# LAB: Time Series Forecasting LAB Forecasting Store Sales

## Download the Dataset

From the dataset folder download the dataset named **Dataset\_Time\_Series\_II\_LAB\_Store\_Sales.csv** and save it in your preferred folder on your computer.

## Sign in to the Azure ML studio

1. Sign in to Azure Machine Learning studio at ml.azure.com.
2. In the left pane, click on **Automated ML** under the **Authoring** section.
3. Click on **+New automated ML job**.

## Create and load dataset

1. On the **Basics settings** tab, enter the **Job** name and **Experiment** name. Enter the names like:   
   **FirstName\_LastName\_Time\_Series\_II\_LAB\_Job**.  
     
   **FirstName\_LastName\_Time\_Series\_II\_LAB\_Experiment**.
2. Click **Next**.

## Task Type and Data

1. In the **Select Task Type** drop down menu, select **Time Series Forecasting**.
2. For **Select Data** click **Create** to create a new data asset from the file on your local computer.
3. On the page **Set the name and type for your data asset**:
4. Enter the **Data asset** name as **Dataset\_Time\_Series\_II\_LAB\_Store\_Sales**
5. For the **Type**, select **Tabular** from the dropdown list.
6. Click **Next**.

## Choose a source for your data asset

1. Select **From local files**
2. Click **Next**.

## Select a datastore

Here we specify the Azure Storage location to upload our data.

1. For the **Datastore type**, select Azure Blob Storage.
2. In the list of datastores, select *workspaceblobstore*.
3. Select **Next**.

## Choose a file or folder

1. Click on the **Upload files or folder** dropdown menu and select the **Upload files** option.
2. Browse to the location where you saved the **Dataset\_Time\_Series\_II\_LAB\_Store\_Sales.csv** file and select **Open**.
3. After the file uploads, select **Next**.

## Settings

1. On the **Settings** page browse your data to check the values and see if anything is unusual or out of order.
2. Click **Next**.

## Schema

For this lab, we choose to ignore the following columns:  
**RowID  
CustomerID  
Country  
ProductID**

Click **Next**.

## Review

Click **Create** to create your dataset.

## Task type & data

When your dataset is ready, Azure ML studio returns back to the Task type & data page.

1. In the **Select task type** dropdown menu the choice **Time Series Forecasting** should already be there. If not, select it.
2. Click on the radio button to the left of the dataset  
   **Dataset\_Time\_Series\_II\_LAB\_Store\_Sales**
3. Click **Next** to continue.

## Task Settings

1. In the **Target column** dropdown list, select the column to use for the model predictions, in this case **Sales**.

## Forecasting Settings

1. For the **Time Column**, select **Order Date**.
2. Leave the **Autodetect time series identifiers checked**.
3. Leave the **Autodetect frequency** checked. This is the setting that we define how often our past observations were collected. If you look into the dataset, data was collected daily.
4. Leave the **Autodetect Forecast Horizon** checked.
5. Leave **Enable deep learning** unchecked.

## Limits

1. For the **Experiment timeout (minutes) enter 65**. We want the experiment to end in 65 minutes. It might take more than 65 minutes.
2. Check to **Enable Early Termination**. We do not want the experiment to run for a long time because we have limited resources available.
3. Leave the rest of the **limits settings** as they are.

## Validate and test

1. **Validation Type**: Enter k-fold cross validation. In k-fold cross validation, the dataset is split into multiple folds. Then, as the model goes through training iterations (cross validations), in this case 5, it is trained on four out of the five folds and the last fold is used as the test fold.
2. **Number of cross validations**: 5
3. **Cross Validations Step Size**: Leave it blank and let Azure AutoML decide the size of the folds.
4. Click **Next**.

## Compute

1. **Select compute type**: From the drop down menu select **Compute Instance**.
2. **Attention**: Make sure your compute instance is running. If it is not running, navigate to **Compute** in the left pane and start your machine. Your compute instance will stop after one hour to save you resources.
3. **Select Azure ML compute instance**: Select the compute instance you have already created. If you do not have a compute instance setup, go to the lab " **Create an Azure Cloud Workspace and compute instance**" and create one.

## Review

Click **Submit Training Job**.

# Second Part of the LAB – Experiment B

In this second part of the lab, create a new Auto ML job. Create a new job within the same experiment you have used in the previous part of the lab.

## Basic Settings

1. In the basic settings, create a new job named:  
   **FirstName\_LastName\_Time\_Series\_II\_LAB\_Job\_B**
2. For Experiment name click Select **Existing**.
3. For the existing experiment, select  
   YourFirstName\_YourLastName\_Time\_Series\_II\_LAB\_Experiment

Continue with the rest of the settings like in the previous Time Series job.

## Forecasting Settings

1. **Time Column**Select **Order Date**.
2. **Autodetect time series identifiers**  
   Uncheck the **Autodetect time series identifiers** and from the list of columns select the **OrderID** and **CustomerName**.
3. **Autodetect frequency**  
   Uncheck the **Autodetect frequency**. This is the setting that we define how often our past observations were collected. If you look into the dataset, data was collected daily. Select **day** from the drop-down menu.
4. **Use a target aggregation function**  
   In this setting we define what will happen to the values of the target variable (dependent/response) after our records have been grouped according to the time series identifiers we have selected. In this case it makes sense to use **sum**() so that the sales amount from all the items in the order are added.
5. **Autodetect Forecast Horizon**  
   Uncheck the **Autodetect Forecast Horizon** and enter 15. Remember, this 15 refers to the number of periods we have defined for the frequency in our time series. Since we defined the frequency as day, 15 means 15 days.
6. Leave **Enable deep learning** unchecked.

## Limits

1. **Number of Trials**  
   For number of trials enter 1000
2. For the **Experiment timeout (minutes) enter 65**.
3. Check **enable early termination**.
4. Leave the rest of the **limits settings** as they are.

## Validate and test

1. **Validation Type**: Enter k-fold cross validation.
2. **Number of cross validations**: 5
3. **Cross Validations Step Size**: Leave it blank and let Azure AutoML decide the size of the folds.
4. Click **Next**.

## Compute

1. **Select compute type**: From the drop down menu select **Compute Instance**.
2. **Select Azure ML compute instance**: Select the compute instance you have already created. If you do not have a compute instance setup, go to the lab " **Create an Azure Cloud Workspace and compute instance**" and create one.

## Review

Click **Submit Training Job**.