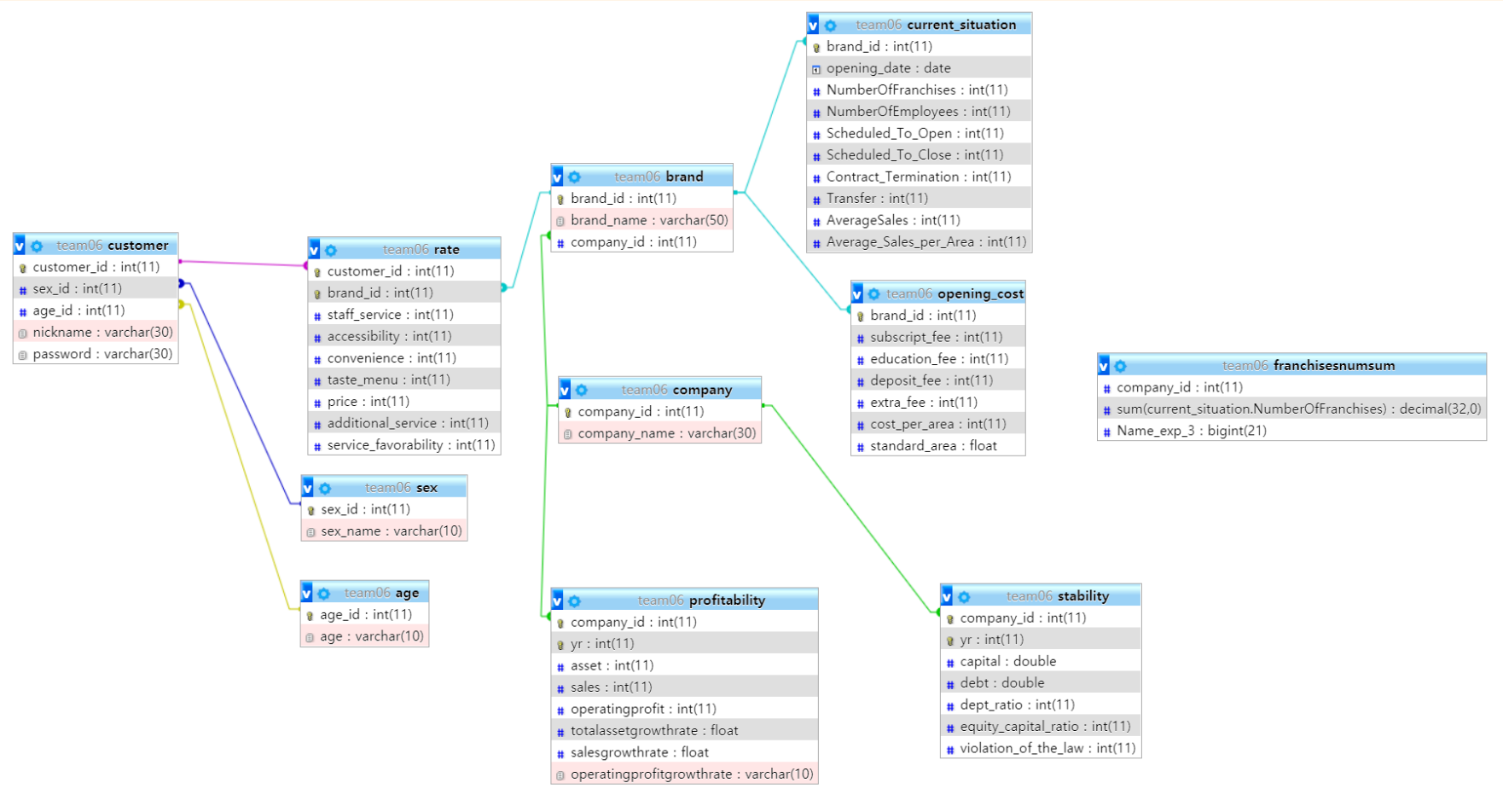
Cafépedia

Big Data Application Project

Analysis of Cafe Brands and Companies Data

1615008 김소림
1615009 김소현
1615049 이서라

Database


team06 customer
customer_id : int(11)
sex_id : int(11)
age_id : int(11)
nickname : varchar(30)
password : varchar(30)
team06 rate
customer_id : int(11)
brand_id : int(11)
staff_service : int(11)
accessibility : int(11)
convenience : int(11)
taste_menu : int(11)
price : int(11)
additional_service : int(11)
service_favorability : int(11)
team06 brand
brand_id : int(11)
brand_name : varchar(50)
company_id : int(11)
team06 current_situation
brand_id : int(11)
opening_date : date
NumberOfFranchises : int(11)
NumberOfEmployees : int(11)
Scheduled_To_Open : int(11)
Scheduled_To_Close : int(11)
Contract_Termination : int(11)
Transfer : int(11)
AverageSales : int(11)
Average_Sales_per_Area : int(11)
team06 opening_cost
brand_id : int(11)
subscript_fee : int(11)
education_fee : int(11)
deposit_fee : int(11)
extra_fee : int(11)
cost_per_area : int(11)
standard_area : float
team06 franchisesnumsum
company_id : int(11)
sum(current_situation.NumberOfFranchises) : decimal(32,0)
Name_exp_3 : bigint(21)
team06 company
company_id : int(11)
company_name : varchar(30)
team06 sex
sex_id : int(11)
sex_name : varchar(10)
team06 age
age_id : int(11)
age : varchar(10)
team06 profitability
company_id : int(11)
yr : int(11)
asset : int(11)
sales : int(11)
operatingprofit : int(11)
totalassetgrowthrate : float
salesgrowthrate : float
operatingprofitgrowthrate : varchar(10)
team06 stability
company_id : int(11)
yr : int(11)
capital : double
debt : double
dept_ratio : int(11)
equity_capital_ratio : int(11)
violation_of_the_law : int(11)

5 Advanced Queries

1. Windowing

```
SELECT company_id, yr, sales, SUM(sales)
OVER(PARTITION BY company_id
ORDER BY yr ROWS BETWEEN UNBOUNDED
PRECEDING AND CURRENT ROW) as accum
FROM profitability WHERE company_id=$company_id
```

2. NTILE

```
SELECT stability.company_id, debt, ntile(5) over(order by
debt asc) as debt_rank
FROM stability
GROUP BY stability.company_id;
```

3. RANK

```
SELECT company.company_id,
stability.capital,company.company_name ,rank() over (order by
stability.capital DESC) num
FROM stability, company
WHERE stability.company_id=company.company_id
AND stability.yr=2020
```

4. DENSE RANK

```
SELECT brand.brand_name, AVG(rate.staff_service) as avg,
DENSE_RANK() OVER (ORDER BY avg DESC) AS rank
FROM rate, customer,brand
WHERE rate.customer_id=customer.customer_id
AND customer.sex_id=1 AND brand.brand_id=rate.brand_id
AND customer.age_id=1
GROUP BY rate.brand_id
ORDER BY AVG(rate.staff_service) DESC
```

5. ROLL UP

```
SELECT brand.company_id, brand.brand_name,
current_situation.opening_date,
SUM(current_situation.NumberOfFranchises) as numfran,
SUM(current_situation.NumberOfEmployees) as numemp,
SUM(current_situation.AverageSales) as avsales,
SUM(current_situation.Average_Sales_per_Area)
as avesalesperarea
FROM current_situation, brand, company
WHERE current_situation.brand_id = brand.brand_id and
company.company_id = brand.company_id and
company.company_id = $company_id
GROUP BY brand.company_id, brand.brand_name WITH ROLLUP
```