# Start coding...

import pandas as pd

import numpy as np

# Read in csv

marketing = pd.read\_csv("bank\_marketing.csv")

# Split into the three tables

client = marketing[["client\_id", "age", "job", "marital",

"education", "credit\_default", "mortgage"]]

campaign = marketing[["client\_id", "number\_contacts", "month", "day",

"contact\_duration", "previous\_campaign\_contacts", "previous\_outcome", "campaign\_outcome"]]

economics = marketing[["client\_id", "cons\_price\_idx", "euribor\_three\_months"]]

## Editing the client dataset

# Clean education column

client["education"] = client["education"].str.replace(".", "\_")

client["education"] = client["education"].replace("unknown", np.NaN)

# Clean job column

client["job"] = client["job"].str.replace(".", "\_")

# Clean and convert client columns to bool data type

for col in ["credit\_default", "mortgage"]:

client[col] = client[col].map({"yes": 1,

"no": 0,

"unknown": 0})

client[col] = client[col].astype(bool)

# Editing the campaign dataset

# Change campaign\_outcome to binary values

campaign["campaign\_outcome"] = campaign["campaign\_outcome"].map({"yes": 1,

"no": 0})

# Convert previous\_outcome to binary values

campaign["previous\_outcome"] = campaign["previous\_outcome"].map({"success": 1,

"failure": 0,

"nonexistent": 0})

# Add year column

campaign["year"] = "2022"

# Convert day to string

campaign["day"] = campaign["day"].astype(str)

# Add last\_contact\_date column

campaign["last\_contact\_date"] = campaign["year"] + "-" + campaign["month"] + "-" + campaign["day"]

# Convert to datetime

campaign["last\_contact\_date"] = pd.to\_datetime(campaign["last\_contact\_date"],

format="%Y-%b-%d")

# Clean and convert outcome columns to bool

for col in ["campaign\_outcome", "previous\_outcome"]:

campaign[col] = campaign[col].astype(bool)

# Drop unneccessary columns

campaign.drop(columns=["month", "day", "year"], inplace=True)

# Save tables to individual csv files

client.to\_csv("client.csv", index=False)

campaign.to\_csv("campaign.csv", index=False)

economics.to\_csv("economics.csv", index=False)