Install Conda on Google Colab

`condacolab` simplifies the setup as much as possible, but there are some gotchas.

⚠️ **Read this before continuing!**

- The `condacolab` commands need to be run as the first Code cell!
- Once you run `condacolab.install()`, the Python kernel will be restarted. This is **normal and expected**. After that, you can continue running the cells below like normal.
- Do not use the `Run all` option. Run the `condacolab` cell *individually* and wait for the kernel to restart. **Only then**, you can run all cells if you want.
- You can only use the `base` environment. Do not try to create new ones; instead update `base` with either:
  - `conda install <packages>`
  - `conda env update -n base -f environment.yml`
- If you want to use GPUs, make sure you are using such an instance before starting!
- If you get an error, please raise an issue here.

```
!pip install -q condacolab
import condacolab
condacolab.install()
```

```
⏬ Downloading https://github.com/jaimergp/miniforge/releases/latest/download/Mambaforge-colab-Linux-x86_64.sh...
📦 Installing...
📌 Adjusting configuration...
🩹 Patching environment...
⏲ Done in 0:00:15
🔁 Restarting kernel...
```

```
import condacolab
condacolab.check()
```

```
✨🍰✨ Everything looks OK!
```

```
!mamba install -q openmm
```

```
  Package           Version  Build                 Channel                    Size
─────────────────────────────────────────────────────────────────────────────────
  Install:
─────────────────────────────────────────────────────────────────────────────────

  cudatoolkit        11.0.3  h15472ef_8            conda-forge/linux-64     952 MB
  fftw                3.3.9  nompi_hcdd671c_101    conda-forge/linux-64       6 MB
  libblas             3.9.0  8_openblas            conda-forge/linux-64      11 KB
  libcblas            3.9.0  8_openblas            conda-forge/linux-64      11 KB
  libgfortran-ng      9.3.0  hff62375_18           conda-forge/linux-64      22 KB
  libgfortran5        9.3.0  hff62375_18           conda-forge/linux-64       2 MB
  liblapack           3.9.0  8_openblas            conda-forge/linux-64      11 KB
  libopenblas        0.3.12  pthreads_h4812303_1   conda-forge/linux-64       9 MB
  numpy              1.20.1  py37haa41c4c_0        conda-forge/linux-64       6 MB
  ocl-icd            2.2.14  h7f98852_0            conda-forge/linux-64     118 KB
  ocl-icd-system      1.0.0  1                     conda-forge/linux-64       4 KB
  openmm              7.5.0  py37h01de88b_6        conda-forge/linux-64      11 MB

  Summary:

  Install: 12 packages

  Total download: 986 MB

─────────────────────────────────────────────────────────────────────────────────

Preparing transaction: ...working... done
Verifying transaction: ...working... done
Executing transaction: ...working... By downloading and using the CUDA Toolkit conda packages, you accept the terms and conditions of the

done
```

For the CUDA platform to be available, make sure you are using a GPU environment.

```
from simtk.testInstallation import main
main()
```

```
OpenMM Version: 7.5
Git Revision: b49b82efb5a253a7c891ca084b3370e181de2ea3

There are 4 Platforms available:
```

```
1 Reference - Successfully computed forces
2 CPU - Successfully computed forces
3 CUDA - Successfully computed forces
4 OpenCL - Successfully computed forces

Median difference in forces between platforms:

Reference vs. CPU: 6.29199e-06
Reference vs. CUDA: 6.73078e-06
CPU vs. CUDA: 7.48056e-07
Reference vs. OpenCL: 6.75891e-06
CPU vs. OpenCL: 8.11491e-07
CUDA vs. OpenCL: 2.68874e-07

All differences are within tolerance.
```

Step 1. IMPORT Libraries

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

Step 2. IMPORT DATASET

```
coupon=pd.read_csv("COUPON.CSV")
```

```
coupon.isnull()
```

	Gender	Item1	Item2	Item3	transactions	Coupon	coupon_discount
0	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...

```
coupon.head()
```

	Gender	Item1	Item2	Item3	transactions	Coupon	coupon_discount
0	Male	Computer Games	Toys	Educational Products	436274.446700	Kids Apparel	117.276864
1	Female	Cosmetics	Womens Wear	Kids Wear	5323.510000	Womens Apparel	70.885046
2	Female	Cosmetics	Womens Wear	Candy	48980.998330	Womens Apparel	342.826195

```
coupon.tail()
```

	Gender	Item1	Item2	Item3	transactions	Coupon	coupon_discount
145	Female	Handbag	Movies	Tie	47223.323000	Womens Apparel	76.259010
146	Male	Computer Games	Toys	Educational Products	70553.563330	Kids Apparel	109.160070
147	Male	Cosmetics	Womens Wear	Candy	56510.570000	Womens Apparel	103.032016

```
coupon[20:30]
```

	Gender	Item1	Item2	Item3	transactions	Coupon	coupon_discount
20	Male	Cosmetics	Movies	Candy	147023.28500	Womens Apparel	109.857869
21	Female	Mens Wear	Womens Wear	Educational Products	54049.13833	Womens Apparel	89.472987
22	Female	Computer Games	Movies	Kids Wear	56796.97833	Kids Apparel	91.595460
23	Female	Mens Wear	Gym Shoes	Tie	32494.81750	Mens Apparel	128.950956
24	Female	Mens Wear	Board Games	Candy	33872.26448	Kids Apparel	42.769784
25	Male	Cosmetics	Movies	Tie	15031.82000	Mens Apparel	43.621101
26	Female	Mens Wear	Board	Tie	21758.84167	Mens Apparel	83.423357

```
coupon.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 7 columns):
 #   Column           Non-Null Count  Dtype  
---  ------           --------------  -----  
 0   Gender           150 non-null    object 
 1   Item1            150 non-null    object 
 2   Item2            150 non-null    object 
 3   Item3            150 non-null    object 
 4   transactions     150 non-null    float64
 5   Coupon           150 non-null    object 
 6   coupon_discount  150 non-null    float64
dtypes: float64(2), object(5)
memory usage: 8.3+ KB
```

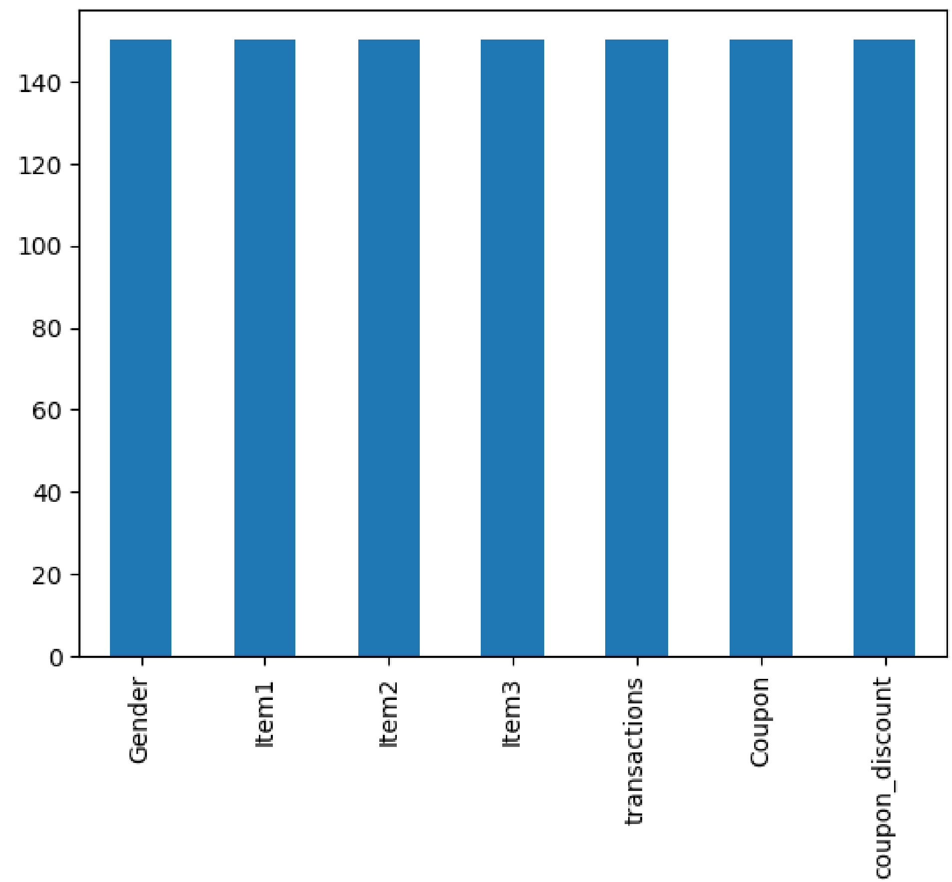
```
coupon.describe()
```

	transactions	coupon_discount
count	1.500000e+02	150.000000
mean	4.643452e+05	165.435743
std	2.194338e+06	136.299236

Step 3.VISUALIZE DATASET

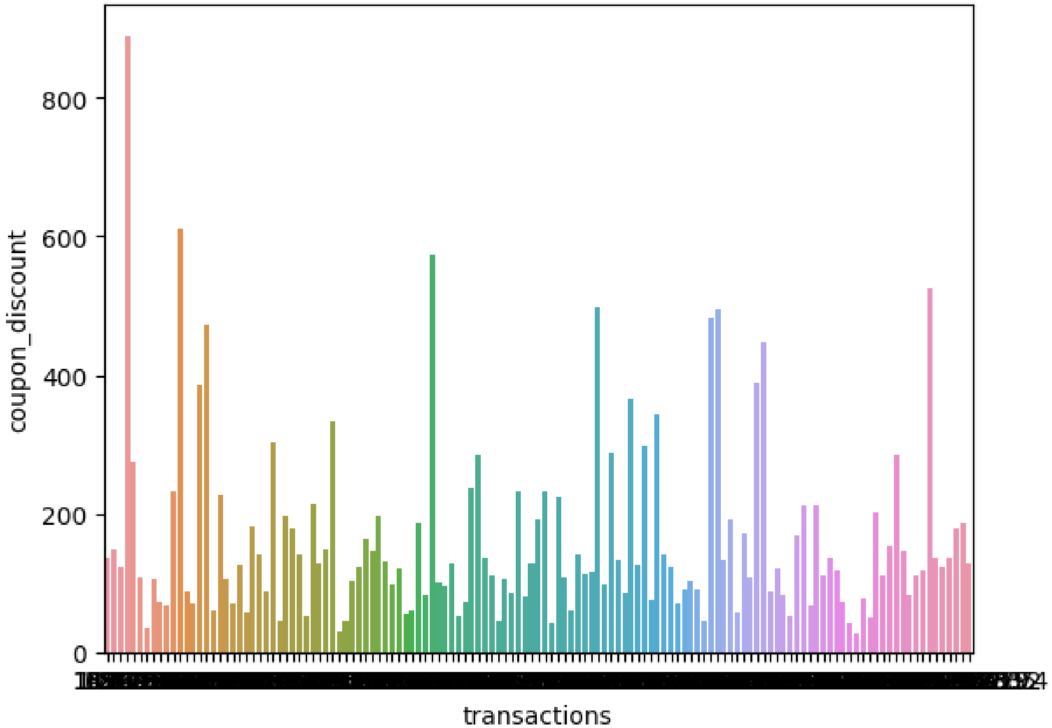
```
coupon.count().plot.bar()
```

```
<Axes: >
```



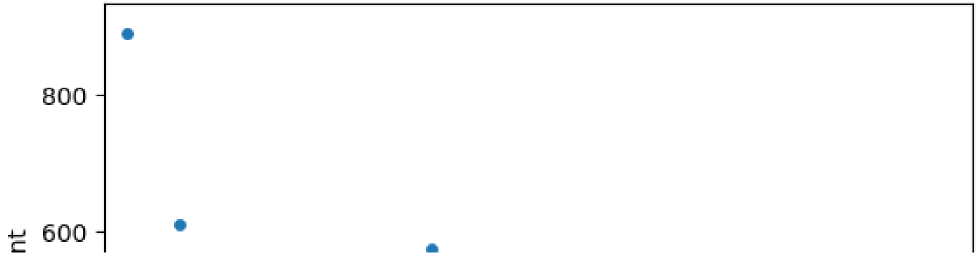
```
sns.barplot(x='transactions',y='coupon_discount',data=coupon)
```

```
<Axes: xlabel='transactions', ylabel='coupon_discount'>
```



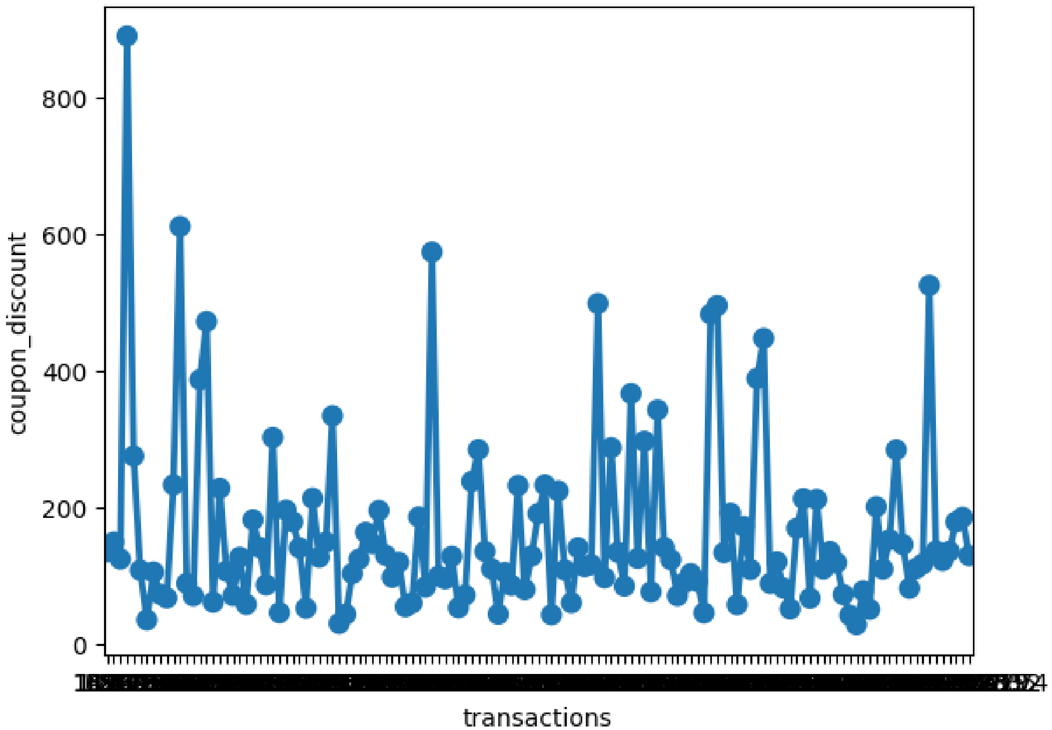

```
sns.stripplot(x=coupon.transactions,y=coupon.coupon_discount,data=coupon)
```

```
<Axes: xlabel='transactions', ylabel='coupon_discount'>
```

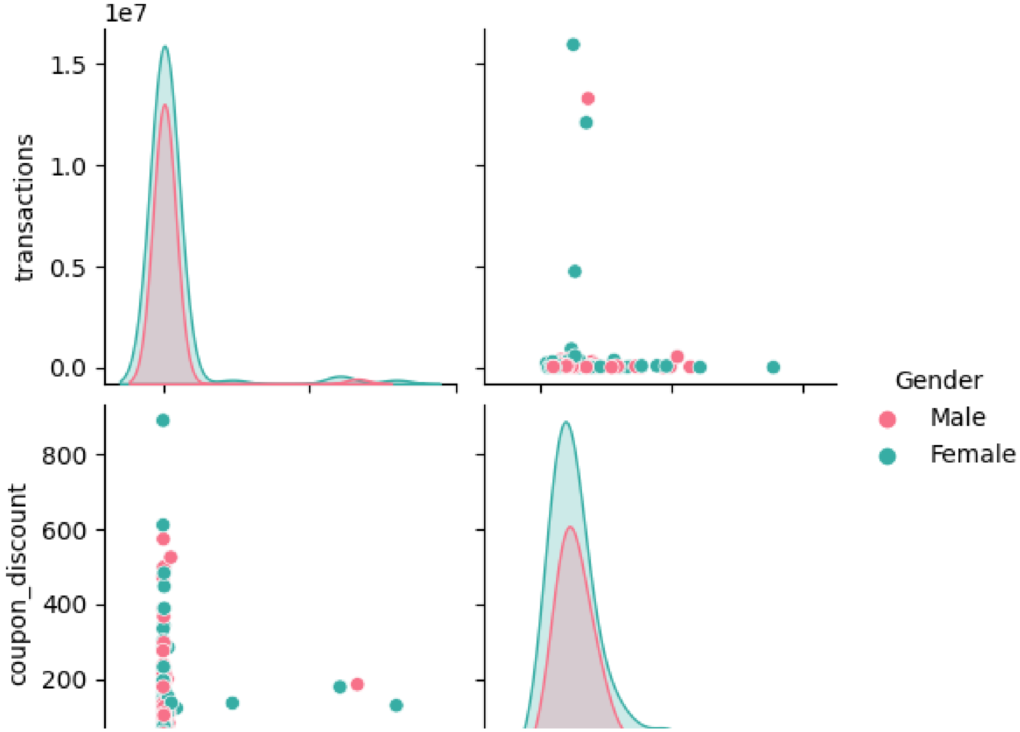


```
sns.pointplot(x='transactions',y='coupon_discount',data=coupon)
```

```
<Axes: xlabel='transactions', ylabel='coupon_discount'>
```




```
sns.pairplot(coupon,hue = 'Gender',diag_kind = "kde",kind = "scatter",palette = "husl")
plt.show()
```



Step 4. CREATE TESTING AND TRAINING DATASET

```
y = coupon['coupon_discount']
X = coupon[['transactions']]
```

```
X
```

	transactions
0	436274.446700
1	5323.510000
2	48980.998330
3	3697.783333
4	289181.848300
...	...

```
y
```

```
0      117.276864
1       70.885046
2      342.826195
3       88.694903
4      200.991719
          ...
145     76.259010
146    109.160070
147    103.032016
148     57.905532
149     53.048012
Name: coupon_discount, Length: 150, dtype: float64
```

```
from sklearn.model_selection import train_test_split
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25)
```

Step 5. TRAIN THE MODEL

```
X_train.shape
from sklearn.linear_model import LinearRegression
```

```
regressor = LinearRegression(fit_intercept = True)
regressor.fit(X_train,y_train)
```

▸ LinearRegression

```
print('Linear Model Coefficient (m): ', regressor.coef_)
print('Linear Model Coefficient (b): ', regressor.intercept_)
```

```
Linear Model Coefficient (m):  [-2.55172129e-07]
Linear Model Coefficient (b):  158.14091228970713
```

Step 6. TEST THE MODEL

```
y_predict = regressor.predict( X_test)
y_predict
```

```
array([158.13795975, 158.13984167, 158.09633953, 158.00748046,
       154.74970747, 158.12433081, 158.13856235, 158.14058008,
       158.12929589, 158.13672825, 158.12541738, 158.07733488,
       158.13856235, 158.14028724, 158.12886221, 158.12962791,
       158.13573854, 158.0871217 , 158.12290899, 158.13820093,
       158.05184326, 158.13280224, 158.0871217 , 158.11647716,
       158.04896804, 158.11163386, 158.04449861, 158.14073824,
       158.13216133, 158.12398963, 158.13280224, 158.00748046,
       158.14044211, 158.13879097, 156.93136694, 158.12641928,
       158.13226144, 158.1224741 ])
```

```
y_test
```

```
44     140.762111
43     387.012500
132     72.370757
17     524.907167
74     185.438313
29     495.274445
142     45.881078
121    275.251875
119    296.931289
61     124.176118
109     45.543524
99      77.790561
7       45.881078
45      72.850639
145     76.259010
95     366.898471
134     97.768613
```

```
77      28.273801
146    109.160070
136    178.931493
126    152.791425
107    232.083176
79      28.273801
110     81.821518
9      284.346347
123    169.525695
53     145.603602
47     124.310000
87     108.356342
92     133.806810
42     232.083176
127    524.907167
36     107.986250
48     302.470667
93     135.619609
22      91.595460
58     224.054986
65     388.746037
Name: coupon_discount, dtype: float64
```

```
plt.scatter(X_train, y_train, color = 'gray')
plt.plot(X_train, regressor.predict(X_train), color = 'red')
plt.ylabel('transactions [dollars]')
plt.xlabel('coupon_discounts[dollars]')
plt.title('transactions vs coupon_discount @chennai shopping mall(Training dataset)')
```

```
Text(0.5, 1.0, 'transactions vs coupon_discount @chennai shopping mall(Training dataset)')
```

Step 7.Evaluate the model

```
from sklearn.metrics import mean_squared_error
acc = mean_squared_error(y_test,y_predict)
print("Mean Square Error : ",acc)
wts = regressor.coef_
incpt = regressor.intercept_
print("Slope :",wts,"\nIntercept ",incpt)
```

```
Mean Square Error :  19554.45459775169
Slope : [-2.55172129e-07]
Intercept  158.14091228970713
```

✓ 0s completed at 11:36 PM