**Course 1: What Is Data Science?**

**10/1/20**

**Defining Data Science and What Data Scientists Do**

* Defining Data Science
* What is Data Science?
* Fundamentals of Data Science
* The Many Paths to Data Science
* Advice for New Data Scientists
* Data Science: The Sexiest Job in the 21st Century

**What Do Data Scientists Do?**

* A day in the Life of a Data Scientist
* Old problems, new problems, Data Science solutions
* Data Science Topics and Algorithms
* What is the cloud?
* What Makes Someone a Data Scientist?

**10/2/20**

**Data Science Topics**

* Foundations of Big Data
* How Big Data is Driving Digital Transformation
* What is Hadoop?
* Data Science Skills & Big Data
* Data Scientists at New York University
* Data Mining
* Quiz: Data Mining

**Deep Learning and Machine Learning**

* What's the difference?
* Neural Networks and Deep Learning
* Applications of Machine Learning
* Regression
* Quiz: Regression

**Data Science in Business**

* Applications of Data Science

**10/3/20**

* How Data Science is Saving Lives
* How Should Companies Get Started in Data Science?
* Applications of Data Science
* The Final Deliverable
* Quiz: The Final Deliverable

**Careers and Recruiting in Data Science**

* How Can Someone Become a Data Scientist?
* Recruiting for Data Science
* Careers in Data Science
* High School Students and Data Science Careers

**The Report Structure**

* The Report Structure
* Quiz: The Report Structure
* Final Assignment

Completed 10/3

10/2 Big Data

“Big Data refers to the dynamic, large and disparate volumes of data being created by people, tools, and machines. It requires new, innovative, and scalable technology to collect, host, and analytically process the vast amount of data gathered in order to derive real-time business insights that relate to consumers, risk, profit, performance, productivity management, and enhanced shareholder value.”

the V's of Big Data

* velocity
* volume
* variety
* veracity
* value

Veracity is the quality and origin of data, and its conformity to facts and accuracy. Attributes include consistency, completeness, integrity, and ambiguity. Drivers include cost and the need for traceability. With the large amount of data available, the debate rages on about the accuracy of data in the digital era. Is the information real, or is it false?

Deep Learning and Machine Learning

Neural networks/deep learning uses linear algebra (matrices)

Based on the videos and the reading material, how would you define a data scientist and data science? **(3 marks)**

**A data scientist is someone who uses data to (usually big data) to find solutions to problems that exist in business and in the world. Data scientists have knowledge of statistics, computer programming, and tools used to store and analyze data. They have good communication skills and are naturally curious, and typically have knowedge about the data for a particular industry or field, such as health care or transportation.**

**Data science is the practice of using structured and unstructured data to gain insights and knowledge. Like other science disciplines, a data science project starts with a hypothesis and then research is done to find conclusions in large amounts of data. Math, computer science, and statistical techniques are used, but data management, data visualization, and communication of results are also important.**

As discussed in the videos and the reading material, data science can be applied to problems across different industries. What industry are you passionate about and would like to pursue a data science career in? **(1 mark)**

**I’m a heath care data analyst and project manager. I started out as a programmer, but it was a long time ago, before R and Python existed. I have programmed in Assembler, COBOL, C and C++. I’ve always loved working with data and have a strong understanding of health care claims, databases and SQL. I would like to learn data science technical skills so I can work on more complex and exciting analytic projects.**

Based on the videos and the reading material, what are the **ten** main components of a report that would be delivered at the end of a data science project? **(5 marks)**

1. **Cover page**
2. **Table of Contents**
3. **Executive Summary or Abstract**
4. **Introduction**
5. **Methodology**
6. **Results**
7. **Discussion**
8. **Conclusion**
9. **Acknowledgements**
10. **References**
11. **Appendices, if needed**

Rubric says:

1. Cover page
2. Table of contents
3. Introductory section
4. Methodology section
5. Results section
6. Discussion section
7. Conclusion section
8. References
9. Acknowledgment
10. Appendix

**2: Tools for Data Science**

**Week 1**

**10/4/20**

### Languages of Data Science

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/uSoGc/course-introduction)** [Lecture Course Introduction](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/uSoGc/course-introduction)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/2LT2g/languages-of-data-science)** [Lecture Languages of Data Science](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/2LT2g/languages-of-data-science)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/HNbm7/introduction-to-python)** [Lecture Introduction to Python](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/HNbm7/introduction-to-python)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/Ve0zo/introduction-to-r-language)** [Lecture Introduction to R Language](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/Ve0zo/introduction-to-r-language)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/UE7VR/introduction-to-sql)** [Lecture Introduction to SQL](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/UE7VR/introduction-to-sql)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/9jW96/other-languages)** [Lecture Other Languages](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/9jW96/other-languages)

**[Practice Quiz:](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/Zue9i/practice-quiz-languages)** [Practice Quiz - Languages](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/Zue9i/practice-quiz-languages)

Data Science Tools

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/nehhw/categories-of-data-science-tools)** [Lecture Categories of Data Science Tools](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/nehhw/categories-of-data-science-tools)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/KtaZO/open-source-tools-for-data-science-part-1)** [Lecture Open Source Tools for Data Science - Part 1](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/KtaZO/open-source-tools-for-data-science-part-1)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/ZScMZ/open-source-tools-for-data-science-part-2)** [Lecture Open Source Tools for Data Science - Part 2](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/ZScMZ/open-source-tools-for-data-science-part-2)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/8ajoX/commercial-tools-for-data-science)** [Lecture Commercial Tools for Data Science](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/8ajoX/commercial-tools-for-data-science)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/PcB72/cloud-based-tools-for-data-science)** [Lecture Cloud Based Tools for Data Science](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/PcB72/cloud-based-tools-for-data-science)

**[Practice Quiz:](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/A8y9q/practice-quiz-tools)** [Practice Quiz - Tools](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/A8y9q/practice-quiz-tools)

Completed 10/5/20

10/6/20

Packages, APIs, Data Sets and Models

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/HKsDX/libraries-for-data-science)** [Lecture Libraries for Data Science](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/HKsDX/libraries-for-data-science)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/1N58k/application-programming-interfaces-api)** [Lecture Application Programming Interfaces (API)10/](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/1N58k/application-programming-interfaces-api)

10//20

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/QfDtE/data-sets-powering-data-science)** [Lecture Data Sets - Powering Data Science](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/QfDtE/data-sets-powering-data-science)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/b4lXn/sharing-enterprise-data-data-asset-exchange)** [Lecture Sharing Enterprise Data - Data Asset eXchange](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/b4lXn/sharing-enterprise-data-data-asset-exchange)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/hG04G/machine-learning-models)** [Lecture Machine Learning Models](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/hG04G/machine-learning-models)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/1vr1M/the-model-asset-exchange)** [Lecture The Model Asset Exchange](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/1vr1M/the-model-asset-exchange)

Completed 10/6/20

**Week 2**

### GitHub

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/y9I8s/overview-of-git-github)** [Lecture Overview of Git/GitHub](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/y9I8s/overview-of-git-github)

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/YWm0i/github-lab-getting-started)** [GitHub Lab - Getting started](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/YWm0i/github-lab-getting-started)

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/VjoQB/branching-merging-and-pull-requests-on-github-optional)** [Branching, Merging and Pull Requests on GitHub (Optional)](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/VjoQB/branching-merging-and-pull-requests-on-github-optional)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/xJOQs/git-and-github-via-command-line-optional)** [Lecture Git and GitHub via command line (Optional)](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/xJOQs/git-and-github-via-command-line-optional)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/WZF5v/branching-and-merging-via-command-line-optional)** [Lecture Branching and merging via command line (Optional)](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/WZF5v/branching-and-merging-via-command-line-optional)

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/RN5Oh/pre-requisites-for-command-line-interface-optional)** [Pre-requisites for command line interface (Optional)](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/RN5Oh/pre-requisites-for-command-line-interface-optional)

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/lVHaf/configuring-ssh-access-to-repository-optional)** [Configuring SSH access to repository (Optional)](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/lVHaf/configuring-ssh-access-to-repository-optional)

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/clE3k/lab-2-branching-and-merging-via-command-line-optional)** [Lab 2: Branching and merging via command line (optional)](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/clE3k/lab-2-branching-and-merging-via-command-line-optional)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/lamfb/contributing-to-repositories-via-pull-request-optional)** [Lecture Contributing to repositories via pull request (Optional)](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/lamfb/contributing-to-repositories-via-pull-request-optional)

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/KgHxH/lab-3-contributing-to-repositories-via-pull-request-optional)** [Lab 3: Contributing to repositories via pull request (Optional)](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/KgHxH/lab-3-contributing-to-repositories-via-pull-request-optional)

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/5l1gp/explore-data-sets-and-models)** [Explore Data Sets and Models](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/5l1gp/explore-data-sets-and-models)

**[Practice Quiz:](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/70eZy/practice-quiz-packages-apis-data-sets-models)** [Practice Quiz - Packages, APIs, Data Sets, Models](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/70eZy/practice-quiz-packages-apis-data-sets-models)

Module 1 Assessment

**[Quiz:](https://www.coursera.org/learn/open-source-tools-for-data-science/exam/YGD4N/graded-quiz)** [Graded Quiz](https://www.coursera.org/learn/open-source-tools-for-data-science/exam/YGD4N/graded-quiz)

### Jupyter Notebook and JupyterLab

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/k5Myi/getting-started-with-jupyter-notebook)** [Lecture Getting Started with Jupyter Notebook](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/k5Myi/getting-started-with-jupyter-notebook)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/dVRyO/getting-started-with-jupyterlab)** [Lecture Getting Started with JupyterLab](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/dVRyO/getting-started-with-jupyterlab)

**Reading:** Hands-on Lab: Jupyter Notebook - The Basics

**[Ungraded External Tool:](https://www.coursera.org/learn/open-source-tools-for-data-science/ungradedLti/jqTx7/lab-jupyter-notebook-the-basics)** [Ungraded External ToolLab - Jupyter Notebook - The Basics](https://www.coursera.org/learn/open-source-tools-for-data-science/ungradedLti/jqTx7/lab-jupyter-notebook-the-basics)

[.](https://www.coursera.org/learn/open-source-tools-for-data-science/ungradedLti/jqTx7/lab-jupyter-notebook-the-basics)**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/4i4q5/lab-jupyter-notebook-more-features)** [Lab - Jupyter Notebook - More Features](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/4i4q5/lab-jupyter-notebook-more-features)

**[Ungraded External Tool:](https://www.coursera.org/learn/open-source-tools-for-data-science/ungradedLti/WH6RD/lab-jupyter-notebook-more-features)** [Ungraded External ToolLab - Jupyter Notebook - More Features](https://www.coursera.org/learn/open-source-tools-for-data-science/ungradedLti/WH6RD/lab-jupyter-notebook-more-features)

**Reading:** Hands-on Lab: Jupyter Notebook - Advanced Features

**[Ungraded External Tool:](https://www.coursera.org/learn/open-source-tools-for-data-science/ungradedLti/OB9w2/lab-jupyter-notebook-advanced-features)** [Ungraded External ToolLab - Jupyter Notebook - Advanced Features](https://www.coursera.org/learn/open-source-tools-for-data-science/ungradedLti/OB9w2/lab-jupyter-notebook-advanced-features)

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/CBjDg/jupyter-notebooks-on-the-internet)** [Jupyter Notebooks on the Internet](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/CBjDg/jupyter-notebooks-on-the-internet)

**[Practice Quiz:](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/i4RTZ/practice-quiz-jupyter-notebook)** [Practice Quiz - Jupyter Notebook](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/i4RTZ/practice-quiz-jupyter-notebook)

### RStudio IDE

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/JtqjY/plotting-within-rstudio-ide)** [Lecture Plotting Within RStudio IDE](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/JtqjY/plotting-within-rstudio-ide)

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/QwqfE/lab-rstudio-the-basics)** [Lab: RStudio - The Basics](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/QwqfE/lab-rstudio-the-basics)

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/TGjKV/lab-creating-an-interactive-map-with-r)** [Lab: Creating an Interactive Map with R](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/TGjKV/lab-creating-an-interactive-map-with-r)

**[Practice Quiz:](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/RQv6O/practice-quiz-rstudio-ide)** [Practice Quiz - RStudio IDE](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/RQv6O/practice-quiz-rstudio-ide)

Module 2 Assessment

**[Quiz:](https://www.coursera.org/learn/open-source-tools-for-data-science/exam/YzYyr/graded-quiz)** [Graded Quiz](https://www.coursera.org/learn/open-source-tools-for-data-science/exam/YzYyr/graded-quiz)

Week 3

### Watson Studio

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/QvnOI/what-is-ibm-watson-studio)** [Lecture What is IBM Watson Studio?](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/QvnOI/what-is-ibm-watson-studio)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/pfhyY/watson-studio-introduction)** [Lecture Watson Studio Introduction](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/pfhyY/watson-studio-introduction)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/Ig5vN/creating-an-account-on-ibm-watson-studio)** [Lecture Creating an Account on IBM Watson Studio](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/Ig5vN/creating-an-account-on-ibm-watson-studio)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/OA5oc/jupyter-notebook-in-watson-studio-part-1)** [Lecture Jupyter Notebook in Watson Studio - Part 1](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/OA5oc/jupyter-notebook-in-watson-studio-part-1)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/yCpia/jupyter-notebook-in-watson-studio-part-2)** [Lecture Jupyter Notebook in Watson Studio - Part 2](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/yCpia/jupyter-notebook-in-watson-studio-part-2)

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/yxtx1/lab-creating-a-watson-studio-project-with-jupyter-notebook)** [Lab: Creating a Watson Studio Project with Jupyter Notebook](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/yxtx1/lab-creating-a-watson-studio-project-with-jupyter-notebook)

**[Practice Quiz:](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/hhF9P/practice-quiz-watson-studio)** [Practice Quiz - Watson Studio](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/hhF9P/practice-quiz-watson-studio)

Other IBM Tools

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/r0MWa/other-ibm-tools-for-data-science)** [Lecture Other IBM Tools for Data Science](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/r0MWa/other-ibm-tools-for-data-science)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/tRvop/data-refinery)** [Lecture Data Refinery](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/tRvop/data-refinery)

**Video:** Lecture SPSS Modeler Flows in Watson Studio

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/4Ei0B/lab-modeler-flows-in-watson-studio)** [Lab: Modeler Flows in Watson Studio](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/4Ei0B/lab-modeler-flows-in-watson-studio)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/Q2rSv/spss-statistics)** [Lecture SPSS Statistics](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/Q2rSv/spss-statistics)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/roMpP/model-deployment-with-watson-machine-learning)** [Lecture Model Deployment with Watson Machine Learning](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/roMpP/model-deployment-with-watson-machine-learning)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/zQAw0/auto-ai-in-watson-studio)** [Lecture Auto AI in Watson Studio](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/zQAw0/auto-ai-in-watson-studio)

**[Video:](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/5wRBQ/ibm-watson-openscale)** [Lecture IBM Watson OpenScale](https://www.coursera.org/learn/open-source-tools-for-data-science/lecture/5wRBQ/ibm-watson-openscale)

**[Practice Quiz:](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/vZaU4/practice-quiz-other-ibm-tools)** [Practice Quiz - Other IBM Tools](https://www.coursera.org/learn/open-source-tools-for-data-science/quiz/vZaU4/practice-quiz-other-ibm-tools)

Module 3 Assessment

**[Quiz:](https://www.coursera.org/learn/open-source-tools-for-data-science/exam/R7O82/graded-quiz)** [Graded Quiz](https://www.coursera.org/learn/open-source-tools-for-data-science/exam/R7O82/graded-quiz)

Week 4

### Create and Share Your Jupyter Notebook

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/mnbnX/instructions-create-and-share-your-jupyter-notebook)** [Instructions: Create and Share Your Jupyter Notebook](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/mnbnX/instructions-create-and-share-your-jupyter-notebook)

**Peer-graded Assignment:** Create and Share Your Jupyter Notebook

**[Review Your Peers:](https://www.coursera.org/learn/open-source-tools-for-data-science/peer/2pLEh/create-and-share-your-jupyter-notebook/give-feedback)** [Create and Share Your Jupyter Notebook](https://www.coursera.org/learn/open-source-tools-for-data-science/peer/2pLEh/create-and-share-your-jupyter-notebook/give-feedback)

Digital Badge

**[Reading:](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/7726U/ibm-digital-badge)** [IBM Digital Badge](https://www.coursera.org/learn/open-source-tools-for-data-science/supplement/7726U/ibm-digital-badge)

Languages quiz (difficult)

Which of the following statements is true?

**1 / 1 point**



Python is the most popular language in data science.



80% of data scientists worldwide use Python.



Python is useful for AI, machine learning, web development, and IoT.



Keras, Scikit-learn, Matplotlib, Pandas, and TensorFlow are all Python libraries



All of the above

Correct

2.

Question 2

Which of the following are SQL databases? (Select all that apply.)

**0 / 1 point**



CouchDB – should not be selected



MongoDB – should not be selected



PostgreSQL - yes



MySQL

Correct



MariaDB

Correct



Oracle

Correct

You didn’t select all the correct answers

3.

Question 3

Which statements are true about Open Source and Free Software? (Select all that apply.)

**0 / 1 point**



Free Software and Open Source can be used interchangeably.

This should not be selected



Free Software can always be run, studied, modified and redistributed with or without changes.

Correct



Most of Free Software licenses also qualify for Open Source.

Correct



Open Source Software can be modified without sharing the modified source code depending on the Open Source license.

Correct

4.

Question 4

Is the following statement true or false: R integrates well with other computer languages like C++, Java, C, .Net and Python.

**1 / 1 point**



True



False

Correct

5.

Question 5

Which of the following languages can be used for data science?

**1 / 1 point**



R



Julia



Java



Javascript



Scala



SQL



All of the above

Correct

6.

Question 6

Which of the following is used to make Artificial Intelligence and Machine Learning possible? (Select all that apply.)

**0 / 1 point**



Oracle – should not be selected



PyTorch yes



TensorFlow.js yes

Correct



Apache Spark

Correct



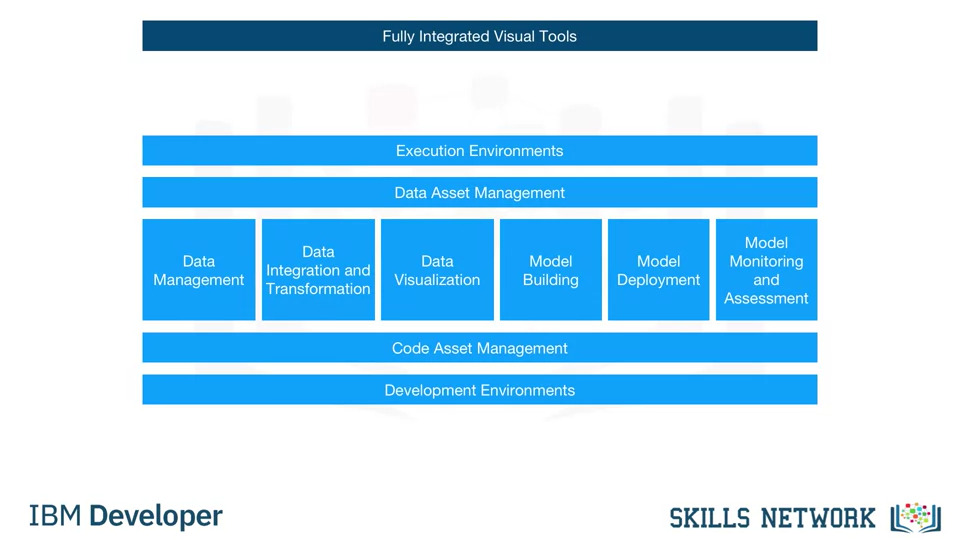
GNU – should not be selected



Caffe

You didn’t select all the correct answers

**Categories of data science tools**



1. Which of the f ollowing are common tasks in data science? (Select all that apply)

Data Management

Data Integration and Transfomation

Data Visualization

Model Building

Model Deployment

Model Monitoring and Assessment

All of the above

2.

Which of the following are data management tools? (Select all that apply.)

GitHub

MySQL

PostgreSQL

KubeFlow (no – ETL)

PixieDust

3.

Which of the following are Data Integration and Transformation tools? (Select all that apply.)

Cassandra – No – Data Management

Apache Kafka

Apache Nifi

Apache AirFlow

Ceph – No – data mangement

4.

Which statement about JupyterLab is correct?

JuypterLab can run R and Python code in addition to other programming languages.

JuypterLab can run R and Python code only.

JuypterLab can run Python code only.

JuypterLab can run R code only.

5.

Which statement about RStudio is correct?

RStudio is the primary choice for development in the R programming language.

RStudio is the primary choice for web development.

RStudio is the primary choice for development in the Python programming language.

6.

Which statements about IBM Watson Studio and OpenScale are correct? (Select all that apply.)

Watson Studio together with Watson OpenScale is a database management system.

Watson Studio together with Watson OpenScale covers the complete development life cycle for all data science, machine learning and AI tasks.

Watson Studio together with Watson OpenScale is available as a Cloud offering as well as a package running on top of Kubernetes/RedHat OpenShift in a local data center called IBM Cloud Pak for Data.



**Data Science Libraries**

Scientific Computing Libraries in Python

Pandas – Data structures and tools (data frame)

NumPy – arrays and matrices

Data Visualization in Python

Matprotlib

Seaborn

Scikit-learn – regression and machine learning regression

Keras – Deep learning neural networks

TensorFlow (production)

PyTorch (research)

Apache Spark (parallel processing)

Scala-library (data engineering

Vegas

Bog DL (Deep learning)

R Libraries

Ggplot2

Keras

Tensorflow

**APIs**

REST APIs – communicate using the internet Representational State Transfer (web service)

JSON file – inscludes data and intructions for operations to be performed

**Data Sets**

