



SNOWPRO® ADVANCED: DATA ANALYST EXAM STUDY GUIDE

Last Updated: October 13, 2025

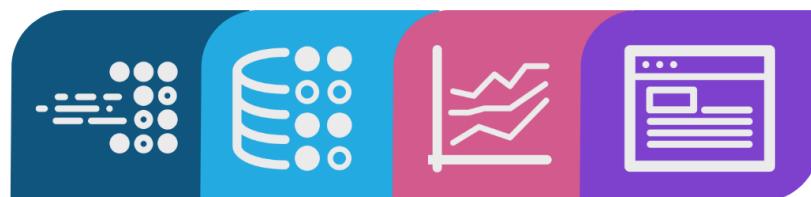


TABLE OF CONTENTS

SNOWPRO ADVANCED: DATA ANALYST STUDY GUIDE OVERVIEW	2
RECOMMENDATIONS FOR USING THE GUIDE	2
SNOWPRO ADVANCED: DATA ANALYST CERTIFICATION OVERVIEW	3
SNOWPRO ADVANCED: DATA ANALYST PREREQUISITE	4
STEPS TO SUCCESS	4
SNOWPRO ADVANCED: DATA ANALYST SUBJECT AREA BREAKDOWN	4
Domain 1.0: Data Ingestion and Data Preparation	5
Domain 1.0 Study Resources	6
Domain 2.0: Data Transformation and Data Modeling	7
Domain 2.0 Study Resources	8
Domain 3.0: Data Analysis	9
Domain 3.0 Study Resources	9
Domain 4.0: Data Presentation and Data Visualization	10
Domain 4.0 Study Resources	10
SNOWPRO ADVANCED: DATA ANALYST SAMPLE QUESTIONS	11
NEXT STEPS	15
REGISTERING FOR YOUR EXAM	15
MAINTAINING YOUR CERTIFICATION	15

SNOWPRO ADVANCED: DATA ANALYST STUDY GUIDE OVERVIEW

This is a self-learning study guide that highlights concepts that may be covered on Snowflake's SnowPro Advanced: Data Analyst Certification exam.

This study guide does not guarantee certification success.

Holding the SnowPro Core certification in good standing is a prerequisite for taking the Advanced: Data Analyst certification.

For an overview and more information on the SnowPro Core Certification exam or SnowPro Advanced Certification series, please navigate [here](#).

RECOMMENDATIONS FOR USING THE GUIDE

This guide will show the Snowflake topics and subtopics covered on the exam. Following the topics will be additional resources consisting of videos, documentation, blogs, or exercises to help you understand the Data Analyst role on the Snowflake AI Data Cloud.

Estimated length of study guide: 10 – 13 hours

Some links may have more value than others, depending on your experience, the same amount of time should not be spent on each link. Some links may appear in more than one domain.

SNOWPRO ADVANCED: DATA ANALYST CERTIFICATION OVERVIEW

The **SnowPro Advanced: Data Analyst** exam tests advanced knowledge and skills to apply comprehensive data analysis principles using Snowflake and its components. The exam will assess skills through scenario-based questions and real-world examples.

This certification will test the ability to:

- Prepare and load data
- Perform data transformations for data analysis
- Build and troubleshoot advanced SQL queries in Snowflake
- Use Snowflake built-in functions and create User-Defined Functions (UDFs)
- Perform descriptive and diagnostic data analyses
- Perform predictive data analysis
- Prepare and present data to meet business requirements

Target Audience:

1+ year of Snowflake data cloud analytics experience, including practical, hands-on use of the Snowflake Data Cloud. In addition, successful candidates may have:

- Fluency with advanced SQL

Knowledge of an additional computer language is a plus but not a requirement.

This exam is designed for:

- Snowflake Data Analysts
- ELT Developers
- Business Intelligence Professionals
- Analytics Engineers

SNOWPRO ADVANCED: DATA ANALYST PREREQUISITE

Eligible individuals must hold an active SnowPro Core Certified credential. If you feel you need more guidance on the Snowflake fundamentals, please see the [SnowPro Core Exam Study Guide](#).

STEPS TO SUCCESS

1. Review the content in this Study Guide
2. [Attend Snowflake's Advanced Analytics Instructor-Led Training](#)
3. [Review the Essential Guide to Data Analytics Ebook](#)
4. [Attend Snowflake's BI Analytics Bootcamp Course](#)
5. [Attend Snowflake Webinars](#)
6. [Review Snowflake's On-Demand Accelerate Series](#)
7. [Attend Snowflake's Virtual Hands-on Labs Live or On-demand](#)
8. [Review Snowflake's Quickstart Tutorials in the Solutions Center](#)
9. Use the links provided in this study guide to review and study the Snowflake documentation
10. Practice with sample questions [here](#)
11. Get hands-on practical experience with relevant business requirements using Snowflake
12. [Schedule your exam](#)
13. Take your exam!

Additional Snowflake Assets to check out for Advanced: Data Analyst

[Snowflake for Dummies Guide Series Books](#)

SNOWPRO ADVANCED: DATA ANALYST SUBJECT AREA BREAKDOWN

This exam guide includes test domains, weightings, and objectives. It is not a comprehensive listing of all the content that will be presented on this examination. The table below lists the main content domains and their weightings.

Domain	Domain Weightings on Exams
1.0 Data Ingestion and Data Preparation	17%
2.0 Data Transformation and Data Modeling	23%
3.0 Data Analysis	32%
4.0 Data Presentation and Data Visualization	28%



► Domain 1.0: Data Ingestion and Data Preparation

1.1 Use a collection system to retrieve data.

- Retrieve data from a source
 - Structured (CSV)
 - Semi-structured (e.g., Parquet, Avro, ORC, JSON, or XML)
 - Unstructured
 - Synthetic Data Generation

1.2 Perform data discovery to identify what is needed from the available datasets.

- Query tables in Snowflake to assess:
 - Data elements including statistics maintained by Snowflake
 - The elements that are required for business goals (using BI reports or SQL analysis)
 - The level of data granularity required
- Evaluate which transformations are required:
 - Perform table joins and set operations (e.g., UNION, UNION ALL, INTERSECT, and MINUS)
 - Perform data filtering and/or transformation
 - ASOF JOINS
- Use commands to read metadata and/or to alter context (e.g., DESCRIBE, SHOW, USE)

1.3 Enrich data by identifying and accessing relevant data from the Snowflake Marketplace.

- Find external data sets that correlate with available data

- Use Secure Data Sharing to enrich existing data sets (e.g., Data from Snowflake Marketplace, The Internal Marketplace, Private Listings, and Listings)
- Create tables and views

1.4 Use best practice considerations relating to data integrity structures.

- Define primary keys for tables
- Perform table joins between parent/child tables
 - Implement constraints

1.5 Implement data processing solutions.

- Cleanse, conform, and enrich data
- Automate and implement data pipelines
 - Scheduling
- Respond to processing failures
 - Use logging and monitoring solutions
 - Auditing
 - Data lineage

1.6 Given a scenario, prepare data and load into Snowflake.

- Load files using Snowsight
- Load data from external/internal stages into a table
- Load different types of data
 - Tabular data/structured data
 - Semi-structured data
 - Unstructured data
- Perform general DML (INSERT, UPDATE, and DELETE)
- Identify and resolve data import errors
- Prepare external tables

1.7 Given a scenario, use Snowflake functions.

- Scalar functions
- Aggregate functions
- Window functions
- Table functions
- System functions
- Geospatial functions

- AI functions
- User-Defined Functions (UDFs)
- ML functions
 - Classification
 - Top Insights
 - Anomaly Detection

Domain 1.0 Study Resources

Additional Assets

[NULL handling in Snowflake \(article\)](#)

Snowflake Documentation Links

[Access History](#)

[Account Usage](#)

[ASOF Joins](#)

[Bulk Loading Using COPY](#)

[COPY INTO <table>](#)

[COPY_HISTORY](#)

[CREATE TABLE](#)

[DATEDIFF](#)

[Data Consumers](#)

[INFER_SCHEMA](#)

[Introduction to External Tables](#)

[Introduction to Unstructured Data Support](#)

[Lateral Join](#)

[Loading Data into Snowflake](#)

[Loading Using the Web Interface \(Limited\)](#)

[ML Functions](#)

[Modifying Constraints](#)

[Object Dependencies](#)

[Overview of Constraints](#)

[Overview of Data Loading](#)

[PARSE_JSON](#)

[PIVOT](#)

[Preparing Your Data Files](#)

[PUT](#)

[Querying Data Using Worksheets](#)

[Querying Metadata for Staged Files](#)

[QUALIFY](#)

[SAMPLE / TABLESAMPLE](#)

[Semi-structured Data Types](#)

[Synthetic Data Generation](#)

[TOP <n>](#)

[TO_TIMESTAMP / TO_TIMESTAMP_*](#)

[Transforming Data During a Load](#)

[User-Defined Functions](#)

[Working with Joins](#)

[Working with Subqueries](#)



Domain 2.0: Data Transformation and Data Modeling

2.1 Prepare different data types into a consumable format.

- CSV
- JSON (query and parse)
- Parquet
- XML

2.2 Given a dataset, clean the data.

- Identify and analyze data quality issues
- Handle erroneous and ambiguous data
 - Handle duplications
 - Handle nulls
- Convert data types
- Use clones as required by specific use-cases
- Use Data Metric Functions (DMFs)
- Use Time Travel and cloning features
- Use built-in functions for traversing, flattening, transforming, and nesting semi-structured data
- Use native data types

2.4 Use data modeling to manipulate the data to meet BI requirements.

- Select and implement an effective data model
- Identify when to use a data model and when to use a flattened data set
- Use different modeling techniques for the consumption layer (e.g., dimensional, Data Vault)

2.3 Given a dataset or scenario, work with and query the data.

- Aggregate and validate the data
- Apply analytic/window functions
- Perform pre-math calculations (e.g., randomization, ranking, grouping, min/max)
- Perform casting - change data types to ensure data can be presented consistently
- Enrich the data
 - Use cartesian joins, sub-queries, CTEs, and union queries
 - Work with hierarchical data
 - Use sampling, approximation, and estimation features

2.5 Optimize query performance.

- Understand how to view and analyze the query execution plan
- Troubleshoot query performance
 - Leverage partition pruning
 - Leverage clustering keys
- Leverage result, metadata, and virtual warehouse caching
- Use search optimization service and virtual warehouse features such as the query acceleration services

Domain 2.0 Study Resources

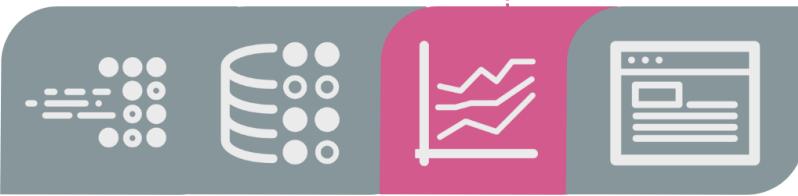
Additional Assets

[NULL handling in Snowflake](#) (article)
[Performance impact from local and remote disk spilling](#) (article)
[How to Analyze JSON with SQL | Snowflake](#) (PDF)

Snowflake Documentation Links

[Analyzing Queries Using Query Profile](#)
[COUNT](#)
[Data Metric Functions](#)
[Data Type Conversion](#)
[DATEDIFF](#)
[FLATTEN](#)
[LAG](#)
[Numeric Data Types](#)

[OBJECT_AGG](#)
[Query Acceleration Service](#)
[Querying Semi-structured Data](#)
[REGEXP_LIKE](#)
[SAMPLE / TABLESAMPLE](#)
[Search Optimization Service](#)
[SPLIT_TO_TABLE](#)
[TRIM](#)
[Understanding & Using Time Travel](#)
[Using Persisted Query Results](#)
[Warehouse Considerations](#)
[Window Functions](#)
[Working with Secure Views](#)
[Working with Temporary and Transient Tables](#)



► Domain 3.0: Data Analysis

3.1 Use SQL extensibility features.

- User-Defined Functions (UDFs)
- User-Defined Table Functions (UDTFs)
- Stored procedures
 - Asynchronous Stored Procedure
- Regular, secure, and materialized views

3.2 Perform descriptive analyses.

- Summarize large data sets using Snowsight dashboards
 - Create a reusable filter

- Perform exploratory ad-hoc analyses using Notebooks and worksheets to describe data

3.3 Perform diagnostic analyses.

- Find reasons/causes of anomalies or patterns in historical data
- Collect related data
- Identify demographics and relationships
- Analyze statistics and trends

3.4 Perform forecasting.

- Use statistics and built-in functions
- Make predictions based on data

Domain 3.0 Study Resources

Snowflake Documentation Links

[APPROX_COUNT_DISTINCT](#)

[AVG](#)

[Constraints](#)

[DENSE_RANK](#)

[Estimating the Number of Distinct Values](#)

[EXPLAIN](#)

[FLATTEN](#)

[Handling Exceptions](#)

[HLL](#)

[NTILE](#)

[Overview of Stored Procedures](#)

[Querying Data Using Worksheets](#)

[RANK](#)

[REGR_INTERCEPT](#)

[REGR_SLOPE](#)

[ROW_NUMBER](#)

[SHA2 , SHA2_HEX](#)

[Snowflake Notebooks](#)

[STDDEV_SAMP](#)

[Visualizing Worksheet Data](#)

[Working with Materialized Views](#)

[Working with Worksheets in Snowsight](#)



Domain 4.0: Data Presentation and Data Visualization

4.1 Given a use case, create reports and dashboards to meet business requirements.

- Evaluate and select the data for building dashboards
 - Set the contexts (e.g., database, schema, virtual warehouse, role)
 - Create and run SQL queries
 - Apply naming conventions to data columns and queries
 - Sort and filter data
- Understand the effects of row access policies and Dynamic Data Masking
- Compare and contrast different chart types (e.g., bar charts, scatter plots, heat grids, scorecards)
- Understand what is required to connect BI tools to Snowflake
- Create charts and dashboard in Snowsight
 - Create and manage custom filters

4.2 Given a use case, maintain reports and dashboards to meet business requirements.

- Build automated and repeatable tasks
- Operationalize data for consumption
- Manage and share Snowsight dashboards
- Configure subscriptions and updates

4.3 Given a use case, incorporate visualizations for dashboards and reports.

- Present data for business-use analyses
- Identify patterns and trends
- Identify correlations among variables
- Troubleshoot common issues with data analytics dashboard and reports
- Customize data presentations using filtering and editing techniques

Domain 4.0 Study Resources

Snowflake Documentation Links

[Account Usage](#)

[CORR](#)

[GENERATOR](#)

[Querying Data Using Worksheets](#)

[ROW_NUMBER](#)

[SEQ1 / SEQ2 / SEQ4 / SEQ8](#)

[Using Row Access Policies](#)

[Visualizing Data With Dashboards](#)

[Visualizing Worksheet Data](#)

SNOWPRO ADVANCED: DATA ANALYST SAMPLE QUESTIONS

1. What is the purpose of the Snowflake geospatial function `ST_TRANSFORM()` ?
 - A. It converts a `GEOMETRY` object from one spatial reference system to another.
 - B. It constructs a `GEOMETRY` object that represents a point with the specified longitude and latitude.
 - C. It returns the `GeoJSON` representation of that value, given a value of type `GEOGRAPHY` or `GEOMETRY`.
 - D. It returns a `GEOMETRY` object that has its Spatial Reference System Identifier (SRID) set to the specified value.
2. The `LOGS` table has a variant column (`V`) , and it contains JSON data with various fields:

ID	V
1	{ "post_age": "45", "post_date": "1975-01-10", "post_name": "Joe", "pre_date": "1975-01-11", "pre_name" : "Joen" }

A Data Analyst wants to parse the `v` column and return only the fields starting with `post` as JSON:

ID	CLEANED
1	{ "post_age": "45", "post_date": "1975-01-10", "post_name": "Joe" }

Which statement will produce the desired output?

- A.

```
SELECT ID, LISTAGG( v2.key || ':' || v2.value, ',') AS cleaned
FROM logs,
LATERAL FLATTEN(input => v) v2
WHERE v2.key LIKE 'post%'
GROUP BY ID;
```
- B.

```
SELECT ID, TO_JSON( v2.key, v2.value) AS cleaned
FROM logs,
LATERAL FLATTEN(input => v) v2
WHERE v2.key LIKE 'post%'
GROUP BY ID;
```

- C. `SELECT ID, ARRAY_AGG(v2.key || ':' || v2.value) AS cleaned
FROM logs,
LATERAL FLATTEN(input => v) v2
WHERE v2.key LIKE 'post%'
GROUP BY ID;`
- D. `SELECT ID, OBJECT_AGG(v2.key, v2.value) AS cleaned
FROM logs,
LATERAL FLATTEN(input => v) v2
WHERE v2.key LIKE 'post%'
GROUP BY ID;`

3. A Data Analyst has a table named `sales_performance` with two columns, `advertising_expense` (independent variable), and `sales_amount` (dependent variable). The Analyst wants to perform a linear regression analysis.

Which query will calculate the required coefficients for the linear regression?

- A.

```
SELECT
    REGR_INTERCEPT(advertising_expense,
sales_amount)
, REGR_SLOPE(advertising_expense, sales_amount)
FROM sales_performance;
```
- B.

```
SELECT
    REGR_VALY(sales_amount, advertising_expense)
, REGR_VALX(sales_amount, advertising_expense)
FROM sales_performance;
```
- C.

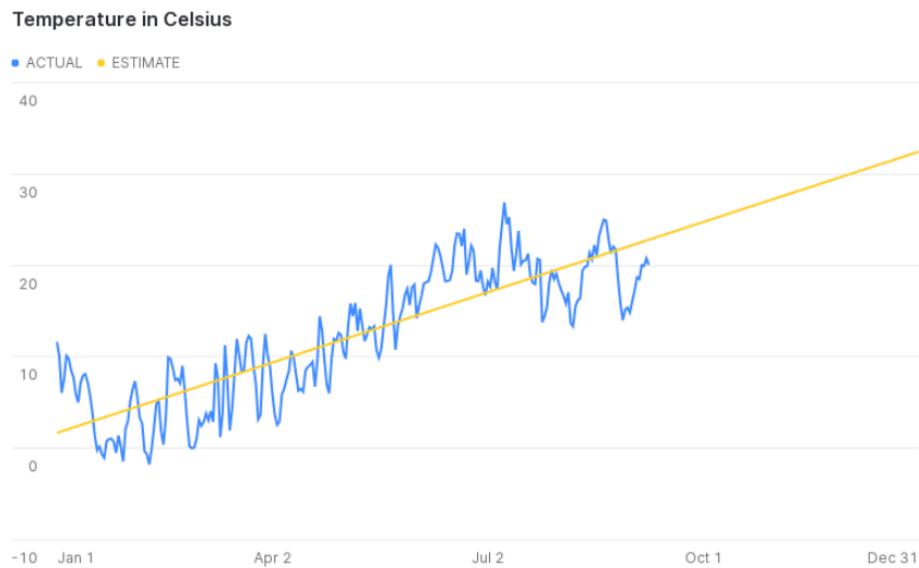
```
SELECT
    REGR_VALY(advertising_expense, sales_amount)
, REGR_VALX(advertising_expense, sales_amount)
FROM sales_performance;
```
- D.

```
SELECT
    REGR_INTERCEPT(sales_amount, advertising_expense)
, REGR_SLOPE(sales_amount, advertising_expense)
FROM sales_performance;
```

4. A Data Analyst created a dashboard using the DATA_ANALYST role. The Analyst shared the dashboard with a user who does not have the DATA_ANALYST role.

How can the user refresh the data on the dashboard if they have a role named REPORT_ENGINEER that has been given access to the underlying tables?

- A. Change the user's default role to REPORT_ENGINEER before opening the dashboard.
 - B. Change the user's role to REPORT_ENGINEER in each worksheet that is used in the dashboard.
 - C. Copy the dashboard and change the user's role from DATA_ANALYST to REPORT_ENGINEER.
 - D. Refresh the data in the dashboard from the custom shared link.
5. Data about temperature is presented on a Snowsight line chart:



What function(s) can be used to construct new data points based on actual data, to add a trendline to the chart based on the Least-Squares Regression (yellow line)?

- A. REGR_SLOPE and REGR_INTERCEPT
- B. COVAR_POP and COVAR_SAMP
- C. STDDEV
- D. CORR

Correct responses for sample questions:

1: a, 2: d, 3: d, 4: c, and 5: a

NEXT STEPS

REGISTERING FOR YOUR EXAM

When you are ready to register for the exam navigate [here](#) to get started. Select the exam you want to take and click “Register Now”. This will take you to our Certification Management system where you will register to take the exam.

MAINTAINING YOUR CERTIFICATION

All Snowflake Certifications expire two (2) years after your certification issue date.

SnowPro Certifications can now be recertified through the Snowflake Continuing Education (CE) program which includes these options -

- Completion of eligible Snowflake [Instructor Led \(ILT\) Training Courses](#)
- Earning of an equivalent or higher-level SnowPro Certification

Note: You must have a valid Certification to participate in the Continuing Education (CE) program.

The information provided in this study guide is provided for your purposes only and may not be provided to third parties.

IN ADDITION, THIS STUDY GUIDE IS PROVIDED “AS IS”. NEITHER SNOWFLAKE NOR ITS SUPPLIERS MAKES ANY OTHER WARRANTIES, EXPRESS OR IMPLIED, STATUTORY OR OTHERWISE, INCLUDING BUT NOT LIMITED TO WARRANTIES OF MERCHANTABILITY, TITLE, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT.