Crowdfunding_ETL

Project 2 - Challenge 13

Building an ETL pipeline using Python, Pandas, and either Python dictionary methods or regular expressions to extract and transform the data. After transform the data, create four CSV files and use the CSV file data to create an ERD and a table schema. And upload the CSV file data into a Postgres database.

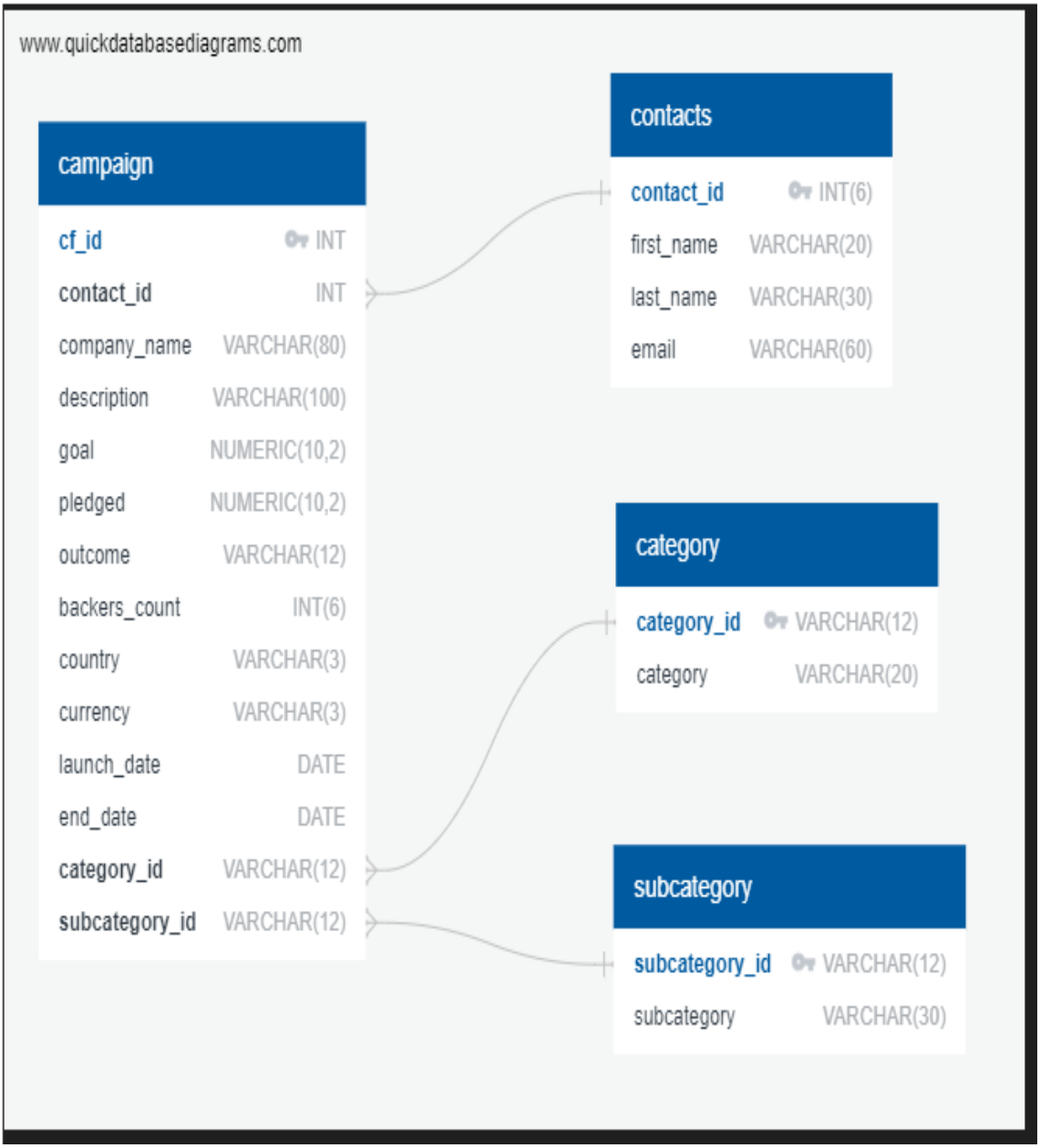
Resources

PgAdmin4

Jupyter

Notebook Visual Studio Code

QuickDBD


www.quickdatabasediagrams.com
campaign
cf_id INT
contact_id INT
company_name VARCHAR(80)
description VARCHAR(100)
goal NUMERIC(10,2)
pledged NUMERIC(10,2)
outcome VARCHAR(12)
backers_count INT(6)
country VARCHAR(3)
currency VARCHAR(3)
launch_date DATE
end_date DATE
category_id VARCHAR(12)
subcategory_id VARCHAR(12)
contacts
contact_id INT(6)
first_name VARCHAR(20)
last_name VARCHAR(30)
email VARCHAR(60)
category
category_id VARCHAR(12)
category VARCHAR(20)
subcategory
subcategory_id VARCHAR(12)
subcategory VARCHAR(30)

	contact_info
0	{"contact_id": 4661, "name": "Cecilia Velasco", "email": "cecilia.velasco@rodrigues.fr"}
1	{"contact_id": 3765, "name": "Mariana Ellis", "email": "mariana.ellis@rossi.org"}
2	{"contact_id": 4187, "name": "Sofie Woods", "email": "sofie.woods@riviere.com"}
3	{"contact_id": 4941, "name": "Jeanette Iannotti", "email": "jeanette.iannotti@yahoo.com"}
4	{"contact_id": 2199, "name": "Samuel Sorgatz", "email": "samuel.sorgatz@gmail.com"}

	contact_id	first_name	last_name	email
0	4661	Cecilia	Velasco	cecilia.velasco@rodrigues.fr
1	3765	Mariana	Ellis	mariana.ellis@rossi.org
2	4187	Sofie	Woods	sofie.woods@riviere.com
3	4941	Jeanette	Iannotti	jeanette.iannotti@yahoo.com
4	2199	Samuel	Sorgatz	samuel.sorgatz@gmail.com
5	5650	Socorro	Luna	socorro.luna@hotmail.com
6	5889	Carolina	Murray	carolina.murray@knight.com
7	4842	Kayla	Moon	kayla.moon@yahoo.de
8	3280	Ariadna	Geisel	ariadna.geisel@rangel.com
9	5468	Danielle	Ladeck	danielle.ladeck@scalfaro.net