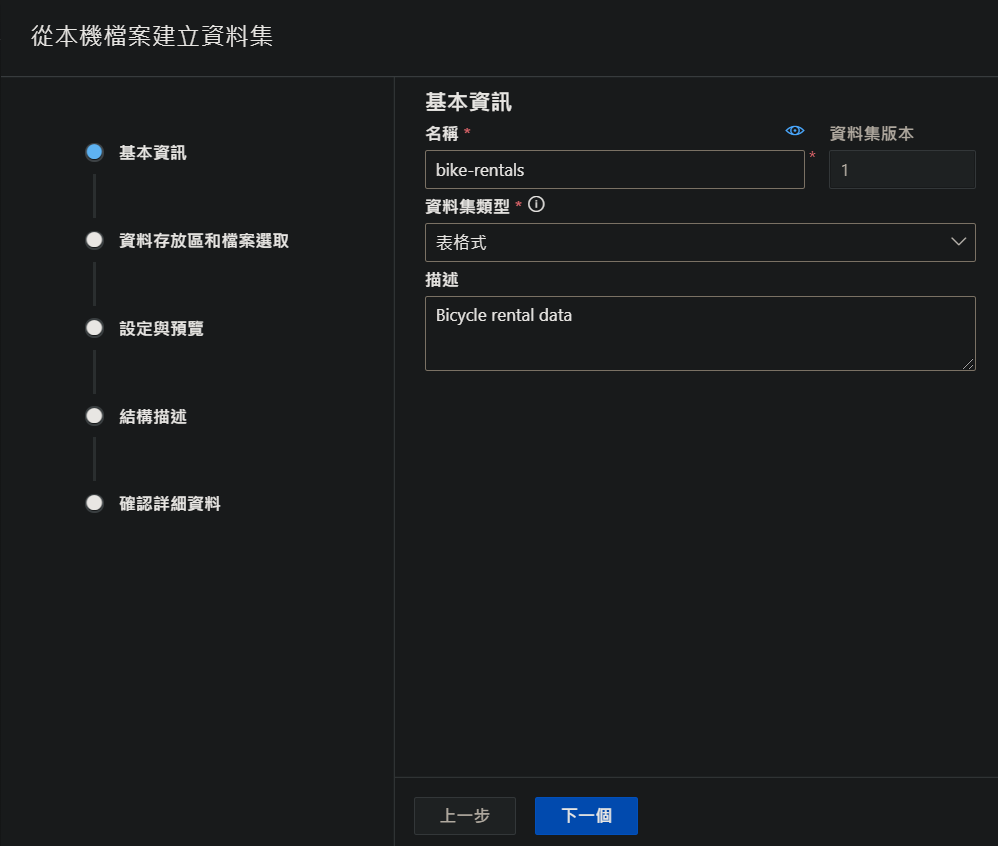
210622(Tues.)

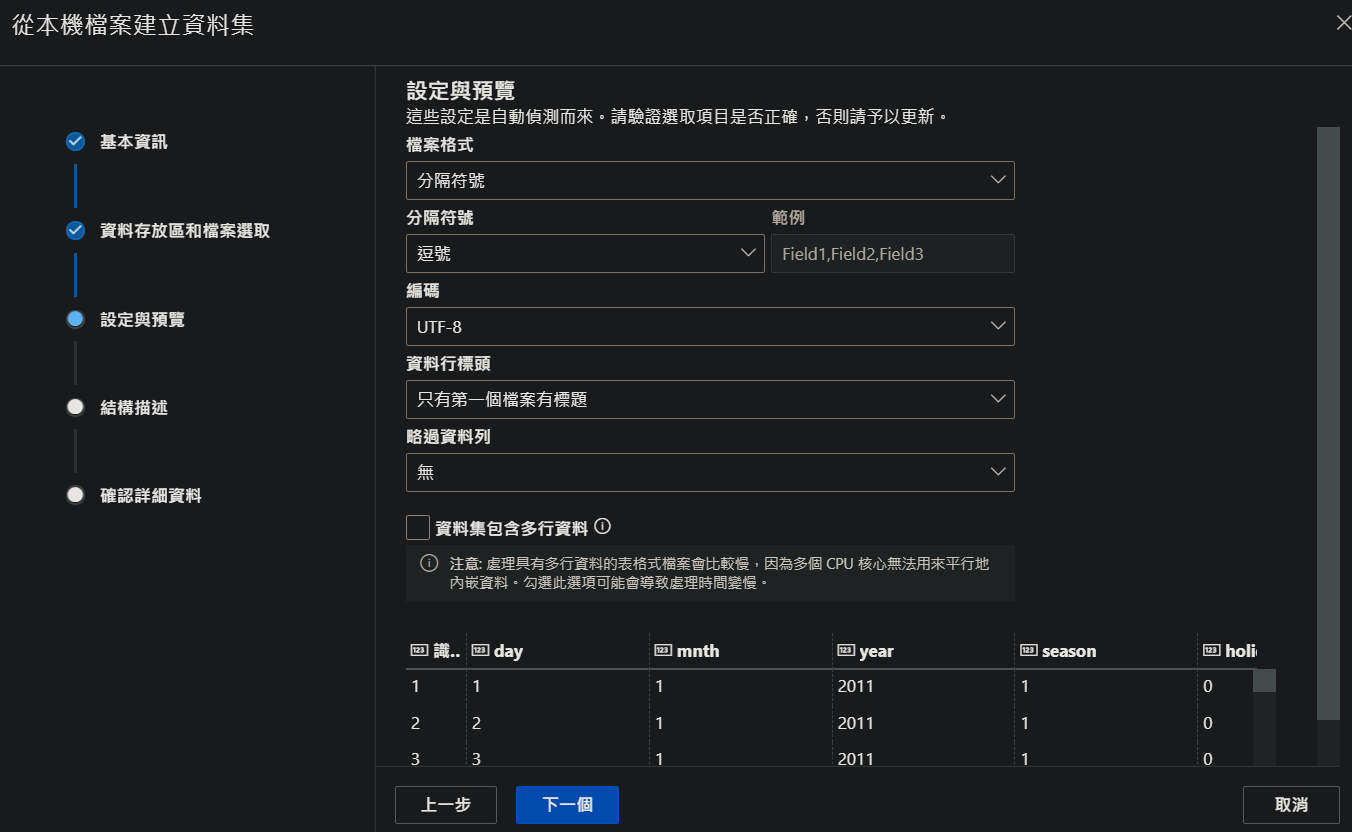
DP100-2

<https://docs.microsoft.com/zh-tw/azure/machine-learning/quickstart-create-resources>

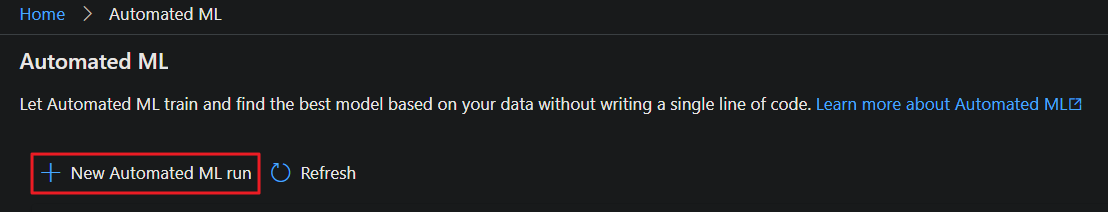


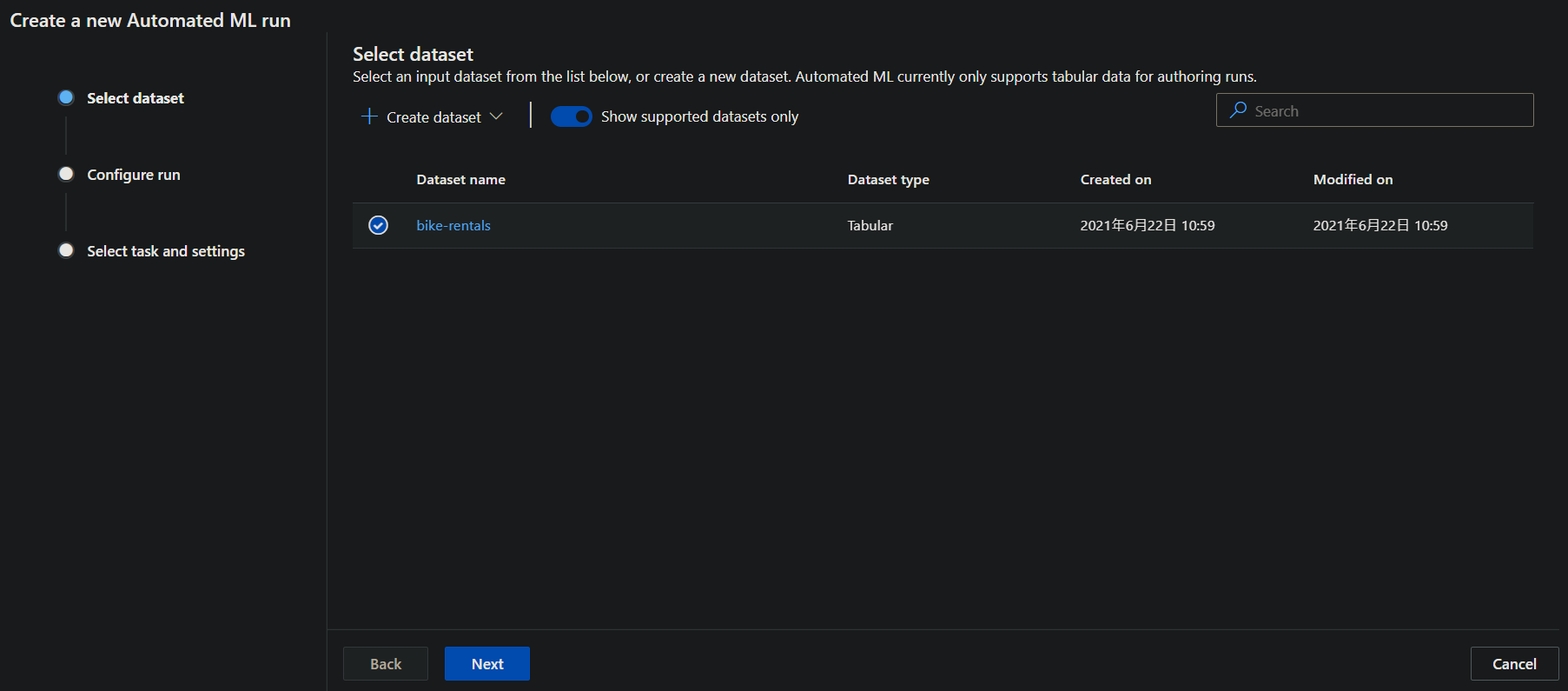


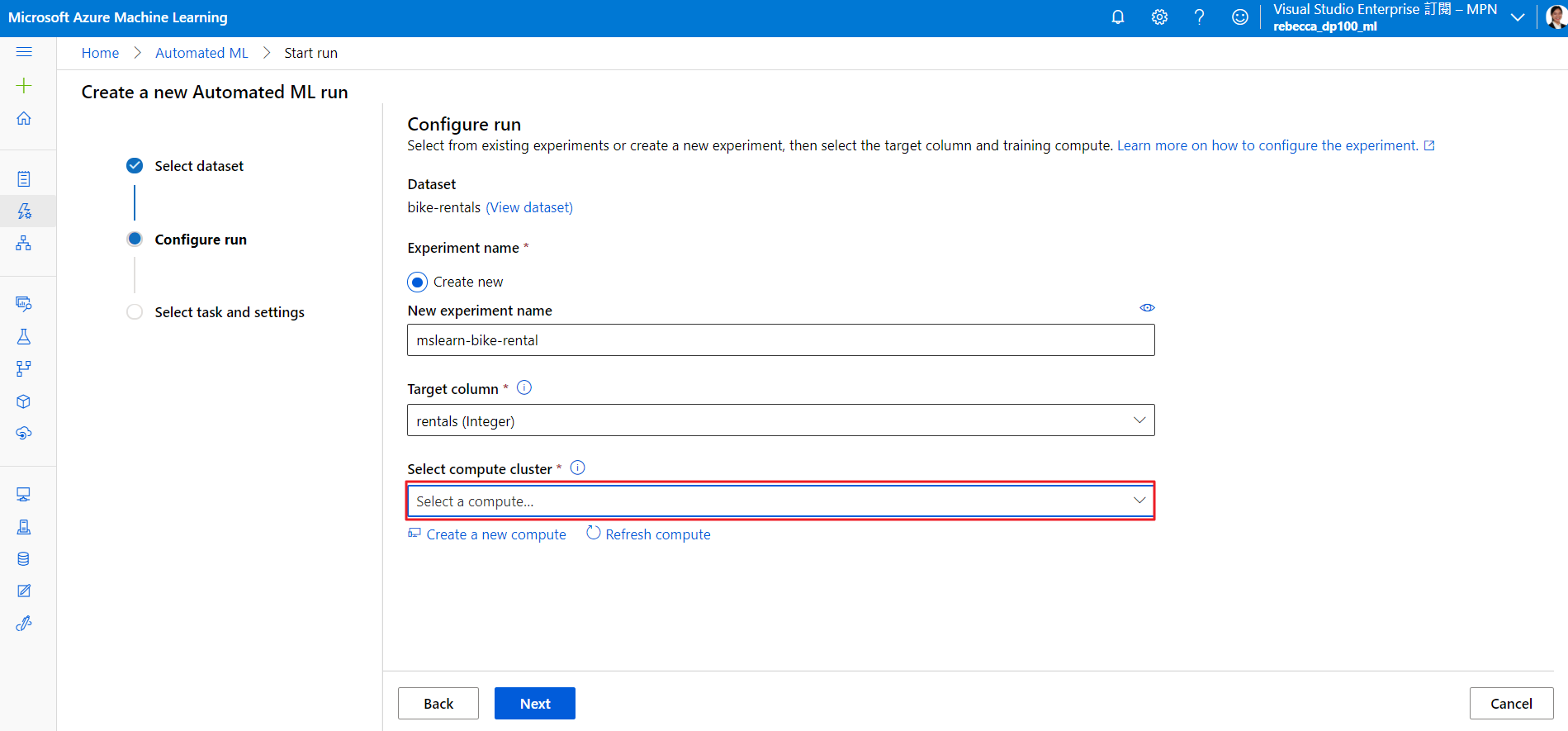


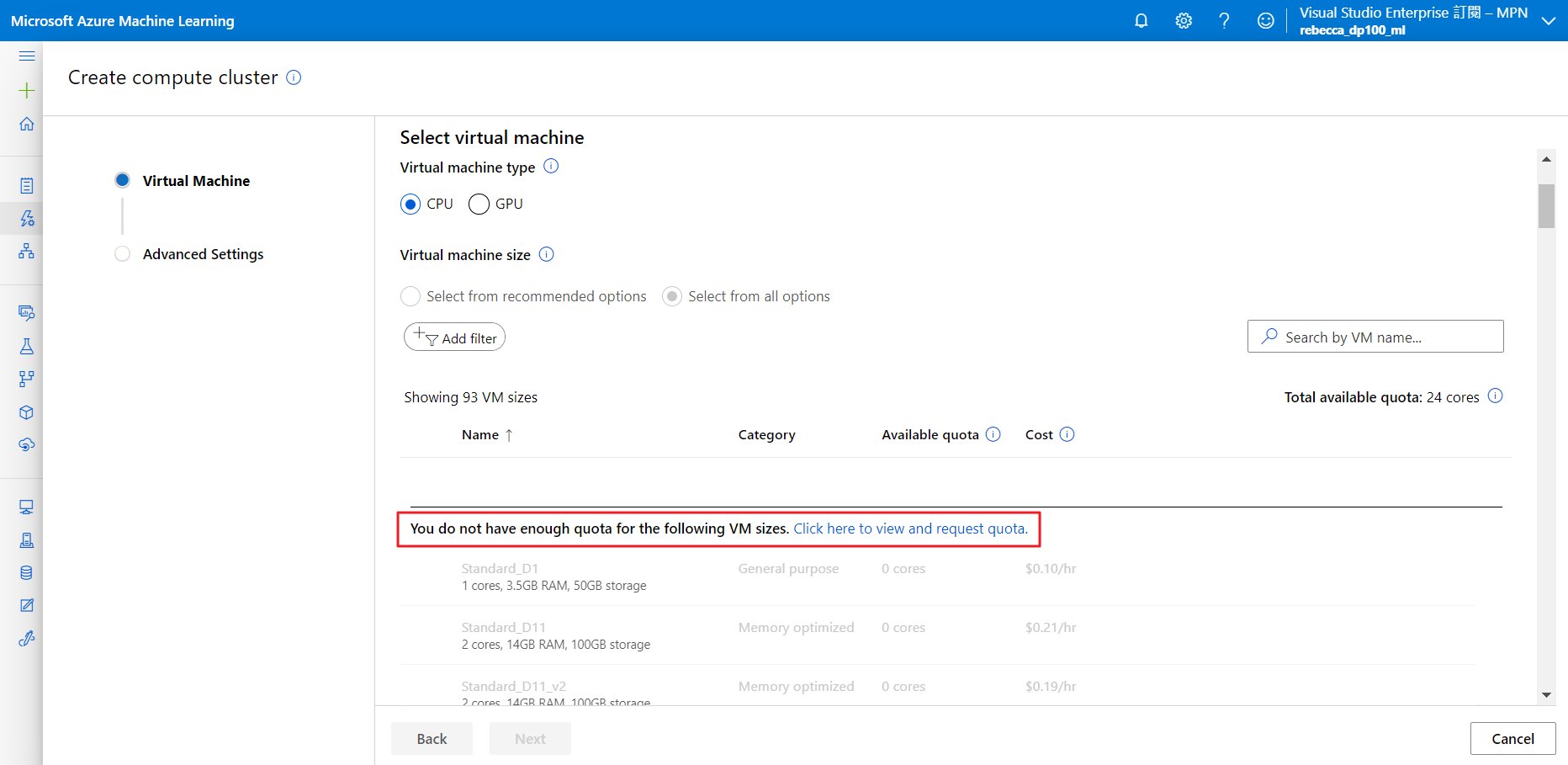








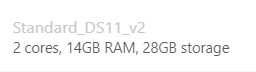






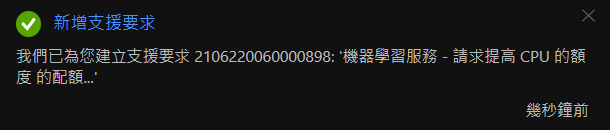












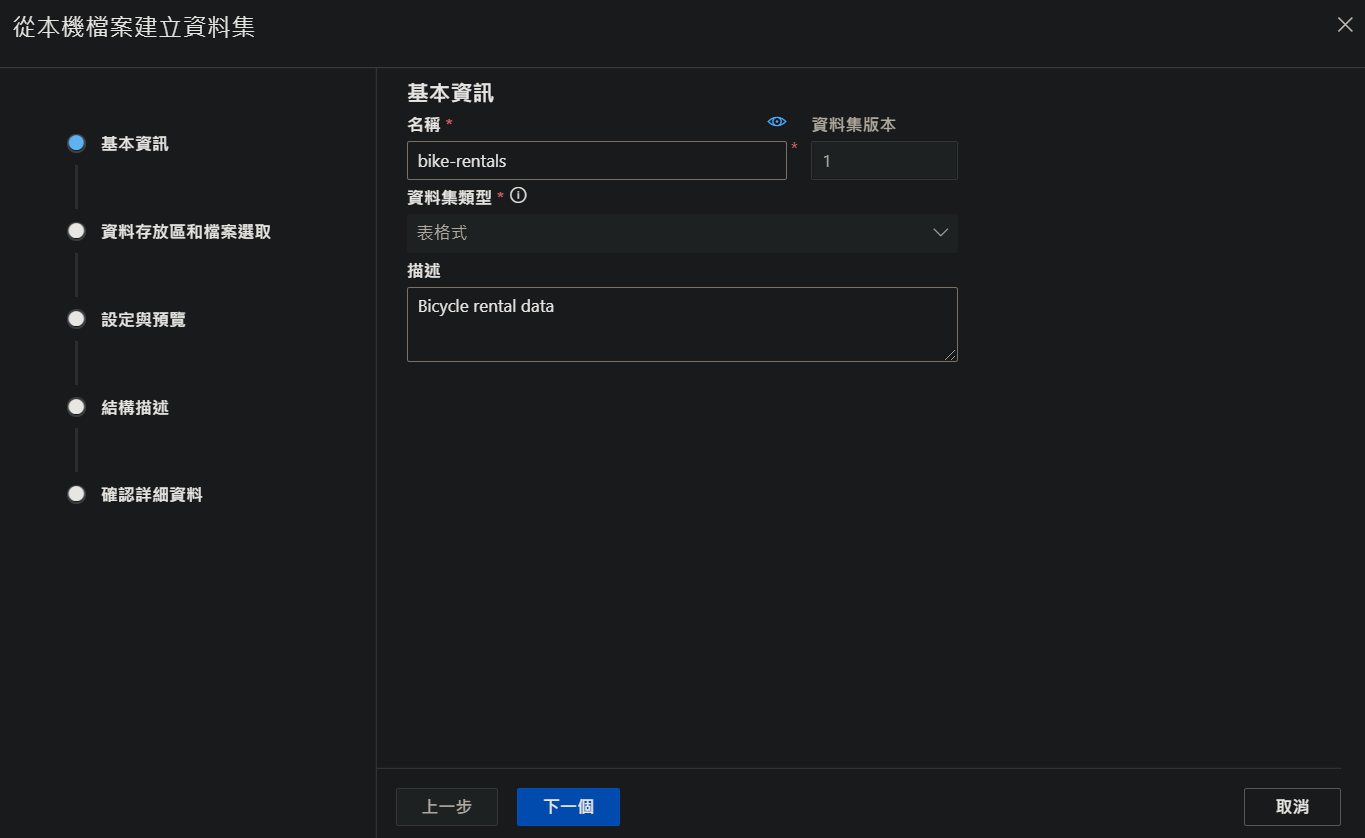
210625

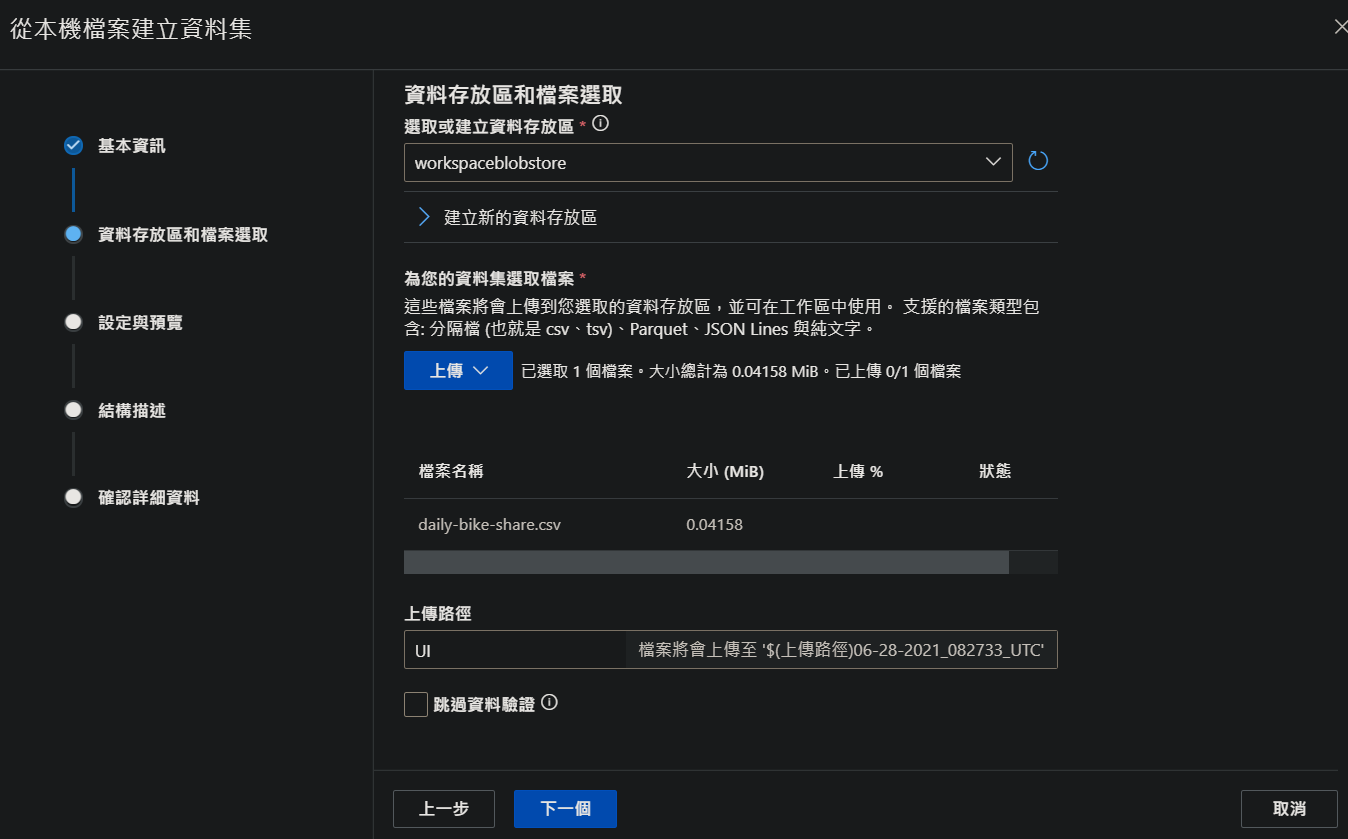


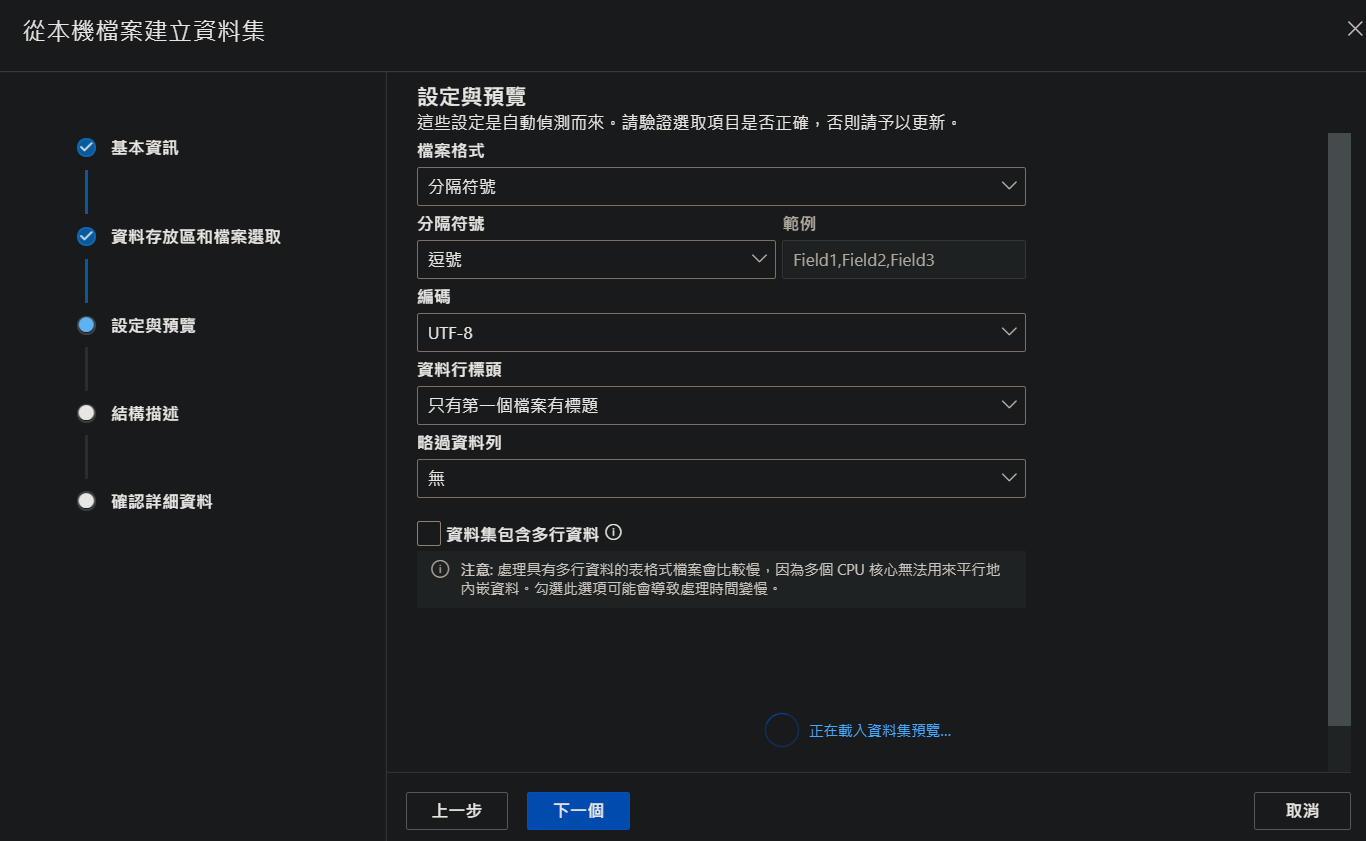




210628(Mon.)

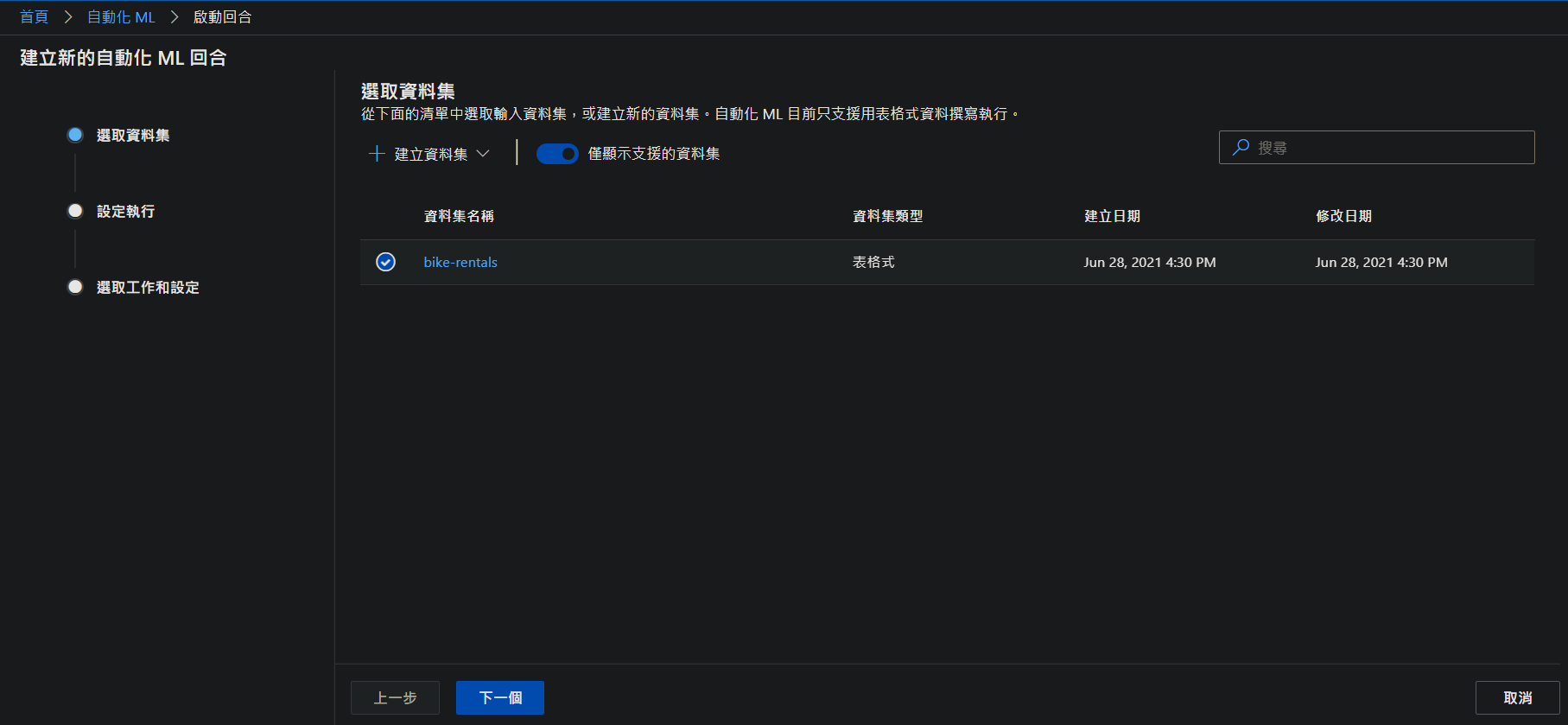














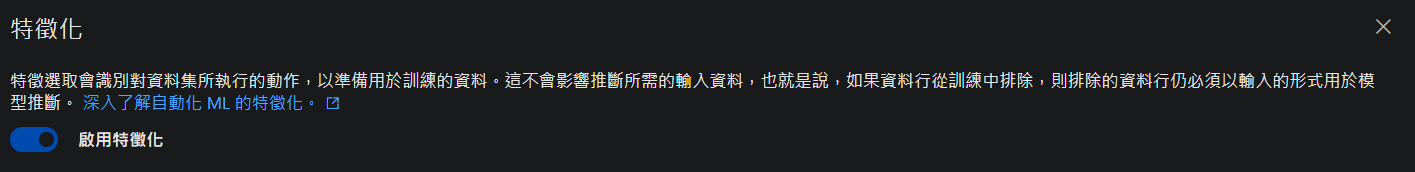




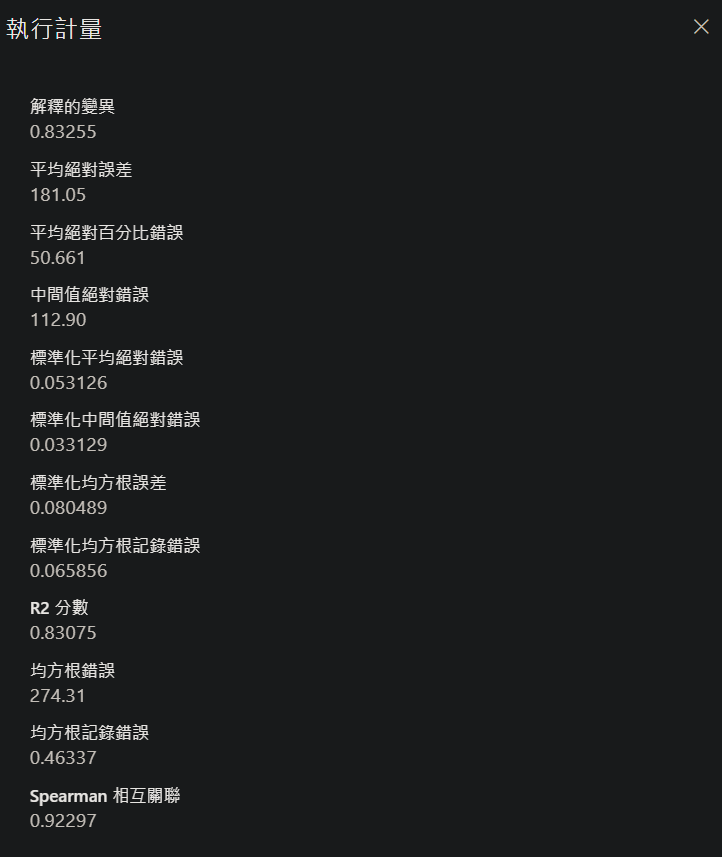


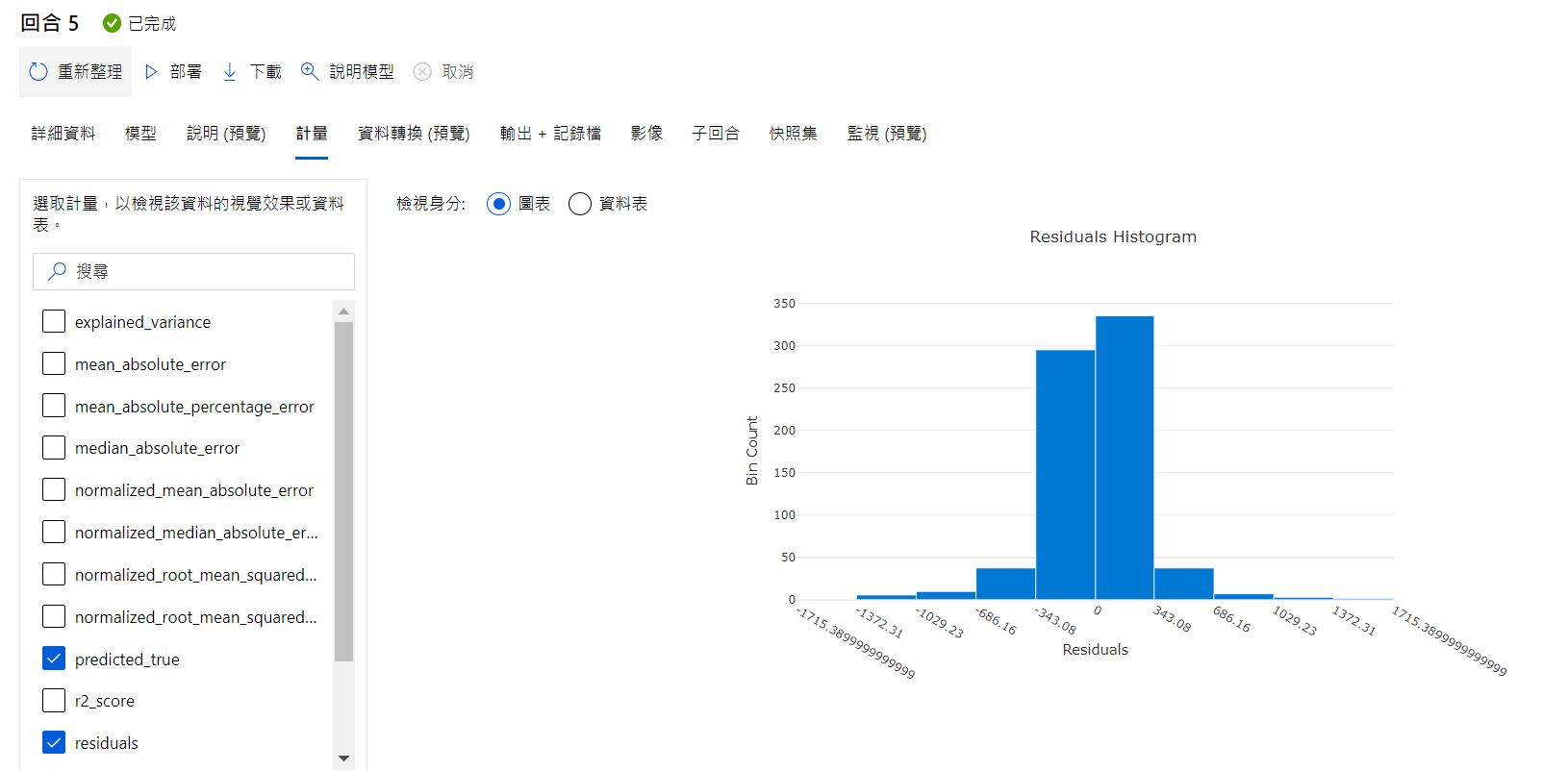


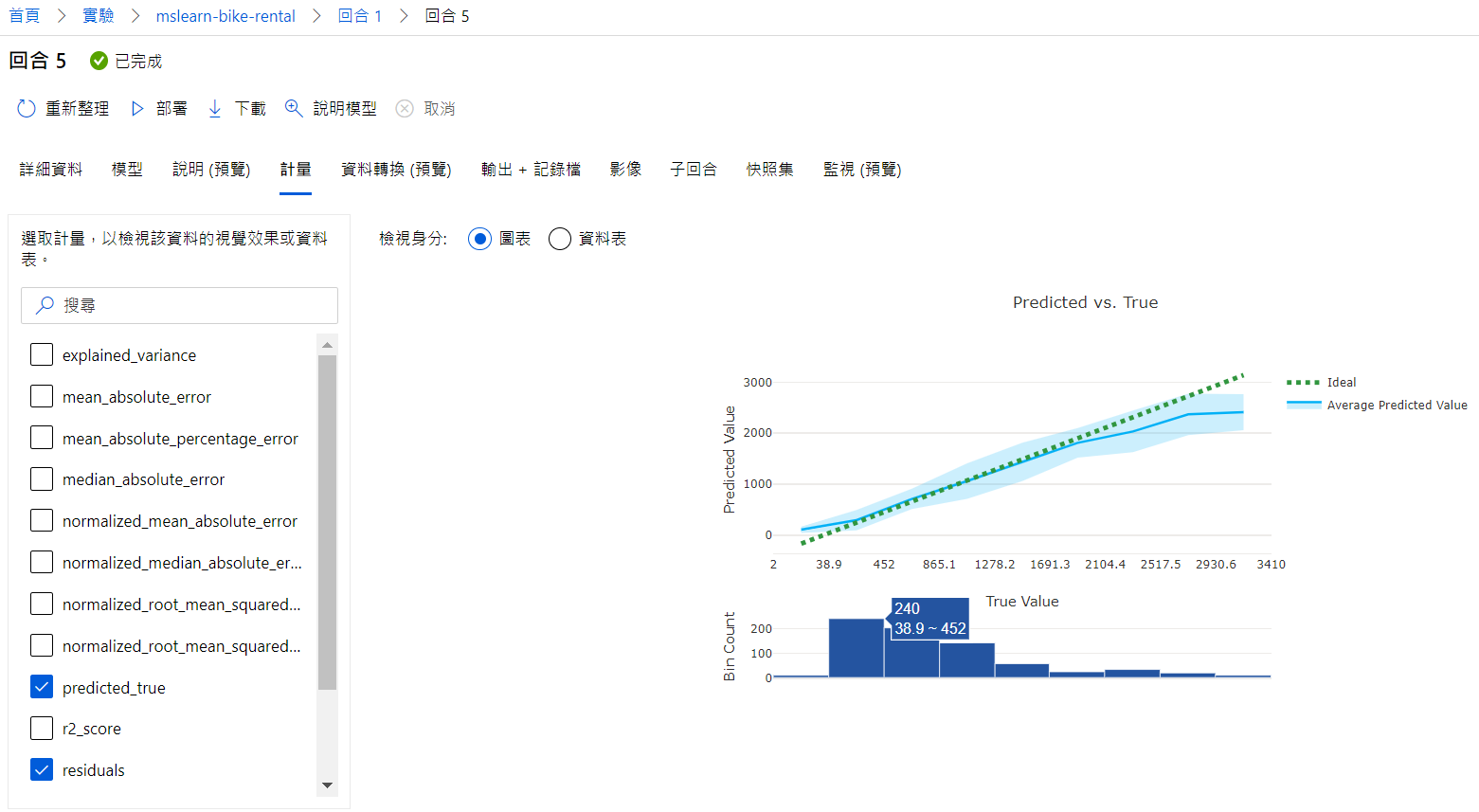


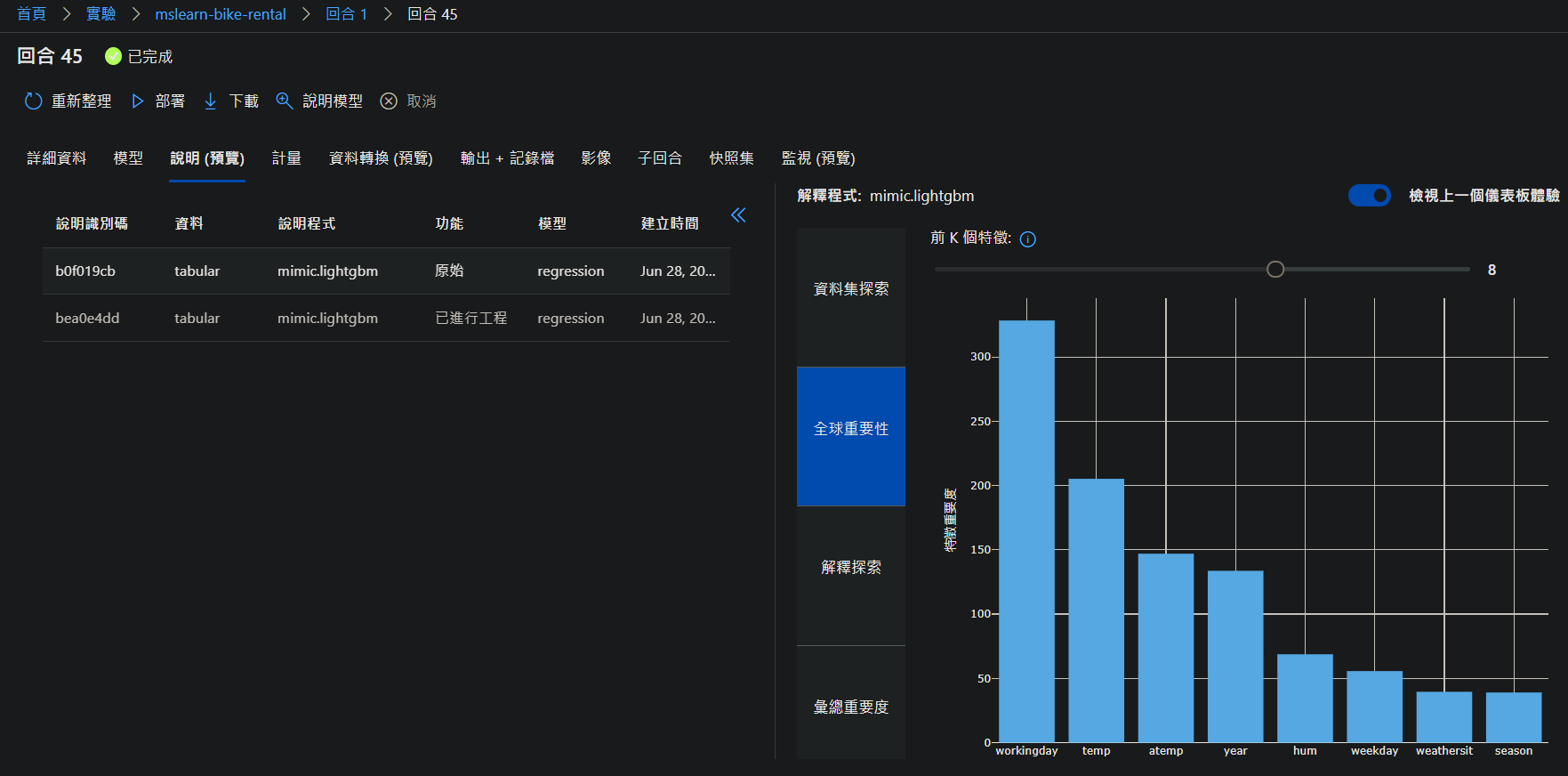


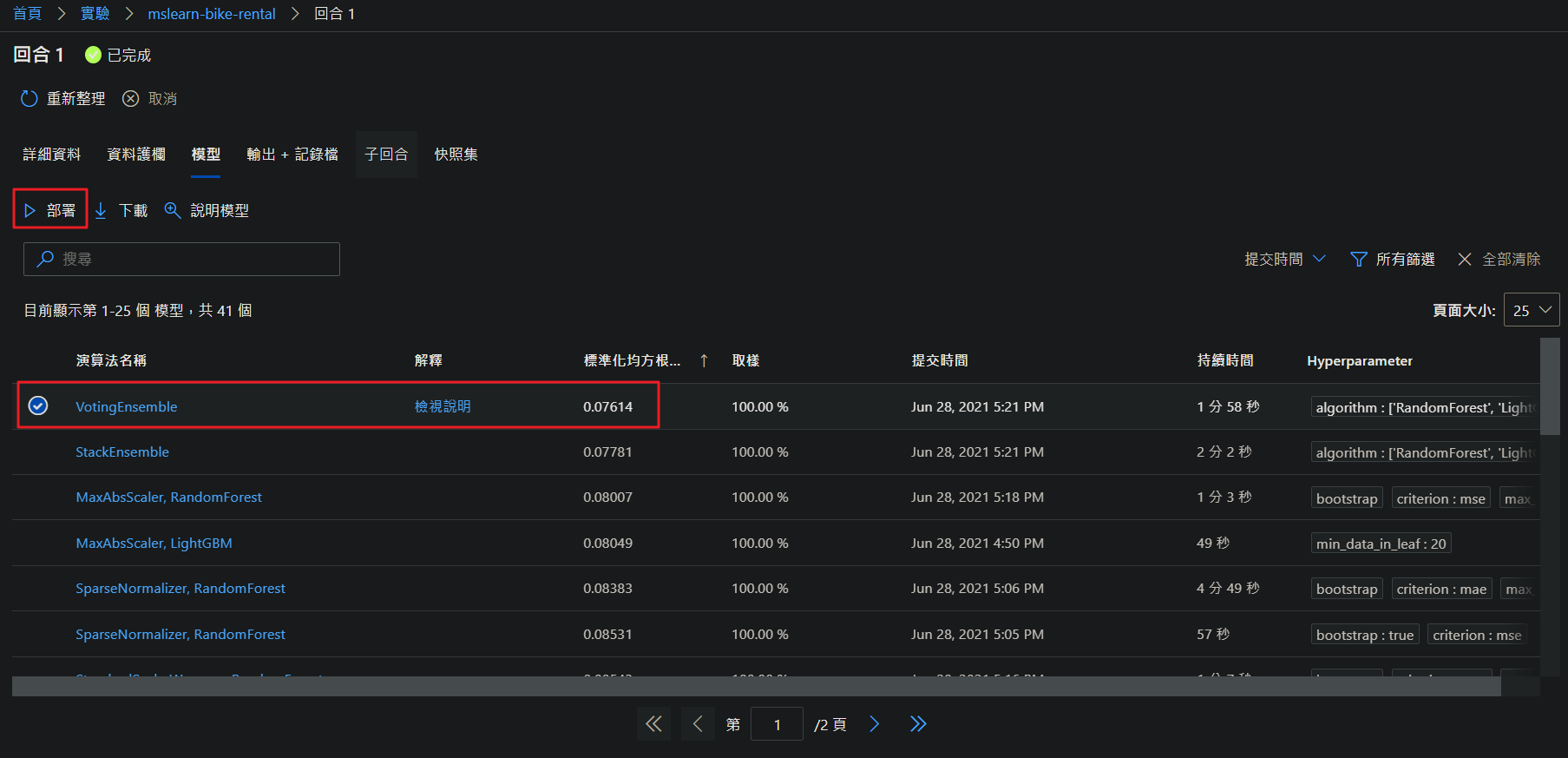




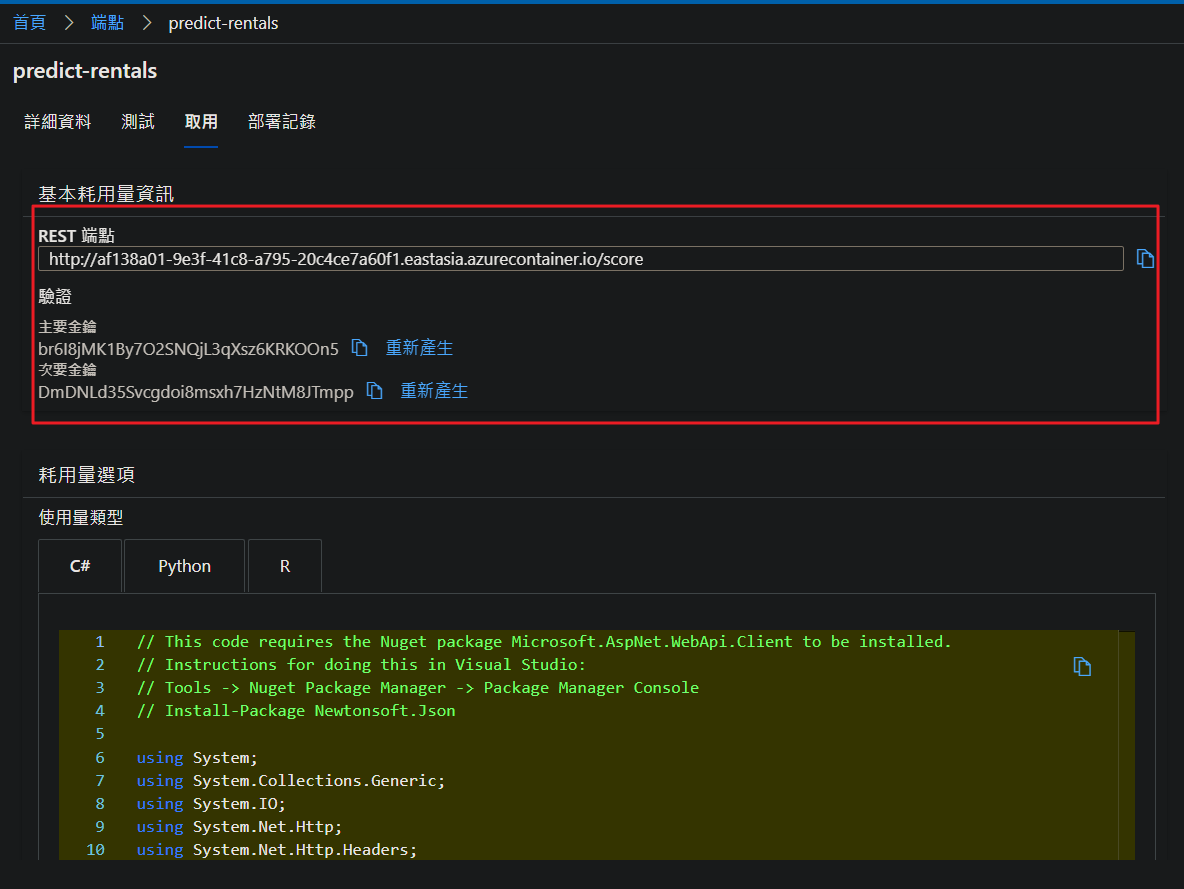








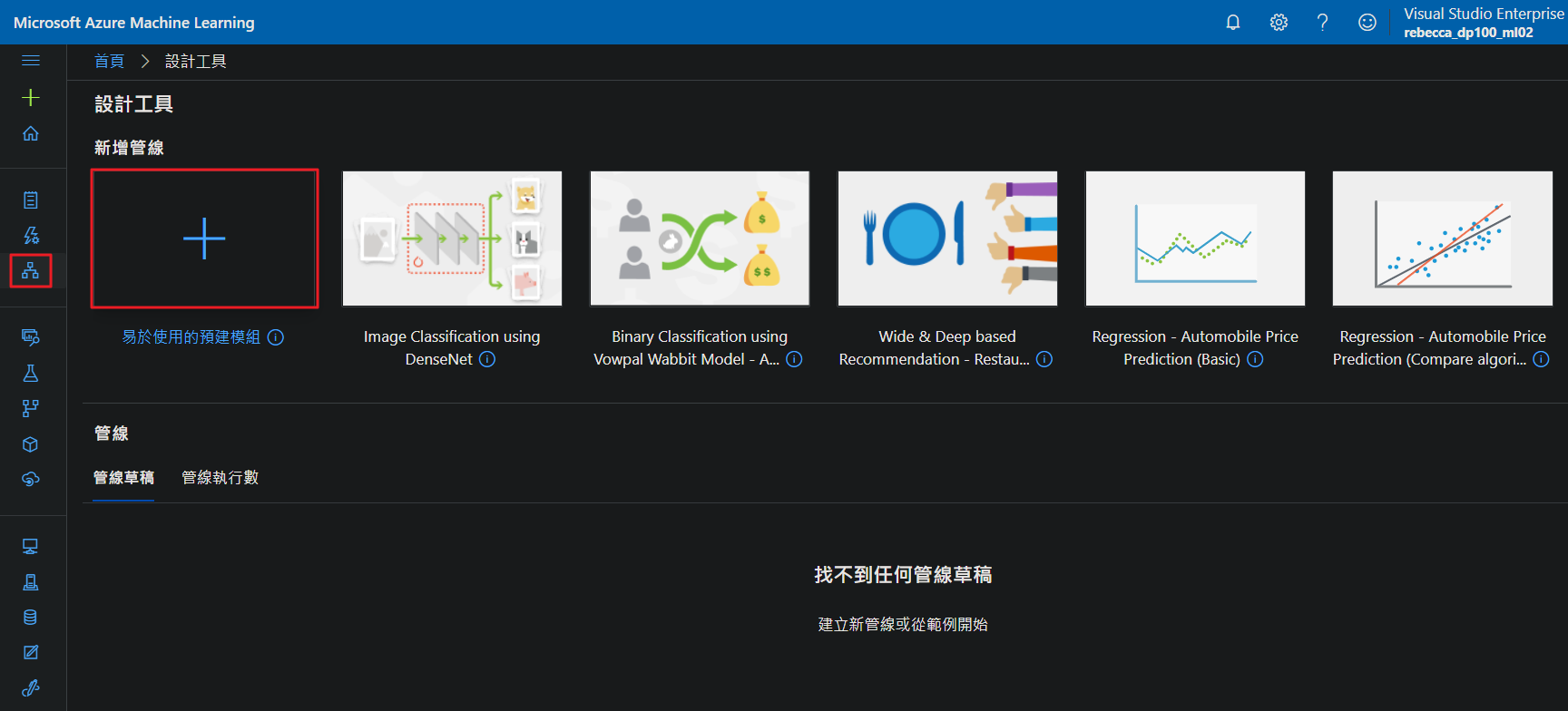


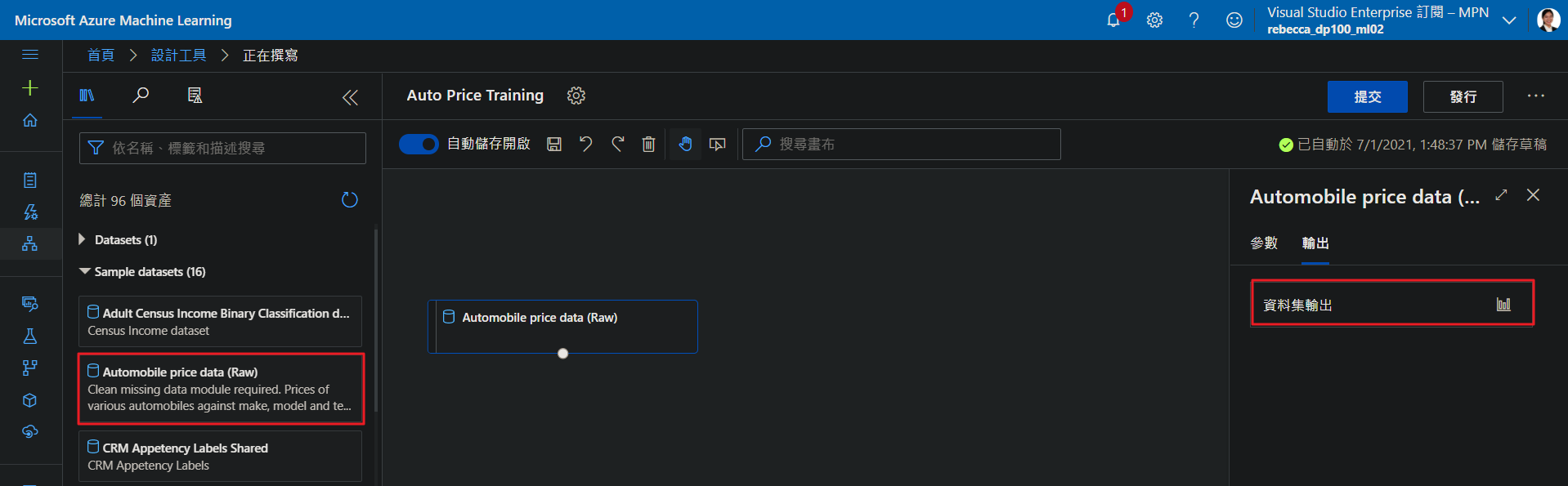






210701(四)









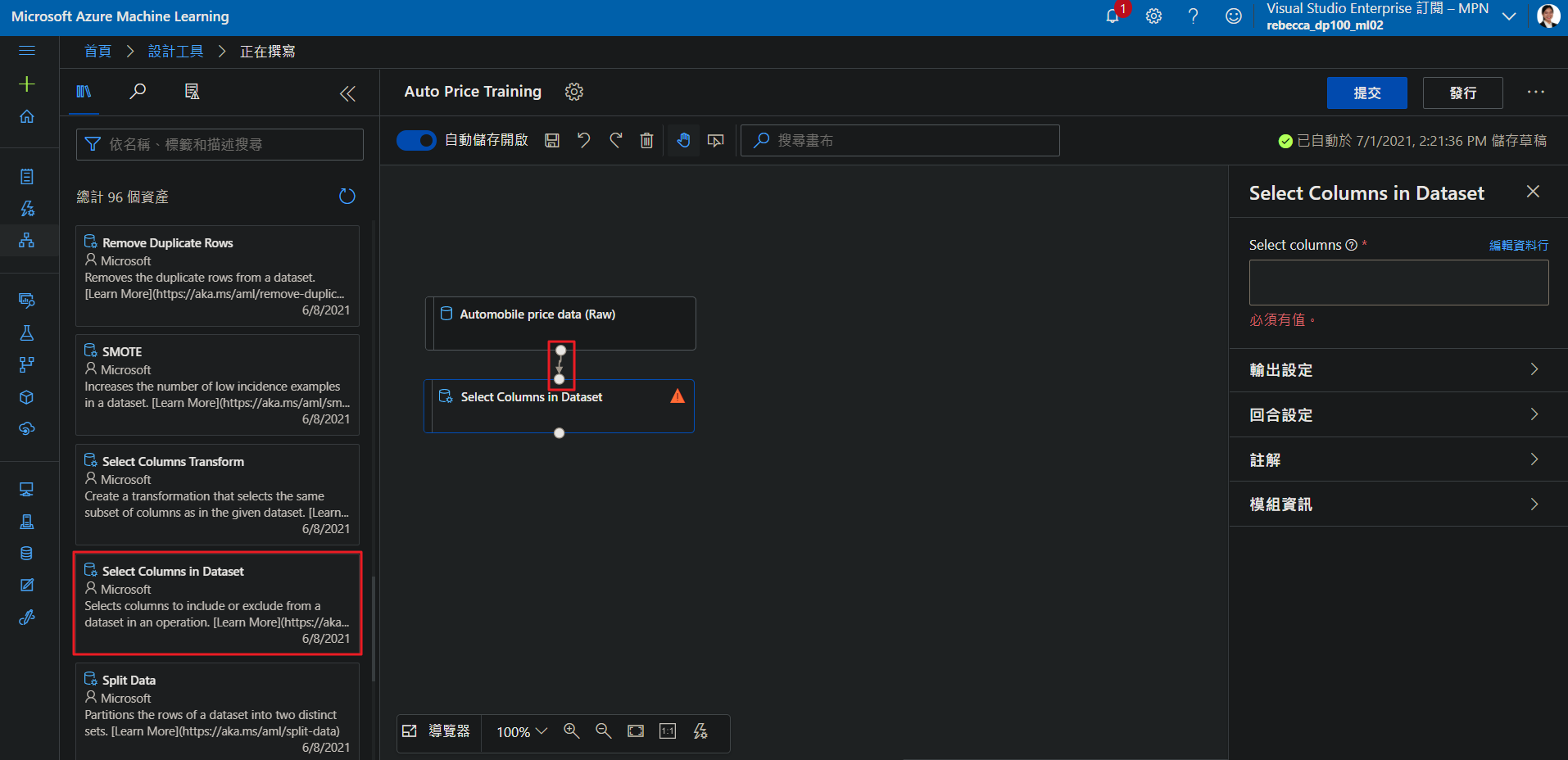
Review the schema of the data, noting that you can see the distributions of the various columns as histograms.

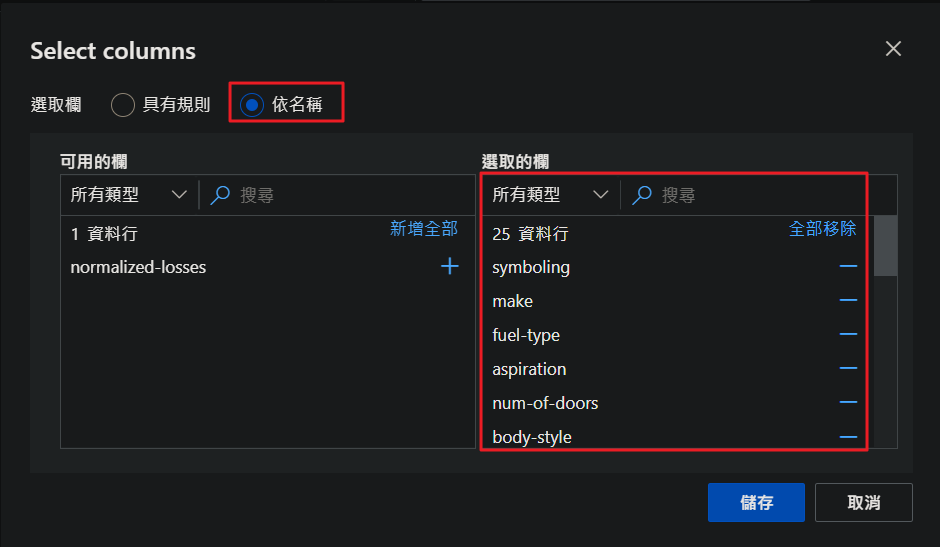
Check various statistics for the column values, and a histogram showing the distribution of the column values.

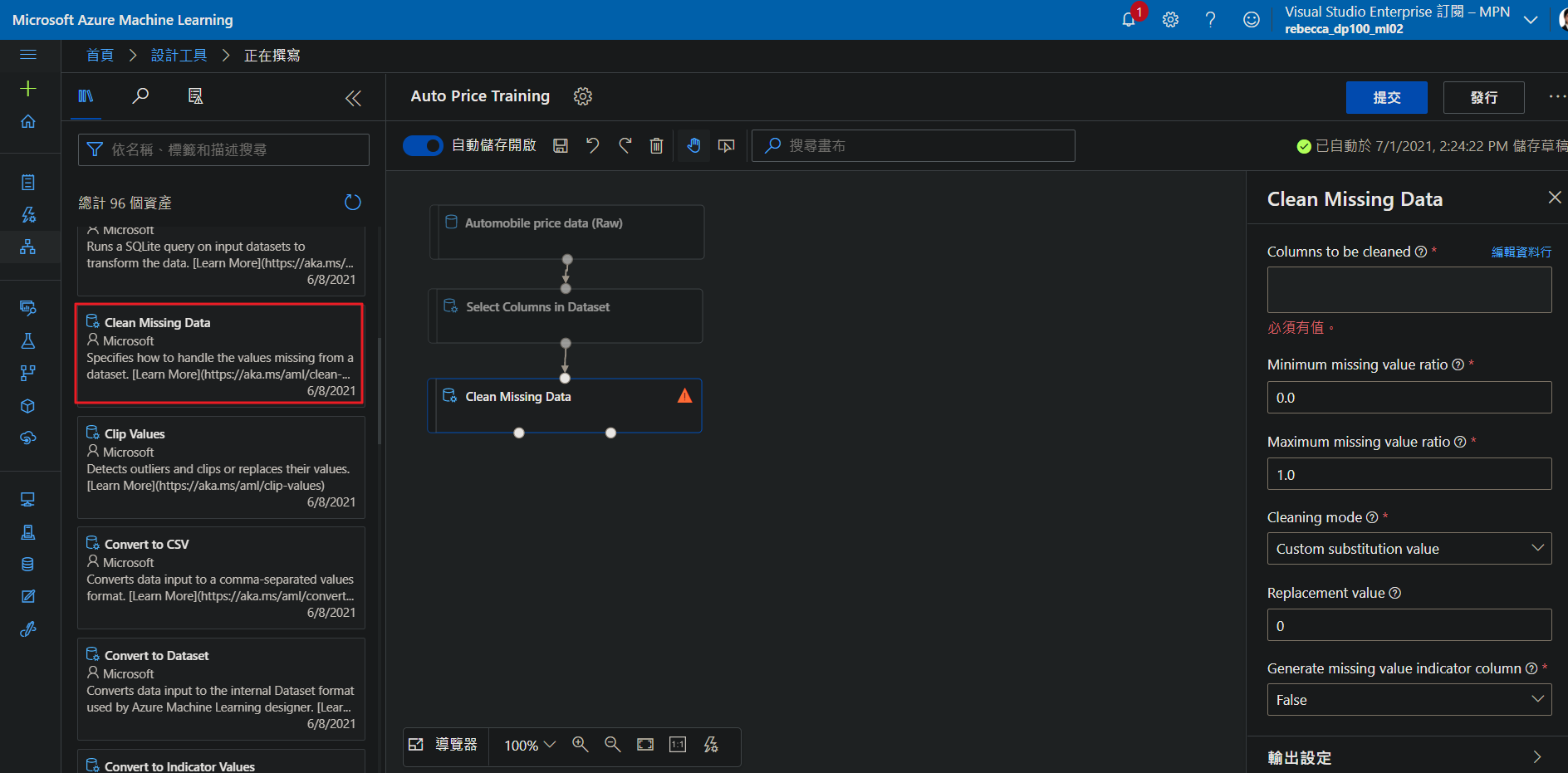
If there are quite a few missing values in a column. This will limit its usefulness in predicting the label; so you might want to exclude it from training.

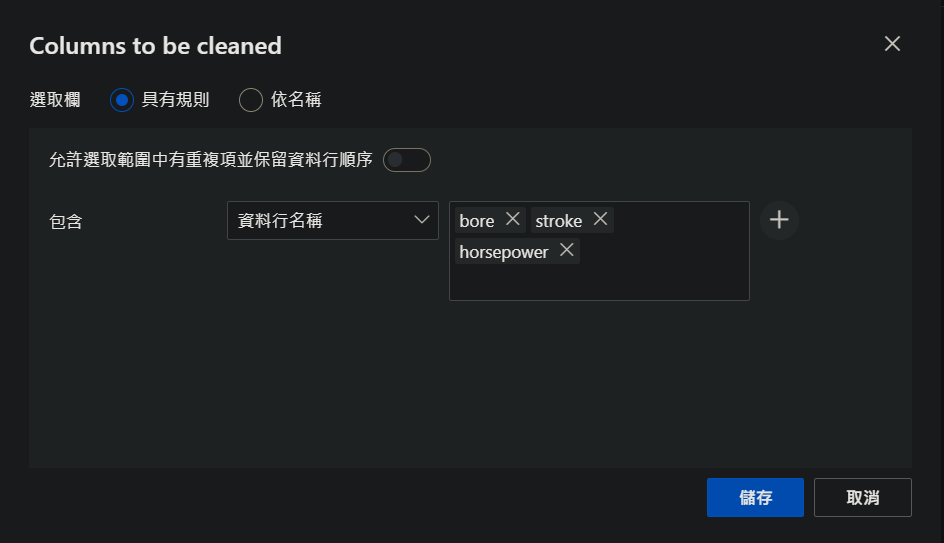
Noting the number of missing values, columns have significantly few missing values, so they may still be useful in predicting label value if you exclude the rows where the values are missing from training.

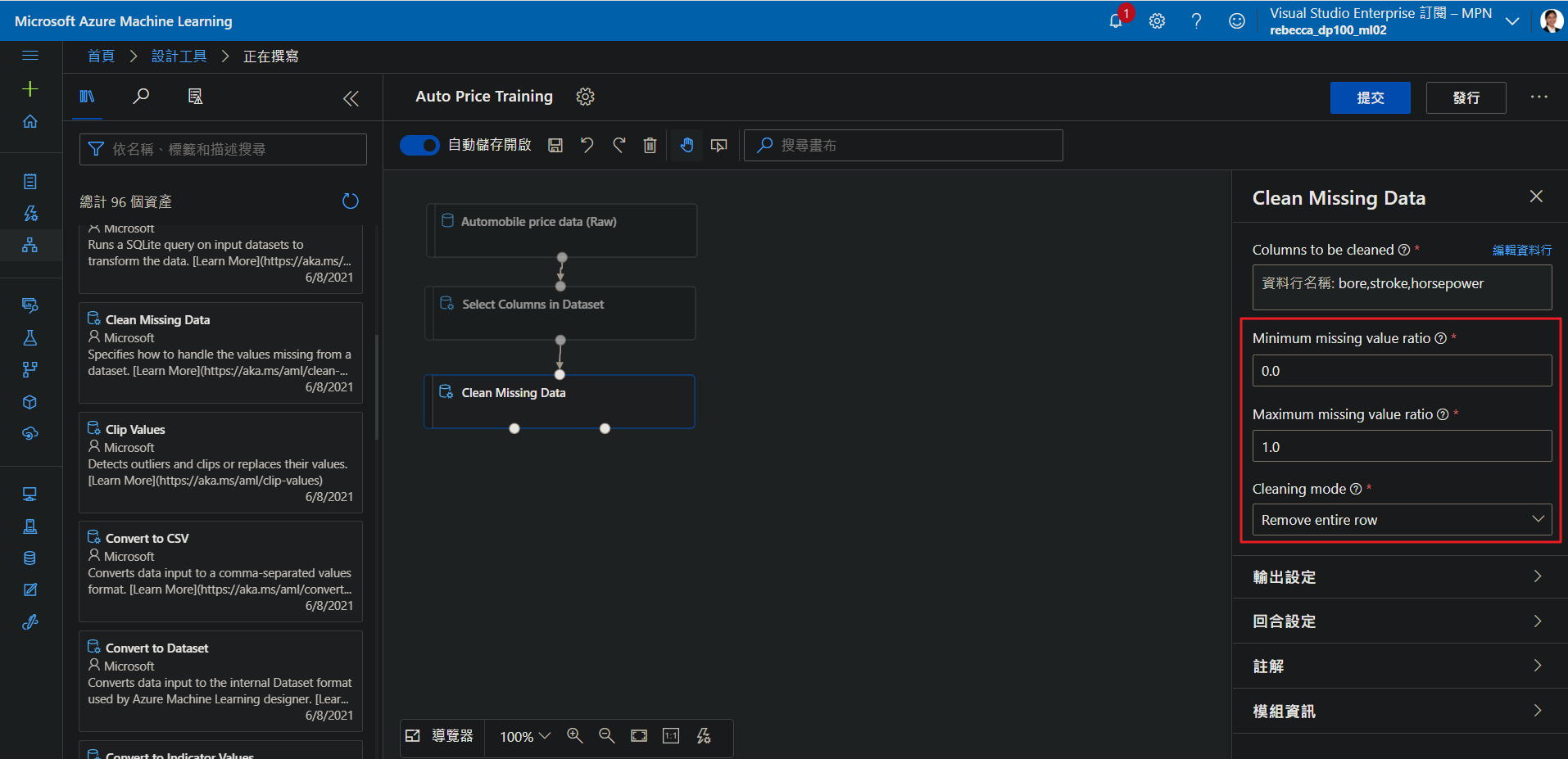
The values in different columns are all measured in different scales, and its possible that the larger values might bias the training algorithm and create an over-dependency on this column compared to columns with lower values. Typically, data scientists mitigate this possible bias by normalizing the numeric columns so they're on the similar scales.



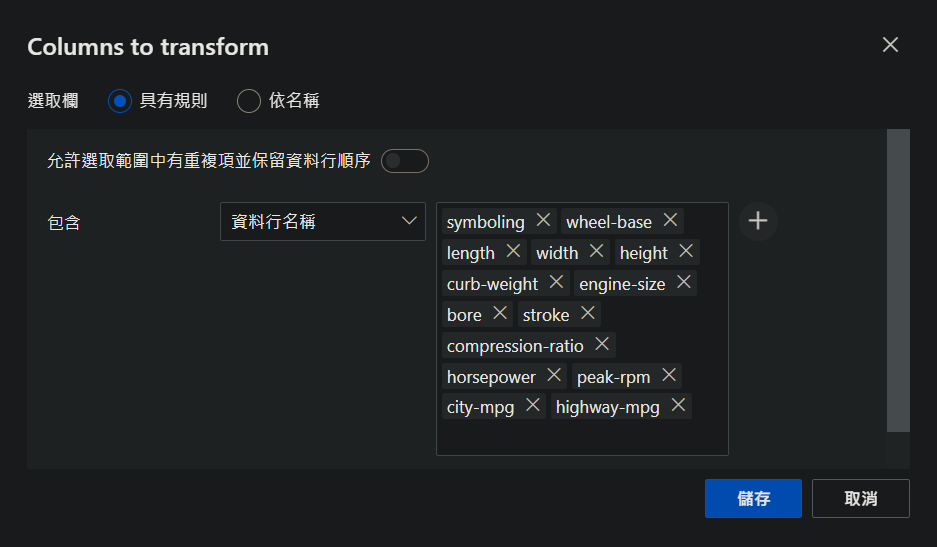


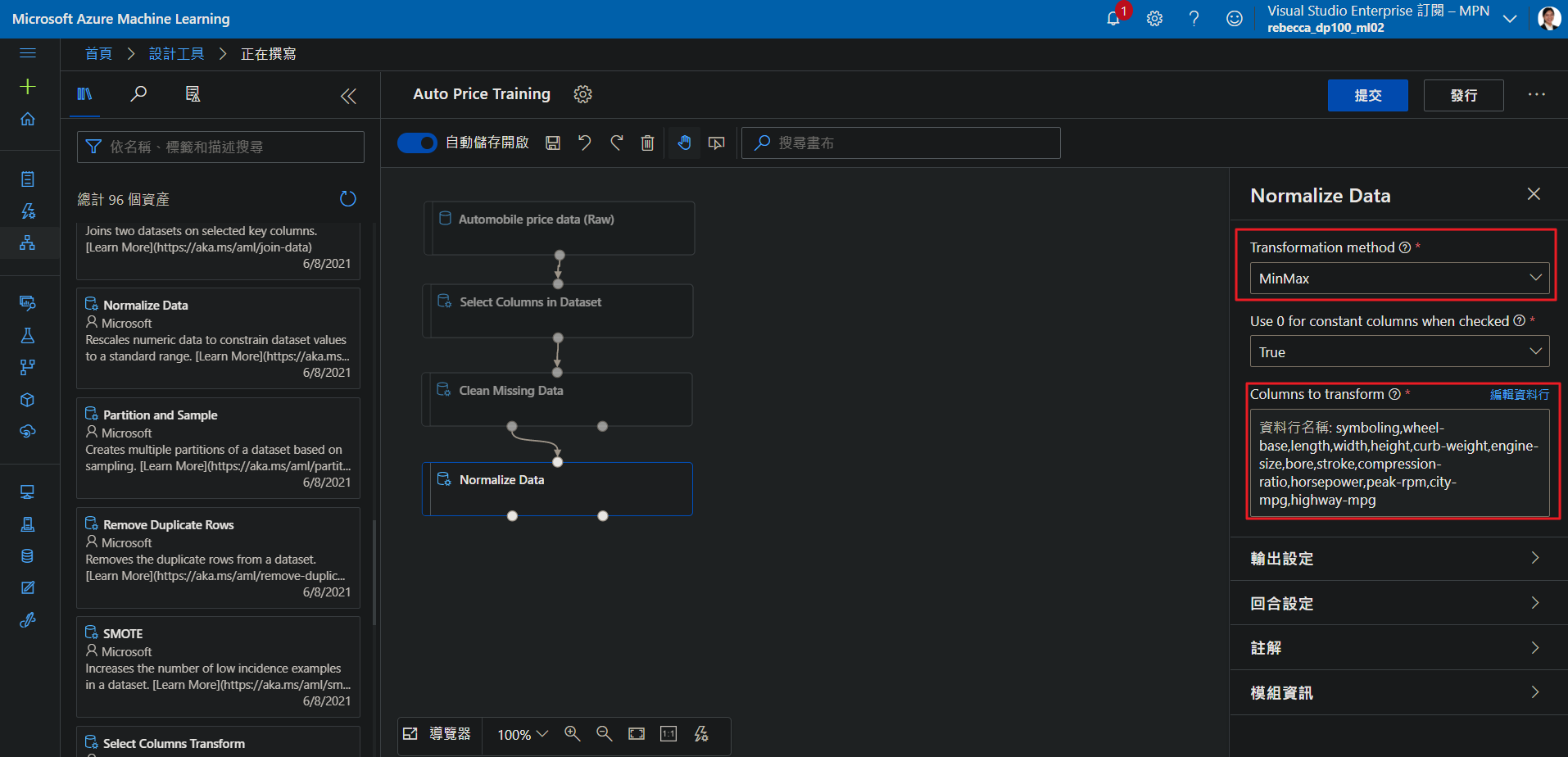


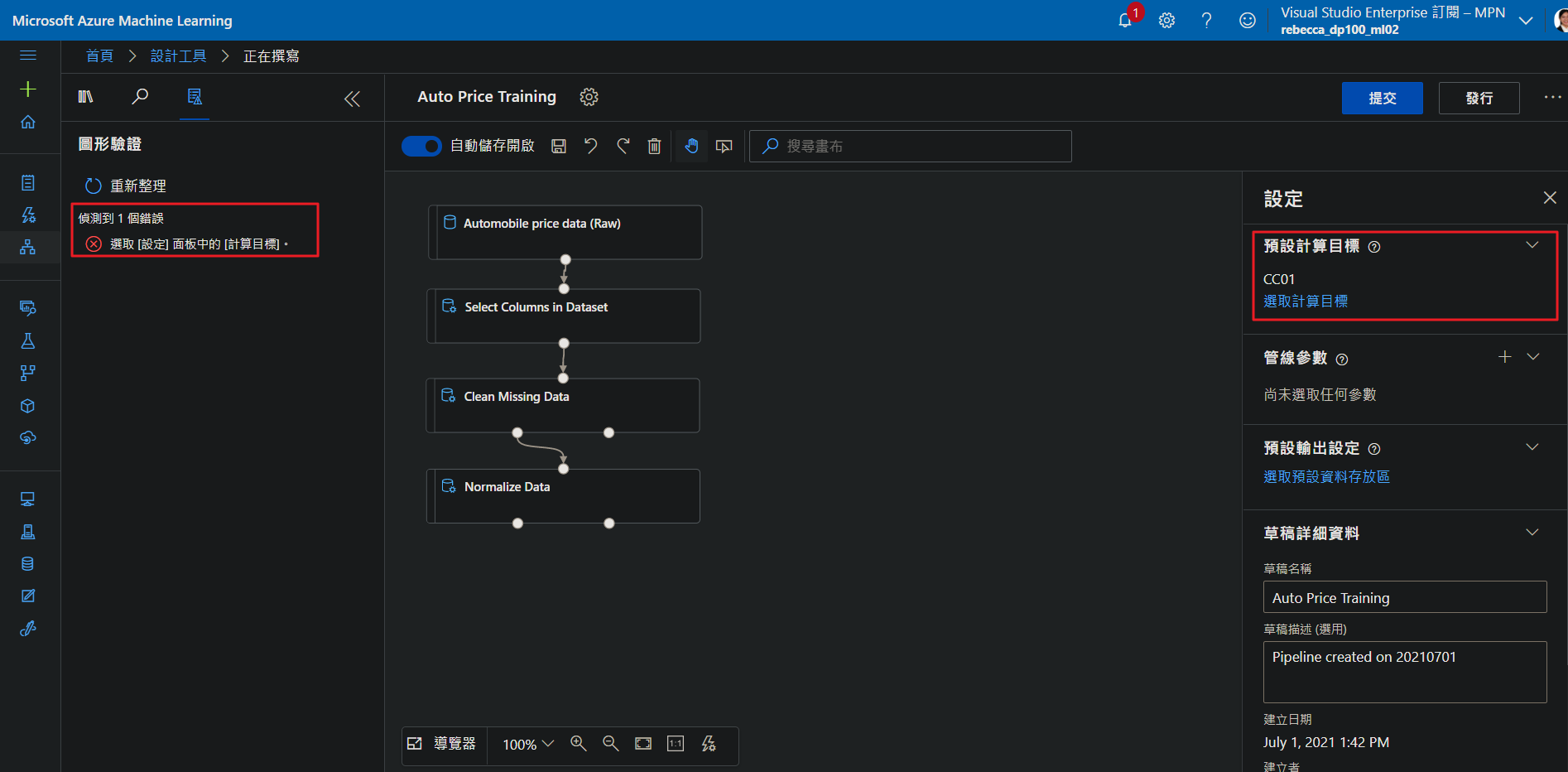


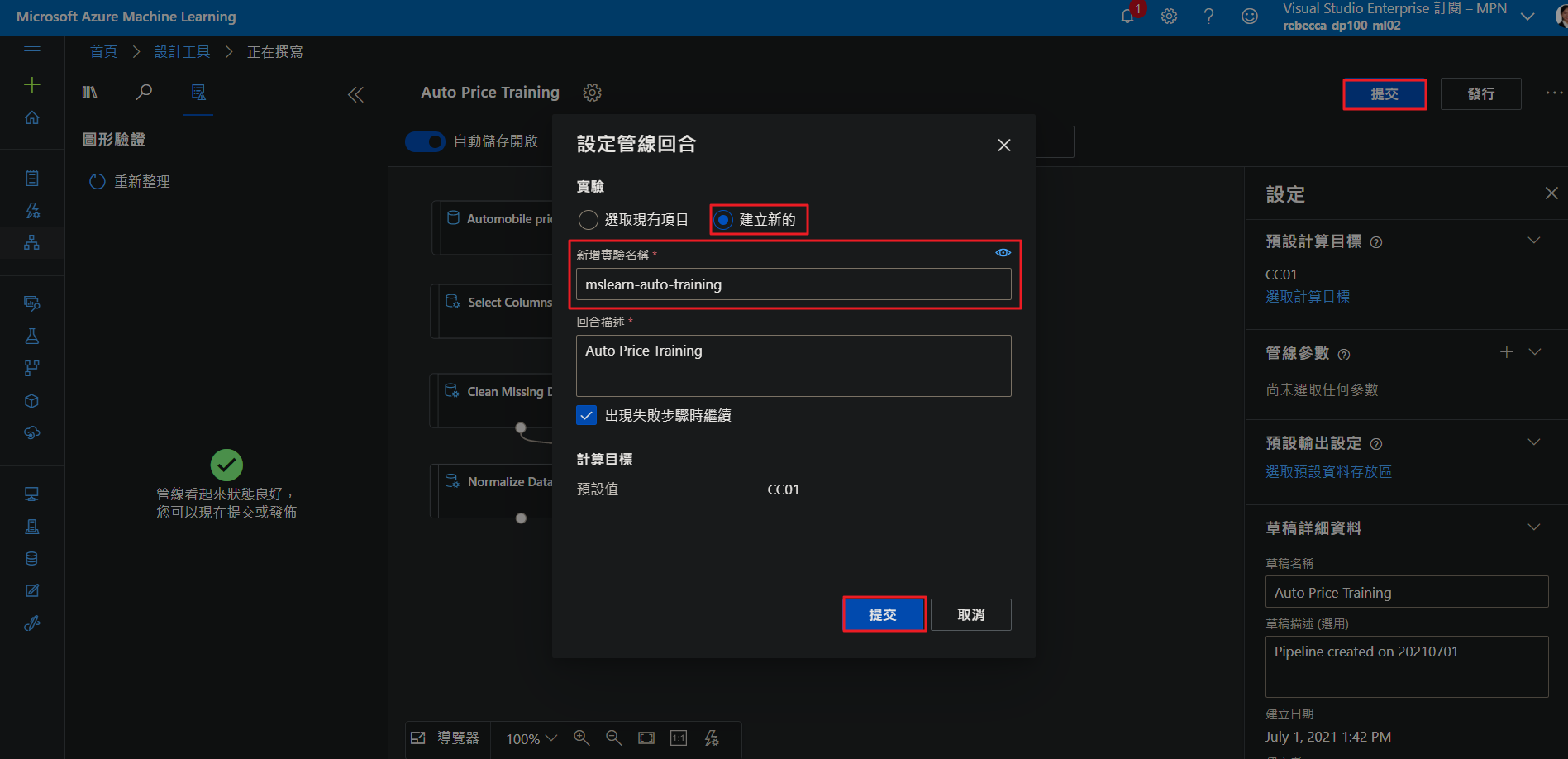


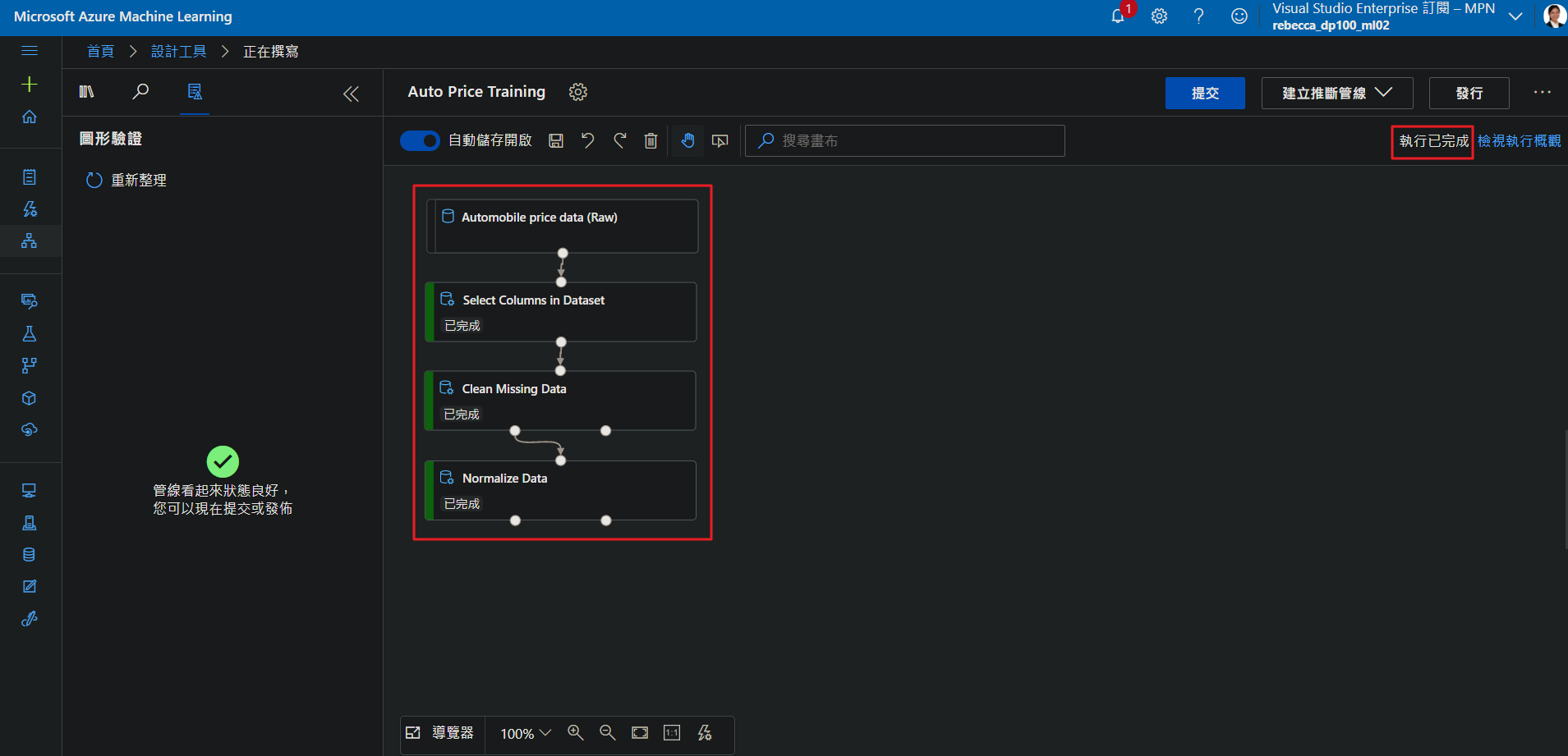




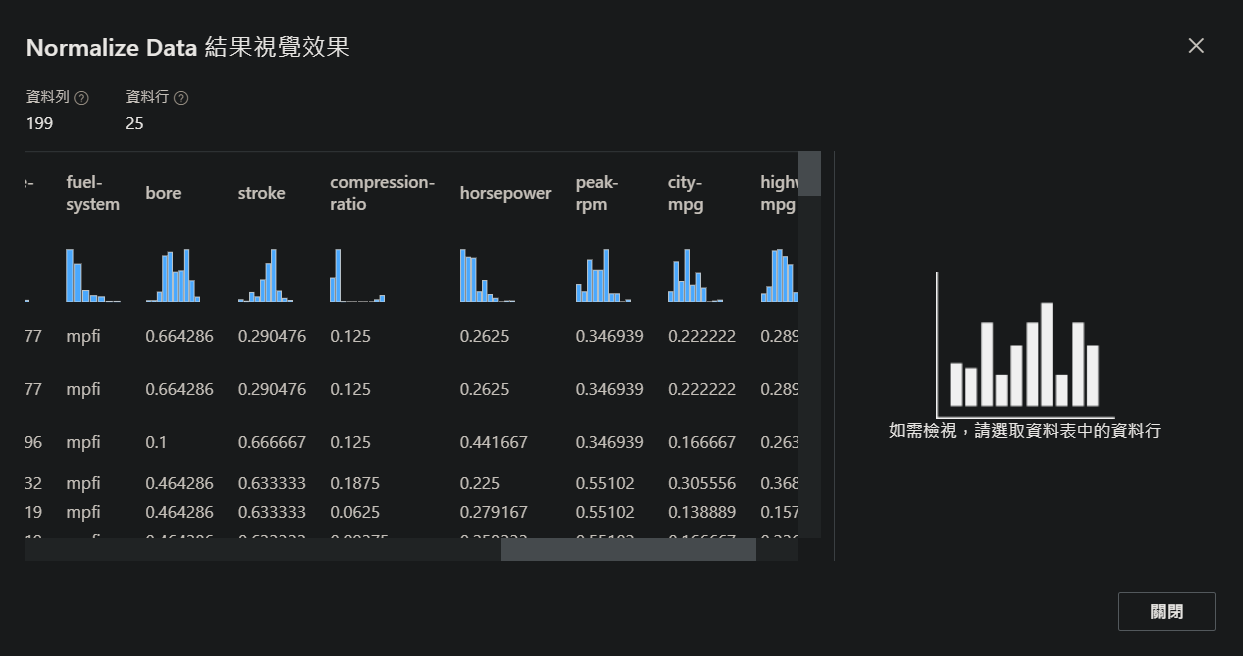


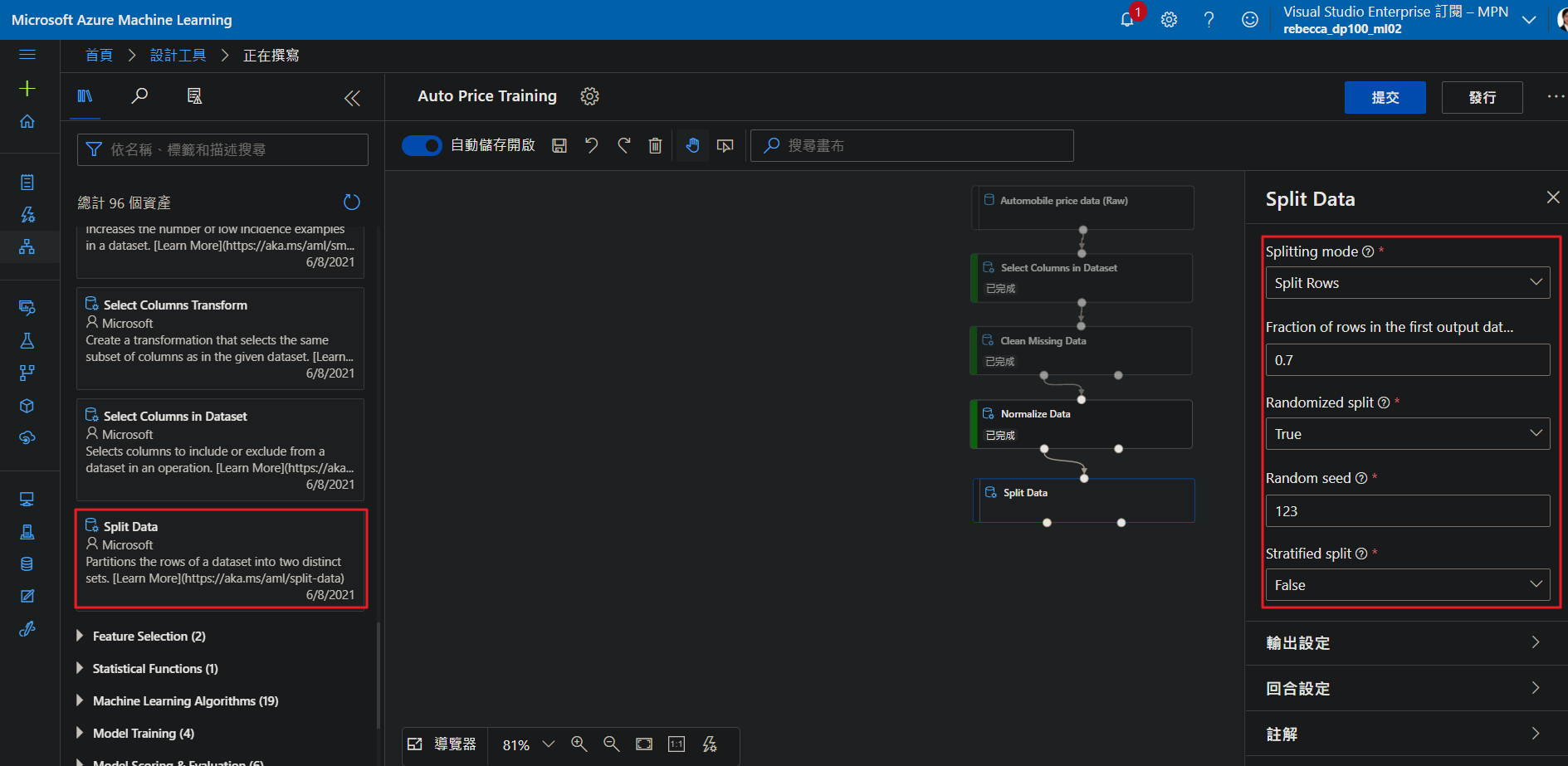


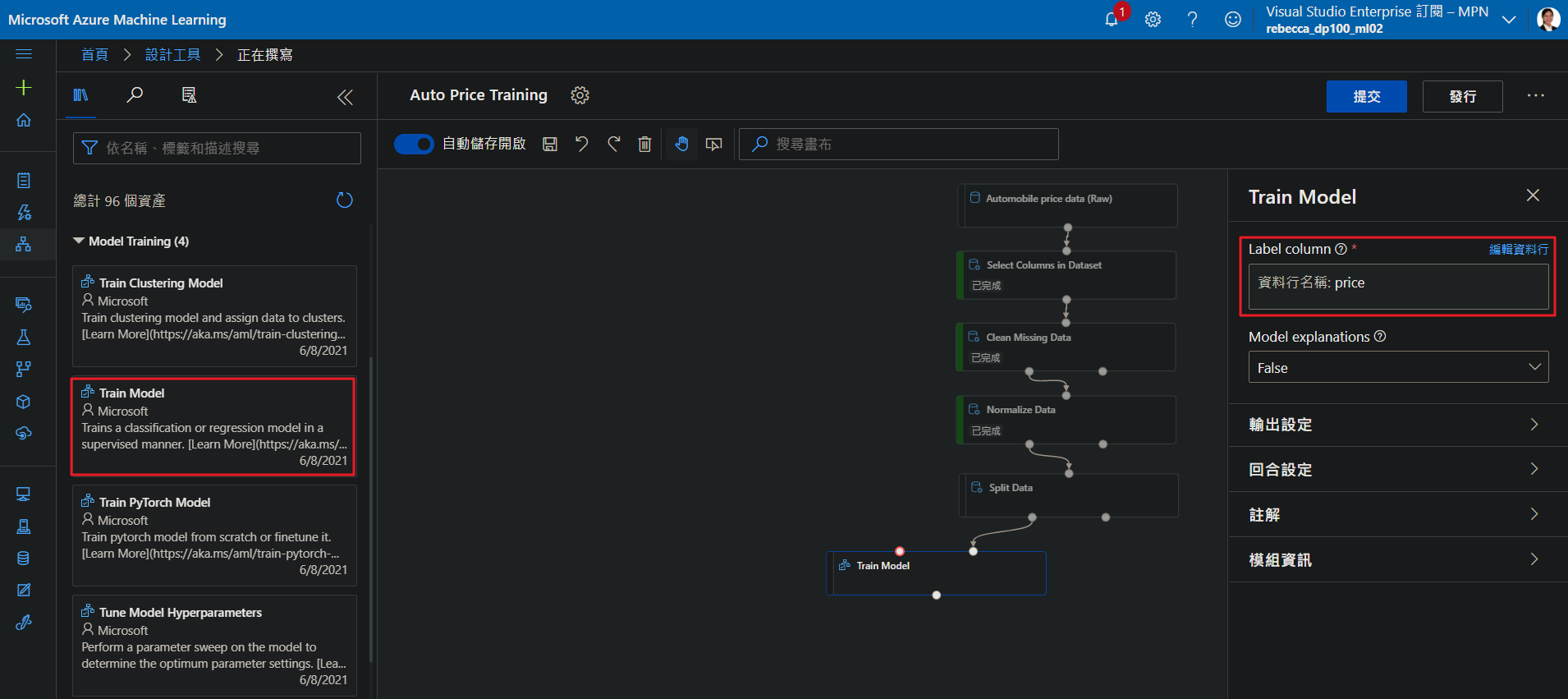


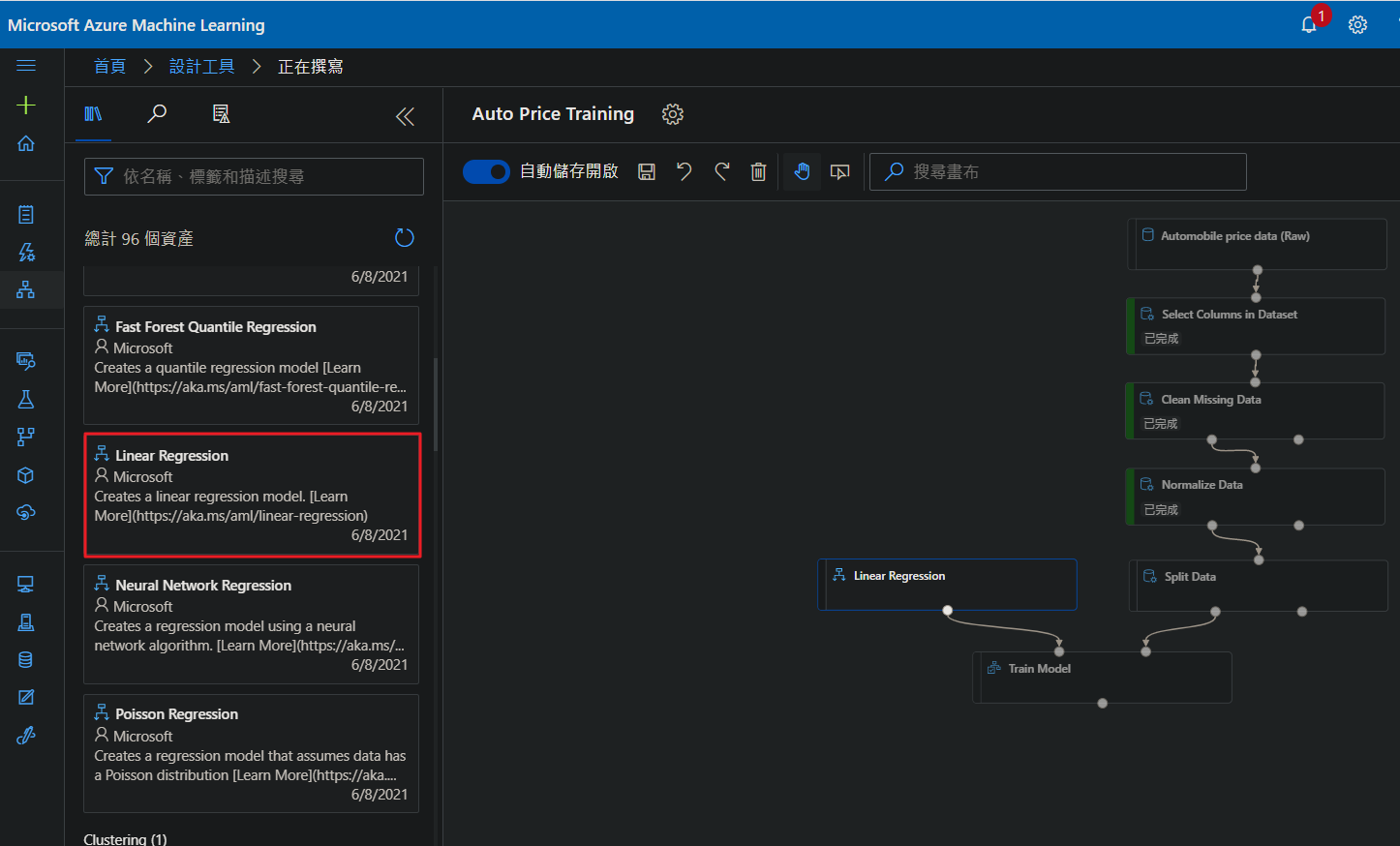


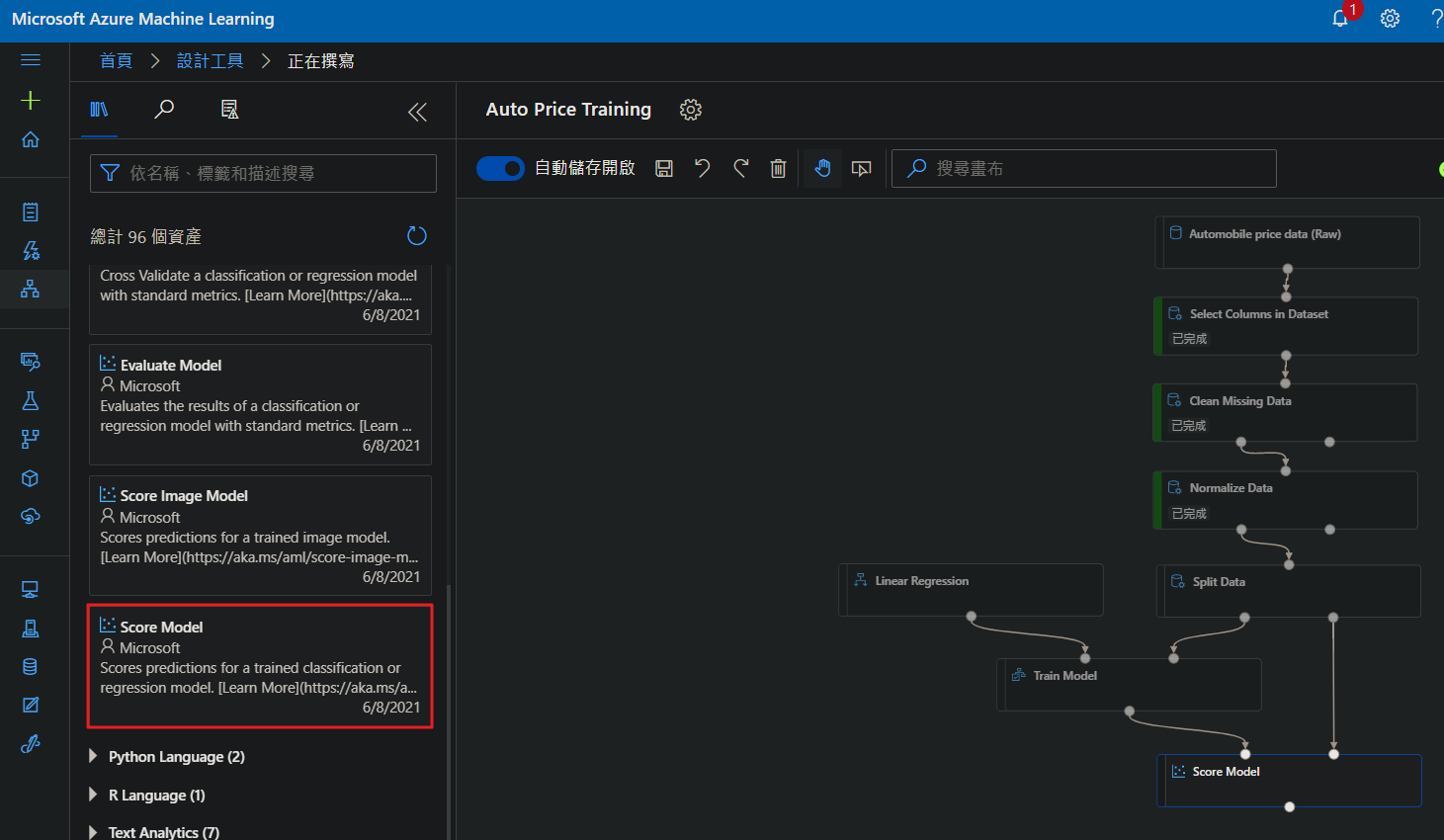


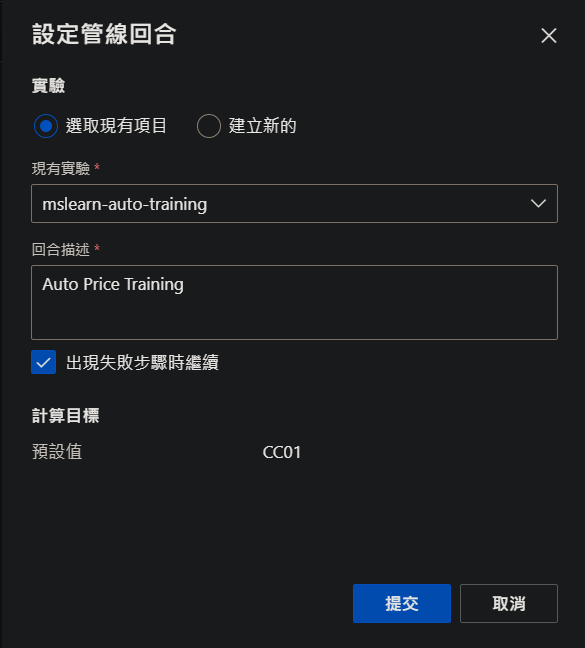




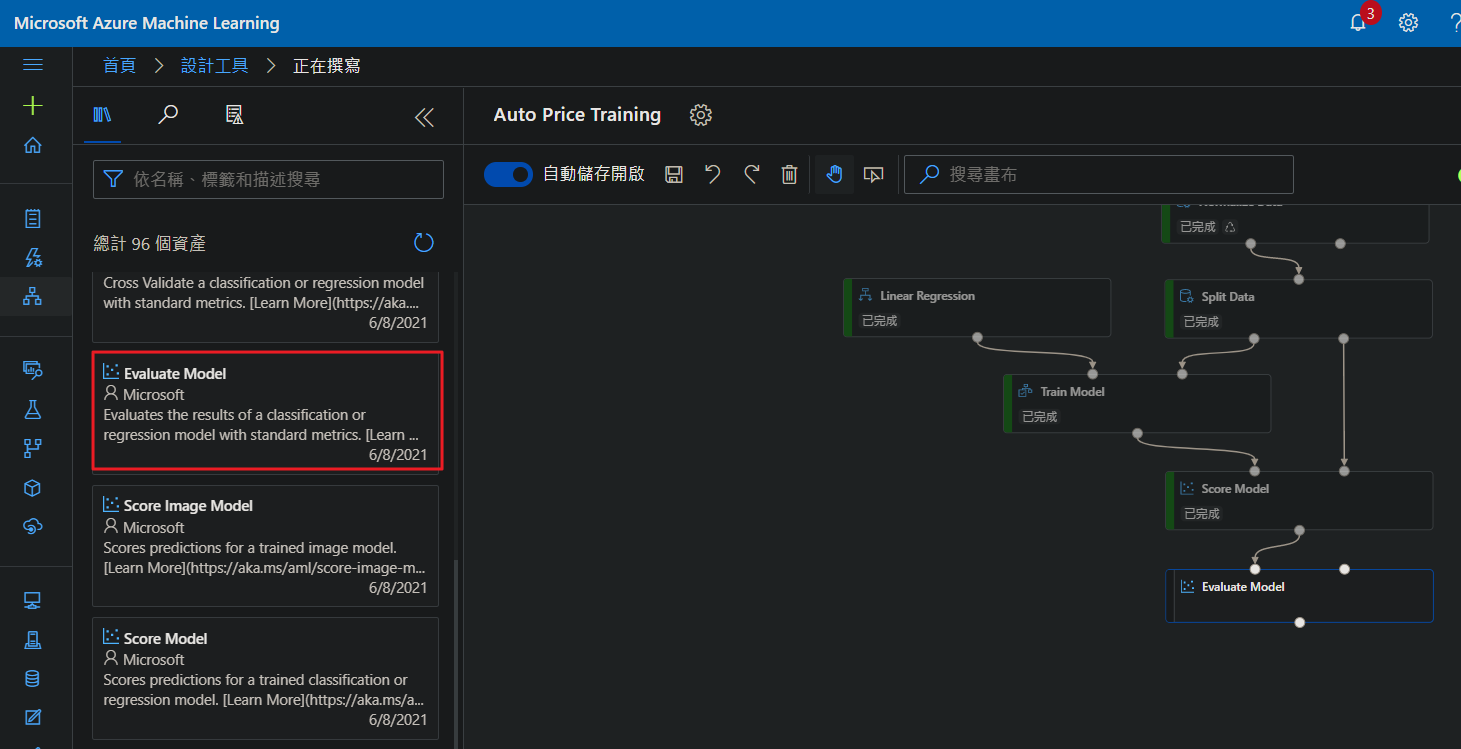




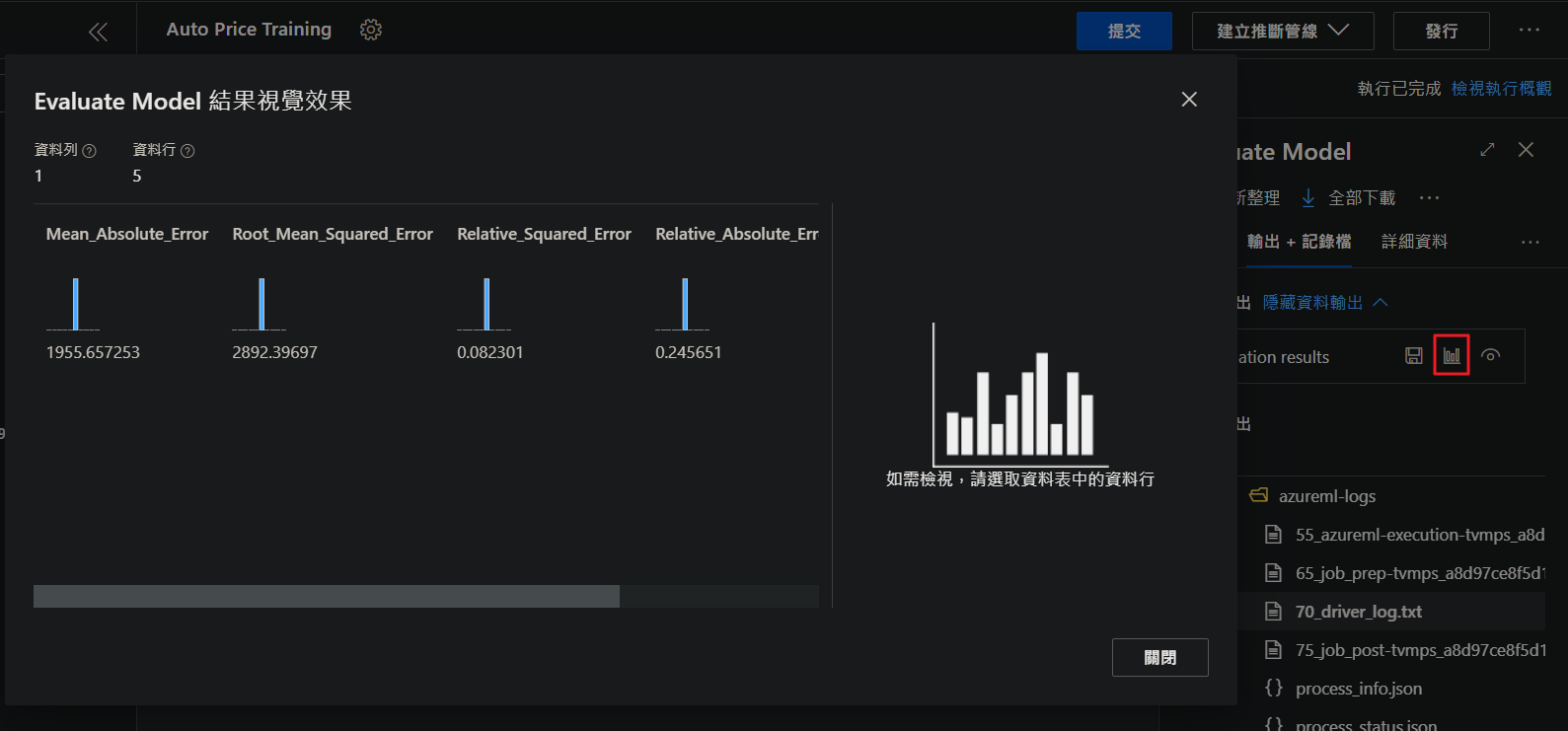








210702(五)



210709(Fri.)

<https://zh.wikipedia.org/wiki/YAML>

YAML（/ˈjæməl/，尾音類似camel駱駝）是一個可讀性高，用來表達資料序列化的格式。YAML參考了其他多種語言，包括：C語言、Python、Perl，並從XML、電子郵件的數據格式（RFC 2822）中獲得靈感。Clark Evans在2001年首次發表了這種語言[1]，另外Ingy döt Net與Oren Ben-Kiki也是這語言的共同設計者[2]。目前已經有數種程式語言或手稿語言支援（或者說解析）這種語言。

YAML是"YAML Ain't a Markup Language"（YAML不是一種標記語言）的遞迴縮寫。在開發的這種語言時，YAML 的意思其實是："Yet Another Markup Language"（仍是一種標記語言），但為了強調這種語言以數據做為中心，而不是以標記語言為重點，而用反向縮略語重新命名。