Realmente solo son dos archivos:

- create_db.py
- data_puller.py

Primero, esta tabla:

```python
class Process(Base):
    __tablename__ = 'processes'

    id = Column(Integer, primary_key=True)
    snapshot_id = Column(Integer, nullable=False)
    pid = Column(Integer, nullable=False)
    name = Column(String, nullable=False)
    username = Column(String, nullable=False)
    exe = Column(String, nullable=True)
    create_time = Column(String, nullable=False)
    cmdline = Column(String, nullable=True)
    ppid = Column(Integer, nullable=False)
    cwd = Column(String, nullable=True)
    net_connections = Column(String, nullable=True)
    status = Column(String, nullable=False)
    memory_info = Column(String, nullable=True)
    hashfile = Column(String, nullable=True)

    def __repr__(self):
        return f"<Process(id={self.id}, pid={self.pid}, name='{self.name}', username='{self.username}')>"
```

Solo obtiene datos con una libreria de procesos, yo lo veo como si estuvieras sacando datos en linux con el comando “ps aux”. Si vamos a analizar bien los procesos del software, obviamente esto no es suficiente, pero vale la pena tener todo esto en una tabla.

```python
def add_process_to_table(snap_id):
    for proc in psutil.process_iter(['pid', 'name', 'username', 'exe', 'create_time', 'cmdline', 'ppid', 'cwd', 'net_connections', 'status', 'mem
        #print(proc.info)
        hashfile = get_file_hash(proc.info['exe']) if proc.info['exe'] else None
        proc.info['hashfile'] = hashfile
        temp = Process(
            snapshot_id=snap_id,
            pid=proc.info['pid'],
            name=proc.info['name'],
            username=proc.info['username'],
            exe=proc.info['exe'],
            create_time=str(proc.info['create_time']),
            cmdline=' '.join(proc.info['cmdline']) if proc.info['cmdline'] else None,
            ppid=proc.info['ppid'],
            cwd=proc.info['cwd'],
            net_connections=str(proc.info['net_connections']),
            status=proc.info['status'],
            memory_info=str(proc.info['memory_info']),
            hashfile=hashfile,
        )
        add_process(temp)

    #display_processes()
```

Para hacer un análisis más profundo nos vamos a “/proc” en el ambiente linux y de ahí las tablas de Process_FDINFO, Sockets, Mounts, Modules, ProcDetails, ProcNetInfo, Maps.

Basicamente todo eso lo ando procesando en la funcion: add_proc_info