SQL Queries
Single Table

SQL Workflow

- CREATE TABLE
- INSERT TUPLES
  - Bulk load: .import
- **Queries**
  - **Data processing**
  - **Data analysis**
  - **Data science**

- PANDAS
  - Create panda object
  - Read CSV file
  - Call functions

SQL Queries

SELECT result_table_schema

FROM input_tables

[WHERE table_predicates AND join_conditions]

[GROUP BY grouping_attributes]

[ORDER BY sorting_attributes]

SQL Queries – Single Table

SELECT result_table_schema

FROM table

[WHERE table_predicates AND join_conditions]

Data from Table

- SQL
  - SELECT *

    FROM Cities_Population
  - * corresponds to the complete schema of the input table
- PANDAS
  - city_pop.head()

Column(s) from Table

- SQL
  - SELECT city

    FROM Cities_Population
  - SELECT city, county

    FROM Cities_Population
- PANDAS
  - city_pop["City"]

Rename Columns in Result

```
SELECT
    city,
    county,
    incorporated AS established,
    pop_2010 AS
    current_population
FROM
    Cities_Population
```

```
SELECT
    city,
    pop_2010 – pop_2000
    AS population_increase
FROM
    Cities_Population
```

No Index Access in SQL

- SQL
  - Only value based access
- PANDAS
  - city_pop["City"][20]
  - city_pop.iloc[20]
  - city_pop.iloc[20][1]
  - city_pop.loc[:10, ['City','County']]

Conditions or Predicates

- SQL
  - SELECT

    *

    FROM

    Cities_Population

    WHERE

    county = ‘Merced’
- PANDAS
  - city_pop.loc[city_pop.County == 'Merced']

Complex Predicates

```
SELECT city, pop_2000, pop_2010
FROM
        Cities_Population
WHERE
        (county = 'Merced' OR county = 'Stanislaus') AND
        pop_2010 > pop_2000
```

Predicates on Strings

- SELECT city
  FROM
  Cities_Population
  WHERE
  city LIKE ‘San %’
- SELECT city
  FROM
  Cities_Population
  WHERE
  city LIKE ‘San%’
- SELECT city
  FROM
  Cities_Population
  WHERE
  city LIKE ‘%San__ %’
- SELECT city
  FROM
  Cities_Population
  WHERE
  city LIKE ‘%San__%’

Check NULL Attributes

```
SELECT
    city,
    incorporated,
    pop_1980,
    pop_1990
FROM Cities_Population
WHERE
    county = 'Los Angeles' AND
    pop_1980 is null
```

```
SELECT city,
    case pop_1980 is null
        when true then pop_1990
        else pop_1990 - pop_1980
    end as change_1980_1990
FROM Cities_Population
WHERE county = 'Los Angeles'
```

ORDER BY Result

- SELECT city, pop_2010

  FROM Cities_Population

  ORDER BY

  pop_2010 [DESC]

- select county, city

  from Cities_Population

  order by county, city

SELECT

city,

pop_2010 - pop_2000 as change_2000_2010

FROM Cities_Population

ORDER BY

change_2000_2010 [desc]

Exercise 6.1.3

- Check the file in the lecture materials for all SQL statements
- Run all the queries on the sample database created and populated in the previous lectures
- f)

```
select model, hd
from pc
where speed = 3.2 and price < 2000
```

Examples

- California_Cities
- Computers
- TPCH