Minitab Model Ops®

Getting Started with Minitab Model Ops®

Minitab

Contents

1. Minitab Model Ops environment

With Minitab Model Ops, you deploy your models online so that you can get your results faster than ever. By making those models accessible through external endpoints, Minitab Model Ops streamlines the incorporation of model results into the applications that you use to make business decisions. Use those predictions to forecast your business with ease.

Minitab Model Ops has the following main sections.

Deployments 📶

> Create and manage model deployments, including Champion and Challenger models. For more information, go to Create a new deployment on page 9.

Model Repository ≣

> The model repository displays status details for all the imported and deployed models across the platform. For more information about adding and deleting models to the repository, go to Import models to the model repository on page 10.

API Keys 🗝

> Displays all API keys by status. From this list, you can sort by expiration date, model, or when the key was created. You can also revoke active keys from this list. For more information, go to API keys on page 4.

System Audit Log 🕑

> Record actions that affect the import and remove status of models. For more information, go to System audit log on page 5.

Note: The **System Audit Log** is not available for users with **Viewer** permissions.

Deployments list

The **Deployments** 📶 page displays a list of all the deployments within the subscription. You can sort this list by any of the columns to organize according to your preference. From the list, you can activate, pause, and delete deployments. For more information, go to Activate, pause, or delete deployments on page 9.

Note: Only paused deployments can be deleted from the deployments list.

Champion Model

Each deployment must have one and only one champion model. Each deployment can have up to two challenger models. For more information on using challenger models, go to Use champion/challenger models in a deployment on page 14.

Status

Each deployment can have one of the following statuses.

Icon	Status	Details
⚪	**Offline**	The deployment has been paused. Paused deployments cannot receive external requests.
🟡	**Pending**	The deployment is being activated or paused.
🔵	**Active**	The deployment is active. Active deployments can receive external requests.
🔴	**Failed**	The deployment has failed. Evaluate any errors in the system audit log and try again.

Drift

Indicates the overall drift status of each deployed champion model so you can quickly see which deployments may need attention.

Icon	Drift	Details
	No status	The model does not have baseline data or has been deployed for less than 50% of the production period.
✓	**Minimal**	The drift is not severe or moderate.
⚠	**Moderate**	The prediction drift is moderate or a minimum number of predictor variables have moderate drift.
⊖	**Severe**	The prediction drift is severe or a minimum number of predictor variables have severe drift.

Model repository

The **Model Repository** ☰ displays a list of all the models within the subscription. You can sort this list by any of the columns to organize according to your preference. From the repository, you can delete models and upload baseline data. For more information, go to Upload baseline data for a model on page **30**.

Note: Only models that are not in a deployment (Ready) can be deleted from the model repository.

Status

Each model can have one of the following statuses.

Icon	Status	Details
	Ready	The model is not in a deployment and is available for a new deployment.
○	**Inactive**	The model is in a deployment, but the deployment is not active.
●	**Deployed (Champion)**	The model is deployed as a champion model.
●	**Deployed**	The model is deployed as a challenger model.
●	**Upload Failed**	The model failed to upload and is unusable. Evaluate any errors in the system audit log and try again.
●	**Failed**	Failure occurred during deployment of the model. This status is not common.

Baseline Data

Indicates whether the baseline data is available or not.

Upload successful

The model has baseline data uploaded.

Not available

The model does not have baseline data uploaded.

API keys

The **API Keys** page lists all the API keys for the subscription.

The list of API keys includes the details for all keys created for the subscription. You can sort this list by any of the columns to organize according to your preference. Each key can be revoked prior to its expiration date.

To create a new key for a deployment, go the **Integration** tab. For more information, go to

System audit log

The audit log for the system records the import and remove actions of models and deployments. Use the following steps to view the audit log for the system.
1. Sign in to Minitab Model Ops.

2. Select **System Audit Log** ⟲.

Note: The audit log for each deployment is found on the **Audit Log** on page **16** tab for the deployment.

Audit log information

The log records the following items for each action. Expand an item to see additional details for the action.

Timestamp
> The date and time of the action.

User
> The name from the **User Information** for the account for Minitab Model Ops.

Action
> This value records one of the following actions.
> * ImportModel
> * UpdateModel
> * RemoveModel
> * CreateDeployment
> * RemoveDeployment
> * UploadDataset

Result
> This value records whether the action succeeds or fails.

Note: Audit logs entries older than 1 year are purged from the system.

Import model failures

The system audit log records the following failures when importing a model.

InvalidColumns
> The model import fails if you use a reserved variable name. For more information, go to Reserved variable names on page **23**.

InvalidSchemaText
> The model import fails if you use an invalid character in the response or class level. For example, the characters that represent a new line of '\n' and '\r' cannot be used.

TooManyNodes
> The model is too large. Reduce the number of trees and/or the node size for your model.

TooManyPredictors
> Models must have 500 predictors or less.

TooManyResponseClasses
> The maximum number of classes is 10 for a multinomial response variable.

Import baseline data failures

The system audit log records the following failures when importing baseline data for a model.

BaselineFileSizeLimit
> The baseline data import fails if the file size is larger than 10 megabytes.

InvalidDataset
> The baseline data file has no valid columns.

TrainingDataNotCurrent
> The training data have been changed since the model was trained; thus, the data are not imported as baseline data. To upload the correct baseline data, go to the **Models** tab in the deployment area or upload the data via the **Model Repository**.

Supported model types

In Minitab Statistical Software, Minitab Model Ops supports all the models from the Predictive Analytics Module and certain models from the Regression menu. For more information on the Predictive Analytics Module, go to the Predictive Analytics Module.

- Analyses from the Predictive Analytics Module (*Minitab 20.1 and later*)
    - CART® Classification
    - CART® Regression
- Analyses from the Predictive Analytics Module (*Minitab 20.2 and later*)
    - Fit Model for TreeNet® Classification
    - Fit Model for TreeNet® Regression
    - Discover Key Predictors for TreeNet® Classification
    - Discover Key Predictors for TreeNet® Regression
    - Random Forests® Classification
    - Random Forests® Regression
- Analyses from the Predictive Analytics Module (*Minitab 21 and later*)
    - Discover Best Model (Binary Response)
    - Discover Best Model (Continuous Response)
- Analyses from the Predictive Analytics Module (*Minitab 21.4 and later*)
    - MARS® Regression

- Analyses from the Regression menu (*Minitab 21 and later*)
    - Fit Regression Model
    - Fit Binary Logistic Model
    - Fit Poisson Model

Important: Models must have 500 predictors or less. For a multinomial response variable, the maximum number of classes is 10.

Language and regional settings

Regional settings affect the format of numeric decimal separators and list separators for your entire subscription. The following table shows the available regions and corresponding separators.

Region	Numeric separator	List separator
United States	period	comma
United Kingdom	period	comma
France	comma	semicolon
Germany	comma	semicolon
Brazil	comma	semicolon
Mexico	period	comma
Japan	period	comma
Korea	period	comma
China	period	comma

Complete the following steps to change your language and regional settings.
1. Sign in to Minitab Model Ops.

2. Select **Language and Regional Settings** ⚙.

3. Under **Region**, select the region to use. Confirm the correct separator for numeric and list values.

4. Select **OK**.

User roles

Minitab Model Ops provides two roles: **Viewer** and **User**. Only license administrators can change user roles.

Permission	Viewer	User
View model and deployment information	X	X
View Integration tab	X	X
View drift and stability reports	X	X
Import and delete models		X
Deployment management: deploy, pause, remove models		X
Upload baseline data		X
Access to deployment settings		X

Permission	Viewer	User
Manage API keys		X
Access to deployment and system audit logs		X

Use Minitab Model Ops with Minitab Connect

Minitab Connect® empowers data users from across the enterprise with tools to transform diverse data into a governed network of data pipelines, feed analytics initiatives, and foster organization-wide collaboration.

Use Minitab Model Ops and Minitab Connect together to create a seamless system where each update to your data in Connect starts a flow that brings near-instant predictions from Model Ops. Use those predictions in your analytics initiatives to deliver insights that inspire action.

Minitab Connect provides two flows:
1. **Model Scoring (Minitab Model Ops®)**

2. **Model Stability (Minitab Model Ops®)**

You can choose the flow trigger for each flow. To automate the process, specify **Scheduled** or **Data Event**. For a scheduled flow, you can choose your start time. For a data event flow, the results are immediate. You can also specify **Manual**.

Note: Model Ops allows 1440 prediction requests per day. When you initiate a manual or scheduled flow, Connect sends all records in the source data in the request. Submitting your data this way causes redundant requests that consume the daily request limit. To submit only unscored records, you can create a data view that filters out previously scored results, then run the flow on the filtered data. Flows triggered by a data event send new or updated records and do not require this workaround.

To see examples of creating Connect tables, creating data views, and establishing flows in Minitab Connect, go to our online tutorials.

2. Create a new deployment

The home page of Minitab Model Ops is the **Deployments** page. This page contains the status of all deployments. From here, you can add new deployments and access existing ones.

1. Sign in to Minitab Model Ops to open the **Deployments** 📶 page.

2. Select **New Deployment**.

3. In **Deployment Name**, enter a descriptive name for the deployment.

4. In **Model Source**, select **Import New Model** to import a new model from your files.

   Note: Use **Select From Repository** to deploy a model from the model repository. For more information, go to Deploy from the model repository on page 10.

5. Select **Next**.

6. Under **Project File**, browse to the file that contains the model to deploy.

7. Under **Model**, select the model to deploy.

8. Select **Use training data for baseline data** to upload the training data with the model.

   For best results, it is good practice to upload the training data as the initial baseline data. If your training data has changed or you don't want to use it as baseline data, you can upload baseline data later. For more information, go to Baseline data on page 21.

9. Under **Model Name**, enter the name of the model.

10. Select **Create Deployment**.

Once the Champion model has been deployed, Minitab Model Ops provides an area that contains the deployment details, including the models, the performance and the settings. For more information, go to Deployment and model details on page 11.

Activate, pause, or delete deployments

From the **Deployments** page, you can activate, pause, or delete a deployment.

Activate 📶

Select the **Activate** button to activate or reactivate the deployment. If your subscription reaches the maximum number of active deployments, you must pause an active deployment to add a new one.

Pause ‖

Select the **Pause** button to pause the deployment. In the paused state, prediction requests, stability requests, service metrics, and storage metrics are unavailable.

Delete 🗑

Select the **Delete** button to remove a deployment. When a deployment is removed, all historical data are also removed.

**Note:** You must pause a deployment before you can delete it.

You can also activate or pause a specific deployment from its deployment area.

Import models to the model repository

The model repository displays status details for all the imported and deployed models across the platform. Learn how to add and delete models to the repository.

Import a model

Importing a model is like opening a file.

1. Sign in to Minitab Model Ops to open the **Model Repository** ☰.

2. Select **Import a model**.

3. In **Project File**, browse to and select the file that contains a model to import.

4. If you import a Minitab project file that contains multiple models, select the model from the dropdown list which contains all the supported models in the file.

   Note: For more information on supported models, go to Supported model types on page **6**.

5. Select **Use training data for baseline data** to upload the training data with the model.

   For best results, it is good practice to upload the training data as the initial baseline data. If your training data has changed or you don't want to use it as baseline data, you can upload baseline data later. For more information, go to Baseline data on page **21**.

6. In **Model Name**, enter a descriptive name for the file.

7. Select **Import Model**.

When you import a binary logistic model from Minitab Statistical Software, Minitab Model Ops requires the binary event probability threshold. In **Binary Event Probability Threshold**, enter a value between 0 and 1. The model assigns cases with an event probability greater than the threshold to the event class. For example, a value of 0.5 means that the model assigns a case to the event class when the probability of the event is higher than 0.5.

Note: The model upload fails when models have more than 500 predictors or for a multinomial response variable, have more than 10 classes. You can check the **System Audit Log** for more information.

Deploy from the model repository

After you import a model to the **Model Repository**, the status is **Ready**. To access a model from its endpoint and make predictions, create a deployment that contains one champion and up to two challenger models.

1. Sign in to Minitab Model Ops to open the **Deployments** page.

2. Select **New Deployment**.

3. In **Deployment Name**, enter a descriptive name for the deployment.

4. Select **Select From Repository** then **Next**.

5. In **Model**, select an initial champion model. You can select only a model that is **Ready** and not used in another deployment. The list of models is in alphabetical order by name, with uppercase letters before lowercase letters.

6. Select **Create Deployment**.

Delete a model

You can delete an imported model from the **Model Repository** ☰ page. Only models that have **Ready** status can be deleted. If a model is already deployed, you must first remove it from the deployment before you can delete it.

Select the **Delete** button to delete a model. When a model is deleted, all historical data are also removed.

Deployment and model details

Deployment details area

From the **Deployments** page, select a deployment name to access the model and deployment details.

The deployment area page includes access to everything you need to access the model details, deployment settings, performance graphs and more. The following tabs include specific information to help you set up and monitor your deployed models.

Models

Displays the champion and challenger models for each deployment and provides access to model information and data and promotion capabilities. This tab also includes the details about the response variable and all predictor variables. For more information on using champion and challenger models, go to Use champion/challenger models in a deployment on page **14**.

Performance

Access the Drift and Stability reports.

Drift report

The Drift report shows drift in the champion model. Drift reports update when new data are received.

Stability report

The Stability report shows the performance of the champion and challenger models. Stability reports update daily.

Settings

Includes the observation ID settings, drift alert thresholds, and email alert settings. For more information, go to Deployment settings on page **15**.

Audit Log

Records actions on the deployment. For more information, go to Deployment audit log on page **16**.

Service Metrics

Service metrics include measures of the health of the endpoint that you use to access the model.

Endpoint Status Summary

The endpoint status summary shows the HTTP status code for each HTTP request. Codes in the range 200 to 299 represent successes. Codes in the range 400 to 599 represent errors. If the graph includes unexpected codes, investigate the cause.

Response Time

The response time shows the time to complete HTTP requests. Usually, the line is relatively flat. If the graph includes an increasing trend or unexpected spikes, investigate the cause.

Storage metrics show the storage utilization for your deployment for the current day. Minitab Model Ops has a limit of 1440 data rows per deployment, per day.

Integration

Includes details about the API keys, as well as code formats to use to generate prediction and stability requests.

View model details

Minitab Model Ops stores basic model information for each model in the deployment. Use the following steps to view the model information.

1. Open the **Deployments** 🔊 page.

2. Select a deployment name to open the deployment details area.

3. Open the **Models** tab.

4. Open the settings of the model •••.

5. Select **Model information**.

The model information includes the following details.

Model Name
   The descriptive name for the model assigned upon import.

Model Type
   The type of model.

Created By
   The name of the user who imported the model.

Created
   The date and time of the model's creation.

Rename an existing model

1. Open the **Deployments** 🔊 page.

2. Select a deployment name to open the deployment details area.

3. Open the **Models** tab.

4. Open the settings of the model •••.

5. Select **Rename model**.

Download baseline data

1. Open the **Deployments** 🔊 page.

2. Select a deployment name to open the deployment details area.

3. Open the **Models** tab.

4. Open the settings of the model •••.

5. Select **Download baseline data**.

6. Select either CSV or MWX format.

Remove a model from a deployment

You can remove a challenger model from a deployment. When you remove a challenger model, the model is still available in the model repository.

1. Open the **Deployments** 🔊 page.

2. Select a deployment name to open the deployment details area.

3. Open the **Models** tab.

4. Open the settings of the challenger model •••.

5. Select **Remove model**.

Model status indicators

From the **Models** tab in the deployment area, you can see the list of the models and whether they are ready to score prediction data and produce Drift reports.

Champion models

Only Champion models can produce Drift reports.

Note: At least ten nonmissing observations are required to plot the baseline data and calculate drift for each variable.

The Champion model is valid and has at least one row of missing or nonmissing values in the baseline data for all predictors.

The Champion model is valid but does not have at least one row of missing or nonmissing values in the baseline data for all predictors. If data for one or more Champion predictors are not present, drift calculations cannot be performed for the missing predictors.

The Champion model is not valid and cannot monitor drift.

Challenger models

Minitab Model Ops does not monitor the drift of challenger models; however, you can still upload baseline data so the model is ready to monitor drift upon promotion to Champion models.

The Challenger model is valid and has at least one row of missing or nonmissing values in the baseline data for all predictors.

The Challenger model is valid but does not have at least one row of missing or nonmissing values in the baseline data for all predictors.

The Challenger model is not valid.

Use champion/challenger models in a deployment

The champion/challenger framework in Minitab Model Ops allows you to use the same production data to compare competing models. Each deployment has a single champion model that is used to produce the predictions for the production environment. The deployment also allows up to two challenger models that can run the same data and provide performance metrics to your team. If and when a challenger model becomes the best model, you can easily promote it to be the new champion model.

Champion and challenger details

Champion model

> The champion model is the best predictive model chosen from a group of models that meets your champion criteria. Upon deployment, the engineer selects the initial champion and sets the response variable and initial predictor variables.

Challenger models

> Challenger models are competing models. Each challenger model must have the same response variable as the champion model. However, challenger models can have different predictors. The challenger model that generates the best results can become the new champion model if you decide to promote it.

All models in the deployment must use the same response variable. Challenger models may share predictors and/or have different predictors. A shared predictor must have the same name and be the same type (numeric or text) across all models. Use the following information for the response and predictor variables from the **Models** tab in the deployment details area to verify that the correct variables are in any data sets you prepare to make predictions, monitor drift, and monitor stability.

Name

> Displays the name of the variable. The production and baseline data sets must use identical names as the variables in the deployment.

Type

> Displays the variable type, either continuous or categorical. The variables in the production and baseline data sets must use the same variable type as the model.

Data Type

> Displays the data type, either text or numeric. Continuous variables will always have a numeric data type and categorical variables can use text or numeric values. The variables in the production and baseline data sets must use the same data type as the model.

Classes

> Displays the class values of the variable. Only categorical variables have classes. Use the class values to make predictions.

Models

> Displays which models contain the predictor variables.

Add challenger models to a deployment

The **Models** tab contains a list of the deployment's competing models. Use the following steps to add up to two challenger models.

1. Open the **Deployments** 🔊 page.

2. Select a deployment name to open the deployment details area.

3. Open the **Models** tab.

4. Select **Add Model**.

5. Choose an available model from the repository.

   Challengers must have the same response variable and cannot be a challenger in any other deployment.

Upload baseline data for a challenger model

For best results, it is good practice to upload the training data as the initial baseline data. Minitab Model Ops monitors drift for the champion model of the deployment. After a challenger is promoted to champion, Minitab Model Ops uses the baseline data for the Drift report.

1. Open the **Deployments** 📡 page.

2. Select a deployment name to open the deployment details area.

3. Open the **Models** tab.

4. Open the settings of the challenger model ⋯.

5. Select **Upload baseline data** or **Replace baseline data**.

Evaluate the performance of challenger models

Use the model performance metrics and graphs on the **Performance** tab to evaluate the champion and challenger models. Only the champion model is live at any given time. All the prediction requests are used by the current champion model in real time. Once a day, the same prediction requests are replayed against the challengers for analytical purposes.

By assessing the challenger models with the same production data as the champion model, you can easily compare the available models in the deployment.

Promote a challenger model

Review the performance metrics to determine whether one of the challenger models has better performance than the champion model. If a challenger model has better performance than the champion model, you can promote it to be the new champion model.

1. Open the **Deployments** 📡 page.

2. Select a deployment name to open the deployment details area.

3. Open the **Models** tab.

4. Open the settings of the challenger model ⋯.

5. Select **Promote to Champion**.

Deployment settings

Find detailed guidance on the **Settings** tab.

The **Settings** tab includes the observation ID settings, alert thresholds, and email alert settings.

Observation ID

ID Name

Specify the name for the unique observation identifier that links the prediction and stability data into a complete record.

Note: The ID Name cannot be changed once a deployment begins processing prediction requests.

Require Observation ID for prediction requests

While the Observation ID is required to generate stability reports, you can also require the Observation ID for prediction requests.

Prediction Timestamp

Timestamp Variable Name

Specify the name for the timestamp variable. When you provide timestamp data with the request, the system uses the provided timestamp in place of the system generated timestamp.

Production Data Period

Select the default production period for the drift status displayed on the deployments page.

Drift Report: Variable Thresholds

Enter the values of PSI that indicate moderate drift and severe drift. The default values are traditional thresholds. The threshold for moderate drift traditionally identifies when to use your judgment regarding continued use of the champion model. The threshold for severe drift traditionally identifies when to retrain the model. The moderate drift threshold must be less than the severe drift threshold.

Enter smaller values to change the drift status in response to smaller differences between the baseline data and the production data. Enter larger values to change the drift status for larger differences between the baseline data and the production data. For more information on the PSI metric, go to Details about the Population Stability Index (PSI) on page 33.

Drift Report: Overall Drift

Specify which drift to include in the overall drift of the champion model.

- Select **Prediction drift** to include drift of the predicted response variable.

- Select **variables show data drift** and enter the number of variables to include drift of the predictor variables.

When the model exhibits moderate or severe overall drift according to your specified thresholds, you will receive an email with details about the model drift status at the end of the production period.

Alerts: Email Configuration

Specify one or more email addresses to receive alerts when the drift status of a champion model changes. Separate multiple email addresses with spaces, commas, or semi-colons.

Deployment audit log

The audit log for the deployment records the actions of the models within the deployment. Use the following steps to view the audit log for the deployment.

1. Open the **Deployments** page.

2. Select a deployment name to open the deployment details area.

3. Open the **Audit Log** tab.

Note: The system audit log for all deployments is found in the **System Audit Log** on page 5 area from the left pane navigation.

Audit log information

The log records the following items for each action. Expand an item to see additional details for the action.

Timestamp

The date and time of the action.

User

The name from the **User Information** for the account for Minitab Model Ops.

Action

This value records one of the following actions.

- ActivateDeployment
- PauseDeployment
- UpdateDeployment
- CreateApiKey
- RevokeApiKey
- AddChallenger
- PromoteChallenger
- RemoveChallenger

Result

This value records whether the action succeeds or fails.

3. Data integration

The **Integration** tab for a deployment includes details about the API keys, as well as code formats to use to generate prediction and stability requests.

Create a new API key

Create an API key for access to the deployment. Access to the endpoint ceases when the API key expires or when an authorized user revokes the API key.

Note: You must have **User** permissions, not **Viewer** permissions to create new API keys.

1. Open the **Deployments** 🔗 page.

2. Select a deployment name to open the deployment details area.

3. Open the **Integration** tab.

4. Under **API Keys**, select **Create API Key**.

5. Enter the key name and an expiration date.

   The maximum expiration date is one year.

6. Select **Create**.

Copy the API Key to the clipboard and save in a safe location. Minitab Model Ops does not save this key, and you will not be able to view or copy it later.

Note: Before you distribute this API key, make sure you are authorized by your organization to publicly share the model data. Anyone with this API key has access to send prediction and stability requests for a model without additional authentication. Minitab disclaims all liability for model data that are publicly shared in violation of your organization's policies and procedures.

Generate code formats

Generate code format for prediction requests

Generate the code required for prediction requests. You can generate the code format for a single row, multiple rows, or an entire file of data. Copy this code and use your API key and data to send to endpoints.

1. Open the 🔗**Deployments** page.

2. Select a deployment name to open the deployment details area.

3. Open the **Integration** tab.

4. Under **Sample Code**, select **Prediction**.

5. In **Data**, select whether to create code for a single row, multiple rows, or a data file. In **Method**, the choice of **cURL GET** or **cURL POST** depends on whether you use a single row or more than a single row.

   - With **Single row**, choose **cURL GET** or **cURL POST**.

   - With **Multiple rows** or **File**, use **cURL POST**.

   You can change the values for the variables. Select **Submit Request** to verify the endpoint.

   Note: Commas and semicolons are not permitted in file paths.

   For information on prediction data requirements, go to Requirements for production data on page 22.

6. Select **Copy**. Add your API key and any data values you want to include, then submit to your terminal.

   Use a leading zero in decimal numbers less than 1. For instance, use 0.25, not .25.

   Note: The sample code uses a period as the decimal separator for numeric values.

To see an example of a python request, go to Sample Python code for a prediction request.

Generate code format for stability requests

Generate the code required for stability requests. You can generate the code format for a single row, multiple rows, or an entire file of data. Copy this code and use your API key and data to send to endpoints.

1. Open the **Deployments** page.

2. Select a deployment name to open the deployment details area.

3. Open the **Integration** tab.

4. Under **Sample Code**, select **Stability**.

5. In **Data**, select whether to create code for a single row, multiple rows, or a data file. The **Method** for all data arrangements is **cURL POST**.

   You can change the values for the variables. Select **Submit Request** to verify the endpoint.

   Note: Commas and semicolons are not permitted in file paths.

   For information on stability data requirements, go to Requirements for stability data on page **23**.

6. Select **Copy**. Add your API key and any data values you want to include, then submit to your terminal.

   Use a leading zero in decimal numbers less than 1. For instance, use 0.25, not .25.

   Note: The sample code uses a period as the decimal separator for numeric values.

To see an example of a python request, go to Sample Python code for a stability request.

Region headers with files

When you use a file to submit a request, you must also specify the region.

```
curl -X POST -F 'data=@<name_of_file>' -H
"Accept-Language: en-US" -H "ApiKey: API_KEY"
'https://modeler-test.minitab.com/api/score'
```

For more information on available regions, go to Language and regional settings on page **7**.

Results from prediction and stability requests

This topic includes results from successful requests and information about errors and warnings.

Prediction results

A successful prediction request returns predictions (and probabilities for classification models), as well as the status code, and the observation identifier and timestamp, if provided. If a timestamp is not provided, the results include the server time in UTC of the prediction request.

Binary
    The results include the predicted class value and both the event and nonevent probabilities.
    Such as: "predictions":[1],"probabilities":{"0":[0.29806847464300],"1":[0.70193152535699]}

Multinomial

The results include the predicted class value and the probabilities for each class.
Such as: "predictions":[4],"probabilities":{"1":[0],"2":[0],"3":[0.1666667],"4":[0.8333333],"5":[0],"6":[0]}

Continuous

The results include the predicted response.
Such as: "predictions":[50.271428571428565]

Stability results

A successful stability request returns the actual response and the observation identifier, as well as the status code, and the date from a timestamp, if one exists. If a timestamp is not provided, the results include the server time in UTC of the prediction request.

Errors

The following table includes the most commonly used error codes with Minitab Model Ops prediction and stability requests.

Note: ERRORS apply to the entire request. If an error is received, no results are logged.

responseCode	Request Type	Details
ERROR_DATA_TYPE	Both	One or more entries has an incorrect data type. Prediction results cannot be logged. One or more entries has an incorrect data type. Stability results cannot be logged.
ERROR_INVALID_PAYLOAD	Prediction	The prediction request has no valid columns. Prediction results cannot be logged.
ERROR_MISSING_COLUMNS	Stability	One or more variable names are incorrect in your request. Stability results cannot be logged.
ERROR_MISSING_OBS_ID	Both	One or more observation ID values are missing. Prediction results cannot be logged because Observation IDs are set as required in the Settings tab. One or more observation ID values are missing. Stability results cannot be logged.
ERROR_INVALID_OBS_ID	Prediction	One or more observation ID values are invalid. Results cannot be logged.
ERROR_UNEQUAL_ROWS	Both	All columns in your request must have the same number of rows. Results cannot be logged.

Warnings

The following table includes the most commonly used warning codes with Minitab Model Ops prediction and stability requests.

Note: WARNINGS apply to individual rows of the request. If a warning is received, partial results may be logged.

responseCode	Request Type	Details
WARNING_OBS_ID	Both	Observation ID value is missing. Prediction results for this row are logged, but are not available for stability reporting. Observation ID value does not match a logged observation ID value within the last six months. Stability results for this row cannot be logged.
WARNING_SCORE_LIMIT	Prediction	You have exceeded the daily maximum number of logged prediction results for this model. Minitab Model Ops has a limit of 1440 data rows per model, per day. Depending on when the daily maximum was reached, some prediction results may not have been logged.
WARNING_STABILITY_LIMIT	Stability	You have exceeded the daily maximum number of logged stability results for this model. Minitab Model Ops has a limit of 2880 data rows per model, per day. Depending on when the daily maximum was reached, some stability results may not have been logged.
WARNING_TIME_RANGE	Prediction	Timestamp range is not valid. Prediction results for this row cannot be logged. For more information on date/time ranges, go to Date/time data requirements on page 23.
WARNING_TIME_VALUE	Prediction	Timestamp value is not valid. Prediction results for this row cannot be logged. For more information on date/time formats, go to Date/time data requirements on page 23.
WARNING_UNKNOWN_REPONSE	Stability	Actual response is missing or contains a level that does not exist in the model. Stability results for this row cannot be logged.

Data requirements

This topic specifies the data requirements for baseline, production, stability, and date/time data.

Baseline data

Drift reports require the baseline data for your model. The baseline data set should include predictor columns that contain the data used for training the model, and may also include an optional column that contains the actual response.

Training data are the data that were used to fit the model. Often, you want to upload the training data after you import the model to use the training data as baseline data.

For more information on importing baseline data, go to Upload baseline and production data on page 30.

Requirements for baseline data

- Use the same variable names and same variable types that the model uses. The names are case-insensitive.

- Use the same number of rows of data in all columns.

- Save baseline data in CSV format. Minitab Model Ops also accepts Minitab worksheet files (MWX), but converts to a CSV format upon upload.
  Note: Use a comma or a semicolon as the delimiter, depending on your regional settings.

- The file size of the CSV must be 10 megabytes or less. Large MWX files may fail if they exceed this limit when converting to CSV.

Example of baseline data

Baseline Data

Age	Sex	Cholesterol	Slope
55	male	233	1
63	male	187	1
27	female	145	3

Table 1 shows sample baseline data. The columns of predictors contain the data for training the model.

Production data

The production data are also used in the drift report. The production data set contains the predictor variables used to make predictions based on the model. The data set is formed by cumulating the predictor variable values from calls to the endpoint. The data set can also include an optional observation ID and an optional date/time. If you do not specify date/time information, Minitab Model Ops uses the current server time.

For more information on importing production data, go to Upload baseline and production data on page **30**.

Requirements for production data

- Use the same variable names and same variable types that the model uses. The names are case-insensitive.
- Use an optional observation ID column, as specified in **Settings**. If you do not include observation ids, Minitab Model Ops cannot generate stability reports with this data. For more information, go to Deployment settings on page **15**.
- Use an optional date/time column, as specified in **Settings**. If you do not specify date/time data, Minitab Model Ops uses the date and time of the prediction request in UTC.
- Use the same number of rows of data in all columns.

Note: When using a file, save the data in CSV format. Use a comma or a semicolon as the delimiter, depending on your regional settings, and use UTF-8, UTF-8-BOM, or UTF-16-LE as the encoding.

Example of production data

Production Data

Age	Sex	Cholesterol	Slope	ObservationID
55	male	233	1	100987
63	male	187	1	100988
27	female	145	3	100989

Table 2 shows sample production data. The columns of predictors contain the data for the calculation of the predictions. If you use the production data to monitor drift, the drift report uses the date and time of the prediction request in the results, unless you add an optional date/time column. You can also use an optional observation ID column that you name on the **Settings** tab. For this example, the observation ID column is *ObservationID*.

Stability data

The stability data are used to create a stability report. The stability data set must contain the response values and the observation IDs.

For more information on importing stability data, go to Send stability data on page **31**.

Requirements for stability data

- Include the name of the response variable. The name is case-insensitive.
- Include the observation ID that corresponds to the prediction for the actual value. Use the same name as specified in **Settings**.
- Use the same number of rows in the columns.

Note: When using a file, save the data in CSV format. Use a comma or a semicolon as the delimiter, depending on your regional settings, and use UTF-8, UTF-8-BOM, or UTF-16-LE as the encoding.

Example of stability data

Stability Data

Heart Disease	ObservationID
0	100987
1	100988
1	100989

Table 3 shows the stability data. The stability data set has a column for the response variable with the name *Heart Disease* and a column with the name *ObservationID* for the unique identifier that links the prediction and stability information into a complete record.

Reserved variable names

Commas and semicolons cannot be used in file paths. Also, the following names are reserved and cannot be used as variable names.

- "prediction-score", "prediction_score"
- "mtb-correlation-id", "mtb_correlation_id"
- "mtb-timestamp", "mtb_timestamp"

Date/time data requirements

Minitab Model Ops requires Coordinated Universal Time (UTC) date/time formats for prediction and stability requests. All date/time information, whether a Z is used or not, are assumed to be in UTC time. If you do not specify date/time information, Minitab Model Ops uses the current server time.

Acceptable date formats

Minitab Model Ops accepts the following date format.

- *yyyy-MM-dd*

To specify dates, use the following format.

- 4 digits for the year (yyyy)
- 2 digits for the month (MM)
- 2 digits for the day (dd)

Separate the year, month, and day with dashes. For example, *2023-08-14* is August 14 of the year 2023. Because this format does not include time, the time is set as midnight and the time zone is Coordinated Universal Time (UTC).

Note: Excluded date information is not permitted. You must specify a complete date.

Acceptable date/time formats

Minitab Model Ops accepts the following date/time formats.

- *yyyy-MM-dd HH:mm:ss*
- *yyyy-MM-dd HH:mm:ssZ*
- *yyyy-MM-ddTHH:mm:ss*
- *yyyy-MM-ddTHH:mm:ssZ*

To specify date and time, use the following format.

- 4 digits for the year (yyyy)
- 2 digits for the month (MM)
- 2 digits for the day (dd)
- Blank space, or T, as separator between the date and the time
- 2 digits for the hours (HH), enter the hours in 24-hour format
- 2 digits for the minutes (mm)
- 2 digits for the seconds (ss)
- Optional Z

Separate the year, month, and day with dashes. Separate the hours, minutes, and seconds with colons.

Note: Any excluded time information is the equivalent of entering 0 for the value. For example, if a time entry does not have minutes and seconds, the minutes and seconds are 0.

For example, use any of the following to represent 8:30 in the morning in Coordinated Universal Time (UTC) on August 14 of the year 2023.

- *2023-08-14 08:30:00*
- *2023-08-14 08:30:00Z*
- *2023-08-14T08:30:00*
- *2023-08-14T08:30:00Z*

Note: When you use a T instead of a blank space, you must also include HH:mm. If a time entry does not have seconds, the seconds are 0.

Using future and past dates

Requests with future date/time values are not supported. Prediction requests with prior date/time values beyond 3 months are not logged; thus, are not available in the drift and stability reports. Stability data are logged only if a matching observation id was found within the last 6 months.

Data formats for missing values

This topic includes details about which missing data are allowed and how to properly indicate missing values.

Acceptable missing value formats

You can represent numeric missing values using the symbols shown in the following table.

Numeric	Example	With missing numeric values
*	77,"male", 145, 1, 0	*,"male", *, 1, 0
'*'	77,"male", 145, 1, 0	'*',"male", '*', 1, 0
"*"	77,"male", 145, 1, 0	"*","male", "*", 1, 0
	77,"male", 145, 1, 0	,"male",, 1, 0
' '	77,"male", 145, 1, 0	' ',"male",' ', 1, 0
" "	77,"male", 145, 1, 0	" ","male"," ", 1, 0

You can represent text missing values using the symbols shown in the following table.

Text	Example	With missing text values
	77,"male", 145, 1, 0	77,, 145, 1, 0
''	77,"male", 145, 1, 0	77,'', 145, 1, 0
""	77,"male", 145, 1, 0	77,"", 145, 1, 0
' '	77,"male", 145, 1, 0	77,' ', 145, 1, 0
" "	77,"male", 145, 1, 0	77," ", 145, 1, 0

Missing baseline data

The baseline data set must include at least one row of data. Minitab Model Ops uploads predictor columns and the actual response column, but ignores other additional columns in your data file. If a predictor is missing from the baseline data file, Minitab Model Ops uploads the file but cannot perform drift calculations for the missing variables.

Missing values of the predictors are allowed if they follow the acceptable formats.

Missing production data

Minitab Model Ops does not require data for all predictors of a model to submit a prediction request.

You can submit missing values for the predictors.

- If the observation ID column is missing, Minitab Model Ops can still provide predictions, but cannot monitor stability.
- If the date/time column is missing, Minitab Model Ops uses the server time, in UTC.

Missing stability data

Minitab Model Ops requires a paired actual response and an observation ID to log the stability data.

You can submit missing values for the actual response and the observation ID, but if the observation ID column is missing, Minitab Model Ops cannot monitor stability.

Details about prediction requests with missing date/time data

Date and time data are optional for prediction requests. If you do not provide date/time data, Minitab Model Ops uses the current server time. To include date and time data, use the name of the **Prediction Timestamp** variable from the **Settings** tab. In this example, we use *Time Column*.

- To submit a single request with a timestamp, use the following format: *"Time Column":["2023-05-13 17:00:00"]*.

  Note: To submit a single request without date/time data, omit the date/time column.

- To submit a request with multiple timestamps, use the following format: *"Time Column":["2023-05-13 17:00:00","2023-05-13 17:30:00","2023-05-13 18:00:00"]*.

Important: Minitab Model Ops does not accept prediction requests that include a mix of timestamps and missing values. If you have missing timestamp data, you must submit a separate request that does not include the date/time column.

4. Monitor model drift

Drift report details

Minitab Model Ops displays the Drift report for the deployed champion model on the **Performance** tab.

When a deployment is paused, you can upload baseline data, but not production data. You must have at least ten observations of baseline data to generate a chart of baseline data and at least ten observations of production data to generate a chart of production data. For information on uploading data and data requirements, go to the following topics.

Once you have baseline and production data for your model, you can select the production period and date range for your report.

	Response Variable	Drift		Baseline Data	Production Data	Prediction Drift (PSI)	Drift Over Time
>	Heart Disease	✓	Minimal			0.073787	

	↓ Predictor Variable	Drift		Baseline Data	Production Data	Data Drift (PSI)	Drift Over Time
>	Age	❗	Severe			0.558643	
>	Chest Pain Type	✓	Minimal			0.038945	
>	Cholesterol	❗	Severe			0.351899	
>	Exercise Angina	✓	Minimal			0.031406	
>	Fasting Blood Sugar	⚠	Moderate			0.190540	

You can open the expander for each variable to view more information.

	Response Variable	Drift		Baseline Data	Production Data	Prediction Drift (PSI)	Drift Over Time
⌄	Heart Disease	✓	Minimal			0.073787	

Data = Production
Heart Disease = Yes
● Production = 63.415%

Data
■ Baseline ■ Production

Data	N	N*
Baseline	250	0
Production	41	0
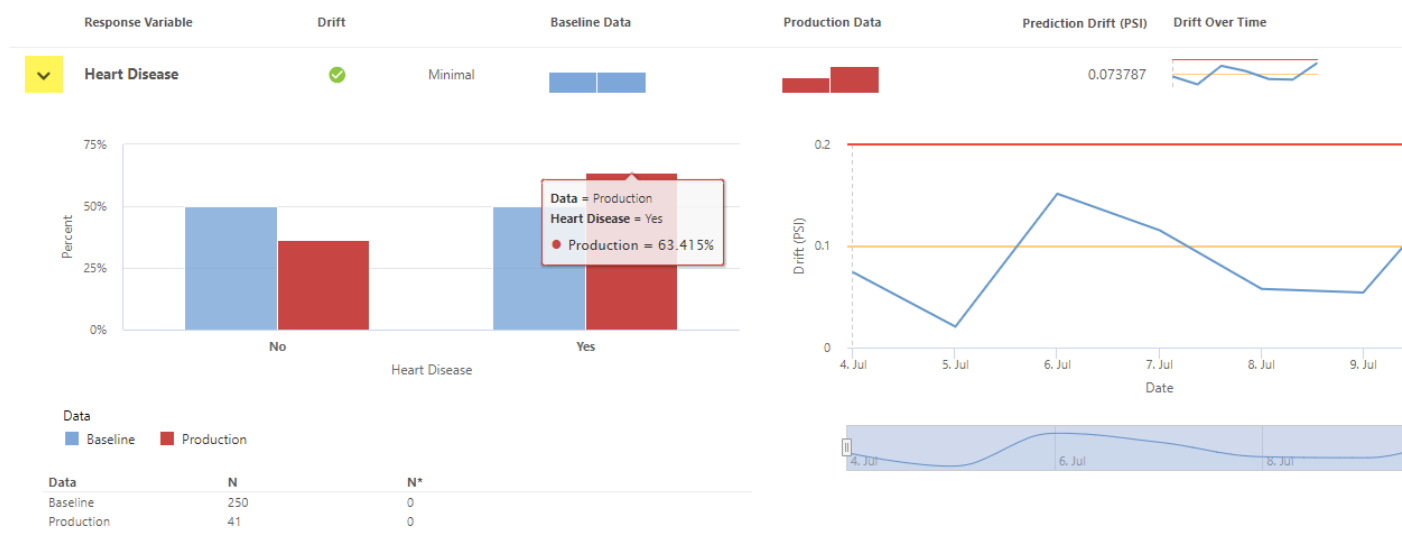
Note: The Drift Over Time plot will not display historical data for a variable when the variable is no longer in the deployment.

Types of model drift

Drift measures the similarity between your baseline data and the production data. Use differences between the baseline data and the production data to determine when to try to improve your model. Minitab Model Ops reports prediction drift and data drift.

Prediction drift
> This type of drift represents a drift in the model's predictions over time. When this drift is substantial, you may want to retrain the model with additional data or replace the model.

Data drift
> This type of drift represents a change in the distribution of the input variables. When this drift is substantial, you may want to retrain the model with additional data.

If there is enough production data to determine the drift status, the drift status is minimal, moderate or severe, according to the definitions in the Settings tab on page **15**.

Examine drift reports

The tutorial has examples of drift reports with baseline data only, minimal drift, and moderate and severe drift. To see these examples, go to Examine drift reports.

Monitor model stability

Minitab Model Ops displays model stability on the Stability report.

Stability report details

Minitab Model Ops displays the Stability report for the deployed champion model and the competing challengers on the **Performance** tab. Stability reports update daily.

When a deployment is paused, stability data are not logged. Stability data are logged only if a matching correlation id was found within the last 6 months. For information on uploading data and data requirements, go to the following topics.

- Requirements for stability data on page **23**
- Send stability data on page **31**

Once you have stability data for your model, you can select the production period and date range for your report.
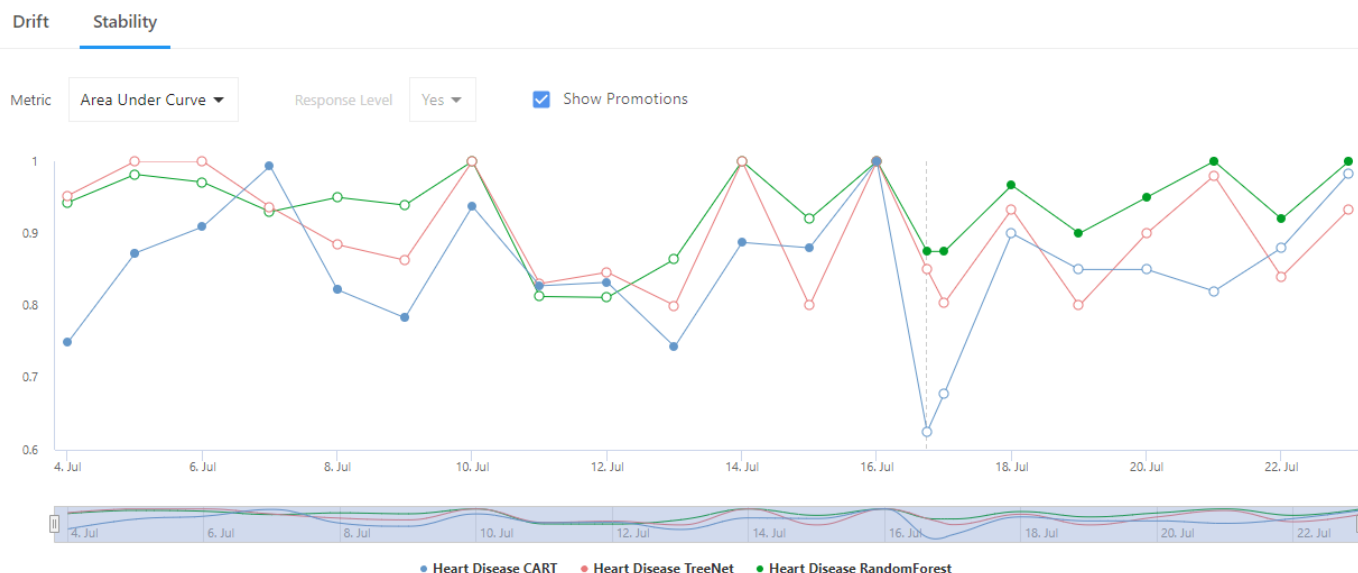
The Random Forest model was promoted on July 16.

Measures of performance for models with categorical response variables

Classification Accuracy

The classification accuracy indicates how often the model accurately classifies the events. Larger values indicate better performance.

AUC

The values for the area under the ROC curve usually range from 0.5 to 1. Larger values indicate a better classification model. When the model can perfectly separate the classes, then the area under the curve is 1. When the model cannot separate the classes better than a random assignment, then the area under the curve is 0.5.

For a multinomial response, the AUC is set to 0 under the following conditions.

- If a class is missing from the actual values, then the AUC for that class is 0. For example, if a model includes the classes 1 to 5 but the stability data do not include any actual values of 4, then the AUC for 4 is 0.

- If every actual value has the same class, then the AUC for all classes is 0. For example, if a model includes the classes 1 to 5 but the actual values are all 2, then the AUC is 0 for all classes.

Average Predicted Probability and Proportion of Event Cases

This graph displays either the average predicted probability of an event or the proportion of cases classified as an event over time.

- When the production data period is daily, the plotted points are daily averages.

- When the production data period is weekly, the plotted points are weekly averages.

- When the production data period is monthly, the plotted points are monthly averages.

Measures of performance for models with continuous response variables

R-squared

The higher the R^2 value, the better the model fits your data.

MAD

The mean absolute deviation (MAD) expresses accuracy in the same units as the data, which helps conceptualize the amount of error. Outliers have less of an effect on MAD than on R^2. Smaller values indicate a better fit.

Average Predicted Response

This graph displays the average predicted response over time.

- When the production data period is daily, the plotted points are daily averages.
- When the production data period is weekly, the plotted points are weekly averages.
- When the production data period is monthly, the plotted points are monthly averages.

Examine stability reports

The tutorial has examples of stability reports. To see these examples, go to Examine stability reports.

Upload baseline and production data

Baseline and production data are required to monitor model drift.

Upload baseline data for a model

Baseline data are required to monitor drift. Follow these steps to add or replace baseline data for an existing model. For information on baseline data requirements, go to Requirements for baseline data on page **21**.

1. From the **Deployments** page, select the deployment to open the tabs on the deployment area page.

2. Open the **Models** tab.

3. Open the settings of the model •••.

4. Select **Upload baseline data** or **Replace baseline data**.

5. Browse to the file that contains the baseline data and specify which columns to upload for the predicted response and predictors.

6. Select **Upload**.

Note: You can also upload baseline data for a model from the model repository.

Baseline data status indicators

From the **Models** tab in the deployment area, you can see the list of your models and whether they have valid baseline data.

Icon	Details
✓	The baseline data upload is valid and you have at least one row of missing or nonmissing values in your baseline data for all predictors in the model.
⚠	The baseline data upload is valid but you are missing data for some of the predictors in the model.

Check the **System Audit Log** on page 5 ⟲ area from the left pane navigation to see specific baseline data errors.

Send production data

To score a model, Minitab Model Ops requires new production data. You can create a Minitab Connect table that contains the new predictions, or you can send the data via Python code, a terminal, or a script in a terminal. For help on generating the required code format for prediction requests, see the sample code from the **Integration** tab.

Note: Prediction requests with prior date/time values beyond three months are not logged; thus, are not available in the drift and stability reports.

For information on production data requirements, go to Requirements for production data on page 22.

Send stability data

To assess a model, Minitab Model Ops requires stability data. You can create a Minitab Connect table that contains the new data, or you can send the data in via Python code, a terminal, or a script in a terminal. For help on generating the required code format for stability requests, see the sample code from the **Integration** tab.

Note: Stability data are logged only if a matching correlation id was found within the last six months.

For information on stability data requirements, go to Requirements for stability data on page 23.

Download baseline and production data

Download baseline data for a model

Follow these steps to download baseline data for a champion or challenger model.
1. From the **Deployments** page, select the deployment to open the tabs on the deployment area page.

2. Open the **Models** tab.

3. Open the settings of the model •••.

4. Select **Download baseline data**.

5. Select the file format, then select **OK**.

Download production data

Follow these steps to download the production data from the drift and stability reports.
1. From the **Deployments** page, select the deployment to open the tabs on the deployment area page.

2. Open the **Performance** tab.

3. Open the settings of the model •••.

4. Select **Download Data**.

5. Select the file format, then select **OK**.

Minitab Model Ops downloads date/time data with this UTC format, *yyyy-MM-ddTHH:mm:ssZ*.

Report period and date settings

Minitab Model Ops displays the model performance graphs on the **Performance** tab. Toggle between **Drift** and **Stability** to see the deployment graphs.

Report period

The deployment drift and stability reports use the same report periods and start and end dates. You can choose whether to represent the data in a daily, weekly, or monthly period.

Daily

Drift and stability metrics are calculated each day and daily values are plotted over time.

Weekly

Drift and stability metrics are calculated each week and weekly values are plotted over time.

Monthly

Drift and stability metrics are calculated each month and monthly values are plotted over time.

Note: The **Production Data Period** setting on the **Settings** tab applies to the overall drift status displayed on the **Deployments** page and does not affect the period on the performance reports.

Start and end dates

The start and end dates of the reports are in UTC time.

The Drift reports contains daily reports over a two-week date range. Select Next tor Previous to scroll to the specific Drift report.

The report start date cannot be earlier than the deployment date unless you have added prediction and/or stability data with dates prior to the deployment date. See the following examples:

- If you create a new deployment on June 1 and have not imported data, the earliest report start date is June 1.

- If you create a new deployment on June 1 and have imported data from May 1 - 31, the earliest report start date is May 1.

Note: Prediction requests with prior date/time values beyond three months are not logged; thus, are not available in the drift and stability reports. Stability data are logged only if a matching correlation id was found within the last six months.

The report end date cannot be later than today's date.

Refresh report data

Minitab Model Ops displays the date and time (in UTC time) of the last report update. If a report is not current, select **Refresh Data** to update.

Details about the Population Stability Index (PSI)

Population Stability Index (PSI) measures the difference in distributions between two samples. A larger value of PSI indicates a greater change, or more serious drift, in the distribution of a variable between two samples. Usually, a PSI value of 0.1 shows that two distributions are moderately different. A PSI value of 0.2 shows that two distributions are severely different.

For moderate or severe drift, investigate the reasons for these differences and their effect on the model's performance. Also, consider whether to change the thresholds based on how well they correspond to model performance. For example, PSI tends to increase as sample sizes increase, so the possibility exists that larger thresholds are appropriate for a process with hundreds of observations per production period.

The calculation of PSI compares the percentages of data in the same categories for two samples of data. Use the following steps to calculate the PSI for a continuous variable.

1. Sort the baseline data in descending order.

2. Find the ideal size of a bin, each containing approximately 10% of the data.
    $$B = \lfloor n / 10 \rfloor + 1$$
    where

B	the number of observations in the bin
n	the number of observations in the data set
$\lfloor n / 10 \rfloor$	is the floor of $n/10$

3. Add B observations to the first bin. Tied values are added to the same bin. Thus, bin sizes may not be equal.

4. Continue for each bin. If the last bin is less than 50% of B, then move those observations to the previous bin.

5. Calculate the percentage of the production data in the same bins as the baseline data.

6. Calculate the PSI statistic with the following equation.
    $$S = \sum_{i=1}^{j} \left[(t - p) \ln \left(\frac{t}{p} \right) \right]$$

where

S	the Population Stability Index (PSI)
j	the number of bins, usually 10
t	the percent of the baseline data in the bin
p	the percent of the production data in the bin

For a categorical variable, each level is a bin. Otherwise, the calculation is the same.

Index

Minitab ⮞ Our mission is to help people discover valuable insights in their data.

Minitab helps companies and institutions to spot trends, solve problems and discover valuable insights in data by delivering a comprehensive and best-in-class suite of data analysis and process improvement tools. Combined with unparalleled ease-of-use, Minitab makes it simpler than ever to get deep insights from data. Plus, a team of highly trained data analytic experts ensure that users get the most out of their analysis, enabling them to make better, faster and more accurate decisions.

For nearly 50 years, Minitab has helped organizations drive cost containment, enhance quality, boost customer satisfaction and increase effectiveness. Thousands of businesses and institutions worldwide use our products to uncover flaws in their processes and improve them. Unlock the value of your data with Minitab.

Automation and Reporting

Minitab Connect®
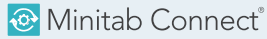
Integrate and transform data for analysis, reporting and monitoring
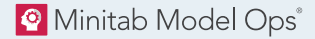

Data Analysis & Predictive Modeling

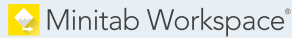**Minitab®**

Powerful statistical software everyone can use

SPM®

Machine Learning and Predictive analytics software
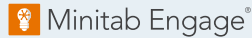

Model Deployment and Monitoring

Minitab Model Ops®

Model lifecycle management on a simple yet powerful platform

Visual Business Tools

Minitab Workspace®

Visual tools to ensure process and product excellence

Project Ideation & Execution

Minitab Engage®

Start, track, manage, and execute innovation and improvement initiatives

---

Self-paced Learning

Education Hub™

Master statistics and Minitab anywhere with online training

Quality Solutions

Real-Time SPC™
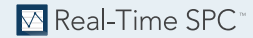
Monitor, respond, and deliver immediate quality and process monitoring

---