Untitled28.ipynb - Colab

colab.research.google.com/drive/1BOXgatkPe8O8SHAjG97WaPN65p95GpsB

Commands + Code + Text ▷ Run all Connect

```python
import pandas as pd
import kagglehub
import os

# Ensure the dataset is downloaded and get its path
path = kagglehub.dataset_download("afnansaifafnan/electric-car-performance-and-battery-dataset")

# Construct the full file path
file_name = os.listdir(path)[0] # Assuming there's only one CSV file in the directory
full_file_path = os.path.join(path, file_name)

# Read the CSV file
data = pd.read_csv(full_file_path)
data.info()
```

```
Downloading from https://www.kaggle.com/api/v1/datasets/download/afnansaifafnan/electric-car-performance-and-battery-dataset?dataset_version_number=1...
100%|██████████| 16.2k/16.2k [00:00<00:00, 15.2MB/s]Extracting files...
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 478 entries, 0 to 477
Data columns (total 22 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   brand                   478 non-null    object
 1   model                   477 non-null    object
 2   top_speed_kmh           478 non-null    int64
 3   battery_capacity_kWh    478 non-null    float64
 4   battery_type            478 non-null    object
 5   number_of_cells         276 non-null    float64
 6   torque_nm               471 non-null    float64
 7   efficiency_wh_per_km    478 non-null    int64
 8   range_km                478 non-null    int64
 9   acceleration_0_100_s    478 non-null    float64
 10  fast_charging_power_kw_dc  477 non-null  float64
 11  fast_charge_port        477 non-null    object
 12  towing_capacity_kg      452 non-null    float64
 13  cargo_volume_l          477 non-null    object
 14  seats                   478 non-null    int64
 15  drivetrain              478 non-null    object
 16  segment                 478 non-null    object
 17  length_mm               478 non-null    int64
 18  width_mm                478 non-null    int64
 19  height_mm               478 non-null    int64
 20  car_body_type           478 non-null    object
 21  source_url              478 non-null    object
dtypes: float64(6), int64(7), object(9)
memory usage: 82.3+ KB
```

Variables Terminal

Type here to search Top Stories ENG 10:20 PM 10-11-2025

Untitled28.ipynb - Colab

colab.research.google.com/drive/1BOXgatkPe8O8SHAjG97WaPN65p95GpsB

Commands + Code + Text ▷ Run all Connect

```
x = data.drop('range_km', axis = 1)
y = data['range_km']
```

```
x
```

	brand	model	top_speed_kmh	battery_capacity_kWh	battery_type	number_of_cells	torque_nm	efficiency_wh_per_km	acceleration_0_100_s	fast_charging_power_kw_dc	...	towing_capacity_kg	cargo_volume_l	seats	drivetrain	segment
0	Abarth	500e Convertible	155	37.8	Lithium-ion	192.0	235.0	166	7.0	67.0	...	0.0	185	4	FWD	B - Compact
1	Abarth	500e Hatchback	155	37.8	Lithium-ion	192.0	235.0	149	7.0	67.0	...	0.0	185	4	FWD	B - Compact
2	Abarth	600e Scorpionissima	200	50.8	Lithium-ion	102.0	345.0	158	5.9	79.0	...	0.0	360	5	FWD	JB - Compact
3	Abarth	600e Turismo	200	50.8	Lithium-ion	102.0	345.0	158	6.2	79.0	...	0.0	360	5	FWD	JB - Compact
4	Always	U5	150	60.0	Lithium-ion	NaN	310.0	156	7.5	78.0	...	NaN	496	5	FWD	JC - Medium
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
473	Zeekr	7X Premium RWD	210	71.0	Lithium-ion	NaN	440.0	148	6.0	240.0	...	2000.0	539	5	RWD	JD - Large
474	Zeekr	X Core RWD (MY25)	190	49.0	Lithium-ion	NaN	343.0	148	5.9	70.0	...	1600.0	362	5	RWD	JB - Compact
475	Zeekr	X Long Range RWD (MY25)	190	65.0	Lithium-ion	NaN	343.0	146	5.6	114.0	...	1600.0	362	5	RWD	JB - Compact
476	Zeekr	X Privilege AWD (MY25)	190	65.0	Lithium-ion	NaN	543.0	153	3.8	114.0	...	1600.0	362	5	AWD	JB - Compact
477	fiefly	NaN	150	41.2	Lithium-ion	112.0	200.0	125	8.1	65.0	...	0.0	404	5	RWD	B - Compact

Variables Terminal

Type here to search Top Stories ENG 10:20 PM 10-11-2025

Untitled28.ipynb - Colab

colab.research.google.com/drive/1BOXgatkPe8O8SHAjG97WaPN65p95GpsB

Commands + Code + Text ▷ Run all Connect

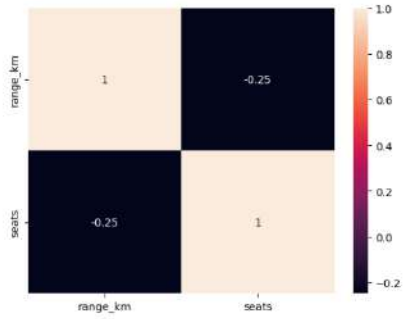
```
data.columns
```

```
Index(['brand', 'model', 'top_speed_kmh', 'battery_capacity_kWh',
       'battery_type', 'number_of_cells', 'torque_nm', 'efficiency_wh_per_km',
       'range_km', 'acceleration_0_100_s', 'fast_charging_power_kw_dc',
       'fast_charge_port', 'towing_capacity_kg', 'cargo_volume_l', 'seats',
       'drivetrain', 'segment', 'length_mm', 'width_mm', 'height_mm',
       'car_body_type', 'source_url'],
      dtype='object')
```

```
import seaborn as sns
import matplotlib.pyplot as plt
```

```
data = data[['range_km','seats']]
```

```
sns.heatmap(data.corr(numeric_only=True), annot = True)
plt.show()
```




```
x = data[['top speed kmh','battery capacity kWh','torque nm']]
```

Variables Terminal

Type here to search Top Stories ENG 10:20 PM 10-11-2025

Untitled28.ipynb - Colab

colab.research.google.com/drive/1BOXgatkPe8O8SHAjG97WaPN65p95GpsB

Commands + Code + Text ▷ Run all Connect

```
x = data[['top_speed_kmh','battery_capacity_kWh','torque_nm']]
y = data[['range_km']]
```

```
x
```

	top_speed_kmh	battery_capacity_kWh	torque_nm
0	155	37.8	235.0
1	155	37.8	235.0
2	200	50.8	345.0
3	200	50.8	345.0
4	150	60.0	310.0
...	...	...	...
473	210	71.0	440.0
474	190	49.0	343.0
475	190	65.0	343.0
476	190	65.0	543.0
477	150	41.2	200.0

478 rows × 3 columns

```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y, test_size = 0.2, random_state = 42)
```

```
from sklearn.linear_model import LinearRegression
model = LinearRegression()
```

```
# Impute missing values in x_train with the mean of each column
x_train_imputed = x_train.fillna(x_train.mean())

# Train the model with the imputed data
model.fit(x_train_imputed, y_train)
```

Variables Terminal

Type here to search Top Stories ENG 10:20 PM 10-11-2025

Untitled28.ipynb - Colab

colab.research.google.com/drive/1BOXgatkPe8O8SHAjG97WaPN65p95GpsB

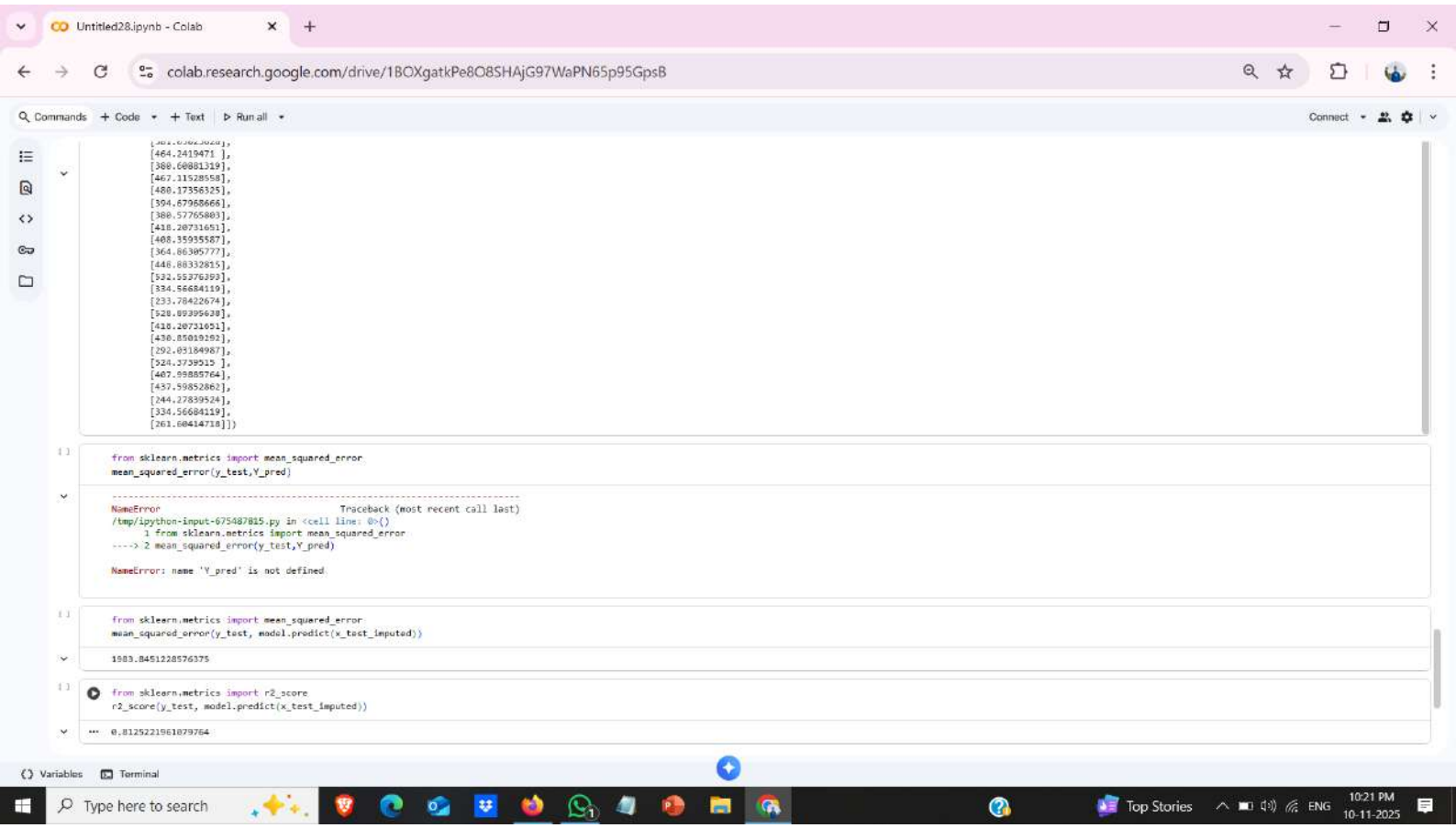
Commands + Code + Text ▷ Run all Connect

```
Y_pred = model.predict(x_test_imputed)
```

```
array([[470.80714829],
       [487.86798168],
       [323.53906361],
       [471.07746484],
       [394.67968666],
       [295.40306121],
       [384.66462086],
       [349.59602558],
       [321.13432154],
       [306.83179946],
       [518.66933877],
       [295.46306121],
       [244.94141529],
       [494.03115375],
       [263.35082974],
       [366.53252453],
       [498.24432973],
       [240.87377056],
       [263.35082974],
       [498.86334472],
       [224.70927643],
       [244.27839524],
       [466.40230351],
       [388.83806595],
       [239.45891745],
       [502.75857377],
       [497.63460127],
       [285.3324194 ],
       [368.87876792],
       [482.4080863 ],
       [452.81298618],
       [274.18086871],
       [308.74996685],
       [301.734142  ],
       [349.14536797],
       [502.17702563],
       [388.57765803],
       [394.67968666],
       [407.99885764],
       [324.31448892],
       [379.42667337],
       [450.8784849 ],
       [350.17600222],
       [478.65775986],
       [394.67968666],
       [371.36064945],
       [503.76965143]
```

Variables Terminal

Type here to search Top Stories ENG 10:21 PM 10-11-2025


Untitled28.ipynb - Colab
colab.research.google.com/drive/1BOXgatkPe8O8SHAjG97WaPN65p95GpsB
Commands
+ Code
+ Text
Run all
Connect
[464.2419471],
[388.60881319],
[467.11528558],
[480.17356325],
[394.67968666],
[380.57765803],
[418.20731651],
[408.35935587],
[364.86305777],
[446.88332815],
[532.55376393],
[334.56684119],
[233.78422674],
[528.89395638],
[418.20731651],
[430.85019202],
[292.03184987],
[524.3739515],
[407.99885764],
[437.59852862],
[244.27839524],
[334.56684119],
[261.68414718]])
from sklearn.metrics import mean_squared_error
mean_squared_error(y_test,Y_pred)
NameError Traceback (most recent call last)
/tmp/ipython-input-675487815.py in <cell line: 0>()
1 from sklearn.metrics import mean_squared_error
----> 2 mean_squared_error(y_test,Y_pred)
NameError: name 'Y_pred' is not defined
from sklearn.metrics import mean_squared_error
mean_squared_error(y_test, model.predict(x_test_imputed))
1983.8451228576375
from sklearn.metrics import r2_score
r2_score(y_test, model.predict(x_test_imputed))
0.8125221961079764
Variables
Terminal
Type here to search
Top Stories
ENG
10:21 PM
10-11-2025