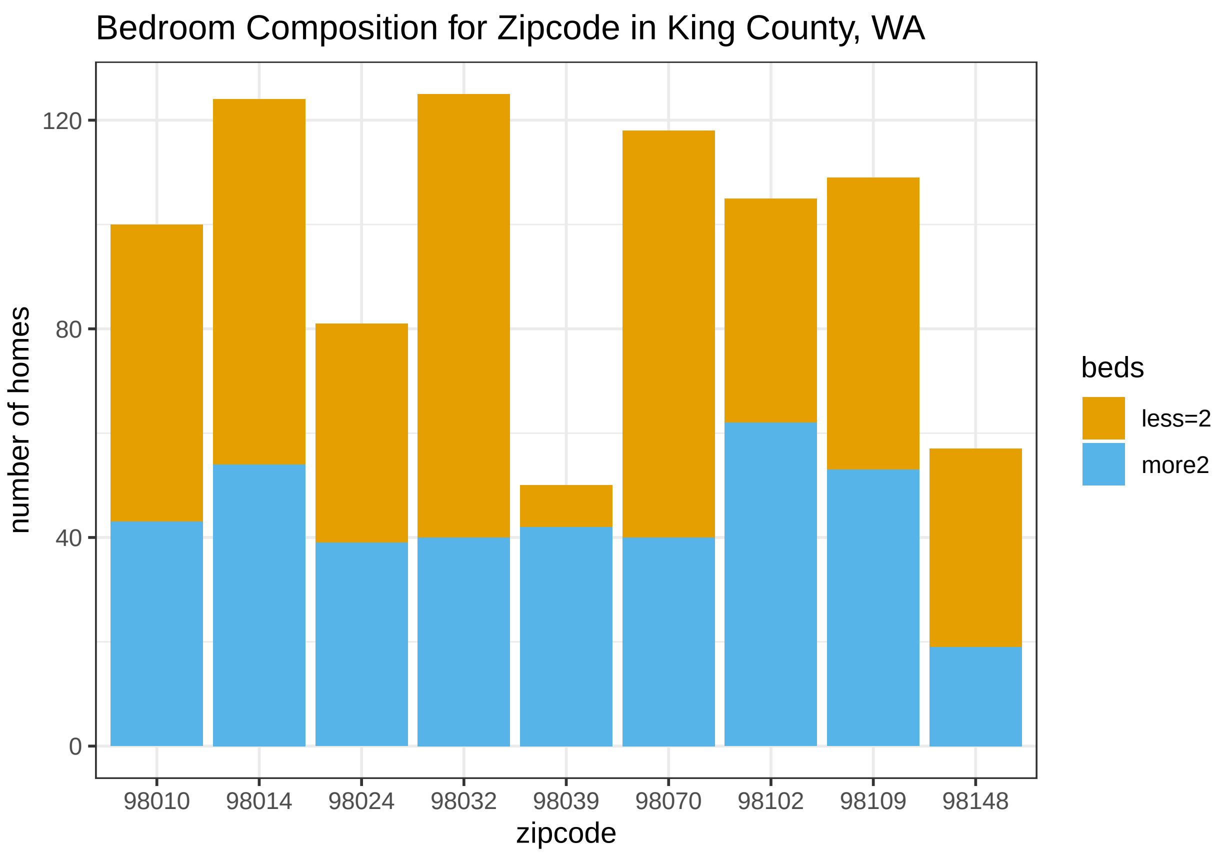
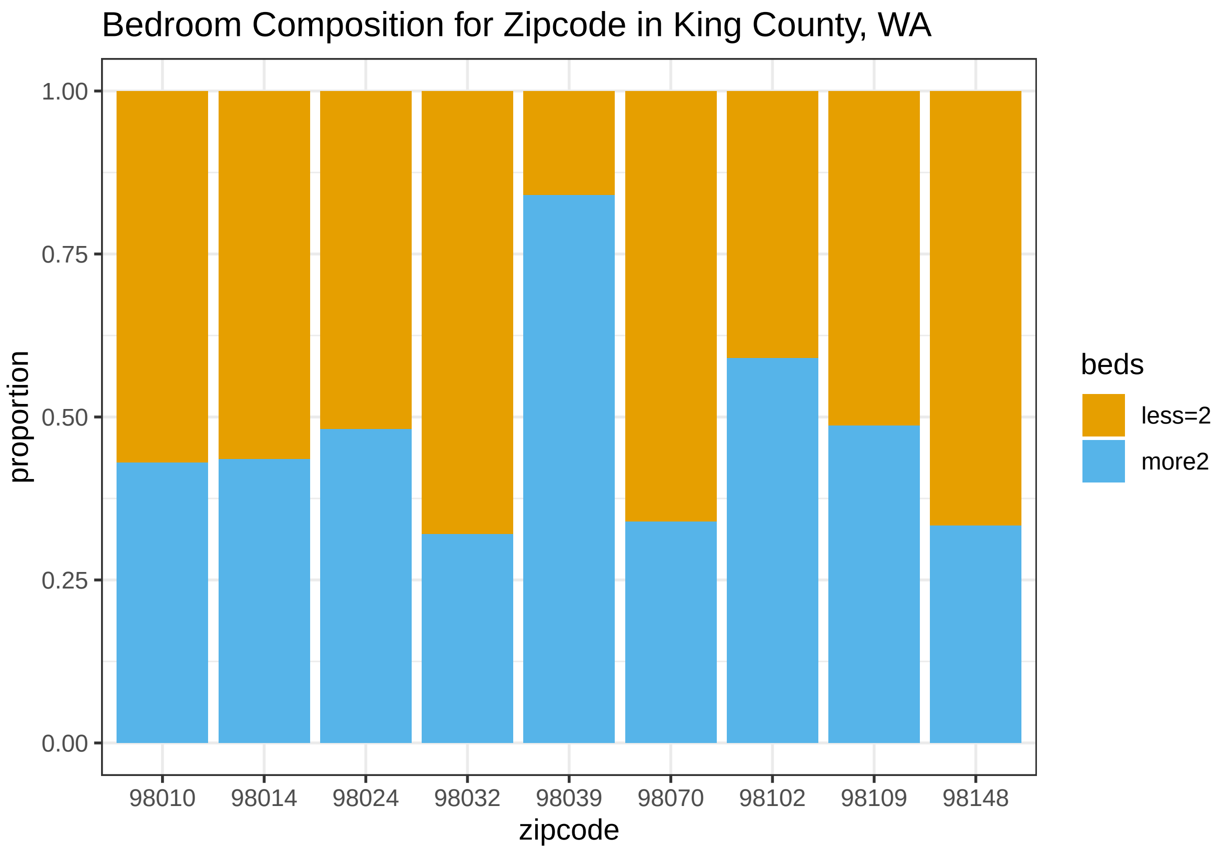
Hierarchical GLMs





Hierarchical GLMs

```r
glm1 <- glm(cbind(more2,lessequal2) ~ 1, data = seattle, family = binomial)
display(glm1)
```

```
## glm(formula = cbind(more2, lessequal2) ~ 1, family = binomial,
##     data = seattle)
##             coef.est coef.se
## (Intercept) -0.20     0.07
## ---
##   n = 869, k = 1
##   residual deviance = 1196.4, null deviance = 1196.4 (difference = 0.0)
```

```r
invlogit(coef(glm1))
```

```
## (Intercept)
##   0.4510932
```

```r
stan1 <- stan_glm(cbind(more2,lessequal2) ~ 1, data = seattle, family = binomial, refresh = 0)
print(stan1)
```

```
## stan_glm
##  family:       binomial [logit]
##  formula:      cbind(more2, lessequal2) ~ 1
##  observations: 869
##  predictors:   1
## ------
##             Median MAD_SD
## (Intercept) -0.2    0.1
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```r
invlogit(coef(stan1))
```

```
## (Intercept)
##   0.4511562
```

```r
seattle %>% summarise(freq = mean(more2))
```

```
## # A tibble: 1 x 1
##    freq
##   <dbl>
## 1 0.451
```

```r
glmer1 <- stan_glmer(cbind(more2,lessequal2) ~ 1 + (1 | zipcode),
                     data = seattle, family = binomial, refresh = 0)
print(glmer1)
```

```
## stan_glmer
##  family:       binomial [logit]
##  formula:      cbind(more2, lessequal2) ~ 1 + (1 | zipcode)
##  observations: 869
## ------
##             Median MAD_SD
## (Intercept) -0.1    0.2
##
## Error terms:
##  Groups  Name        Std.Dev.
##  zipcode (Intercept) 0.75
## Num. levels: zipcode 9
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```r
fixef(glmer1)
```

```
## (Intercept)
##  -0.1027465
```

```r
ranef(glmer1)
```

```
## $zipcode
##       (Intercept)
## 98010 -0.16656036
## 98014 -0.14945066
## 98024  0.01306696
## 98032 -0.60303269
## 98039  1.34142591
## 98070 -0.51789418
## 98102  0.41693212
## 98109  0.04043826
## 98148 -0.50169135
##
## with conditional variances for "zipcode"
```

```r
coef(glmer1)
```

```
## $zipcode
##       (Intercept)
## 98010 -0.26930690
## 98014 -0.25219720
## 98024 -0.08967958
## 98032 -0.70577923
## 98039  1.23867937
## 98070 -0.62064073
## 98102  0.31418558
## 98109 -0.06230828
## 98148 -0.60443789
##
```

```
## attr(,"class")
## [1] "coef.mer"
```

```
seattle %>% group_by(zipcode) %>% summarise(freq = mean(more2), n = n()) %>%
  ungroup() %>%
  bind_cols(tibble(glmer_est = invlogit(coef(glmer1)$zipcode[[1]]))
)
```

```
## # A tibble: 9 x 4
##   zipcode  freq     n glmer_est
##   <fct>   <dbl> <int>     <dbl>
## 1 98010   0.43    100     0.433
## 2 98014   0.435   124     0.437
## 3 98024   0.481    81     0.478
## 4 98032   0.32    125     0.331
## 5 98039   0.84     50     0.775
## 6 98070   0.339   118     0.350
## 7 98102   0.590   105     0.578
## 8 98109   0.486   109     0.484
## 9 98148   0.333    57     0.353
```

```r
glmer2 <- stan_glmer(cbind(more2,lessequal2) ~  scale_sqft + (1 + scale_sqft | zipcode),
                 data = seattle, family = binomial, refresh = 0)
print(glmer2)
```

```
## stan_glmer
##  family:       binomial [logit]
##  formula:      cbind(more2, lessequal2) ~ scale_sqft + (1 + scale_sqft | zipcode)
##  observations: 869
## ------
##             Median MAD_SD
## (Intercept) 0.0    0.1
## scale_sqft  2.6    0.3
##
## Error terms:
##  Groups  Name        Std.Dev. Corr
##  zipcode (Intercept) 0.36
##          scale_sqft  0.73     -0.41
## Num. levels: zipcode 9
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```r
fixef(glmer2)
```

```
## (Intercept)  scale_sqft
## 0.008075054 2.562810019
```

```r
ranef(glmer2)
```

```
## $zipcode
##       (Intercept)  scale_sqft
## 98010 -0.19355120  0.32465085
## 98014 -0.11171548  0.37711644
## 98024 -0.09345333  0.30622201
## 98032 -0.05358081 -0.08713374
## 98039  0.01501471 -0.01143682
## 98070 -0.22437035  0.27526049
## 98102  0.43709917 -1.20124400
## 98109  0.13083105 -0.30786904
## 98148  0.06550678  0.25440726
##
## with conditional variances for "zipcode"
```

```r
coef(glmer2)
```

```
## $zipcode
##       (Intercept) scale_sqft
## 98010 -0.18547615   2.887461
## 98014 -0.10364042   2.939926
## 98024 -0.08537828   2.869032
## 98032 -0.04550575   2.475676
## 98039  0.02308976   2.551373
## 98070 -0.21629529   2.838071
## 98102  0.44517422   1.361566
## 98109  0.13890610   2.254941
```

```
## 98148  0.07358183   2.817217
##
## attr(,"class")
## [1] "coef.mer"
```

Stan

```
data {
  int<lower=1> D;
  int<lower=0> N;
  int<lower=1> L;
  int<lower=0,upper=1> y[N];
  int<lower=1,upper=L> ll[N];
  row_vector[D] x[N];
}
parameters {
  real mu[D];
  real<lower=0> sigma[D];
  vector[D] beta[L];
}
model {
  for (d in 1:D) {
    mu[d] ~ normal(0, 100);
    for (l in 1:L)
      beta[l,d] ~ normal(mu[d], sigma[d]);
  }
  for (n in 1:N)
    y[n] ~ bernoulli(inv_logit(x[n] * beta[ll[n]]));
}
```