

SQL Workshop Day 1

UC Davis DSI

2018

1 Selecting and Sorting

1. How many tables are in the database?

Solution: 7. Look under 'Database Structure' tab.

2. Select all fields in the 'company' table.

Solution:

```
SELECT *  
FROM company  
;
```

3. Which table has the most rows? The fewest rows?

Solution:

Count rows using:

```
SELECT COUNT(*)  
FROM daily_share_prices -- replace with name of each table  
;
```

4. Write a query so that the fields in `financial_ratios` table appear in alphabetical order.

Solution:

```
SELECT beta  
    , dividend_yield  
    , market_cap_in_millions  
    , price_earnings_ratio  
    , price_to_book_ratio  
    , return_on_equity  
    , share_price  
    , ticker  
FROM financial_ratios  
;
```

5. Which company has the least expensive stock in the `financial_ratios` table?

Solution:

```
SELECT ticker, share_price
FROM financial_ratios
ORDER BY share_price
LIMIT 1
;
```

6. Which company has the most expensive stock in the `financial_ratios` table?

Solution:

```
SELECT ticker, share_price
FROM financial_ratios
ORDER BY share_price DESC
LIMIT 1
;
```

7. How many different stocks do we have daily share prices for?

Solution:

```
SELECT COUNT(DISTINCT ticker)
FROM daily_share_prices
;
```

8. How many different days do we have daily share prices for?

Solution:

```
SELECT COUNT(DISTINCT date)
FROM daily_share_prices
;
```

9. Which 5 states have the highest population in 2017? lowest?

Solution:

```
SELECT *
FROM state_populations
ORDER BY popnEstimate2017 -- DESC for lowest
LIMIT 5
;
```

10. How many different 4 digit industry sector codes are possible?

Solution:

```
SELECT COUNT(DISTINCT sic)
FROM sic
;
```

11. How many different industry sectors appear in the company information?

Solution:

```
SELECT COUNT(DISTINCT sic_code)
FROM company_info
;
```

12. What are all the different prices that each company closed at?

Solution:

```
SELECT DISTINCT ticker, close
FROM daily_share_prices
;
```

13. Which states (in this database) have companies in them?

Solution:

```
SELECT DISTINCT state
FROM company_locations
```

14. How many companies are in each of the different tables? What can you conclude?

Solution:

```
SELECT COUNT(DISTINCT ticker)
FROM company_info -- Similar for others
;
```

company_info: 450, company_locations: 500, company_name: 505, daily_share_prices: 505. Different numbers mean that not all companies are in all of the tables.

2 Filtering and Calculations

1. Which company is located in Oregon?

Solution:

```
SELECT *  
FROM company_locations  
WHERE state = 'Oregon'  
;
```

2. Which companies have a price-earnings ratio below 7.0?

Solution:

```
SELECT ticker, share_price  
FROM financial_ratios  
WHERE price_earnings_ratio < 7  
;
```

3. Which companies have an asset turnover (revenue/assets) greater than 5.0?

Solution:

```
SELECT ticker, (revenue / asset) AS asset_turnover  
FROM company_info  
WHERE asset_turnover > 5.0  
;
```

4. Which states had a 2010 population of at least 10,000,000?

Solution:

```
SELECT *  
FROM state_populations  
WHERE popn2010 >= 10000000  
;
```

5. How many days did Apple's (AAPL) share price close above 150?

Solution:

```
SELECT COUNT(*)  
FROM daily_share_prices  
WHERE ticker == 'AAPL'  
AND close > 150  
;
```

6. Which New York companies are not located in New York city?

Solution:

```
SELECT *
FROM company_locations
WHERE state = 'New York'
AND city != 'New York'
;
```

7. Which California companies are located in Irvine, San Jose, or Thousand Oaks?

Solution:

```
SELECT *
FROM company_locations
WHERE state = 'California'
AND city IN ('Irvine', 'San Jose', 'Thousand Oaks')
;
```

8. Which companies have assets of more than \$1,000,000 or are in the software industry?

Solution:

```
SELECT *
FROM company_info
WHERE asset > 1000000000000
OR industry = 'Software'
;
```

9. Which division has SIC codes that mention the word “tree”?

Solution:

```
SELECT division, description
FROM sic
WHERE description LIKE '%tree %' -- The space is important!
;
```

10. Which companies have web pages that do not end in “.com”? It’s okay if your query includes 1-5 false positives.

Solution:

```
SELECT *
FROM company_info
WHERE TRIM(web_page, '/') NOT LIKE '%.com'
;
```

11. Which companies are classified as construction (SIC codes 1500 – 1799)?

Solution:

```
SELECT *
FROM company_info
WHERE sic_code BETWEEN 1500 AND 1799
;
```

12. Which companies are in Silicon Valley, which is approximately bounded by the box (-122.996, 38.092) to (-121.479, 37.039)?

Solution:

```
SELECT *
FROM company_locations
WHERE xcoord BETWEEN -122.996 AND -121.479
AND ycoord BETWEEN 37.039 AND 38.092
;
```

13. Are any rows in the daily share prices table missing the opening or closing price?

Solution:

```
SELECT *
FROM daily_share_prices
WHERE open IS NULL
OR close IS NULL
;
```

14. Are any states missing population estimates?

Solution:

```
SELECT *
FROM state_populations
WHERE popn2010 IS NULL
OR popnestimate2017 IS NULL
;
```

15. How many California cities have a company whose address is on a “street” or “road”?

Solution:

```
SELECT DISTINCT city
FROM company_locations
WHERE state = 'California'
AND (
  -- Be careful: some streets might not end with '.'
  street LIKE '%st.'
```

```
OR street LIKE '%rd.'  
)  
;
```