# Frågor – Kategorisk Data

Läs länken: <https://machinelearningmastery.com/one-hot-encoding-for-categorical-data/>

I Scikit-learn måste all data konverteras till numeriska variabler för att kunna användas. Besvara nedanstående instuderingsfrågor efter att ha läst länken ovan:

1. **Vad kännetecknar en nominell variabel?**

Svar:

1. **Vad kännetecknar en ordinal variabel?**

Svar:

1. **Förklara (gärna med ett exempel): Ordinal encoding, one-hot encoding, dummy variable encoding.**

Svar:

**Om vi använder vanlig linjär regression, skall vi använda one-hot-encoding eller dummy-variable encoding?**

Svar:

1. Läs dokumentationen i Pandas nedan. Notera, oftast är Pandas smidigare att använda än scikit-learn. <https://pandas.pydata.org/docs/reference/api/pandas.get_dummies.html>  
   Notera, "***drop\_first****=False"*styr om det blir "one-hot-encoding" eller dummy variable encoding.