SEATTLE RENTAL PROPERTY STATISTICS

Group AA1: Irene Hu, Joseph Tran,

Rebecca Jessup, Ty Okazaki

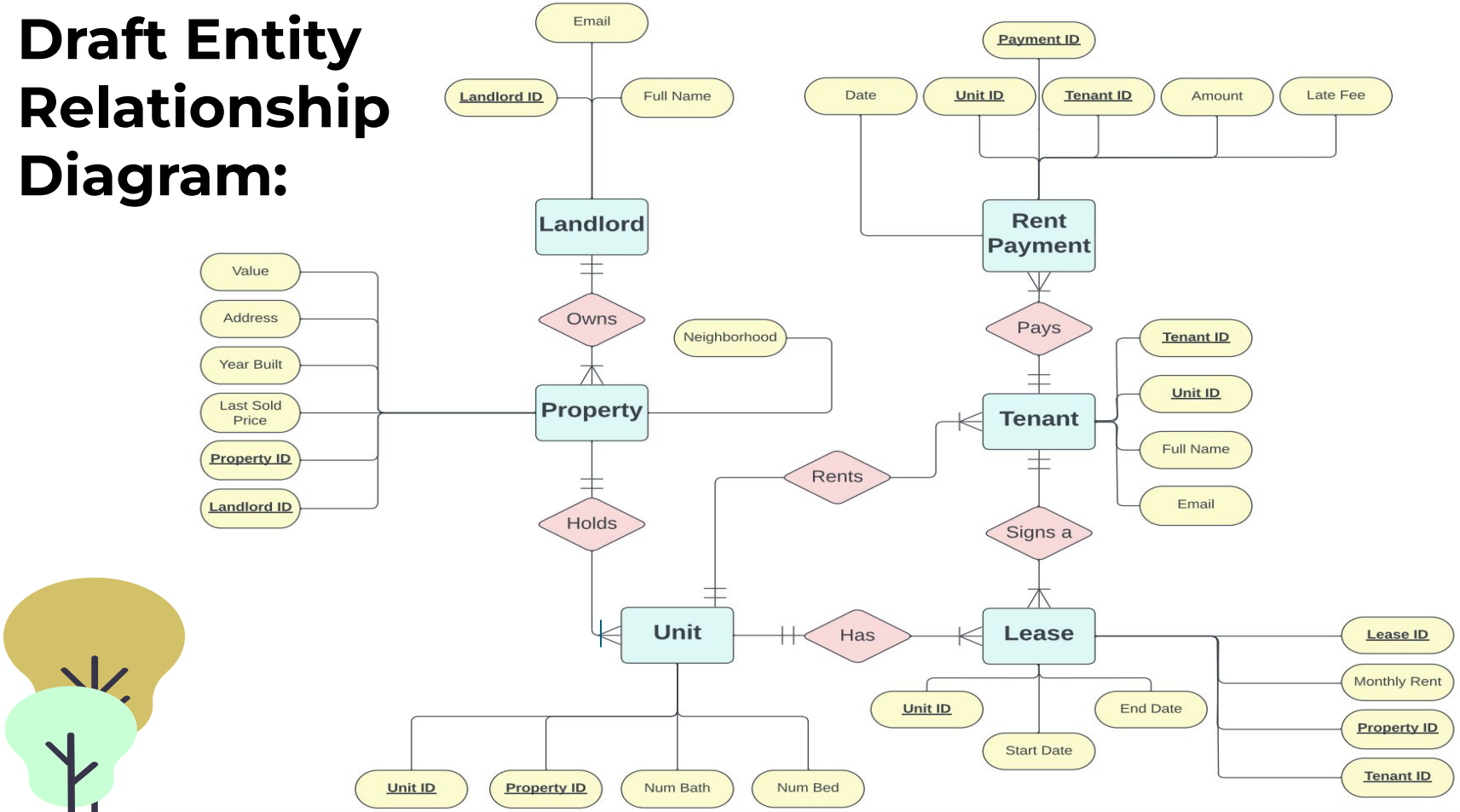
BUSINESS PROBLEM:

For our project, we are interested in focusing on the **rental property industry in Seattle** to understand **market trends** regarding **property values and rental rates.** From our data, we can investigate how properties differ in pricing across **neighborhoods,** which properties have the most **units** available, which **landlords** own the most properties, and other questions as well. By leveraging this information and database into our project, renters and real estate professionals can make informed decisions that can lead to more successful transactions and better returns on investments.

PROJECT STAKEHOLDERS:

- Potential tenants: which areas are most affordable? How do similar properties in different areas compare?
- Landlords: which tenants have good rental history? Which landlords own the most properties?
- Real Estate Professionals: what types of rental properties are the most common? How are property values changing?
- Developers: which areas have the most expensive rental properties? Which areas are properties the oldest?

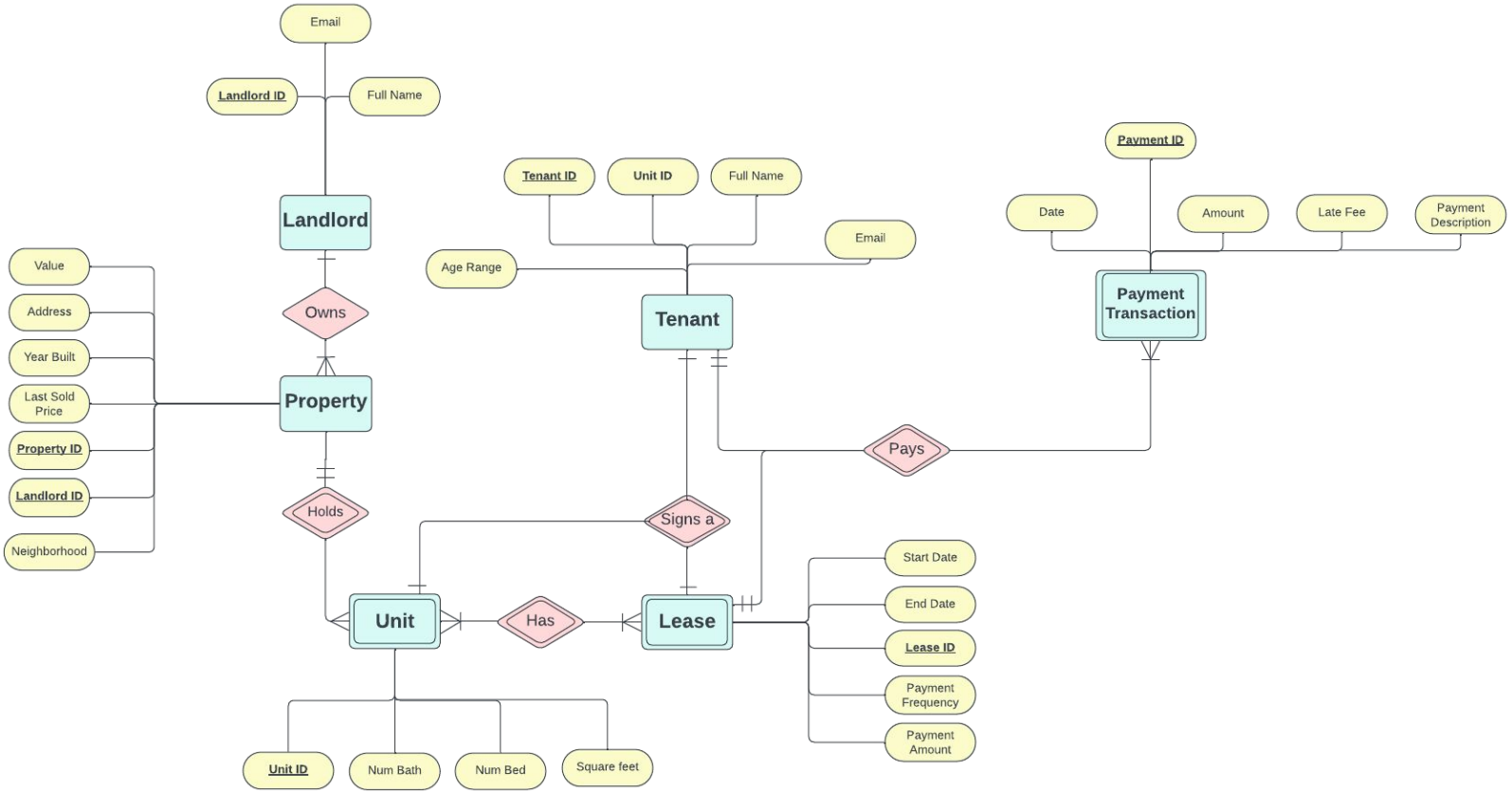
Draft Entity Relationship Diagram:



THOUGHT PROCESS:

- Represent the key entities in a rental property (tenant, landlord, property, unit, lease, rent)
- Need to include IDs as foreign key attributes to connect entities, for example landlord ID as property attribute
- Different types of properties (homes and apartments) could have one or multiple units
- Each property has only one landlord but one landlord can own multiple properties
- The term “rent” could be used as both an entity and a relationship

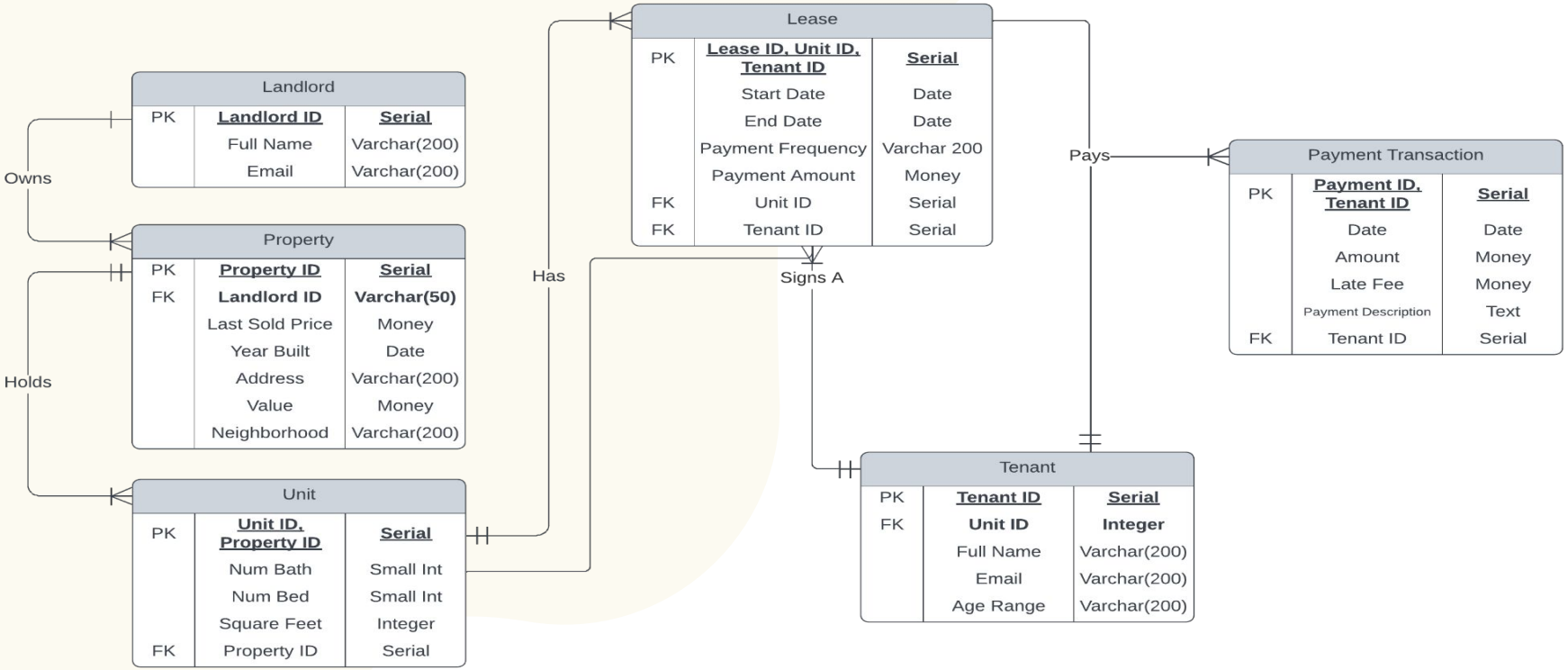
Final Entity Relationship Diagram:



CHANGES WE MADE:

- We removed duplicate entities that weren't necessary and implemented weak entity sets
- Updated cardinalities for the different entities
- Ternary relationships: "signs_a" and "pays" to better connect entities
- Revised "monthly rent" attribute in lease entity to payment frequency and amount

TABLE DESIGN:



DATABASE IMPLEMENTATION

```sql
-- CREATE TABLES
CREATE TABLE Lease (
    lease_id SERIAL,
    unit_id SERIAL,
    tenant_id SERIAL,
    start_date DATE,
    end_date DATE,
    payment_frequency VARCHAR(200),
    payment_amount MONEY,
    PRIMARY KEY (lease_id, unit_id, tenant_id)
);


CREATE TABLE PaymentTransaction (
    payment_id SERIAL,
    tenant_id SERIAL,
    date DATE,
    amount MONEY,
    late_fee MONEY,
    payment_desc TEXT,
    PRIMARY KEY (payment_id, tenant_id)
);


CREATE TABLE Tenant (
    tenant_id SERIAL,
    unit_id INT,
    full_name VARCHAR(200),
    email VARCHAR(200),
    age_range VARCHAR(200),
    PRIMARY KEY (tenant_id)
);
```

```sql
-- INSERT INTO Data for Lease
INSERT INTO Lease (unit_id, tenant_id, start_date, end_date, payment_frequency, payment_amount)
VALUES
    (1, 1, '2022-01-01', '2022-12-31', 'Monthly', 1000),
    (2, 2, '2022-01-01', '2022-12-31', 'Monthly', 1000),
    (3, 3, '2022-02-01', '2023-01-31', 'Monthly', 1200),
    (4, 4, '2022-02-01', '2023-01-31', 'Monthly', 1200),
    (5, 5, '2022-03-01', '2023-02-28', 'Monthly', 1500),
    (6, 6, '2022-03-01', '2023-02-28', 'Monthly', 1500),
    (7, 7, '2022-01-01', '2022-06-30', 'Monthly', 4500),
    (8, 8, '2022-01-01', '2022-03-31', 'Monthly', 3500),
    (9, 9, '2022-02-01', '2024-01-31', 'Monthly', 2000),
    (10, 10, '2022-02-01', '2023-01-31', 'Monthly', 3200),
    (11, 11, '2022-03-01', '2023-02-28', 'Monthly', 2100),
    (12, 12, '2022-03-01', '2023-02-28', 'Monthly', 1650),
    (13, 13, '2022-04-01', '2023-03-31', 'Monthly', 1800),
    (14, 14, '2022-05-01', '2023-04-30', 'Monthly', 1900),
    (15, 15, '2022-06-01', '2023-05-31', 'Monthly', 2000),
    (16, 16, '2022-07-01', '2023-06-30', 'Monthly', 2200),
    (17, 17, '2022-08-01', '2023-07-31', 'Monthly', 2300),
    (18, 18, '2022-09-01', '2023-08-31', 'Monthly', 2400),
    (19, 19, '2022-10-01', '2023-09-30', 'Monthly', 2500),
    (20, 20, '2022-11-01', '2023-10-31', 'Monthly', 2600),
    (21, 21, '2022-12-01', '2023-11-30', 'Monthly', 2700),
    (22, 22, '2023-01-01', '2023-12-31', 'Monthly', 2800),
    (23, 23, '2023-02-01', '2024-01-31', 'Monthly', 2900),
    (24, 24, '2023-03-01', '2024-02-29', 'Monthly', 3000);
```

```sql
-- INSERT INTO Data for Property
INSERT INTO Property (landlord_id, last_sold_price, year_built, address, current_value, neighborhood)
VALUES
    (1, 200000, '2016-01-01', '1023 East Alder St', 250000, 'Capitol Hill'),
    (2, 300000, '2021-01-01', '4722 Fauntleroy Way SW', 350000, 'West Seattle'),
    (3, 400000, '2010-01-01', '526 Yale Ave N', 500000, 'Capitol Hill'),
    (1, 200000, '2020-01-01', '400 Queen Anne Ave N', 1250000, 'Queen Anne'),
    (1, 300000, '2010-01-01', '701 5th Ave N', 1500000, 'Queen Anne'),
    (3, 500000, '1989-01-01', '210 8th Ave N, Seattle', 750000, 'South Lake Union'),
    (4, 849000, '1999-03-06', '2718 31st Ave S', 1250000, 'Mount Baker'),
    (5, 250000, '2015-01-01', '1205 E Pike St', 350000, 'Capitol Hill'),
    (2, 350000, '2012-01-01', '6501 5th Ave NE', 450000, 'Green Lake'),
    (4, 500000, '1990-01-01', '4400 4th Ave SW', 650000, 'West Seattle'),
    (3, 200000, '2005-01-01', '1000 4th Ave', 800000, 'Downtown Seattle'),
    (6, 450000, '2018-01-01', '1618 15th Ave', 550000, 'Central District'),
    (2, 357000, '2009-08-24', '7346 15th Ave NW', 925000, 'U-District'),
    (2, 900000, '1968-02-21', '6516 47th Ave NE', 975000, 'Ballard'),
    (5, 389900, '1992-10-07', '2225 Eastlake Ave E', 850000, 'Capitol Hill'),
    (5, 292000, '1964-10-03', '3042 34th Ave SW', 799900, 'West Seattle'),
    (6, 520000, '1973-06-24', '716 NW 60th St', 720000, 'Wedgwood'),
    (2, 340000, '1938-03-17', '1015 W Fulton St', 889000, 'Ballard'),
    (2, 4322000, '2020-08-04', '3104 Western Ave', 4400000, 'Downtown Seattle'),
    (4, 600000, '1984-03-21', '3429 33rd Ave W', 1110000, 'Ballard');
```

SAMPLE QUESTIONS + QUERIES:

-- How many units are currently available to rent for each property?

```
SELECT Property.property_id, Property.address, COUNT(Unit.unit_id) -
COUNT(Lease.unit_id) AS vacant_units
FROM Property
JOIN Unit
    ON Property.property_id = Unit.property_id
LEFT JOIN Lease
    ON Unit.unit_id = Lease.unit_id
GROUP BY Property.property_id
ORDER BY vacant_units DESC
LIMIT 10;
```

-- What is average monthly rent for properties in each neighborhood?

```
SELECT neighborhood, ROUND(AVG(payment_amount::NUMERIC), 2) AS
avg_monthly_rent
FROM Property
JOIN Unit
    ON Property.property_id = Unit.property_id
JOIN Lease
    ON Unit.unit_id = Lease.unit_id
GROUP BY neighborhood
LIMIT 10;
```

-- What are the names and email addresses of tenants that have had to pay late fees?

```
SELECT full_name, email
FROM Tenant
JOIN PaymentTransaction
   ON Tenant.tenant_id = PaymentTransaction.tenant_id
WHERE late_fee > 0::money
GROUP BY Tenant.tenant_id;
```

-- Which properties have the highest property value ordered by year built?

```
SELECT property_id, address, current_value, year_built
FROM Property
ORDER BY current_value DESC, year_built ASC;
```

-- Which landlords own the most properties?

```
SELECT l.full_name, COUNT(*) AS property_count
FROM Landlord l
JOIN Property p
   ON l.landlord_id = p.landlord_id
GROUP BY l.full_name
ORDER BY property_count DESC;
```

DEMO QUERY RESULTS:

```sql
-- How many units are currently available to rent for each property? (Real estate agent, Landlord, Tenant)
SELECT Property.property_id, Property.address, COUNT(Unit.unit_id) - COUNT(Lease.unit_id) AS vacant_units
FROM Property
JOIN Unit
    ON Property.property_id = Unit.property_id
LEFT JOIN Lease
    ON Unit.unit_id = Lease.unit_id
GROUP BY Property.property_id
ORDER BY vacant_units DESC
LIMIT 10;


-- Output
-- property_id      address              vacant_units
-- 12           "1618 15th Ave"              2
-- 19           "3104 Western Ave"           1
-- 13           "7346 15th Ave NW"           1
-- 17           "716 NW 60th St"             1
-- 14           "6516 47th Ave NE"           1
-- 16           "3042 34th Ave SW"           1
-- 18           "1015 W Fulton St"           1
-- 15           "2225 Eastlake Ave E"        1
-- 1            "1023 East Alder St"         0
-- 5            "701 5th Ave N"              0
```

```sql
-- Which neighborhood has the highest number of tenants?
SELECT neighborhood, COUNT(*) AS num_tenants
FROM Unit
JOIN Property
    ON Unit.property_id = Property.property_id
GROUP BY neighborhood
ORDER BY num_tenants DESC
LIMIT 10;

-- Output
-- neighborhood     num_tenants
-- "Capitol Hill"       7
-- "West Seattle"       5
-- "Queen Anne"         4
-- "Central District"   3
-- "South Lake Union"   3
-- "Downtown Seattle"   3
-- "Ballard"            2
-- "Mount Baker"        2
-- "Green Lake"         2
-- "Wedgwood"           1
```

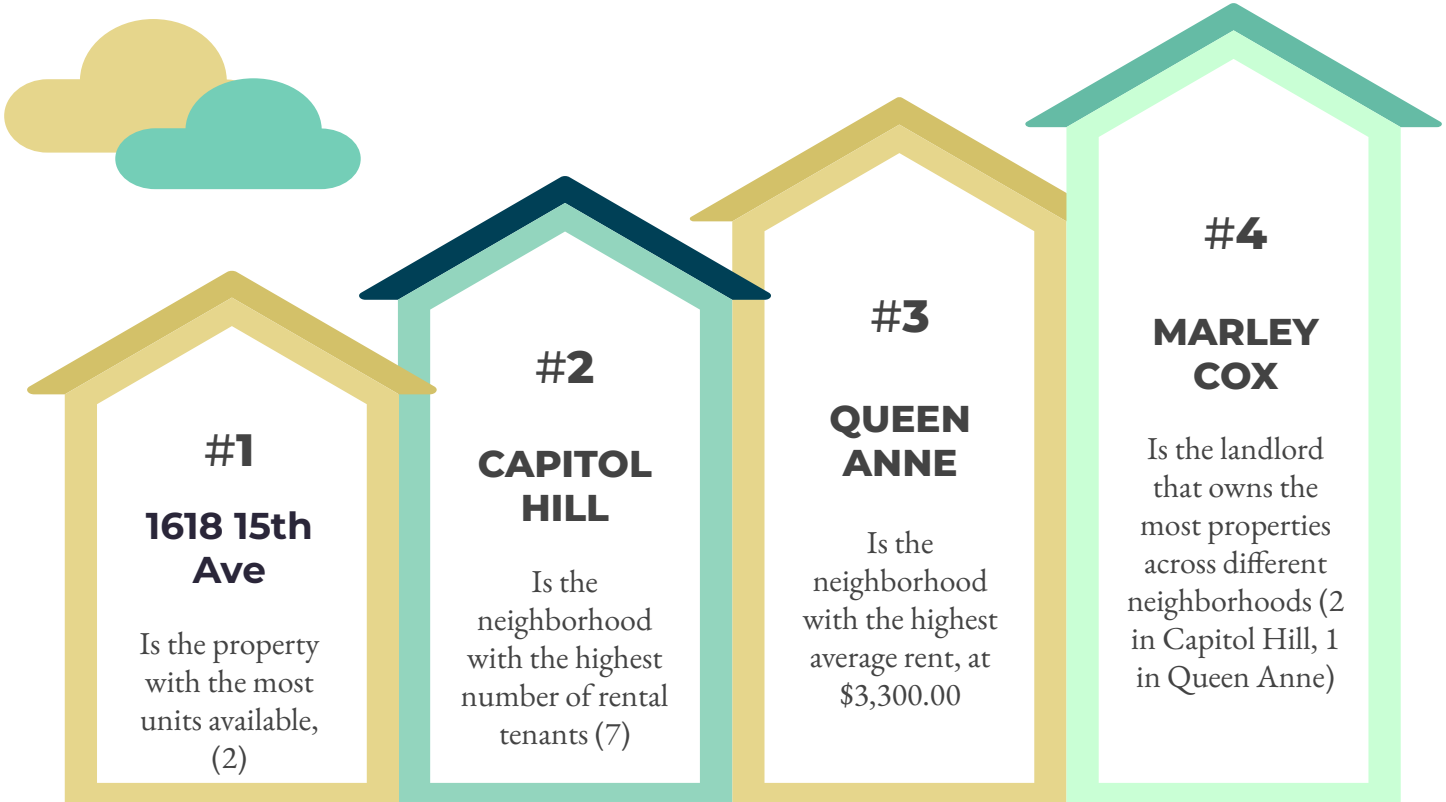
```sql
-- What is the average monthly rent for properties in each neighborhood?
SELECT neighborhood, ROUND(AVG(payment_amount::NUMERIC), 2) AS avg_monthly_rent
FROM Property
JOIN Unit
    ON Property.property_id = Unit.property_id
JOIN Lease
    ON Unit.unit_id = Lease.unit_id
GROUP BY neighborhood
LIMIT 10;

-- Output
-- neighborhood     avg_monthly_rent
-- "Green Lake"         2450.00
-- "Downtown Seattle"   2850.00
-- "Mount Baker"        1950.00
-- "Central District"   3000.00
-- "Queen Anne"         3300.00
-- "South Lake Union"   1850.00
-- "Capitol Hill"       1583.33
-- "West Seattle"       1925.00
```

```sql
-- How many properties are owned by the same landlord in a certain neighborhood?
SELECT p.neighborhood, l.full_name AS owner, COUNT(*) as num_property
FROM Property p
JOIN Landlord l ON p.landlord_id = l.landlord_id
GROUP BY p.neighborhood, owner
ORDER BY num_property DESC
LIMIT 10;

-- Output
-- neighborhood             owner           num_property
-- "Queen Anne"         "Marley Cox"            2
-- "Capitol Hill"       "Adam Patel"            2
-- "Ballard"            "Grayson Simpson"       2
-- "West Seattle"       "Grayson Simpson"       1
-- "Capitol Hill"       "Robert Davis"          1
-- "South Lake Union"   "Robert Davis"          1
-- "Ballard"            "Emily Wilson"          1
-- "Wedgwood"           "Sarah Kim"             1
-- "West Seattle"       "Adam Patel"            1
-- "Capitol Hill"       "Marley Cox"            1
```

PROJECT FINDINGS:



#1

1618 15th Ave

Is the property with the most units available, (2)

#2

CAPITOL HILL

Is the neighborhood with the highest number of rental tenants (7)

#3

QUEEN ANNE

Is the neighborhood with the highest average rent, at $3,300.00

#4

MARLEY COX

Is the landlord that owns the most properties across different neighborhoods (2 in Capitol Hill, 1 in Queen Anne)

Names and Email Address of those with Late Fees

	Full Name	Email Address
1	Alice Lee	mia.anderson@gmail.com
2	Emily Chen	emily.chen@gmail.com
3	Jane Doe	jane.doe@gmail.com
4	Mia Anderson	mia.anderson@gmail.com
5	Olivia Nguyen	olivia.nguyen@gmail.com
6	Alex Kim	alex.kim@gmail.com
7	Emily Wright	emily.wright@gmail.com
8	Bob Johnson	bob.johnson@gmail.com
9	Ava Taylor	ava.taylor@gmail.com
10	John Smith	john.smith@gmail.com

Discussion & Summary

-- What was interesting?

The process of refining the ERD took lots thoughtfulness and time, but was critical in helping us create our database.. It was interesting to go through the entire process of creating a database and working with the data, we now have a better understanding and appreciation of working with complex databases and the work that is done create them.

-- What did we learn?

We learned the complexities associated with sketching and honing the relations within a database - one that models systems present in reality. This project also helped us conceptualize how complex databases are, especially ones that are dealing with large amount of data, like the rental housing market.