Assignment 1 Report

2025-02-13

Exercise 2

a)

We first plot the data:

Crops by County and Related Factor
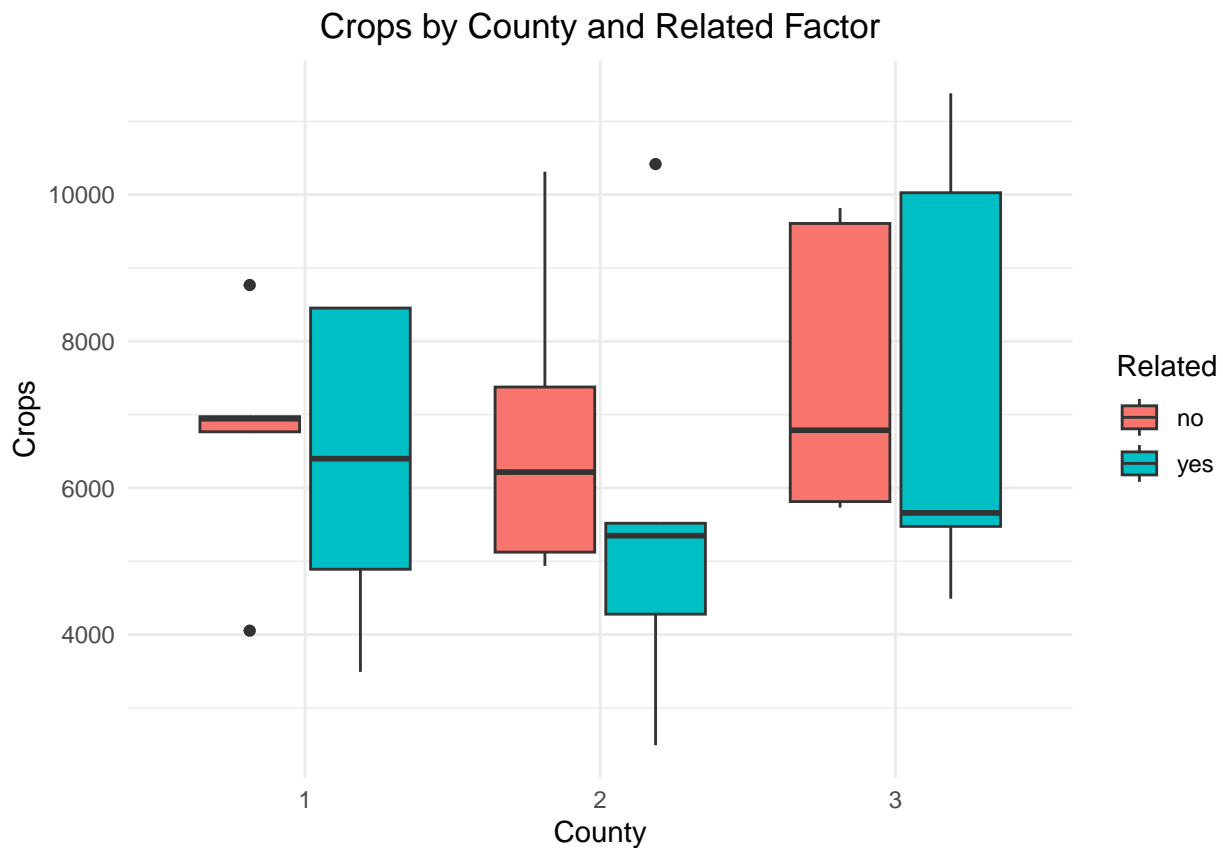
We then perform ANOVA only with County and Related:

```
model = lm(Crops ~ County + Related, data=data); anova(model)
```

```
## Analysis of Variance Table
##
## Response: Crops
##           Df    Sum Sq Mean Sq F value Pr(>F)
## County     2   8841441 4420721  0.8192 0.4518
## Related    1   2378957 2378957  0.4409 0.5126
## Residuals 26 140303437 5396286
```

We follow by including the interaction between the two:

```
model_int = lm(Crops ~ County * Related, data=data); anova(model_int)
```

```
## Analysis of Variance Table
##
## Response: Crops
##                Df    Sum Sq  Mean Sq F value Pr(>F)
## County          2   8841441  4420721  0.7644 0.4766
## Related         1   2378957  2378957  0.4113 0.5274
## County:Related  2   1497573   748786  0.1295 0.8792
## Residuals      24 138805865  5783578
```

The interaction terms are not significant, let's use the first model; are its residuals normally distributed?

Normal Q–Q Plot
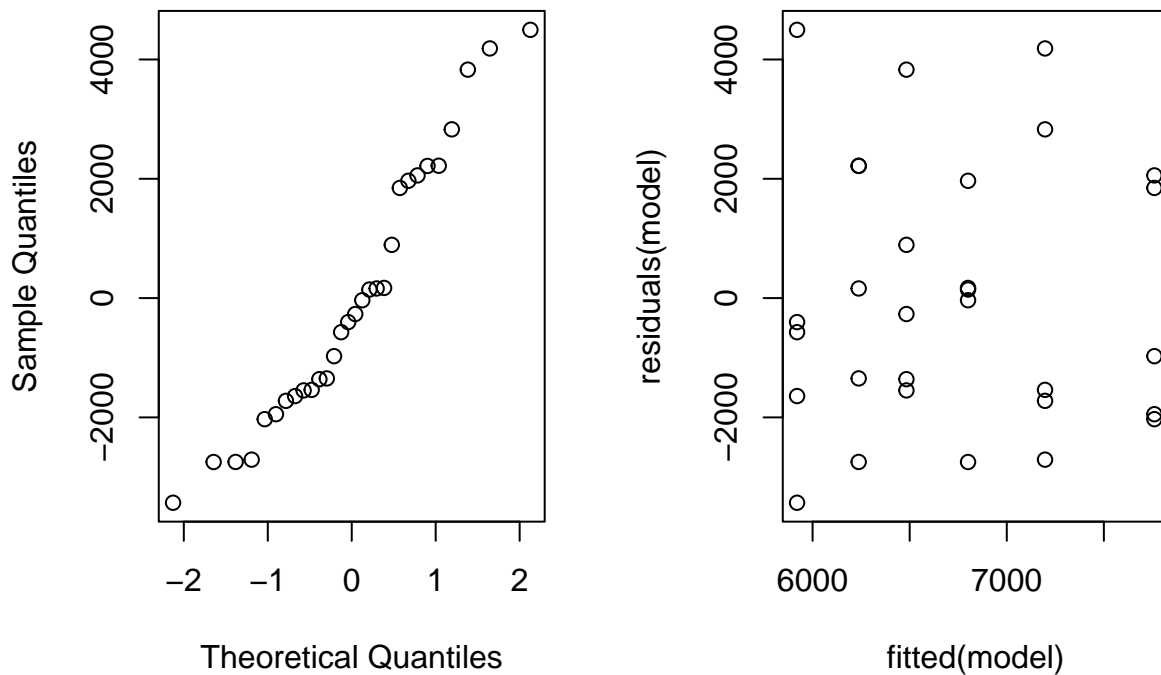
Now we want to estimate the crops for a typical farm in County 3 for which landlord and tenant are not related. Let's look at the model results:

```
summary(model)$coefficients
```

```
##              Estimate Std. Error     t value      Pr(>|t|)
## (Intercept)    6800.6   848.2363   8.0173415 1.698496e-08
## County2        -317.3  1038.8730  -0.3054271 7.624728e-01
## County3         959.7  1038.8730   0.9237895 3.640904e-01
## Relatedyes     -563.2   848.2363  -0.6639659 5.125576e-01
```

Thus, to calculate the desired value we compute:

$$Crops_{cou.=3,rel.=no} = 6800.6 + 959.7 = 7760.3$$