**Membuat model regresi dengan perancang Azure Machine Learning**

Regresi adalah teknik pembelajaran mesin yang diawasi yang digunakan untuk memprediksi nilai numerik. Pelajari cara membuat model regresi menggunakan perancang Azure Machine Learning.

**Tujuan pembelajaran :**

Pelajari cara melatih dan menerbitkan model regresi dengan perancang Azure Machine Learning.

1. **Pengantar**

Anda dapat menggunakan perancang Microsoft Azure Machine Learning untuk membuat model regresi menggunakan antarmuka visual seret dan lepas, tanpa perlu menulis kode apa pun.

Dalam modul ini, Anda akan mempelajari cara:

* Mengidentifikasi skenario pembelajaran mesin regresi.
* Gunakan perancang Azure Machine Learning untuk melatih model regresi.
* Gunakan model regresi untuk menyimpulkan.
* Terapkan model regresi sebagai layanan.

Untuk menyelesaikan modul ini, Anda memerlukan langganan Microsoft Azure. Jika belum memilikinya, Anda dapat mendaftar untuk uji coba gratis di <https://azure.microsoft.com.>

1. **Mengidentifikasi skenario pembelajaran mesin regresi**

Regresi adalah bentuk pembelajaran mesin yang digunakan untuk memahami hubungan antara variabel untuk memprediksi hasil yang diinginkan. Regresi memprediksi label atau hasil numerik berdasarkan variabel, atau fitur. Misalnya, perusahaan penjualan mobil mungkin menggunakan karakteristik mobil (seperti ukuran mesin, jumlah kursi, jarak tempuh, dan sebagainya) untuk memprediksi kemungkinan harga jualnya. Dalam hal ini, karakteristik mobil adalah fitur, dan harga jualnya adalah label.

Regresi adalah contoh teknik pembelajaran mesin yang diawasi tempat Anda melatih model menggunakan data yang menyertakan fitur dan nilai yang diketahui untuk label, sehingga model belajar untuk menyesuaikan kombinasi fitur dengan label. Selanjutnya, setelah pelatihan selesai, Anda dapat menggunakan model terlatih untuk memprediksi label untuk item baru yang labelnya tidak diketahui.

1. **Skenario untuk model pembelajaran mesin regresi**

Model pembelajaran mesin regresi digunakan di banyak industri. Beberapa skenarionya adalah:

* Menggunakan karakteristik rumah, seperti luas persegi dan jumlah kamar, untuk memprediksi harga rumah.
* Menggunakan karakteristik kondisi peternakan, seperti cuaca dan kualitas tanah, untuk memprediksi hasil panen.
* Menggunakan karakteristik kampanye sebelumnya, seperti log iklan, untuk memprediksi klik iklan di masa mendatang.

1. **Apa itu Azure Machine Learning?**

Melatih dan menyebarkan model pembelajaran mesin yang efektif melibatkan banyak pekerjaan, sebagian besar memakan waktu dan padat sumber daya. Azure Machine Learning adalah layanan berbasis cloud yang membantu menyederhanakan beberapa tugas yang diperlukan untuk menyiapkan data, melatih model, dan menyebarkan layanan prediktif. Model pembelajaran mesin regresi dapat dibangun menggunakan Azure Machine Learning.

**Yang paling penting**, Azure Machine Learning membantu ilmuwan data meningkatkan efisiensi mereka dengan mengotomatiskan banyak tugas yang memakan waktu, yang terkait dengan model pelatihan. Hal ini memungkinkan mereka untuk menggunakan sumber daya komputasi berbasis cloud yang dapat diskalakan secara efektif untuk menangani volume data berukuran besar dan mengeluarkan biaya hanya jika benar-benar digunakan.

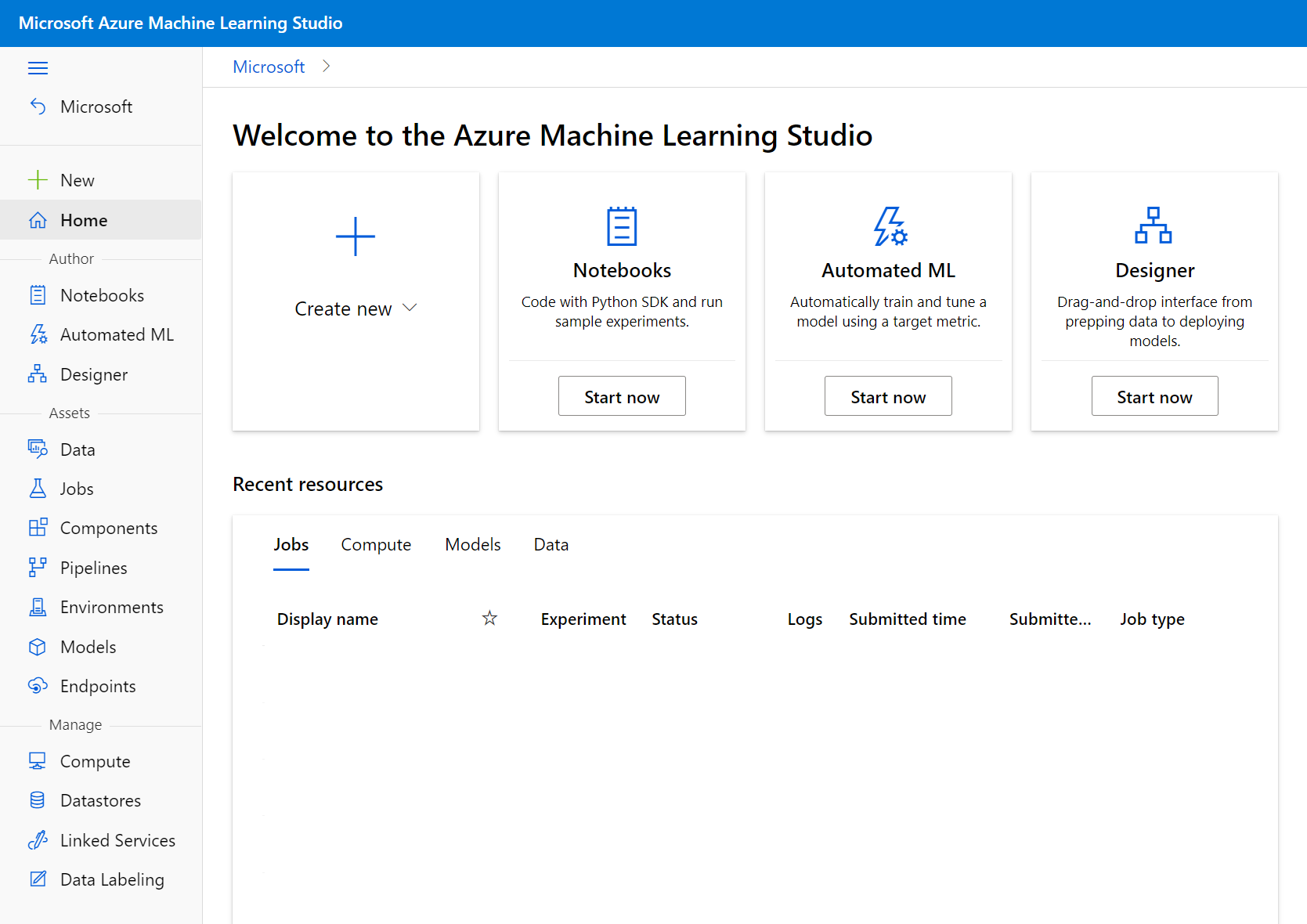
1. **Ruang kerja Azure Machine Learning**

Untuk menggunakan Azure Machine Learning, buat sumber daya ruang kerja terlebih dahulu di langganan Azure Anda. Kemudian, Anda dapat menggunakan ruang kerja ini untuk mengelola data, sumber daya komputasi, kode, model, dan artefak lainnya yang terkait dengan beban kerja pembelajaran mesin Anda.

Setelah membuat ruang kerja Azure Machine Learning, Anda dapat mengembangkan solusi dengan layanan pembelajaran mesin Azure, baik dengan alat pengembang atau portal web studio Azure Machine Learning.

1. **Studio Azure Machine Learning**

Studio Azure Machine Learning adalah portal web untuk solusi pembelajaran mesin di Azure. Ini termasuk berbagai fitur dan kemampuan yang membantu ilmuwan data menyiapkan data, melatih model, menerbitkan layanan prediktif, dan memantau penggunaannya. Untuk mulai menggunakan portal web, Anda perlu menetapkan ruang kerja yang dibuat di portal Azure ke studio Azure Machine Learning.



1. **Komputasi Azure Machine Learning**

Pada intinya, Azure Machine Learning adalah layanan untuk melatih dan mengelola model pembelajaran mesin, dan diperlukan sumber daya komputasi untuk menjalankan proses pelatihan tersebut. Target komputasi adalah sumber daya berbasis cloud tempat Anda dapat menjalankan pelatihan model dan proses eksplorasi data.

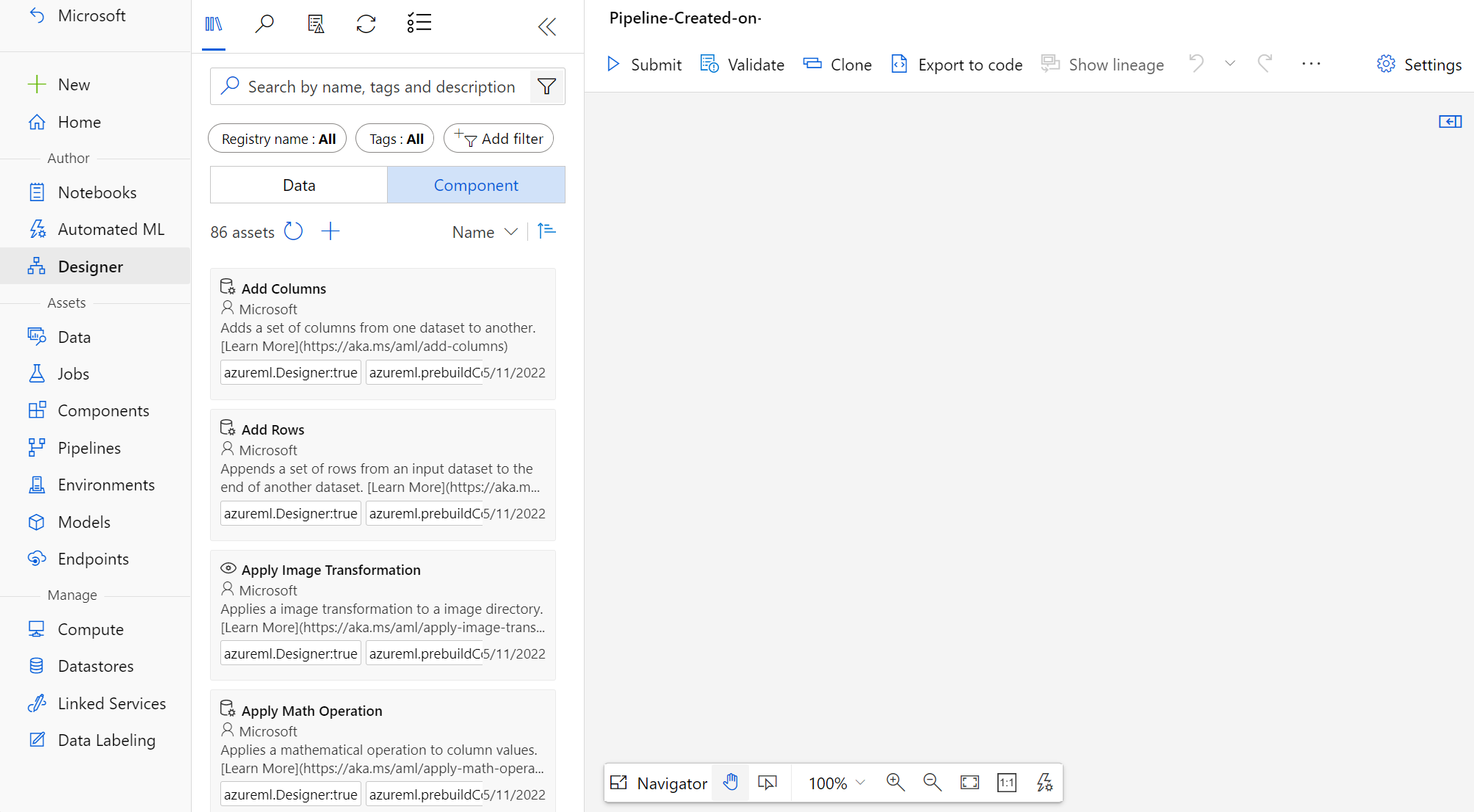
Di studio Azure Machine Learning : https://ml.azure.com/, Anda dapat mengelola target komputasi untuk aktivitas ilmu data. Ada empat jenis sumber daya komputasi yang dapat dibuat:

* **Instans Komputasi:** Stasiun kerja pengembangan yang dapat digunakan oleh ilmuwan data untuk bekerja dengan data dan model.
* **Kluster Komputasi:** Kluster komputer virtual yang dapat diskalakan untuk pemrosesan kode eksperimen sesuai permintaan.
* **Kluster Inferensi:** Target penyebaran untuk layanan prediktif yang menggunakan model terlatih Anda.
* **Komputasi Terlampir:** Tautan ke sumber daya komputasi Azure yang ada, seperti kluster Virtual Machines atau Azure Databricks.

1. **Apa itu Perancang / Designer Azure Machine Learning?**

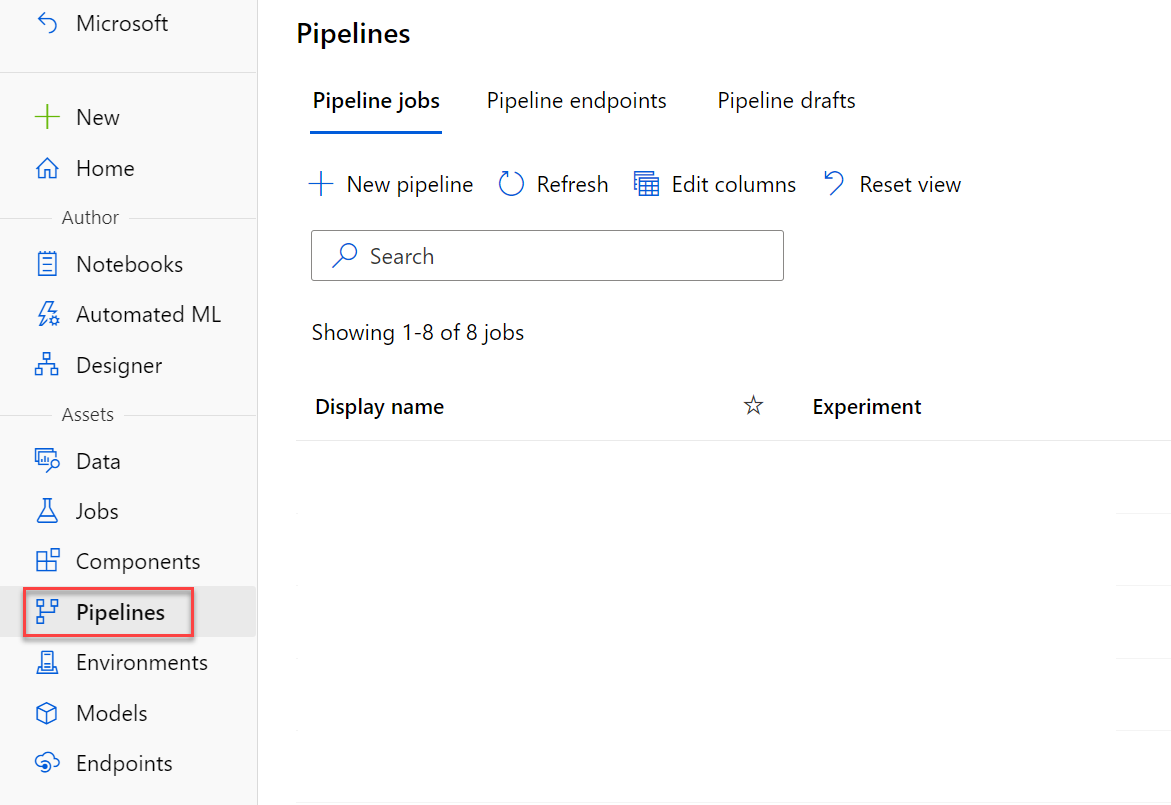
Di studio Azure Machine Learning, ada beberapa cara untuk membuat model pembelajaran mesin regresi. Salah satu caranya adalah dengan menggunakan antarmuka visual yang disebut perancang yang dapat Anda gunakan untuk melatih, menguji, dan menyebarkan model pembelajaran mesin. Antarmuka seret dan letakkan menggunakan input dan output yang ditentukan dengan jelas yang dapat dibagikan, digunakan kembali, dan dikontrol versi.

Setiap proyek perancang, yang dikenal sebagai alur, memiliki panel kiri untuk navigasi dan kanvas di sisi kanan Anda. Untuk menggunakan perancang, identifikasi blok penyusun, atau komponen, yang diperlukan untuk model Anda, tempatkan dan sambungkan di kanvas Anda, dan jalankan pekerjaan pembelajaran mesin.



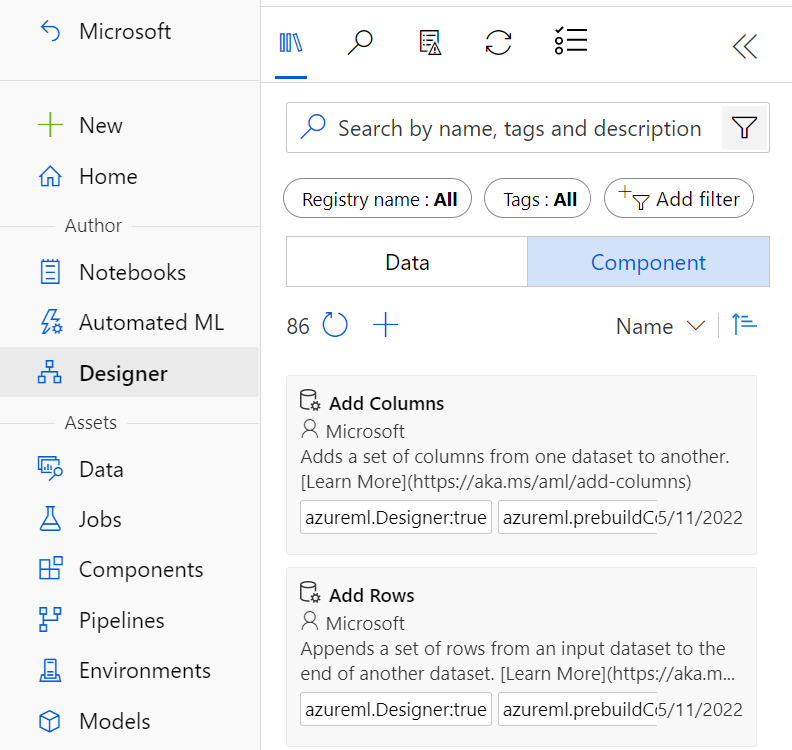
1. **Pipelines**

Alur memungkinkan Anda mengatur, mengelola, dan menggunakan kembali alur kerja pembelajaran mesin yang kompleks di seluruh proyek dan pengguna. Alur dimulai dengan himpunan data tempat Anda ingin melatih model. Setiap kali Anda menjalankan alur, konfigurasi alur dan hasilnya disimpan di ruang kerja Anda sebagai pekerjaan alur.



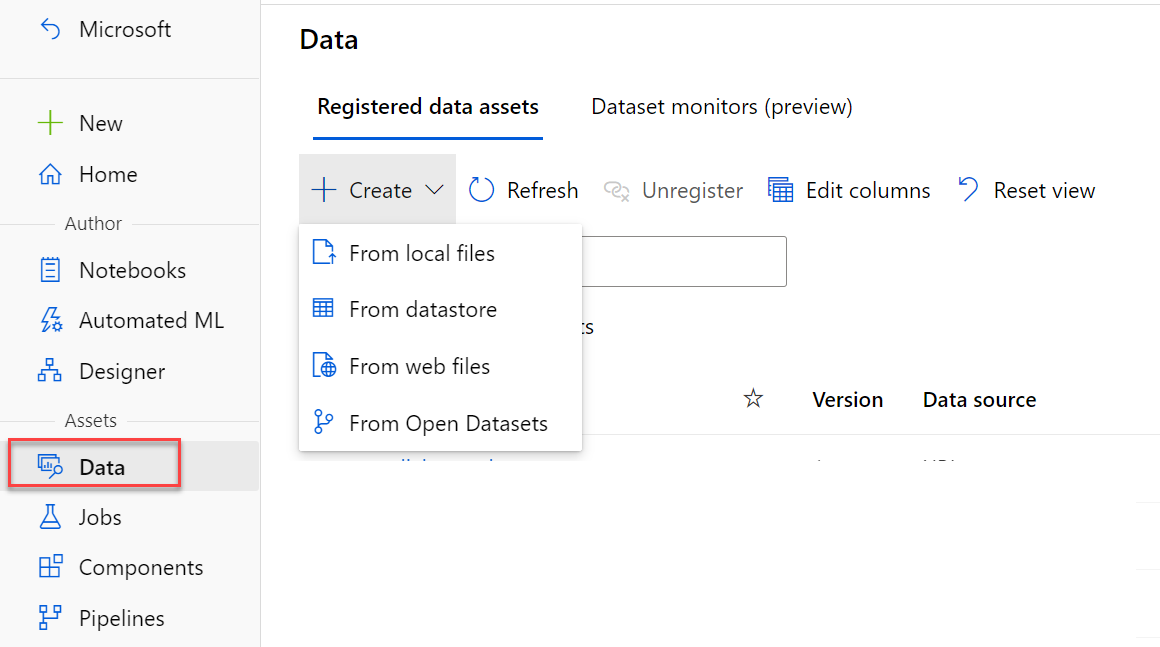
1. **Komponen**

Komponen Azure Machine Learning merangkum satu langkah dalam alur pembelajaran mesin. Anda dapat menganggap komponen sebagai fungsi pemrograman dan sebagai blok penyusun untuk alur Azure Machine Learning. Dalam proyek alur, Anda dapat mengakses aset data dan komponen dari tab **Pustaka Aset** panel kiri.



1. **Himpunan Data**

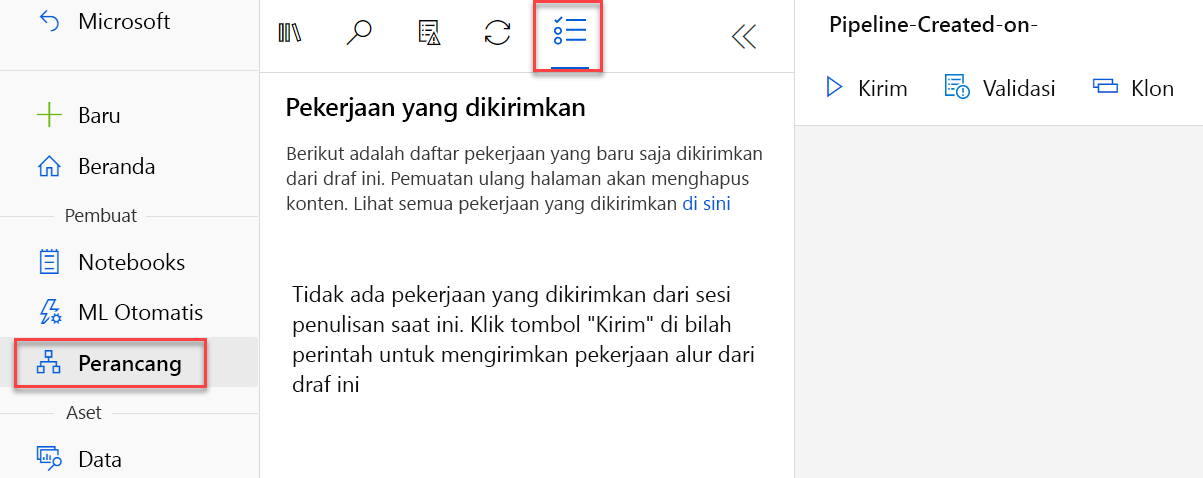
Anda dapat membuat aset data pada halaman **Data** dari file lokal, penyimpanan data, file web, dan Open Datasets. Aset data ini akan muncul bersama dengan himpunan data sampel standar di **Pustaka Aset** perancang.



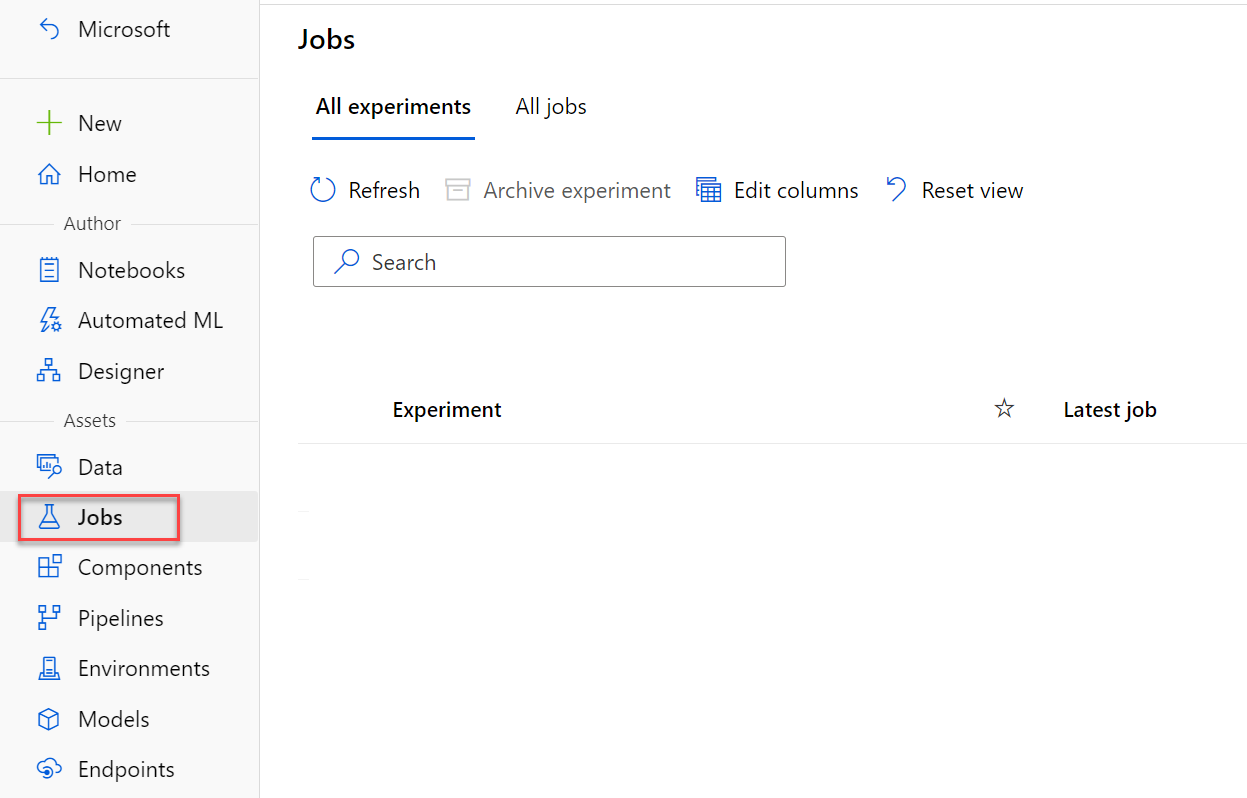
1. **Pekerjaan Azure Machine Learning**

Pekerjaan Azure Machine Learning (ML) menjalankan tugas terhadap target komputasi yang ditentukan. Pekerjaan memungkinkan pelacakan sistematis untuk eksperimen dan alur kerja pembelajaran mesin. Setelah pekerjaan dibuat, Azure ML mempertahankan rekaman eksekusi untuk pekerjaan tersebut. Semua rekaman eksekusi pekerjaan Anda dapat dilihat di studio Azure ML.

Dalam proyek perancang, Anda dapat mengakses status pekerjaan alur menggunakan tab **Pekerjaan yang dikirimkan** di panel kiri.



Anda dapat menemukan semua pekerjaan yang telah dijalankan di ruang kerja pada halaman **Pekerjaan**.



1. **Memahami langkah-langkah untuk regresi**

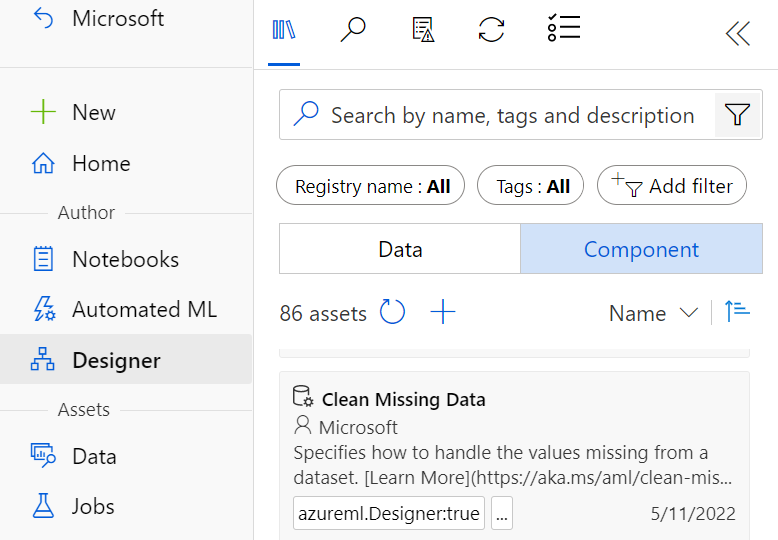
Anda dapat membayangkan langkah-langkah untuk melatih dan mengevaluasi model pembelajaran mesin regresi seperti:

1. **Menyiapkan data:** Mengidentifikasi fitur dan label dalam himpunan data. Pra-proses, atau membersihkan dan mentransformasi, data seperlunya.
2. **Melatih model:** Membagi data menjadi dua grup, pelatihan, dan set validasi. Latih model pembelajaran mesin menggunakan himpunan data pelatihan. Uji model pembelajaran mesin untuk performa menggunakan himpunan data validasi.
3. **Mengevaluasi performa:** Bandingkan seberapa dekat prediksi model dengan label yang diketahui.
4. **Menyebarkan layanan prediktif:** Setelah melatih model pembelajaran mesin, Anda perlu mengonversi alur pelatihan menjadi alur inferensi real time. Kemudian Anda dapat menyebarkan model sebagai aplikasi di server atau perangkat sehingga orang lain dapat menggunakannya.

Mari kita ikuti empat langkah ini saat muncul di perancang Azure.

1. **Siapkan data**

Perancang Azure Machine Learning memiliki beberapa komponen bawaan yang dapat digunakan untuk menyiapkan data untuk pelatihan. Komponen-komponen ini memungkinkan Anda membersihkan data, menormalkan fitur, menggabungkan tabel, dan banyak lagi.

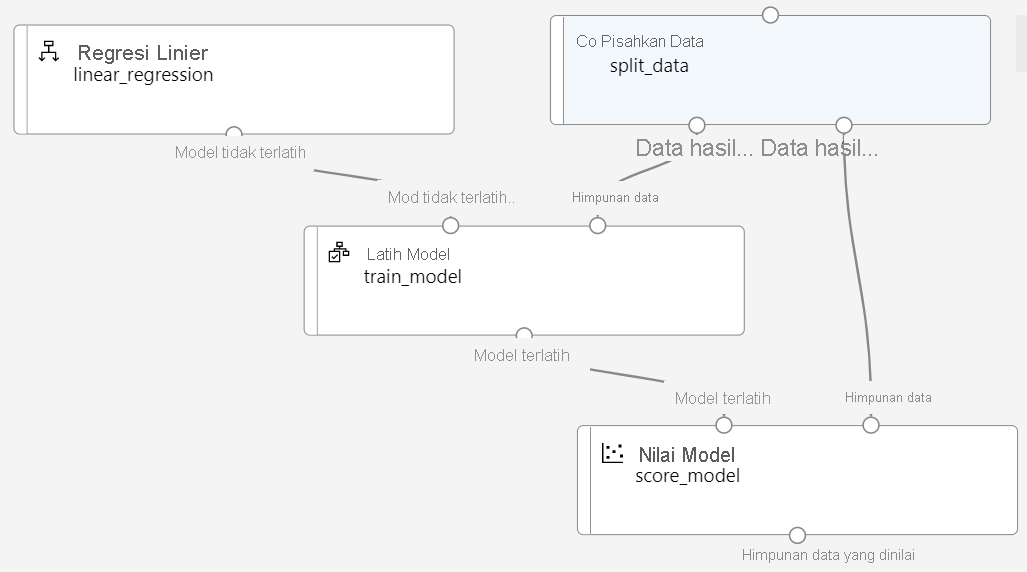


1. **Model latihan**

Untuk melatih model regresi, Anda memerlukan himpunan data yang menyertakan fitur historis, karakteristik entitas yang ingin Anda prediksi, dan nilai label yang diketahui. Label adalah kuantitas yang ingin Anda latih model untuk diprediksi.

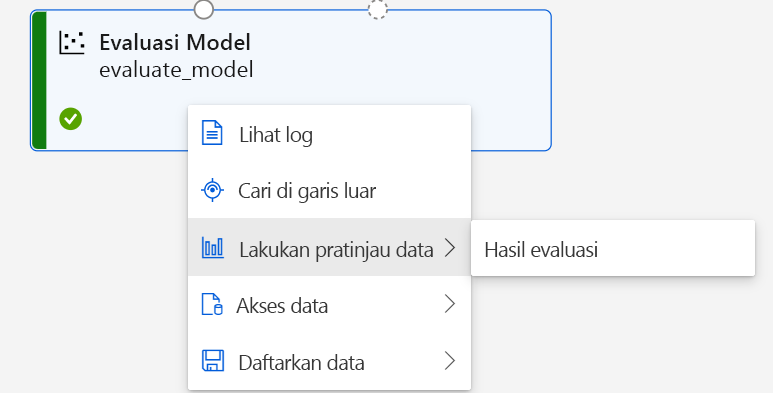
Melatih model menggunakan subset data, sambil menahan beberapa data untuk menguji model terlatih merupakan praktik yang umum. Ini memungkinkan Anda membandingkan label yang diprediksi model dengan label yang diketahui secara aktual dalam himpunan data asli.

Anda akan menggunakan komponen Model Skorperancang untuk menghasilkan nilai label kelas yang diprediksi. Setelah Anda menghubungkan semua komponen, Anda akan ingin menjalankan eksperimen, yang akan menggunakan aset data di kanvas untuk melatih dan menilai model.



1. **Evaluasi performa**

Setelah melatih model, penting untuk mengevaluasi performanya. Ada banyak metrik dan metodologi performa untuk mengevaluasi seberapa baik model membuat prediksi. Anda dapat meninjau metrik evaluasi pada halaman pekerjaan yang telah selesai dengan mengklik kanan komponen **Evaluasi model.**



* **Kesalahan Absolut Rata-Rata (MAE):** Perbedaan rata-rata antara nilai yang diprediksi dan nilai yang benar. Nilai ini didasarkan pada unit yang sama dengan label, dalam hal ini dolar. Semakin rendah nilai ini, semakin baik prediksi model.
* **Kesalahan Kuadrat Rata-Rata Akar (RMSE):** Akar kuadrat dari perbedaan kuadrat rata-rata antara nilai yang diprediksi dan yang benar. Hasilnya adalah metrik berdasarkan unit yang sama dengan label (dolar). Jika dibandingkan dengan MAE (di atas), perbedaan yang lebih besar menunjukkan variansi yang lebih besar dalam kesalahan individu (misalnya, dengan beberapa kesalahan menjadi sangat kecil, sementara yang lain besar).
* **Kesalahan Kuadrat Relatif (RSE):** Metrik relatif antara 0 dan 1 berdasarkan kuadrat perbedaan antara nilai yang diprediksi dan yang benar. Semakin metrik ini mendekati 0, semakin baik performa model. Karena bersifat relatif, metrik ini dapat digunakan untuk membandingkan model tempat label berada di unit yang berbeda.
* **Kesalahan Absolut Relatif (RAE):** Metrik relatif antara 0 dan 1 berdasarkan perbedaan absolut antara nilai yang diprediksi dan yang sebenarnya. Semakin metrik ini mendekati 0, semakin baik performa model. Seperti RSE, metrik ini dapat digunakan untuk membandingkan model tempat label berada di unit yang berbeda.
* **Koefisien Penentuan (R2):** Metrik ini lebih sering disebut sebagai R Kuadrat, dan meringkas berapa banyak variansi antara nilai yang diprediksi dan yang benar yang dijelaskan oleh model. Semakin nilai ini mendekati 1, semakin baik performa model.

1. **Menyebarkan layanan prediktif**

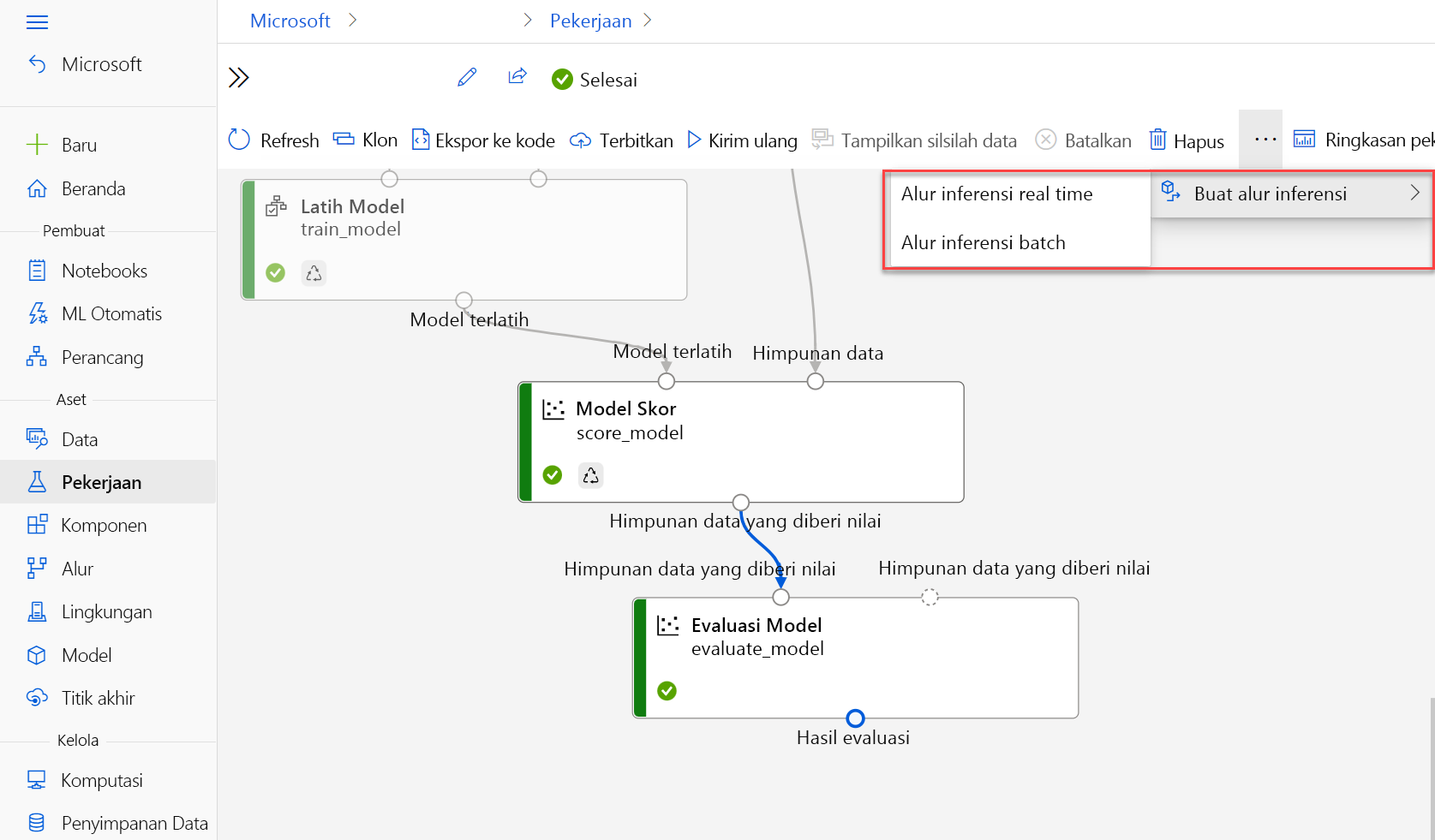
Anda memiliki kemampuan untuk menyebarkan layanan yang dapat digunakan secara real-time. Untuk mengotomatiskan model Anda ke dalam layanan yang membuat prediksi berkelanjutan, Anda perlu membuat dan menyebarkan alur inferensi.

1. **Alur inferensi**

Untuk menyebarkan alur, Anda harus terlebih dahulu mengonversi alur pelatihan menjadi alur inferensi real time. Proses ini akan menghapus komponen pelatihan serta menambahkan input dan output layanan web untuk menangani permintaan.

Alur inferensi melakukan transformasi data yang sama seperti alur pertama untuk data baru. Kemudian menggunakan model terlatih untuk menyimpulkan, atau memprediksi, nilai label berdasarkan fitur-fiturnya. Model ini akan menjadi dasar untuk layanan prediktif yang dapat Anda terbitkan untuk digunakan oleh aplikasi.

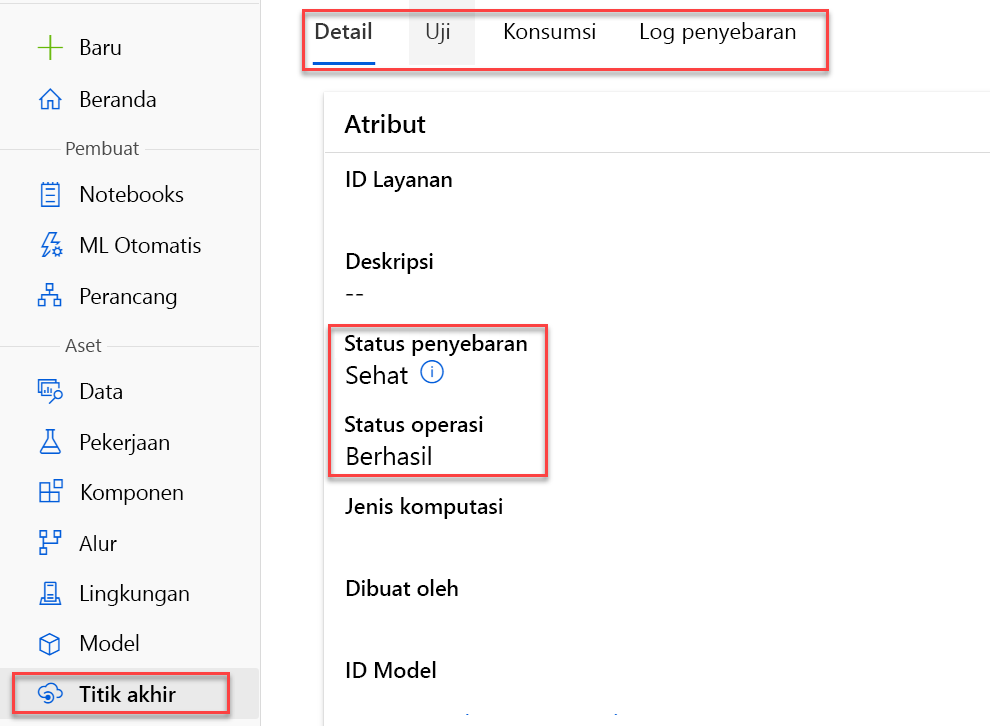
Anda dapat membuat alur inferensi dengan memilih menu di atas pekerjaan yang telah selesai.



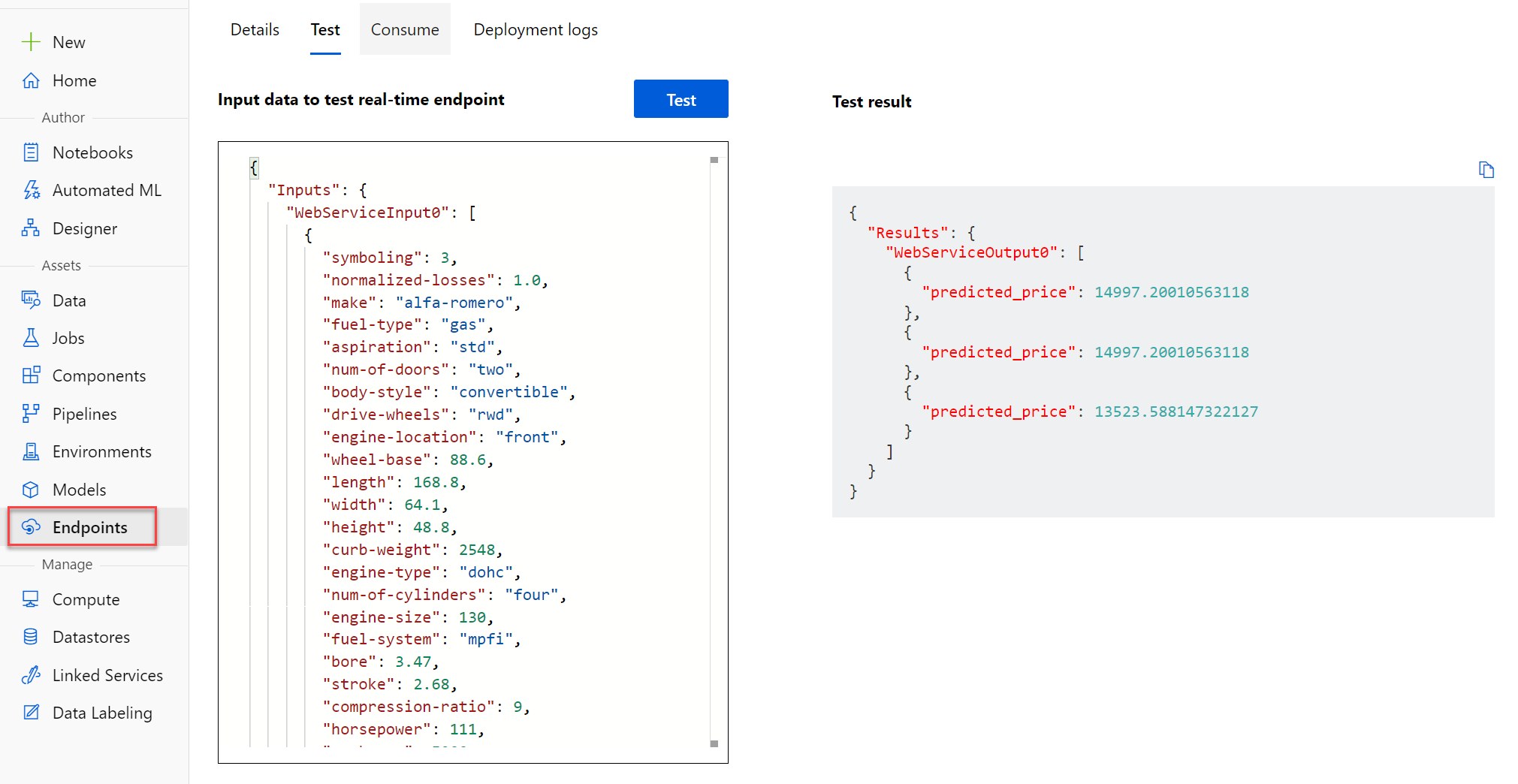
1. **Penyebaran**

Setelah membuat alur inferensi, Anda dapat menyebarkannya sebagai titik akhir. Di halaman titik akhir, Anda dapat melihat detail penyebaran, menguji layanan alur Anda dengan data sampel, dan menemukan info masuk untuk menyambungkan layanan alur Anda ke aplikasi klien.

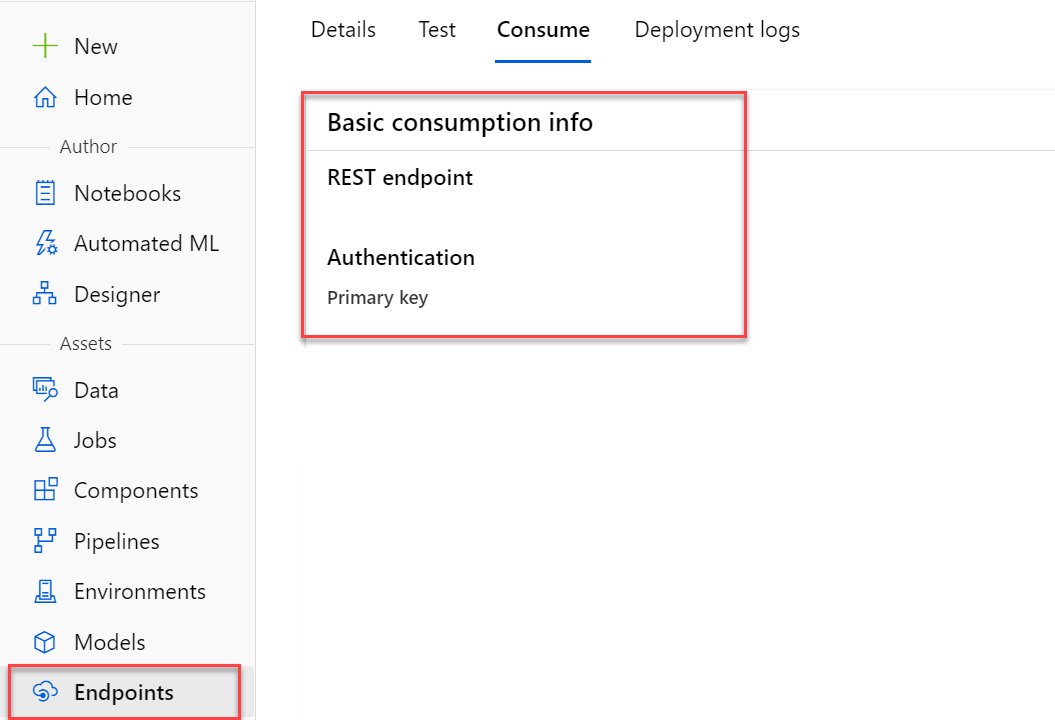
Ini akan memakan waktu beberapa saat agar titik akhir Anda disebarkan. Status Penyebaran pada tab **Detail** akan menunjukkan keadaan Sehat saat penyebaran berhasil.



Pada tab **Pengujian**, Anda dapat menguji layanan yang disebarkan dengan data sampel dalam format JSON. Tab pengujian adalah alat yang dapat Anda gunakan untuk memeriksa dengan cepat untuk melihat apakah model Anda berperilaku seperti yang diharapkan. Biasanya alat ini sangat membantu untuk menguji layanan sebelum menghubungkannya ke aplikasi.



Anda dapat menemukan info masuk untuk layanan Anda di tab **Pemakaian**. Info masuk ini digunakan untuk menghubungkan model pembelajaran mesin terlatih Anda sebagai layanan ke aplikasi klien.



**6. Latihan - Menjelajahi regresi dengan perancang Azure Machine Learning**

Dalam latihan ini, Anda akan **melatih model regresi yang memprediksi harga mobil berdasarkan karakteristiknya.**

**Catatan :**

Untuk menyelesaikan lab ini, Anda memerlukan langganan Azure tempat Anda memiliki akses administratif.

**Link :**

<https://microsoftlearning.github.io/AI-900-AIFundamentals/instructions/02a-create-regression-model.html>

1. **Ringkasan**

Dalam modul ini, Anda **belajar cara**:

* Identifikasi skenario pembelajaran mesin regresi.
* Gunakan perancang Azure Machine Learning untuk melatih model regresi.
* Gunakan model regresi untuk menyimpulkan.
* Terapkan model regresi sebagai layanan.