Create database insurance_claims;

Use insurance_claims;

Create table claims(months_as_customer integer, age integer, policy_number integer, policy_bind_date varchar(15), policy_state char(5), policy_csl varchar(10), policy_deductable integer,policy_annual_premium integer,umbrella_limit integer,insured_zip integer,insured_sex varchar(10), insured_education_level varchar(15),insured_occupation varchar(25),insured_hobbies varchar(15),insured_relationship varchar(15), capital_gains integer, capital_loss integer, incident_date varchar(15), incident_type varchar(25),collision_type varchar(20),incident_severity varchar(20), authorities_contacted varchar(10),incident_state varchar(2),incident_city varchar(15),incident_location varchar(25),incident_hour_of_the_day integer, number_of_vehicles_involved integer, property_damage varchar(10),bodily_injuries integer, witnesses integer, police_report_available varchar(15), total_claim_amount integer, injury_claim integer, property_claim integer, vehicle_claim integer, auto_make varchar(15));

Select * from claims

limit 15;

/*1.   Calculate the proportion of claim spend on injury, property and vehicle (total).*/

Select round((sum(injury_claim)/sum(total_claim_amount))*100,2) as percentage_of_injury_claim,

round((sum(property_claim)/sum(total_claim_amount))*100,2) as percentage_of_property_claim,

round((sum(vehicle_claim)/sum(total_claim_amount))*100,2) as percentage_of_vehicle_claim

from claims;

/*2.   Calculate the proportion of claim spend on injury, property and vehicle (for top 10 total claims).*/

With cte as (select policy_number, injury_claim, property_claim, vehicle_claim, total_claim_amount, dense_rank() over (order by total_claim_amount desc) as rnk from claims)

Select round((sum(injury_claim)/sum(total_claim_amount))*100,2) as percentage_of_injury_claim,

round((sum(property_claim)/sum(total_claim_amount))*100,2) as percentage_of_property_claim,

round((sum(vehicle_claim)/sum(total_claim_amount))*100,2) as percentage_of_vehicle_claim,rnk

from cte

group by policy_number

having rnk<11

order by rnk;

/*3.   Create a visualization that provides a breakdown between the male and female insurers,

along with education level each year, starting from 1990.*/

Alter table claims

add column New_bind_date date;

Update claims

set new_bind_date=str_to_date(policy_bind_date,"%d-%m-%Y");

SET SQL_SAFE_UPDATES = 0;

select new_bind_date from claims;

Select insured_sex,insured_education_level,year(new_bind_date) as insurance_year,count(insured_sex) as total_number

from claims

group by 1,2,3

order by 1,2,3;

/*4.   Compare the number of insurers regionwise.*/

Select incident_state, count(*) as total_insurers from claims

group by incident_state

order by total_accidents desc;

/* 5.   Comment on the relationship between deductible and premium.*/

Select (count(*) * SUM(policy_deductable*policy_annual_premium) - SUM(policy_deductable) * SUM(policy_annual_premium)) /

(SQRT(count(*) * SUM(policy_deductable*policy_deductable) - SUM(policy_deductable) * SUM(policy_deductable)) *

SQRT(count(*) * SUM(policy_annual_premium*policy_annual_premium) -

SUM(policy_annual_premium) * SUM(policy_annual_premium)))

AS correlation_coefficient

        from claims;

/* 6.   Which date had the maximum number of accidents? */

Alter table claims

add column new_incident_date date;

Update claims

set new_incident_date=str_to_date(incident_date,"%d-%m-%Y");

SET SQL_SAFE_UPDATES = 0;

select new_incident_date, incident_date from claims;

Select new_incident_date, count(*) as number_of_accidents

from claims

where incident_type != "Vehicle Theft"

group by new_incident_date

order by number_of_accidents desc

limit 1;

/* 7.   Which age group is most likely to meet an accident? */

Select

Case when age<=29 then "19-29"

   when age<=39 then "30-39"

        when age<=49 then "40-49"

   when age<=59 then "50-59"

   Else "59+"

   end as "age_bin", count(*) as count

```
        from claims
        group by age_bin
        order by age_bin;
```

/* 8.   Compare capital gain and capital loss and comment on profit.*/

```
Select sum(capital_gains) as total_gains, sum(capital_loss) as total_loss,
(capital_gains- capital_loss) as profit
from claims;
```

/* 9.   Are females more likely to take benefit of automobile insurance? */

```
select
sum(case when insured_sex= "Male" then 1 else 0 end) as number_of_males,
sum(case when insured_sex= "Female" then 1 else 0 end) as number_of_females
from claims;
```

/* 10.   Which auto making company had the most accidents? */

```
Select auto_make, count(*) as number_of_accidents
from claims
where collision_type like "%Collision%"
group by auto_make
order by number_of_accidents desc
limit 1;
```