Interview questions BI Developer position for Athlon International

Exercise 1.

You receive a bug report from your frontend colleague that duplicate records seem to appear for cars in the sold_cars dashboard.

The data flow into this dashboard is as follows:

Extract CRM database tables for a country >
Load tables in Snowflake >
Historized tables in Snowflake inserts new, updated, and deleted records based on the Load table >
Mart view of sold cars combines data from historized tables >
Sold cars data gets loaded into QlikSense and served as a dashboard

One of your colleagues has applied the following fix in the Mart view, removing the duplicate issue in the report:

```sql
WITH dimdate AS (
    SELECT * FROM DWH_DATA_MODEL.DWH_DATA_MODEL.dim_date
      WHERE date = last_day(add_months(current_date,-1))

    ),

 ...

  rel AS (
    SELECT *
    FROM (
    SELECT row_number() OVER (PARTITION BY nr ORDER BY ifnull(dss_update_time,
'1800-01-01') DESC) AS rownum
        ,rel.*
    FROM dimdate dd
    INNER JOIN "DATASTORE_BE"."DATASTORE"."DS_S_BE_ATLAS_REL" rel ON dd.load_date
BETWEEN trunc(rel.dss_start_date, 'day')
            AND dateadd(day, - 1, trunc(rel.dss_end_date, 'day'))
        AND rel.dss_crud <> 'D'
    )
    WHERE rownum = 1
  ),

  ...
```

- Can you explain what this fix does to solve the duplicate issue? Do you like this solution?
- Can you think of possible underlying problems in the DS_S_BE_ATLAS_REL (Historized) table that's causing the duplicates?
- How would you go about addressing these issues and what are things you should be wary of when applying a fix to the historized table?

Exercise 2.

At the end of this document, you find an output of columns found in the active portfolio table.

- As the market is moving to electric vehicles, we are interested to see how far along this shift in fuel type in Athlon's fleet:
    - Write a query that shows the number of new EV contracts from year to year, as a percentage of the total cars in that country's active portfolio for each country.
    - Also show how much each country contributes to the electrification of Athlon's total fleet per year.
- In 2019 Dutch licence plates changed their pattern from 9-XX-999 to 999-XX-9 (9 for numbers x for letters). The new pattern gets applied as soon as the old pattern has been fully used up per model.
    - Write a query that pinpoints the date of the first occurrence of the new pattern per brand in the Netherlands (NL).
- Write a query that shows the percentual movement of the average list price for the different Mercedes models per year.

- CONTRACTMILEAGE
- IDCONTRACT
- MONTHLYROADASSISTANCE
- INSURANCETYPE
- COUNTRYCODE
- MILEAGEEXCESS
- CONTRACTRETURNEDDATE
- CURRENCY
- DISCOUNTVEHICLE
- FLEETNAME
- FUELTYPE
- MONTHLYINTEREST
- NETLISTPRICE
- VEHICLETYPE
- MONTHLYRM
- RMSPEND
- VEHICLEBRAND
- TRANSMISSION
- VEHICLESTATUS
- MONTHLYDEPRECIATION
- DATAPROVIDERCODE
- DATEFIRSTREGISTRATION
- MONTHLYCASCO
- STOCKSINCE
- BUYBACK
- CUSTOMERHOLDINGNAME
- ISRENTAL
- TAX
- CONTRACTSTARTDATE
- DOORS
- MONTHLYMGTFEE
- ACCESSORIES
- TRACTION
- MONTHLYOTHER
- BOOKVALUE
- NETLISTPRICEOPTIONS
- IDCOUNTRYVEHICLE
- RESIDUALVALUE
- INVESTMENT
- CONTRACTSHARE
- MONTHLYOTHERINSURANCE
- RESIDUALVALUENONUPLIFT
- MONTHLYROADTAX
- MONTHLYRENTAL
- MONTHLYTIRES
- MODELNAME
- MONTHLYTPL
- RMINVOICED
- HORSEPOWER
- CONTRACTDURATION

- INTERESTRATE
- EMISSIONNORM
- RECALC
- STARTMILEAGE
- CUSTOMERID
- FRAMENUMBER
- MONTHLYTERM
- PORTFOLIOPERIOD
- DESCRIPTION
- MONTHLYFUEL
- DISCOUNTSERVICE
- LASTKNOWNMILEAGEDATE
- CO2
- IMPAIRMENT
- INTERESTTYPE
- FUNDINGPERCENTAGE
- BODYTYPE
- CONTRACTENDDATE
- CONTRACTSORT
- CUSTOMERNAME
- LASTKNOWNMILEAGE
- MONTHLYVAT
- LICENSEPLATE