Write Hive Queries for the Below analysis.

**1) To Find the impact of education qualification and occupation on total purchase-
for each education qualification under an occupation, find the average
totalpurchaseytd**

<mark>(Sol)</mark>
select Education, Occupation, avg(totalpurchaseytd) num_of_occurance from individual
group by 1,2;

Bachelors	Professional	7479.308333333334
Bachelors	Management	1854.8718867924529
Partial College	Skilled Manual	3204.7428571428572
High School	Skilled Manual	2829.4961363636367
Partial College	Clerical	5297.515555555556
Partial High School	Clerical	3735.447
Graduate Degree	Management	3504.454
Partial College	Professional	2882.080384615386
High School	Professional	3177.532380952381
Partial High School	Skilled Manual	3430.7933333333335

**2) Use grouping sets OR Roll-ups OR Cubes to find avg of totalchildren based on all
records, yearly income brackets, and (occupation and qualification). Use grouping IDs
to segregate the reports**

<mark>(Sol)</mark>
select Occupation, Education, YearlyIncome, avg(TotalChildren) from individual
group by Occupation, Education, YearlyIncome
with rollup;

Clerical	Partial College	25001-50000	2.1111
Clerical	Partial College		2.1111
Clerical	Partial High School	0-25000	2.0000
Clerical	Partial High School	25001-50000	0.0000
Clerical	Partial High School		1.0000
Clerical			1.5263
Management	Bachelors	25001-50000	2.0000
Management	Bachelors	50001-75000	3.9474
Management	Bachelors	75001-100000	2.4667
Management	Bachelors	greater than 100000	1.6429
Management	Bachelors		2.7358
Management	Graduate Degree	25001-50000	2.0000
Management	Graduate Degree	50001-75000	3.3846
Management	Graduate Degree	75001-100000	2.2500
Management	Graduate Degree	greater than 100000	0.9375
Management	Graduate Degree		2.0571
Management			2.4659
Professional	Bachelors	50001-75000	1.1053
Professional	Bachelors	75001-100000	3.0000

```
Professional    Bachelors                  1.5000
Professional    High School    25001-50000  2.0000
Professional    High School    50001-75000  2.1250
Professional    High School    75001-100000 2.0000
Professional    High School    greater than 100000   3.0000
Professional    High School                 2.1905
Professional    Partial College50001-75000  2.7308
Professional    Partial College75001-100000 2.8125
Professional    Partial Collegegreater than 100000    2.3000
Professional    Partial College             2.6731
Professional                                 2.2784
Skilled ManualHigh School      0-25000              3.8571
Skilled ManualHigh School      25001-50000  0.6000
Skilled ManualHigh School      50001-75000  2.0000
Skilled ManualHigh School      75001-100000 2.0000
Skilled ManualHigh School                   1.3409
Skilled ManualPartial College25001-50000    0.4118
Skilled ManualPartial College50001-75000    1.6923
Skilled ManualPartial College75001-100000   1.6667
Skilled ManualPartial College                1.2449
Skilled ManualPartial High School    25001-50000   2.0000
Skilled ManualPartial High School    75001-100000  2.0000
Skilled ManualPartial High School                2.0000
Skilled Manual                 1.3125
                          1.9767
```

3) Sort the number of cars owned after grouping the records by yearly income brackets

(Sol)
select YearlyIncome, sum(NumberCarsOwned) cars_owned from individual
group by YearlyIncome
order by 2;

```
0-25000          24
75001-100000 100
greater than 100000    127
25001-50000  135
50001-75000  190
```

4) use row_number() to rank the commuteDistance in demographics data partitioned by yearly income

(Sol)

select * from (select CustomerID, gender, row_number() over (partition by YearlyIncome
order by CommuteDistance) as rnk from individual) j where j.rnk < 3;

NOTE
-> For each question mentioned above, provide the Query and the output screenshots as Solutions.
-> The queries can be created based on personal understanding and in any method within the boundaries of the question, but the output must include the desired results.