# AutoML Time Series Forecasting LAB I Data: Bike Rental Data

## Download the Dataset

From datasets folder download the dataset named **Dataset\_Time\_Series\_I\_LAB\_BikeRentals.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\_I\_LAB\_Job**.  
   **FirstName\_LastName\_Time\_Series\_I\_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\_I\_LAB\_BikeRentals**
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 file **Dataset\_Time\_Series\_I\_LAB\_BikeRentals.csv**   
   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, choose to ignore the **casual** and **registered** columns. These columns are a breakdown of the cnt column and we will not include them.

Click **Next**.

## Review

Click **Create** to create your dataset.

## Task type & data

When your dataset is ready, the 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. The options include classification, regression, time series forecasting, natural language processing (NLP), or computer vision.
2. Click on the radio button to the left of the dataset  
   **Dataset\_Time\_Series\_I\_LAB\_BikeRentals**
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 **cnt**. The numbers in this column indicate the number of bikes rented by day.
2. Click on **View additional configuration settings** to see the choices but leave the default values.
3. Click on **View Featurization settings – Read through will not adjust**

## Forecasting Settings

1. For the **Time Column**, select **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. For the **Autodetect Forecast Horizon** uncheck it and enter 14, that is two weeks.
5. Leave **Enable deep learning** unchecked. This setting will examine in detail the features of the model but it will take a lot more time and resources to run.

## Limits

1. For the **Experiment timeout (minutes) enter 60**. We want the experiment to end after this time interval. It will take about 10-15 minutes to setup and run the experiment and it will take another 3 minutes for each iteration.
2. Check to activate the **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

Leave all the choices as they are on Azure ML and click **Next**.

## Review

Click **Submit Training Job**.