Welcome to week 5.

This week, we will shift gears a bit and talk about categorical data.

So far we've been focused on quantitative data.

But this year we'll look at what happens when we want to

predict a categorical variable such as whether a customer

will make a purchase or commit fraud as

a function of a large number of quantitative variables.

We'll look at linear and quadratic discriminant analysis which let

us look at our categories as multivariate normal distributed groups.

Then we'll look at an approach that does not make this type of assumption called

support vector machines which are very powerful and flexible approach,

although at the cost of interpretability.

Welcome to week 5.