When I was looking over the columns of data I was able to see different correlations between some of them from the get go which kind of made it easier to create ERD and to normalize the data into 3nf. I find it easier to be able to look at the data in excel and the descriptions that the website gives of the columns helpful to create my ERD's in my head and on paper rather than on the computer since I think too much and too fast to be able to spew them out by typing and trying to find certain things.

4 Questions

The questions that I thought of were the very basic ones that could be used in other types data manipulation for other data sets.

1. Which patron type has the lowest/highest checkout activity?

```
select patron_type, count(*) as checkout_count
from checkout_activities
group by patron_type
order by checkout_count asc
limit 1;

select patron_type, count(*) as checkout_count
from checkout_activities
group by patron_type
order by checkout_count desc
limit 1;
```

Understanding which patron types are most and least active helps in tailoring services and resources accordingly. It informs decisions on membership types and program offerings.

2. Is there change between different years on number of checkouts/renwals?

```
select extract(year from checkout_date) as year, count(*) as num_checkouts
from checkout_activities
group by year
order by year;

select extract(year from renewal_date) as year, count(*) as num_renewals
from checkout_activities
where renewal_date is not null
group by year
order by year;
```

Monitoring trends in library usage over years provides insights into the library's relevance and popularity over time, aiding in resource allocation and strategic planning.

3. What are the most popular months and years for library usage?

```
select extract(month from checkout_date) as month, count(*) as num_checkouts
from checkout_activities
group by month
order by num_checkouts desc
limit 1;

select extract(year from checkout_date) as year, count(*) as num_checkouts
from checkout_activities
group by year
order by num_checkouts desc
limit 1;
```

Identifying peak periods informs staffing and resource management, making sure that when ruch hour comes there is enough staff available.

4. Is there a relationship between patron age and their library use?

```
select round(avg(ca.checkout_count)) as avg_checkouts, p.age_group
from patrons p
join (
select patron_id, count(*) as checkout_count
from checkout_activities
group by patron_id
) ca on p.patron_id = ca.patron_id
group by p.age_group;
```

Analyzing age helps in understanding usage patterns among different age groups which help find the right target audience and who to tailor programs for as well.