Mynd Data Challenge 2023

In this directory you will find a database called `mynd.db` containing 3 tables:

PROPERTIES: data of properties we manage.
- `property_id`
- `market`
- `bedrooms`
- `bathrooms`
- `square_footage`
- `year_built`
- `onboard_date` (management started)
- `offboard_date` (churn date)

TENANTS: contains tenancy data
- `tenant_id`
- `property_id`
- `move_in_date`
- `move_out_date`

CONTRACTS: lease agreements with a start date and a stipulated end date. At the end of a contract, a tenant can create a new one with a new id ("renew") or move out.
- `contract_id`
- `property_id`
- `tenant_id`
- `start_date`
- `end_date`
- `rent`

You can access it in Python like so:

```python
import sqlite3
import pandas as pd

def run_query(query):
    conn = sqlite3.connect('mynd.db')
    cur = conn.cursor()
    df = pd.read_sql_query(query, conn)
    cur.close()
    return df

run_query('''SELECT *
             FROM properties''')
```

Questions

1. **Plot monthly time series for:**
   a. **Occupancy rate**: occupied / total units managed
   b. **Renewal rate**: % of contracts expiring each month that are renewed

2. **What effect does vacancy duration have on investor retention?**

   Assume that concluded vacancies have no influence on future churn risk.

3. We have a group of potential investors, each with a different probability p of converting into paying customers. We can spend $100 in a customer representative call and make $p = 25\%$, or we can send them a personalized email at a cost of $20, which makes $p = 9\%$. What is your recommendation for the team?

   Hint: Try to think how your answer changes with different values of p and customer LTV.

Notes

You can assume the following:
- The data is clean and has no errors
- All properties are onboarded vacant
- 1 property = 1 investor (Offboarding a property = Investor churning)
- If a contract is "renewed", a new contract is created with a start date equal to the end date of the prior contract.