

```
#import library
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score

# import data
data = pd.read_csv('house_prices.csv')

# describe data
print(data.head())
print(data.info())
print(data.describe())

# data visualisation
plt.hist(data['price'], bins=20)
plt.xlabel('House Price')
plt.ylabel('Frequency')
plt.title('Distribution of House Prices')
plt.show()

# data processing
data.dropna(inplace=True)
data = pd.get_dummies(data, columns=['neighborhood'], drop_first=True)

# define target(y) and features(x)
X = data.drop('price', axis=1)
y = data['price']

# train test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

model = LinearRegression()
model.fit(X_train, y_train)

y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
print(f'Mean Squared Error: {mse}')
print(f'R-squared: {r2}')
```

```
new_house_features = pd.DataFrame({
    'bedrooms': [3],
    'sqft': [1500],
    'neighborhood_Suburb': [1]
})
predicted_price = model.predict(new_house_features)
print(f'Predicted Price: ${predicted_price[0]:.2f}')
```


Please follow our [blog](#) to see more information about new features, tips and tricks, and featured notebooks such as [Analyzing a Bank Failure with Colab](#).

2024-01-29

- New [Kaggle Notebooks <> Colab updates](#)! Now you can:
 - Import directly from Colab without having to download/re-upload
 - Upload via link, by pasting Google Drive or Colab URLs
 - Export & run Kaggle Notebooks on Colab with 1 click
- Try these notebooks that talk to Gemini:
 - [Gemini and Stable Diffusion](#)
 - [Learning with Gemini and ChatGPT](#)
 - [Talk to Gemini with Google's Speech to Text API](#)
 - [Sell lemonade with Gemini and Sheets](#)
 - [Generate images with Gemini and Vertex](#)
- Python package upgrades
 - google-cloud-aiplatform 1.38.1 -> 1.39.0
 - bigframes 0.18.0 -> 0.19.2
 - polars 0.17.3 -> 0.20.2
 - gdown 4.6.6 -> 4.7.3 ([GitHub issue](#))
 - tensorflow-hub 0.15.0 -> 0.16.0
 - flax 0.7.5 -> 0.8.0
- Python package inclusions
 - sentencepiece 0.1.99

2024-01-08

- Avoid nested scrollbars for large outputs by using `google.colab.output.no_vertical_scroll()`
[Example notebook](#)
- Fix [bug](#) where downloading models from Hugging Face could freeze
- Python package upgrades
 - huggingface-hub 0.19.4 -> 0.20.2
 - bigframes 0.17.0 -> 0.18.0

2023-12-18

- Expanded access to AI coding has arrived in Colab across 175 locales for all tiers of Colab users
- Improvements to display of ML-based inline completions (for eligible Pro/Pro+ users)
- Started a series of [notebooks](#) highlighting Gemini API capabilities
- Enable ⌘/Ctrl+L to select the full line in an editor
- Fixed [bug](#) where we weren't correctly formatting output from multiple execution results
- Python package upgrades
 - CUDA 11.8 to CUDA 12.2
 - tensorflow 2.14.0 -> 2.15.0
 - tensorboard 2.14.0 -> 2.15.0
 - keras 2.14.0 -> 2.15.0
 - Nvidia drivers 525.105.17 -> 535.104.05
 - tensorflow-gcs-config 2.14.0 -> 2.15.0
 - bigframes 0.13.0 -> 0.17.0
 - geemap 0.28.2 -> 0.29.6
 - pyarrow 9.0.0 -> 10.0.1
 - google-generativeai 0.2.2 -> 0.3.1
 - jax 0.4.20 -> 0.4.23
 - jaxlib 0.4.20 -> 0.4.23
- Python package inclusions
 - kagglehub 0.1.4
 - google-cloud-aiplatform 1.38.1

2023-11-27

- Removed warning when calling `await` to make it render as code
- Added "Run selection" to the cell context menu
- Added highlighting for the `%%python` cell magic
- Launched AI coding features for Pro/Pro+ users in more locales
- Python package upgrades
 - bigframes 0.12.0 -> 0.13.0
- Python package inclusions
 - transformers 4.35.2
 - google-generativeai 0.2.2

2023-11-08

- Launched Secrets, for safe storage of private keys on Colab ([tweet](#))
- Fixed issue where TensorBoard would not load ([#3990](#))
- Python package upgrades
 - lightgbm 4.0.0 -> 4.1.0

- bigframes 0.10.0 -> 0.12.0
- bokeh 3.2.2 -> 3.3.0
- duckdb 0.8.1 -> 0.9.1
- numba 0.56.4 -> 0.58.1
- tweepy 4.13.0 -> 4.14.0
- jax 0.4.16 -> 0.4.20
- jaxlib 0.4.16 -> 0.4.20

2023-10-23

- Updated the **Open notebook** dialog for better usability and support for smaller screen sizes
- Added smart paste support for data from Google Sheets for R notebooks
- Enabled showing release notes in a tab
- Launched AI coding features for Pro/Pro+ users in Australia 🇦🇺 Canada 🇨🇦 India 🇮🇳 and Japan 🇯🇵 ([tweet](#))
- Python package upgrades
 - earthengine-api 0.1.357 -> 0.1.375
 - flax 0.7.2 -> 0.7.4
 - geemap 0.27.4 -> 0.28.2
 - jax 0.4.14 -> 0.4.16
 - jaxlib 0.4.14 -> 0.4.16
 - keras 2.13.1 -> 2.14.0
 - tensorboard 2.13.0 -> 2.14.1
 - tensorflow 2.13.0 -> 2.14.0
 - tensorflow-gcs-config 2.13.0 -> 2.14.0
 - tensorflow-hub 0.14.0 -> 0.15.0
 - tensorflow-probability 0.20.1 -> 0.22.0
 - torch 2.0.1 -> 2.1.0
 - torchaudio 2.0.2 -> 2.1.0
 - torchtext 0.15.2 -> 0.16.0
 - torchvision 0.15.2 -> 0.16.0
 - xgboost 1.7.6 -> 2.0.0
- Python package inclusions
 - bigframes 0.10.0
 - malloy 2023.1056

2023-09-22

- Added the ability to scope an AI generated suggestion to a specific Pandas dataframe ([tweet](#))
- Added Colab link previews to Docs ([tweet](#))
- Added smart paste support for data from Google Sheets
- Increased font size of dropdowns in interactive forms
- Improved rendering of the notebook when printing
- Python package upgrades
 - tensorflow 2.12.0 -> 2.13.0
 - tensorboard 2.12.3 -> 2.13.0
 - keras 2.12.0 -> 2.13.1
 - tensorflow-gcs-config 2.12.0 -> 2.13.
 - scipy 1.10.1-> 1.11.2
 - cython 0.29.6 -> 3.0.2
- Python package inclusions
 - geemap 0.26.0

2023-08-18

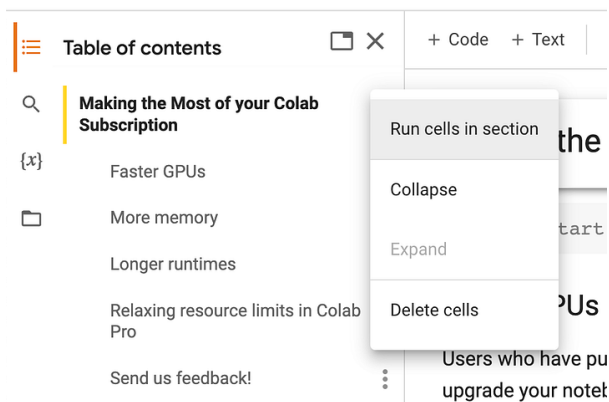
- Added "Change runtime type" to the menu in the connection button
- Improved auto-reconnection to an already running notebook ([#3764](#))
- Increased the specs of our highmem machines for Pro users
- Fixed add-apt-repository command on Ubuntu 22.04 runtime ([#3867](#))
- Python package upgrades
 - bokeh 2.4.3 -> 3.2.2
 - cmake 3.25.2 -> 3.27.2
 - cryptography 3.4.8 -> 41.0.3
 - dask 2022.12.1 -> 2023.8.0
 - distributed 2022.12.1 -> 2023.8.0
 - earthengine-api 0.1.358 -> 0.1.364
 - flax 0.7.0 -> 0.7.2
 - ipython-sql 0.4.0 -> 0.5.0
 - jax 0.4.13 -> 0.4.14
 - jaxlib 0.4.13 -> 0.4.14
 - lightgbm 3.3.5 -> 4.0.0
 - mkl 2019.0 -> 2023.2.0
 - notebook 6.4.8 -> 6.5.5
 - numpy 1.22.4 -> 1.23.5
 - opencv-python 4.7.0.72 -> 4.8.0.76
 - pillow 8.4.0 -> 9.4.0
 - plotly 5.13.1 -> 5.15.0
 - prettytable 0.7.2 -> 3.8.0
 - pytensor 2.10.1 -> 2.14.2
 - spacy 3.5.4 -> 3.6.1
 - statsmodels 0.13.5 -> 0.14.0
 - xarray 2022.12.0 -> 2023.7.0
- Python package inclusions
 - PyDrive2 1.6.3

2023-07-21

- Launched auto-plotting for dataframes, available using the chart button that shows up alongside datatables ([post](#))



- Added a menu to the table of contents to support running a section or collapsing/expanding sections ([post](#))



- Added an option to automatically run the first cell or section, available under Edit -> Notebook settings ([post](#))

Notebook settings

Runtime type

Python 3 ▾

Hardware accelerator

None ▾ ⓘ

☒ Automatically run the first cell or section

☐ Omit code cell output when saving this notebook

Cancel Save

- Launched Pro/Pro+ to Algeria, Argentina, Chile, Ecuador, Egypt, Ghana, Kenya, Malaysia, Nepal, Nigeria, Peru, Rwanda, Saudi Arabia, South Africa, Sri Lanka, Tunisia, and Ukraine ([tweet](#))
- Added a command, "Toggle tab moves focus" for toggling tab trapping in the editor (Tools -> Command palette, "Toggle tab moves focus")
- Fixed issue where `files.upload()` was sometimes returning an incorrect filename ([#1550](#))
- Fixed f-string syntax highlighting bug ([#3802](#))
- Disabled ambiguous characters highlighting for commonly used LaTeX characters ([#3648](#))
- Upgraded Ubuntu from 20.04 LTS to [22.04 LTS](#)
- Updated the Colab Marketplace VM image
- Python package upgrades:
 - autograd 1.6.1 -> 1.6.2
 - drivefs 76.0 -> 77.0
 - flax 0.6.11 -> 0.7.0
 - earthengine-api 0.1.357 -> 0.1.358
 - GDAL 3.3.2->3.4.3
 - google-cloud-bigquery-storage 2.20.0 -> 2.22.2
 - gspread-dataframe 3.0.8 -> 3.3.1
 - holidays 0.27.1 -> 0.29
 - jax 0.4.10 -> jax 0.4.13
 - jaxlib 0.4.10 -> jax 0.4.13
 - jupyterlab-widgets 3.0.7 -> 3.0.8
 - nbformat 5.9.0 -> 5.9.1
 - opencv-python-headless 4.7.0.72 -> 4.8.0.74
 - pygame 2.4.0 -> 2.5.0
 - spacy 3.5.3 -> 3.5.4
 - SQLAlchemy 2.0.16 -> 2.0.19
 - tabulate 0.8.10 -> 0.9.0
 - tensorflow-hub 0.13.0 -> 0.14.0

2023-06-23

- Launched AI coding features to subscribed users starting with Pro+ users in the US ([tweet](#), [post](#))
- Added the Kernel Selector in the Notebook Settings ([tweet](#))
- Fixed double space trimming issue in markdown [#3766](#)
- Fixed run button indicator not always centered [#3609](#)
- Fixed inconsistencies for automatic indentation on multi-line [#3697](#)
- Upgraded Python from 3.10.11 to 3.10.12
- Python package updates:
 - duckdb 0.7.1 -> 0.8.1
 - earthengine-api 0.1.350 -> 0.1.357
 - flax 0.6.9 -> 0.6.11
 - google-cloud-bigquery 3.9.0 -> 3.10.0
 - google-cloud-bigquery-storage 2.19.1 -> 2.20.0
 - grpcio 1.54.0 -> 1.56.0
 - holidays 0.25 -> 0.27.1

- nbformat 5.8.0 -> 5.9.0
- prophet 1.1.3 -> 1.1.4
- pydata-google-auth 1.7.0 -> 1.8.0
- spacy 3.5.2 -> 3.5.3
- tensorboard 2.12.2 -> 2.12.3
- xgboost 1.7.5 -> 1.7.6
- Python package inclusions:
 - gcsfs 2023.6.0
 - geopandas 0.13.2