

# Database Management Systems

Lecture 6

Azure Machine Learning\*

Azure Stream Analytics\*

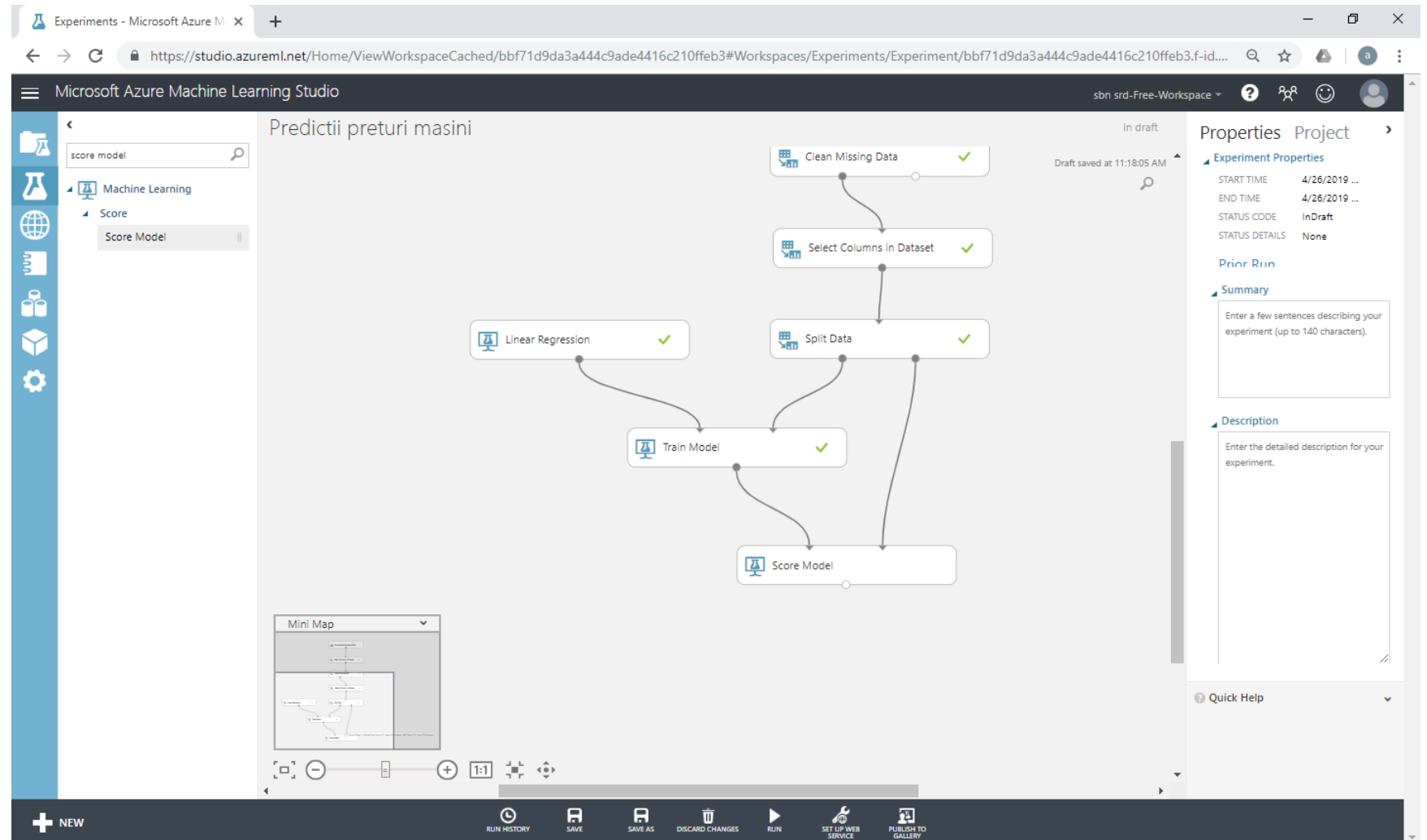
\* not among the exam topics

# Azure Machine Learning

# Car Price Prediction

\* testing the model - *Score Model* module

• *Run experiment*

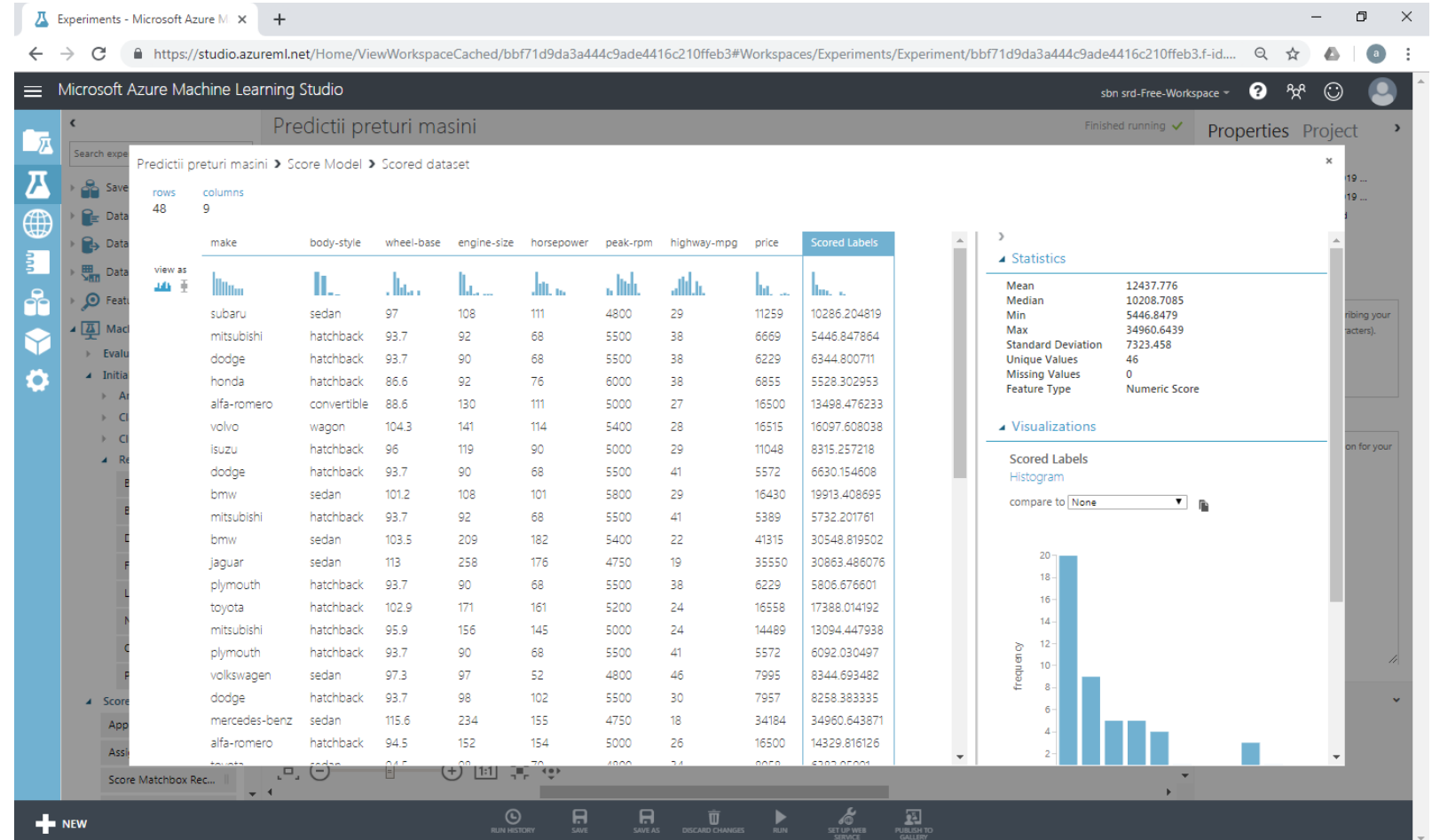


# Car Price Prediction

\* testing the model

- *Score Model* output port-> *Visualize*

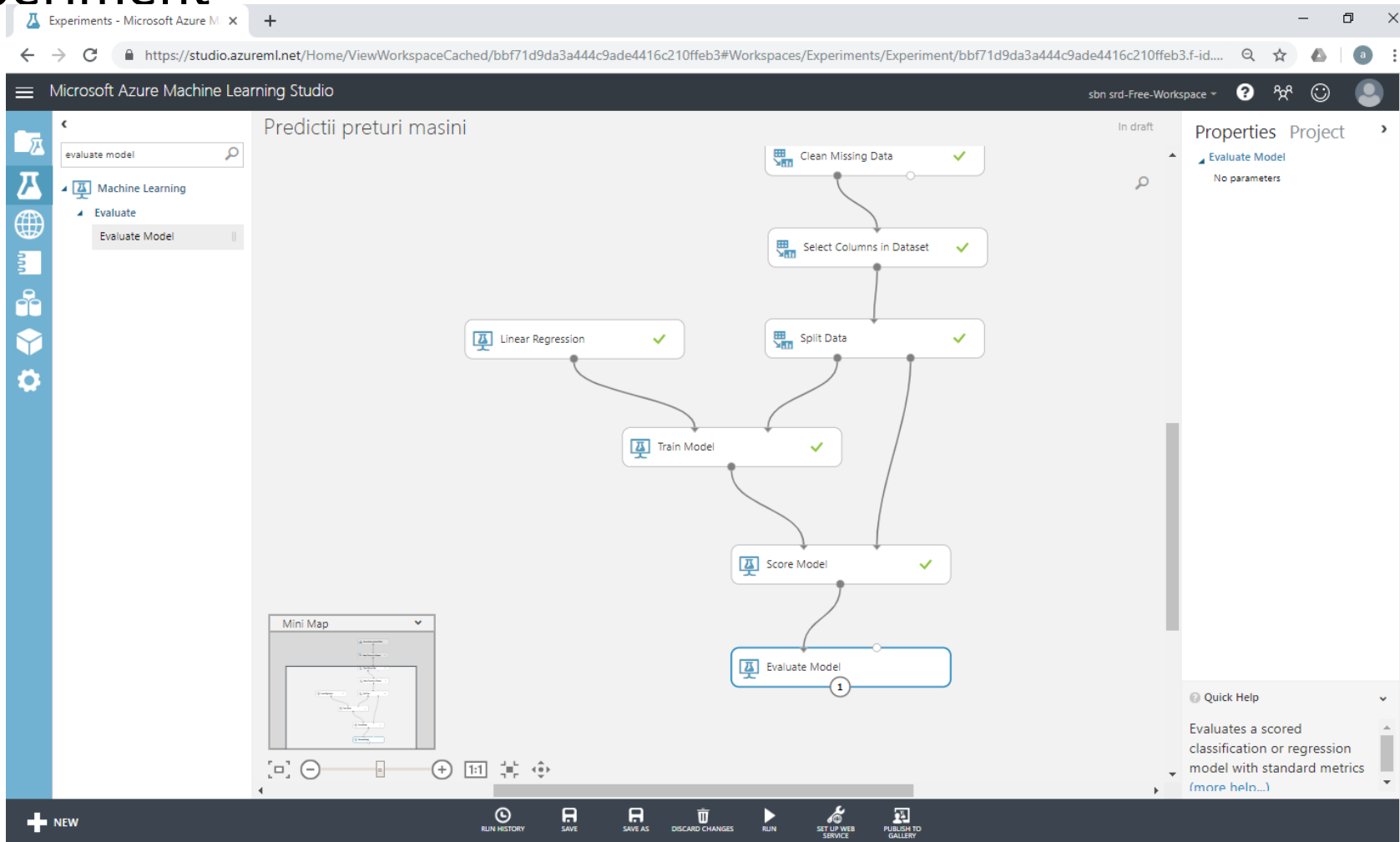
- estimated / actual values for the *price* column



# Car Price Prediction

\* testing the model

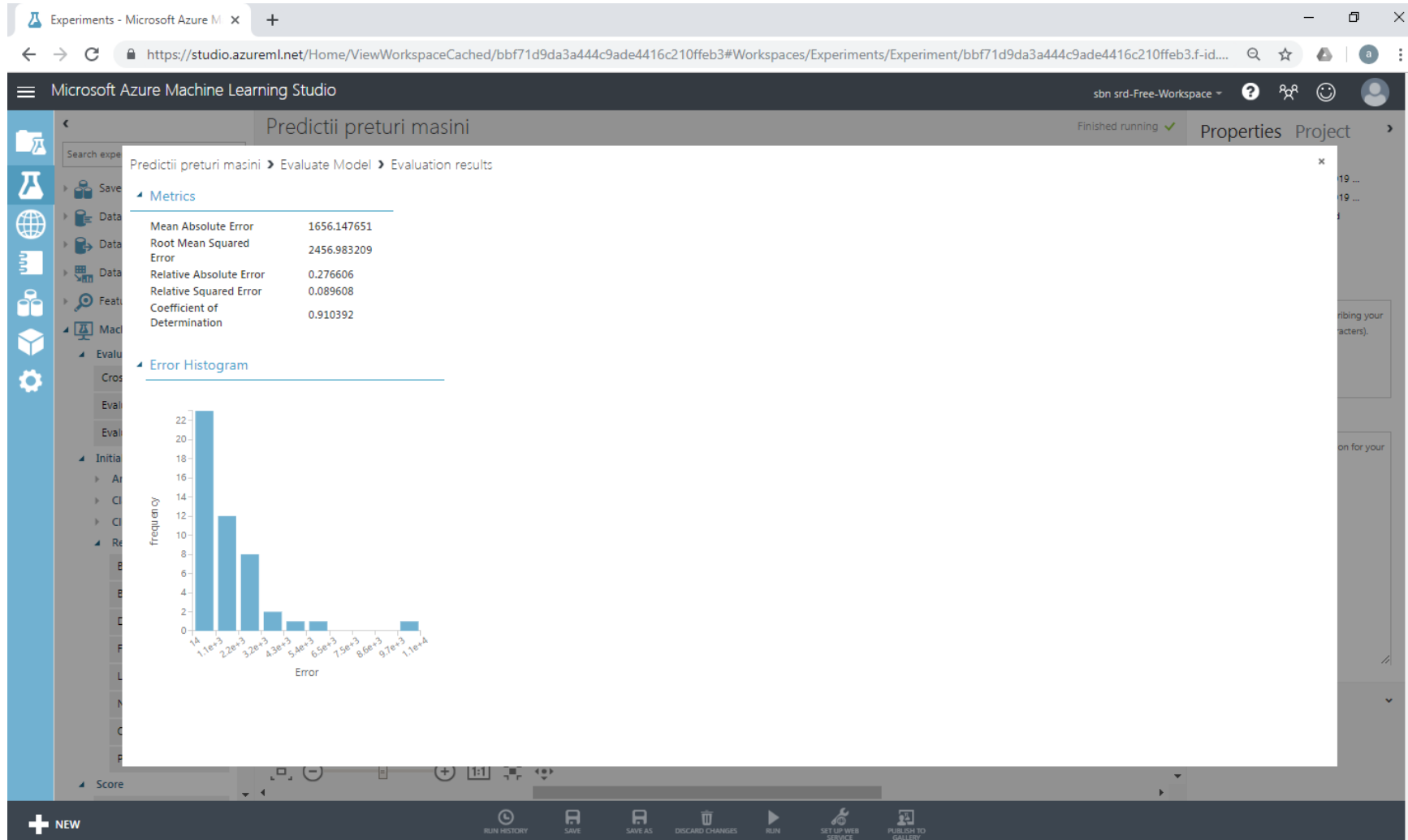
- *Evaluate Model* module
- *Run experiment*



# Car Price Prediction

\* testing the model

- *Evaluate Model* output port -> *Visualize*



# Car Price Prediction

\* eliminate resources

The screenshot displays the Microsoft Azure Machine Learning Studio interface. The left sidebar contains navigation options: PROJECTS, EXPERIMENTS, WEB SERVICES, NOTEBOOKS, DATASETS, TRAINED MODELS, and SETTINGS. The main area is titled 'experiments' and shows a table of experiments. The table has columns for NAME, AUTHOR, STATUS, LAST EDITED, and PROJECT. Two experiments are listed: 'Predictive Experiment - Mini ...' and 'Predictii preturi masini'. The second experiment is selected. To the right of the table, a workflow diagram is shown, illustrating the process of training and evaluating a linear regression model. The workflow steps are: Automobile price data (Raw), Select Columns in Dataset, Clean Missing Data, Select Columns in Dataset, Split Data, Linear Regression, Train Model, Score Model, and Evaluate Model. Each step is represented by a box with a green checkmark, indicating successful completion.

	NAME	AUTHOR	STATUS	LAST EDITED	PROJECT
<input type="checkbox"/>	Predictive Experiment - Mini ...	surdusabina	Finished	4/26/2019 3:13:51 PM	None
<input checked="" type="checkbox"/>	Predictii preturi masini	surdusabina	Finished	4/26/2019 11:28:46 AM	None

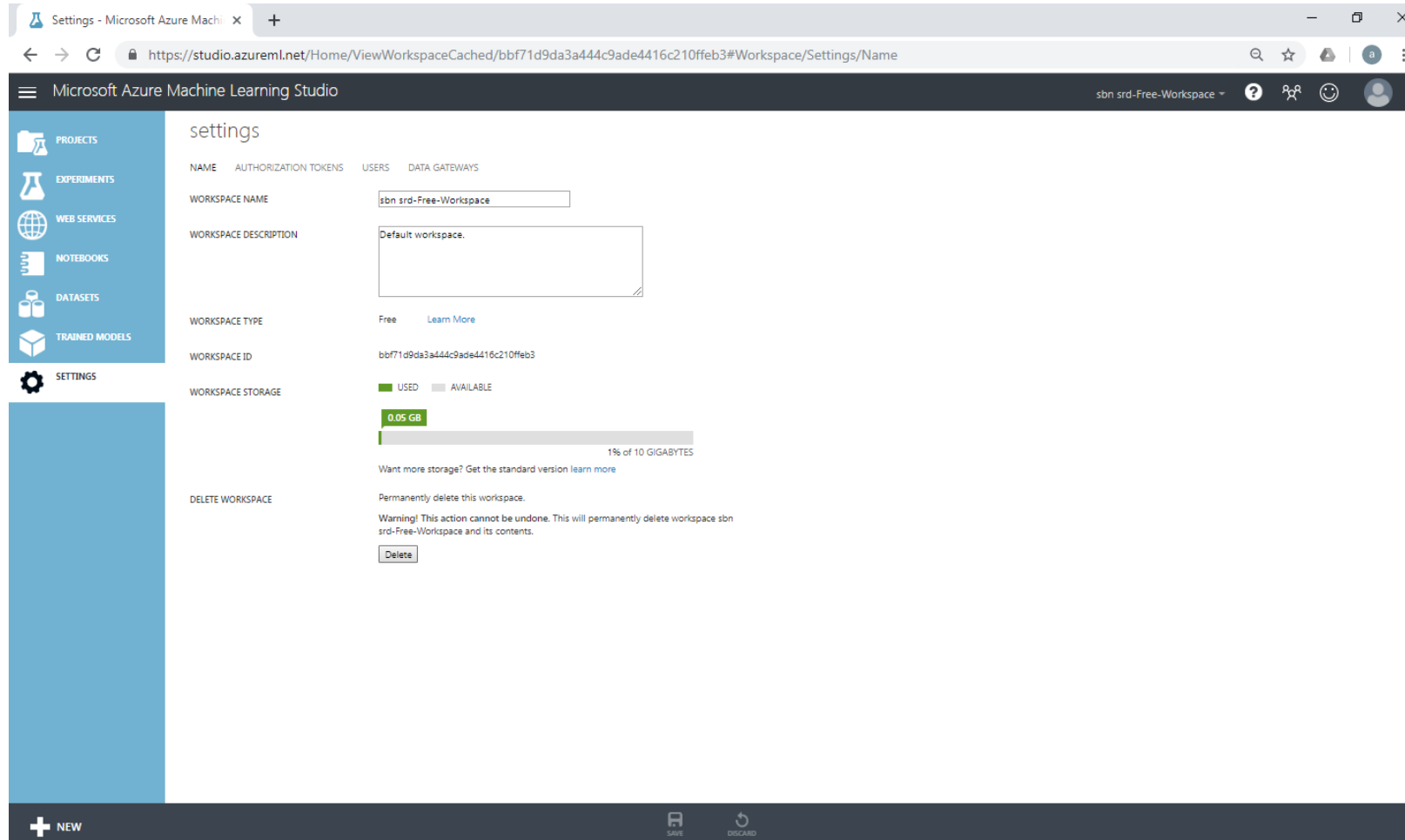
Workflow Diagram:

```
graph TD; A[Automobile price data (Raw)] --> B[Select Columns in Dataset]; B --> C[Clean Missing Data]; C --> D[Select Columns in Dataset]; D --> E[Split Data]; E --> F[Linear Regression]; E --> G[Train Model]; F --> G; G --> H[Score Model]; H --> I[Evaluate Model];
```

# Car Price Prediction

\* eliminate resources

- delete workspace: *Settings -> Delete*



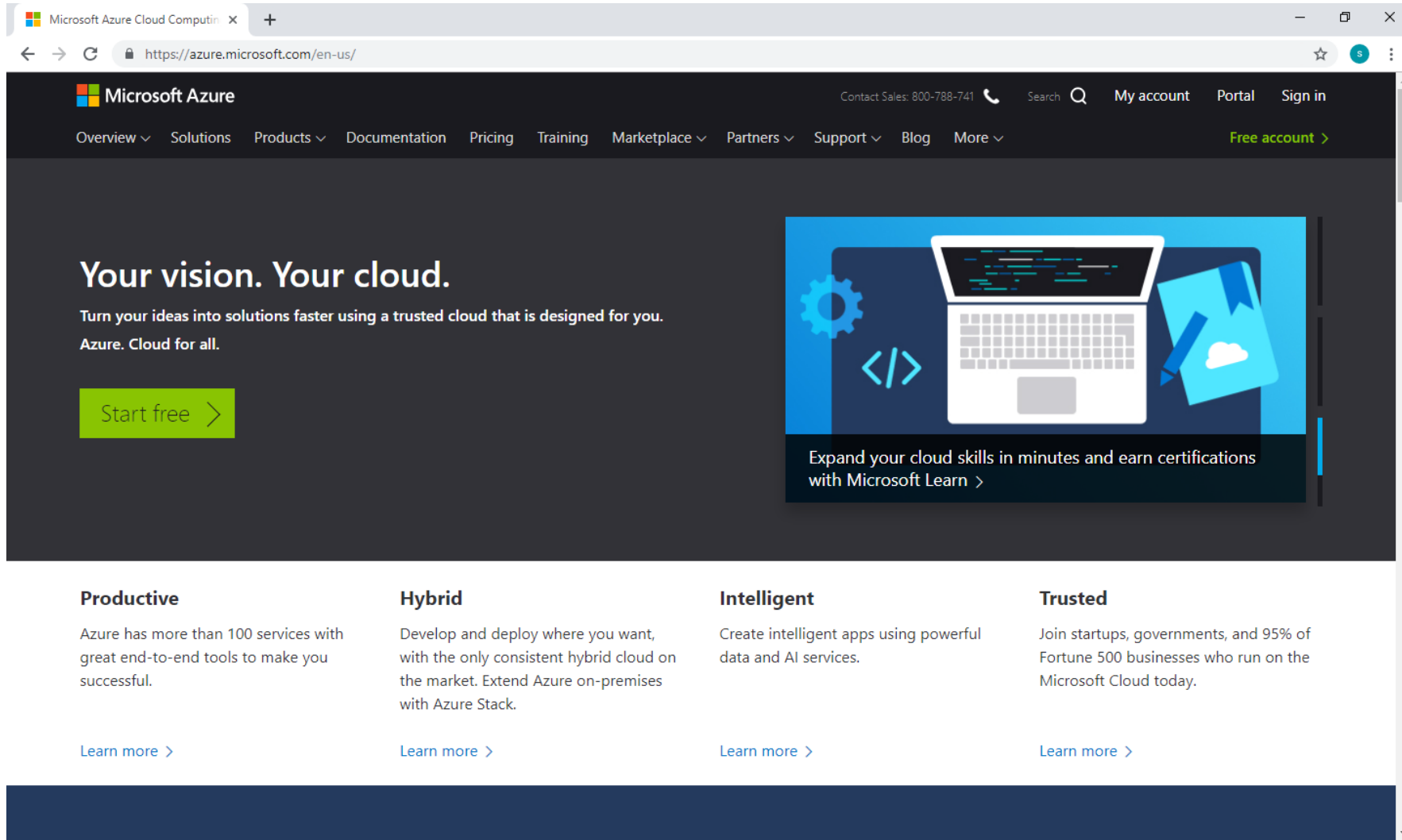


# Sentiment Analysis

## Azure Machine Learning and Azure Stream Analytics

# Azure Account – Azure for Students

- Azure portal
  - <https://portal.azure.com/#home>



# Sentiment Analysis: Azure Stream Analytics + Azure ML

- input
  - .csv file with tweets in Blob storage
- deploy sentiment analytics model as a web service
- create job
  - input – .csv file with tweets in Blob storage
  - output – another .csv file (same Blob storage)
    - for each tweet t – store the sentiment of t (*positive, neutral or negative*) and the probability that t is positive
  - Azure Machine Learning function – call web service as a function on each input tweet

# Input

\* *Create a resource -> Storage -> Storage account*

The screenshot shows the 'Create storage account' page in the Microsoft Azure portal. The left sidebar contains navigation links: 'Create a resource', 'Home', 'Dashboard', 'All services', 'FAVORITES', 'All resources', 'Resource groups', 'App Services', 'Function Apps', 'SQL databases', 'Azure Cosmos DB', 'Virtual machines', 'Load balancers', 'Storage accounts', 'Virtual networks', 'Azure Active Directory', 'Monitor', 'Advisor', 'Security Center', 'Cost Management + Billing', and 'Help + support'. The main content area is titled 'Create storage account' and includes tabs for 'Basics', 'Advanced', 'Tags', and 'Review + create'. The 'Basics' tab is active, showing a description of Azure Storage and a list of project details. The 'PROJECT DETAILS' section includes a 'Subscription' dropdown set to 'Pay-As-You-Go' and a 'Resource group' dropdown set to '(New) samlrg'. The 'INSTANCE DETAILS' section includes a 'Storage account name' field with 'samlsa', a 'Location' dropdown set to '(US) South Central US', a 'Performance' section with 'Standard' selected, an 'Account kind' dropdown set to 'StorageV2 (general purpose v2)', a 'Replication' dropdown set to 'Read-access geo-redundant storage (RA-GRS)', and an 'Access tier (default)' section with 'Hot' selected. At the bottom, there are buttons for 'Review + create', 'Previous', and 'Next : Advanced >'.

Create storage account - Microsoft

https://portal.azure.com/#create/Microsoft.StorageAccount-ARM

Microsoft Azure

Search resources, services, and docs

Home > New > Create storage account

### Create storage account

Basics Advanced Tags Review + create

Azure Storage is a Microsoft-managed service providing cloud storage that is highly available, secure, durable, scalable, and redundant. Azure Storage includes Azure Blobs (objects), Azure Data Lake Storage Gen2, Azure Files, Azure Queues, and Azure Tables. The cost of your storage account depends on the usage and the options you choose below. [Learn more](#)

**PROJECT DETAILS**

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

\* Subscription Pay-As-You-Go

\* Resource group (New) samlrg [Create new](#)

**INSTANCE DETAILS**

The default deployment model is Resource Manager, which supports the latest Azure features. You may choose to deploy using the classic deployment model instead. [Choose classic deployment model](#)

\* Storage account name samlsa

\* Location (US) South Central US

Performance ☒ Standard ☐ Premium

Account kind StorageV2 (general purpose v2)

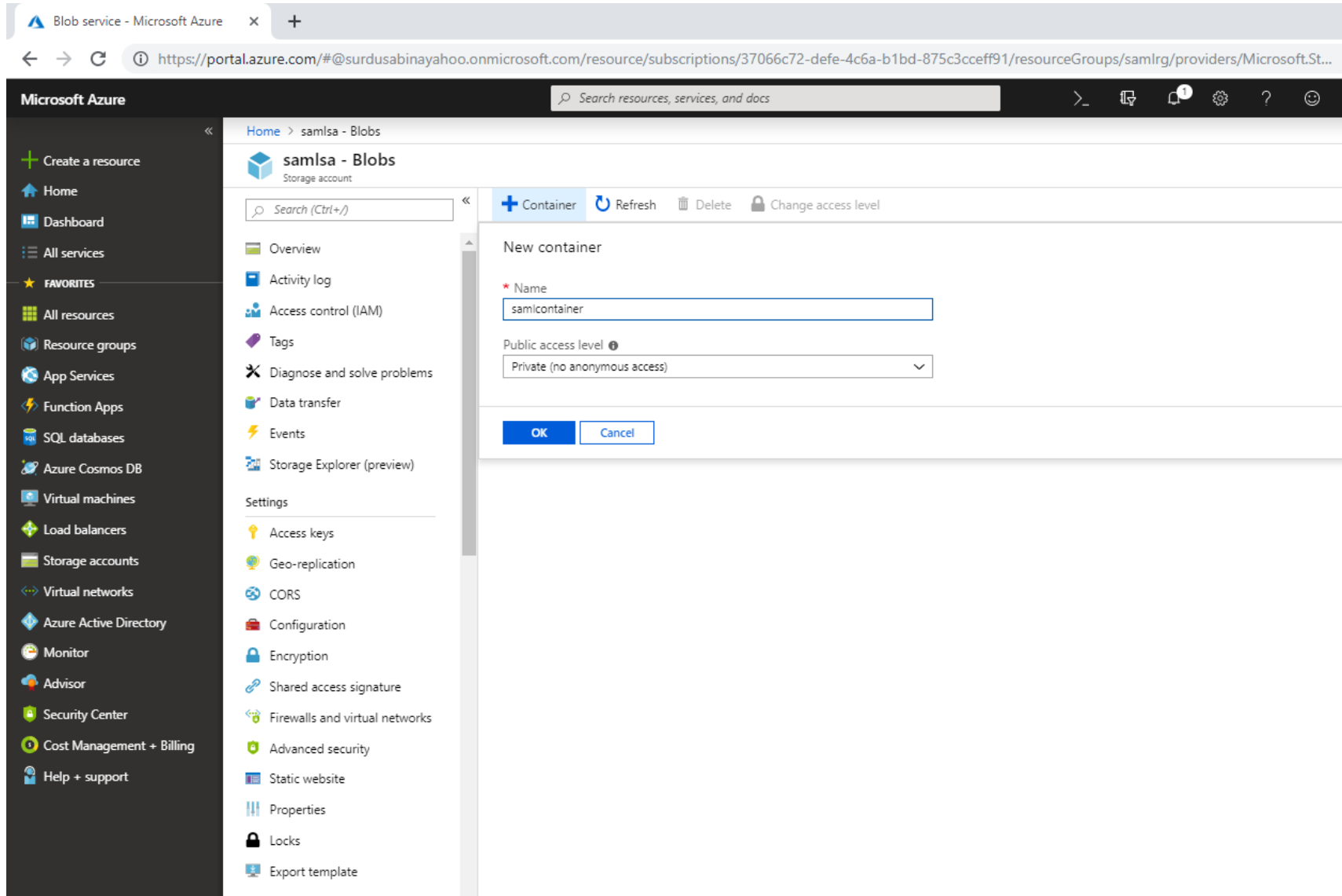
Replication Read-access geo-redundant storage (RA-GRS)

Access tier (default) ☐ Cool ☒ Hot

[Review + create](#) [Previous](#) [Next : Advanced >](#)

# Input

\* load file -> *Storage account -> Containers -> + Container*



# Input

\* load file -> *container* -> *Upload*

The screenshot shows the Microsoft Azure portal interface. The left sidebar contains navigation links: Create a resource, Home, Dashboard, All services, FAVORITES, All resources, Resource groups, App Services, Function Apps, SQL databases, Azure Cosmos DB, Virtual machines, Load balancers, Storage accounts, Virtual networks, Azure Active Directory, Monitor, Advisor, Security Center, Cost Management + Billing, and Help + support. The main content area displays the 'samlcontainer' storage container overview. The breadcrumb path is 'Home > samlsa - Blobs > samlcontainer'. The container name is 'samlcontainer' and it is of type 'Container'. The 'Overview' tab is selected in the left sidebar. The main content area shows the 'Upload' button, 'Refresh', 'Change access level', 'Delete', 'Acquire lease', 'Break lease', 'View snapshots', and 'Create snapshot' actions. The 'Authentication method' is 'Access key (Switch to Azure AD User Account)' and the 'Location' is 'samlcontainer'. A search bar for blobs is present with the placeholder text 'Search blobs by prefix (case-sensitive)'. Below the search bar is a table with columns: NAME, MODIFIED, ACCESS TIER, BLOB TYPE, and SIZE. The table currently displays 'No blobs found.'

# Input

\* load file -> choose file -> *Upload*

The screenshot displays the Azure portal interface for managing a storage container. On the left is a dark sidebar with navigation options: 'Create a resource', 'Home', 'Dashboard', 'All services', 'FAVORITES', 'All resources', 'Resource groups', 'App Services', 'Function Apps', 'SQL databases', 'Azure Cosmos DB', 'Virtual machines', 'Load balancers', 'Storage accounts', 'Virtual networks', 'Azure Active Directory', 'Monitor', 'Advisor', 'Security Center', 'Cost Management + Billing', and 'Help + support'.

The main content area shows the 'samlcontainer' container page. The breadcrumb path is 'Home > samlsa - Blobs > samlcontainer'. The container name 'samlcontainer' is displayed with the label 'Container'. Below this is a search bar and a set of action buttons: 'Upload', 'Refresh', 'Change access level', 'Delete', 'Acquire lease', 'Break lease', and 'View'. The 'Authentication method' is 'Access key' with a link to 'Switch to Azure AD User Account'. The 'Location' is 'samlcontainer'. A search bar for blobs is present with the placeholder 'Search blobs by prefix (case-sensitive)'. Below this is a table with columns 'NAME', 'MODIFIED', and 'ACCESS TIER', which currently shows 'No blobs found.'.

On the right, the 'Upload blob' pane is open for the 'samlcontainer/' path. It includes a 'Files' section with a text input containing 'input\_saml.csv' and a file selection icon. There is a checkbox for 'Overwrite if files already exist'. An 'Advanced' section is expanded, showing 'Authentication type' with options 'Azure AD user account' and 'Account key' (selected). The 'Blob type' is set to 'Block blob'. A checkbox 'Upload .vhd files as page blobs (recommended)' is checked. The 'Block size' is set to '4 MB'. There is an 'Upload to folder' text input. At the bottom of the pane is a blue 'Upload' button.

# Input

\* input\_saml.csv

*Text*

*They love ice cream*

*The sky is blue*

*Machine Learning is cool*

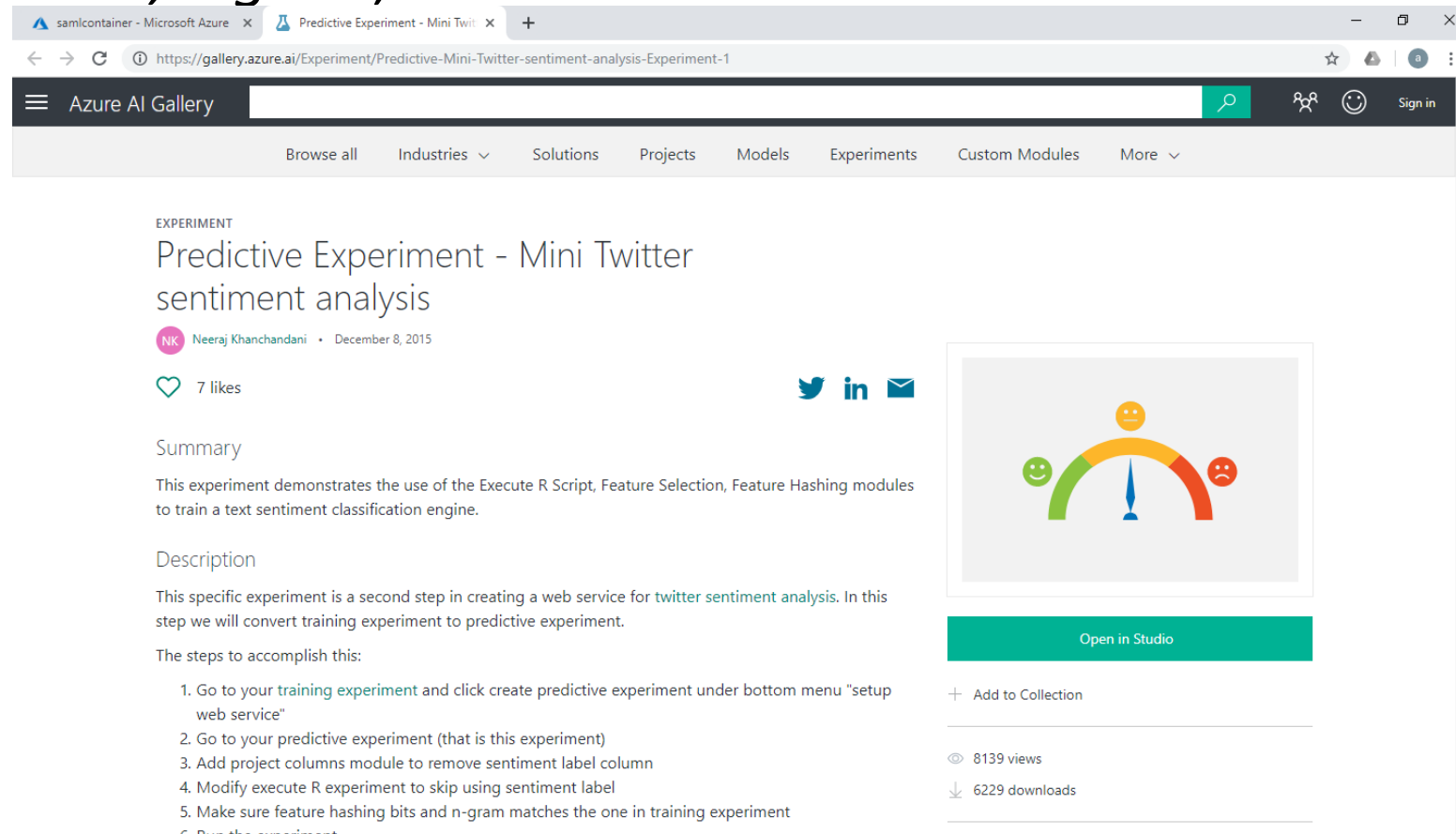
*Machine Learning is not cool*



# Machine Learning Model -> Web Service

## \* sentiment analytics model (Azure AI Gallery)

- <https://gallery.azure.ai/Experiment/Predictive-Mini-Twitter-sentiment-analysis-Experiment-1>
- *Open in Studio, Sign in, select location*



The screenshot shows a web browser window displaying the Azure AI Gallery page for an experiment titled "Predictive Experiment - Mini Twitter sentiment analysis". The page header includes the Azure AI Gallery logo and navigation links: "Browse all", "Industries", "Solutions", "Projects", "Models", "Experiments", "Custom Modules", and "More". The experiment details section shows the title, author "Neeraj Khanchandani" (with a profile icon), date "December 8, 2015", and "7 likes". There are social media share icons for Twitter, LinkedIn, and Email. The "Summary" section states: "This experiment demonstrates the use of the Execute R Script, Feature Selection, Feature Hashing modules to train a text sentiment classification engine." The "Description" section explains: "This specific experiment is a second step in creating a web service for twitter sentiment analysis. In this step we will convert training experiment to predictive experiment. The steps to accomplish this:" followed by a numbered list of six steps. To the right of the text is a graphic of a sentiment gauge with a needle pointing to the neutral position, flanked by a green smiley face and a red frowny face. Below the graphic is a prominent green "Open in Studio" button. At the bottom right, there are links for "Add to Collection", "8139 views", and "6229 downloads".

EXPERIMENT

### Predictive Experiment - Mini Twitter sentiment analysis

NK Neeraj Khanchandani • December 8, 2015

7 likes

Summary

This experiment demonstrates the use of the Execute R Script, Feature Selection, Feature Hashing modules to train a text sentiment classification engine.

Description

This specific experiment is a second step in creating a web service for twitter sentiment analysis. In this step we will convert training experiment to predictive experiment.

The steps to accomplish this:

1. Go to your training experiment and click create predictive experiment under bottom menu "setup web service"
2. Go to your predictive experiment (that is this experiment)
3. Add project columns module to remove sentiment label column
4. Modify execute R experiment to skip using sentiment label
5. Make sure feature hashing bits and n-gram matches the one in training experiment
6. Run the experiment

Open in Studio

+ Add to Collection

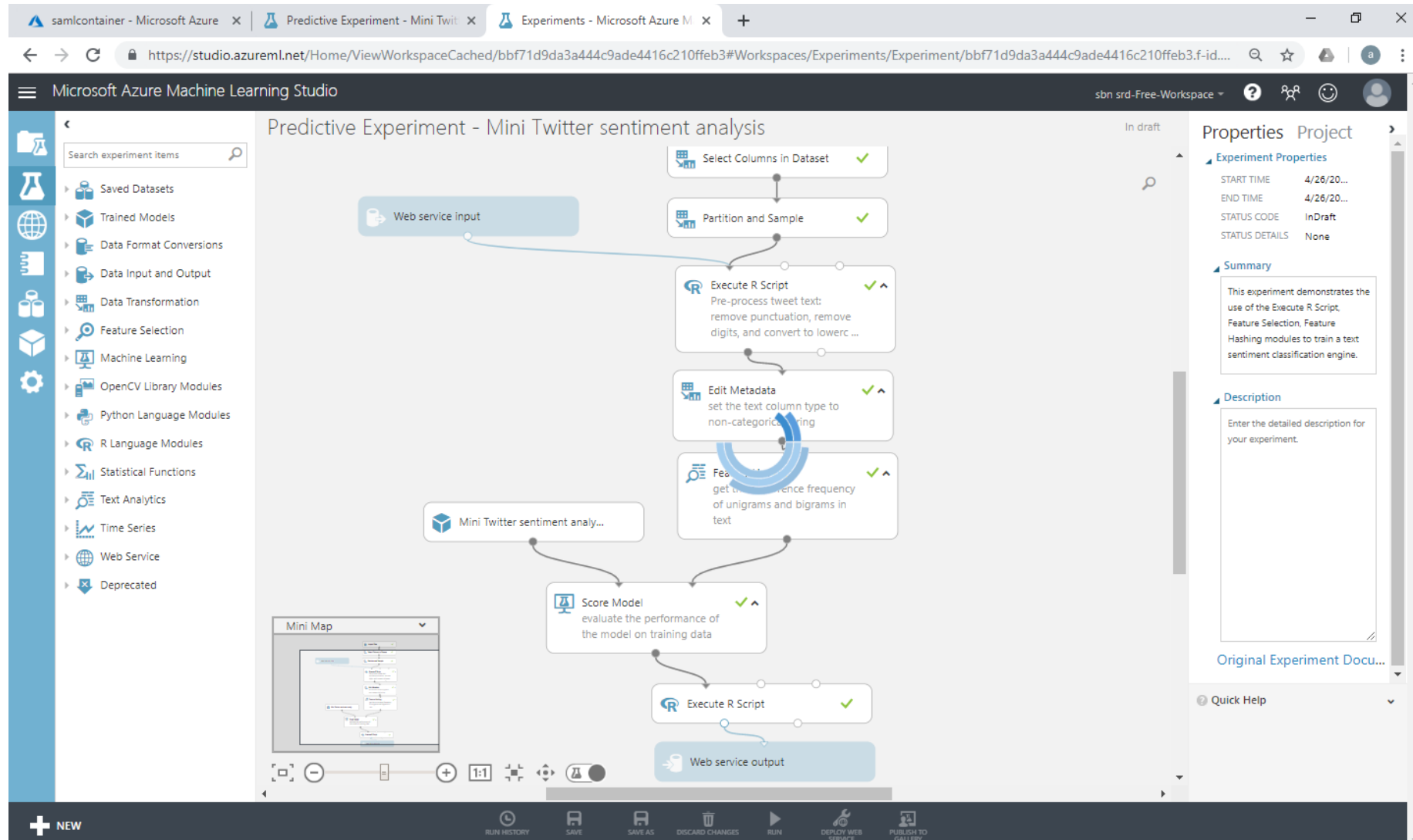
8139 views

6229 downloads

# Machine Learning Model -> Web Service

\* sentiment analytics model

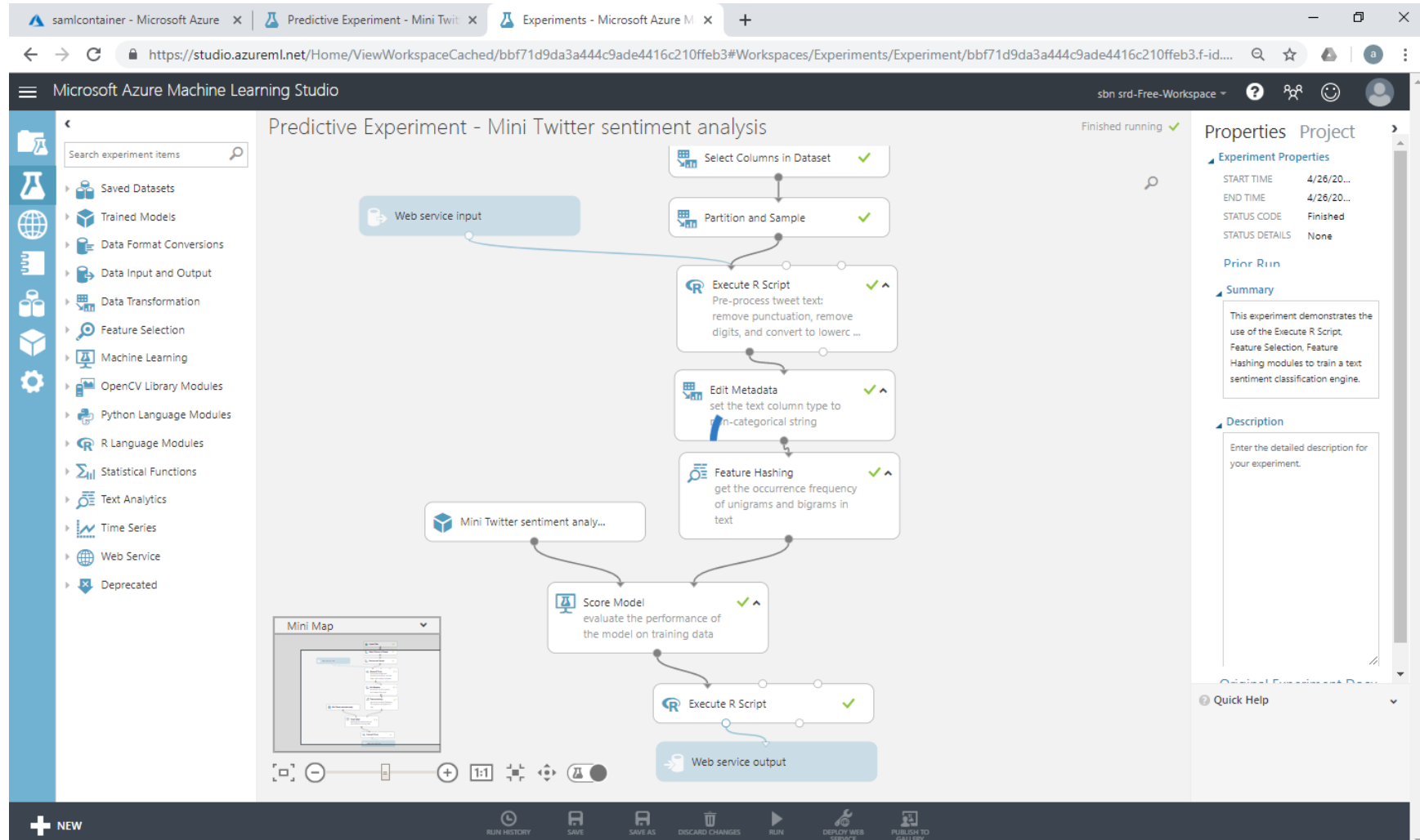
- Run



# Machine Learning Model -> Web Service

\* sentiment analytics model

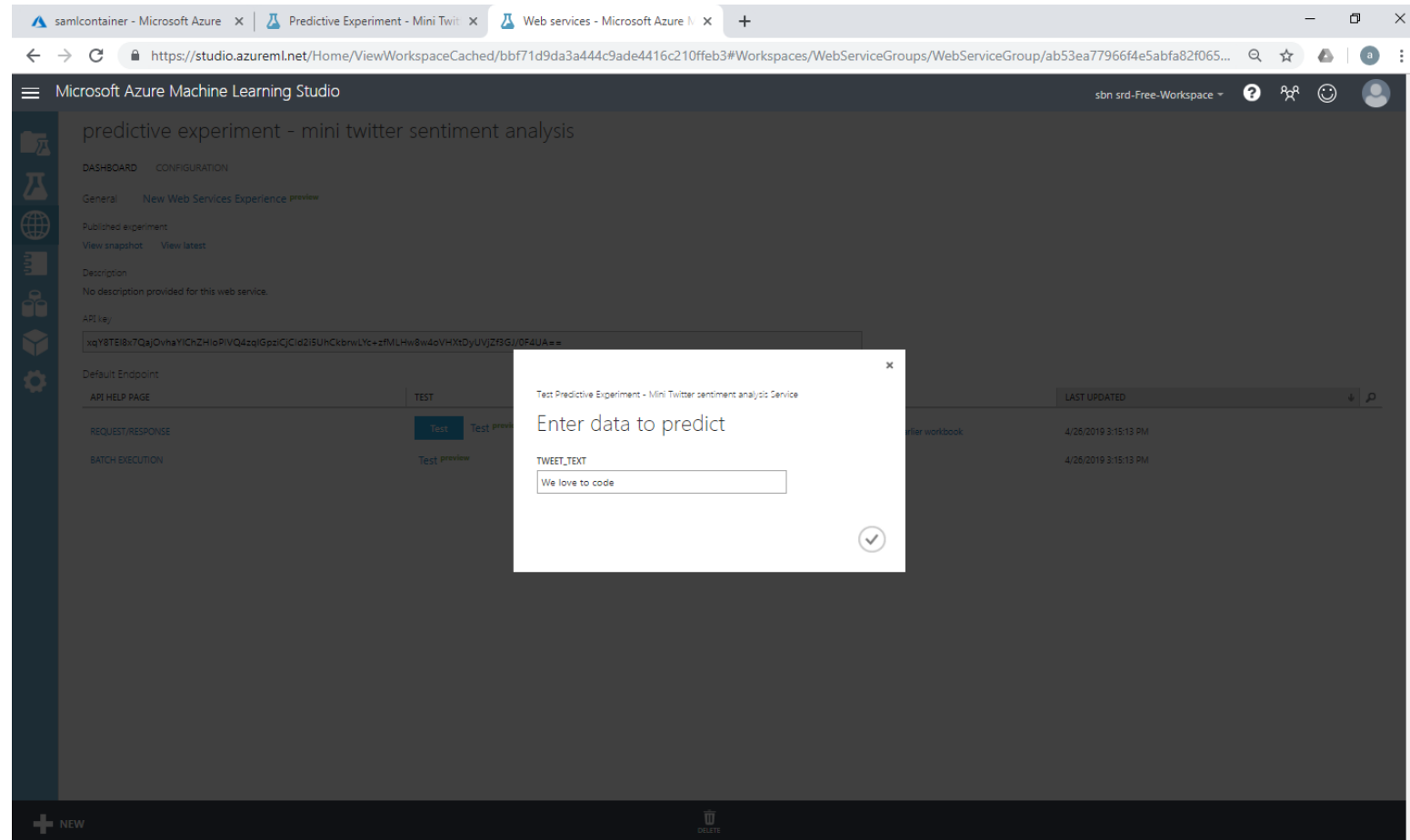
- *Deploy Web Service*



# Machine Learning Model -> Web Service

\* sentiment analytics model

- testing
  - Test -> input text



# Machine Learning Model -> Web Service

## \* sentiment analytics model

- testing
  - result
- download Excel workbook (*Excel 2010 or earlier workbook*) – it contains the API key and the web service URL

The screenshot displays the Microsoft Azure Machine Learning Studio interface. The browser address bar shows the URL: <https://studio.azureml.net/Home/ViewWorkspaceCached/bbf71d9da3a444c9ade4416c210feb3#Workspaces/WebServiceGroups/WebServiceGroup/ab53ea77966f4e5abfa82f065...>

The main header of the studio is "Microsoft Azure Machine Learning Studio" with a user profile icon and the text "sbn srd-Free-Workspace".

The left sidebar contains navigation icons for Dashboard, Configuration, Published experiment, View snapshot, View latest, Description, No description provided for this web service, API key, Default Endpoint, API HELP PAGE, TEST, APPS, and LAST UPDATED.

The main content area shows the configuration for the "predictive experiment - mini twitter sentiment analysis". It includes a "General" tab, a "New Web Services Experience" link, and a "Description" section. The "API key" field is populated with a long alphanumeric string: `xqY8TEl8x7QajOvhaYlChZHoPiVQ4zqGpzCjCId2ISUHCKbrwLYc-zfMLHw@w4oVHXtDyUVjZf3Gj/0F4UA==`.

Below the API key, there is a table with columns: REQUEST/RESPONSE, BATCH EXECUTION, TEST, APPS, and LAST UPDATED. The table contains two rows of data.

REQUEST/RESPONSE	BATCH EXECUTION	TEST	APPS	LAST UPDATED
Test	Test preview	Test	Excel 2013 or later   Excel 2010 or earlier workbook	4/26/2019 3:15:13 PM
Test	Test preview	Test	Excel 2013 or later workbook	4/26/2019 3:15:13 PM

At the bottom of the interface, a status bar shows a green checkmark and the message: "Predictive Experiment - Mini Twitter sentiment analysis" test returned [{"positive":0.704714417457581}].

# Job

*\* Create a resource -> Analytics -> Stream Analytics Job*

The screenshot displays the Microsoft Azure portal interface for creating a new Stream Analytics job. The browser address bar shows the URL `https://portal.azure.com/#create/Microsoft.StreamAnalyticsJob`. The left-hand navigation pane lists various Azure services, with 'Create a resource' at the top. The main content area is titled 'New Stream Analytics job' and contains the following configuration fields:

- Job name:** A text input field containing 'samjob' with a green checkmark indicating it is valid.
- Subscription:** A dropdown menu set to 'Pay-As-You-Go'.
- Resource group:** A dropdown menu set to 'samrg', with a 'Create new' link below it.
- Location:** A dropdown menu set to 'South Central US'.
- Hosting environment:** Two buttons, 'Cloud' (selected) and 'Edge'.
- Streaming units (1 to 192):** A slider control set to '1'.

At the bottom of the configuration panel, there is a blue 'Create' button and a link for 'Automation options'. The right side of the page is a large, solid blue area.

# Job

\* configure the input – previously loaded .csv file

The screenshot displays the Microsoft Azure portal interface. The left sidebar contains navigation options such as 'Create a resource', 'Home', 'Dashboard', 'All services', and 'FAVORITES'. The main content area is titled 'samIjob - Inputs' and shows a list of resources with 'samIjob' selected. A dropdown menu is open, showing options to 'Add stream input' and 'Add reference input'. The 'Add stream input' dropdown is expanded, showing 'Event Hub', 'IoT Hub', and 'Blob storage'. The 'Blob storage' option is highlighted, and a 'Blob storage' input type is visible in the table below.

SOURCE TYPE	SOURCE
Blob storage	

# Job

\* configure the input

- *Event serialization format -> CSV*

The screenshot displays the Azure portal interface for configuring a Stream Analytics job. The left sidebar shows the navigation menu with 'All resources' selected. The main area is titled 'samjob - Inputs' and shows a table with one input named 'inputdata'. The 'Blob storage' configuration panel is open on the right, showing settings for the 'inputdata' input. The 'Event serialization format' is set to 'CSV'.

**Left Sidebar (Navigation):**

- Create a resource
- Home
- Dashboard
- All services
- FAVORITES
- All resources
- Resource groups
- App Services
- Function Apps
- SQL databases
- Azure Cosmos DB
- Virtual machines
- Load balancers
- Storage accounts
- Virtual networks
- Azure Active Directory
- Monitor
- Advisor
- Security Center
- Cost Management + Billing
- Help + support

**Top Bar (Breadcrumbs):** Home > All resources > samjob - Inputs

**Left Panel (All resources):**

- Filter by name...
- NAME
- samljob
- samlisa

**Right Panel (samjob - Inputs):**

- Overview
- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Settings
- Locks
- Job topology
- Inputs
- Functions
- Query
- Outputs
- Configure
- Storage account settings
- Scale
- Locale
- Event ordering
- Error policy
- Compatibility level
- Managed Identity
- General

**Table (Inputs):**

NAME	SOURCE TYPE
inputdata	

**Right Panel (Blob storage configuration):**

- Input alias: inputdata
- Provide Blob storage settings manually (selected)
- Subscription: Pay-As-You-Go
- Storage account: samisa
- Storage account key: [Redacted]
- Container: samcontainer
- Path pattern: [Redacted]
- Date format: YYYY/MM/DD
- Time format: HH
- Partitions: [Redacted]
- Event serialization format: CSV

**Buttons:** Add stream input, Add reference input, Save