Assignment 7: optional catchup assignment 2 - VERTEX AI - for midterm and quiz - this will catch up midterm.

Time Series Forecasting with Vertex AI and BigQuery ML
Reference: https://codelabs.developers.google.com/codelabs/time-series-forecasting-with-cloud-ai-platform#0

Objectives:

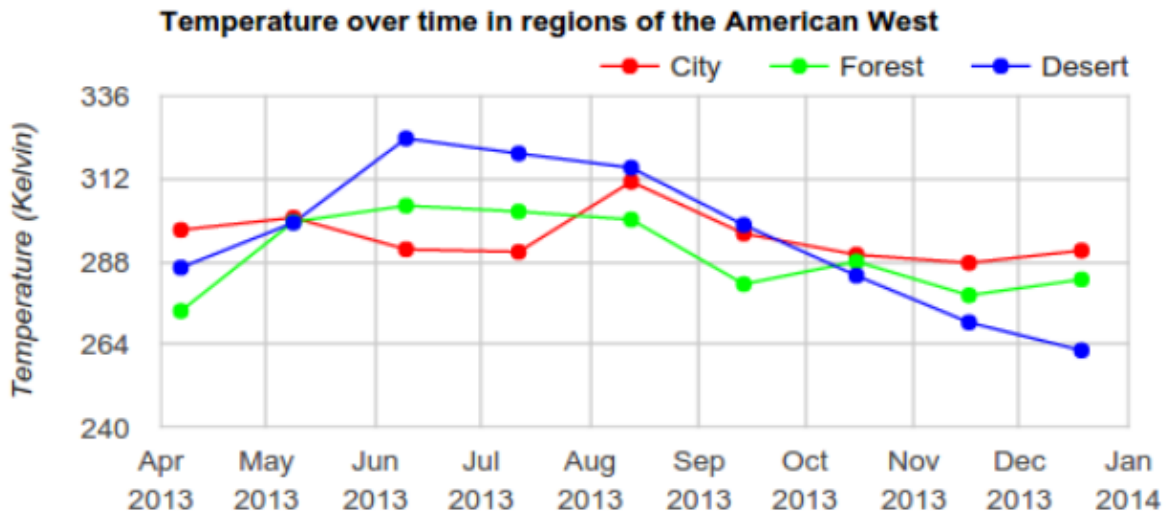
- Transform data so that it can be used in an ML model
- Visualize and explore data
- Use BigQuery ML to create a time-series forecasting model
- Build a time-series forecasting model with TensorFlow using LSTM and CNN architectures

Introduction to Time-Series Forecasting

The focus of this codelab is on how to *apply* time-series forecasting techniques using the Google Cloud Platform

Time Series Data

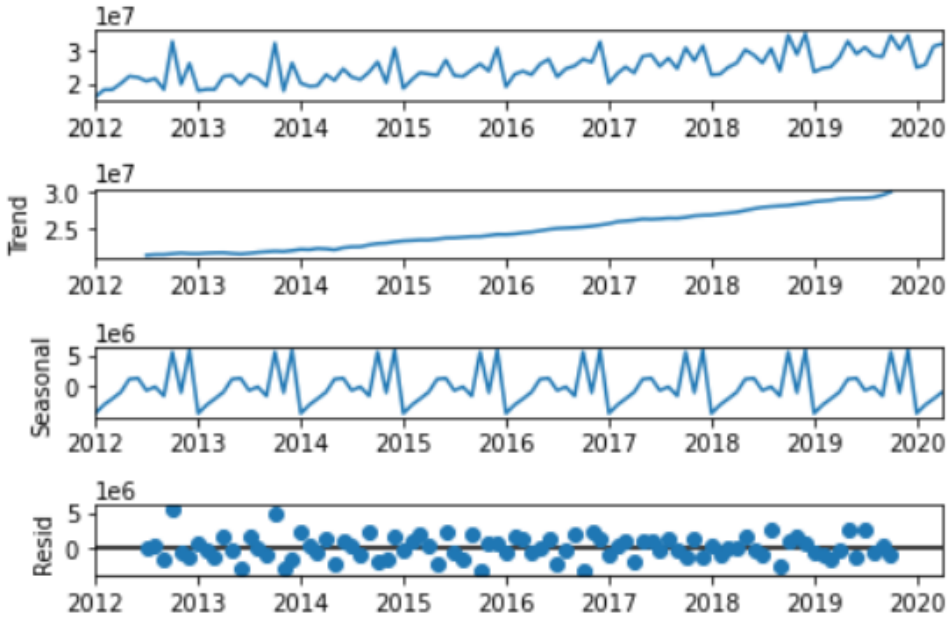
It's a dataset with data recorded at regular time intervals. A time-series dataset contains both time and at least one variable that is dependent on time.



Components

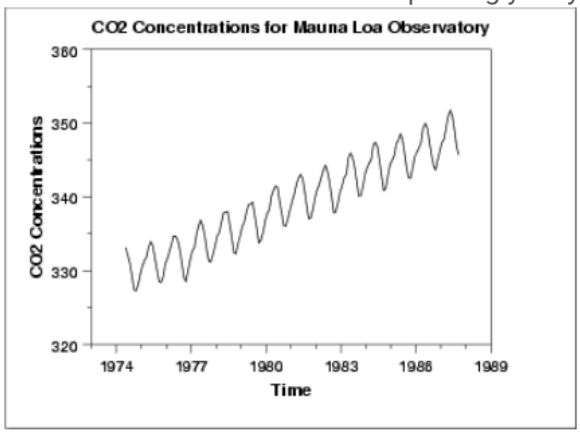
A time-series can be decomposed into components:

- **Trend**: moves up or down in a reasonably predictable pattern
- **Seasonal**: repeats over a specific period such as a day, week, month, season, etc.
- **Random**: residual fluctuations

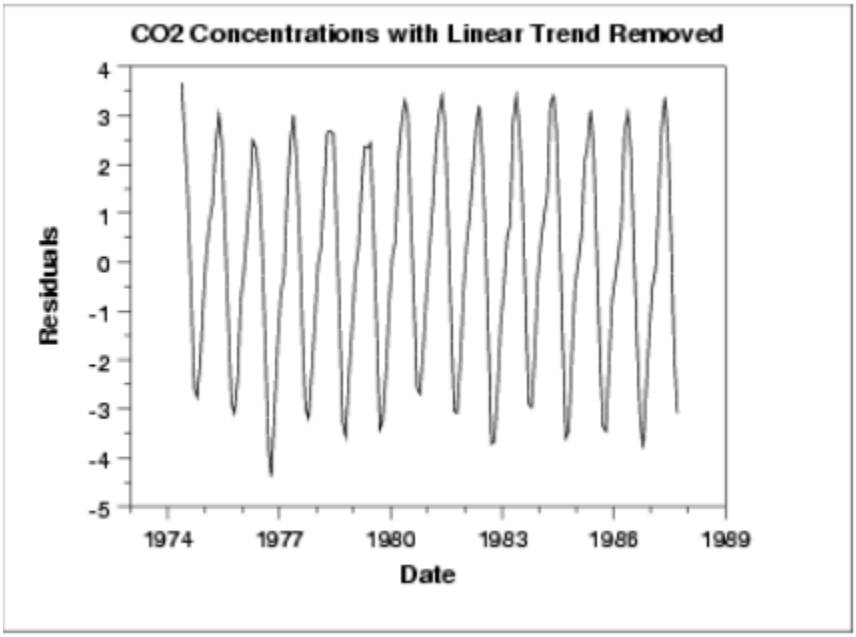


Stationarity

For best results in forecasting, time-series data should be made stationary, where statistical properties such as mean and variance are constant over time. Techniques such as differencing and detrending can be applied to raw data to make it more stationary.
The plot below of CO2 concentration shows a repeating yearly pattern with an upward trend



After removing the linear trend, the data is more suitable for forecasting, as it now has a constant mean.



Using Time Series Data for Machine Learning

To use time-series data in a machine learning problem, it needs to be transformed so that previous values can be used to predict future values. This table shows an example of how lagged variables are created to help predict the target.

Original Data

Date	Value
Jan/2000	11
Feb/2000	12
Mar/2000	13
Apr/2000	14
May/2000	15
Jun/2000	16
Jul/2000	16
Aug/2000	15
Sep/2000	14
Oct/2000	13
Nov/2000	12
Dec/2000	11

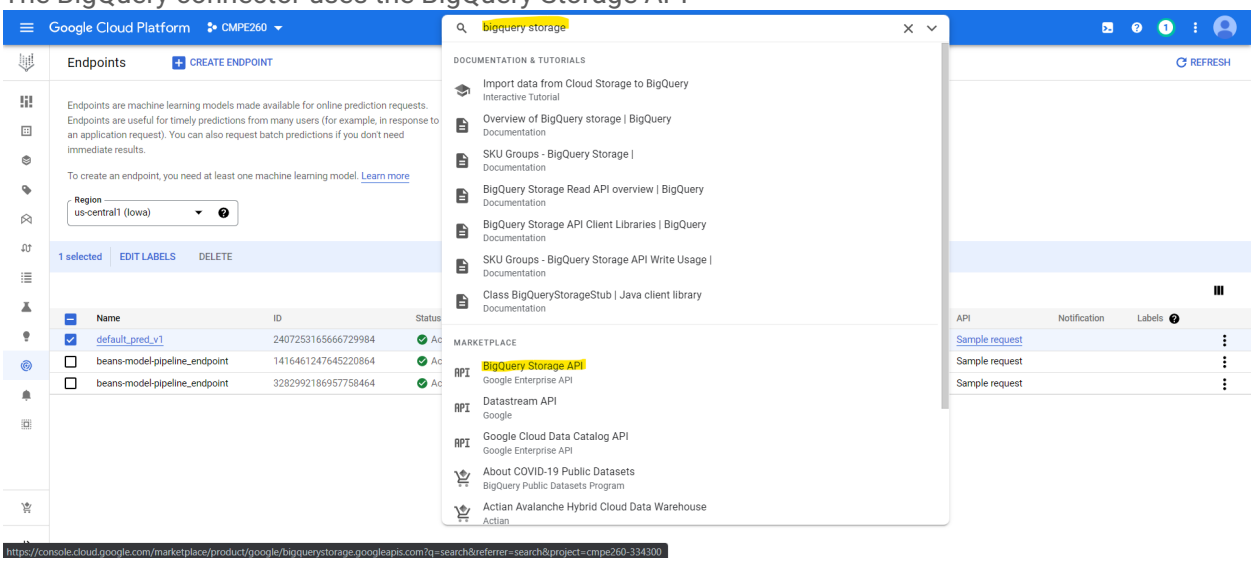
LAGGED VARIABLES (FEATURES) TARGET

Date	t-2	t-1	Actual
Jan/2000			11
Feb/2000		11	12
Mar/2000	11	12	13
Apr/2000	12	13	14
May/2000	13	14	15
Jun/2000	14	15	16
Jul/2000	15	16	16
Aug/2000	16	16	15
Sep/2000	16	15	14
Oct/2000	15	14	13
Nov/2000	14	13	12
Dec/2000	13	12	11

Setup your Notebook environment

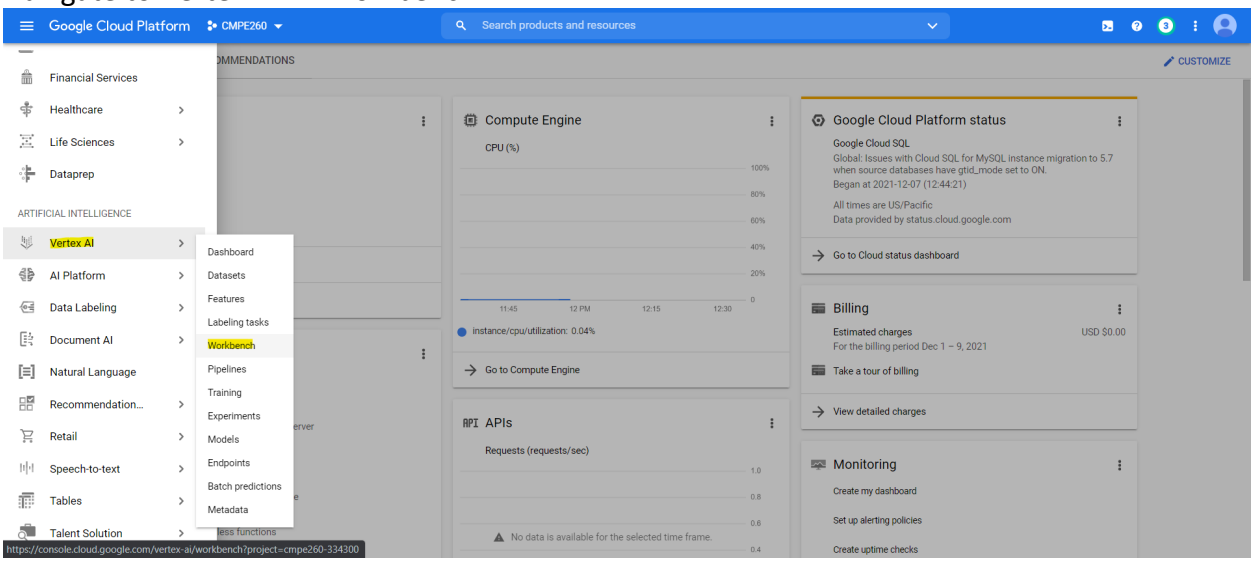
Step 1: Enable APIs

The BigQuery connector uses the BigQuery Storage API

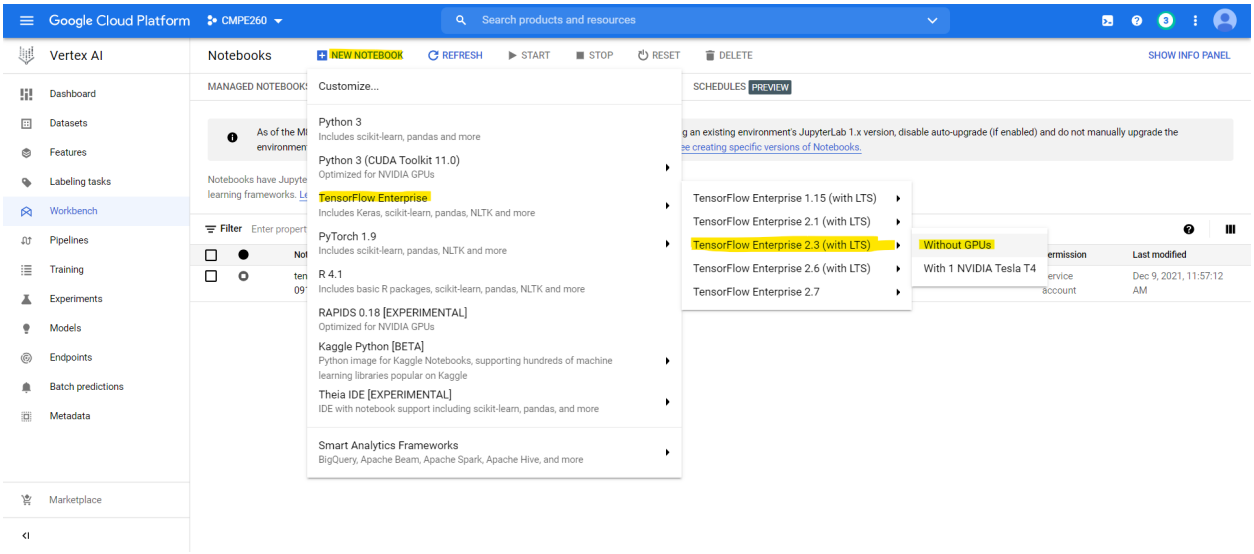


Step 2: Create a Vertex AI Workbench notebook

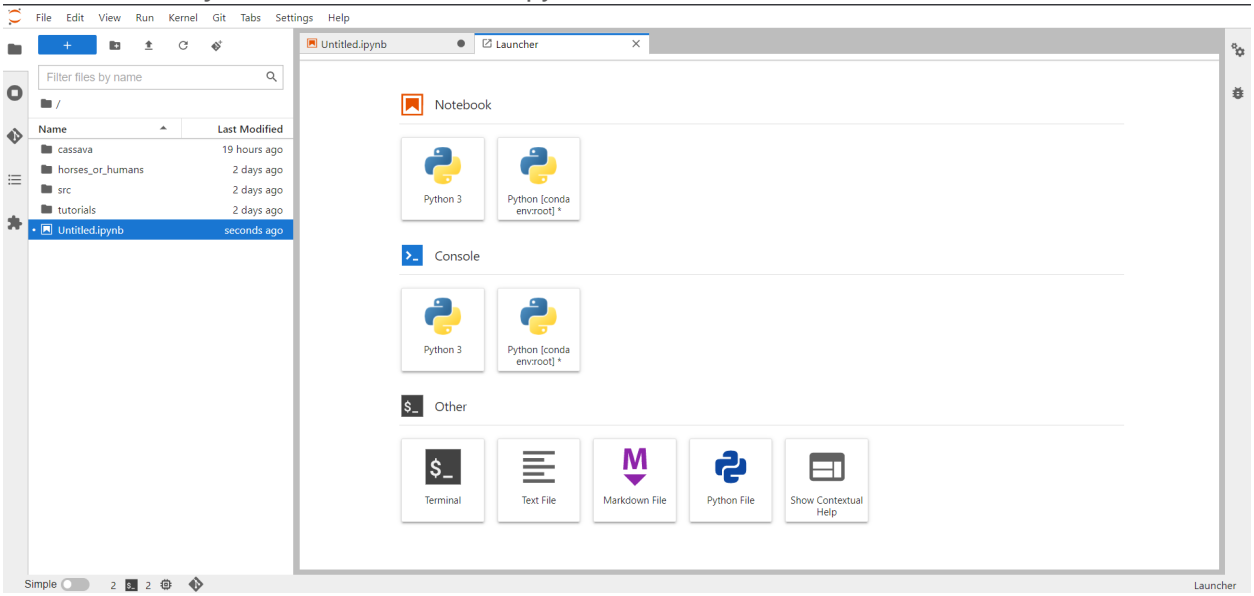
Navigate to Vertex AI → Workbench



Select **TensorFlow Enterprise 2.3 (with LTS)** instance type **without GPUs**:



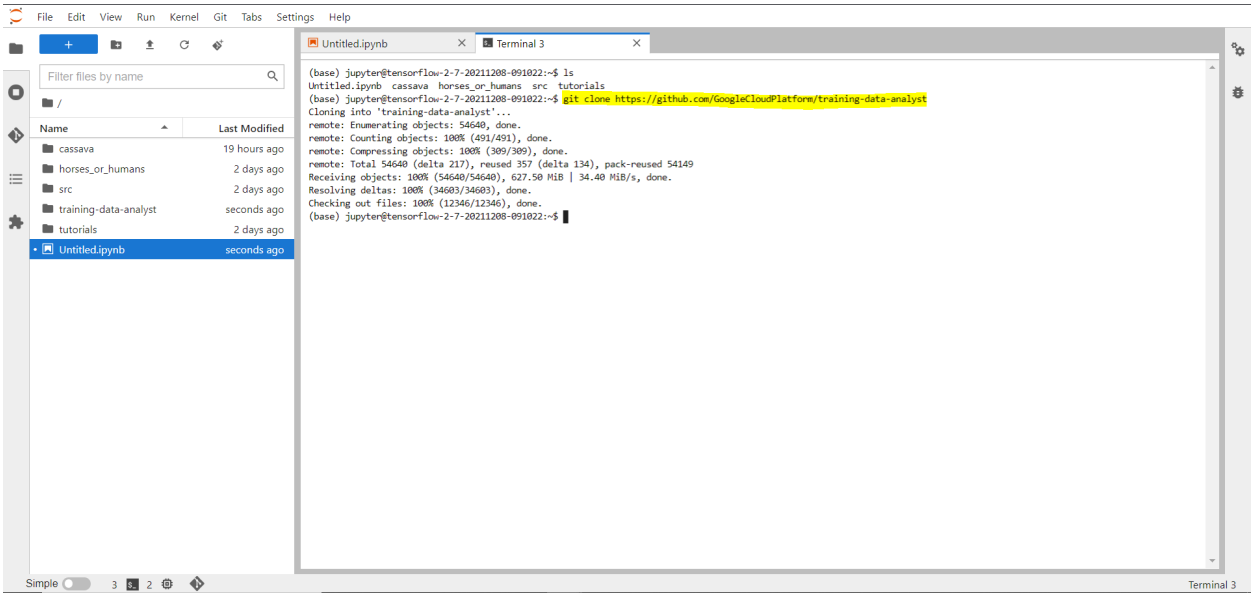
Then, create a **Python 3** notebook from JupyterLab:



Step 3: Download lab materials

Run the following command in the Terminal:

git clone https://github.com/GoogleCloudPlatform/training-data-analyst



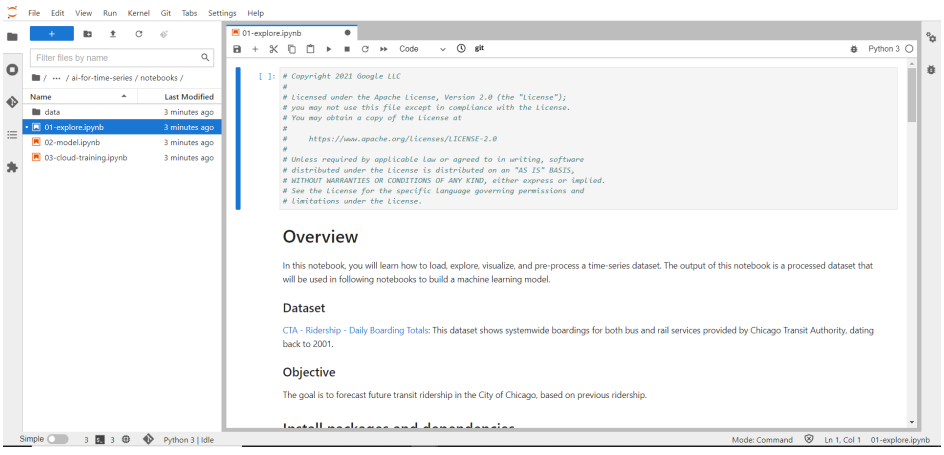
Explore and Visualize Data

The following are the tasks in this section:

- Create a query that groups data into a time-series
- Fill missing values
- Visualize data
- Decompose time-series into trend and seasonal components

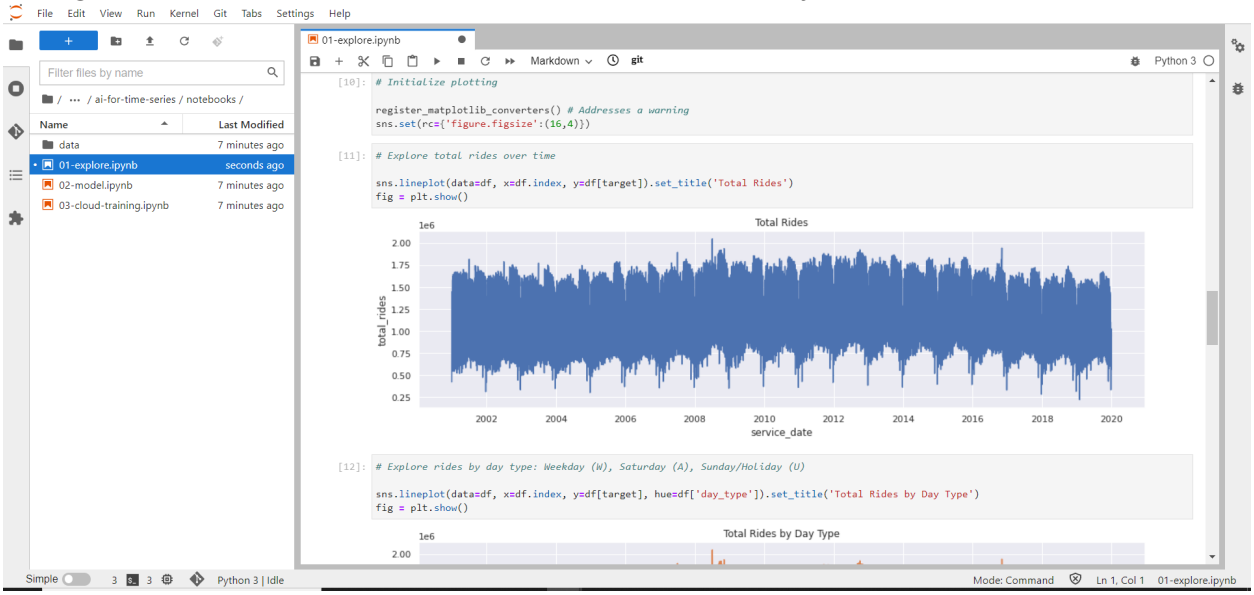
Step 1

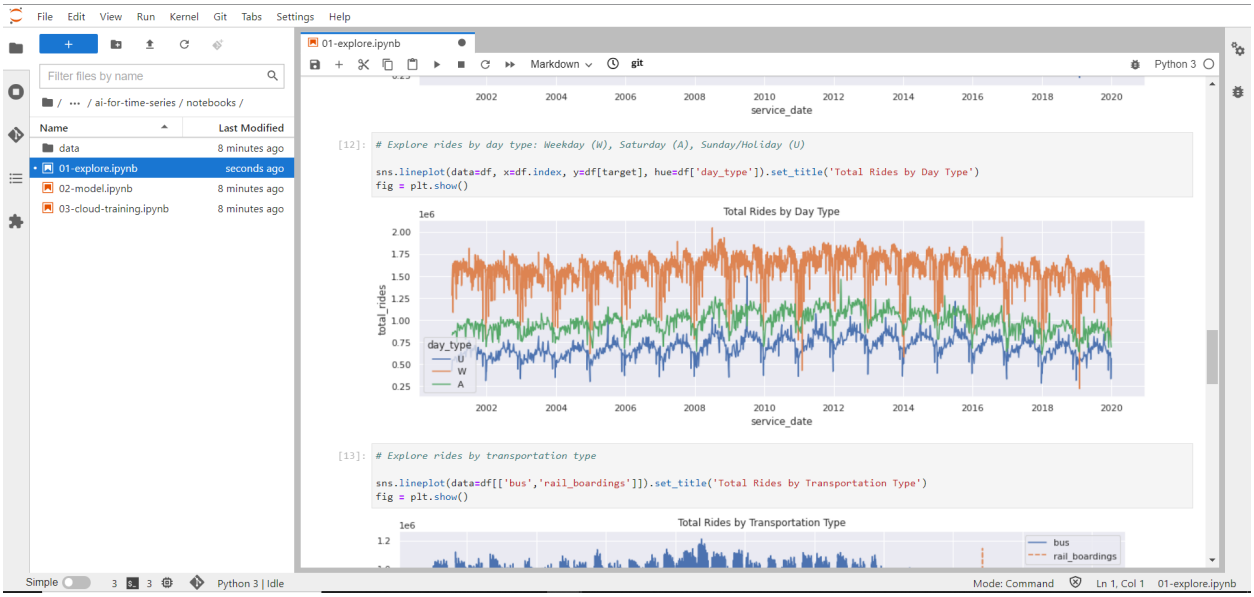
In Vertex AI Workbench, navigate to `training-data-analyst/courses/ai-for-time-series/notebooks` and open `01-explore.ipynb`.

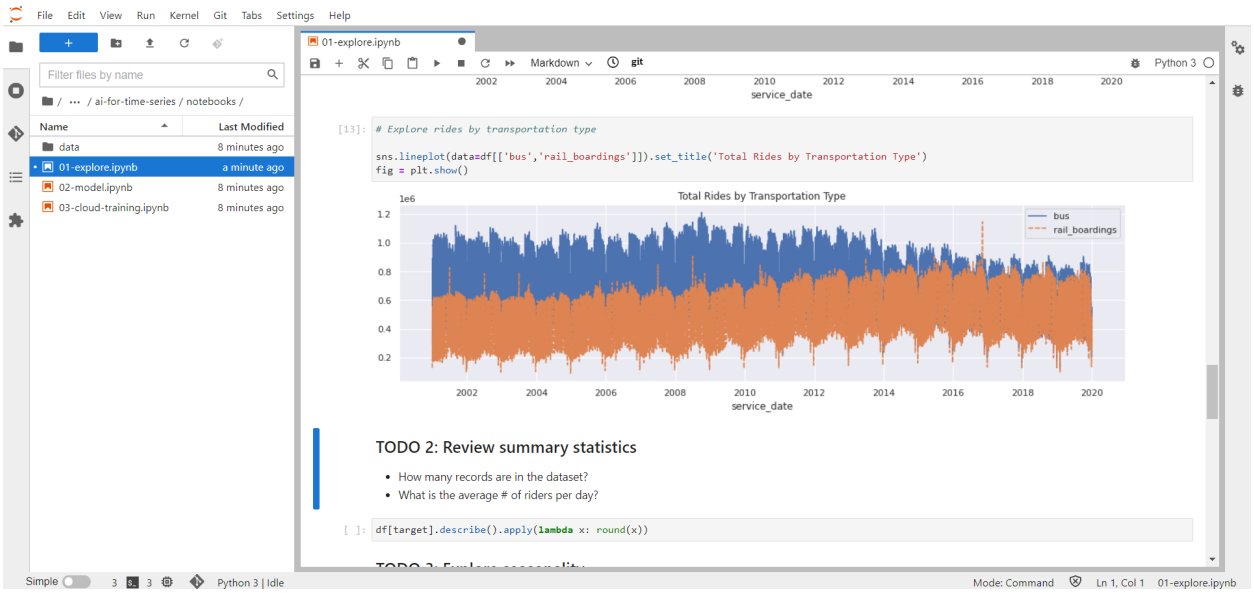


Step 2

Clear all the cells in the notebook (Edit > Clear All Outputs), change the region, project and bucket settings in one of the first few cells, and then Run the cells one by one.







Create a Model with BigQuery Time Series Forecasting

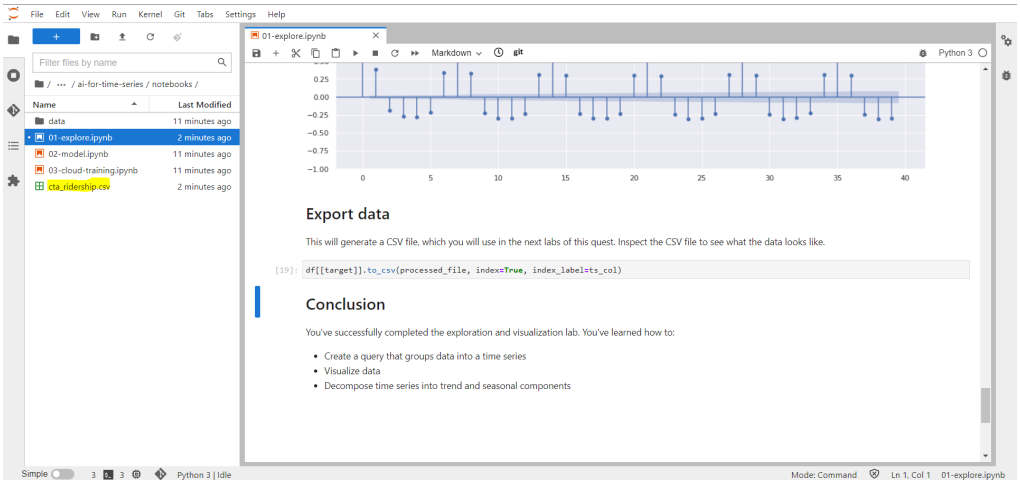
Following are the tasks in this section:

- Import your time series input data into a BigQuery table
- Create a time series model using BQML syntax
- Learn how to evaluate your model parameters and accuracy
- Forecast using your model

Step 1

We are going to create a BigQuery table with the raw data from the CSV we just explored

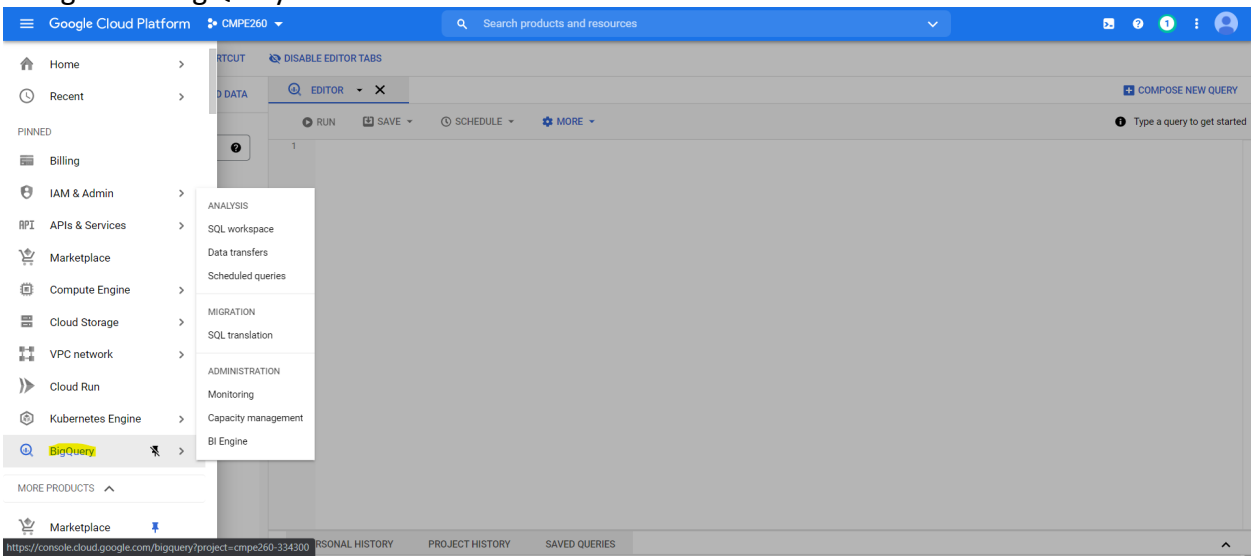
From the `training-data-analyst/courses/ai-for-time-series/notebooks/data` directory, right-click on `cta_ridership.csv` and **Download** it to your local environment.



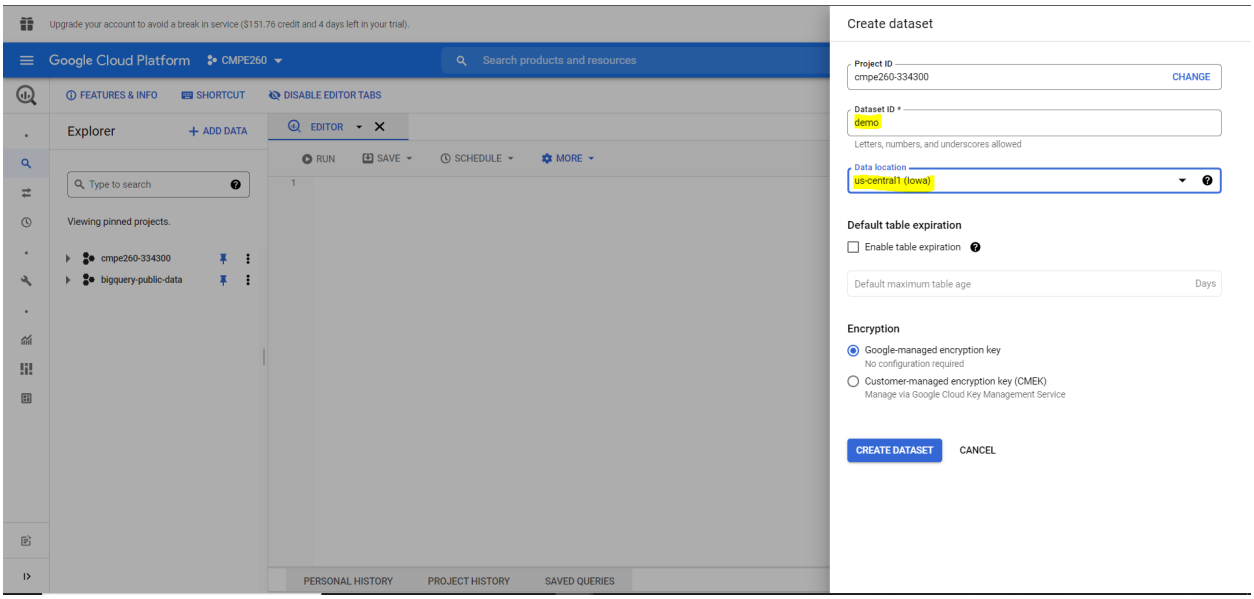
Step 2

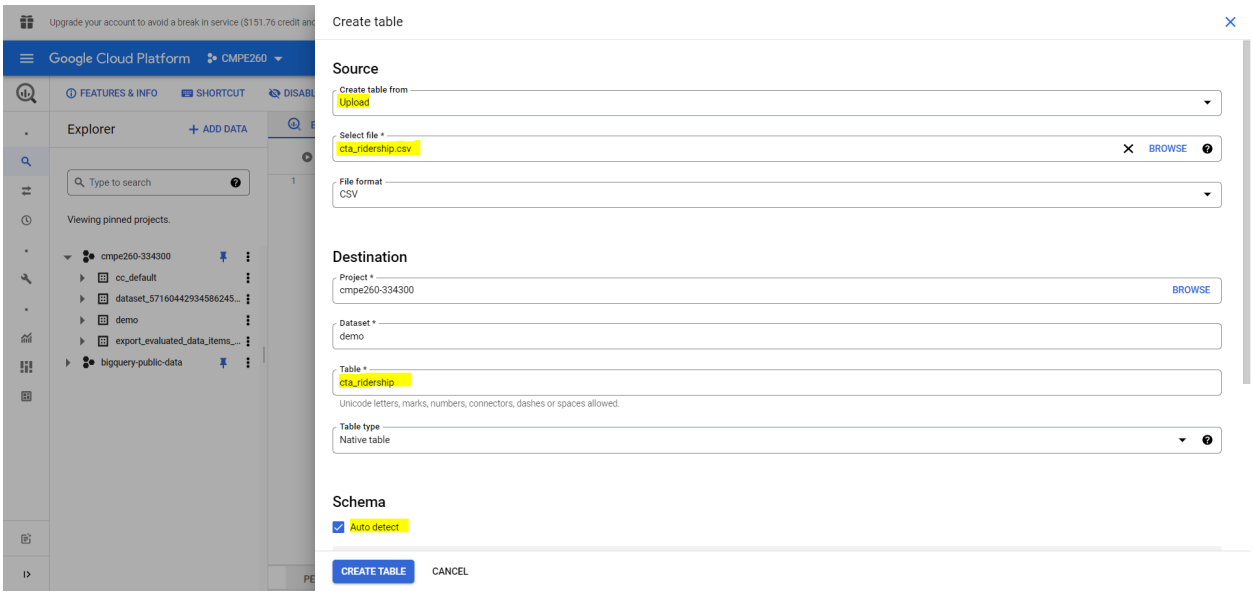
We will upload this data into a BigQuery table.

Navigate to ‘BigQuery’



Create a Dataset ‘demo’:





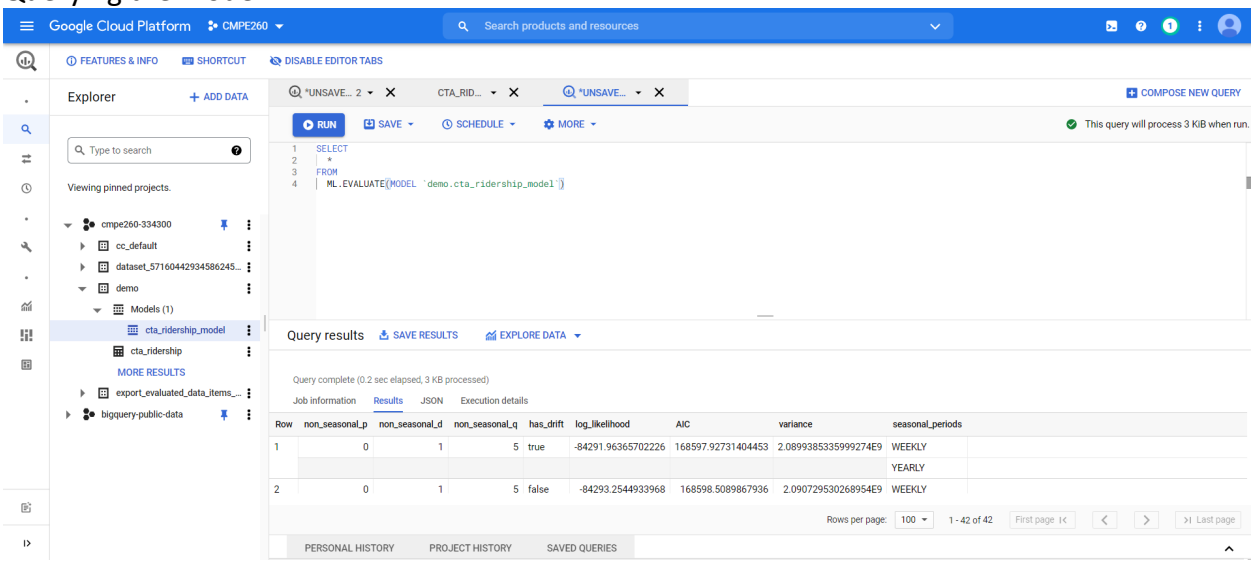
Step 3

Create our model

Run the below query in the query editor

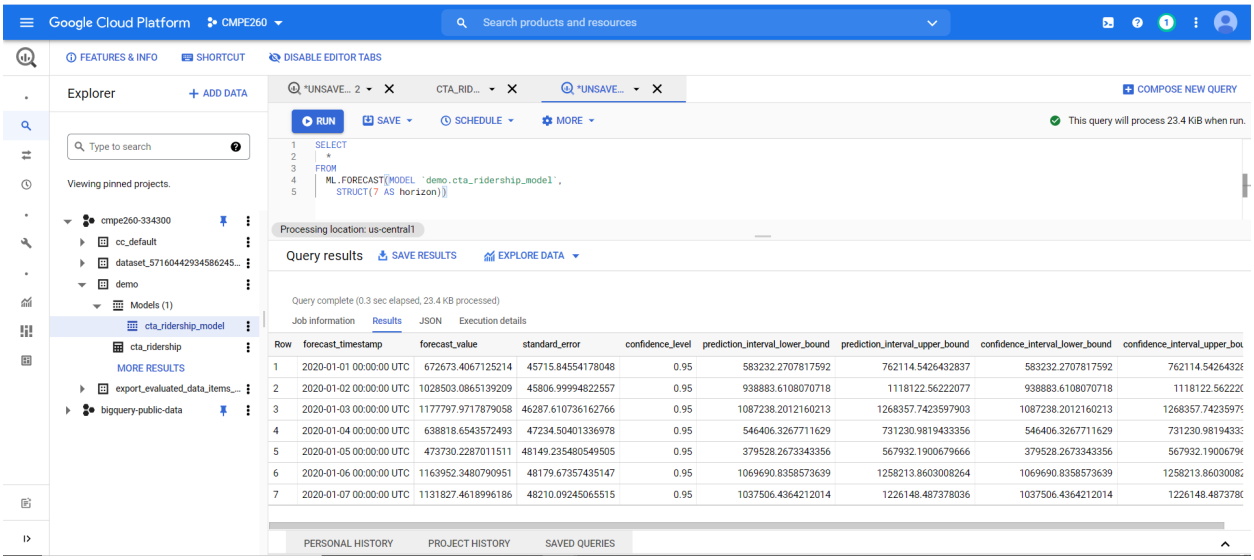
Step 4

Querying the model



Step 5

Ready to forecast with the ML.FORECAST function

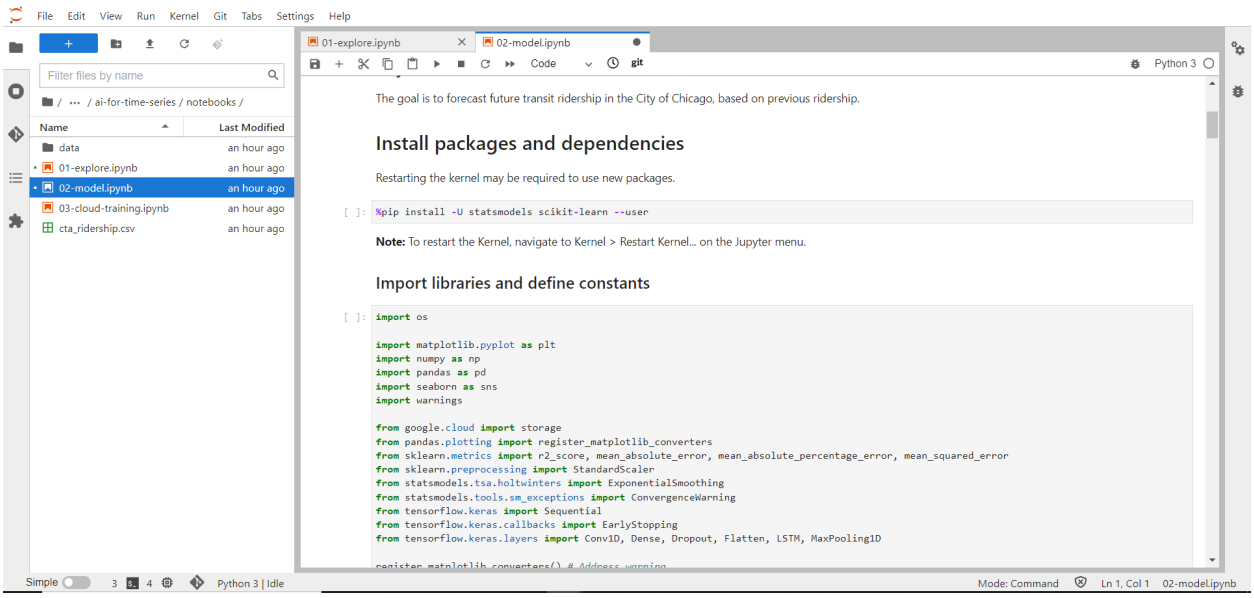


Build a Custom Forecasting Model

- Remove outliers from the data
- Perform multi-step forecasting
- Include additional features in a time-series model
- Learn about neural network architectures for time-series forecasting: LSTM and CNN
- Learn about statistical models, including Holt-Winters Exponential Smoothing
- Ensemble models

Step 1

In Vertex AI Workbench, navigate to `training-data-analyst/courses/ai-for-time-series/notebooks` and open `02-model.ipynb`.



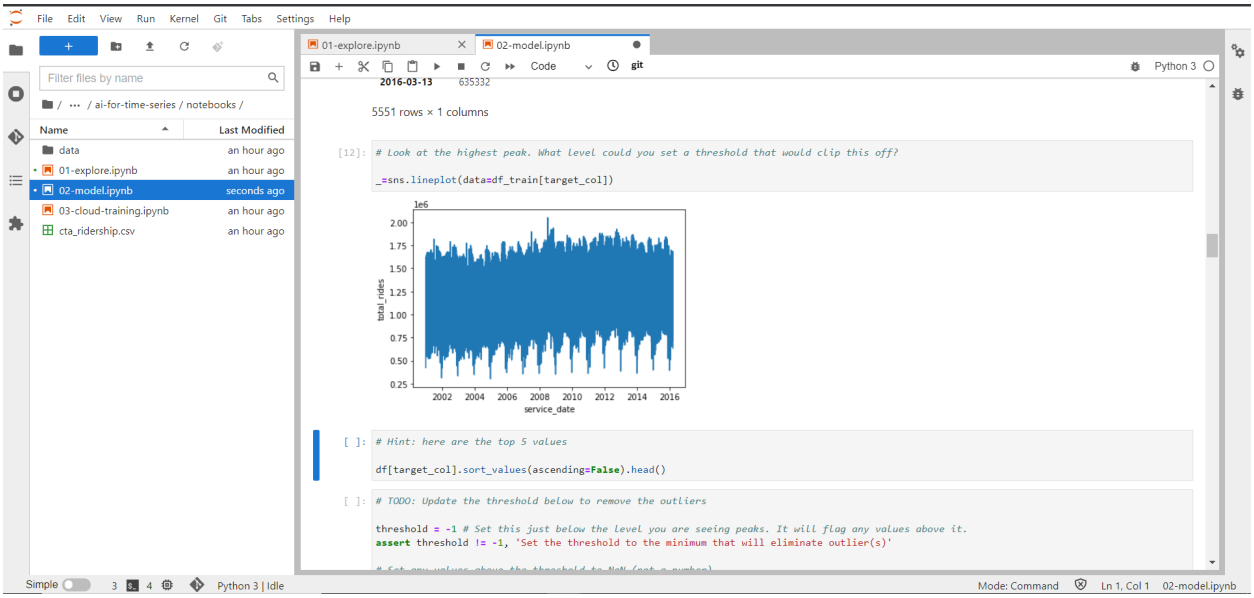
Step 2

Clear all the cells in the notebook (Edit > Clear All Outputs)

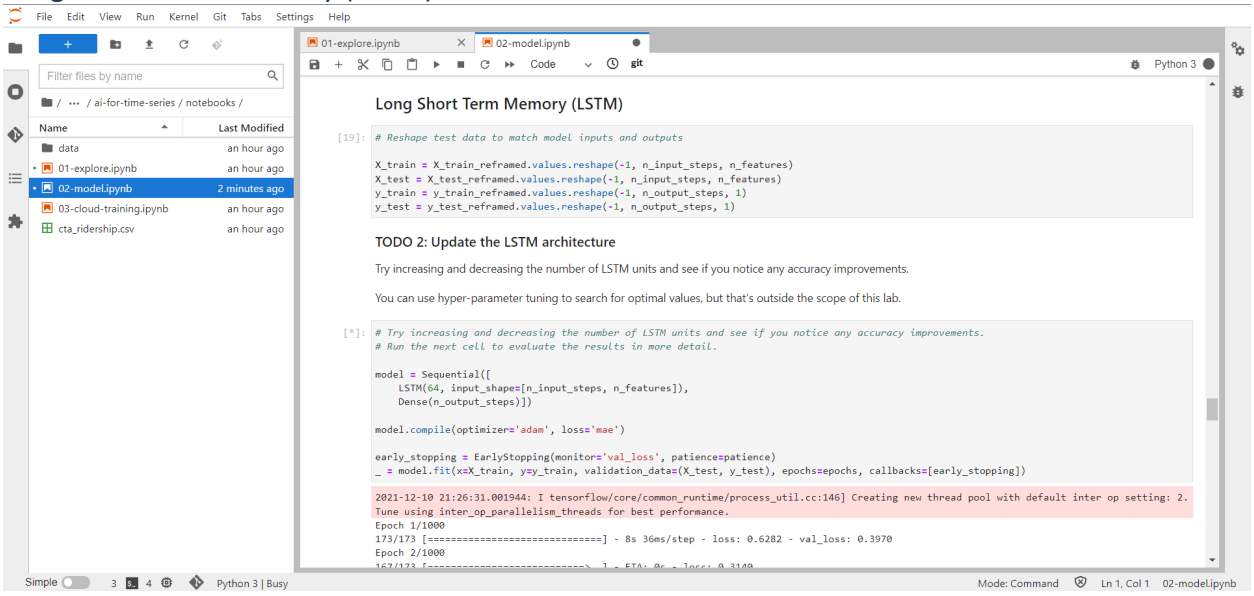
Step 3

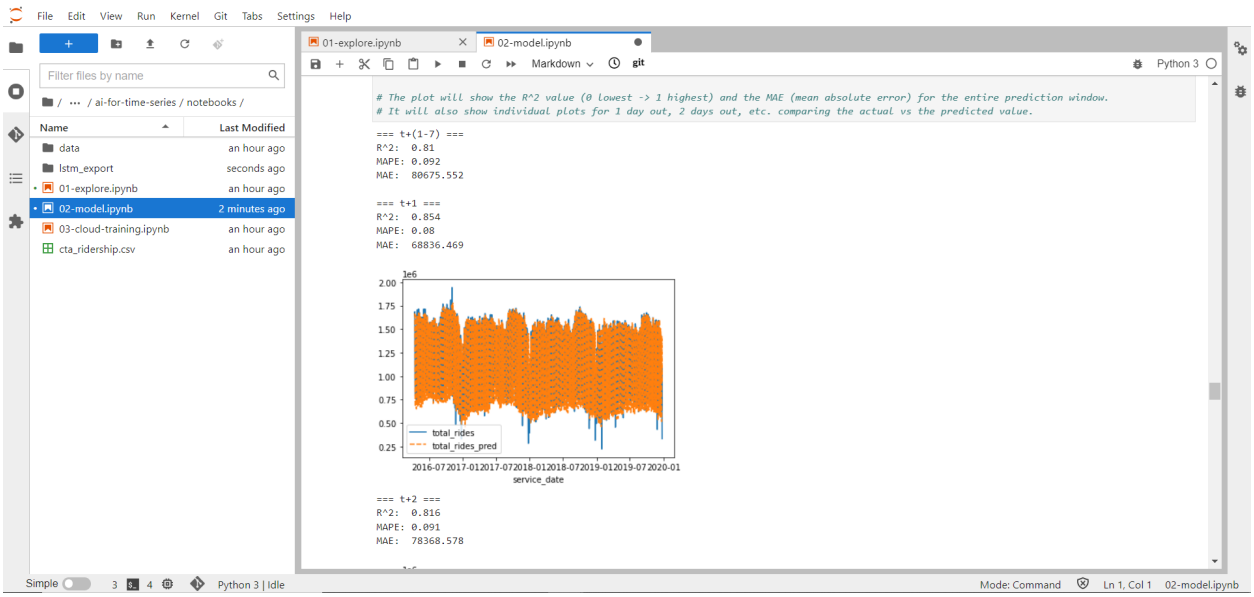
Then Run the cells one by one

Results:

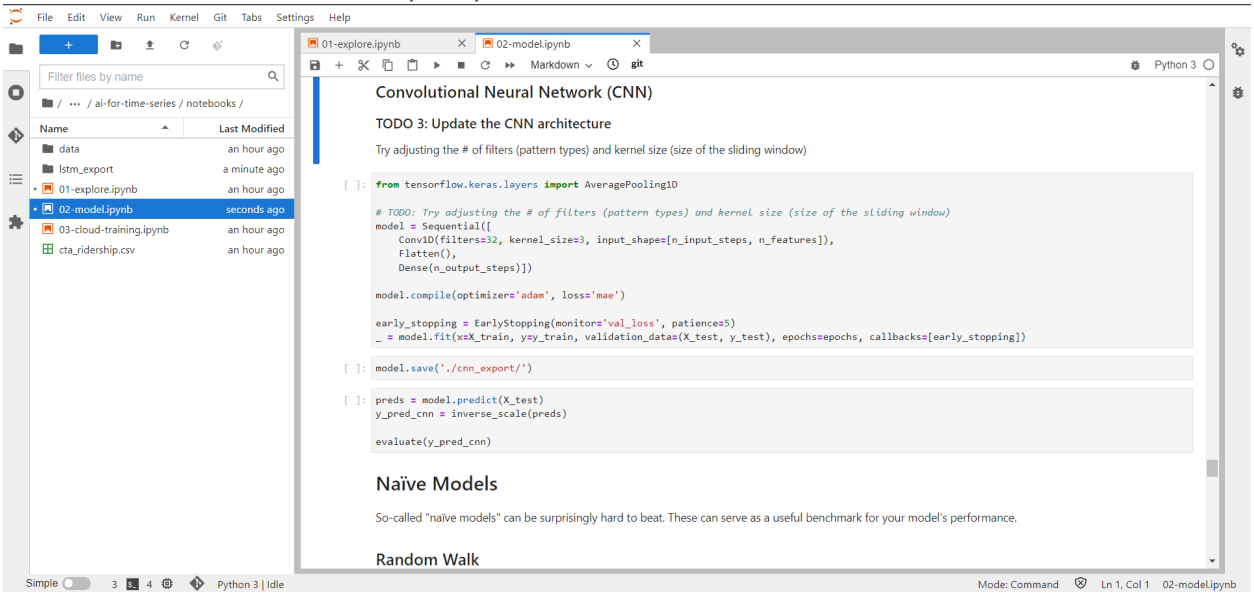


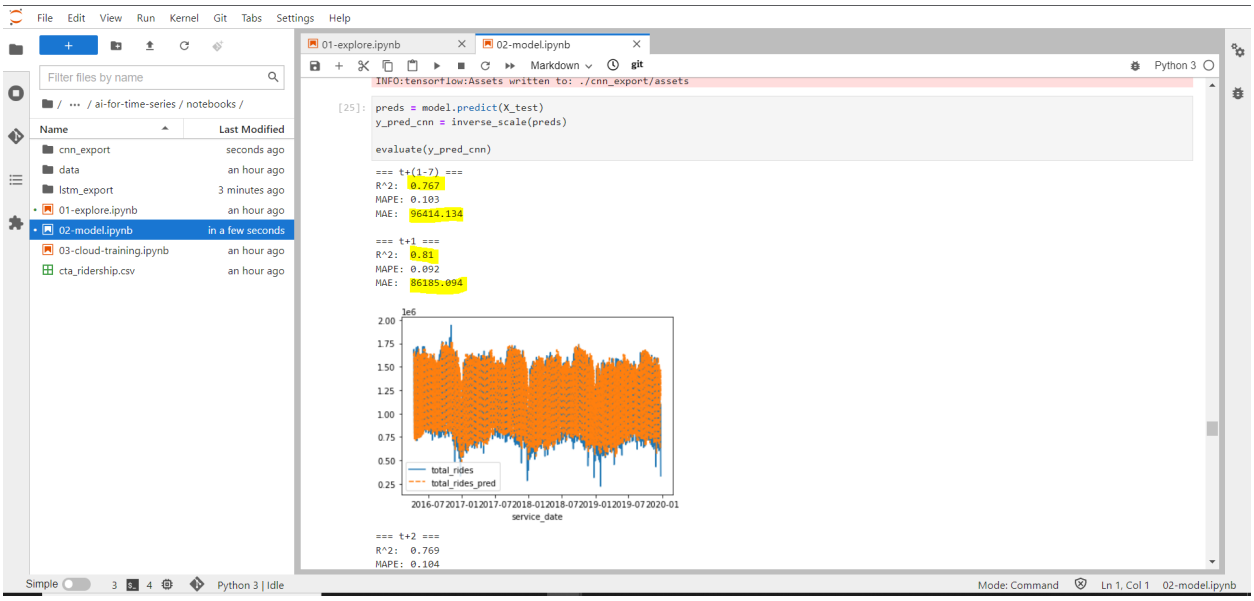
Long Short Term Memory (LSTM)



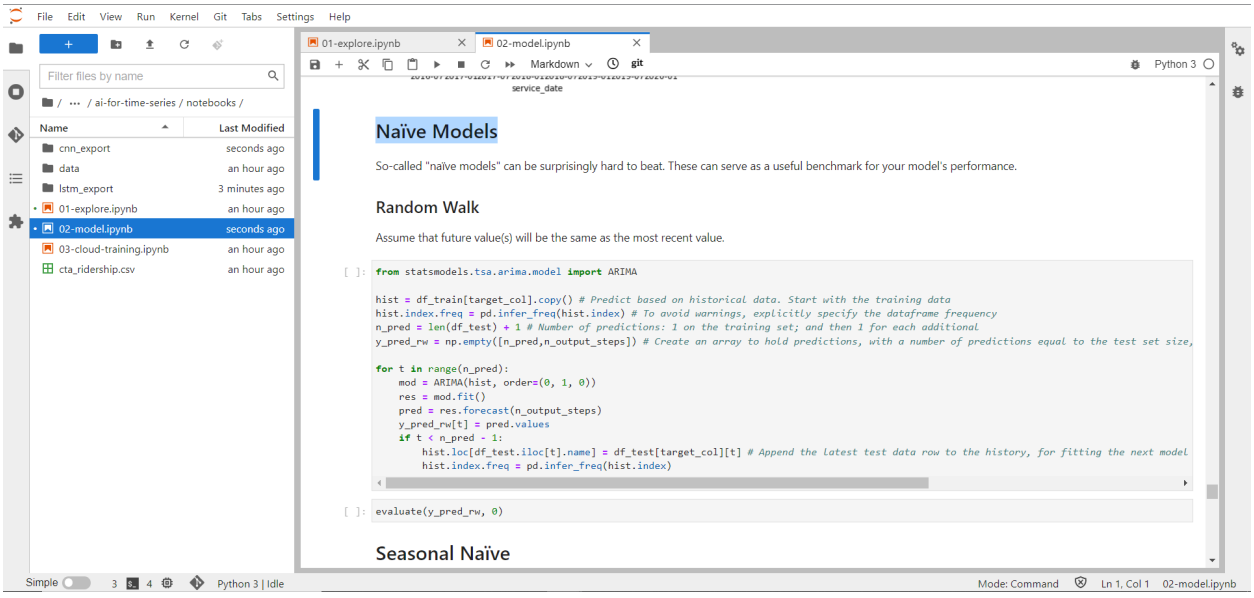


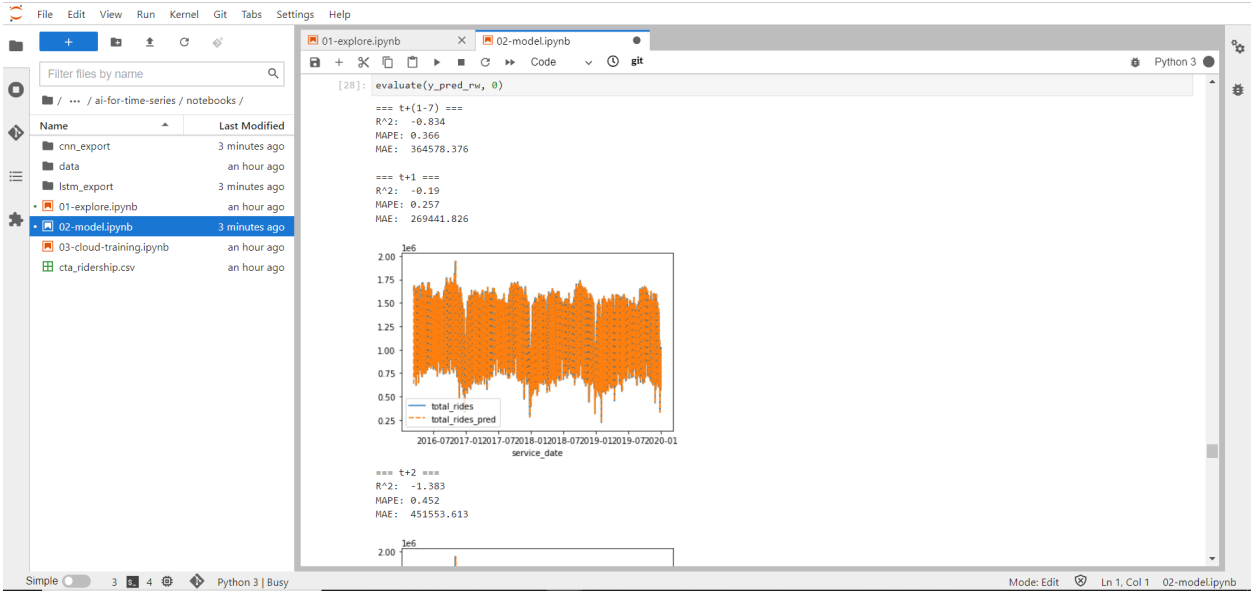
Convolutional Neural Network (CNN)



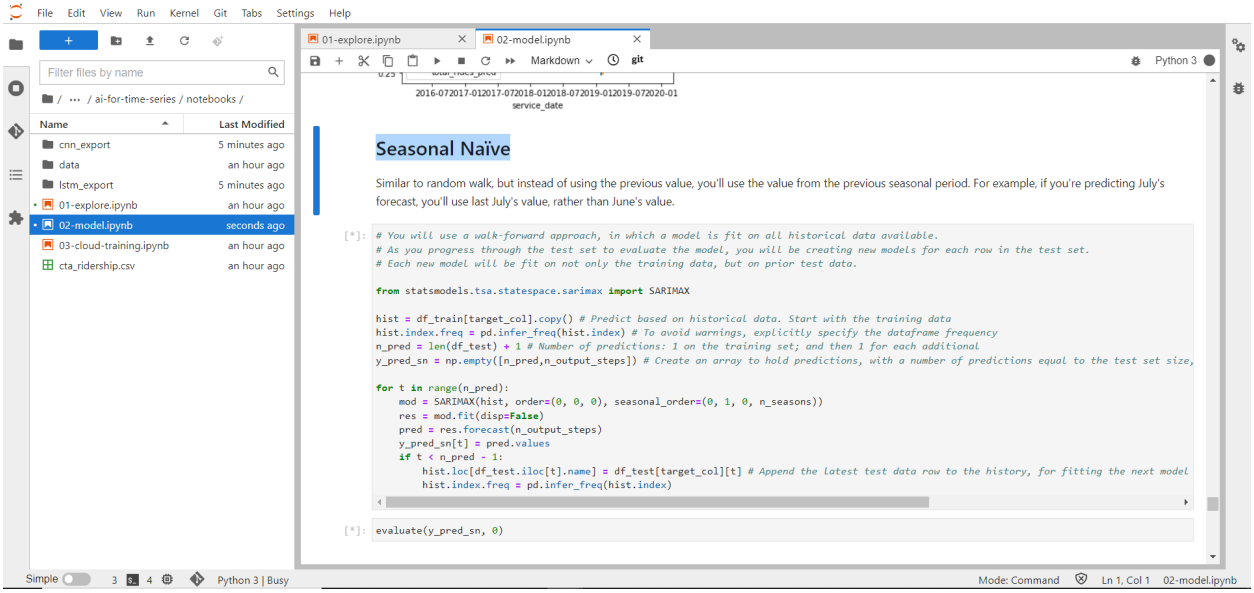


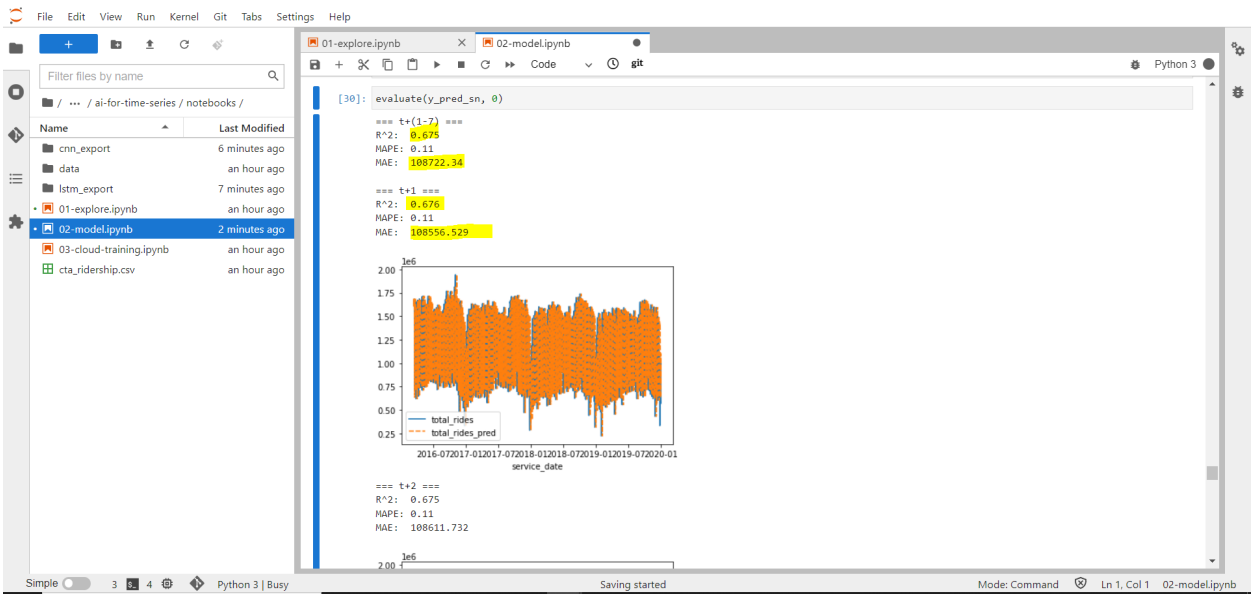
Naïve Models



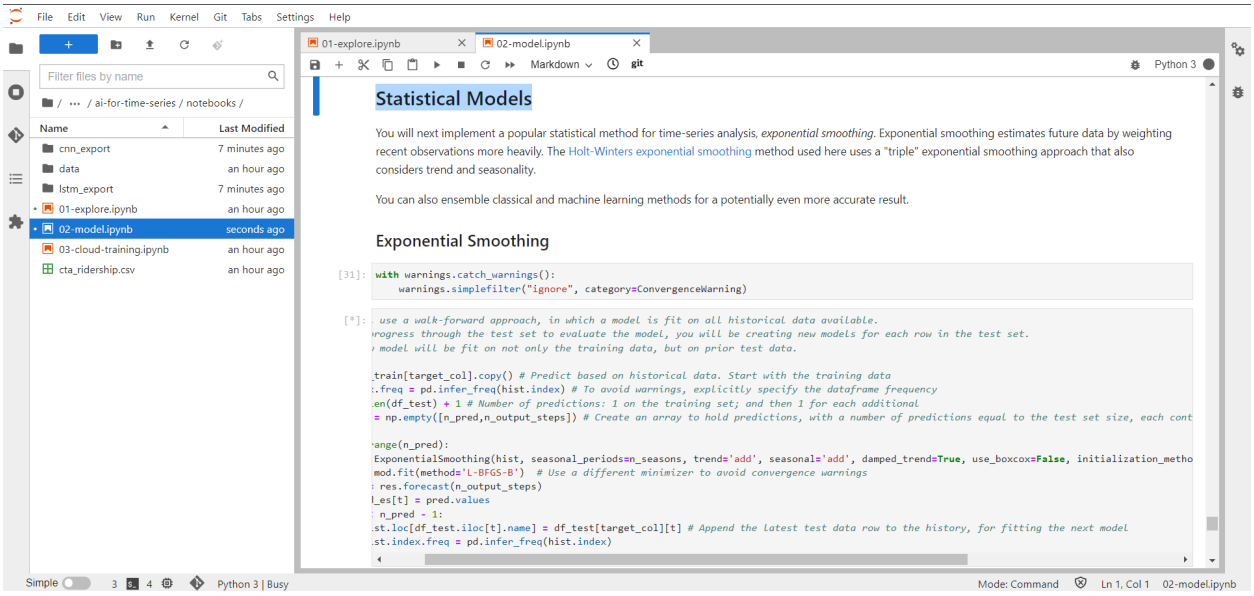


Seasonal Naïve





Statistical Models

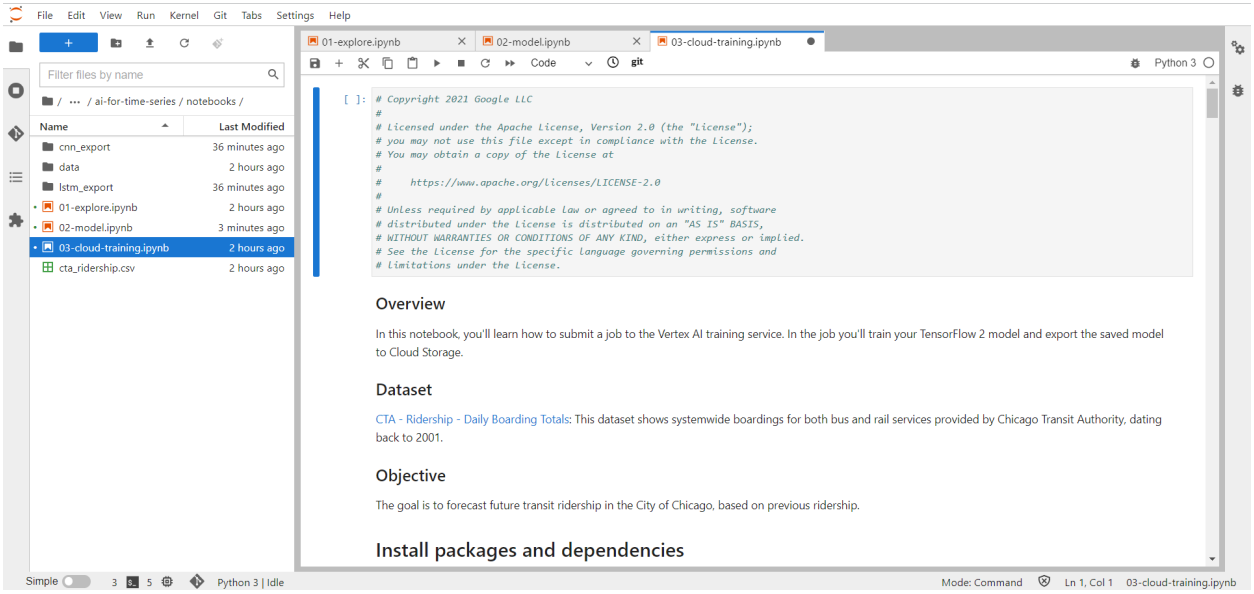


Train and Predict in the Cloud

- Prepare data and models for training in the cloud
- Train your model and monitor the progress of the job with AI Platform Training
- Predict using the model with AI Platform Prediction

Step 1

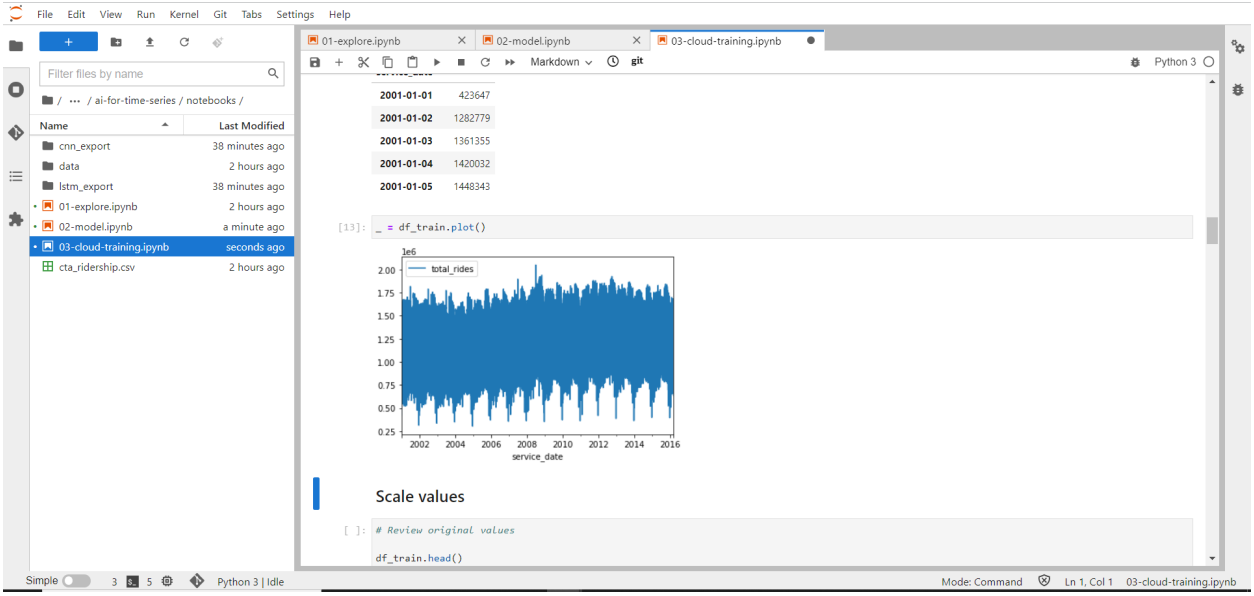
In Vertex AI Workbench, navigate to `training-data-analyst/courses/ai-for-time-series/notebooks` and open `03-cloud-training.ipynb`.



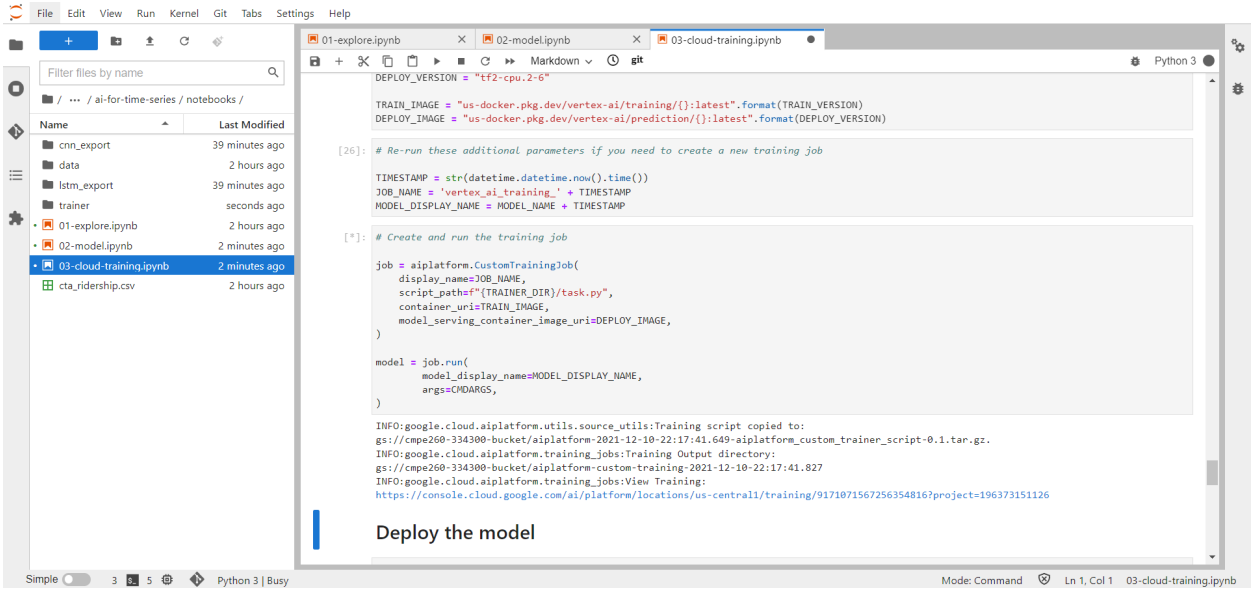
Step 2

Clear all the cells in the notebook (Edit > Clear All Outputs), change the region, project and bucket settings in one of the first few cells, and then Run the cells one by one.

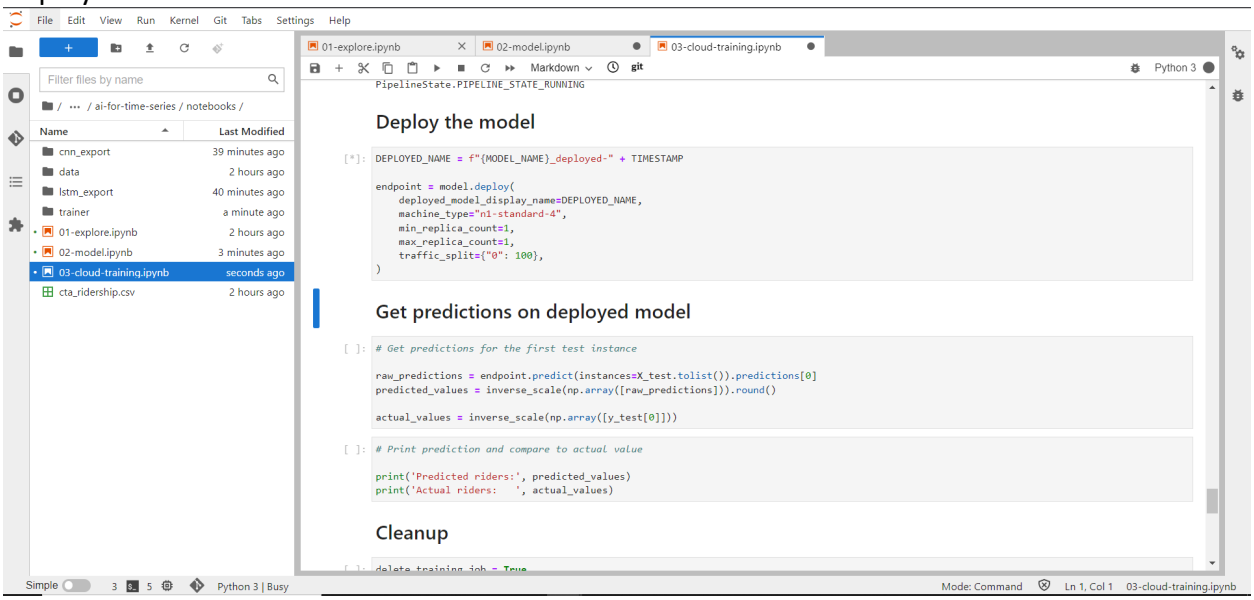
Step 3

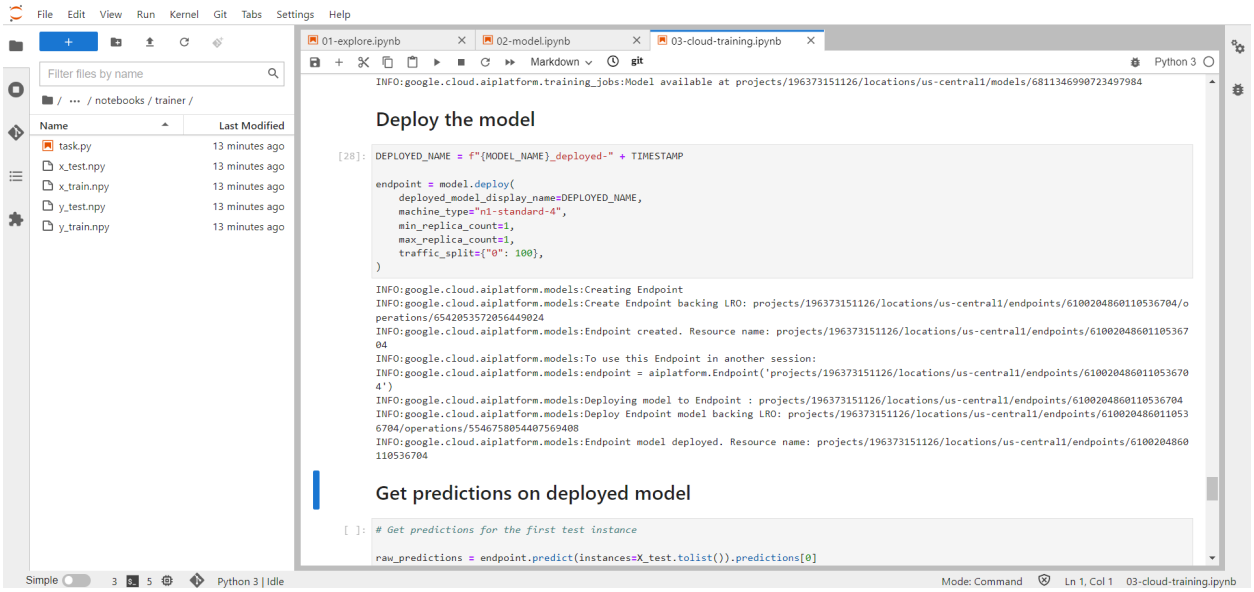


Create and run the training job

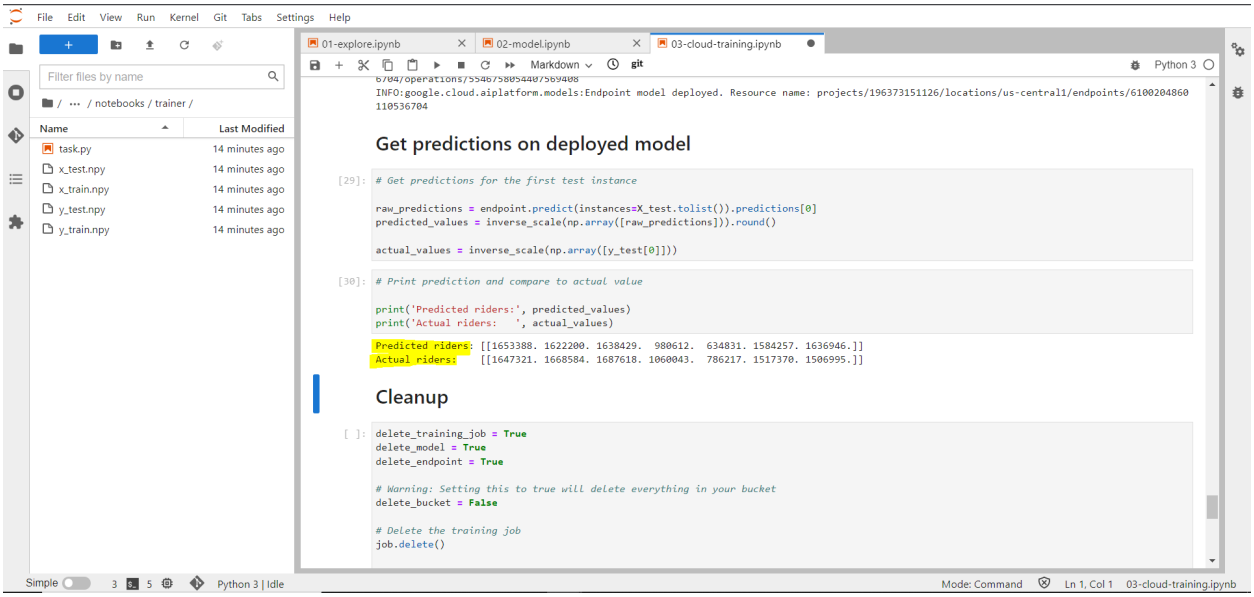


Deploy the Model

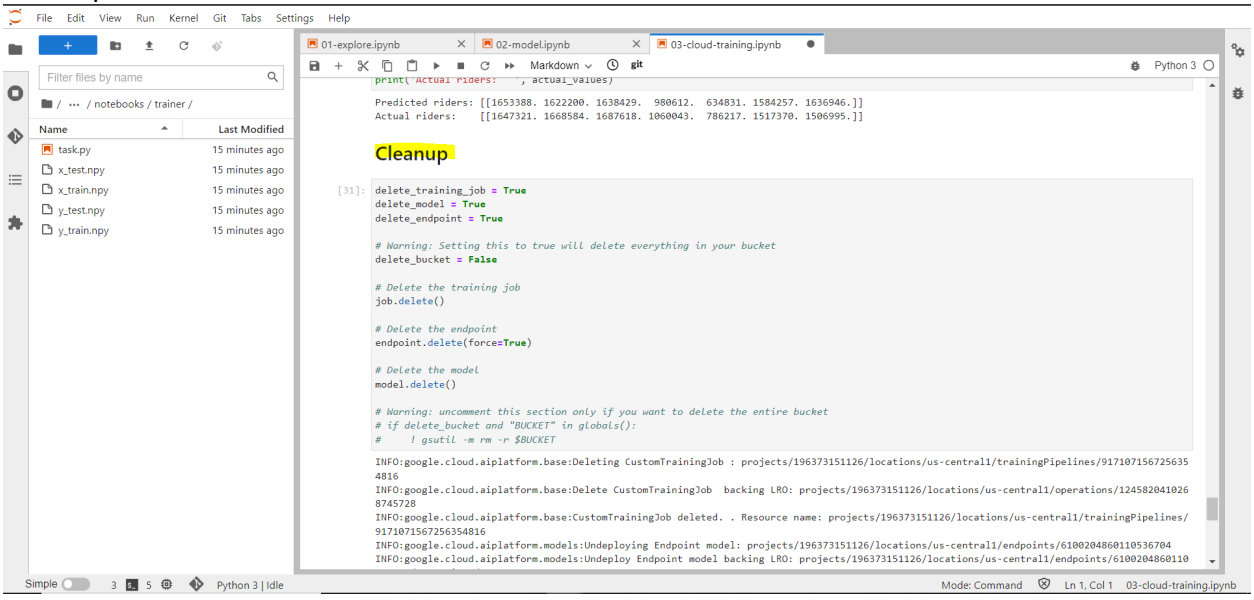




Prediction Results



Cleanup



Cleanup

From the Workbench UI in your Cloud Console, select the notebook and then select **Stop**: