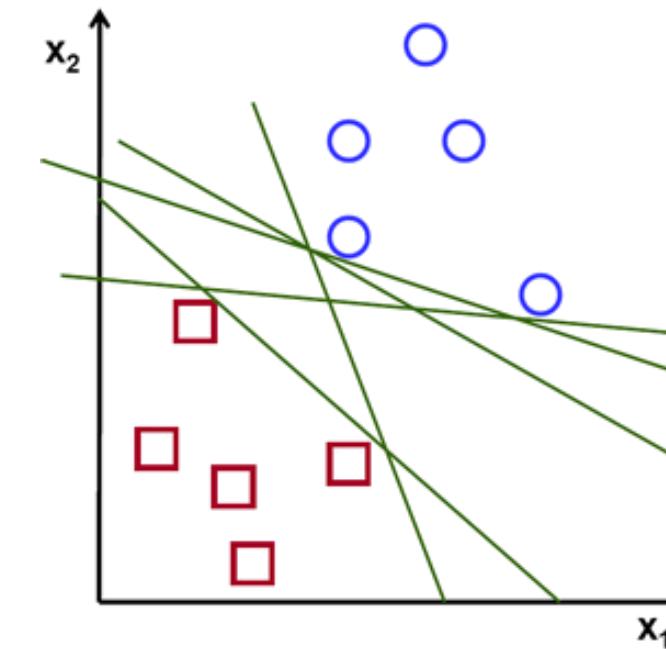


Image source: <https://towardsdatascience.com/svm-feature-selection-and-kernels-840781cc1a6c>

# Random Forest

Get started at: <https://ibm.biz/WebinarClassification>

- Ensemble model that grows multiple tree and classify objects based on the “votes” of all the tree
- Fits decision trees on sub-samples of datasets
- Uses average to improve accuracy of the model and controls over-fitting
- Slow real time prediction
- Examples: Fraud prediction, stock market analysis

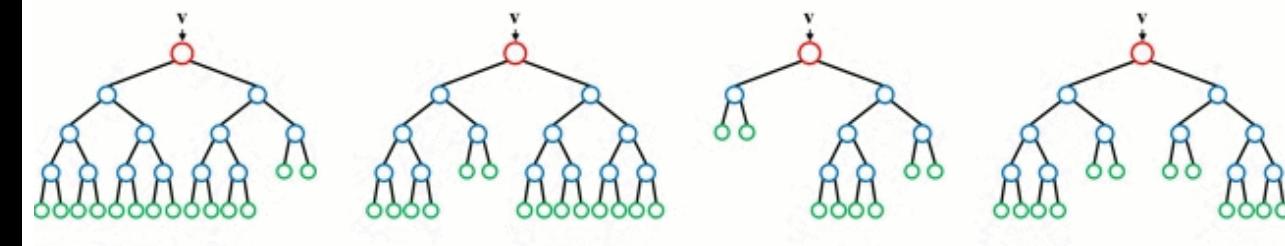


Image source: <https://towardsdatascience.com/a-quick-and-dirty-guide-to-random-forest-regression-52ca0af157f8>

# Why use Watson Studio?

- Very easy to learn
- Explore collection of data sets
- All tools in one place!
- All computation happens in the cloud!
- Easily integrate the models with websites



**Connect & Access Data**  
Connect and discover content from multiple data sources in the cloud or on premises. Bring **structured** and **unstructured** data to one toolkit.

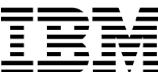
**Search and Find Relevant Data**  
Find data (structured, unstructured) and AI assets (e.g., ML/DL models, notebooks, Watson Data Kits) in the **Knowledge Catalog** with intelligent search and giving the right access to the right users.

**Prepare Data for Analysis**  
Clean and prepare your data with **Data Refinery**, a tool to create data preparation pipelines visually. Use popular open source libraries to prepare unstructured data.

**Build and Train ML/DL Models**  
**Democratize** the creation of ML and DL models. Design your AI models **programmatically** or **visually** with the most popular **open source** and IBM ML/DL frameworks or leverage transfer learning on **pre-trained** models using **Watson tools** to adapt to your business domain. Train at scale on **GPUs** and **distributed** compute

**Deploy Models**  
Deploy your models easily anywhere and have them **scale automatically** for online, batch or streaming use cases

**Monitor, Analyze and Manage**  
Monitor the performance of the models in production and trigger automatic retraining and redeployment of models. Build **Enterprise Trust** with Bias Detection, Mitigation Model **Robustness** and Testing Service Model **Security**.



# Jupyter Notebook

A web-based environment for interactive computing. You can run small pieces of code that process your data, and you can immediately view the results of your computation.

Notebooks include all of the building blocks you need to work with data:

- The data
- The code computations that process the data
- Visualizations of the results
- Text and rich media to enhance understanding

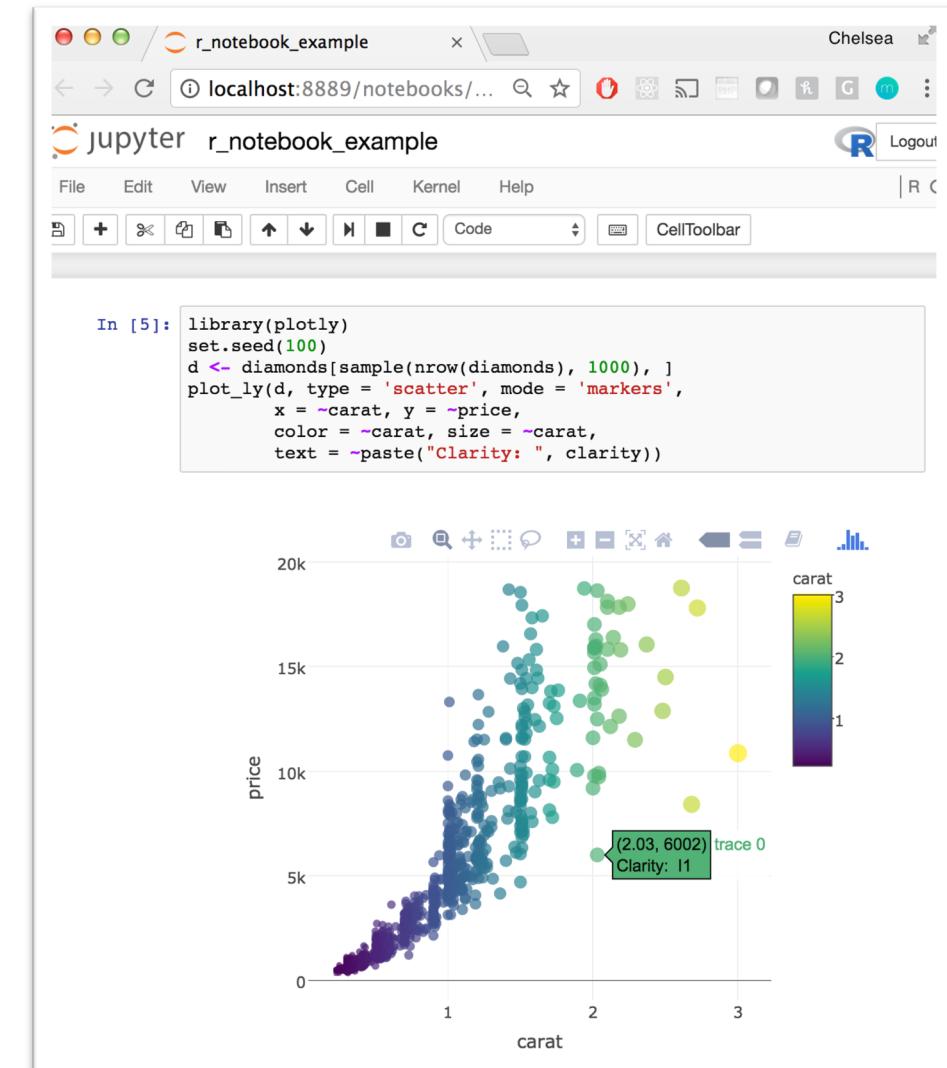
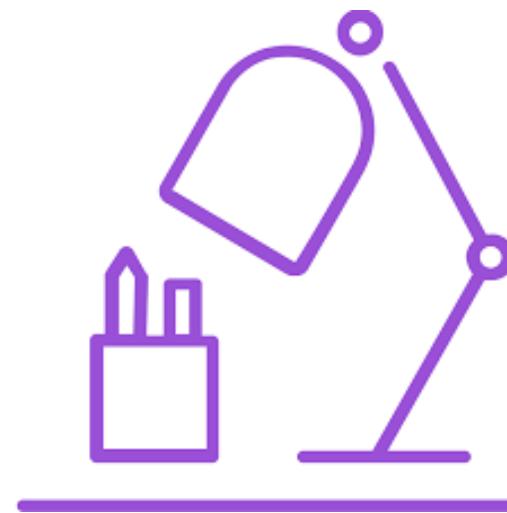


Image source: <https://plotly.com/python/ipython-notebook-tutorial/>



# Hands-On

- Sign up/Log in to your IBM Cloud Account  
<http://ibm.biz/JupyterML>
- Follow along for the hands on:  
<https://github.com/ozzael-codes/CustomerChurnAnalysis>



# Wrap Up

- **Data Science** refers to the discipline of extracting knowledge and insights from data and sharing discoveries in all this data.
- **Data Analytics types:** Descriptive, Predictive and Perspective.
- **Machine Learning** is a subset of AI techniques which uses statistical methods to enable machines to improve with experience.
- **Machine Learning** has two main types of Algorithms: Supervised and Unsupervised.
- The advantages of using **Watson Studio**.
- **Jupyter Notebook**, as web-based environment for interactive computing. You can run small pieces of code that process your data, and you can immediately view the results of your computation.
- Lastly, we have concluded the presentation with a demo for performing data preparation and visualization Model training using Jupyter Notebook.

# Which courses to take?

Get started at: <http://ibm.biz/JupyterML>

All courses are available on [cognitiveclass.ai](https://cognitiveclass.ai)

[Statistics 101](#)

**The courses for Data Science Program:**

**1- Data Science Introduction and concepts:**

[Data Science 101](#)

[Data Science Tools](#)

[Data Science Methodology](#)

**2- Getting hands-on with Python for Data Science:**

[Python for Data Science](#)

[Data Analysis with Python](#)

[Data Visualization with Python](#)

**3- Machine Learning in Detail:**

<https://www.coursera.org/learn/machine-learning>

<https://cognitiveclass.ai/courses/machine-learning-with-python>

**4- Taking one step further:**

Text Analytics - <https://cognitiveclass.ai/courses/systemt>

Data Science for Business - <https://cognitiveclass.ai/learn/data-science-business>

Build your own Chatbot - <https://cognitiveclass.ai/courses/how-to-build-a-chatbot>

Node-RED: Basics to Bots - <https://cognitiveclass.ai/courses/node-red-basics-to-bots>

IBM Cloud Essentials - <https://cognitiveclass.ai/courses/ibm-cloud-essentials>

# Useful Links

Learn – develop – connect

**IBM Developer** [developer.ibm.com](https://developer.ibm.com)

**Meetup Page** (<https://www.meetup.com/IBM-Cloud-MEA/>)

**Link to hands on** (<https://github.com/ozzael-codes/CustomerChurnAnalysis/>)

# Digital Developer Conference Data & AI

4 TRACKS

6 HOURS

5 HANDS-ON LABS

1 DATA & AI ESSENTIALS COURSE

NEW DATA COMPETITION LAUNCH

LIVE INTERVIEWS AND DISCUSSIONS

SPEAKERS FROM INDUSTRY, ACADEMIA, & OPENSOURCE

SPEAKER PANELS

COURSERA DISCOUNTS ON SPECIALIZATIONS

Americas & Europe  
November 10

Asia & Asia Pacific  
November 24

**Receive an exclusive discount on Coursera specializations and certifications after attending the conference.**

**Register Today:** [ibm.biz/devcon-ai](http://ibm.biz/devcon-ai)

**On-demand replays available post-event**



**IBM Research**



Explore AI, Data Science, & Machine Learning Certifications from IBM Developer. Expand your knowledge with resources that are valuable to your resume and the work you do every day.



Data & AI  
Essentials

CP4D Journey  
Intermediate



# Join IBM Developer Community

## IBM Champions Program

- Find more details and apply at: [developer.ibm.com/champions/](https://developer.ibm.com/champions/)

### Selection Criteria:

- Chosen for each country annually
- Expertise and innovation with IBM products and technology
- **Significant** contributions to advocacy, community, or influence
- Outside one's own job or position (organization or customer engagements)
- Over the previous 12 months

### Need more information and guidance?

Email me : [qamar.n@ibm.com](mailto:qamar.n@ibm.com)

Ask on Slack:



**Survey**

<https://www.surveymonkey.com/r/YWG7MFL>



# Thank you

**Sidra Ahmed**

Developer Advocate, UAE  
Sidra.Ahmed@ibm.com

**Worood Dabbas**

Client Technical Specialist, UAE  
Worood.al.dabbas1@ibm.com



