



70-773^{Q&As}

Analyzing Big Data with Microsoft R

Pass Microsoft 70-773 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.pass4itsure.com/70-773.html>

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft
Official Exam Center

-  **Instant Download** After Purchase
-  **100% Money Back** Guarantee
-  **365 Days** Free Update
-  **800,000+** Satisfied Customers



**QUESTION 1**

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while

others might not have a correct solution.

After you answer a question in this sections, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Microsoft SQL Server instance that has R Services (In-Database) installed.

You need to monitor the R jobs that are sent to SQL Server.

Solution: You call a function from the RevoPemaR package.

Does this meet the goal?

A. Yes

B. No

Correct Answer: B

References: <https://docs.microsoft.com/en-us/machine-learning-server/r-reference/revopemar/pemar>

QUESTION 2

You have a Microsoft SQL Server instance that has R Services (In-Database) installed. The server has a comma-separated values (CSV) file stored in the local file system.

For analytic purposes, you need to read the CSV file into a database table in the SQL Server instance.

You connect to the SQL Server instance by using SQL Server Management Studio.

What should you use from `sp_execute_external_script`?

A. `RxSqlServerData` and specify the CSV file path in the connection string

B. `rxDataStep` and specify the CSV file path as the `inFile` argument

C. `rxImportToXdf` and specify specify the CSV file as the input

D. `read.csv` and specify the CSV file path as the parameter

Correct Answer: D

QUESTION 3

Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is



repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in this series.

Start of repeated scenario

You are developing a Microsoft R Open solution that will leverage the computing power of the database server for some of your datasets.

You are performing feature engineering and data preparation for the datasets.

The following is a sample of the dataset.

```
rxGetInfo(df)
head(df)

  age incwage perwt wkswork1 state
1  50    9000    30      48  Indiana
2  41   35000    20      48  Indiana
3  55   40400    21      52  Indiana
4  56   45000    30      52  Indiana
5  46   17200    60      52  Indiana
6  49   35000    21      52  Indiana
```

End of repeated scenario.

You need to sort the data from the dataset sample and to remove duplicates by using wkswork1.

Which R code segment should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

```
rxSort(inData = [sampleInData], outFile = [sampleOutData], sortByVars = c("incwage", "perwt", "age", "wkswork1"),
```

<input type="checkbox"/>	<input type="checkbox"/>	= "wkswork1")
removeDupKeys = FALSE	dupFreqVar	
removeDupKeys = TRUE	varsToDrop	
rxMerge = FALSE	varsToKeep	
rxMerge = TRUE		

Correct Answer:

**Answer Area**

```
rxSort(inData = [sampleInData], outFile = [sampleOutData], sortByVars = c("incwage", "perwt", "age", "wkswork1"),
```

removeDupKeys = FALSE		dupFreqVar	
removeDupKeys = TRUE		varsToDrop	
rxMerge = FALSE		VarsToKeep	
rxMerge = TRUE			

QUESTION 4

Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series.

Information and details provided in a question apply only to that question.

You have a dataset that contains the physical characteristics of people. You need to visualize a relationship between height and weight for a subset of observations in the dataset.

What should you use?

- A. the Describe package
- B. the rxHistogram function
- C. the rxSummary function
- D. the rxQuantile function
- E. the rxCube function
- F. the summary function
- G. the rxCrossTabs function
- H. the ggplot2 package

Correct Answer: H

QUESTION 5

Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series.

Information and details provided in a question apply only to that question.

You need to estimate a model where the outcome variable is continuous, is in the range of $[0, \infty]$, and has a substantial mass at an exact value of 0.



Which function should you use?

- A. rxPredict
- B. rxLogit
- C. summary
- D. rxLinMod
- E. rxTweedie
- F. stepAic
- G. rxTransform
- H. rxDataStep

Correct Answer: F

QUESTION 6

You need to build a model that looks at the probability of an outcome. You must regulate between L1 and L2. Which classification method should you use?

- A. Two-Class Neural Network
- B. Two-Class Support Vector Machine
- C. Two-Class Decision Forest
- D. Two-Class Logistic Regression

Correct Answer: D

References: <https://msdn.microsoft.com/en-us/library/azure/dn905994.aspx>

QUESTION 7

Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series.

Information and details provided in a question apply only to that question.

You need to evaluate the significance of coefficients that are produced by using a model that was estimated already.

Which function should you use?

- A. rxPredict
- B. rxLogit
- C. summary



- D. rxLinMod
- E. rxTweedie
- F. stepAic
- G. rxTransform
- H. rxDataStep

Correct Answer: D

References: <https://docs.microsoft.com/en-us/machine-learning-server/r/how-to-revoscaler-linear-model>

QUESTION 8

You have an Apache Hadoop Hive data warehouse. RevoScalerR is not installed.

You need to sort the data according to the variables in the dataset.

What should you do?

- A. Connect to the database by using an ODBC connection, and then use the rxSort function.
- B. Create a table in the ORC file format.
- C. Connect to the database by using an ODBC connection, and then use the rxDataStep function.
- D. Execute a Hive query that sorts the data, and then reads the results.

Correct Answer: D

QUESTION 9

Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is

exactly the same in each question in this series.

Start of repeated scenario

You are developing a Microsoft R Open solution that will leverage the computing power of the database server for some of your datasets.

You are performing feature engineering and data preparation for the datasets.

The following is a sample of the dataset.



```
rxGetInfo(df)
head(df)
```

	age	incwage	perwt	wkswork1	state
1	50	9000	30	48	Indiana
2	41	35000	20	48	Indiana
3	55	40400	21	52	Indiana
4	56	45000	30	52	Indiana
5	46	17200	60	52	Indiana
6	49	35000	21	52	Indiana

End of repeated scenario.

You plan to score some data to create data features to address empty rows.

You have the following R code.

```
xdPath <- file.path(rxGetOption("[sampleInData], "), "inputfile.xdf")
xdfLagged <- [sampleOutDataincludingFeatures](fileext = ".xdf")
rxSort(inData = xdfPath,
       outFile = xdfLagged,
       sortByVars = "Date")
rxDataStep(inData = xdfLagged,
           outFile = xdfLagged,
           transformObjects = list(
             varToLag = "Open",
             newName = "previousOpen"),
           transformFunc = lagVar,
           append = "cols",
           overwrite = TRUE)
rxDataStep(xdfLagged,
           varsToKeep = c("Date", "Open", "previousOpen"),
           numRows = 10)
```

You need to transform the data and overwrite the current dataset.

Which R code segment should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:



Answer Area

(inData = [sampleInData], outFile=[sampleOutDataincludingFeatures],

rxCube
rxDataStep
rxExec
transform

 transformFunc = computeNonLagFeatures,

overwrite=FALSE
overwrite=TRUE
varsToDrop=All
varsToDrop=NONE

)

Correct Answer:

Answer Area

(inData = [sampleInData], outFile=[sampleOutDataincludingFeatures],

rxCube
rxDataStep
rxExec
transform

 transformFunc = computeNonLagFeatures,

overwrite=FALSE
overwrite=TRUE
varsToDrop=All
varsToDrop=NONE

)

QUESTION 10

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while

others might not have a correct solution.

After you answer a question in this sections, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You use dplyrXdf, and you discover that after you exit the session, the output files that were created were deleted.

You need to prevent the files from being deleted.

Solution: You use dplyrXdf with the persist verb.

Does this meet the goal?

A. Yes

B. No

Correct Answer: A

References: <http://blog.revolutionanalytics.com/2016/12/dplyrxdf-090-now-available.html>



QUESTION 11

You have one-class support vector machines (SVMs).

You have a large dataset, but you do not have enough training time to fully test the model.

What is an alternative method to validate the model?

- A. Use Principal Components Analysis (PCA)-Based Anomaly Detection.
- B. Replace the SVMs with two-class SVMs.
- C. Perform feature selection.
- D. Use outlier detection.

Correct Answer: A

QUESTION 12

You are running a parallel function that uses the following R code segment. (Line numbers are included for reference only.)

```
01 cp
```



To Read the [Whole Q&As](#), please purchase the [Complete Version](#) from [Our website](#).

Try our product !

100% Guaranteed Success

100% Money Back Guarantee

365 Days Free Update

Instant Download After Purchase

24x7 Customer Support

Average 99.9% Success Rate

More than 800,000 Satisfied Customers Worldwide

Multi-Platform capabilities - [Windows](#), [Mac](#), [Android](#), [iPhone](#), [iPod](#), [iPad](#), [Kindle](#)

We provide exam PDF and VCE of Cisco, Microsoft, IBM, CompTIA, Oracle and other IT Certifications. You can view Vendor list of All Certification Exams offered:

<https://www.pass4itsure.com/allproducts>

Need Help

Please provide as much detail as possible so we can best assist you.

To update a previously submitted ticket:



 <p>One Year Free Update Free update is available within One Year after your purchase. After One Year, you will get 50% discounts for updating. And we are proud to boast a 24/7 efficient Customer Support system via Email.</p>	 <p>Money Back Guarantee To ensure that you are spending on quality products, we provide 100% money back guarantee for 30 days from the date of purchase.</p>	 <p>Security & Privacy We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information & peace of mind.</p>
---	---	--

Any charges made through this site will appear as Global Simulators Limited.

All trademarks are the property of their respective owners.

Copyright © pass4itsure, All Rights Reserved.