Attempt 1

<div style="float:right; border:1px solid #ccc; padding:8px;">**All questions**</div>

Question 1: Correct

Snowflake is available on premise

True

False	(Correct)

Explanation

Snowflake's data warehouse is a true SaaS offering. More specifically:

There is no hardware (virtual or physical) for you to select, install, configure, or manage.

There is no software for you to install, configure, or manage.

Ongoing maintenance, management, and tuning is handled by Snowflake.

Snowflake runs completely on cloud infrastructure. All components of Snowflake's service (other than an optional command line client), run in a public cloud infrastructure.

Snowflake uses virtual compute instances for its compute needs and a storage service for persistent storage of data. Snowflake cannot be run on private cloud infrastructures (on-premises or hosted).

Snowflake is not a packaged software offering that can be installed by a user. Snowflake manages all aspects of software installation and updates.

https://docs.snowflake.com/en/user-guide/intro-key-concepts.html

Question 2: Incorrect

Pick the true statement for snowflake architecture

Shared nothing architecture	(Incorrect)

Shared disk architecture

| Multi-Cluster Shared Data architecture | (Correct) |

Explanation
Snowflake's architecture is a hybrid of traditional shared-disk database architectures and shared-nothing database architectures. Similar to shared-disk architectures, Snowflake uses a central data repository for persisted data that is accessible from all compute nodes in the data warehouse. But similar to shared-nothing architectures, Snowflake processes queries using MPP (massively parallel processing) compute clusters where each node in the cluster stores a portion of the entire data set locally. This approach offers the data management simplicity of a shared-disk architecture, but with the performance and scale-out benefits of a shared-nothing architecture.

https://docs.snowflake.com/en/user-guide/intro-key-concepts.html

Question 3: **Correct**

The three key layers of snowflake are

| Extraction, Ingestion, Load | |

| Database storage, Query Processing, Cloud Services | (Correct) |

| Database, Virtual Warehouse, Data Experience | |

Explanation
Snowflake's unique architecture consists of three key layers:

Database Storage

Query Processing

Cloud Services

https://docs.snowflake.com/en/user-guide/intro-key-concepts.html

Question 4: **Correct**

How is data loaded into snowflake?

| Snowflake loads the data in parquet format on the underlying cloud storage |

underlying cloud storage

Snowflake loads the data in JSON format on the
underlying cloud storage

Snowflake reorganizes the data into its
internal optimized, compressed, columnar
format at stores on the underlying cloud   **(Correct)**
storage

Explanation

When data is loaded into Snowflake, Snowflake reorganizes that data into its internal
optimized, compressed, columnar format. Snowflake stores this optimized data in cloud
storage.

Snowflake manages all aspects of how this data is stored — the organization, file size,
structure, compression, metadata, statistics, and other aspects of data storage are handled
by Snowflake. The data objects stored by Snowflake are not directly visible nor accessible by
customers; they are only accessible through SQL query operations run using Snowflake.

https://docs.snowflake.com/en/user-guide/intro-key-concepts.html

Question 5: **Correct**

Query processing in snowflake is done by...

Snowflake processes queries using AWS EMR running
spark

Snowflake processes queries running spark on a EC2
instance

Snowflake process queries using 'Virtual
Warehouses'   **(Correct)**

Explanation

Query execution is performed in the processing layer. Snowflake processes queries using
"virtual warehouses". Each virtual warehouse is an MPP compute cluster composed of
multiple compute nodes allocated by Snowflake from a cloud provider.

Each virtual warehouse is an independent compute cluster that does not share compute resources with other virtual warehouses. As a result, each virtual warehouse has no impact on the performance of other virtual warehouses.

https://docs.snowflake.com/en/user-guide/intro-key-concepts.html

Question 6: **Incorrect**

Which of the below services are provided by cloud services

Metadata management	**(Correct)**

Authentication	**(Correct)**

Infrastructure management	**(Correct)**

Query execution	**(Incorrect)**

Explanation
The cloud services layer is a collection of services that coordinate activities across Snowflake. These services tie together all of the different components of Snowflake in order to process user requests, from login to query dispatch. The cloud services layer also runs on compute instances provisioned by Snowflake from the cloud provider.

Among the services in this layer:

Authentication

Infrastructure management

Metadata management

Query parsing and optimization

Access control

Question 7: **Correct**

How will you store JSON data in snowflake

Using a column with datatype as JSON

Using a column with datatype as VARCHAR

Using a column with datatype as VARIANT   **(Correct)**

Explanation
Snowflakes variant datatype allows you to store semi-structured non relational data like JSON, AVRO, XML

Question 8: **Correct**

You have two virtual warehouses in your snowflake instance. You have updated the data in the storage layer using one of the warehouses. When will the other warehouse be able to see the data

Immediately   **(Correct)**

After 30 minutes once snowflake completes data synchronization

You will need to trigger the data synchronization process for the other warehouse to see the data

Explanation
Virtual warehouses uses the same data storage layer. So If one of the warehouses updates the data, it is immediately available to all the warehouses

Question 9: **Incorrect**

Zero-copy cloning in snowflake is powered by which service?

Metadata store of the service layer   **(Correct)**

SSD cache of Virtual warehouse   **(Incorrect)**

Query result cache

Explanation

Metadata store is a key component of the services layer. It supports zero copy cloning, data sharing and time travel

Question 10: **Correct**

What influences snowflake pricing?

Amount of data queried from snowflake

Amount of data scanned during querying snowflake

Snowflake pricing is based on usage and it charges only for storage and compute	**(Correct)**

Explanation

There are only two const components for which Snowflake charges its customers. They are storage and compute. This is very important to remember when you optimize snowflake for cost.

Question 11: **Correct**

Compute cost in snowflake depends on

The actual query execution time

The query execution time and the time the query waits for the resource

The warehouse size and how long the warehouse runs	**(Correct)**

Explanation
This is the beauty of snowflake. The cost does not depend on how many queries you run on the warehouse. It depends on which warehouse size you have chose and how long the warehouse was alive.

Question 12: **Correct**

You are an account administrator and you want to use a third party product in snowflake. Where will you go to enable the third party product?

The third party product's webpage and contact them through contact us

Call snowflake support to enable the product

Enable the product through Partner Connect in Snowflake web console	**(Correct)**

Explanation
You can enable 3rd party products by navigating to partner connect on Snowflake web console

Question 13: **Correct**

Compute in snowflake is provided by

The cloud providers VM instances

Only EC2 instances on AWS

Virtual warehouses	**(Correct)**

Explanation
Snowflake is a SaaS and runs on all the three major cloud providers AWS, Azure and GCP. It abstracts the computing resources through its virtual warehouse concept. Virtual warehouses are one or more cluster of servers that provide compute resources

Question 14: **Correct**

Once you have selected a warehouse size, you cannot resize the size of the warehouse

True

False	**(Correct)**

Explanation
You can always resize a warehouse to make it bigger or smaller. However the queries that are already in process will not be able to use the resized warehouse. The resized warehouse will be used by all future queries

Question 15: **Correct**

What is the easiest way to monitor the queries that are run on snowflake?

Create a tableau dashboard and connect to snowflake

All queries go to cloudwatch and use cloudwatch to monitor

Click on the the History tab and monitor all queries that are executed in the last 14 days	**(Correct)**

Explanation
The history page displays all queries that are executed in last 14 days.

Question 16: **Correct**

You are a snowflake architect hired by an organization. They want you to design their warehouse strategy. Which one of the strategy will you pick up.

Both loading and analysis of data will be done by a single warehouse to reduce cost

You do not need to use warehouse for loading data, only query analysis will require a warehouse

You will recommend to use a multi-warehouse strategy. The load workload will be done by one ware house and the query analysis workload will be done by another warehouse

(Correct)

Explanation

Multi-warehouse strategy allows you to isolate workloads by the workload type. These also helps manage access, cost accounting in a better way. Hence you should recommend multi-warehouse strategy. Please note that cost does not depend on how many warehouses you have. Compute cost depends on the size of the warehouse and the amount of time they are running.

Question 17: **Correct**

In your snowflake environment, you have a medium warehouse. The medium warehouse is used by business to run adhoc queries. The warehouse has auto suspend set at 15 minutes. You have noticed that all of the queries that run on this warehouse finishes within a minute. What will you do to optimize cost of compute in this case?

Since all the queries are completing in a minute, you should delete this warehouse and tell the users to use another existing warehouse

You will tell the users that they do not need snowflake to run their queries, they should load their data in another on premise database and run query from there

You will recommend to reduce the auto suspend time to 1 minute

(Correct)

Explanation

All virtual warehouses can be setup with 'Auto Suspend' setting. This setting suspends the warehouse if it is idle for the time that has been setup for the configuration. In this case, all the queries are completing in 1 minute but since the auto suspend is turned on, the warehouse will keep on running for 15 minutes. Compute cost depends on how long the warehouse is running. So, you should reduce the auto suspend time setting

Question 18: **Correct**

When a warehouse does not have enough resources available to process the queries, what happens to the incoming queries?

Queries are aborted immediately

Snowflake automatically resizes the warehouse

The queries are queued and then executed when the resources are available again	**(Correct)**

Explanation

If the warehouse is already overloaded with queries, the new queries are queued for execution. As and when the resources get available, the warehouse executes the query. You can monitor if your warehouse is overloaded by tracking the query overload parameter or going to the warehouse view in the snowflake web UI.

Question 19: **Incorrect**

In you organization, you have snowflake Enterprise addition. You notice that consistently, your queries are getting queued on the warehouses and delaying your ETL process. What are the possible solution options that you can think of?

Resize the warehouse	**(Correct)**

Use multi-cluster warehouse	**(Correct)**

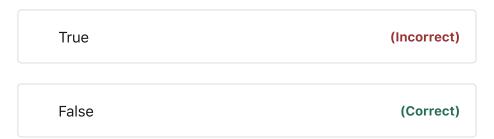| Set the auto-resize parameter of the warehouse	**(Incorrect)**

Contact snowflake support to increase the number of servers for the warehouse

Explanation

When you see that you are experiencing query overload on a consistent basis, you know that the warehouse is getting overworked. In this case you can either resize your warehouse to make it bigger or use a multi-cluster warehouse. In a multi-cluster ware house you can mention the minimum and maximum servers to instantiate. Multi-cluster warehouse will then spin up additional servers based on the query load. Multi-cluster warehouses are an enterprise edition feature

Question 20: **Incorrect**

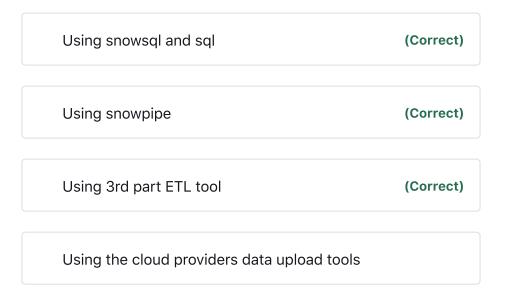You can load data using PUT command through worksheets in snowflake web ui

True	**(Incorrect)**

False	**(Correct)**

Explanation
You cannot use worksheets to load data into snowflake using PUT. PUT only works from SNOWSQL

Question 21: **Correct**

What are the valid data loading options in snowflake?

Using snowsql and sql	**(Correct)**

Using snowpipe	**(Correct)**

Using 3rd part ETL tool	**(Correct)**

Using the cloud providers data upload tools

Explanation
You can bulk load data into snowflake using SNOWSQL, Snowpipe or 3rd party ETL tools like Informatica, Talend, Pentaho etc.

Question 22: **Correct**

Snowflake allows only loading of structured data

True

False	(Correct)

Explanation
In snowflake, you can load and store both structured and semi-structured data like XML, JSON, AVRO, Parquet and ORC. Variant datatype is usually used store semi-structured data

Question 23: **Incorrect**

What are the usual data loading steps in snowflake?

Source -> Snowflake Stage -> Snowflake table	(Correct)

Source -> Cloud storage -> Snowflake stage -> Snowflake table	(Incorrect)

Source -> Snowflake temp table -> Snowflake transient table-> Snowflake permanent table	

Explanation
Data from source is usually batch loaded to Snowflake stage which is a snowflake object that loads the data as raw files in the underlying cloud storage object. From the snowflake stage, the data is copied into snowflake tables for further processing.

Question 24: **Incorrect**

You have several CSV files loaded into your named snowflake internal stage. You want to load files from the stage into a table using pattern matching to only load uncompressed CSV files whose names include the string `sales`. Which is the command that you will use to do the same?

```
copy into mytable
```

```
      from @my_int_stage
      file_format = (format_name = myformat)    (Correct)
      pattern='.*sales.*[.]csv';
```

```
copy into mytable
    from @my_int_stage
    regex='.*sales.*[.]csv';                    (Incorrect)
```

```
copy into mytable
    from @my_int_stage
    match_pattern='.*sales.*[.]csv';
```

Explanation

To load data from staged files to an existing table, the files must already be staged in one of the following locations:

1. Named internal stage (or table/user stage). Files can be staged using the PUT command.

2. Named external stage that references an external location (Amazon S3, Google Cloud Storage, or Microsoft Azure).

3. External location (Amazon S3, Google Cloud Storage, or Microsoft Azure).

In this case we have the file staged in a named internal stage. To load files matching a pattern, we will need to use the PATTERN keyword.

```
[ PATTERN = '<regex_pattern>' ]
```

https://docs.snowflake.com/en/sql-reference/sql/copy-into-table.html

Question 25: **Correct**

Snowflake is compliant with which certifications?
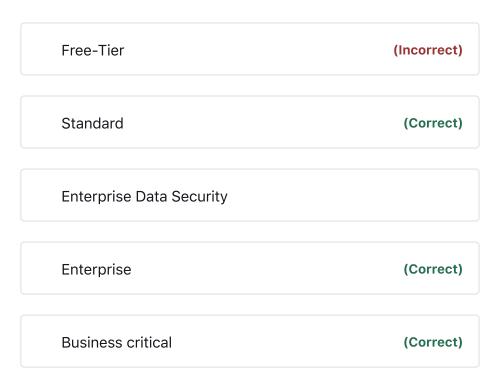
| HIPAA, PCI DSS, HIPAA, PCI DSS and FedRAMP | **(Correct)** |

| HIPAA, PCI DSS |

| Only HIPAA |

Only FedRAMP

Question 26: **Incorrect**

When you sign up for a snowflake trial account, what are three snowflake editions that are offered

Free-Tier	**(Incorrect)**

Standard	**(Correct)**

Enterprise Data Security	

Enterprise	**(Correct)**

Business critical	**(Correct)**

Question 27: **Incorrect**

What are the snowflake editions available as of today?

Standard Edition,Enterprise Edition,Business Critical Edition,Virtual Private Snowflake	**(Correct)**

Standard Edition,Enterprise Edition,ESD,Business Critical Edition,Virtual Private Snowflake	

Standard Edition,Enterprise Edition	**(Incorrect)**

Explanation

Snowflake offers multiple editions to chose from.

The Snowflake Edition that your organization chooses determines the unit costs for the credits and the data storage you use. Other factors that impact unit costs are the regions where your Snowflake account is located and whether it is an *On Demand* or *Capacity* account:

On Demand: Usage-based pricing with no long-term licensing requirements.

Capacity: Discounted pricing based on an up-front Capacity commitment.

https://docs.snowflake.com/en/user-guide/intro-editions.html#

Question 28: **Incorrect**

Which snowflake edition supports private communication between Snowflake and your other VPCs through AWS PrivateLink

Standard	

Premier	

Enterprise	**(Incorrect)**

Business critical	**(Correct)**

Explanation

Only business critical and above supports private communication between Snowflake and your other VPCs through AWS PrivateLink or your other VNets through Azure PrivateLink

Question 29: **Correct**

Which cloud providers are supported by snowflake

Google Cloud Platform	**(Correct)**

Azure	**(Correct)**

AWS	(Correct)

Explanation

A Snowflake account can be hosted on any of the following cloud platforms:

Amazon Web Services (AWS)

Google Cloud Platform (GCP)

Microsoft Azure (Azure)

On each platform, Snowflake provides one or more regionswhere the account is provisioned.

Note : GCP support has started from 2020 - https://www.snowflake.com/news/snowflake-announces-general-availability-on-google-cloud/

https://docs.snowflake.com/en/user-guide/intro-cloud-platforms.html

Question 30: **Incorrect**

Snowflake automatically stores data in encrypted form in all editions

True	(Correct)

False	(Incorrect)

Explanation

Protecting customer data is one of Snowflake's highest priorities. Snowflake encrypts all customer data by default, using the latest security standards, at no additional cost. Snowflake provides best-in-class key management, which is entirely transparent to customers. This makes Snowflake one of the easiest to use and most secure data warehouses

https://docs.snowflake.com/en/user-guide/security-encryption.html

Question 31: **Incorrect**

You have created a lot of shares with role as ACCOUNTADMIN. You want to create the shares with SYSADMIN as role, so you have granted 'CREATE SHARE' privilege to SYSADMIN. How do you change the ownership of the existing shares?

Execute 'GRANT OWNERSHIP ON <SHARE NAME> TO ROLE SYSADMIN'	(Incorrect)

NAME> TO ROLE SYSADMIN

Execute 'GRANT OWNERSHIP TO ROLE SYSADMIN ON
<SHARE NAME>

The only option is to drop and recreate the
share. Ownership of a share cannot be   **(Correct)**
granted to another role

Explanation

```
The ACCOUNTADMIN role has the privileges to create a share.
You can also grant CREATE SHARE and IMPORT SHARE to other roles, enabling the ta
```

Ownership of a share cannot be transferred to another role. It needs to be deleted and
recreated. This is by design and is done for security reasons. Users with this role can expose
any object they own.

```
Ownership of a share, as well as the objects in the share, may be either through
Also, it's possible for the same role to own a share and the objects in the shar
```

https://docs.snowflake.com/en/user-guide/security-access-privileges-shares.html#enabling-
non-accountadmin-roles-to-perform-data-sharing-tasks

Question 32: **Incorrect**

Select all options that are true for ORDER BY in snowflake

All data is sorted according to the numeric
byte value of each character in the ASCII   **(Correct)**
table. UTF-8 encoding is supported

For numeric values, leading zeros before the
decimal point and trailing zeros (**0**) after the   **(Correct)**
decimal point have no effect on sort order.

Unless specified otherwise, NULL values are
considered to be higher than any non-NULL
values. As a result, the ordering for NULLS
depends on the sort order:

(Correct)

1. If the sort order is `ASC`, NULLS are returned last; to force NULLS to be first, use `NULLS FIRST`.
2. If the sort order is `DESC`, NULLS are returned first; to force NULLS to be last, use `NULLS LAST`.
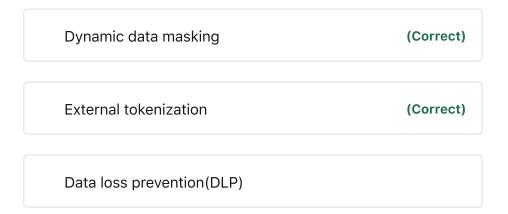
Explanation
You should at least remember the below two points. Even if this does not appear in certification, you need to know this when you will work with snowflake

1. If the sort order is `ASC`, NULLS are returned last; to force NULLS to be first, use `NULLS FIRST`.

2. If the sort order is `DESC`, NULLS are returned first; to force NULLS to be last, use `NULLS LAST`.

https://docs.snowflake.com/en/sql-reference/constructs/order-by.html#usage-notes

Question 33: **Correct**

Column level security in Snowflake allows the application of a masking policy to a column within a table or view. Which two features are related to column level security

Dynamic data masking	**(Correct)**

External tokenization	**(Correct)**

Data loss prevention(DLP)	

Explanation
Column level security is still in preview feature and is available in enterprise edition or above. This is a very useful feature and you will need this while designing your solution on snowflake

https://docs.snowflake.com/en/user-guide/security-column.html#column-level-security

Question 34: **Incorrect**

In snowflake, what are the two major cost categories

Storage	**(Correct)**

| Compute | (Correct) |

| Cloud services | (Incorrect) |

| Network Traffic | |

Explanation
This is a very important thing to remember. In snowflake you pay only for storage and compute. So, everything that you do in snowflake, you must keep this in perspective. For example while designing queries, check if your query can be satisfied by the metadata store. Metadata store in snowflake is in the services layer and does not charge you for compute or storage. Running MIN() and MAX() functions on your field doesn't require any warehouse compute as it is stored in the metadata store, hence it does not cost you anything. Also queries like 'SHOW COLUMNS' does not require a running warehouse to execute as the data comes from the metadata store

Question 35: **Correct**

What are the installment options available for snowflake?

| Snowflake hosted accounts on AWS | (Correct) |

| Hybrid on-premise and cloud | |

| On-premise | |

| Snowflake hosted accounts on Azure | (Correct) |

Explanation
This is very important to know. Snowflake does not have an on-premise option. Since last year it was available on AWS and Azure. From this year onwards, it is also available on GCP

Snowflake's data warehouse is a true SaaS offering. More specifically:

1. There is no hardware (virtual or physical) for you to select, install, configure, or manage.

2. There is no software for you to install, configure, or manage.

3. Ongoing maintenance, management, and tuning is handled by Snowflake.

Snowflake **runs completely on cloud infrastructure**. All components of Snowflake's service (other than an optional command line client), run in a public cloud infrastructure.

https://docs.snowflake.com/en/user-guide/intro-key-concepts.html#data-warehouse-as-a-cloud-service

Question 36: **Incorrect**

How is the data storage cost computed for snowflake?

Based on amount of compressed data	**(Correct)**

Based on amount of un-compressed data	**(Incorrect)**

Based on daily average of data stored	**(Correct)**

Amount stored on first day of month

Amount stored on last day of month

Explanation
Usage for **data storage** is calculated on the daily average amount of data (in bytes) stored in the system for:

1. Files staged for bulk data loading/unloading (can be stored compressed or uncompressed).

2. Database tables, including historical data for Time Travel (always compressed by Snowflake).

3. Fail-safe for database tables (always compressed by Snowflake).

4. Clones of database tables that reference data deleted in the table that owns the clones.

Please look at point #3 above. Snowflake has a storage cost for fail-safe. Now by this time you may already know that Transient tables do not have fail-safe. So for use cases where it makes sense to use transient tables, you can choose transient table to save some costs. Note that storage costs are not that much so weigh in on what you are gaining by going with transient table.

Question 37: Incorrect

Which type of data incur snowflake storage cost?

Data stored in permanent tables	**(Correct)**

Data retained to enable data recovery(fail-safe and time travel)	**(Correct)**

Cached results	**(Incorrect)**

Semi structured data stored in the cloud storage(like AWS S3, GCS)

Explanation
Storage cost is charged for below type of data

1. Files stored in Snowflake locations (i.e. user and table stages or internal named stages) for bulk data loading/unloading. The user who stages a file can choose whether or not to compress the file to reduce storage.

2. Data stored in database tables, including historical data maintained for Time Travel. Snowflake automatically compresses all data stored in tables and uses the compressed file size to calculate the total storage used for an account.

3. Historical data maintained for Fail-safe.

Question 38: Correct

How often does snowflake release new feature

Daily

| Weekly | (Correct) |

| Monthly |

| Annually |

Explanation

Snowflake is committed to providing a seamless, always up-to-date experience for our users while also delivering ever-increasing value through rapid development and continual innovation.

To meet this commitment, we deploy new releases and patch releases each week. This allows us to regularly deliver service improvements in the form of new features, enhancements, and fixes. The deployments happen transparently in the background; users experience no downtime or disruption of service, and are always assured of running on the most-recent release with access to the latest features.

https://docs.snowflake.com/en/user-guide/intro-releases.html#snowflake-releases

Question 39: **Correct**

Snowflake can carry out transformations after loading files staged by partner software (ELT)

| TRUE | (Correct) |

| FALSE |

Explanation

Snowflake, the cloud data platform, offers secure data sharing that eliminates the need for data extraction or transformation between departments, geographies, or partners. For primary data source loading, Snowflake works with a range of data integration partners and allows users to choose either ETL or transform data after loading (ELT). Snowflake removes the worry from data integration and allows you to focus on results.

Question 40: **Correct**

From a snowflake perspective what is common about Fivetran, Informatica, Matillion, Segment, Stitch and Talend

They are all snowflake data integration partner **(Correct)**

They are all snowflake competitors

They are all programming interface patterns

Explanation
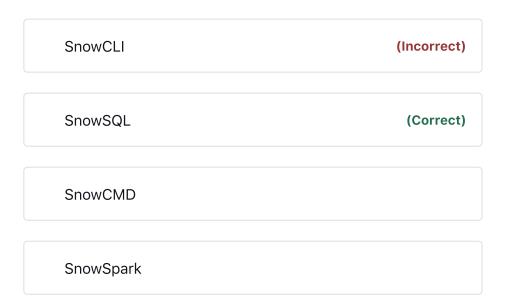These products are data integration partner of snowflake and can be used to load data into snowflake

Question 41: Correct

Looker is a snowflake business intelligence partner

TRUE   **(Correct)**

FALSE

Question 42: Incorrect

What is the name of the Snowflake's Command Line Interface tool?

SnowCLI   **(Incorrect)**

SnowSQL   **(Correct)**

SnowCMD

SnowSpark

Explanation
SnowSQL is the next-generation command line client for connecting to Snowflake to execute SQL queries and perform all DDL and DML operations, including loading data into and unloading data out of database tables.

SnowSQL (`snowsql` executable) can be run as an interactive shell or in batch mode through `stdin` or using the `-f` option.

https://docs.snowflake.com/en/user-guide/snowsql.html#snowsql-cli-client

Question 43: **Correct**

What is the best way to get the latest ODBC connector for use with Snowflake?
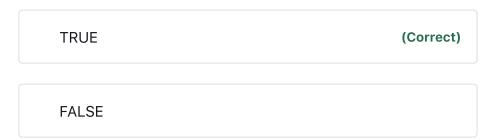
| Download it from snowflake web UI | **(Correct)** |

| Search google and download any version of ODBC |

| Compile in .NET |

Question 44: **Correct**

Snowflake has scalar and tabular user-defined functions (UDFs), with support for both SQL and JavaScript.

| TRUE | **(Correct)** |

| FALSE |

Explanation
Snowflake currently supports two types of UDFs, *SQL* and *JavaScript*:

A SQL UDF evaluates an arbitrary SQL expression and returns either scalar or tabular results.

A JavaScript UDF lets you use the JavaScript programming language to manipulate data and return either scalar or tabular results.

https://docs.snowflake.com/en/sql-reference/user-defined-functions.html#udfs-user-defined-functions

Question 45: **Correct**

This object in snowflake returns a set of rows instead of a single, scalar value, and can be accessed in the FROM clause of a query

UDTF	**(Correct)**

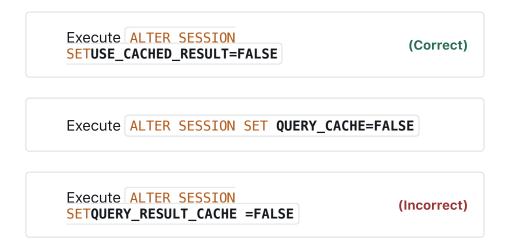| |
|---|
| UDF |

Stored procedure

Explanation
A UDTF returns a set of rows instead of a single, scalar value, and can be accessed in the FROM clause of a query. Snowflake supports both SQL and JavaScript UDTFs. This topic covers SQL UDTFs.

https://docs.snowflake.com/en/sql-reference/udf-table-functions.html#sql-udtfs-user-defined-table-functions

Question 46: **Incorrect**

By default, result reuse is enabled in snowflake. if you want to disable it what will you do?

Execute `ALTER SESSION` `SET`**USE_CACHED_RESULT=FALSE**	**(Correct)**

Execute `ALTER SESSION SET` **QUERY_CACHE=FALSE**

Execute `ALTER SESSION` `SET`**QUERY_RESULT_CACHE =FALSE**	**(Incorrect)**

Explanation
You will need to change it with the alter session query. This may not come in certification. But you will probably using it a lot when you actually work on a snowflake project. You will use it to test the performance once you have optimized your query. For example if a table is not partitioned well and you want to check the query performance after properly partitioning the table(either through clustering or reloading the table using natural order)

Question 47: **Correct**

Each time a persisted result for a query is reused, Snowflake resets the 24-hour retention period for the result, up to a maximum of 31 days from the date and time that the query was first executed. After 31 days, the result is purged and the next time the query is submitted, a new result is generated and persisted.

TRUE	**(Correct)**

FALSE	

Explanation
This is a very important point to remember. The query result of an already run query is persisted for 24 hours and every time the query is run the 24 hour window is reset for another 24 hours till 31 days

Question 48: **Correct**

Snowpipe guarantees that files are loaded in the same order they are staged

TRUE	

FALSE	**(Correct)**

Explanation
For each pipe object, Snowflake establishes a single queue to sequence data files awaiting loading. As new data files are discovered in a stage, Snowpipe appends them to the queue. However, multiple processes pull files from the queue; and so, while Snowpipe generally loads older files first, there is no guarantee that files are loaded in the same order they are staged.

Question 49: **Correct**

Snowpipe is recommended to load data in which of the below scenario?

You have small volume of frequent data	**(Correct)**

You have a huge volume of data generated as part of a

batch schedule

Explanation
Please not the key word 'SMALL VOLUME' and 'FREQUENT'. This question can be framed in multiple ways.

Question 50: Incorrect

What should be your minimum snowflake edition for data sharing capability

Standard	**(Correct)**

Enterprise	**(Incorrect)**

Business critical

Explanation

Data Sharing 🔗

Feature/Service	Standard	Enterprise	Business Critical	VPS
As a data provider, securely share data with other accounts.	✔	✔	✔	✔
As a data consumer, query data shared with your account by data providers.	✔	✔	✔	✔
Secure data sharing across regions and cloud platforms (through data replication)	✔	✔	✔	✔
Snowflake Data Marketplace, where providers and consumers meet to securely sharing data.	✔	✔	✔	✔
Data Exchange, a private hub of administrators, providers, and consumers that you invite to securely collaborate around data.	✔	✔	✔	✔

Question 51: Correct

Select the snowflake edition that allow only a maximum of 1 day of time travel

Standard	**(Correct)**

Enterprise

VPS

Business Critical

Explanation

Standard Time Travel (up to 1 day) for accessing/restoring modified and deleted data is available for all editions. Extended Time Travel (up to 90 days) is available for all editions except Standard

Question 52: **Correct**

Files that are already copied from the stage to the source table can be loaded again into a table cloned from the source table

TRUE	**(Correct)**

FALSE

Explanation

You do not have to mug up the answer.

Just remember what prevents the loaded files to be reloaded again. It is the load metadata. Snowflake maintains detailed metadata for each table into which data is loaded. This data expires after 64 days. Through this data, Snowflake knows that the file has been already loaded to the table. When you clone the table, this metadata is not cloned and hence you will be able to load the data again to the cloned table.

Refresh your memory on load metadata

Load Metadata

Snowflake maintains detailed metadata for each table into which data is loaded, including:

Name of each file from which data was loaded

1. File size

2. ETag for the file

3. Number of rows parsed in the file

4. Timestamp of the last load for the file

5. Information about any errors encountered in the file during loading

This load metadata expires after 64 days. If the LAST_MODIFIED date for a staged data file is less than or equal to 64 days, the COPY command can determine its load status for a given table and prevent reloading (and data duplication). The LAST_MODIFIED date is the timestamp when the file was initially staged or when it was last modified, whichever is later.

If the LAST_MODIFIED date is older than 64 days, the load status is still known if **either** of the following events occurred less than or equal to 64 days prior to the current date:

The file was loaded successfully.

The initial set of data for the table (i.e. the first batch after the table was created) was loaded.

However, the COPY command cannot definitively determine whether a file has been loaded already if the LAST_MODIFIED date is older than 64 days **and** the initial set of data was loaded into the table more than 64 days earlier (**and** if the file was loaded into the table, that also occurred more than 64 days earlier). In this case, to prevent accidental reload, the command skips the file by default.

Question 53: Incorrect

What is the technique called which snowflake uses to limit the number of micro-partitions retrieved as part of a query?

Pruning	**(Correct)**

Selective Filter	

Indexing	

Clustering	**(Incorrect)**

Explanation
Query Pruning

The micro-partition metadata maintained by Snowflake enables precise pruning of columns in micro-partitions at query run-time, including columns containing semi-structured data. In other words, a query that specifies a filter predicate on a range of values that accesses 10% of the values in the range should ideally only scan 10% of the micro-partitions.

5/29/2021   Snowflake Certification Preparation | Udemy

For example, assume a large table contains one year of historical data with date and hour columns. Assuming uniform distribution of the data, a query targeting a particular hour would ideally scan 1/8760th of the micro-partitions in the table and then only scan the portion of the micro-partitions that contain the data for the hour column; Snowflake uses columnar scanning of partitions so that an entire partition is not scanned if a query only filters by one column.

In other words, the closer the ratio of scanned micro-partitions and columnar data is to the ratio of actual data selected, the more efficient is the pruning performed on the table.

For time-series data, this level of pruning enables potentially sub-second response times for queries within ranges (i.e. "slices") as fine-grained as one hour or even less.

Not all predicate expressions can be used to prune. For example, Snowflake does not prune micro-partitions based on a predicate with a subquery, even if the subquery results in a constant.

Question 54: Incorrect

Select two choices that are true about snowflake roles

Snowflake users has a limit on the number of roles that they can assume	
Snowflake user can have one or more roles	**(Correct)**
Privileges can be directly assigned to a user	**(Incorrect)**
For a particular session, only one role can be active at a given time	**(Correct)**

Question 55: Correct

Snowflake waits till all the servers are provisioned for a new virtual warehouse, before it executes query on that warehouse

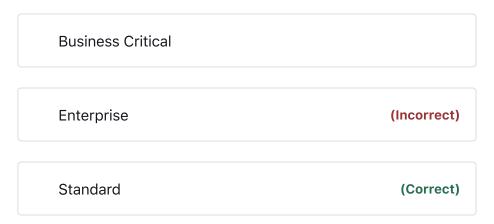|	
TRUE	**(Correct)**
FALSE	

Explanation
Snowflake does not begin executing SQL statements submitted to a warehouse until all of the servers for the warehouse are successfully provisioned, unless any of the servers fail to provision:

1. If any of the servers for the warehouse fail to provision during start-up, Snowflake attempts to repair the failed server(s).

2. During the repair process, the warehouse starts processing SQL statements once 50% or more of the requested servers are successfully provisioned.

Question 56: **Incorrect**

If you want customer dedicated virtual warehouse, which is the lowest snowflake edition that you should opt for
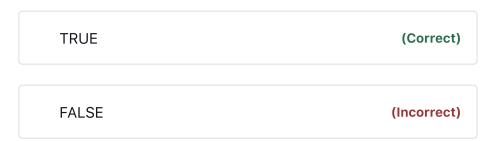
Business Critical	

Enterprise	**(Incorrect)**

Standard	**(Correct)**

Explanation
Customer dedicated virtual warehouse is available in all editions

Question 57: **Incorrect**

You need to contact Snowflake Support team if retrieval of data is required from fail safe

TRUE	**(Correct)**

FALSE	**(Incorrect)**

Explanation
There is nothing to explain here, you need to remember this.

Question 58: **Incorrect**

After how many days do the load metadata of a table expire?

64	**(Correct)**

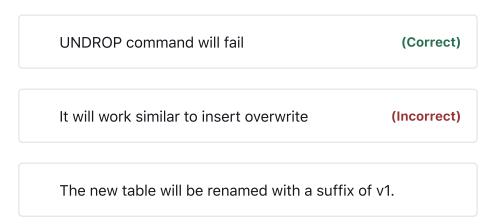|	
14	**(Incorrect)**

7

365

Explanation

This load metadata `expires after 64 days.` If the LAST_MODIFIED date for a staged data file is less than or equal to 64 days, the COPY command can determine its load status for a given table and prevent reloading (and data duplication). The LAST_MODIFIED date is the timestamp when the file was initially staged or when it was last modified, whichever is later.

Question 59: **Incorrect**

Let us imagine you have a permanent table named EMPLOYEE, you have dropped the table. Then you created another table with the same name. If you execute the UNDROP command now, what will happen?

UNDROP command will fail	**(Correct)**

It will work similar to insert overwrite	**(Incorrect)**

The new table will be renamed with a suffix of v1.

Explanation

Let us do a hand's on and learn it ourselves.

```
CREATE OR REPLACE TABLE EMPLOYEE(EMPLOYEE_NAME VARCHAR);

INSERT INTO EMPLOYEE VALUES ('AKSHAY'),('RUPAK'),('JEETU'); -- Please note the s

DROP TABLE EMPLOYEE;

CREATE OR REPLACE TABLE EMPLOYEE(EMPLOYEE_NAME VARCHAR);

UNDROP TABLE EMPLOYEE;
```

You will get the below error

```
SQL compilation error: Object 'EMPLOYEE' already exists.
```

Question 60: **Incorrect**

Which of the below are true with respect to snowflake web ui?

Each worksheet can have different role, warehouse, schema and database	**(Correct)**

Each worksheet is a separate session	**(Correct)**

You will be able to run a 'PUT' command from worksheet	**(Incorrect)**

You must use the same role and warehouse for all worksheets	

Explanation
These are the easy set of questions in the exam. You should not miss it. Please log on to the WEB UI and validate each statement. Try to run the below. Can you run it

```
CREATE STAGE CSV_STAGE;
PUT 'D:/MYFILE.CVS' @CSV_STAG;
```

This will not run and you will get below error

```
SQL compilation error: The command is not supported from the UI: PUT
```

Question 61: **Correct**

When a network policy includes values in both the allowed and blocked IP address lists, Snowflake applies the **blocked** IP address list first.

TRUE	**(Correct)**

FALSE

Explanation
When a network policy includes values in both the allowed and blocked IP address lists, Snowflake applies the **blocked** IP address list first.

Do **not** add `0.0.0.0/0` to the blocked IP address list. `0.0.0.0/0` is interpreted to be "all IPv4 addresses on the local machine". Because Snowflake resolves this list first, this would block your own access. Also, note that it is not necessary to include this IP address in the allowed IP address list.

Question 62: **Incorrect**

Lets say you executed a transaction in a snowflake session. Due to some reason the session disconnects and the transaction now is in a detached state. The transaction cannot be committed or rolled back. The object on which the transaction was applied is also now locked. if you do not do anything and let snowflake eventually abort the transaction, how long will you need to wait?

60 minutes

15 minutes	**(Incorrect)**

4 hours	**(Correct)**

Explanation
If a transaction is running in a session and the session disconnects abruptly, preventing the transaction from committing or rolling back, the transaction is left in a detached state, including any locks that the transaction is holding on resources. If this happens, you might need to abort the transaction.

To abort a running transaction, the user who started the transaction or an account administrator can call the system function, SYSTEM$ABORT_TRANSACTION.

`If the transaction is left open. Snowflake typically rolls back the transaction after it has been idle for four hours.`

Question 63: **Incorrect**

Snowflake does not support nested transactions

TRUE	**(Correct)**

FALSE	**(Incorrect)**

Explanation
Overlapping Transactions

This section describes overlapping transactions.

A stored procedure that contains a transaction can be called from within another transaction. The outer transaction can be in an outer stored procedure or can be outside any stored procedure.

The inner transaction is *not* treated as nested; instead, the inner transaction is *a separate transaction*. Snowflake calls these "autonomous scoped transactions" (or simply "scoped transactions"), because each transaction executes in a conceptually *independent scope*.

Note

Terminology note:

The terms "inner" and "outer" are commonly used when describing nested operations, such as nested stored procedure calls. Although Snowflake supports nested procedure calls, `Snowflake does not support nested transactions`;

https://docs.snowflake.com/en/sql-reference/transactions.html#overlapping-transactions

Question 64: **Correct**

You have cloned a table. Which of the below queries will work on the cloned table?

DROP TABLE <TABLE_NAME>

SELECT * FROM <TABLE_NAME>

SHOW TABLES LIKE '<TABLE_NAME>'

ALL OF THE ABOVE   **(Correct)**

Explanation

Cloning creates a copy of an existing object in the system. This command is primarily used for creating zero-copy clones of databases, schemas, and non-temporary tables; however, it can also be used to quickly/easily create clones of other schema objects (i.e. external stages, file formats, and sequences).

The command is a variation of the object-specific CREATE <object> commands with the addition of the `CLONE` keyword.

When you clone a table, it is actually a regular table which shares the micro-partitions with the table from which it has been cloned

Question 65: **Correct**

Select the term that is associated with compute layer?

Query optimization

Query planning

Query processing   **(Correct)**

Explanation

Compute layer has virtual warehouses which are responsible for query processing
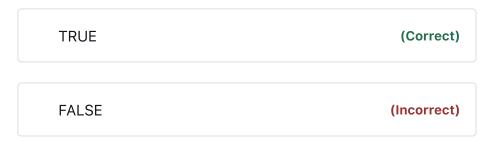
Question 66: **Correct**

You can use the query profiler view only for completed queries

TRUE

FALSE   **(Correct)**

Question 67: Incorrect

You need snowflake instances in different regions. You will need to have separate account for each region

| TRUE | (Correct) |

| FALSE | (Incorrect) |

Question 68: Correct

In a multi-cluster warehouse, you will need to manually scale the warehouses(minimum cluster to maximum cluster)

| TRUE |

| FALSE | (Correct) |

Explanation

Multi-cluster warehouses enable you to scale compute resources to manage your user and query concurrency needs as they change, such as during peak and off hours. And it is 100% automated.

Question 69: Correct

Which of the below statement will you use to recreate a specified object?

| GET_DDL | (Correct) |

| GET_SCHEMA |

| SHOW SCHEMA |

Explanation
GET_DDL

Returns a DDL statement that can be used to recreate the specified object. For databases and schemas, GET_DDL is recursive, i.e. it returns the DDL statements for recreating all supported objects within the specified database/schema.

GET_DDL currently supports the following object types:

Databases (see CREATE DATABASE)

Schemas (see CREATE SCHEMA)

Tables (see CREATE TABLE) (not external tables)

Views (see CREATE VIEW)

Streams (see CREATE STREAM)

Tasks (see CREATE TASK)

Sequences (see CREATE SEQUENCE)

File formats (see CREATE FILE FORMAT)

Pipes (see CREATE PIPE)

UDFs (see CREATE FUNCTION)

Stored Procedures (see CREATE PROCEDURE)

Try it out

```
CREATE OR REPLACE TABLE EMPLOYEE(EMPLOYEE_NAME VARCHAR);
select get_ddl('TABLE', 'EMPLOYEE', true);
```

What do you get?

Question 70: **Correct**

This command can be used to list streams for the current/specified database or schema, or across your entire account.

SHOW STREAMS	**(Correct)**

DISPLAY STREAMS

LIST STREAMS

Explanation
SHOW STREAMS

Lists the streams for which you have access privileges. The command can be used to list streams for the current/specified database or schema, or across your entire account.

The output returns stream metadata and properties, ordered lexicographically by database, schema, and stream name (see Output in this topic for descriptions of the output columns). This is important to note if you wish to filter the results using the provided filters.

https://docs.snowflake.com/en/sql-reference/sql/show-streams.html#show-streams

Question 71: **Incorrect**

When you use this parameter with SHOW STREAMS, The output also includes an additional `dropped_on` column, which displays

1. Date and timestamp (for dropped streams).

2. `NULL` (for active streams).

HISTORY	**(Correct)**

TERSE

DROPPED=TRUE	**(Incorrect)**

Explanation
Parameters

`TERSE`

Returns only a subset of the output columns:

`created_on`

`name`

`kind` (rename of `type` column in full set of columns)

`database_name`

`schema_name`

`tableOn` (rename of `table_name` column in full set of columns)

`HISTORY`

Optionally includes dropped streams that have not yet been purged (i.e. they are still within their respective Time Travel retention periods). If multiple versions of a dropped stream exist, the output displays a row for each version. The output also includes an additional `dropped_on` column, which displays:

Date and timestamp (for dropped streams).

`NULL` (for active streams).

Default: No value (dropped streams are *not* included in the output)

https://docs.snowflake.com/en/sql-reference/sql/show-streams.html#parameters

Question 72: **Incorrect**

The table functions in INFORMATION_SCHEMA can be used to return account-level usage and historical information for storage, warehouses, user logins, and queries

TRUE	**(Correct)**

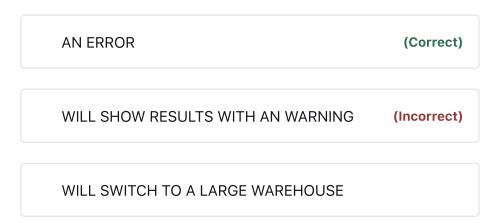| FALSE	**(Incorrect)**

Explanation
List of Table Functions

The table functions in INFORMATION_SCHEMA can be used to return account-level usage and historical information for storage, warehouses, user logins, and queries:

https://docs.snowflake.com/en/sql-reference/info-schema.html#list-of-table-functions

Question 73: **Incorrect**

What can you expect if the filters specified in an INFORMATION_SCHEMA query are not sufficiently selective?

AN ERROR	**(Correct)**

WILL SHOW RESULTS WITH AN WARNING	**(Incorrect)**

WILL SWITCH TO A LARGE WAREHOUSE	

Explanation
To prevent performance issues, the following error is returned if the filters specified in an INFORMATION_SCHEMA query are not sufficiently selective:

`Information schema query returned too much data. Please repeat query with more selective predicates.`

https://docs.snowflake.com/en/sql-reference/info-schema.html#general-usage-notes

Question 74: **Incorrect**

Information Schema Views will require a warehouse to execute whereas SHOW command does not require one.

| TRUE | **(Correct)** |

| FALSE | **(Incorrect)** |

Explanation
The INFORMATION_SCHEMA views provides a SQL interface to the same information provided by the SHOW <objects>commands. You can use the views to replace these commands; however, there are some key differences to consider before switching:

Considerations	SHOW Commands	Information Schema Views
Warehouses	Not required to execute.	Warehouse must be running and currently in use to query the views.
Pattern matching/filtering	Case-insensitive (when filtering using LIKE).	Standard (case-sensitive) SQL semantics. Snowflake automatically converts unquoted, case-insensitive identifiers to uppercase internally, so unquoted object names must be queried in uppercase in the Information Schema views.
Query results	Most SHOW commands limit results to the current schema by default.	Views display all objects in the current/specified database. To query against a particular schema, you must use a filter predicate (e.g. `... WHERE table_schema = CURRENT_SCHEMA()...`). Note that Information Schema queries lacking sufficiently selective filters return an error and do not execute (see General Usage Notes in this topic).

Question 75: **Incorrect**

When unloading data into multiple files, you will use this copy option to specify the maximum size of each file created

| MAX_FILE_SIZE | **(Correct)** |

| MAX_SIZE | |

MAX_FILE_BYTES	**(Incorrect)**

Explanation
Bulk Unloading into Single or Multiple Files

The COPY INTO <location> command provides a copy option (SINGLE) for unloading data into a single file or multiple files. The default is SINGLE = FALSE (i.e. unload into multiple files).

Snowflake assigns each file a unique name. The location path specified for the command can contain a filename prefix that is assigned to all the data files generated. If a prefix is not specified, Snowflake prefixes the generated filenames with `data_`.

Snowflake appends a suffix that ensures each file name is unique across parallel execution threads; e.g. `data_stats_0_1_0`.

When unloading data into multiple files, use the `MAX_FILE_SIZE` copy option to specify the maximum size of each file created.

Question 76: **Correct**

How do you truncate a date (from a timestamp) down to the year, month, and day

Use concatenation of the date part

Split the date into year, month and day

Use date_trunc and execute a query as below
```
select to_date('2015-05-08T23:39:20.123-07:
       date_trunc('YEAR', "DATE1") as "TRUN
       date_trunc('MONTH', "DATE1") as "TRU
       date_trunc('DAY', "DATE1") as "TRUNC
```
(Correct)

Explanation
You do not need to remember the query(just remember the function **date_trunc**) and it will most probably not come in your certification, but this will be an useful function when you actually start working on snowflake

https://docs.snowflake.com/en/sql-reference/functions/date_trunc.html#examples

Question 77: **Correct**

You are trying to set a variable by suing the set variable function `SET MY_VARIABLE ='XXXX'` . You got an error as below

`Assignment to 'MY VARIABLE' not done because value exceeds size limit of variables. Its size is 312; the limit is 256`

What is the reason for this error?

The size of string or binary variables is limited to 256 bytes	**(Correct)**

This is a temporary error due to insufficient memory in VM instance

Your warehouse is not big enough to accomodate this

Explanation

Variables can be set by executing the SQL statement `SET` or by setting the variables in the connection string when you connect to Snowflake.

The size of string or binary variables is limited to 256 bytes.

https://docs.snowflake.com/en/sql-reference/session-variables.html#initializing-variables

Question 78: **Incorrect**

Snowflake supports specifying a SELECT statement instead of a table in the COPY INTO <location> command

TRUE	**(Correct)**

FALSE	**(Incorrect)**

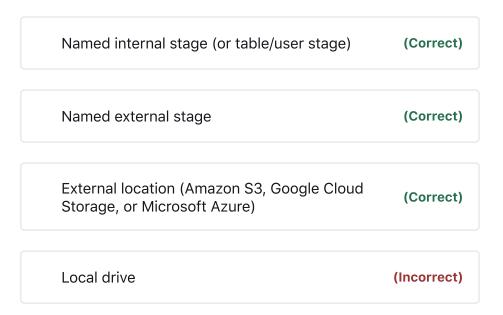**Explanation**
Bulk Unloading Using Queries

`Snowflake supports specifying a SELECT statement instead of a table in the` `COPY INTO <location>` `command`. The results of the query are written to one or more files as specified in the command and the file(s) are stored in the specified location (internal or external).

SELECT queries in COPY statements support the full syntax and semantics of Snowflake SQL queries, including JOIN clauses, which enables downloading data from multiple tables.

https://docs.snowflake.com/en/user-guide/data-unload-overview.html#bulk-unloading-using-queries

Question 79: **Incorrect**

Using **COPY INTO *<location>*** command, you can unload data from a table (or query) into the below locations.

Named internal stage (or table/user stage)	**(Correct)**

Named external stage	**(Correct)**

External location (Amazon S3, Google Cloud Storage, or Microsoft Azure)	**(Correct)**

Local drive	**(Incorrect)**

Explanation
COPY INTO *<location>*

Unloads data from a table (or query) into one or more files in one of the following locations:

1. Named internal stage (or table/user stage). The files can then be downloaded from the stage/location using the **GETcommand**.

2. Named external stage that references an external location (Amazon S3, Google Cloud Storage, or Microsoft Azure).

3. External location (Amazon S3, Google Cloud Storage, or Microsoft Azure).

Question 80: **Correct**

To download files from the stage/location loaded through COPY INTO <LOCATION> command, you will use

GET	(Correct)

UNLOAD

COPY INTO

Explanation
GET

Downloads data files from one of the following Snowflake stages to a local directory/folder on a client machine:

Named internal stage.

Internal stage for a specified table.

Internal stage for the current user.

Typically, this command is executed after using the COPY INTO <location> command to unload data from a table into a Snowflake stage.

https://docs.snowflake.com/en/sql-reference/sql/get.html#get

Question 81: **Incorrect**

GET does *not* support downloading files from external stages

FALSE	(Incorrect)

TRUE	(Correct)

Explanation
`GET does `***not***` support downloading files from external stages.` To download files from external stages, use the utilities provided by the cloud service.

The following Snowflake clients do not support GET:

Go Snowflake Driver

.NET Driver

Node.js Driver

The ODBC driver supports GET with Snowflake accounts hosted on the following platforms:
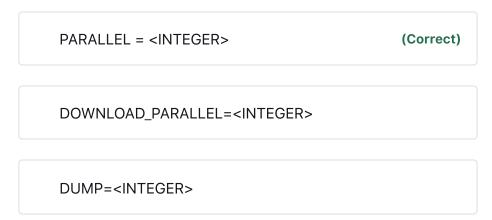
Amazon Web Services (using ODBC Driver Version 2.17.5 and higher).

Google Cloud Platform (using ODBC Driver Version 2.21.5 and higher).

Microsoft Azure (using ODBC Driver Version 2.20.2 and higher).

Question 82: **Correct**

You will use this parameter to specify the number of threads to use for downloading the files using GET command

PARALLEL = <INTEGER>   **(Correct)**

DOWNLOAD_PARALLEL=<INTEGER>

DUMP=<INTEGER>

Explanation
Optional Parameters

`PARALLEL = integer`

Specifies the number of threads to use for downloading the files. The granularity unit for downloading is one file.

Increasing the number of threads can improve performance when downloading large files.

Supported values: Any integer value from `1` (no parallelism) to `99` (use 99 threads for downloading files).

Default: `10`

`PATTERN = 'regex_pattern'`

Specifies a regular expression pattern for filtering files to download.

Default: No value (all files in the specified stage are downloaded)

https://docs.snowflake.com/en/sql-reference/sql/get.html#optional-parameters

Question 83: **Incorrect**

Following commands cannot be executed from worksheets

PUT   **(Correct)**

PUT   **(Correct)**

GET   **(Correct)**

SHOW

LIST <STAGE>   **(Incorrect)**

Question 84: Correct

If file format options are specified in multiple locations, the load operation applies the options in the following order of precedence.

1. COPY INTO TABLE statement.
2. Stage definition.   **(Correct)**
3. Table definition.

1. Stage definition.
2. COPY INTO TABLE statement.
3. Table definition.

1. COPY INTO TABLE statement.
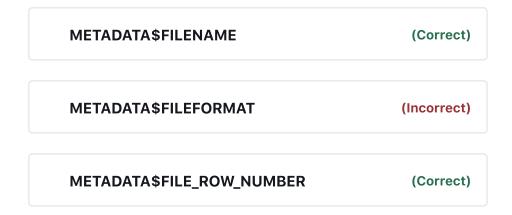2. Table definition.
3. Stage definition.

Explanation

If file format options are specified in multiple locations, the load operation applies the options in the following order of precedence:

1. COPY INTO TABLE statement.

2. Stage definition.

3. Table definition.

Question 85: Incorrect

Which are the two metadata columns for staged files

METADATA$FILENAME	(Correct)

METADATA$FILEFORMAT	(Incorrect)

METADATA$FILE_ROW_NUMBER	(Correct)

Explanation
Metadata Columns

Currently, the following metadata columns can be queried or copied into tables:

METADATA$FILENAME

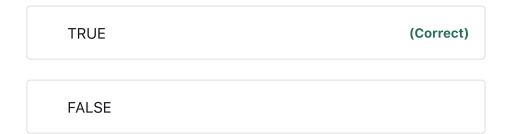Name of the staged data file the current row belongs to. Includes the path to the data file in the stage.

METADATA$FILE_ROW_NUMBER

Row number for each record in the container staged data file.

https://docs.snowflake.com/en/user-guide/querying-metadata.html#metadata-columns

Question 86: **Correct**

Metadata columns for staged files can only be queried by name

TRUE	(Correct)

FALSE	

Explanation
Query Limitations

Metadata cannot be inserted into existing table rows.

Metadata columns can only be queried by name; as such, they are not included in the output of any of the following statements:

SELECT *

SHOW <objects>

DESCRIBE <object>

Queries on INFORMATION_SCHEMA views

https://docs.snowflake.com/en/user-guide/querying-metadata.html#query-limitations

Question 87: Incorrect

Only named stages (internal or external) and user stages are supported for COPY transformations

TRUE	(Correct)

FALSE	(Incorrect)

Explanation
Usage Notes

This section provides usage information for transforming staged data files during a load.

Supported Stages

```
Only named stages (internal or external) and user stages are supported for COPY
```

Supported File Formats

The following file format types are supported for COPY transformations:
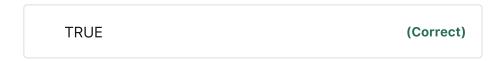
CSV

JSON

Avro

ORC

Parquet

XML

Question 88: Incorrect

The VALIDATION_MODE parameter does not support COPY statements that transform data during a load.

FALSE	(Incorrect)

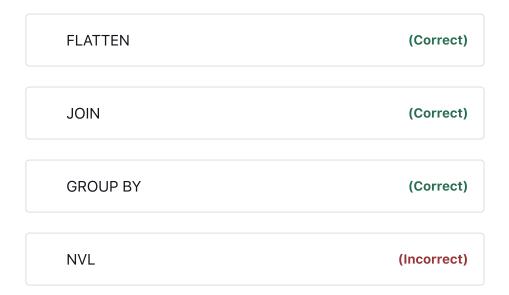| TRUE	(Correct)

Explanation
VALIDATION_MODE Parameter

The VALIDATION_MODE parameter does not support COPY statements that transform data during a load.

https://docs.snowflake.com/en/user-guide/data-load-transform.html#validation-mode-parameter

Question 89: **Incorrect**

Following transformations are not supported in a COPY command. Select three.

FLATTEN	(Correct)

JOIN	(Correct)

GROUP BY	(Correct)

NVL	(Incorrect)

Explanation
Note that COPY transformations do ***not*** support the FLATTEN function, or JOIN or GROUP BY (aggregate) syntax.

This is a very important topic, please read it

https://docs.snowflake.com/en/user-guide/data-load-transform.html#transforming-data-during-a-load

Question 90: **Correct**

Snowflake recommends to compress your data files when you are loading large data sets.

TRUE	(Correct)

FALSE

Explanation

` We recommend that you compress your data files when you are loading large data sets. ` See CREATE FILE FORMAT for the compression algorithms supported for each data type.

When loading compressed data, specify the compression method for your data files. The COMPRESSION file format option describes how your data files are *already* compressed in the stage. Set the COMPRESSION option in one of the following ways:

1. As a file format option specified directly in the COPY INTO <table> statement.

2. As a file format option specified for a named file format or stage object. The named file format/stage object can then be referenced in the COPY INTO *<table>* statement.

Question 91: Correct

Snowpipe is a serverless function

TRUE	**(Correct)**

FALSE

Explanation
Automated data loads leverage event notifications for cloud storage to inform Snowpipe of the arrival of new data files to load. Snowpipe copies the files into a queue, from which they are loaded into the ` target table in a continuous, serverless fashion ` based on parameters defined in a specified pipe object.

Question 92: Incorrect

When calling the rest endpoints in snowpipe, below authentication methods are supported

key pair authentication with JSON Web Token	**(Correct)**

user id and password

Both	(Incorrect)

Explanation
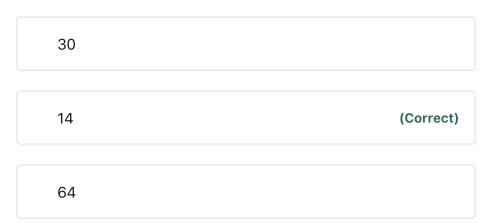Authentication

Bulk data load

Relies on the security options supported by the client for authenticating and initiating a user session.

Snowpipe

When calling the REST endpoints: Requires key pair authentication with JSON Web Token (JWT). JWTs are signed using a public/private key pair with RSA encryption.

Question 93: **Correct**

Load history of Snowpipe expires after how many days.

30	

14	(Correct)

64	

Explanation
Load History

Bulk data load

Stored in the metadata of the target table for 64 days. Available upon completion of the COPY statement as the statement output.

Snowpipe

Stored in the metadata of the pipe for 14 days. Must be requested from Snowflake via a REST endpoint, SQL table function, or ACCOUNT_USAGE view.

Question 94: **Incorrect**

SnowPipe can load a file with same name if it has been modified later.

TRUE	**(Incorrect)**

FALSE	**(Correct)**

Explanation
Data Duplication
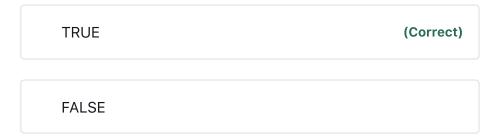
Snowpipe uses file loading metadata associated with each ***pipe object*** to prevent reloading the same files (and duplicating data) in a table. This metadata stores the path (i.e. prefix) and name of each loaded file, and prevents loading files with the same name even if they were later modified (i.e. have a different eTag).

Question 95: **Correct**

SNOWPIPE REST APIs support both internal and external stage

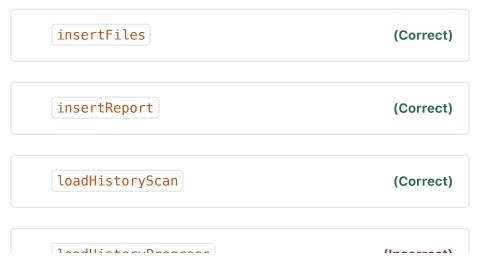| TRUE	**(Correct)**

FALSE	

Explanation
Snowpipe supports both internal (Snowflake) stages and external stages, i.e. S3 buckets. Please note that this is true for REST APIs. SNOWPIPE Auto ingest supports only external stage

Question 96: **Incorrect**

Which of the below are SNOWPIPE REST APIs. Select three.

`insertFiles`	**(Correct)**

`insertReport`	**(Correct)**

`loadHistoryScan`	**(Correct)**

`loadHistoryProgress`	**(Incorrect)**

loadHistoryProgress   (Incorrect)

Explanation
SNOWPIPE has three APIs

Data File Ingestion

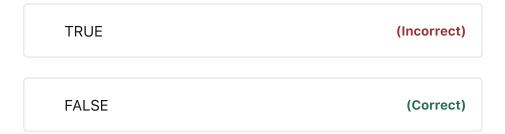1. Endpoint: `insertFiles`

Load History Reports

1. Endpoint: `insertReport`

2. Endpoint: `loadHistoryScan`

https://docs.snowflake.com/en/user-guide/data-load-snowpipe-rest-apis.html#snowpipe-rest-api

Question 97: Incorrect

A successful response from the SNOWPIPE insertFiles API means that the files are ingested.

TRUE	(Incorrect)

FALSE	(Correct)

Explanation
Endpoint: `insertFiles`

Informs Snowflake about the files to be ingested into a table. A successful response from this endpoint means that Snowflake has recorded the list of files to add to the table. `It does not necessarily mean the files have been ingested`

Question 98: Incorrect

The insertReport SNOWPIPE API can retrieve file ingestion events and report it. The events are retained for infinite time until deleted manually.

TRUE	(Incorrect)

FALSE	(Correct)

Explanation

The SNOWPIPE REST API topic is relevant for both SNOWPRO CORE and SNOWPRO ARCHITECT certifications. Please read this chapter thoroughly

https://docs.snowflake.com/en/user-guide/data-load-snowpipe-rest-apis.html#snowpipe-rest-api

Endpoint: `insertReport`

Retrieves a report of files submitted via `insertFiles` whose contents were recently ingested into a table. Note that for large files, this may only be part of the file.

Note the following limitations for this endpoint:

1. The 10,000 most recent events are retained.

2. Events are retained for a maximum of 10 minutes.

An event occurs when data from a file submitted via `insertFiles` has been committed to the table and is available to queries. The `insertReport` endpoint can be thought of like the UNIX command tail. By calling this command repeatedly, it is possible to see the full history of events on a pipe over time. Note that the command must be called often enough to not miss events. How often depends on the rate files are sent to `insertFiles`.

Question 99: **Incorrect**

To help avoid exceeding the rate limit (error code 429), snowflake recommends relying more heavily on `insertReport` than `loadHistoryScan`

TRUE	**(Correct)**

FALSE	**(Incorrect)**

Explanation
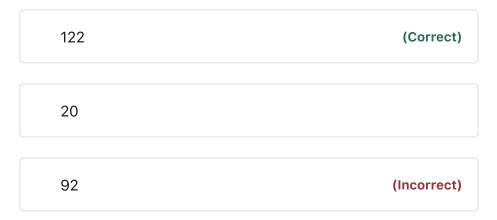Endpoint: `loadHistoryScan`

Fetches a report about ingested files whose contents have been added to table. Note that for large files, this may only be part of the file. This endpoint differs from `insertReport` in that it views the history between two points in time. There is a maximum of 10,000 items returned, but multiple calls can be issued to cover the desired time range.

Important

This endpoint is rate limited to avoid excessive calls. To help avoid exceeding the rate limit (error code 429), we recommend relying more heavily on `insertReport` than `loadHistoryScan`. When calling `loadHistoryScan`, specify the most narrow time range that includes a set of data loads. For example, reading the last 10 minutes of history every 8 minutes would work well. Trying to read the last 24 hours of history every minute will result in 429 errors indicating a rate limit has been reached. The rate limits are designed to allow each history record to be read a handful of times.

Question 100: **Incorrect**

You have a warehouse. It ran for 62 seconds and then was suspended. After that it resumed and ran for 20 seconds. For how many seconds will you be billed

122	**(Correct)**

20	

92	**(Incorrect)**

Explanation
When a virtual warehouse starts for the first time, the first one minute is anyway charged after that the billing is per second

Warehouses are only billed for credit usage when they are running. When a warehouse is suspended, it does not accrue any credit usage.

The credit numbers shown here are for a full hour of usage; however, credits are billed per-second, with a 60-second (i.e. 1-minute) minimum:

Each time a warehouse is started or resized to a larger size, the warehouse is billed for 1 minute's worth of usage based on the hourly rate shown above.

After 1 minute, all subsequent billing is per-second.

Stopping and restarting a warehouse within the first minute does not change the amount billed; the minimum billing charge is 1 minute.

When a warehouse is increased in size, credits are billed only for the *additional* servers that are provisioned. For example, changing from Small (2) to Medium (4) results in billing charges for 1 minute's worth of 2 credits.

https://docs.snowflake.com/en/user-guide/credits.html#virtual-warehouse-credit-usage