

+

+

↑

↺

Filter files by name

🔍

/

Name

▲

...

📁 archive

📁 data

📁 marketing

📄 df_final.pkl

📄 df_final.xlsx

📄 df_test.pkl

📄 df_test.xlsx

📄 df_train.pkl

📄 df_train.xlsx

📄 EDA.pdf

🔗 keys.py

📄 load_data.ipynb

📄 model.ipynb

keys.py

load_data.ipynb

model.ipynb

Python (bw_unit_4)

🔌

🔌

[351]:

```
from keys import *
import json
import pandas as pd
import numpy as np
from json.decoder import JSONDecodeError
from pandas.io.json import json_normalize
import os
from datetime import datetime
import matplotlib.pyplot as plt
import scipy.stats as stats
import statistics as st
```

[370]:

```
def read_csvs(directory, df_combined):
    for filename in os.listdir(directory):
        if filename.endswith('.csv'):
            csv_path = os.path.join(directory, filename)
            df_file = pd.read_csv(csv_path)
            df_combined = pd.concat([df_combined, df_file])
        else:
            continue
    return df_combined

def expand_multi_level_cols(df_combined):
    types = ["creator", "location", "profile", "category" ]

    n = len(df_combined)
    df_out = pd.DataFrame()

    for type_i in types:
        df = df_combined[type_i]
        if type_i == 'photo':
            df_loop = pd.DataFrame(columns=photo_keys)

        elif type_i == 'creator':
            df_loop = pd.DataFrame(columns=creator_keys)

        elif type_i == 'location':
            df_loop = pd.DataFrame(columns=location_keys)
```

Simple

0

\$

2

🔌

Python (bw_unit_4) | Idle

Saving completed

Mode: Command

🔒

Ln 1, Col 1

load_data.ipynb