Which of the following are common tasks in data science? (Select all that apply)

**1 point**



Data Management



Data Integration and Transformation



Data Visualization



Model Building



Model Deployment



Model Monitoring and Assessment

**2.**

Question 2

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

**1 point**



GitHub



MySQL



PostgreSQL



KubeFlow



PixieDust

**3.**

Question 3

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

**1 point**



Cassandra



Apache Kafka



Apache Nifi



Apache AirFlow



Ceph

**4.**

Question 4

Which statement about JupyterLab is correct?

**1 point**



JuypterLab can run Python code only.



JuypterLab can run R code only.



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



JuypterLab can run R and Python code only.

**5.**

Question 5

Which statement about RStudio is correct?

**1 point**



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.**

Question 6

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

**1 point**



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.

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

Question 1

According to the Module 1 reading, “The Sexiest Job in the 21st Century”, a report by the McKinsey Global Institute, by 2018, it is projected that there will be a shortage of people with deep analytical skills in the United States. What is the size of this shortage?

**1 point**



140,000 – 190,000 people



800,000 – 900,000 people



120,000 people



3 – 6 million people



20,000 – 50,000 people

**2.**

Question 2

According to the Module 1 reading, “The Sexiest Job in the 21st Century”, data Science was called the sexiest job of what century by Harvard Business Review?

**1 point**



19th century



20th century



21st century



22nd century

**3.**

Question 3

According to the Module 1 reading "What Makes Someone a Data Scientist”, this chief economist at Google declared that "the sexy job in the next ten years will be statisticians”. What is his name:

**1 point**



Larry Page



Sundar Pichai



Hal Varian



Sal Barian

**4.**

Question 4

According to the reading Module 1 "What Makes Someone a Data Scientist", the author defines a data scientist as someone who finds solutions to what, by analyzing data using appropriate tools and then tells stories to communicate their findings to the relevant stakeholders?

**1 point**



Problems



Drawbacks



Complications



Questions

**5.**

Question 5

According to the Module 2 reading “Data Mining”, the output of a data mining exercise largely depends on the quality of what?

**1 point**



The data



The project



The data scientist



The material

**6.**

Question 6

According to the Module 2 reading, “Data Mining”, when data is missing in a systematic way, you should do what?

**1 point**



Find out who was managing the database



Determine the impact of the missing data on the results and whether missing data can be excluded from the analysis



Extrapolate and average the data



Stop immediately

**7.**

Question 7

Based on the Module 2 reading, "Regression”, the real added value of the author's research on what type of properties is quantifying the magnitude of relationships between housing prices and different determinants?

**1 point**



Foreclosed



Vacant



Commercial real estate



Residential real estate

**8.**

Question 8

Based on the Module 2 reading, "Regression”, adding what type of room impacted the price of a home the most?

**1 point**



Washroom



Theater room



Game room



Bedroom

**9.**

Question 9

According to the Module 3 reading, "The Final Deliverable”, the ultimate purpose of analytics is to communicate findings to what people to formulate policy or strategy?

**1 point**



**Stakeholders**



Salespeople



Marketing



CEO’s

**10.**

Question 10

Based on The Module 3 reading, "The Final Deliverable”, it mentions a common role of a data scientist is to use what to build a narrative to communicate findings to stakeholders?

**1 point**



Spreadsheets



**Analytics insights**



Strategies



Big data

**11.**

Question 11

Based on the Module 3 reading, “The Report Structure”, regardless of the length of the final deliverable, the author recommends that it includes a \_\_\_\_\_, table of contents, executive summary, a methodology section, and a discussion section.

**1 point**



**Cover page**



Copy of the data set



Project scope statement



Dedication

**12.**

Question 12

An introductory section is always helpful in introducing the research methods and presenting the statistical calculations.

**1 point**



***False.***



True.

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

Which of the following are common tasks in data science? (Select all that apply)

**1 point**



**Data Management**



**Data Integration and Transformation**

****

**Data Visualization**

****

**Model Building**

****

**Model Deployment**

****

**Model Monitoring and Assessment**

**2.**

Question 2

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

**1 point**



GitHub



**MySQL**



**PostgreSQL**



KubeFlow



PixieDust

**3.**

Question 3

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

**1 point**



Cassandra



**Apache Kafka**

****

**Apache Nifi**

****

**Apache AirFlow**



Ceph

**4.**

Question 4

Which statement about JupyterLab is correct?

**1 point**



JuypterLab can run Python code only.



JuypterLab can run R code only.



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



JuypterLab can run R and Python code only.

**5.**

Question 5

Which statement about RStudio is correct?

**1 point**



**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.**

Question 6

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

**1 point**



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.**

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

**1.**

Question 1

Which of the following statements is true?

**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

**2.**

Question 2

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

**1 point**



CouchDB



Oracle



PostgreSQL



MongoDB



MariaDB



MySQL

**3.**

Question 3

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

**1 point**



Free Software and Open Source can be used interchangeably.



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



Most of Free Software licenses also qualify for Open Source.



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

**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 point**



True



False

**5.**

Question 5

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

**1 point**



R



Julia



Java



Javascript



Scala



SQL



All of the above

**6.**

Question 6

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

**1 point**



Oracle



PyTorch



TensorFlow.js



Apache Spark



GNU



Caffe

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

## Practice Quiz - Packages, APIs, Data Sets, Models

**Total points**6

### 1.

Question 1

Which scientific computing library provides data structures and data analysis tools for Python?

**1 point**



TensorFlow



**Pandas**



YumPies



Seahorse

### 2.

Question 2

What does the acronym API stand for?

**1 point**



Abstract Programming Interface



Abstract Python Interface



**Application Programming Interface**



Algorithmic Programming Interface

### 3.

Question 3

True or False: Open data is always distributed under a Community Data License Agreement.

**1 point**



True



**False**

### 4.

Question 4

Which of the following is not a type of Machine Learning?

**1 point**



**Supervised teaching**



Supervised learning



Reinforcement learning



Unsupervised learning

### 5.

Question 5

Which of the following is NOT a deep learning framework?

**1 point**



TensorFlow



PyTorch



Keras



**Tommy**

### 6.

Question 6

Fill in the blank: The MAX model-serving microservices expose a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ that applications use to consume a model.

**1 point**



**REST API**



Java API



Scala API



Python API

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

### 1.

Question 1

Which are the three most used languages for data science? (Select all that apply.)

**1 point**



**R**

****

**Python**

****

**SQL**



Scala



Java

### 2.

Question 2

Is it possible to use machine learning within a web browser with Javascript?

**1 point**

****

**Yes**



No

### 3.

Question 3

Comma Separated Values (CSV) is a commonly used format to store:

**1 point**

****

Hierarchical or network data



**Tabular data**



All of the above

### 4.

Question 4

Classification models can be used to determine whether:

**1 point**



An email is likely spam.



A video contains a specific sound.



An image contains a dog.



**All of the above.**

### 5.

Question 5

Fill in the blank: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ is the heart of every organization.

**1 point**



Open source



Integration



**Data**



The cloud

### 6.

Question 6

What does the “BI” in BI Tools stand for?

**1 point**



Business Intelligence



Build Information



Build Integration



**Business Integration**

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

**Practice Quiz - Jupyter Notebook**

**Total points**3

**1.**

Question 1

Which of the following functions do Jupyter Notebook unify?

**1 point**



Editing and execution of source code.



Editing and display of documentation.



Visualization of charts.



**All of the above.**

**2.**

Question 2

Which statement is true about Jupyter Notebook?

**1 point**



Jupyter Notebook is a commercial product of IBM.



**Jupyter Notebook is free and open source.**

**3.**

Question 3

What is a Jupyter Notebook kernel?

**1 point**



**It is a wrapper running on the Jupyter server encapsulating the programming language interpreter.**



It is part of the operating system the Jupyter server runs on.

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

## Practice Quiz - RStudio IDE

**Total points**3

### 1.

Question 1

Which of the following functions does RStudio unify? (Select all that apply.)

**1 point**



Storing of data.



**Editing and execution of source code.**



**Display of the R Console.**



**Visualization of plots.**



**Visualization of data in table form.**

### 2.

Question 2

Which statement is true about the RStudio IDE?

**1 point**



RStudio is a commercial product of IBM.



**RStudio is free and open source.**

### 3.

Question 3

Which statement about R packages is correct?

**1 point**



R doesn't require any packages to be installed since it contains all functionality necessary which a data scientists ever requires.



**R currently supports more than 15,000 packages which can be installed to extend R's functionality**.

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

## Practice Quiz - GitHub

**Total points**3

### 1.

Question 1

Which of the following statements are true? (Select all that apply.)

**1 point**



**Git is very useful for data science as well, since data science often involves a lot of source code to be written and managed.**



Git is an integrated development environment for data science.



**Git is a system for version control of source code**.

### 2.

Question 2

Which of the following statements about repositories are correct? (Select all that apply.)

**1 point**



The remote repository is only accessible by myself.

****

**The local repository is only accessible by myself.**



**The staging is only accessible by myself.**

****

**The remote repository is accessible by all contributors.**



The local repository is accessible by all contributors.

### 3.

Question 3

What is the best process for contributing a bugfix to a foreign repository?

**1 point**



Ask the repository owner for write access to the repository.



Send the fix via email to the author.



**Fork the repository, update the fork and create a pull request.**

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

### 1.

Question 1

What tool do most Python developers use?

**1 point**



**Jupyter Notebooks / JupyterLab**



RStudio

### 2.

Question 2

True or false? Jupyter Notebooks / JupyterLab support development in R.

**1 point**



**True**



False

### 3.

Question 3

Which tool unifies documentation, source code and data visualizations into a single document?

**1 point**



VSCode



Notepad



**Jupyter Notebooks / JupyterLab**

### 4.

Question 4

Which command is used to install packages in R?

**1 point**



**install.packages("package name")**



install("package name")



package("package name")



install.package("package name")

### 5.

Question 5

True or False: The Jupyter Notebook kernel must be installed on a local server.

**1 point**



True



**False**

### 6.

Question 6

True or false? RStudio supports development in Python.

**1 point**



**True**



False

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

## Practice Quiz - Watson Studio

**Total points**10

### 1.

Question 1

Fill in the blank: In Watson Studio, a \_\_\_\_\_\_\_\_\_\_\_\_ is how you organize your resources to achieve a particular goal. Resources can include data, collaborators, and analytic assets like notebooks and models.

**1 point**



Job



**Project**



Notebook



Asset

### 2.

Question 2

Fill in the blank: It's a best practice to remove or replace \_\_\_ **Credentials**

\_\_\_\_\_\_\_\_\_\_ before publishing to GitHub.

**1 point**



Markdown text



**Credentials**



Charts



Code cells

### 3.

Question 3

Which of the following do you need to create in order to publish a notebook to your GitHub repository?

**1 point**



Apps



Profile



**Access token**



Login credential

### 4.

Question 4

Fill in the blank: If you'd like to schedule a notebook in Watson Studio to run at a different time you can create a(n) \_\_\_\_\_\_\_\_\_\_\_\_\_.

**1 point**



**Job**



Asset



API



Markdown cell

### 5.

Question 5

Fill in the blank: On the environments tab you can define the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

**1 point**



Hardware size.



Software configuration.



Runtime configuration for notebook editor.



Runtime configuration for flow editor.



**All of the above.**

### 6.

Question 6

Fill in the blank: When sharing a read only version of a notebook, you can choose to share \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

**1 point**



Only text and output.



All content, excluding sensitive code cells.



All content including code.



A permalink.



**All of the above.**

### 7.

Question 7

Fill in the blank: When working in a Jupyter Notebook, before returning to a project, it's important to \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

**1 point**



Insert cells.



**Save your notebook.**



Insert to code.



Run cells.

### 8.

Question 8

Fill in the blank: Before running a notebook, it's a best practice to \_\_\_\_\_\_\_\_\_\_\_\_ to describe what the notebook does.

**1 point**



Insert a cell at the bottom of the notebook.



Refresh your page.



**Insert a cell at the top of the notebook.**



Delete notebook cells.

### 9.

Question 9

Fill in the blank: In the \_\_\_\_\_\_\_\_\_\_\_\_\_ tab you can define the hardware size and software configuration for the runtime associated with Watson Studio tools such as notebooks.

**1 point**



Overview



Assets



**Environments**



Settings

### 10.

Question 10

Fill in the blank: IBM Cloud uses \_\_ **Resource groups**

\_\_\_\_\_\_\_\_\_\_\_\_ as a way for you to organize your account resources in customizable groupings so that you can quickly assign users access to more than one resource at a time.

**1 point**



Services



Projects



Catalogs



**Resource groups**

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

## Practice Quiz - Other IBM Tools

**Total points**12

### 1.

Question 1

Which products (of those we covered) allow you to build data pipelines using graphical user interface and no coding?

**1 point**



Only IBM SPSS Statistics.



Only IBM SPSS Modeler.



OpenScale



**IBM SPSS Modeler and Modeler Flows in Watson Studio.**



All of the above.

### 2.

Question 2

Which features of Data Refinery help save hours and days of data preparation?

**1 point**



Flexibility of using Intuitive user interface and coding templates enabled with powerful operations to shape and clean data.



Data visualization and profiles to spot the difference and guide data preparation steps.



Incremental snapshots of the results allowing the user to gauge success with each iterative change.



Saving, editing and fixing the steps provide ability to iteratively fix the steps in the flow.



**All of the above.**

### 3.

Question 3

Watson Knowledge Catalog provides what functionality?

**1 point**



Process data, build and deploy models.



Create data and deploy models into production.



**Catalog data and ML assets, help to find relevant assets, keep track of asset lineage, enforce data governance.**



Build data and water pipelines.



Catalog all books mentioning Doctor Watson and Sherlock Holmes.

### 4.

Question 4

Fill in the blank: PMML, PFA, and ONNX are \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

**1 point**



Robots that are plotting to take over the planet.



**Open standards for predictive model serialization, exchange, and deployment**.



Abbreviations for machine learning algorithm names.



Codes for getting rid of undesired data or models.



Passwords for some super-secret system.

### 5.

Question 5

Which node must be used in Modeler flows before any modeling node?

**1 point**



Output node



**Type node**



Derive node



Auto Numeric node



All of the above

### 6.

Question 6

Fill in the blank: Auto Classification node can be used for data with \_\_\_\_\_\_\_\_\_\_\_\_\_\_.

**1 point**



No target variables.



**A categorical target variable.**



A continuous target variable.



Any target variable.



All of the above

### 7.

Question 7

Fill in the blank: Auto Numeric node can be used for data with \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

**1 point**



No target variables.



A categorical target variable.



**A continuous target variable.**



Any target variable.



All of the above.

### 8.

Question 8

IBM SPSS Modeler evolved from which product?

**1 point**



**Clementine**



IBM DB2



SPSS



Netezza



Oracle

### 9.

Question 9

Fill in the blank: IBM SPSS Statistics syntax can be created using \_\_\_\_\_\_\_\_\_\_\_.

**1 point**



AutoAI



Watson Studio Modeler flows.



IBM SPSS Modeler streams.



**Graphical user interface of IBM SPSS Statistics product or syntax editor.**

****

OpenScale

### 10.

Question 10

AutoAI provides which of the following services?

**1 point**



Monitoring for fairness, bias, and model drift.



**Automatic finding of optimal data preparation steps, model selection, and hyperparameter optimization.**



Cataloging data and model assets.



Creating SPSS syntax.



All of the above.

### 11.

Question 11

OpenScale provides which of the following services?

**1 point**



Creating SPSS syntax.



Automatic finding of optimal data preparation steps, model selection, and hyperparameter optimization.



Cataloging data and model assets.



**Monitoring for fairness, bias, and model drift.**



All of the above.

### 12.

Question 12

Predictive Model Markup Language (PMML) was created by which entity?

**1 point**



IBM



Microsoft



Oracle



SPSS



**The Data Mining Group**

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

### 1.

Question 1

Data Refinery provides which of the following services?

**1 point**



Catalog the data assets.



Monitor for bias and model drift.



Visualize and prepare data.



Automatically build models.



All of the above.

### 2.

Question 2

IBM SPSS Modeler includes what kind of models?

**1 point**



Classification models (for data with a categorical target).



Regression models (for data with a continuous target).



Clustering models (for data with no target variables).



Other kinds of models.

****

**All of the above.**

### 3.

Question 3

Open Neural Network eXchange (ONNX) was originally created for what models?

**1 point**



Regression models



Decision trees



Clustering models



Support Vector Machines (SVM).



**Deep learning models.**

### 4.

Question 4

Fill in the blank: If you'd like to schedule a notebook in Watson Studio to run at a different time, you can create a(n) \_\_\_\_\_\_\_\_.

**1 point**



markdown cell



API



asset



**job**

### 5.

Question 5

Fill in the blank: In the \_\_\_\_\_\_\_\_\_\_ tab you can define the hardware size and software configuration for the runtime associated with Watson Studio tools such as Notebook.

**1 point**



**environments**



settings



overview



assets

### 6.

Question 6

Fill in the blank: It's a best practice to remove or replace \_\_\_\_\_\_\_\_\_\_\_\_\_ before publishing to GitHub.

**1 point**



markdown text



charts



code cells

****

**credentials**

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)

**RUBRIC**

Is a Jupyter Notebook viewable on Watson Studio at the provided URL?



0 pts

No URL for a Jupyter Notebook has been submitted.



1 pt

URL for a Jupyter Notebook has been submitted and the Notebook contents can be viewed by going to the URL. However, it is not viewable in Watson Studio.



2 pts

Yes, the Jupyter Notebook is viewable on Watson Studio using the provided URL.

**OVERALL ASSIGNMENT RUBRIC**

Is the title of the notebook similar to "My Jupyter Notebook on IBM Watson Studio" (spelling or grammar mistakes should not be penalized). To the best of your ability, does the title appear to be in H1 header styling? (**2 marks**)



0 pts

None of the criteria are fulfilled.



1 pt

Some of the criteria are fulfilled.



2 pts

All of the criteria are fulfilled.

Does Cell 2 appear to include the **author's name in bold**, as well as **an occupation in a separate line below in regular font? (2 marks)**



0 pts

None of the criteria are fulfilled.



1 pt

Some of the criteria are fulfilled.



2 pts

All of the criteria are fulfilled.

Does Cell 3 appear to include a sentence describing the author's interest in data science, in *italic* formatting? (The content or coherence of the sentence should not be evaluated.) (**2 marks**)



0 pts

None of the criteria are fulfilled.



1 pt

Some of the criteria are fulfilled.



2 pts

All of the criteria are fulfilled.

Does Cell 4 appear to have some text describing the code in Cell 5, and does the text in Cell 4 appear to be in **H3 header** formatting? (**2 marks**)



0 pts

None of the critera are fulfilled.



1 pt

Some of the criteria are fulfilled.



2 pts

All of the criteria are fulfilled.

Does Cell 5 contain code, and is there an output displayed from the code executed in Cell 5? (Please judge the presence of the output of the code, not its content. For example, error outputs should not be penalized.) (**2 marks**)



0 pts

None of the criteria are fulfilled.



1 pt

Some of the criteria are fulfilled.



2 pts

All of the criteria are fulfilled.

Using Markdown or HTML, does the final cell(s) include at least 3 of the following: horizontal rule, bulleted list, numbered list, tables, hyperlinks, images, code/syntax highlighting, blocked quotes, strikethrough? (**3 marks**)



0 pts

None of the criteria are fulfilled.



1 pt

The final cell(s) only include **ONE** of the mentioned formatting styles.



2 pts

The final cell(s) only include **TWO** of the mentioned formatting styles.



3 pts

The final cell(s) include **THREE or more** of the mentioned formatting styles.

<https://eu-gb.dataplatform.cloud.ibm.com/analytics/notebooks/v2/a189df3a-e4e5-4630-828a-a8ea12101fb7/view?access_token=612e72e7aa3f46c55b59cf634a90cdc77b6a0589fae552032cd129c8ab369b00>

# My Jupiter notebook on IBM Watson Studio

Djamel Benkhettou I wish to be a data scientist

I am interested in data science because in simple word it's THE FUTURE

**The following code describe what three brothers Mark, Jhon, Edward, buyed when they were in the sweet shop**

Mark=2

Jhon=1

Edward=8

AmountTheCandies=Mark+Jhon+Edward

AmountTheCandies

**The Table**[**¶**](https://eu-gb.dataplatform.cloud.ibm.com/data/notebooks/converter/assets/a189df3a-e4e5-4630-828a-a8ea12101fb7?access_token=612e72e7aa3f46c55b59cf634a90cdc77b6a0589fae552032cd129c8ab369b00&project=4db88985-b873-42de-aa92-745e44bb8ac2#The-Table)

| **Name** |  | **Age** |  | **Type of candy** |  | **Quantity of candy** |
| --- | --- | --- | --- | --- | --- | --- |
| Mark |  | 7 |  | lollipops candy |  | 2 |
| Jhon |  | 11 |  | Chocolate bar |  | 1 |
| Edward |  | 9 |  | Chewing gums |  | 8 |
| ----- |  | ----- |  | ----- |  | ----- |
| Quantity of candy bayed |  | ----- |  | ----- |  | 11 |

### 1.

Question 1

SQL is what type of database management system?

**1 point**



Object-oriented



Network



Hierarchical



**Relational**

### 2.

Question 2

PyTorch is what type of Python library?

**1 point**



**Machine learning**



Data visualization



HTTP



Linear algebra

### 3.

Question 3

What type of model would you use if you wanted to find the relationship between dependent and independent variables?

**1 point**



Clustering model



**Regression model**



Classification model

### 4.

Question 4

What is Data Asset Management also known as?

**1 point**



Perimeter security



**Data governance**



Data stewardship



Data strategy

### 5.

Question 5

When learning data science what open source tools are the most used?

**1 point**



VSCode



Notepad



**RStudio**



Spyder



**Jupyter Notebooks / JupyterLab**

### 6.

Question 6

RStudio is the tool most R developers use.

**1 point**



No



**Yes**

### 7.

Question 7

What type of environment is RStudio?

**1 point**



Software Development Kit (SDK)



Integrated Development Environment (IDE)



Framework



API

### 8.

Question 8

Which of the following does Jupyter notebooks support?

**1 point**



Numerous programming languages



All of the above



GitHub



Data visualizations

### 9.

Question 9

What type of assets does the Watson Knowledge Catalog let you discover in Watson Studio?

**1 point**



AutoAI



**Machine Learning**



Modeler Flows



Data Refinery

### 10.

Question 10

What are some of the Data Refinery abilities when working with data?

**1 point**



Can only analyze small amounts of data at a time



**Analyzes and transforms data quickly**



Analyzes data but cannot be saved for future use



Analyzes data but takes longer to prepare

### 11.

Question 11

What type of node is used to define metadata for features in Modeler flows?

**1 point**



All of the above



A modeling node



**A type node**



An output node



A data source node

### 12.

Question 12

What software package is used specifically to analyze data related to social sciences?

**1 point**



SPSS



Modelers



Charts



Graphical user interface

**Coursera Honor Code**  [Learn more](https://learner.coursera.help/hc/articles/209818863" \t "_blank)



**Final Exam**

**Latest Submission Grade 75%**

**1.**

Question 1

Which of these is a database query language?

**1 / 1 point**



Julia



Python



**SQL**



All of the Above

**Correct**

**2.**

Question 2

PyTorch is what type of Python library?

**1 / 1 point**



Linear algebra



**Machine learning**



HTTP



Data visualization

**Correct**

**3.**

Question 3

What type of model would you use if you wanted to find the relationship between a set of variables?

**1 / 1 point**



**Regression model**



Clustering model



Classification model

**Correct**

**4.**

Question 4

Which of the following are true about Data Asset Management?

**0 / 1 point**



Also known as data governance.



A crucial part of data science at the enterprise level.



To be done effectively data must be versioned and annotated with meta data.



**All of the above.**

**Incorrect**

**5.**

Question 5

What open source tool was developed and built by statisticians?

**0 / 1 point**



VSCode



RStudio



Spyder

****

**Notepad**



Jupyter Notebooks / JupyterLab

**Incorrect**

**6.**

Question 6

RStudio is the tool most R developers use.

**1 / 1 point**



**Yes**



No

**Correct**

**7.**

Question 7

Storing data in tables in a function that RStudio provides.

**0 / 1 point**



Yes



**No**

**Incorrect**

**8.**

Question 8

What does the acronym “Jupyter” mean in Jupyter Notebooks?

**1 / 1 point**



Julia, Python and Ruby



Javascript, Python and R



Julia, PHP and R



**Julia, Python and R**

**Correct**

**9.**

Question 9

Which feature lets you discover Machine Learning assets in Watson Studio?

**1 / 1 point**



**Watson Knowledge Catalog**



Modeler Flows



AutoAI



OpenScale

**Correct**

**10.**

Question 10

What process can help a Data Refinery build a repeatable Data Pipeline?

**1 / 1 point**



Manually create jobs



**Create a scheduled job on different workloads**



Manually write APIs



Create a fixed workload

**Correct**

**11.**

Question 11

What type of node is used to define metadata for features in Modeler flows?

**1 / 1 point**



A modeling node



**A type node**



A data source node



An output node



All of the above

**Correct**

**12.**

Question 12

What Modeler flow includes data management capabilities and visualization?

**1 / 1 point**



Charts



All of the above



SPSS Modeler streams



SPSS syntax



Graphical user interface

**Correct**

## Final Exam

**Latest Submission Grade 83.33%**

### 1.

Question 1

SQL is what type of database management system?

**1 / 1 point**



Object-oriented



Network



Hierarchical



**Relational**

**Correct**

### 2.

Question 2

Which of these is not a machine learning or deep learning library for Python?

**0 / 1 point**



Scikit-learn



NumPy



**Keras**



PyTorch

**Incorrect**

### 3.

Question 3

Generally speaking, which type of model is used to predict a numerical value, such as the potential sales price of a used car?

**1 / 1 point**



Clustering model



**Regression model**



Classification model

**Correct**

### 4.

Question 4

What is Data Asset Management also known as?

**1 / 1 point**



Perimeter security



**Data governance**



Data stewardship



Data strategy

**Correct**

### 5.

Question 5

When learning data science what open source tools are the most used?

**1 / 1 point**



**RStudio**

**Correct**



Notepad



Spyder



**Jupyter Notebooks / JupyterLab**

**Correct**



VSCode

### 6.

Question 6

What tool do most R developers use?

**1 / 1 point**



**RStudio**



Jupyter Notebooks / JupyterLab

**Correct**

### 7.

Question 7

Which of the following functions does RStudio provide?

**1 / 1 point**



Creating relationships between data tables.



**Editing and execution of R code.**



Documenting R code applications.



Storing data in tables.

**Correct**

### 8.

Question 8

Which of the following statements about Jupyter Notebook is correct?

**1 / 1 point**



Jupyter Notebook is a commercial product of IBM.



Jupyter Notebook is only available if installed locally on your computer.



**Jupyter Notebook supports the Visualization of data in charts.**



Jupyter Notebook provides storage of massive quantities of data in data lakes.

**Correct**

### 9.

Question 9

Which feature in Watson Studio helps to keep track of and discover relevant Machine Learning assets?

**1 / 1 point**



OpenScale



All of the above



**Watson Knowledge Catalog**



AutoAI



Modeler Flows

**Correct**

### 10.

Question 10

Creating a scheduled Job and using a custom environment to run a data flow/pipeline on different workloads can help build what?

**1 / 1 point**



APIs with automation



**Repeatable Data Pipelines for workloads of almost any size**



Not supported



Fixed workloads

**Correct**

### 11.

Question 11

What type of node is used to partition the data into a training and testing set in Modeler flows?

**0 / 1 point**



All of the above



A partition node



**A type node**



A data source node



Auto data prep node

**Incorrect**

### 12.

Question 12

What Modeler flow includes data management capabilities and visualization?

**1 / 1 point**



Graphical user interface



All of the above



Charts



SPSS syntax



**SPSS Modeler streams**

**Correct**