



# 70-773<sup>Q&As</sup>

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.passapply.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 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 generate a residual based on two columns. The solution must build a trend indicator. Which function should you use?

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

Correct Answer: A

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

---

### QUESTION 2

You have a dataset that has multiple blocks and only numeric variables.

You are computing in a local compute context.

You plan to lag a variable named `x` to create a new variable named `x_lagged` by using a transform function. You will create a new element in the output of the function.

You need to minimize the number of missing values.

Which three actions should you perform? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Assign a value to the first value of `x_lagged` in the current block.
- B. Use `rxSet` to store the last value of `x_lagged` in the current block.
- C. Use `rxSet` to store the last value of `x` in the current block.
- D. Use `rxGet` to retrieve the first value of `x` in the next block to be processed.
- E. Use `rxGet` to retrieve a value stored in processing of the prior block.



Correct Answer: ACD

---

### QUESTION 3

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 `rxSetComputeContext` with the local parameter before performing operations that save results.

Does this meet the goal?

A. Yes

B. No

Correct Answer: B

---

### QUESTION 4

You need to run a large data tree model by using `rxDForest`. The model must use cross validation. Which `rxDForest` option should you use?

A. `maxSurrogate`

B. `maxNumBins`

C. `maxDepth`

D. `maxCompete`

E. `xVal`

Correct Answer: E

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

---

### 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 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 6

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 calculate a measure of central tendency and variability for the variables in a dataset that is grouped by using another categorical variable.

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: C

---



### QUESTION 7

You have cloud and on-premises resources that include Microsoft SQL Server and a big data environment in Apache Hadoop.

You have 50 billion fact records.

You need to build time series models to execute forecasting reports on the fact records.

What should you use?

- A. RxSpark on the Hadoop cluster
- B. RxHadoopMR on the Hadoop cluster
- C. RxLocalseq on the SQL Server database
- D. RxLocalParallel on the SQL Server database

Correct Answer: A

---

### QUESTION 8

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

Correct Answer: D

---

### QUESTION 9

You are using `rxPredict` for a logistic regression model.

You need to obtain prediction standard errors and confidence intervals.

Which R code segment should you use? To answer, drag the appropriate values to the correct targets. Each value may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.



NOTE: Each correct selection is worth one point.

Select and Place:

Values	Answer Area
<input type="text" value="confidence"/>	<pre>model &lt;- <input type="text" value="Value"/> (default ~ year + creditScore + yearsEmploy + ccDebt, data = trainingDataFileName, blocksPerRead = 2, verbose = 1, reportProgress=2, covCoef= <input type="text" value="Value"/> ) rxPredict(model, data = targetDataFileName, outData = targetDataFileName, computeStdErr = <input type="text" value="Value"/> , interval = "<input type="text" value="Value"/> ", overwrite=TRUE)</pre>
<input type="text" value="FALSE"/>	
<input type="text" value="glm"/>	
<input type="text" value="none"/>	
<input type="text" value="rxLinMod"/>	
<input type="text" value="rxLogit"/>	
<input type="text" value="TRUE"/>	

Correct Answer:

Values	Answer Area
<input type="text" value="confidence"/>	<pre>model &lt;- <input type="text" value="rxLogit"/> (default ~ year + creditScore + yearsEmploy + ccDebt, data = trainingDataFileName, blocksPerRead = 2, verbose = 1, reportProgress=2, covCoef= <input type="text" value="TRUE"/> ) rxPredict(model, data = targetDataFileName, outData = targetDataFileName, computeStdErr = <input type="text" value="TRUE"/> , interval = "<input type="text" value="confidence"/> ", overwrite=TRUE)</pre>
<input type="text" value="FALSE"/>	
<input type="text" value="glm"/>	
<input type="text" value="none"/>	
<input type="text" value="rxLinMod"/>	
<input type="text" value="rxLogit"/>	
<input type="text" value="TRUE"/>	

### 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

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 get all of the deciles for a variable in a data frame.

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: D

---

### QUESTION 12

You are running a large logistic regression for 1,000 feature variables by using the `LoisticRegression()` function in the `MicrosoftML` package. All of the predictor variables are numeric.

Currently, you specify the input variables separately by using the following formula.

`Outcome ~ Feature000 + Feature001 + Feature002 + ... + Feature999`

You discover that it takes 20 minutes to estimate each model.

You need to reduce the amount of time required to estimate each model without losing any information in the predictors.

What should you do?



- A. Use `stepControl()` to perform stepwise regression to limit the number of variables that contribute to the model.
- B. Use `selectFeatures()` to select the features that provide the most information about the outcome variable.
- C. Use `princomp()` on the correlation matrix of Features, and then use only the first 100 principle components to reduce the number of input variables.
- D. Use `concat()` to create a single array variable named Features, and then specify a new formula named Outcome ~ Features.

Correct Answer: D

[70-773 VCE Dumps](#)

[70-773 Practice Test](#)

[70-773 Study Guide](#)





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.passapply.com/allproducts>

## Need Help

Please provide as much detail as possible so we can best assist you.  
To update a previously submitted ticket:



 <p><b>One Year Free Update</b> 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><b>Money Back Guarantee</b> 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><b>Security &amp; Privacy</b> We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information &amp; 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 © passapply, All Rights Reserved.