

## ✓ Project : EDA on Superstore Dataset

```
#Python basics
```

```
#Libraries: numpy[numeric calcuations], pandas [data manipulation], matplotlib, seaborn
```

```
#Objective: Exploratory Data Analysis [EDA]
```

```
# I performed Exploratory data analysis [EDA] on superstore dataset using Python.  
# In Data Analysis advanced libraries of python Numpy, Pandas, Matplotlib, Seaborn  
# In EDA Univariate Analysis and Bivariate Analysis is performed and results are visualized
```

```
#Import libraries
```

```
import numpy as np  
import pandas as pd  
import matplotlib.pyplot as plt  
import seaborn as sns
```

```
import warnings  
warnings.filterwarnings('ignore')
```

```
pip install xlrd
```

```
➡ Collecting xlrd  
  Downloading xlrd-2.0.1-py2.py3-none-any.whl.metadata (3.4 kB)  
  Downloading xlrd-2.0.1-py2.py3-none-any.whl (96 kB)  
----- 0.0/96.5 kB ? eta -:--:--  
----- 10.2/96.5 kB ? eta -:--:--
```

```

----- 30.7/96.5 kB 435.7 kB/s eta 0:00:
----- 41.0/96.5 kB 393.8 kB/s eta 0:00:
----- 71.7/96.5 kB 491.5 kB/s eta 0:00:
----- 92.2/96.5 kB 476.3 kB/s eta 0:00:
----- 96.5/96.5 kB 394.0 kB/s eta 0:00:

```

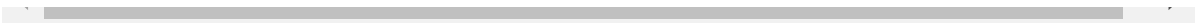
Installing collected packages: xlrd

Successfully installed xlrd-2.0.1

Note: you may need to restart the kernel to use updated packages.

```
#read dataset
```

```
df=pd.read_excel('superstore.xls')
```



```
#display data
```

```
df
```

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](#).

## 2024-08-20

- TPU memory usage and utilization can now be checked with `!tpu-info`
- Gemini Chat responses are now grounded in relevant sources
- Added a new "Create Gemini API key" link in the user secrets panel
- Added a new "Gemini: Creating a prompt" snippet and touched up the existing "Gemini: Connecting to Gemini" snippet
- Added the ability to specify custom placeholder text for various interactive form params (see [examples](#))
- Keyboard navigation a11y improvements to comments UI
- Various minor rendering improvements to interactive forms UI
- A11y improvements for the run button and header
- Updated tooltip styling
- A11y improvements for the file browser's disk usage bar
- On mobile, tooltips now trigger on long press
- On mobile, release notes updates will no longer display automatically
- Python package upgrades
  - astropy 5.3.4 -> 6.1.2
  - bigframes 1.11.1 -> 1.14.0
  - bokeh 3.3.4 -> 3.4.3
  - dask 2023.8.1 -> 2024.7.1
  - earthengine-api 0.1.412 -> 0.1.416
  - geopandas 0.13.2 -> 0.14.4
  - kagglehub 0.2.8 -> 0.2.9



- keras 2.15.0 -> 3.4.1
- lightgbm 4.1.0 -> 4.4.0
- malloy 2023.1067 -> 2024.1067
- numba 0.58.1 -> 0.60.0
- numpy 1.25.2 -> 1.26.4
- opencv-python 4.8.0.76 -> 4.10.0.84
- pandas 2.0.3 -> 2.1.4
- pandas-gbq 0.19.2 -> 0.23.1
- panel 1.3.8 -> 1.4.5
- requests 2.31.0 -> 2.32.3
- scikit-learn 1.2.2 -> 1.3.2
- scipy 1.11.4 -> 1.13.1
- tensorboard 2.15.2 -> 2.17.0
- tensorflow 2.15.0 -> 2.17.0
- tf-keras 2.15.1 -> 2.17.0
- xarray 2023.7.0 -> 2024.6.0
- xgboost 2.0.3 -> 2.1.1
- Python package inclusions
  - einops 0.8.0

## 2024-07-22

- You can now embed Google sheets directly into Colab to streamline interactions with data with InteractiveSheet.

Example:

```
from google.colab import sheets
sh = sheets.InteractiveSheet()
df = sh.as_df()
```

- Fixed multiple rendering bugs in cell editors with wide text content (i.e. text is no longer hidden or clipped)
- Fixed multiple accessibility issues in Colab's comments feature (e.g. proper keyboard focus management, added accessibility landmarks, etc)
- Fixed bug where AI code generation would fail for extremely long broken code snippets

#2. Display total number of rows and columns

```
df.shape
print("Total Number of Rows: ",df.shape[0])
print("Total Number of Coulmns: ",df.shape[1])
```

```
↔ Total Number of Rows: 9994
   Total Number of Coulmns: 21
```

```
#3. Display first rows. ---- head()
#4. Display last 10 rows ---- tail()
#5. Display random 10 rows. ---- sample()
```

```
#3. Display first rows. ---- head()
```

```
df.head(10)
```

- Fixed multiple scrollbar bugs in the user secrets panel
- Added the ability for workspace admin to purchase Colab Pro and Pro+ Subscriptions for users
- Fixed bug where user secrets couldn't be moved to a tab
- Fixed several focus management accessibility issues in tabs, the table of contents, the left toolbar, and the run button
- Fixed bug where overflowing cells may be omitted when pasting from Google Sheets
- Fixed bug where the generate code button did not activate on touch
- Python package upgrades
  - bigframes 1.9.0 -> 1.11.1
  - cvxpy 1.3.4 -> 1.5.2
  - earthengine-api 0.1.408 -> 0.1.412
  - google-api-core 2.11.1 -> 2.19.1
  - google-api-python-client 2.84.0 -> 2.137.0
  - google-cloud-aiplatform 1.56.0 -> 1.59.0
  - google-cloud-bigquery 3.21.0 -> 3.25.0
  - google-cloud-core 2.3.3 -> 2.4.1
  - google-cloud-datastore 2.15.2 -> 2.19.0
  - google-cloud-firestore 2.11.1 -> 2.16.1
  - google-cloud-functions 1.13.3 -> 1.16.4
  - google-generativeai 0.5.4 -> 0.7.2
  - kagglehub 0.2.5 -> 0.2.8
  - pip 23.1.2 -> 24.1.2
  - setuptools 67.7.2 -> 71.0.4
  - sympy 1.12.1 -> 1.13.1
  - torch 2.3.0 -> 2.3.1
  - transformers 4.41.2 -> 4.42.4
- Python package inclusions
  - accelerate 0.32.1

**2024-06-18**

- Inline AI completions are now available to users on the free-of-charge tier
- Reduced latency for LSP and terminal connections
- Improved quality of inline completions
- Visual improvements to switch controls across Colab
- Various bug fixes, performance and a11y improvements to the user secrets panel
- Improved tooltip UX behavior
- Improved behavior when copying data from Google Sheets and pasting in Colab
- Scroll to cell fixes for single tabbed view and jump to cell command
- Improved tab header behavior
- A11y improvements for notebook-focused cells
- Python package upgrades
  - torch 2.2.1 -> 2.3.0
  - torchaudio 2.2.1 -> 2.3.0
  - torchvision 0.17.1 -> 0.18.0
  - torchtext 0.17.1 -> 0.18.0
  - google-cloud-aiplatform 1.51.0 -> 1.56.0
  - bigframes 1.5.0 -> 1.8.0
  - regex 2023.12.25 -> 2024.5.15

## 2024-05-13

- Code actions are now supported to automatically improve and refactor code. Code actions can be triggered by the keyboard shortcut "Ctrl/⌘ + ."
- Python package upgrades
  - bigframes 1.0.0 -> 1.5.0
  - google-cloud-aiplatform 1.47.0 -> 1.51.0
  - jax[tpu] 0.4.23 -> 0.4.26
- Python package inclusions
  - cudf 24.4.1



## 2024-04-15

- TPU v2 runtime is now available
- L4 runtime is now available for paid users
- New distributed fine-tuning Gemma tutorial on TPUs ([GitHub](#))
- Symbol rename is now supported with keyboard shortcut F2
- Fixed bug causing inability to re-upload deleted files
- Fixed breaking bug in colabtools %upload\_files\_async
- Added syntax highlighting to %%writefile cells
- Cuda dependencies that come with Torch are cached for faster downloads for packages that require Torch and its dependencies ([GitHub issue](#))
- Python package upgrades
  - bigframes 0.24.0 -> 1.0.0
  - duckdb 0.9.2 -> 0.10.1
  - google-cloud-aiplatform 1.43.0 -> 1.47.0
  - jax 0.4.23 -> 0.4.26

## 2024-03-13

- Fixed bug that sometimes caused UserSecrets to move / disappear
- Improved messaging for mounting drive in an unsupported environment ([GitHub issue](#))
- Python package upgrades
  - torch 2.1.0 -> 2.2.1
  - torchaudio 2.1.0 -> 2.2.1
  - torchvision 0.16.0 -> 0.17.1
  - torchtext 0.16.0 -> 0.17.1
  - PyMC 5.7.2 -> 5.10.4
  - BigFrames 0.21.0 -> 0.24.0
  - google-cloud-aiplatform 1.42.1 -> 1.43.0
  - tornado 6.3.2 -> 6.3.3

```
#4. Display last 10 rows ---- tail()
```

```
df.tail(10)
```

## 2024-02-21

- Try out Gemma on [Colab](#)!
- Allow unicode in form text inputs
- Display documentation and link to source when displaying functions
- Display image-like ndarrays as images
- Improved UX around quick charts and execution error suggestions
- Released Marketplace image for the month of February ([GitHub issue](#))
- Python package upgrades
  - bigframes 0.19.2 -> 0.21.0
  - regex 2023.6.3 -> 2023.12.25
  - spacy 3.6.1 -> 3.7.4
  - beautifulsoup4 4.11.2 -> 4.12.3
  - tensorflow-probability 0.22.0 -> 0.23.0
  - google-cloud-language 2.9.1 -> 2.13.1
  - google-cloud-aiplatform 1.39.0 -> 1.42.1
  - transformers 4.35.2 -> 4.37.2
  - pyarrow 10.0.1 -> 14.0.2

## 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



```
#5. Display random 10 rows. ---- sample()

df.sample(10)
```

- 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 AU Canada CA India IN and Japan JP ([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

# 6. Display all column names

df.columns

```
Index(['Row ID', 'Order ID', 'Order Date', 'Ship Date', 'Ship Mode',
      'Customer ID', 'Customer Name', 'Segment', 'Country', 'City',
      'State',
      'Postal Code', 'Region', 'Product ID', 'Category', 'Sub-Category',
      'Product Name', 'Sales', 'Quantity', 'Discount', 'Profit'],
      dtype='object')
```

#7. Display datatypes of each column

df.dtypes

```
Row ID          int64
Order ID        object
Order Date      datetime64[ns]
Ship Date       datetime64[ns]
Ship Mode       object
Customer ID     object
Customer Name   object
Segment         object
Country         object
City           object
```

- 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

```

State                object
Postal Code          int64
Region               object
Product ID           object
Category             object
Sub-Category         object
Product Name         object
Sales                float64
Quantity             int64
Discount             float64
Profit               float64
dtype: object

```

#8. Identify whether any columns has null values

```
df.isnull().sum()
```

```

Row ID      0
Order ID    0
Order Date  0
Ship Date   0
Ship Mode   0
Customer ID  0
Customer Name 0
Segment     0
Country     0
City        0
State       0
Postal Code  0
Region      0
Product ID  0
Category    0
Sub-Category 0
Product Name 0
Sales       0
Quantity    0
Discount    0
Profit      0
dtype: int64

```

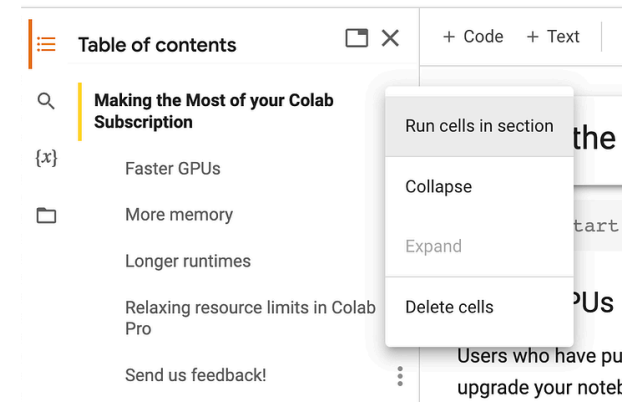
- lightgbm 3.3.5 -> 4.0.0
- mkl 2019.0 -> 2023.2.0
- notebook 6.4.8 -> 6.5.5
- numpy 1.22.4 -> 1.23.5
- opencv-python 4.7.0.72 -> 4.8.0.76
- pillow 8.4.0 -> 9.4.0
- plotly 5.13.1 -> 5.15.0
- prettytable 0.7.2 -> 3.8.0
- pytensor 2.10.1 -> 2.14.2
- spacy 3.5.4 -> 3.6.1
- statsmodels 0.13.5 -> 0.14.0
- 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

#9.

#Categorical Features


#Continous Features

#1. Order Id


#unique

#nunique--count of unique

df['Order ID'].nunique()

 5009

df.nunique()



Row ID	9994
Order ID	5009
Order Date	1237
Ship Date	1334
Ship Mode	4
Customer ID	793
Customer Name	793
Segment	3
Country	1
City	531
State	49
Postal Code	631
Region	4
Product ID	1862
Category	3
Sub-Category	17
Product Name	1850
Sales	6144
Quantity	14
Discount	12

(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

```
Profit          7545
dtype: int64
```

```
#Identify categorical and continous features.
```

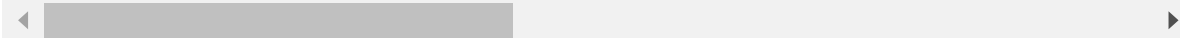
```
cat=[]
cont=[]
```

```
threshold=20
```

```
for i in df.columns:
    if df[i].nunique() >= 20:
        cont.append(i)
    else :
        cat.append(i)
```

```
print("Categorical Features: ",cat)
print()
print("Continous Features: ",cont)
```

```
➦ Categorical Features: ['Ship Mode', 'Segment', 'Country', 'Region', 'Categor
Continous Features: ['Row ID', 'Order ID', 'Order Date', 'Ship Date', 'Custo
```



```
#10. Display Unique values for each categorical feature.
```

```
for i in cat:
    print(i,"\t",df[i].unique())
    print()
```

```
➦ Ship Mode          ['Second Class' 'Standard Class' 'First Class' 'Same Day']

Segment              ['Consumer' 'Corporate' 'Home Office']

Country              ['United States']
```

- 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
  - google-cloud-bigquery-connection 1.12.0
  - google-cloud-functions 1.13.0

```

Region      ['South' 'West' 'Central' 'East']

Category     ['Furniture' 'Office Supplies' 'Technology']

Sub-Category ['Bookcases' 'Chairs' 'Labels' 'Tables' 'Storage' 'Furnishin
'Phones' 'Binders' 'Appliances' 'Paper' 'Accessories' 'Envelopes'
'Fasteners' 'Supplies' 'Machines' 'Copiers']

Quantity     [ 2  3  5  7  4  6  9  1  8 14 11 13 10 12]

Discount     [0.   0.45 0.2  0.8  0.3  0.5  0.7  0.6  0.32 0.1  0.4  0.15

```

# 11. Display unique values in city and state

```

print("city: ")
print(df['City'].unique())

```

```

city:
['Henderson' 'Los Angeles' 'Fort Lauderdale' 'Concord' 'Seattle'
'Fort Worth' 'Madison' 'West Jordan' 'San Francisco' 'Fremont'
'Philadelphia' 'Orem' 'Houston' 'Richardson' 'Naperville' 'Melbourne'
'Eagan' 'Westland' 'Dover' 'New Albany' 'New York City' 'Troy' 'Chicago'
'Gilbert' 'Springfield' 'Jackson' 'Memphis' 'Decatur' 'Durham' 'Columbia'
'Rochester' 'Minneapolis' 'Portland' 'Saint Paul' 'Aurora' 'Charlotte'
'Orland Park' 'Urbandale' 'Columbus' 'Bristol' 'Wilmington' 'Bloomington'
'Phoenix' 'Roseville' 'Independence' 'Pasadena' 'Newark' 'Franklin'
'Scottsdale' 'San Jose' 'Edmond' 'Carlsbad' 'San Antonio' 'Monroe'
'Fairfield' 'Grand Prairie' 'Redlands' 'Hamilton' 'Westfield' 'Akron'
'Denver' 'Dallas' 'Whittier' 'Saginaw' 'Medina' 'Dublin' 'Detroit'
'Tampa' 'Santa Clara' 'Lakeville' 'San Diego' 'Brentwood' 'Chapel Hill'
'Morristown' 'Cincinnati' 'Inglewood' 'Tamarac' 'Colorado Springs'
'Belleville' 'Taylor' 'Lakewood' 'Arlington' 'Arvada' 'Hackensack'
'Saint Petersburg' 'Long Beach' 'Hesperia' 'Murfreesboro' 'Layton'
'Austin' 'Lowell' 'Manchester' 'Harlingen' 'Tucson' 'Quincy'
'Pembroke Pines' 'Des Moines' 'Peoria' 'Las Vegas' 'Warwick' 'Miami'
'Huntington Beach' 'Richmond' 'Louisville' 'Lawrence' 'Canton'
'New Rochelle' 'Gastonia' 'Jacksonville' 'Auburn' 'Norman' 'Park Ridge'

```

- grpc-google-iam-v1 0.12.6
- multidict 6.0.4
- tensorboard-data-server 0.7.1

## 2023-06-02

- Released the new site [colab.google](https://colab.google)
- Published Colab's Docker runtime image to us-docker.pkg.dev/colab-images/public/runtime ([tweet](#) [instructions](#))
- Launched support for Google children accounts ([tweet](#))
- Launched DagsHub integration ([tweet](#), [post](#))
- Upgraded to Monaco Editor Version 0.37.1
- Fixed various Vim keybinding bugs
- Fixed issue where the N and P letters sometimes couldn't be typed ([#3664](#))
- Fixed rendering support for compositional inputs ([#3660](#), [#3679](#))
- Fixed lag in notebooks with lots of cells ([#3676](#))
- Improved support for R by adding a Runtime type notebook setting (Edit -> Notebook settings)
- Improved documentation for connecting to a local runtime (Connect -> Connect to a local runtime)
- Python package updates:
  - holidays 0.23 -> 0.25
  - jax 0.4.8 -> 0.4.10
  - jaxlib 0.4.8 -> 0.4.10
  - pip 23.0.1 -> 23.1.2
  - tensorflow-probability 0.19.0 -> 0.20.1
  - torch 2.0.0 -> 2.0.1
  - torchaudio 2.0.1 -> 2.0.2
  - torchdata 0.6.0 -> 0.6.1
  - torchtext 0.15.1 -> 0.15.2
  - torchvision 0.15.1 -> 0.15.2
  - tornado 6.2 -> 6.3.1

## 2023-05-05

- Released GPU type selection for paid users, allowing them to choose a preferred NVIDIA GPU

```
'Amarillo' 'Lindenhurst' 'Huntsville' 'Fayetteville' 'Costa Mesa'
'Parker' 'Atlanta' 'Gladstone' 'Great Falls' 'Lakeland' 'Montgomery'
'Mesa' 'Green Bay' 'Anaheim' 'Marysville' 'Salem' 'Laredo' 'Grove City'
'Dearborn' 'Warner Robins' 'Vallejo' 'Mission Viejo' 'Rochester Hills'
'Plainfield' 'Sierra Vista' 'Vancouver' 'Cleveland' 'Tyler' 'Burlington'
'Waynesboro' 'Chester' 'Cary' 'Palm Coast' 'Mount Vernon' 'Hialeah'
'Oceanside' 'Evanston' 'Trenton' 'Cottage Grove' 'Bossier City'
'Lancaster' 'Asheville' 'Lake Elsinore' 'Omaha' 'Edmonds' 'Santa Ana'
'Milwaukee' 'Florence' 'Lorain' 'Linden' 'Salinas' 'New Brunswick'
'Garland' 'Norwich' 'Alexandria' 'Toledo' 'Farmington' 'Riverside'
'Torrance' 'Round Rock' 'Boca Raton' 'Virginia Beach' 'Murrieta'
'Olympia' 'Washington' 'Jefferson City' 'Saint Peters' 'Rockford'
'Brownsville' 'Yonkers' 'Oakland' 'Clinton' 'Encinitas' 'Roswell'
'Jonesboro' 'Antioch' 'Homestead' 'La Porte' 'Lansing' 'Cuyahoga Falls'
'Reno' 'Harrisonburg' 'Escondido' 'Royal Oak' 'Rockville' 'Coral Springs'
'Buffalo' 'Boynton Beach' 'Gulfport' 'Fresno' 'Greenville' 'Macon'
'Cedar Rapids' 'Providence' 'Pueblo' 'Deltona' 'Murray' 'Middletown'
'Freeport' 'Pico Rivera' 'Provo' 'Pleasant Grove' 'Smyrna' 'Parma'
'Mobile' 'New Bedford' 'Irving' 'Vineland' 'Glendale' 'Niagara Falls'
'Thomasville' 'Westminster' 'Coppell' 'Pomona' 'North Las Vegas'
'Allentown' 'Tempe' 'Laguna Niguel' 'Bridgeton' 'Everett' 'Watertown'
'Appleton' 'Bellevue' 'Allen' 'El Paso' 'Grapevine' 'Carrollton' 'Kent'
'Lafayette' 'Tigard' 'Skokie' 'Plano' 'Suffolk' 'Indianapolis' 'Bayonne'
'Greensboro' 'Baltimore' 'Kenosha' 'Olathe' 'Tulsa' 'Redmond' 'Raleigh'
'Muskogee' 'Meriden' 'Bowling Green' 'South Bend' 'Spokane' 'Keller'
'Port Orange' 'Medford' 'Charlottesville' 'Missoula' 'Apopka' 'Reading'
'Broomfield' 'Paterson' 'Oklahoma City' 'Chesapeake' 'Lubbock'
'Johnson City' 'San Bernardino' 'Leominster' 'Bozeman' 'Perth Amboy'
'Ontario' 'Rancho Cucamonga' 'Moorhead' 'Mesquite' 'Stockton'
'Ormond Beach' 'Sunnyvale' 'York' 'College Station' 'Saint Louis'
'Manteca' 'San Angelo' 'Salt Lake City' 'Knoxville' 'Little Rock'
'Lincoln Park' 'Marion' 'Littleton' 'Bangor' 'Southaven' 'New Castle'
'Midland' 'Sioux Falls' 'Fort Collins' 'Clarksville' 'Sacramento'
'Thousand Oaks' 'Malden' 'Holyoke' 'Albuquerque' 'Sparks' 'Coachella'
'Elmhurst' 'Passaic' 'North Charleston' 'Newport News' 'Jamestown'
'Mishawaka' 'La Quinta' 'Tallahassee' 'Nashville' 'Bellingham'
'Woodstock' 'Haltom City' 'Wheeling' 'Summerville' 'Hot Springs'
'Englewood' 'Las Cruces' 'Hoover' 'Frisco' 'Vacaville' 'Waukesha'
```

- Upgraded R from 4.2.3 to 4.3.0
- Upgraded Python from 3.9.16 to 3.10.11
- Python package updates:
  - attrs 22.2.0 -> attrs 23.1.0
  - earthengine-api 0.1.349 -> earthengine-api 0.1.350
  - flax 0.6.8 -> 0.6.9
  - grpcio 1.53.0 -> 1.54.0
  - nbclient 0.7.3 -> 0.7.4
  - tensorflow-datasets 4.8.3 -> 4.9.2
  - termcolor 2.2.0 -> 2.3.0
  - zict 2.2.0 -> 3.0.0

## 2023-04-14

- Python package updates:
  - google-api-python-client 2.70.0 -> 2.84.0
  - google-auth-oauthlib 0.4.6 -> 1.0.0
  - google-cloud-bigquery 3.4.2 -> 3.9.0
  - google-cloud-datastore 2.11.1 -> 2.15.1
  - google-cloud-firestore 2.7.3 -> 2.11.0
  - google-cloud-language 2.6.1 -> 2.9.1
  - google-cloud-storage 2.7.0 -> 2.8.0
  - google-cloud-translate 3.8.4 -> 3.11.1
  - networkx 3.0 -> 3.1
  - notebook 6.3.0 -> 6.4.8
  - jax 0.4.7 -> 0.4.8
  - pandas 1.4.4 -> 1.5.3
  - spacy 3.5.1 -> 3.5.2
  - SQLAlchemy 1.4.47 -> 2.0.9
  - xgboost 1.7.4 -> 1.7.5

## 2023-03-31

- Improve bash ! syntax highlighting ([GitHub issue](#))
- Fix bug where VIM keybindings weren't working in the file editor
- Upgraded R from 4.2.2 to 4.2.3
- Python package updates:
  - arviz 0.12.1 -> 0.15.1



```
print("State : ")
print(df['State'].unique())
```

```
⇒ State :
['Kentucky' 'California' 'Florida' 'North Carolina' 'Washington' 'Texas'
 'Wisconsin' 'Utah' 'Nebraska' 'Pennsylvania' 'Illinois' 'Minnesota'
 'Michigan' 'Delaware' 'Indiana' 'New York' 'Arizona' 'Virginia'
 'Tennessee' 'Alabama' 'South Carolina' 'Oregon' 'Colorado' 'Iowa' 'Ohio'
 'Missouri' 'Oklahoma' 'New Mexico' 'Louisiana' 'Connecticut' 'New Jersey'
 'Massachusetts' 'Georgia' 'Nevada' 'Rhode Island' 'Mississippi'
 'Arkansas' 'Montana' 'New Hampshire' 'Maryland' 'District of Columbia'
 'Kansas' 'Vermont' 'Maine' 'South Dakota' 'Idaho' 'North Dakota'
 'Wyoming' 'West Virginia']
```

## ✓ Uni Variate Analysis

```
cat
```

```
⇒ ['Ship Mode',
   'Segment',
   'Country',
   'Region',
   'Category',
   'Sub-Category',
   'Quantity',
   'Discount']
```

#1. Display count of orders through each shipment mode.

```
df['Ship Mode'].value_counts()
```

```
⇒ Ship Mode
Standard Class    5968
Second Class      1945
First Class       1538
```

- astropy 4.3.1 -> 5.2.2
- dopamine-rl 1.0.5 -> 4.0.6
- gensim 3.6.0 -> 4.3.1
- ipykernel 5.3.4 -> 5.5.6
- ipython 7.9.0 -> 7.34.0
- jax 0.4.4 -> 0.4.7
- jaxlib 0.4.4 -> 0.4.7
- jupyter\_core 5.2.0 -> 5.3.0
- keras 2.11.0 -> 2.12.0
- lightgbm 2.2.3 -> 3.3.5
- matplotlib 3.5.3 -> 3.7.1
- nltk 3.7 -> 3.8.1
- opencv-python 4.6.0.66 -> 4.7.0.72
- plotly 5.5.0 -> 5.13.1
- pymc 4.1.4 -> 5.1.2
- seaborn 0.11.2 -> 0.12.2
- spacy 3.4.4 -> 3.5.1
- sympy 1.7.1 -> 1.11.1
- tensorboard 2.11.2 -> 2.12.0
- tensorflow 2.11.0 -> 2.12.0
- tensorflow-estimator 2.11.0 -> 2.12.0
- tensorflow-hub 0.12.0 -> 0.13.0
- torch 1.13.1 -> 2.0.0
- torchaudio 0.13.1 -> 2.0.1
- torchtext 0.14.1 -> 0.15.1
- torchvision 0.14.1 -> 0.15.1

## 2023-03-10

- Added the [Colab editor shortcuts](#) example notebook
- Fixed triggering of @-mention and email autocomplete for large comments ([GitHub issue](#))
- Added View Resources to the Runtime menu
- Made file viewer images fit the view by default, resizing to original size on click
- When in VIM mode, enable copy as well as allowing propagation to monaco-vim to escape visual mode ([GitHub issue](#))
- Upgraded CUDA 11.6.2 -> 11.8.0 and cuDNN 8.4.0.27 -> 8.7.0.84
- Upgraded Nvidia drivers 525.78.01 -> 530.30.02

```
Same Day          543
Name: count, dtype: int64
```

```
#Count of shipment mode in percentages
```

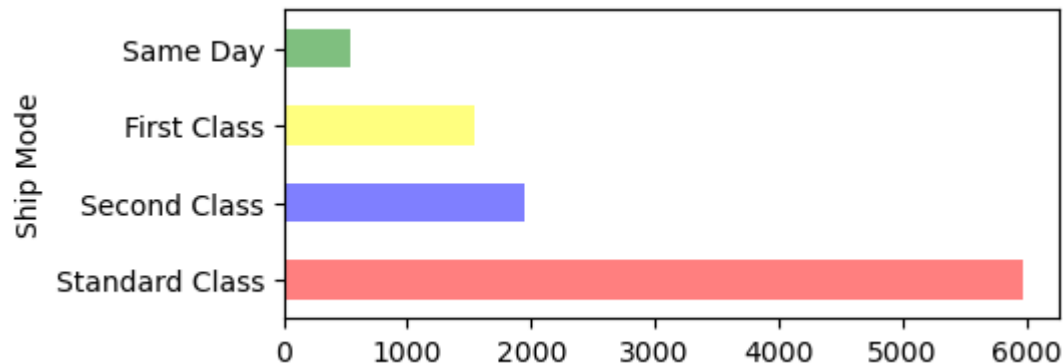
```
np.round(df['Ship Mode'].value_counts() / df.shape[0] * 100 , 2)
```

```
⇒ Ship Mode
Standard Class    59.72
Second Class      19.46
First Class       15.39
Same Day          5.43
Name: count, dtype: float64
```

```
# Visualization
```

```
plt.figure(figsize=(5,2))
df['Ship Mode'].value_counts().plot.barh(color=['red','blue','yellow','green'],al
```

```
⇒ <Axes: ylabel='Ship Mode'>
```



```
# 1.Display region wise order count
```

```
df['Region'].value_counts()
```

- Upgraded Python 3.8.10 -> 3.9.16
- Python package updates:
  - beautifulsoup4 4.6.3 -> 4.9.3
  - bokeh 2.3.3 -> 2.4.3
  - debugpy 1.0.0 -> 1.6.6
  - Flask 1.1.4 -> 2.2.3
  - jax 0.3.25 -> 0.4.4
  - jaxlib 0.3.25 -> 0.4.4
  - Jinja2 2.11.3 -> 3.1.2
  - matplotlib 3.2.2 -> 3.5.3
  - nbconvert 5.6.1 -> 6.5.4
  - pandas 1.3.5 -> 1.4.4
  - pandas-datareader 0.9.0 -> 0.10.0
  - pandas-profiling 1.4.1 -> 3.2.0
  - Pillow 7.1.2 -> 8.4.0
  - plotnine 0.8.0 -> 0.10.1
  - scikit-image 0.18.3 -> 0.19.3
  - scikit-learn 1.0.2 -> 1.2.2
  - scipy 1.7.3 -> 1.10.1
  - setuptools 57.4.0 -> 63.4.3
  - sklearn-pandas 1.8.0 -> 2.2.0
  - statsmodels 0.12.2 -> 0.13.5
  - urllib3 1.24.3 -> 1.26.14
  - Werkzeug 1.0.1 -> 2.2.3
  - wrapt 1.14.1 -> 1.15.0
  - xgboost 0.90 -> 1.7.4
  - xlrd 1.2.0 -> 2.0.1

## 2023-02-17

- Show graphs of RAM and disk usage in notebook toolbar
- Copy cell links directly to the clipboard instead of showing a dialog when clicking on the link icon in the cell toolbar
- Updated the [Colab Marketplace VM image](#)
- Upgraded CUDA to 11.6.2 and cuDNN to 8.4.0.27
- Python package updates:
  - tensorflow 2.9.2 -> 2.11.0
  - tensorboard 2.9.1 -> 2.11.2

```

Region
West      3203
East      2848
Central    2323
South     1620
Name: count, dtype: int64

```

# 2. % Terms

```
np.round(df['Region'].value_counts() / df.shape[0] * 100 , 2)
```

```

Region
West      32.05
East      28.50
Central    23.24
South     16.21
Name: count, dtype: float64

```

# 3.Display ordre count by each segment

```
df['Segment'].value_counts()
```

```

Segment
Consumer      5191
Corporate     3020
Home Office   1783
Name: count, dtype: int64

```

```
df['Segment'].value_counts() / df.shape[0] * 100
```

```

Segment
Consumer      51.941165
Corporate     30.218131
Home Office   17.840704
Name: count, dtype: float64

```

- keras 2.9.0 -> 2.11.0
- tensorflow-estimator 2.9.0 -> 2.11.0
- tensorflow-probability 0.17.0 -> 0.19.0
- tensorflow-gcs-config 2.9.0 -> 2.11.0
- earthengine-api 0.1.339 -> 0.1.341
- flatbuffers 1.12 -> 23.1.21
- platformdirs 2.6.2 -> 3.0.0
- pydata-google-auth 1.6.0 -> 1.7.0
- python-utils 3.4.5 -> 3.5.2
- tenacity 8.1.0 -> 8.2.1
- tifffile 2023.1.23.1 -> 2023.2.3
- notebook 5.7.16 -> 6.3.0
- tornado 6.0.4 -> 6.2
- aiohttp 3.8.3 -> 3.8.4
- charset-normalizer 2.1.1 -> 3.0.1
- fastai 2.7.0 -> 2.7.1
- soundfile 0.11.0 -> 0.12.1
- typing-extensions 4.4.0 -> 4.5.0
- widgetsnbextension 3.6.1 -> 3.6.2
- pydantic 1.10.4 -> 1.10.5
- zipp 3.12.0 -> 3.13.0
- numpy 1.21.6 -> 1.22.4
- drivefs 66.0 -> 69.0
- gdal 3.0.4 -> 3.3.2 [GitHub issue](#)
- Added libudunits2-dev for smoother R package installs [GitHub issue](#)

## 2023-02-03

- Improved tooltips for pandas series to show common statistics about the series object
- Made the forms dropdown behave like an autocomplete box when it allows input
- Updated the nvidia driver from 460.32.03 to 510.47.03
- Python package updates:
  - absl-py 1.3.0 -> 1.4.0
  - bleach 5.0.1 -> 6.0.0
  - cachetools 5.2.1 -> 5.3.0
  - cmdstanpy 1.0.8 -> 1.1.0

#4. Display top 5 states by order count

```
df['City'].value_counts()[:10]
```

```

City
New York City    915
Los Angeles      747
Philadelphia     537
San Francisco    510
Seattle          428
Houston          377
Chicago          314
Columbus         222
San Diego        170
Springfield      163
Name: count, dtype: int64

```

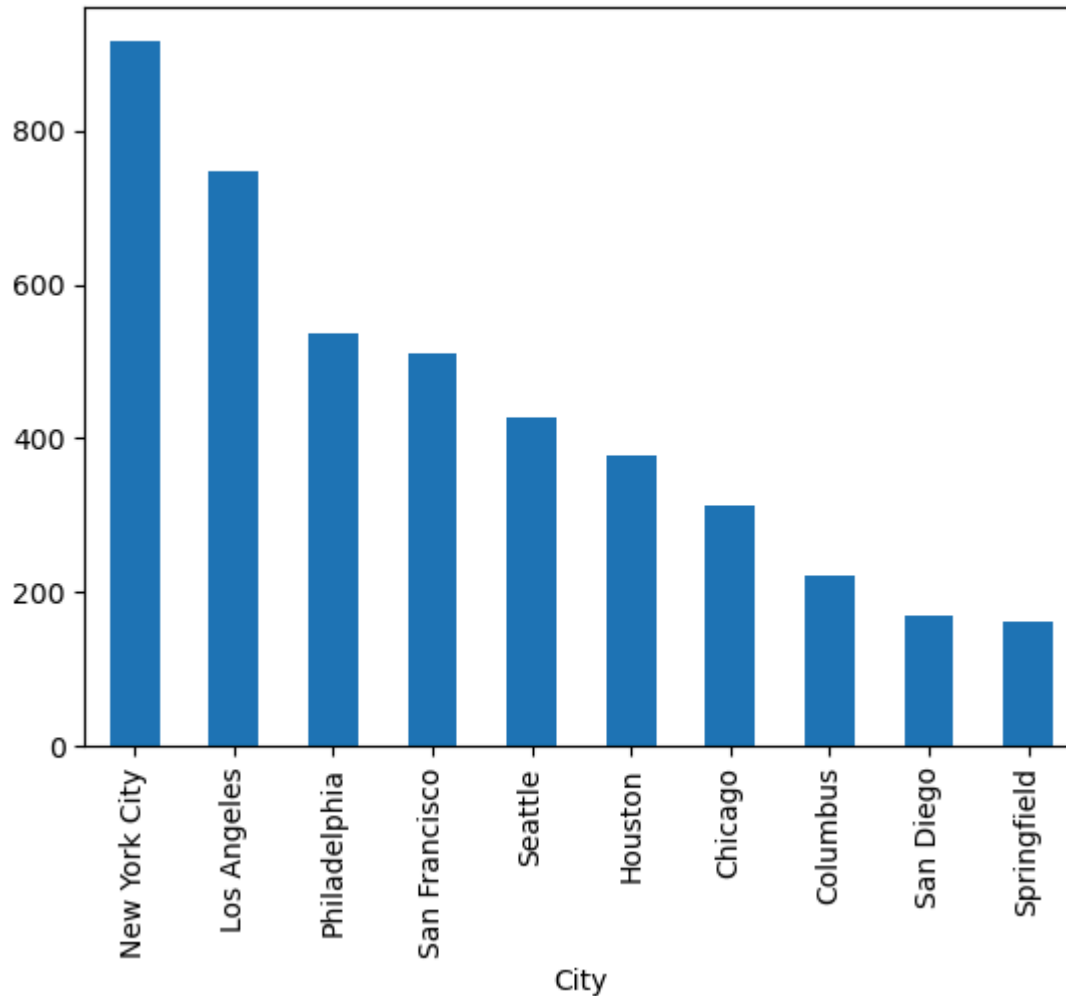
```
df['City'].value_counts()[:10].plot.bar()
```

- dnspython 2.2.1 -> 2.3.0
- fsspec 2022.11.0 -> 2023.1.0
- google-cloud-bigquery-storage 2.17.0 -> 2.18.1
- holidays 0.18 -> 0.19
- jupyter-core 5.1.3 -> 5.2.0
- packaging 21.3 -> 23.0
- prometheus-client 0.15.0 -> 0.16.0
- pyct 0.4.8 -> 0.5.0
- pydata-google-auth 1.5.0 -> 1.6.0
- python-slugify 7.0.0 -> 8.0.0
- sqlalchemy 1.4.46 -> 2.0.0
- tensorflow-io-gcs-filesystem 0.29.0 -> 0.30.0
- tifffile 2022.10.10 -> 2023.1.23.1
- zipp 3.11.0 -> 3.12.0
- Pinned sqlalchemy to version 1.4.46

## 2023-01-12

- Added support for @-mention and email autocomplete in comments
- Improved errors when GitHub notebooks can't be loaded
- Increased color contrast for colors used for syntax highlighting in the code editor
- Added terminal access for custom GCE VM runtimes
- Upgraded Ubuntu from 18.04 LTS to 20.04 LTS ([GitHub issue](#))
- Python package updates:
  - GDAL 2.2.2 -> 2.2.3.
  - NumPy from 1.21.5 to 1.21.6.
  - attrs 22.1.0 -> 22.2.0
  - chardet 3.0.4 -> 4.0.0
  - cloudpickle 1.6.0 -> 2.2.0
  - filelock 3.8.2 -> 3.9.0
  - google-api-core 2.8.2 -> 2.11.0
  - google-api-python-client 1.12.11 -> 2.70.0
  - google-auth-httpplib2 0.0.3 -> 0.1.0
  - google-cloud-bigquery 3.3.5 -> 3.4.1

↳ <Axes: xlabel='City'>



cont

↳ ['Row ID',  
'Order ID',  
'Order Date',  
'Ship Date',  
'Customer ID',

- google-cloud-datastore 2.9.0 -> 2.11.0
- google-cloud-firestore 2.7.2 -> 2.7.3
- google-cloud-storage 2.5.0 -> 2.7.0
- holidays 0.17.2 -> holidays 0.18
- importlib-metadata 5.2.0 -> 6.0.0
- networkx 2.8.8 -> 3.0
- opencv-python-headless 4.6.0.66 -> 4.7.0.68
- pip 21.1.3 -> 22.04
- pip-tools 6.2.0 -> 6.6.2
- prettytable 3.5.0 -> 3.6.0
- requests 2.23.0 -> 2.25.1
- termcolor 2.1.1 -> 2.2.0
- torch 1.13.0 -> 1.13.1
- torchaudio 0.13.0 -> 0.13.1
- torchtext 0.14.0 -> 0.14.1
- torchvision 0.14.0 -> 0.14.1

## 2022-12-06

- Made fallback runtime version available until mid-December ([GitHub issue](#))
- Upgraded to Python 3.8 ([GitHub issue](#))
- Python package updates:
  - jax from 0.3.23 to 0.3.25, jaxlib from 0.3.22 to 0.3.25
  - pyarrow from 6.0.1 to 9.0.0
  - torch from 1.12.1 to 1.13.0
  - torchaudio from 0.12.1 to 0.13.0
  - torchvision from 0.13.1 to 0.14.0
  - torchtext from 0.13.1 to 0.14.0
  - xldr from 1.1.0 to 1.2.0
  - DriveFS from 62.0.1 to 66.0.3
- Made styling of markdown tables in outputs match markdown tables in text cells
- Improved formatting for empty interactive table rows
- Fixed syntax highlighting for variables with names that contain Python keywords ([GitHub issue](#))

## 2022-11-11

```
'Customer Name',
'City',
'State',
'Postal Code',
'Product ID',
'Product Name',
'Sales',
'Profit']
```

```
# Display total sales.      sum()
```

```
print("Total Sales Values: ",df['Sales'].sum().round())
```

```
➦ Total Sales Values: 2297201.0
```

```
#Display Total Profit
```

```
print(" Total Profit Value: ",df['Profit'].sum().round())
```

```
➦ Total Profit Value: 286397.0
```

```
#Display % Profit
```

```
print("% Profit : ", np.round(df['Profit'].sum().round() / df['Sales'].sum().round(), 2))
```

```
➦ % Profit : 12.47 %
```

## ✓ 2. Bi Variate Analysis

```
#group by()
```

- Added more dark editor themes for Monaco (when in dark mode, "Editor colorization" appears as an option in the Editor tab of the Tools → Settings dialog)
- Fixed bug where collapsed forms were deleted on mobile [GitHub issue](#)
- Python package updates:
  - rpy2 from 3.4.0 to 3.5.5 ([GitHub issue](#))
  - notebook from 5.5.0 to 5.7.16
  - tornado from 5.1.1 to 6.0.4
  - tensorflow\_probability from 0.16.0 to 0.17.0
  - pandas-gbq from 0.13.3 to 0.17.9
  - protobuf from 3.17.3 to 3.19.6
  - google-api-core[grpc] from 1.31.5 to 2.8.2
  - google-cloud-bigquery from 1.21.0 to 3.3.5
  - google-cloud-core from 1.0.1 to 2.3.2
  - google-cloud-datastore from 1.8.0 to 2.9.0
  - google-cloud-firestore from 1.7.0 to 2.7.2
  - google-cloud-language from 1.2.0 to 2.6.1
  - google-cloud-storage from 1.18.0 to 2.5.0
  - google-cloud-translate from 1.5.0 to 3.8.4

## 2022-10-21

- Launched a single-click way to get from BigQuery to Colab to further explore query results ([announcement](#))
- Launched [Pro, Pro+, and Pay As You Go](#) to 19 additional countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, Greece, Hungary, Latvia, Lithuania, Norway, Portugal, Romania, Slovakia, Slovenia, and Sweden ([tweet](#))
- Updated jax from 0.3.17 to 0.3.23, jaxlib from 0.3.15 to 0.3.22, TensorFlow from 2.8.2 to 2.9.2, CUDA from 11.1 to 11.2, and cuDNN from 8.0 to 8.1 ([backend-info](#))
- Added a readonly option to [drive.mount](#)
- Fixed bug where Xarray was not working ([GitHub issue](#))

cat

```

➞ ['Ship Mode',
   'Segment',
   'Country',
   'Region',
   'Category',
   'Sub-Category',
   'Quantity',
   'Discount']

```

#1. Display Region wise --- i. Total sales      ii.Total profit

#Context: 'Region' , 'Sales'

#1. Total Region wise Sales

```
df.groupby('Region')['Sales'].sum().round()
```

```

➞ Region
Central    501240.0
East       678781.0
South      391722.0
West       725458.0
Name: Sales, dtype: float64

```

# 2. % Total Sales

```
np.round(df.groupby('Region')['Sales'].sum().round(2) / df['Sales'].sum().rou
```

```

➞ Region
Central    21.82
East       29.55
South      17.05
West       31.58
Name: Sales, dtype: float64

```

- Modified Markdown parsing to ignore block quote symbol within MathJax ([GitHub issue](#))

## 2022-09-30

- Launched [Pay As You Go](#), allowing premium GPU access without requiring a subscription
- Added vim and tcllib to our runtime image
- Fixed bug where open files were closed on kernel disconnect ([GitHub issue](#))
- Fixed bug where the play button/execution indicator was not clickable when scrolled into the cell output ([GitHub issue](#))
- Updated the styling for form titles so that they avoid obscuring the code editor
- Created a GitHub repo, [backend-info](#), with the latest apt-list.txt and pip-freeze.txt files for the Colab runtime ([GitHub issue](#))
- Added [files.upload\\_file\(filename\)](#) to upload a file from the browser to the runtime with a specified filename

## 2022-09-16

- Upgraded pymc from 3.11.0 to 4.1.4, jax from 0.3.14 to 0.3.17, jaxlib from 0.3.14 to 0.3.15, fsspec from 2022.8.1 to 2022.8.2
- Modified our save flow to avoid persisting Drive filenames as titles in notebook JSON
- Updated our [Terms of Service](#)
- Modified the Jump to Cell command to locate the cursor at the end of the command palette input (Jump to cell in Tools → Command palette in a notebook with section headings)
- Updated the styling of the Drive notebook comment UI
- Added support for terminating your runtime from code: python from google.colab import runtime runtime.unassign()
- Added regex filter support to the Recent notebooks dialog

## #3. Total Region wise Profit

```
df.groupby('Region')['Profit'].sum().round(2)
```

```

→ Region
Central    39706.36
East       91522.78
South      46749.43
West       108418.45
Name: Profit, dtype: float64

```

## #4. Region wise Profit contribution

```
np.round(df.groupby('Region')['Profit'].sum() / df['Profit'].sum() * 100 , 2 )
```

```

→ Region
Central    13.86
East       31.96
South      16.32
West       37.86
Name: Profit, dtype: float64

```

## # 6. Display 'Segment' wise

```
# i. Total Sales
```

```
# ii Total Profit.
```

## # Segment wise 'Sales'

```
df.groupby('Segment')['Sales'].sum().round(2)
```

```

→ Segment
Consumer    1161401.34
Corporate    706146.37
Home Office  429653.15
Name: Sales, dtype: float64

```

- Inline google.colab.files.upload JS to fix files.upload() not working ([GitHub issue](#))

**2022-08-26**

- Upgraded PyYAML from 3.13 to 6.0 ([GitHub issue](#)), drivefs from 61.0.3 to 62.0.1
- Upgraded TensorFlow from 2.8.2 to 2.9.1 and ipywidgets from 7.7.1 to 8.0.1 but rolled both back due to a number of user reports ([GitHub issue](#), [GitHub issue](#))
- Stop persisting inferred titles in notebook JSON ([GitHub issue](#))
- Fix bug in background execution which affected some Pro+ users ([GitHub issue](#))
- Fix bug where Download as .py incorrectly handled text cells ending in a double quote
- Fix bug for Pro and Pro+ users where we weren't honoring the preference (Tools → Settings) to use a temporary scratch notebook as the default landing page
- Provide undo/redo for scratch cells
- When writing ipynb files, serialize empty multiline strings as [ ] for better consistency with JupyterLab

**2022-08-11**

- Upgraded ipython from 5.5.0 to 7.9.0, fbprophet 0.7 to prophet 1.1, tensorflow-datasets from 4.0.1 to 4.6.0, drivefs from 60.0.2 to 61.0.3, pytorch from 1.12.0 to 1.12.1, numba from 0.51 to 0.56, and lxml from 4.2.0 to 4.9.1
- Loosened our requests version requirement ([GitHub issue](#))
- Removed support for TensorFlow 1
- Added Help → Report Drive abuse for Drive notebooks
- Fixed indentation for Python lines ending in [
- Modified styling of tables in Markdown to left-align them rather than centering them



```
# Segment wise 'Profit'
```

```
df.groupby('Segment')['Profit'].sum().round(2)
```

```

→ Segment
Consumer      134119.21
Corporate      91979.13
Home Office    60298.68
Name: Profit, dtype: float64

```

```
# 7. Display city wise sales.
```

```
df.groupby('City')['Sales'].sum().nlargest(5)
```

```

→ City
New York City    256368.161
Los Angeles      175851.341
Seattle          119540.742
San Francisco    112669.092
Philadelphia     109077.013
Name: Sales, dtype: float64

```

```
# 8. Display Regiowise total sales of each Segment.
```

```
df.groupby(['Region', 'Segment'])['Sales'].sum().unstack().round()
```

```

→
Segment  Consumer  Corporate  Home Office
Region
Central  252031.0   157996.0   91213.0
East     350908.0   200409.0   127464.0
South    195581.0   121886.0   74255.0
West     362881.0   225855.0   136722.0

```

- Fixed special character replacement when copying interactive tables as Markdown
- Fixed ansi 8-bit color parsing ([GitHub issue](#))
- Configured logging to preempt transitive imports and other loading from implicitly configuring the root logger
- Modified forms to use a value of None instead of causing a parse error when clearing raw and numeric-typed form fields

## 2022-07-22

- Update scipy from 1.4.1 to 1.7.3, drivefs from 59.0.3 to 60.0.2, pytorch from 1.11 to 1.12, jax & jaxlib from 0.3.8 to 0.3.14, opencv-python from 4.1.2.30 to 4.6.0.66, spaCy from 3.3.1 to 3.4.0, and dlib from 19.18.0 to 19.24.0
- Fix Open in tab doc link which was rendering incorrectly ([GitHub issue](#))
- Add a preference for the default tab orientation to the Site section of the settings menu under Tools → Settings
- Show a warning for USE\_AUTH\_EPHEM usage when running authenticate\_user on a TPU runtime ([code](#))

## 2022-07-01

- Add a preference for code font to the settings menu under Tools → Settings
- Update drivefs from 58.0.3 to 59.0.3 and spacy from 2.2.4 to 3.3.1
- Allow [display\\_data](#) and [execute\\_result](#) text outputs to wrap, matching behavior of JupyterLab (does not affect stream outputs/print statements).
- Improve LSP handling of some magics, esp. %%writefile ([GitHub issue](#)).
- Add a [FAQ entry](#) about the mount Drive button behavior and include link buttons for each FAQ entry
- Fix bug where the notebook was sometimes hidden behind other tabs on load when in single pane view.

# The End

- Fix issue with inconsistent scrolling when an editor is in multi-select mode.
- Fix bug where clicking on a link in a form would navigate away from the notebook
- Show a confirmation dialog before performing Replace all from the Find and replace pane.

## 2022-06-10

- Update drivefs from 57.0.5 to 58.0.3 and tensorflow from 2.8.0 to 2.8.2
- Support more than 100 repos in the GitHub repo selector shown in the open dialog and the clone to GitHub dialog
- Show full notebook names on hover in the open dialog
- Improve the color contrast for links, buttons, and the ipywidgets.Accordion widget in dark mode

## 2022-05-20

- Support URL params for linking to some common pref settings: [force\\_theme=dark](#), [force\\_corgi\\_mode=1](#), [force\\_font\\_size=14](#). Params forced by URL are not persisted unless saved using Tools → Settings.
- Add a class markdown-google-sans to allow Markdown to render in Google Sans
- Update monaco-vim from 0.1.19 to 0.3.4
- Update drivefs from 55.0.3 to 57.0.5, jax from 0.3.4 to 0.3.8, and jaxlib from 0.3.2 to 0.3.7

## 2022-04-29

- Added 🦄 mode (under Miscellaneous in Tools → Settings)
- Added "Disconnect and delete runtime" option to the menu next to the Connect button
- Improved rendering of filter options in an interactive table
- Added git-lfs to the base image

Row ID	Order ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Customer Segment	Region
1250	1251	CA-115602	2017-12-18	2017-12-24	Standard Class	DJ-13630	Doug Jacobs	Consumer	Un Sta
9985	9986	CA-15215	2015-05-17	2015-05-23	Standard Class	CG-12520	Stewart Gwinski	Consumer	Un Sta
9944	9945	CA-145415	2015-04-12	2015-04-17	Second Class	RD-19660	Robert Dilbeck	Home Office	Un Sta
9986	9987	CA-125794	2016-08-29	2016-10-03	Standard Class	ML-17410	Maris LaWare	Consumer	Un Sta
2437	2438	CA-164210	2014-11-18	2014-11-20	Second Class	PW-19240	Pierre Wener	Consumer	Un Sta
		US-							..

- Updated torch from 1.10.0 to 1.11.0, jupyter-core from 4.9.2 to 4.10.0, and cmake from 3.12.0 to 3.22.3
- Added more details to our [FAQ](#) about unsupported uses (using proxies, downloading torrents, etc.)
- Fixed [issue](#) with apt-get dependencies

2022-04-15

- Add an option in the file browser to show hidden files.
- Upgrade gdown from 4.2.0 to 4.4.0, google-api-core[grpc] from 1.26.0 to 1.31.5, and pytz from 2018.4 to 2022.1

2022-03-25

- Launched [Pro/Pro+](#) to 12 additional countries: Australia, Bangladesh, Colombia, Hong Kong, Indonesia, Mexico, New Zealand, Pakistan, Philippines, Singapore, Taiwan, and Vietnam
- Added [google.colab.auth.authenticate\\_service\\_account](#) to support using [Service Account keys](#)