STAT415P

AUTHOR
Alejandro Gomez

STAT415 Project

Are Some Nations Stronger Than Others?

Abstract

This report analyzes the average maximum back squat, in Kilograms, for Male Power Lifters around the world. In particular, we will be focusing on official competition results of lifters from the United States, China, Russia, Poland, and United Kingdom. A Bayesian Hierarchical Linear Model, assuming random intercepts, is performed to investigate if the population average maximum back squat for male professional power lifters is the same for all 5 nations. After observing data, pairwise comparisons and other Bayesian techniques suggest the contrary. Particularly, it is estimated that China and Poland are the strongest of the nations, while the UK and USA are the weakest. While Bayesian Statistics will not always provide direct results of a distinct difference between National Means, we will use a combination of Frequentest and Bayesian analyses to do this.

Research Question

In particular, we will be investigating whether these 5 nations share the same population mean maximum back squat for Power Lifters with the criteria listed below:

Inclusion Criteria:

1. Male Lifters
2. Ages 24-34
3. 100-110 Kg body weight classes
4. No equipment used
5. Lifts from competitions no earlier than year 2000.

It is important the inclusion criteria for this analysis is implemented, specifically because of how many weight classes there are for lifters, and the strong correlation between Body Weight and the amount of weight one can lift. We want to narrow down the analysis to continue focus on a large pool of lifters while also strengthening the credibility and practical application of our results.

Data

The data of interest data contains over 3 Million observations on world wide Power Lifting competitions for a wide range of body weight classes, ages, and nations. Data is collected from 1964 to 2019. For purposes of our study, we will be focusing primarily on the Maximum Back Squat of 3 attempts per Lifter in official competitions. For context, in an official power lifting competition, lifters are given 3 attempts for back squat, and the highest score is recorded. A snap shot of the original and cleaned data set is provided below:

Original Data

▶ Code

Name	Sex	Event	Equipment	Age	AgeClass	BirthYearClass	Division	BodyweightKg	WeightClassKg
Alona Vladi	F	SBD	Raw	33.0	24-34	24-39	O	58.30	60
Galina Solovyanova	F	SBD	Raw	43.0	40-44	40-49	M1	73.10	75
Daniil Voronin	M	SBD	Raw	15.5	16-17	14-18	T	67.40	75
Aleksey Krasov	M	SBD	Raw	35.0	35-39	24-39	O	66.65	75
Margarita Pleschenkova	M	SBD	Raw	26.5	24-34	24-39	O	72.45	75

Summary of Cleaned Data

▶ Code

Country	Avg Max Squat (Kgs)
China	244.8600
Poland	238.1300
Russia	228.5750
UK	219.5250
USA	221.4815

- The cleaned data above provides the observed Average Maximum Back Squat for each of the 5 nations

Validity Conditions

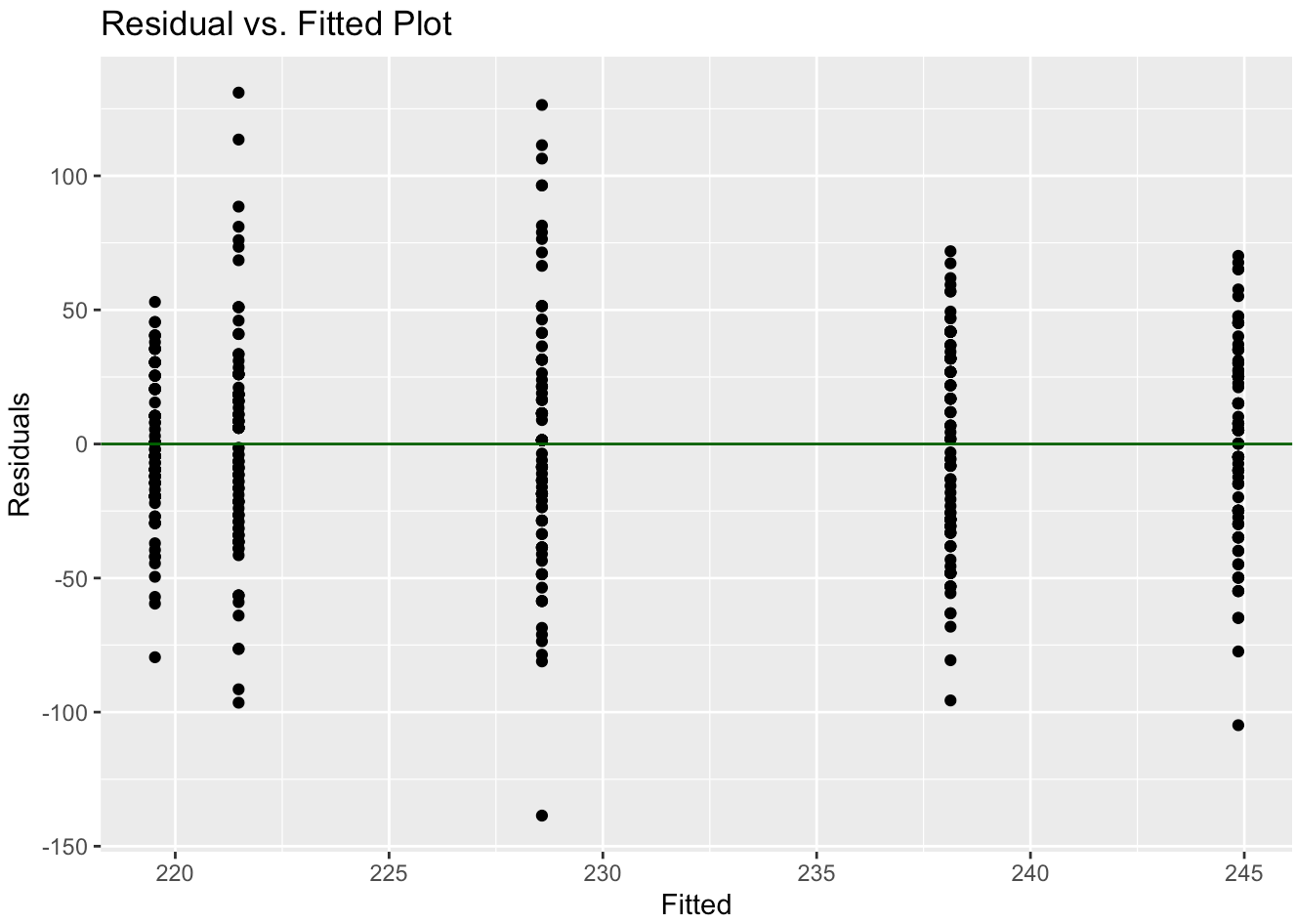
Under the implementation of any Linear Model, it is important that the Validity Conditions of our data are met. The visualizations provided below help verify Independence, Normality, and Equal Variance.

Independence

While this data does not come from an experiment, I have randomly selected 100 lifters from each of the five nations, from a larger pool of lifters. We know that our data contains information on unique lifters, and we will assume Independence for the purposes of our analysis.

Equal Variance

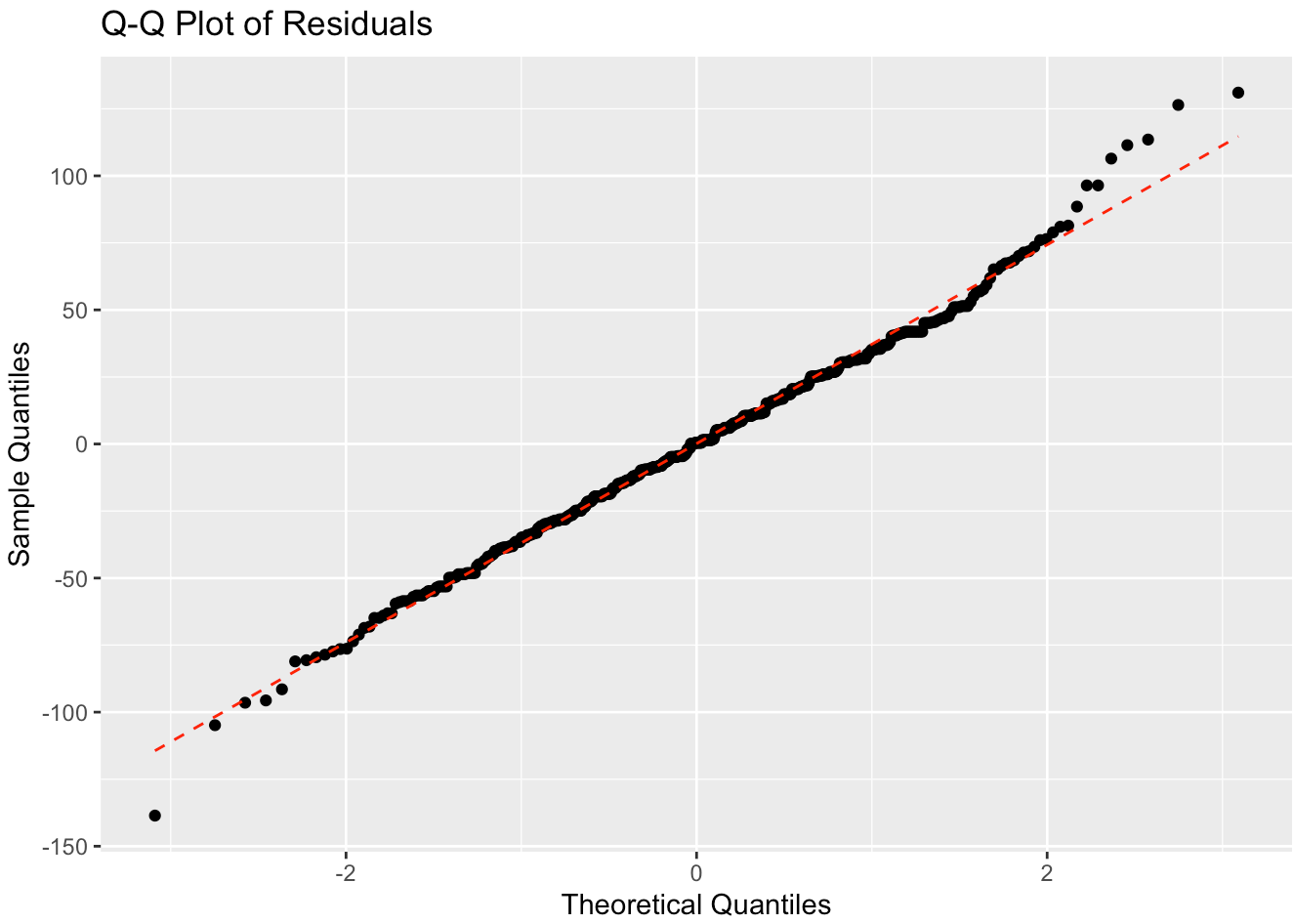
▶ Code



Based on the residuals vs. fitted plot, the vast majority of points fall in between expected residual bands/distances, so we can assume equal variance.

Normality

▶ Code



After observation of the QQPlot, all points follow very strong linear path, so normality may be assumed.

Prior Assumptions

Our analysis will implement the use of a Bayesian Hierarchical linear model.

In doing so, we will assume each individual Max Back Squat for Power Lifters in the same Nation follow a normal distribution. Each within-nation observation is also assumed to be independent of each other.

$y_{ij} \sim N(\mu_j, \sigma^2)$

Where...

y_{ij} : The Max Back Squat of Power Lifter i from Nation j

μ_j : Population Mean Max Back Squat for Nation j

σ : Population within-nation variability in squat performance - the expected deviation of a Power Lifter's Max Back Squat from the national average

We will also assume the True National Average Max Back Squats are independently drawn from a common population distribution. However, in the posterior, these estimates may be correlated due to shared information across nations.

$\mu_j \sim N(\mu, \tau^2)$

Where...

μ : Overall Population Average Maximum Back Squat for all 5 nations

τ: Population between-nation variability in squat performance - the expected deviation of a Nation's population mean back squat from the International Average (of the 5 nations).

Prior Distributions

Prior Distribution for the Internation Population Mean Back Squat (across all 5 Nations):

$\mu \sim N(mean = 220, sd = 20)$

- Based on my Prior knowledge of Power Lifting, I would expect the overall mean max squat for the 5 nations to be around 220 Kilograms with a prior uncertainty of about 20 Kilograms.

Prior Distribution for Within Nation Variability:

$\sigma \sim N(mean = 15, sd = 8)$

- Based on my Prior knowledge, I expect the Population Within-Nation variation in max back squats for the five nations to follow a normal distribution with mean 15 and a prior uncertainty of about 8 Kilograms.

Prior Distribution for Between Nation Variability:

$\tau \sim N(mean = 0, sd = 15)$

- Based on my prior knowledge, I expect the Population Between-Nation variation in Population Average Max Back Squats to follow a normal distribution with mean 0 and standard deviation 15. This distribution reflects my prior hypothesis that the Nations' population means do not differ from one another.

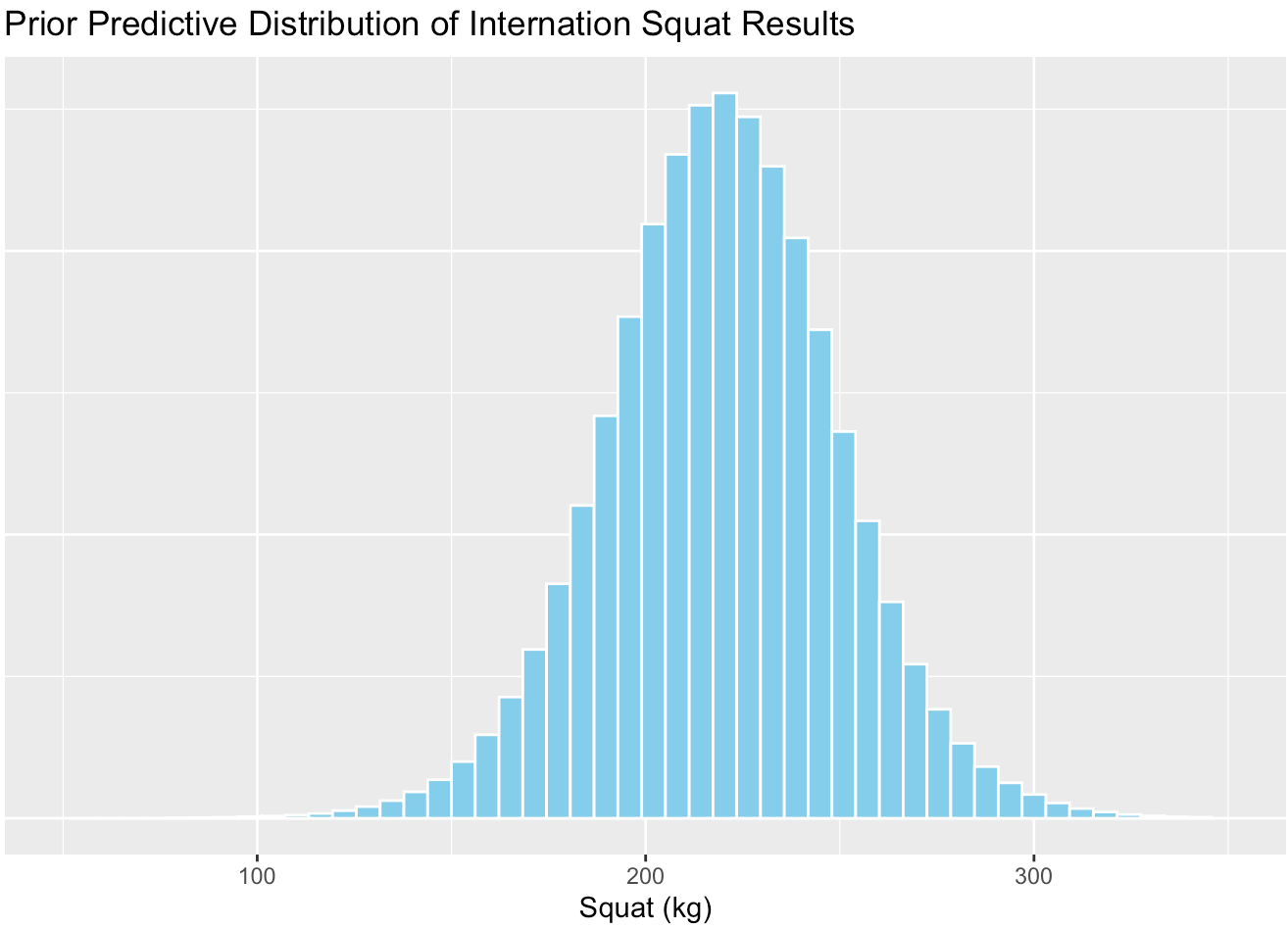
Prior Predictive Tuning

Before jumping in the the analysis, it is important to ensure that our Prior Distributions provide reasonable results that won't interfere with our procedure. For these reasons, we will conduct Prior Predictive checks to ensure a reliable analysis.

- In this case, we will use brms to simulate many simulated Maximum Back Squat results across all 5 Nations, using only our prior distributions and assumptions. We will then analyze the resulting Prior Predictive Distribution.

▶ Code

▶ Code



▶ Code

Mean	Median	Std.Dev
219.6664	219.7543	30.22689

▶ Code

```
      1%       10%       25%       50%       75%       90%       99% 
145.1476 182.1249 200.3778 219.7543 239.1231 257.0588 293.2889 
```

The Prior Predictive Distribution follows a closely normal distribution with a mean of about 220 KGs and a loose standard deviation of about 30KGs. I believe this to provide strong evidence of reliable prior distributions because there contains no unreasonable squat values that I wouldn't expect from a Professional Power Lifter.

98% Credible Interval: Before observing data it is 49 times more likely that the True Population Mean Back Squat for Professional Male Power Lifters in these 5 nations under the specified inclusion criteria is between 145.148 and 293.289, than not.

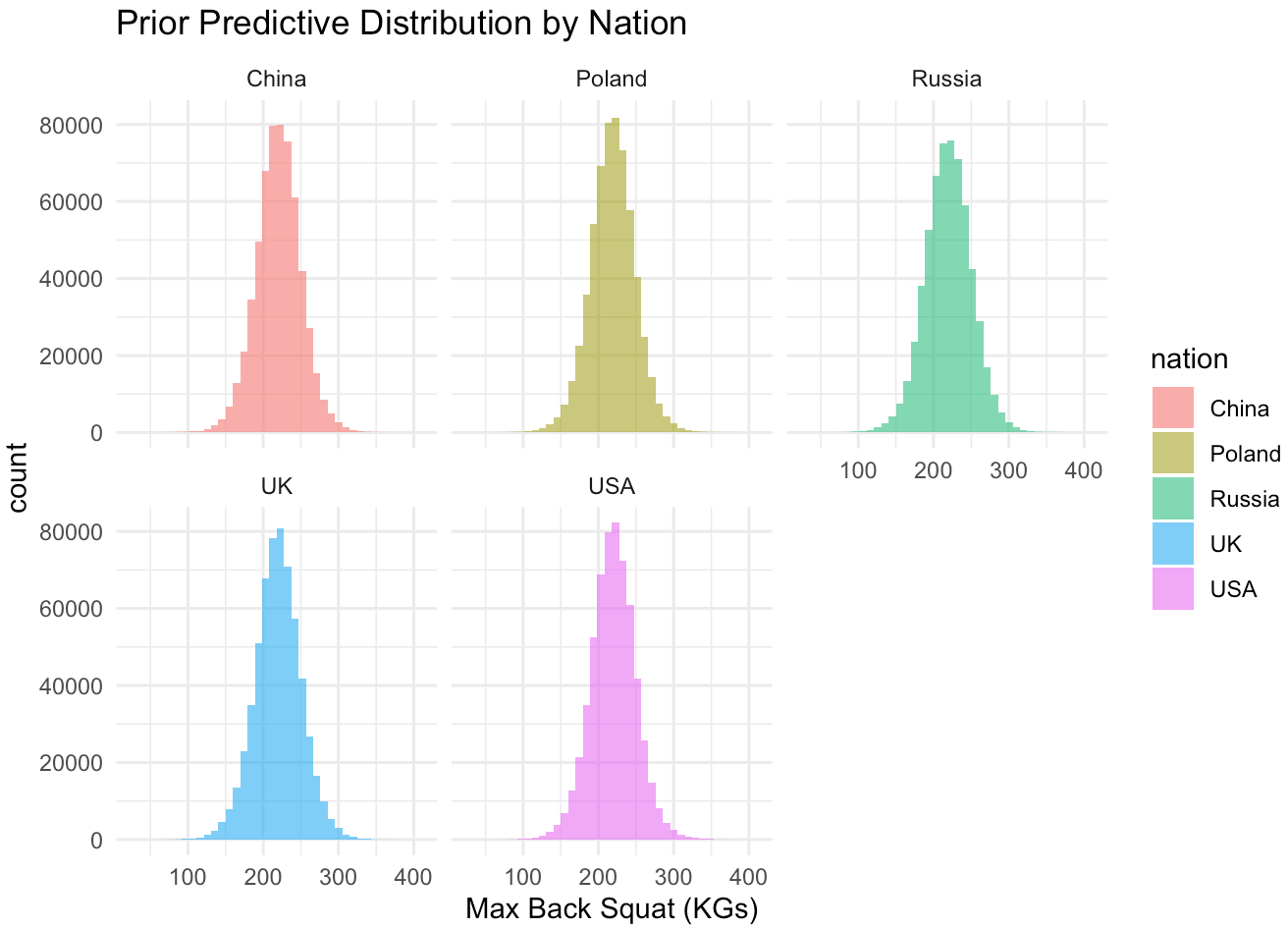
80% CI: Before observing data it is 4 times more likely that the True Population Mean Back Squat for Professional Male Power Lifters in these 5 nations under the specified inclusion criteria is between 182.125 and 257.059, than not.

50% CI: Before observing data it is equally likely that the True Population Mean Back Squat for Professional Male Power Lifters in these 5 nations under the specified inclusion criteria is between 200.378 and 239.123, than not.

National Prior Predictive Distributions

The plot below displays the Prior Predictive Distribution for each Individual Nation. All distributions are very similar to each other, and show no signs of significant, practically unrealistic outliers. This visualization only provides more evidence of sufficient prior distributions.

▶ Code



Posterior Distribution

After approval of our prior distributions, we will now move to Posterior Inference. In our model, we will be treating the intercept as random (A Frequentest interpretation would be treating Nation as a Random Effect).

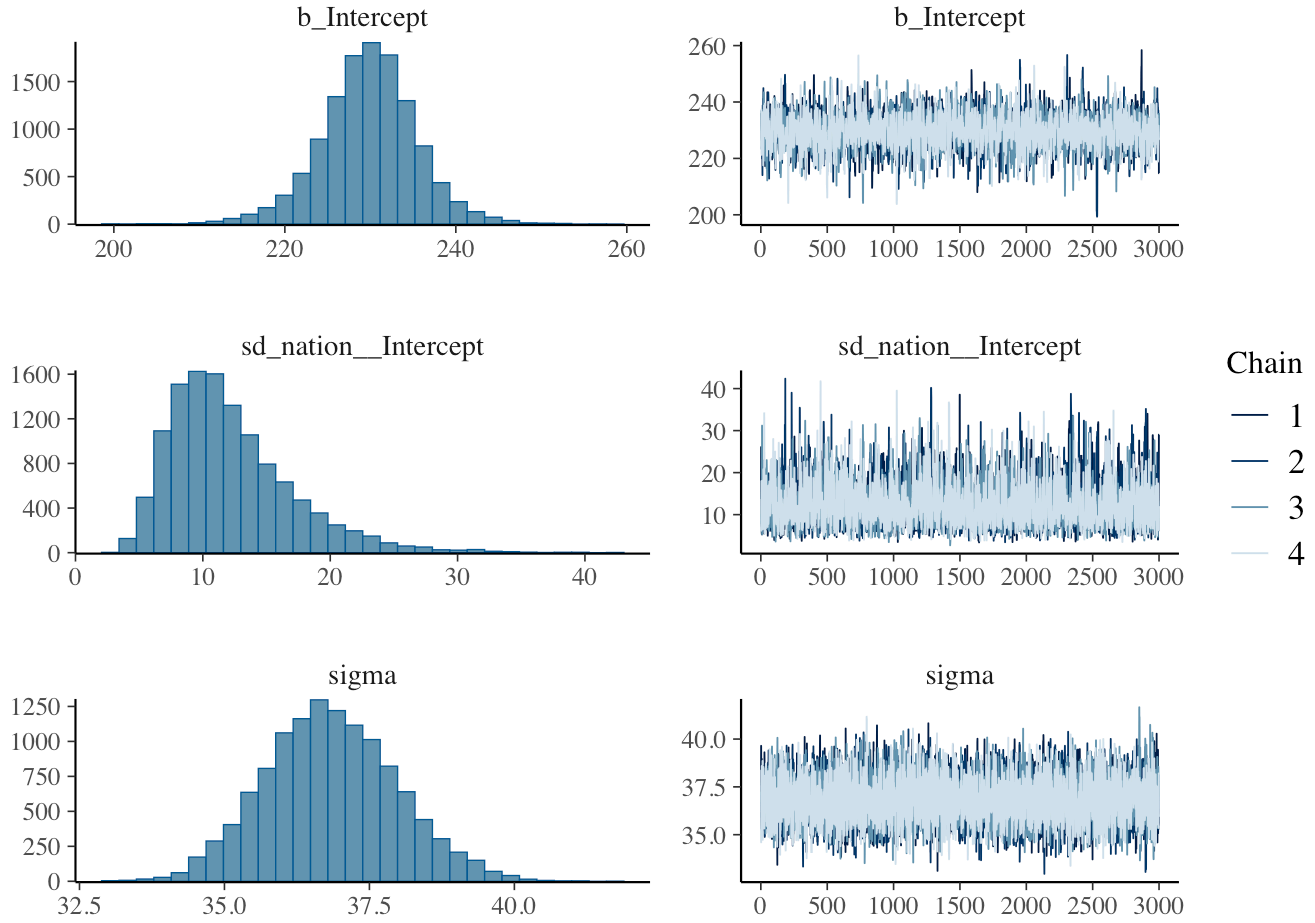
▶ Code

▶ Code

```
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: squat ~ 1 + (1 | nation)
   Data: my_lifts (Number of observations: 500)
  Draws: 4 chains, each with iter = 4000; warmup = 1000; thin = 1;
         total post-warmup draws = 12000

Multilevel Hyperparameters:
~nation (Number of levels: 5)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)    12.16      4.89     5.38    24.17 1.00     3074     5292

Regression Coefficients:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept   229.81      5.66   218.09   241.02 1.00     2985     3993

Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma    36.84      1.13    34.70    39.17 1.00     9721     7737

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

▶ Code

▶ Code



μ = 229.807

- After observing data, we estimate the true International mean max back squat for Power Lifters in all 5 nations to be about 229.807 Kilograms.

σ = 36.836

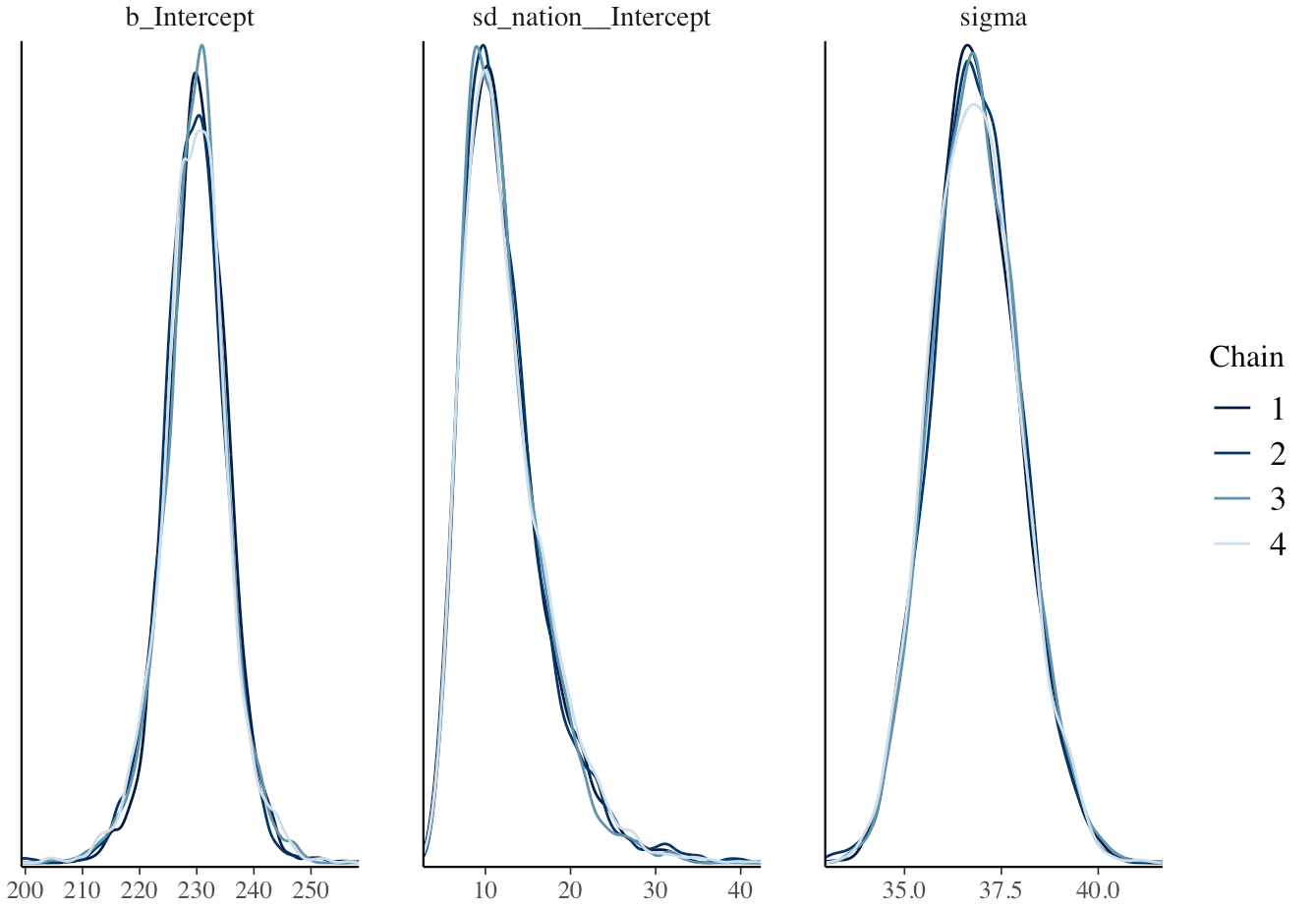
- After observing data, we estimate the typical deviation of a Power Lifter's max back squat from his nation's population mean max back squat to be about 36.836 Kilograms.

τ = 12.164

- After observing data, we estimate the typical deviation of a nation's true mean max back squat from the International national mean max back squat to be about 12.164 Kilograms.

The plot below measures how well the MCMC sampling worked. Because the chains follow each other relatively well, we can conclude that we have obtained quality posterior data.

▶ Code

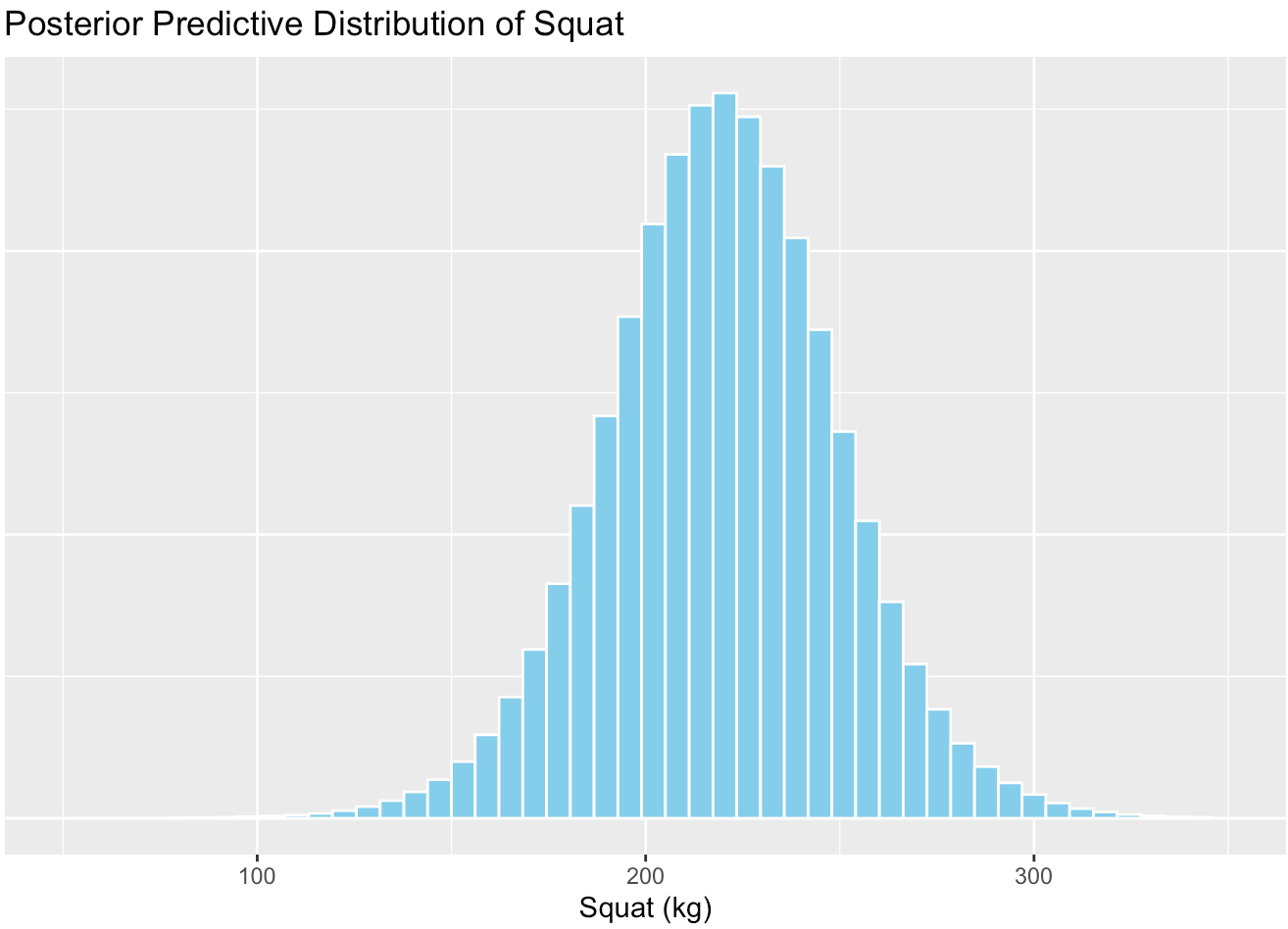


Posterior Prediction and Checks

To ensure our model assumptions suit our posterior data and validate our inference, we will also conduct some Posterior predictive checks. The following displays the Posterior predictive Distribution of Max Squats across all nations.

▶ Code

▶ Code



▶ Code

Mean	Median	Std.Dev
230.4483	230.4246	38.01434

▶ Code

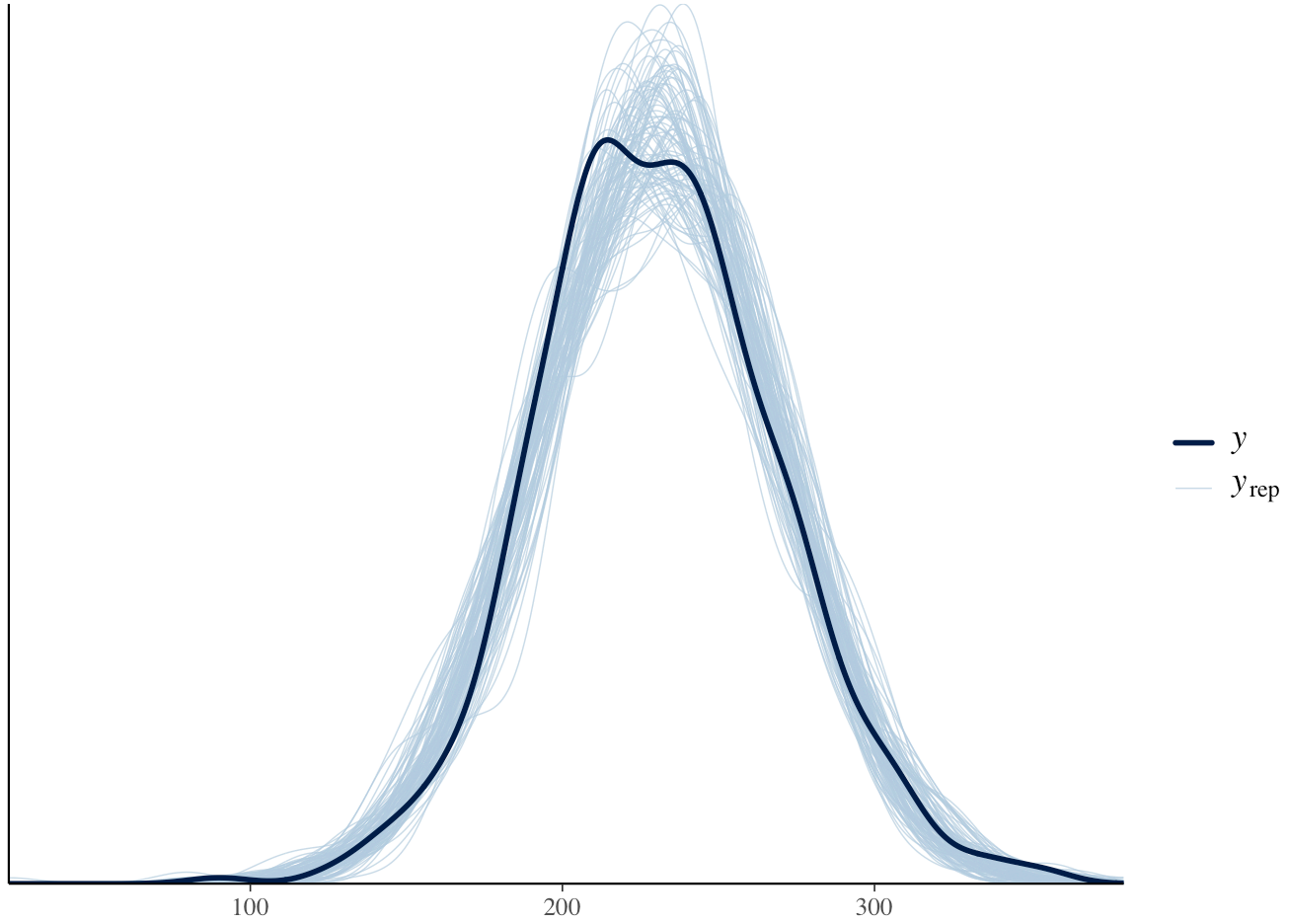
The Posterior Predictive Distribution follows a closely normal distribution with a mean of about 231 KGs and a loose standard deviation of about 38KGs.

98% Credible Interval: After observing data it is 49 times more likely that the True Population Mean Back Squat for Professional Male Power Lifters in these 5 nations under the specified inclusion criteria is between 142.02 and 318.966, than not.

80% CI: Before observing data it is 4 times more likely that the True Population Mean Back Squat for Professional Male Power Lifters in these 5 nations under the specified inclusion criteria is between 181.766 and 279.158, than not.

50% CI: Before observing data it is equally likely that the True Population Mean Back Squat for Professional Male Power Lifters in these 5 nations under the specified inclusion criteria is between 204.817 and 256.083, than not.

▶ Code



Based on the plot above, our model seems to fit our data pretty well. There is one small deviation from the cloud of simulations, but I do not believe it is enough to completely invalidate our model. For this reason, we can assume that our model fits our data, and our results are reasonable and relatively reliable.

Comparing Nations

Posterior Fixed Effects Table

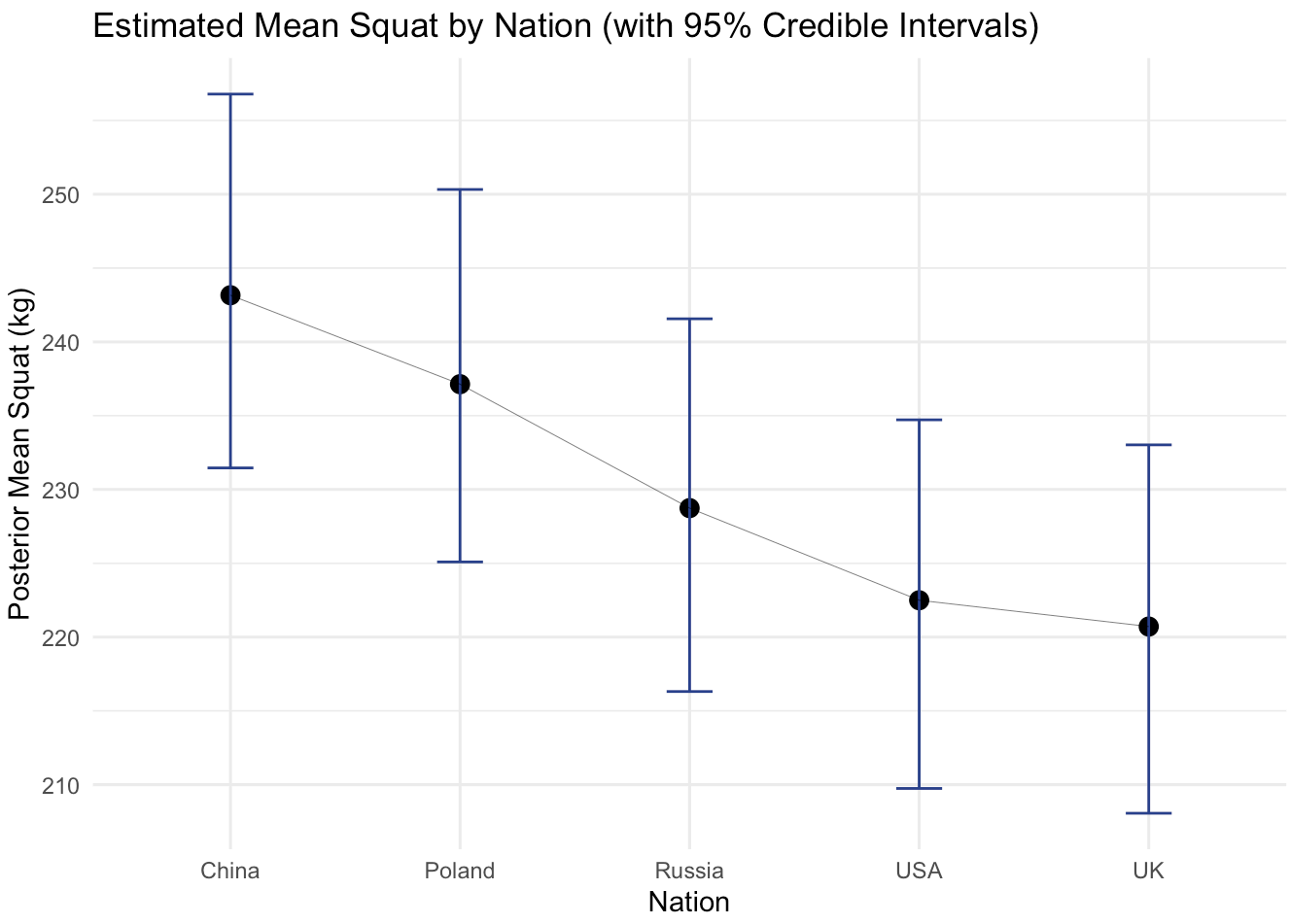
The table below displays the posterior effect of each Nation on the True International Average Max Back Squat, followed by their corresponding Standard Errors, and 95% Credible Intervals.

▶ Code

Fixed Effects	Estimate	Std. Error	Lwr 95% CI	Upr 95% CI
China	13.352061	6.389758	1.652266	26.978850
Poland	7.324779	6.353447	-4.717270	20.516491
Russia	-1.070916	6.295518	-13.496011	11.754421
UK	-9.094997	6.306652	-21.743582	3.214214
USA	-7.313644	6.245168	-20.062717	4.910480

To visualize the effects table above, this plot below displays the Posterior Mean Max Back Squat for each Nation, with each of their corresponding 95% Credible Intervals.

▶ Code



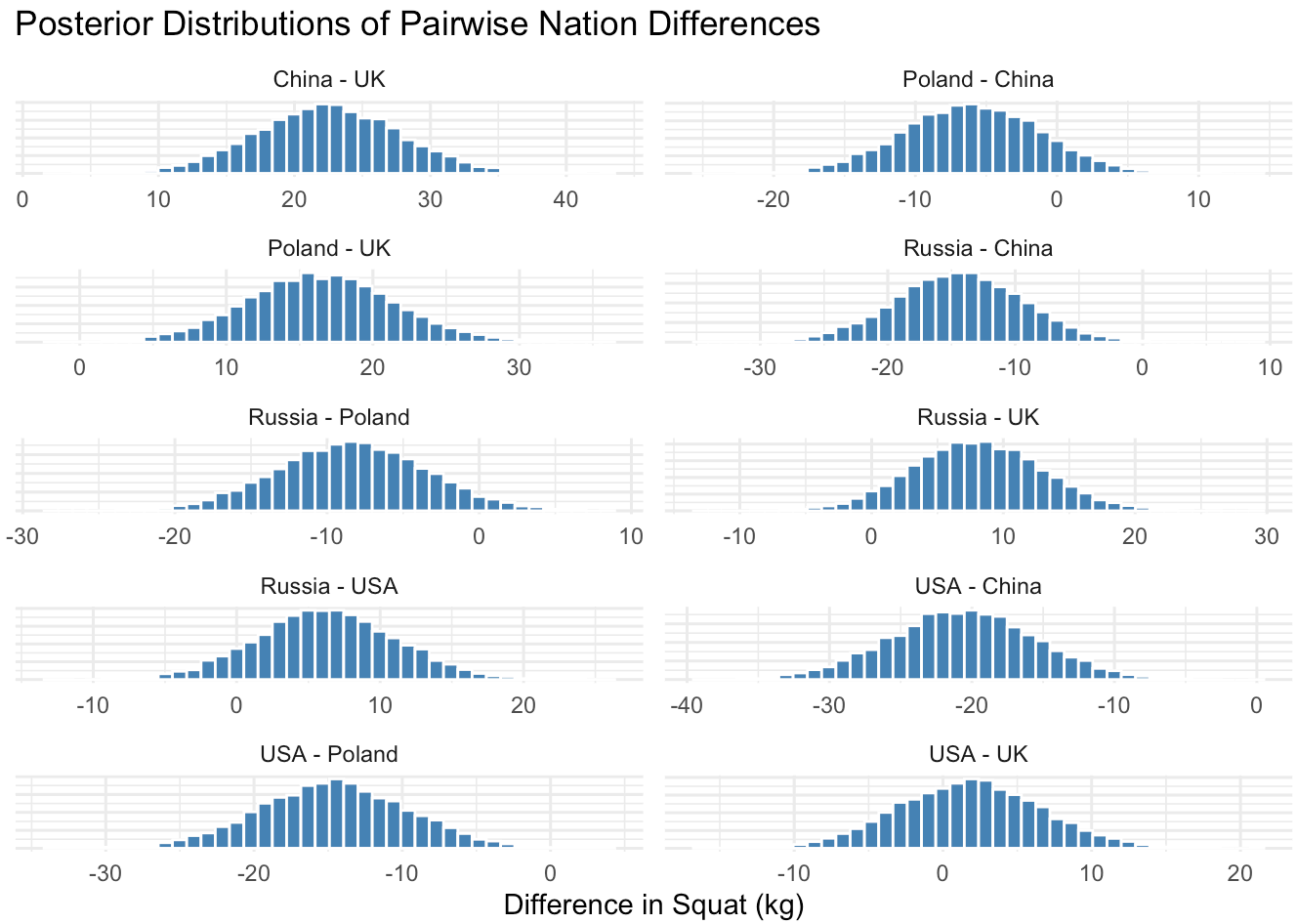
▶ Code

▶ Code

▶ Code

The grid of curves below represents the Posterior Distributions for the differences in True National Mean Max Back Squat between pairs of 2 Nations.

▶ Code



Using what we know from these tables and plots, it is reasonable enough to say that the True National Mean Max back Squats between the 5 nations are likely not the same. This is supported by the varying largely effects estimates (both positive and negative values) and the many Posterior Difference Distributions that do not contain 0, like USA and Poland.

In this investigation, we will assume a region of practical significance to be $\pm$ 2.5 KGs, meaning we define two Nations to have distinctly different National Mean Max Back Squats if their means differ by more than 2.5 KGs.

The table below displays the following Information:

Contrast: Specific Posterior Difference Distribtuion between 2 Nations

Mean_Diff: Pairwise Difference in Average Max Back Squat between 2 nations

Lower/Upper_95: 95% Credible Interval measuring Posterior Probability that the True Difference in Average Squat between 2 Nations is in the specified interval

Prob_in_Rope: The Posterior probability that the True Differnce in average Squat between 2 Nations is less than $\pm$ 2.5 Kgs.

Decision: Using Prob_in_Rope, it is decided whether the 2 Nations' Population Average Max Back Squats are distinctly different from each other.

▶ Code

contrast	mean_diff	lower_95	upper_95	prob_in_rope	decision
Russia - USA	6.243	-3.312	15.935	0.184	Unclear
Russia - China	-14.423	-24.384	-4.527	0.010	Second > First
Russia - Poland	-8.396	-18.016	1.248	0.103	Second > First
Russia - UK	8.024	-1.533	17.914	0.113	Unclear
USA - China	-20.666	-30.839	-10.502	0.000	Second > First
USA - Poland	-14.638	-24.594	-4.751	0.007	Second > First
USA - UK	1.781	-7.674	11.484	0.366	Unclear
Poland - China	-6.027	-15.689	3.470	0.199	Unclear
Poland - UK	16.420	6.442	26.651	0.003	First > Second
China - UK	22.447	12.061	32.824	0.000	First > Second

The letters table below summarizes the table above, helping us understand which Nations are estimated to have practically different National Average Max Back Squats from each other, after observing data.

▶ Code

Nation	Letters
Poland	A
China	AB
Russia	__C
USA	__CD
UK	___D

In conclusion, observing data revealed that Poland and China have the largest True National Average Max Back Squats, standing out from the other 3 Nations. With a region of practical significance of $\pm$ 2.5 KGs, we are unable to differentiate the results of Poland and China from one another. On the other end, observing data also revealed that the United Kingdom had the lowest National Average Max Back Squat, which is also statistically indistinguishable from the United States.

Sensitivity Analysis

Now we will be analyzing the how sensitive our Posterior Data is to different Prior Distributions. To understand this, we will lets brms provide the prior distributions and conduct the analysis.

▶ Code

▶ Code

```
Warning: There were 6 divergent transitions after warmup. Increasing
adapt_delta above 0.8 may help. See
http://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup

 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: squat ~ 1 + (1 | nation)
   Data: my_lifts (Number of observations: 500)
  Draws: 4 chains, each with iter = 4000; warmup = 1000; thin = 1;
         total post-warmup draws = 12000

Multilevel Hyperparameters:
~nation (Number of levels: 5)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)    14.39      7.40     5.69    33.26 1.00     1983     2490

Regression Coefficients:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept   230.17      6.91   215.17   243.66 1.00     2455     2076

Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma    37.24      1.17    35.00    39.60 1.00     7891     6692

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

▶ Code

μ estimate: 230.173

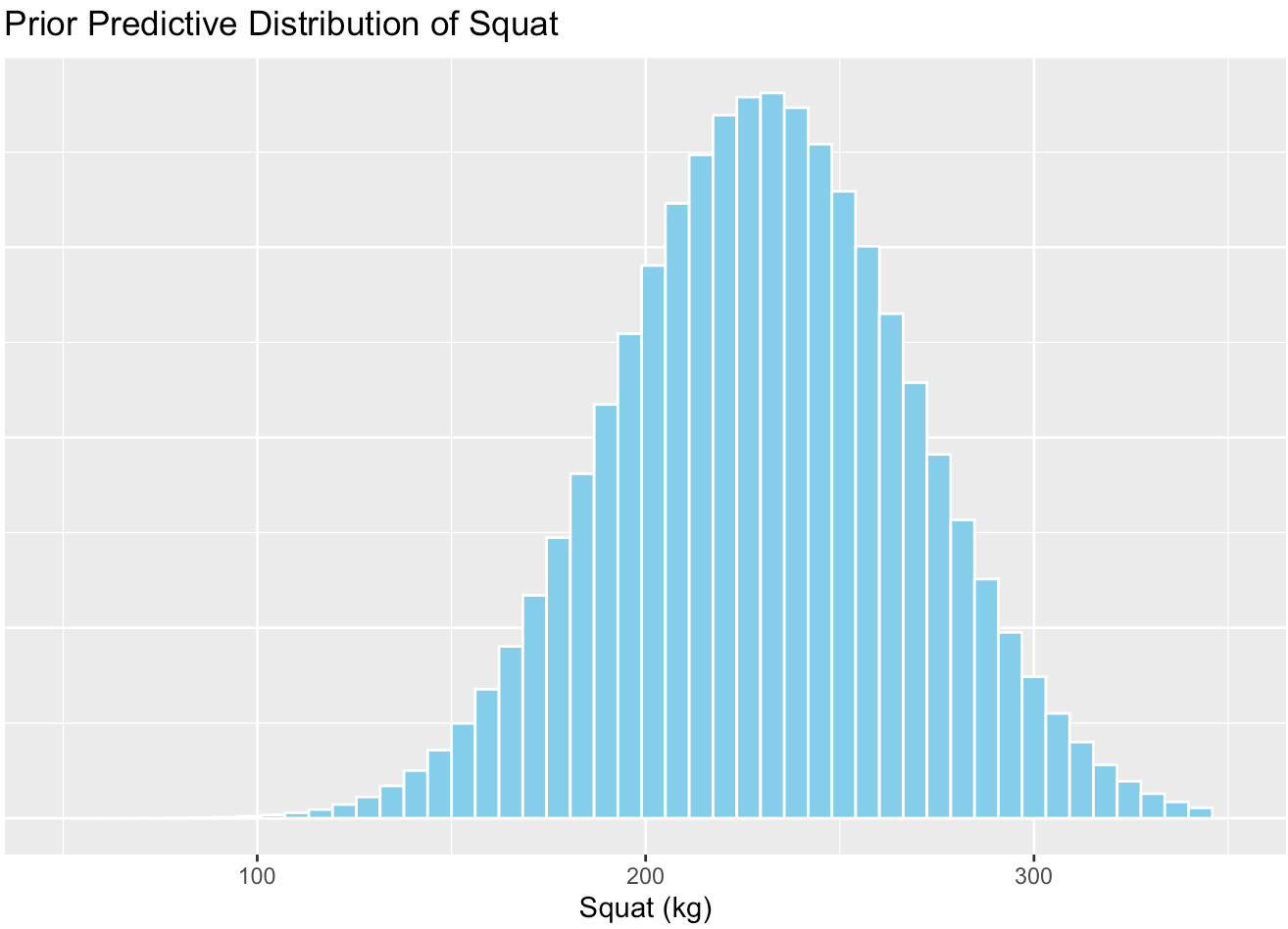
σ estimate: 37.244

τ estimate: 14.389

- These inference results are not exact, but are very similar to our original analysis when we chose our own prior distributions.

▶ Code

▶ Code



▶ Code

Mean	Median	Std.Dev
230.5174	230.4931	38.41984

▶ Code

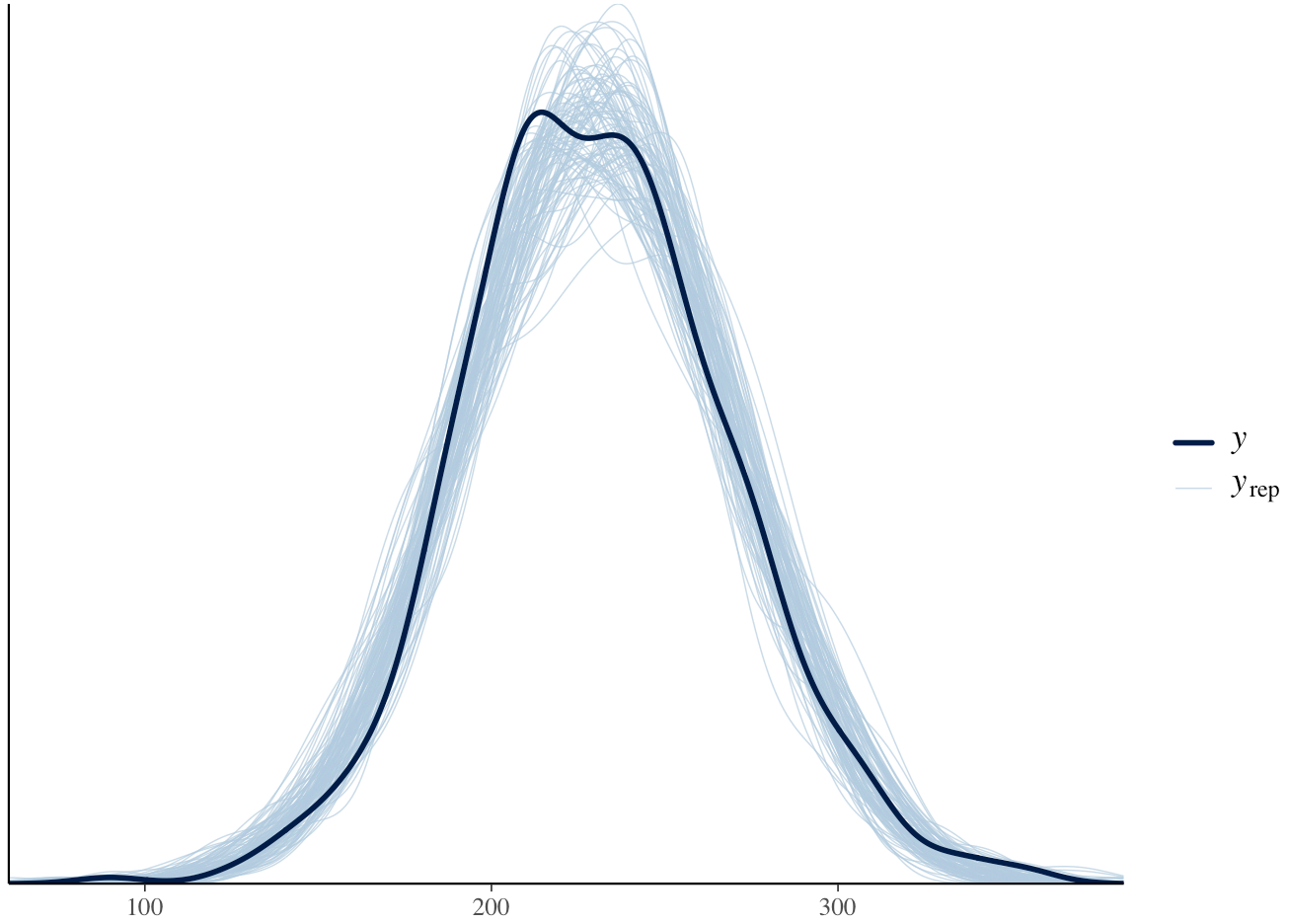
The Posterior Predictive Distribution follows a closely normal distribution with a mean of about 231 KGs and a loose standard deviation of about 39KGs. This distribution is very identical to our previous posterior predictive distribution.

98% Credible Interval: After observing data it is 49 times more likely that the True Population Mean Back Squat for Professional Male Power Lifters in these 5 nations under the specified inclusion criteria is between 141.173 and 320.106, than not.

80% CI: Before observing data it is 4 times more likely that the True Population Mean Back Squat for Professional Male Power Lifters in these 5 nations under the specified inclusion criteria is between 181.334 and 279.742, than not.

50% CI: Before observing data it is equally likely that the True Population Mean Back Squat for Professional Male Power Lifters in these 5 nations under the specified inclusion criteria is between 204.62 and 256.407, than not.

▶ Code



This visual is also almost exactly identical to the prior one, ensuring relatively reliable assumptions and data.

In Conclusion, we have reason to believe that Prior Distributions have very little influence on Posterior Data.

Frequentest One Way Anova

This analysis can also be performed under a frequentest perspective. More specifically, we can run a one way ANOVA, treating nation as a random effect.

▶ Code

```
Linear mixed model fit by REML ['lmerMod']
Formula: squat ~ (1 | nation)
   Data: my_lifts

REML criterion at convergence: 5039.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
```

```
-3.7313 -0.6301 -0.0220  0.6582  3.4934

Random effects:
 Groups   Name        Variance Std.Dev.
 nation   (Intercept)  103.6   10.18
 Residual             1383.8   37.20
Number of obs: 500, groups:  nation, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)  230.514      4.847   47.56
```

μ: 230.514

σ: 37.2

τ: 10.18

- These results are very similar to our Bayesian Analysis and provide identical conlcusions.

Pairwise Comparisons

▶ Code

	(Intercept)
China	12.655880
Poland	6.718625
Russia	-1.710864
UK	-9.694840
USA	-7.968801

- China and Poland Are the strongest with UK and USA being the weakest. These results are extremely identical under our Bayesian analysis.

Conclusions

Reflecting on our analysis, we can summarize our conclusions into 3 sections:

International Conclusions

- After observing data, it is estimated that the Nationwide Population Average Maximum Back Squat for Power Lifters with the specified criteria to be 229.807 Kilograms, with the 5 Nations varying from the Nationwide Mean by around 36.836 Kilograms.
- It is also estimated that within each Nation, individual Maximum Back Squats vary from the National Average Maximum by around 12.164 Kilograms.

- We also have evidence to conclude that the National Population Average Maximum Back Squat for each of the 5 nations are NOT the same.

Nation-Specific Conclusions

- After observing data, Poland and China are the Strongest Back Squatting nations of the 5, as their respective estimated Population Average Maximum Back Squats are about 244 and 248 Kilograms. After analysis of posterior differences and implementing a region of practical significance of $\pm$ 2.5 Kilograms, it cannot be distinguished if either Poland or China is stronger than the other. On the other hand, our analysis suggests that the United Kingdom is the weakest of the 5 nations, and is statistically indistinguishable from the United States. Russia takes home the middle of the 5 nations, but is also statistically indistinguishable from the United States.

Model Specific Conclusions

- Under a Bayesian Hierarchical Linear Model, the choice of prior distribution does not have a significant impact on the outcome of the posterior data.
- Our model will result in comparable results to the Frequentest One-Way ANOVA, treating Nation as a random effect.