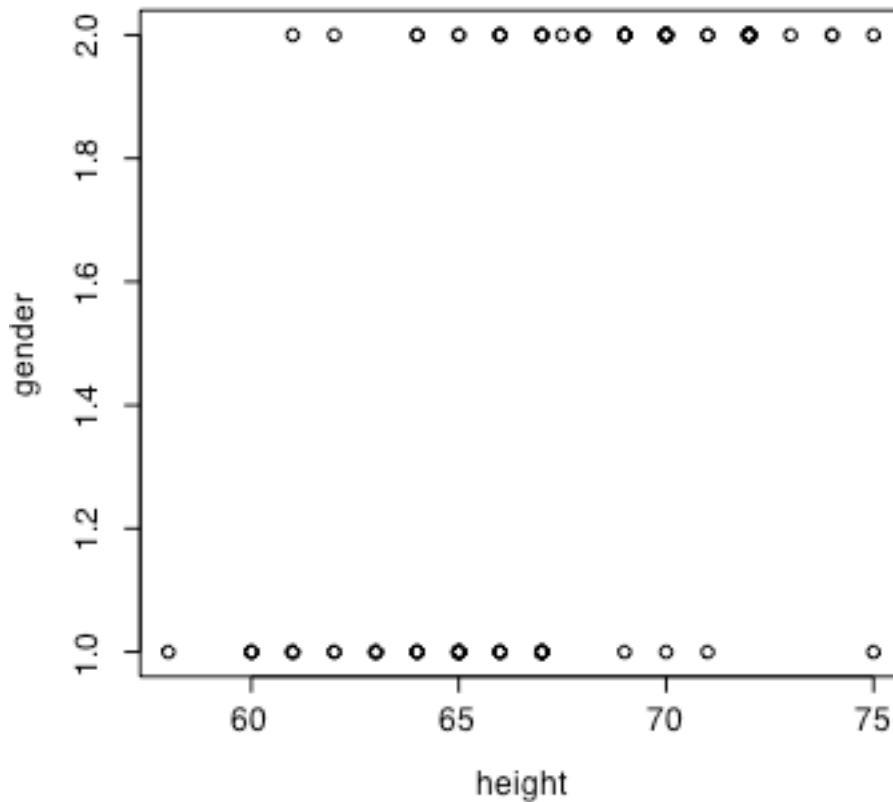


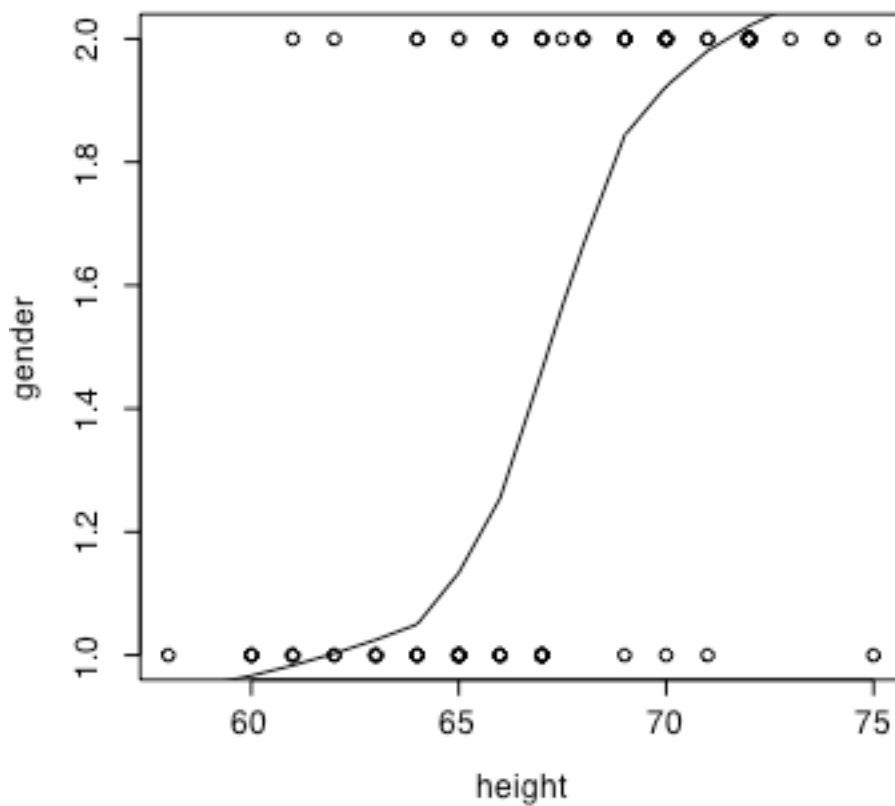
R notes for Logistic Regression

```
> survey.table <- read.table("survey.txt", header=T, sep="\t")
> names(survey.table)
[1] "type"      "gender"    "weight"    "smoke"     "ghosts"    "age"
[7] "attractive" "circumference" "bushage"   "ticket"    "socialsci" "flu"
[13] "tire"
> survey.table <- survey.table[-c(58,59),] #Remove outliers
> plot(height, gender)
```



Can we predict gender from height? We'll try to fit a model that will estimate the probability a randomly selected person from those of a certain height is male. For example, what's the probability a randomly selected 65" tall person is male?

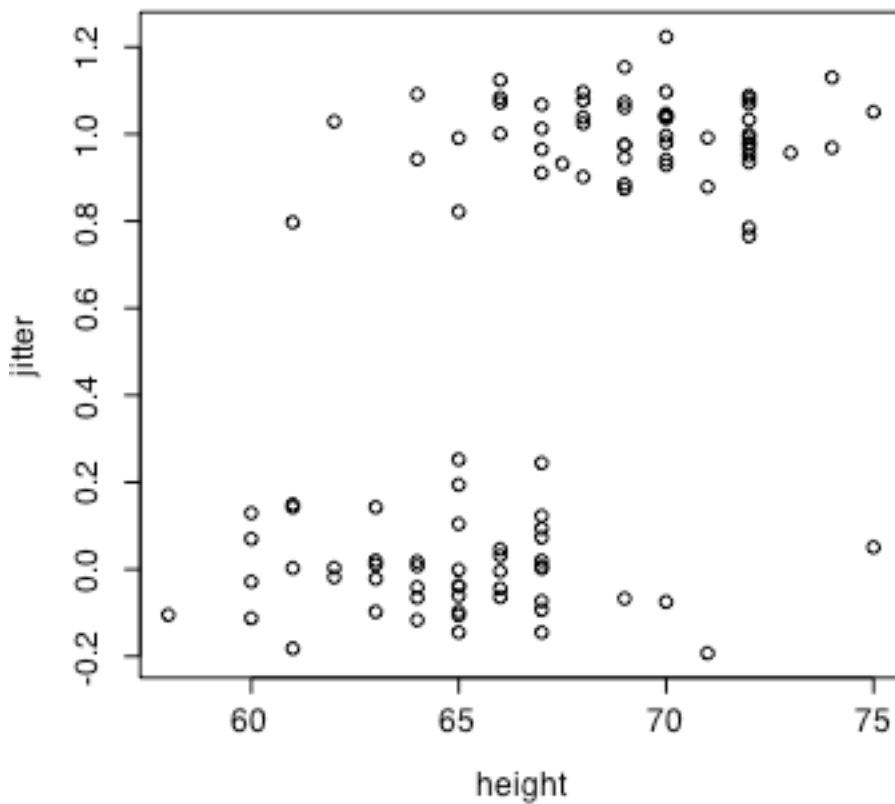
```
> lines(lowess(height, gender))
```



```

> gender2 <- as.numeric(gender)
> gender2[1:10]
[1] 1 1 2 1 1 2 2 2 1 1
> gender[1:10]
[1] f f m f f m m m f f
Levels: f m
> gender2[gender2==1]<- 0
> gender2[gender2==2] <- 1
> jitter <- gender2 + rnorm(length(gender2), 0, .1)
> plot(height, jitter)

```



Fitting the model requires some inputs that we don't have time to explain in this course.

```
> height.glm <- glm(gender2~height, family=binomial(link=logit))
> summary(height.glm)
```

Call:

```
glm(formula = gender2 ~ height, family = binomial(link = logit))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.8428	-0.8508	0.3442	0.7543	2.3798

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-32.37299	6.31080	-5.13	2.9e-07 ***
height	0.48528	0.09422	5.15	2.6e-07 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 149.13 on 107 degrees of freedom
Residual deviance: 101.88 on 106 degrees of freedom
AIC: 105.88

Number of Fisher Scoring iterations: 5

The important points are the coefficient of height and the fact that it is statistically significant, which means that height is helpful in determining gender. To translate this, we need to exponentiate the coefficient

```
> exp(.48528)  
[1] 1.62463
```

This means that a 1" difference in height is associated with a 1.6 increase in odds. That is, a person who is one inch taller than another is 1.6 times more likely to be male than is the shorter person.