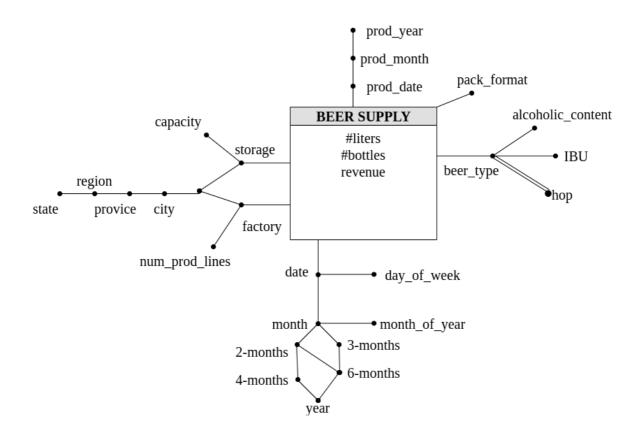


# **Data Warehousing**

Politecnico di Torino

Craft beer

## **Conceptual design**



### Logical design

Primary keys are underlined.

```
BEERSUPPLY (<u>TProdId</u>, <u>TSupId</u>, <u>FId</u>, <u>SId</u>, <u>BId</u>, <u>packFormat</u>, #liters, #bottles, revenue)

PRODUCTIONTIME (<u>TProdId</u>, pdate, pmonth, pyear)

SUPPLYTIME (<u>TSupId</u>, sdate, smonth, s2m, s3m, s4m, s6m, syear, smoy, sdow)

LOCATION (<u>LId</u>, city, province, region, state)

FACTORY(<u>FId</u>, factory, numproductionlines, LId)

STORAGE(<u>SId</u>, storage, capacity, LId)

BEER(<u>BId</u>, beertype, alcoholiccontent, ibu)

HOP(HId, hoptype)

BEERHOP(<u>BId</u>, HId)
```

#### Queries

### **Query A**

Consider beers distributed in packs with the format: 6 bottles. Separately for production month and factory, analyze: the cumulative monthly number of produced liters from the beginning of the year, the monthly percentage of liters produced with respect to the total annual liters for the considered factory, assign a rank to the factories according to the total number of produced liters in decreasing order, separately for each production month.

```
SELECT
    sum(sum(#liters)) over (partition by pyear, fid
        order by pmonth rows unbounded preceding),
    100* sum(#liters)/sum(sum(#liters)) over (partition by pyear, fid),
    rank() over (partition by pmonth order by sum(#liters) desc )
FROM
    productiontime pt, factory f, beersupply bs
WHERE
    bs.tprodid = pt.tprodid and bs.fid = f.fid
    and packformat=6
GROUP BY
    pmonth, pyear, fid
```

## **Query B**

Consider beers produced in 2018. Run the analysis separately for beer type, storage and month of arrival at the storage. Analyze: the average revenue for a bottle, the average daily (consider the day of arrival at the storage) number of produced bottles, the percentage of bottles of the considered beer type with respect to the total considering all the beer types.

```
SELECT
    sum(revenue)/sum(#bottles),
        sum(#bottles)/count(distinct sdate),
        100*sum(#bottles)/sum(sum(#bottles)) over(partition by sid, smonth)
FROM
    beersupply bs, productiontime pt, beer b, storage s, supplytime st
WHERE
    bs.tprodid = pt.tprodid and bs.bid = b.bid and bs.sid = s.sid and
    bs.tsupid = st.tsupdid
    and pyear=2018
GROUP BY
    beertype, sid, smonth
```