Jacob, Mia, Christian

Brainstorm:

1. En virksomhed skal have overblik over hvilke medarbejder der sjællendt møder ind til tiden.
2. En virksomhed der på baggrund af en masse data om salg skal finde en løsning på hvordan de kan optimere salget.
    a. https://www.kaggle.com/datasets/divyeshardeshana/warehouse-and-retail-sales
    b. https://www.kaggle.com/datasets/rutuspatel/walmart-dataset-retail

    Spørgsmål:
    - Hvor meget kan vi sælge hvis vejret er godt?
    - Hvornår sælger vi mest?
    - Hvor meget er det højeste vi har solgt for?
    - Hvad kan man gøre for at sælge mere?
3. Et system for en dyrehandel som skal kunne afgøre hvem der har bedst mulighed til at adoptere et bestemt dyr (ML, KNN).
4. En betting virksomhed der bruger statistikker fra sportsresultater til at afgøre hvilke odds de skal sætte.
5. En bilforhandler der lavet statistik på brugte biler for hvor meget de skal koste og hvad standen er.
    a. https://www.kaggle.com/datasets/harikrishnareddyb/used-car-price-predictions
    b. https://www.kaggle.com/datasets/rinichristy/2022-fuel-consumption-ratings
    c. https://www.kaggle.com/datasets/sudiwilliams/19951999-fuel-consumption-ratings
6. statistik på hvad for nogle biler der forurener mest pr kørt kilometer i forhold til størrelse og vægt.
7. Tjek om man har råd til hus af en hvis størrelse.
    a. https://www.kaggle.com/datasets/harlfoxem/housesalesprediction
    b. https://www.kaggle.com/datasets/shree1992/housedata
    c. https://www.kaggle.com/datasets/dumburanjith/house-sales-prediction-and-classification

    Spørgsmål:
    - Hænger størrelse og pris sammen? Og er det en lineær sammenhæng? (Regression)
    - Hænger lokation og udsigt sammen med prisen? (Regression)
    - Hvis jeg har et bestemt antal penge, hvor stort et hus kan jeg så købe? (ML)

Jacob, Mia, Christian

Cars:

The context of this assignment is to find connections between the car models, their price and how much they pollute in Canada in the years 1995-1999 and 2022. This will be valuable for environmental activist groups. It will also be helpful for the car model companies, so that they can use the data to sell more cars. For instance if a car in 1999 pollutes 100% and in 2022 it only pollutes 50% then they can use that as advertisement, because people will often buy cars that are more environmentally good.
We have three datasets, the first has data on the car models and prices and where the car was sold. The second has data on the model-specific fuel consumption ratings and estimated carbon dioxide emissions for new cars in Canada 2022. The last dataset has also data on the fuel consumption ratings and estimated carbon dioxide emissions in Canada in the years 1995-1999.

Research questions:

- What models pollute the most?
- Is there a connection between the car model and it's price?
- Is there a connection between price and the city it is from? with taxes and more perhaps.
- What models have improved in preventing pollution?
- What year, between 1995 and 1999, was there the most pollution?

Hypotheses:

Our hypothesis is that the older models pollutes more, and the price on a car will be the same if it is being sold in the USA, Canada or US. We expect the models to pollute less in 2022 than in 1999 and earlier.