Why RA?

- Formalism for describing queries
- Basis of relational databases
- Will make you a SQL wizard!
Notes on RA

- Multiple possible query plans
- Logical vs. Physical query plans

Query optimization = finding cheaper, equivalent expressions

A physical query plan is a logical query plan annotated with physical implementation details
Example: RA-to-SQL

Person(id, name, countryid)
Country(id, name, continent)

```
SELECT C.id
FROM Person P, Country C
WHERE P.countryid = C.id
AND C.continent = 'Africa'
GROUP BY C.id
HAVING COUNT(*) > 10000000

Can we make a more efficient plan?
```

Equivalently in equation form: $\Pi_{C.id}(\sigma_{\text{pop} > 10,000,000}(\gamma_{C.id,COUNT(*) \rightarrow \text{pop}}(\sigma_{\text{continent} = \text{Africa}}(\bowtie_{P.countryid = C.id}(\text{Person} P \bowtie_{P.countryid = C.id} \text{Country} C))))))$
RA Reference Sheet

<table>
<thead>
<tr>
<th>Name</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection</td>
<td>$\sigma$</td>
</tr>
<tr>
<td>Projection</td>
<td>$\pi$</td>
</tr>
<tr>
<td>Join</td>
<td>$\bowtie$</td>
</tr>
<tr>
<td>Group By</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>Set Difference</td>
<td>$-$</td>
</tr>
<tr>
<td>Duplicate Elimination</td>
<td>$\delta$</td>
</tr>
</tbody>
</table>