Commodify

Database Population and Queries

Guest User

Follow these steps for connecting to "commodify-db" on AWS with READ ONLY access:

Username: guest Password: wheat

Port: 3306

Endpoint: commodify-db.cn5sga8k6aq8.us-east-1.rds.amazonaws.com

Run this line in your terminal:

```
mysql --host=commodify-db.cn5sga8k6aq8.us-east-1.rds.amazonaws.com --port=3306 --user=guest -p
```

Input password at prompt.

Queries

Query 1   Analysts and traders dealing with commodity statistics will most commonly use a resource such as this to obtain a schedule of the major supply and demand attributes for a given commodity, for a given country or region, over a given period. For example the below query requests beginning stocks, yield, production, imports, domestic consumption, exports, and ending stocks, for corn, in the United States, year by year from from 2011 to the most recent year.

```
SELECT c.year, c.beginning_stocks, c.yield, c.production,
       c.imports, c.domestic_consumption, c.exports, c.ending_stocks
FROM Commodity c JOIN Political_Entity p ON c.pe_id = p.id
WHERE c.name = 'Corn' AND p.name = 'United States' AND c.year > 2010;
```

Query 2   Users may want to obtain information on the largest producers of commodities. For instance, the query belows requests for the 10 biggest producers of soybeans in 2020, counting states as well as countries.

```
SELECT p.name, c.production
FROM Political_Entity p JOIN Commodity c on p.id = c.pe_id
WHERE c.name = 'Oilseed, Soybean' AND c. year = 2020
ORDER BY c.production DESC
LIMIT 10;
```

Query 3   It is human nature to compare oneself with others. Users may want to find low-performing states that are producing less commodities than another country. An example is the query below which returns all the (USA) states with wheat yields smaller than Canada's last year:

```
SELECT p.name
FROM Political_Entity p JOIN Commodity c on p.id = c.pe_id
WHERE c.name = 'Wheat' AND
      c.year = 2019 AND
      p.is_country = 0 AND
      c.yield < (
        SELECT c2.yield
        FROM Political_Entity p2 JOIN Commodity c2 on p2.id = c2.pe_id
        WHERE p2.name = 'Canada' AND c2.year = 2020 AND c2.name = 'Wheat'
      );
```

Query 4 In view of the threat of climate change, analysts and traders may wish to identify extreme weather-resistant commodities for which the global production levels increase with the occurrence of harsh weather, using U.S. weather data as a proxy for the global weather. The query below requests for a list of commodities that 'thrive' in harsh weather conditions - experienced increase in global production in years of extreme weather (given by top 10 years with highest average temperature)

```
WITH top10highesttemp AS (
    SELECT year FROM (
      SELECT year, avg(temp) AS avg_year_temp
      FROM Weather
      GROUP BY year
    ) a
    ORDER BY avg_year_temp DESC
    LIMIT 10
)
SELECT c.name
FROM (
  SELECT A.name, A.year, A.sum - LAG(A.sum) OVER (ORDER BY A.name, A.year ) AS diff
  FROM (
    SELECT DISTINCT c.name, c.year, SUM(c.production) AS sum
    FROM Commodity c
    GROUP BY c.name, c.year
  ) A
) c JOIN top10highesttemp ON c.year = top10highesttemp.year
WHERE c.diff > 0;
```

Query 5 Weather conditions may be useful in the prediction of production and consumption levels of commodities. For instance, the query below requests for the average temperature and rainfall in top 5 best producing year/month in U.S. states.

```
SELECT AVG(w.temp) AS avg_best_temp, AVG(w.rainfall) AS avg_best_rainfall
FROM Commodity c JOIN Weather w on c.pe_id = w.pe_id
WHERE c.name = 'Oilseed, Soybean'
ORDER BY c.production DESC
LIMIT 5;
```