Relatório Solução Etapa 2

Programa IT Academy - Processo Seletivo - Edição #15

Nome: Lucas Mariano Leipnitz de Fraga

Instruções para Execução do Programa

Versão do python: 3.9

Todas as bibliotecas utilizadas são nativas do python

Para rodar o programa é preciso executar o arquivo manager.py a partir da pasta program, para que ele consiga encontrar o caminho correto para os arquivos csv e db na pasta data. No caso de UnicodeDecodeError é preciso salvar o arquivo csv no formato utf-8 com BOM.

Introdução

Na etapa dois do processo seletivo, tivemos que desenvolver um programa responsável por ler uma base de dados, disponibilizada em um arquivo csv, e executar diversas tarefas em cima dos dados da base. Como muitas das tarefas podiam ser executadas facilmente por SQL query, eu decidi utilizar o pacote SQLite3, um pacote de python que permite fazer um banco de dados SQL local, para fazer um bando de dados com os dados do arquivo csv.

Após isso, basta chamar o banco de dados para consultar de forma rápida e segura os dados, tratar eles e mostrar para o usuário. Abaixo, cada consulta do banco de dados feita para sua respectiva tarefa:

- [Consultar medicamento pelo nome]:
 "SELECT * FROM Medicamento WHERE SUBSTANCIA LIKE '%{name}%' AND COMERCIALIZACAO_2020 == 'Sim'"
- [Buscar pelo código de barras]:
 SELECT * FROM Medicamento WHERE EAN_1 == '{code}' ORDER BY PMC_0
- [Comparativa PIS/COFINS]:
 "SELECT count(PIS_COFINS) FROM Medicamento where PIS_COFINS == '{param}'"
 Onde param pode ser "Positiva", "Negativa" ou "Neutra"

A partir deste raciocínio, criei uma arquitetura em camadas do projeto, para que o usuário possa solicitar uma tarefa e, sem com que ele tem acesso ao banco, busque as

informações no banco, faça operações lógicas em cima das informações e imprima o resultado para o usuário.

Arquitetura do Projeto

A arquitetura do projeto é uma arquitetura em camadas, onde existem seis módulos, um que se comunica com o usuário, dois que se comunicam com a base de dados, dois que fazem a intermediação entre o usuário e a base de dados e um que serve para armazenar os dados em memória. Escolhi fazer a aplicação com esta arquitetura pois acho ela simples e fácil de implementar, mas também bem modular por ter seis módulos, e bem robusta, já que poucos módulos se comunicam entre si, diminuindo a possibilidade de passarem bugs uns para os outros e segura, já que o usuário só tem comunicação com o módulo de interface. Abaixo, um diagrama UML da arquitetura:

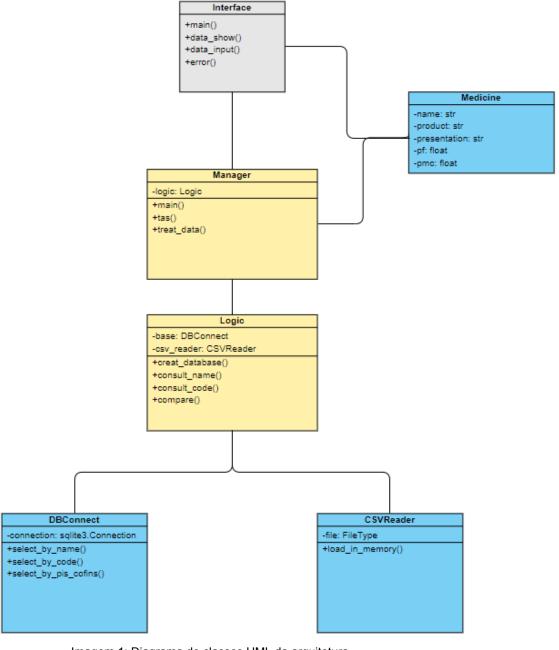


Imagem 1: Diagrama de classes UML da arquitetura.

Módulos

Interface

A classe Interface faz parte do pacote interface e tem o objetivo de se comunicar com o usuário final, imprimindo as opções de tarefas e os resultados paro o usuário:

Imagem 2: Comunicação entre a interface e o usuário via terminal.

Manager

A classe Manager é responsável por fazer a administração entre a interface com o usuário (frontend) e a lógica do programa (backend), armazenar e tratar os dados de entrada e de pesquisa e controlar em que estado o programa está. Ele é o arquivo principal do programa e será o primeiro módulo a ser criado e em seguida, será chamado a sua função main(). Ela começa verificando se o meds.db está na pasta data e, se não estiver, ela chama a função de Logic que será responsável por criar e preencher o banco de dados a partir do arquivo csv. É nela, também, onde está o laço principal do programa, que será executado até o usuário escolher a opção 4- Terminar o programa.

```
main():
entrada: nenhuma
saída: nenhuma
saria: nenh
```

Imagem 3: Função main() de Manager, o laço principal do programa.

No laço principal está a chamada da função main() de Interface, que vai imprimir as mensagens mostradas na imagem 1 e esperar que o usuário digite algo. Além disso, Manager tem as funçõs treat_option() e treat_data que tratam os dados de entrada do usuário, para que ele não digite opções inválidas, letras no lugar de números e caracteres que possam causar erro na consulta ao banco de dados.

Caso os dados inseridos pelo usuário forem válidos, ela chamará uma das três funções task() que vão, então, chamar as funções de Logic com os dados passados pelo usuário para fazer o processamento lógico da requisição. Após isso, Logic vai retornar uma lista de tuplas para representar medicamentos, que será então transformada em um objeto Medicine por Menager, ou três números representando as porcentagens da tarefa 3, e serão enviados para a Interface apresentar o resultado para o usuário.

Logic

A classe Logic é responsável pela lógica das tarefas do programa. É ela quem se comunica com o banco de dados (DBConnect) e o arquivo csv (CSVReader) e não Manager, adicionando uma camada a mais de segurança entre o usuário final e os dados reais no banco de dados, evitando possíveis invasões, além de distribuir melhor o acoplamento do programa.

A classe Logic tem uma série de funções responsáveis por se comunicar com o banco de dados e pedir requisições para ele e se comunicar com o arquivo CSV para povoar o banco de dados com cada linha do arquivo.

```
create_database()
entrada: nenhuma
saida: nenhuma
objetivo: responsável por criar o arquivo SQLite do banco de dados e a tabela Medicamentos, chamar csv_reader para ler o arquivo csv e salvar as linhas em memória e adicionar cada linha ao banco dados.

def create_database(self):
    self.base.open_connection()
    self.base.init_db()
    self.csv_reader.load_in_memory()
    for row in self.csv_reader.load_in_memory()
    for row in self.csv_reader.data:
        | self.base.add_row(row)
        self.base.comsit()
        self.base.comsit()
        self.base.comsit()
```

Imagem 4: Função create_database() de Logic, responsável por preencher o banco de dados. Podemos ver que ela chama a função load_in_memory() de csv_reader para isso, e depois percorre os dados adicionando no banco de dados com a função add row() de base.

Além disso, é o módulo Logic quem vai solicitar as consultas do banco, verificar se a consulta não encontrou resultados e informar Manager. Abaixou, algumas funções de Logic e seus objetivos:

```
consult_name():
entrada: string que contém o nome a ser consultado
saída: lista de tuplas que representa a consulta
objetivo: chama base para fazer a consulta de medicamentos que completam o nome inserido.
def consult_name(self, name):
   self.base.open_connection()
   consult = self.base.select_by_name(name)
   self.base.close_connection()
   meds = []
    for medicine in consult:
       new_med = []
       new_med.append(medicine[0])
       new_med.append(medicine[8])
       new_med.append(medicine[9])
       new_med.append(medicine[13])
       new_med.append(medicine[23])
        meds.append(tuple(new_med))
    return meds
```

Imagem 5: Função consult_name() de Lógic responsável por fazer a consulta por nome

```
consult_code():
entrada: string que contém o código a ser consultado
saída: lista de tuplas que representa a consulta
objetivo: chama base para fazer a consulta de medicamentos com o código inserido.
def consult_code(self, code):
    self.base.open_connection()
    consult = self.base.select_by_code(code)
    self.base.close_connection()
    print(consult)
    if(len(consult) == 0):
        return None
        meds = []
        new_med = []
        new_med.append(consult[0][0])
        new_med.append(consult[0][8])
        new_med.append(consult[0][9])
        new_med.append(float(consult[0][13].replace(',','.')))
        new_med.append(float(consult[0][23].replace(',','.')))
        meds.append(tuple(new_med))
        new\_med = []
        new_med.append(consult[-1][0])
        new med.append(consult[-1][8])
        new_med.append(consult[-1][9])
        new_med.append(float(consult[-1][13].replace(',','.')))
        new_med.append(float(consult[-1][23].replace(',','.')))
        meds.append(tuple(new_med))
    return meds
compare()):
entrada:
saída: três inteiros que prepresentam as porcentagens
objetivo: faz a comparação de avaliações PIS/COFINS
def compare(self):
    self.base.open_connection()
    neutral_counter = self.base.count_by_pis_cofins("Neutra")[0][0]
    positive_counter = self.base.count_by_pis_cofins("Positiva")[0][0]
    negative_counter = self.base.count_by_pis_cofins("Negativa")[0][0]
    self.base.close_connection()
    total = negative_counter + neutral_counter + positive_counter
    negative_percent = negative_counter*100/total
    positive_percent = positive_counter*100/total
    neutral_percent = neutral_counter*100/total
    return negative_percent, neutral_percent, positive_percent
```

Imagem 6: Funções consult_code() e compare()

DBConnect, CSVReader e Medicine

Por último, temos os módulos que fazem parte do pacote data. Eles tem o objetivo de se comunicar com os dados persistentes do programa (DBConnect com o banco de dados e CSVReader com o arquivo csv) e armazenar os dados em memória de forma clara e de fácil compreensão (Medicine). Falaremos de cada um.

Medicine é uma dataclass, uma classe específica para armazenar dados e não fazer operações com eles. Implementei essa classe para armazenar os dados de forma clara para debugar o programa mais facilmente. Abaixo, uma imagem da classe com seus atributos:

```
@dataclasss
@dataclass
class Medicine:
    name: str
    product: str
    presentation: str
    pf: float
    pmc: float

def __init__(self, name="", product="", presentation="", pf=0.0, pmc=0.0):
        self.name = name
        self.product = product
        self.presentation = presentation
        self.pf = pf
        self.pmc = pmc
```

Imagem 7: Classe Medicine, ela armazena todos os 5 dados que serão necessários.

DBConnect é responsável por fazer a conexão com o banco de dados e solicitar que ele faça uma busca. O método init_db() é responsável por criar a tabela no banco, quando for solicitado por Logic e Manager que isso seja feito.

As funções open_connection(), close_connection() mudarão o estado do atributo con, que representa a conexão do banco, para aberto ou fechado. A função commit() tem o objetivo de fazer o commit das alterações no banco.

Por último, os métodos de seleção são chamados por Logic, dependente de qual tarefa foi solicitado por Manager, e receberão uma string nome ou ano ou um parâmetro de seleção.

```
Horse Monnect

Class Monnect

Exact consists on a passes a consist do bases

on splits (consection

Exact consists on a passes

of consection

Exact consists on a passes

of consection(consection)

exact consists a strength on a passes

of consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(consection(co
```

Imagem 8: Funções de seleção de DBConnect

O módulo CSVReader é responsável por ler o arquivo csv e salvar as informações dele em memória. Ele é chamado por Logic apenas que Manager detecta que o banco de dados não existe ainda no sistema, para que todo o conteúdo do arquivo csv seja copiado para o banco de dados.

```
Classe CSVReader

Class (CSVReader):

csvfile: fileType
absolute_name: str
data = []

#inicializa a classe com o caminho absoluto do arquivo csv.
def __init__(self, file_name):

self.absolute_name = os.getcud() + '\\data\\' + file_name

#abbre a arquivo csv para leitura.

def open_file(self):

self.csvfile = open(self.absolute_name, newline='')

#fecha o arquivo csv para leitura.

#compared to the self.csvfile open(self.absolute_name, newline='')

#fecha o arquivo csv
def close_file(self):

self.csvfile.close()

#carrega em data uma lista de todas as linhas, onde cada linha é um dicionário com a chave sendo a informação do header do arquivo e o valor sendo o valor de cada columa da informação.
def load_in_memory(self):

self.open_file()

reader = csv.DictReader(self.csvfile, delimiter=";")

for row in reader:
self.data.append(row)
self.close_file()
```

Imagem 9: A classe CSVReader e seus atributos e métodos

Testes

Para os testes dos módulos CSVReader e DBConnect, foram feitos testes de unidades utilizando a biblioteca unittest nativa do python. Além disso, foram feitos testes de integração do sistema inteira ou parte dele rodando o programa e vendo as saídas e erros que ele retornava.

```
Classes responsáveis pelos testes de unidade de CSVReader
class TestReaderMethods(unittest.TestCase):
   csv_reader = rd.CSVReader('TA_PRECO_MEDICAMENTO.csv')
   def test_absolute_name(self):
       self.assertEqual(self.csv_reader.absolute_name, answers.TEST_1_READER_ANSWER)
   def test_load_in_memory_first_element(self):
       index = 0
       attribute = 'CNPJ'
       self.csv_reader.load_in_memory()
       self.assertEqual(self.csv_reader.data[index][attribute], answers.TEST_2_READER_ANSWER)
   def test_load_in_memory_last_element(self):
       index = -1
       attribute = 'CNPJ'
       self.csv_reader.load_in_memory()
       self.assertEqual(self.csv_reader.data[index][attribute], answers.TEST_3_READER_ANSWER)
   def test_load_in_memory_intermediary_element(self):
       index = 10
       attribute = 'CNPJ'
       self.csv_reader.load_in_memory()
       self.assertEqual(self.csv_reader.data[index][attribute], answers.TEST_4_READER_ANSWER)
   def test_attribute_1(self):
       index = 10
       attribute = "SUBSTÂNCIA"
       self.csv_reader.load_in_memory()
       self.assertEqual(self.csv_reader.data[index][attribute], answers.TEST_5_READER_ANSWER)
   def test_attribute_2(self):
       index = 10
       attribute = "PRODUTO"
       self.csv_reader.load_in_memory()
       self.assertEqual(self.csv_reader.data[index][attribute], answers.TEST_6_READER_ANSWER)
```

Imagem 10: Testes de unidade da classe CSVReader

Projeto Final

Por fim, temos o programa inteiro dentro da pasta program, que é o pacote principal. Dentro dele estão os arquivos manager.py que implementa a classe Manager, logic.py que implementa a classe Logic, além do arquivo que implementa os testes de unidade (teste.py) e das respostas para cada teste (tests_answer.py). Dentro do pacote program, temos mais dois pacotes, data que abriga os arquivos que implementam as classes CSVReader (arquivo csvreader.py), DBConnect (database.py) e Medicine (medicine.py).

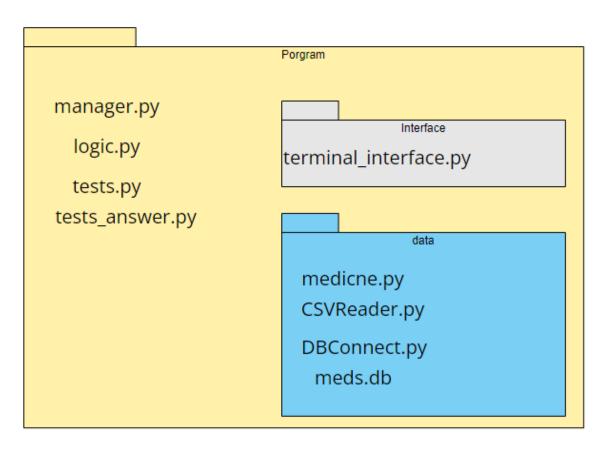


Imagem 11: Diagrama de pacotes do projeto

Resultados

Para a primeira tarefa, podemos ver que o programa encontrou os medicamentos Montelucaste e Montelucaste de sódio ao ser inserido o nome "monte".

Imagem 12: Testes da primeira tarefa

Para o segundo teste, testamos o código 7891317421618, o único código que encontrei no csv que pertence a dois remédios. Podemos ver que ele encontrou os dois remédios e os imprimiu, com a diferença de valores que é zero, já que os dois tem o mesmo valor.

Imagem 13: Testes da segunda tarefa

Para a terceira tarefa, os resultados foram 30,2% de avaliações negativas, 69,3% de avaliações positivas e 0,3% de neutras. Podemos ver que nem um asterisco foi impresso para neutras pois ela é menor que 1%.

```
Selecione uma das opções
1 - Consultar medicamento pelo nome
2 - Buscar pelo código de barras
3 - Comparativo PIS/COFINS
  - Terminar o programa
Insira aqui: 3
CLASSIFICACAO PERCENTUAL TG
Negativa 30.26605853287723
                        TGRAFICO
                              ***********
            0.33447358418852147
Positiva
           69.39946788293425
Selecione uma das opções
1 - Consultar medicamento pelo nome2 - Buscar pelo código de barras3 - Comparativo PIS/COFINS
  - Terminar o programa
Insira aqui: _
```

Imagem 14: Resultados encontrados pelo programa

Conclusões e Autoavaliação

Portanto, o programa parece estar bem robusto e funcional. Aparentemente ele consegue resolver todas as 3 tarefas e não foi encontrado nenhum bug na execução dos testes.

Fiquei bem satisfeito com meu resultado, considerando que sou novato com SQLite e testes de unidade para python, consegui ganhar bastante conhecimento nestas duas áreas. Consegui aplicar vários conceitos que aprendi e cadeiras relacionadas a engenharia de software como modularidade e arquitetura em camadas, que é uma área que tenho bastante interesse.

Por fim, gostaria só de ter implementado uma interface gráfica para fazer a comunicação com o usuário, mas infelizmente não deu tempo.