

# Predicting Heart Disease

Michael Davies (mld9s) and Akeem Wells (ajw3rg)

University of Virginia

DS6040

Henry, Teague Rhine

August 13 2021

## Table Of Contents

<b>Abstract</b>	<b>1</b>
<b>Problem Statement</b>	<b>1</b>
Data . . . . .	2
Imputing . . . . .	6
Scaling: Z-Score vs. Mean Absolute Scaling . . . . .	7
Z-Score . . . . .	7
Maximum Absolute Scaling . . . . .	7
<b>Approach</b>	<b>8</b>
Bayesian Logistic Model . . . . .	9
Bayesian Model Averaging . . . . .	9
Bayesian Hierarchical Model . . . . .	9
<b>Prior Rationale</b>	<b>10</b>
<b>Findings</b>	<b>10</b>
Bayesian Logistic Model . . . . .	10
Bayesian Model Averaging . . . . .	13
Bayesian Hierarchical Model . . . . .	14
Summary . . . . .	18
<b>Additional Resources</b>	<b>19</b>
<b>Supplemental Images</b>	<b>20</b>
<b>Supplemental Tables</b>	<b>33</b>

## Figures

1	Pair plot of continuous features by country . . . . .	4
2	Pair plot of continuous features by "target" (response) . . . . .	5
3	Correlation Matrix of continuous features . . . . .	6
4	Bayesian Logistic Model Trace . . . . .	11
5	Bayesian Hierarchical Model Trace . . . . .	16
6	Age Histogram . . . . .	20
7	Ca Histogram . . . . .	21
8	Chol Histogram . . . . .	21
9	Country Histogram . . . . .	22
10	Cp Histogram . . . . .	22
11	Exang Histogram . . . . .	23
12	Fbs Histogram . . . . .	23
13	Oldpeak Histogram . . . . .	24
14	Restecg Histogram . . . . .	24
15	Sex Histogram . . . . .	25
16	Slope Histogram . . . . .	25
17	Target Histogram . . . . .	26
18	Thal Histogram . . . . .	26
19	Thalach Histogram . . . . .	27
20	Trestbps Histogram . . . . .	27
21	Trace Plot Of Logistic Model . . . . .	28
22	Trace Plot Of Mus Hierarchical Model . . . . .	29
23	Trace Plot Of Sigmas Hierarchical Model . . . . .	30
24	Trace Plot Of Offsets Hierarchical Model . . . . .	31
25	Trace Plot Of Marginals Hierarchical Model . . . . .	32

## Tables

1	Summary Table of Descriptive Statistics . . . . .	3
2	Summary Table of Scaled Descriptive Statistics . . . . .	8
3	Summary Table Bayesian Logistic Model . . . . .	12
4	Summary Table statsmodels . . . . .	13
5	Summary Table Bayesian Model Averaging . . . . .	14
6	Summary Table Bayesian Hierarchical Model Mu . . . . .	17
7	Summary Table Hierarchical Offset . . . . .	33
8	Summary Table Hierarchical Sigmas . . . . .	34
9	Summary Table Hierarchical Marginals . . . . .	35

## **Abstract**

Given the risk of heart disease in modern society, detection of cardiovascular disease and identifying its risk level for adults is a critical task. Therefore, we implemented three Bayesian classification models to classify whether a patient's heart is normal or if there is the presence of heart disease. Given that all three of our models obtain consistent results, we have confidence that specific features from this data have value in predicting the presence of heart disease. Results from these models suggest that old peak (ST depression induced by exercise relative to rest), thalach (maximum heart rate achieved), and ca (number of major coronary artery calcification vessels (0-3) colored by fluoroscopy) play a large role in determining the likelihood that a patient has heart disease. We can also find some differences in likelihood of having heart disease between countries, which is an interesting observation as well as justifies the use of Hierarchical Bayesian Analysis for problems of this nature.

## **Problem Statement**

“Heart disease has become a major health problem in both developed and developing countries, and it is cited as the number one cause of death throughout the world each year.” Given the risk of heart disease in modern society, detection of cardiovascular disease and identifying its risk level for adults is a critical task. We implemented three Bayesian classification models to classify whether a patient's heart is normal or if there is the presence of heart disease. More specifically, we developed binary classification models that predict the posterior probability that an individual has heart disease (given our data and model).

## Data

The data we selected comes from [here](#) and [here](#):

- Hungarian Institute of Cardiology. Budapest
- University Hospital, Zurich, Switzerland
- University Hospital, Basel, Switzerland
- V.A. Medical Center, Long Beach and Cleveland Clinic Foundation

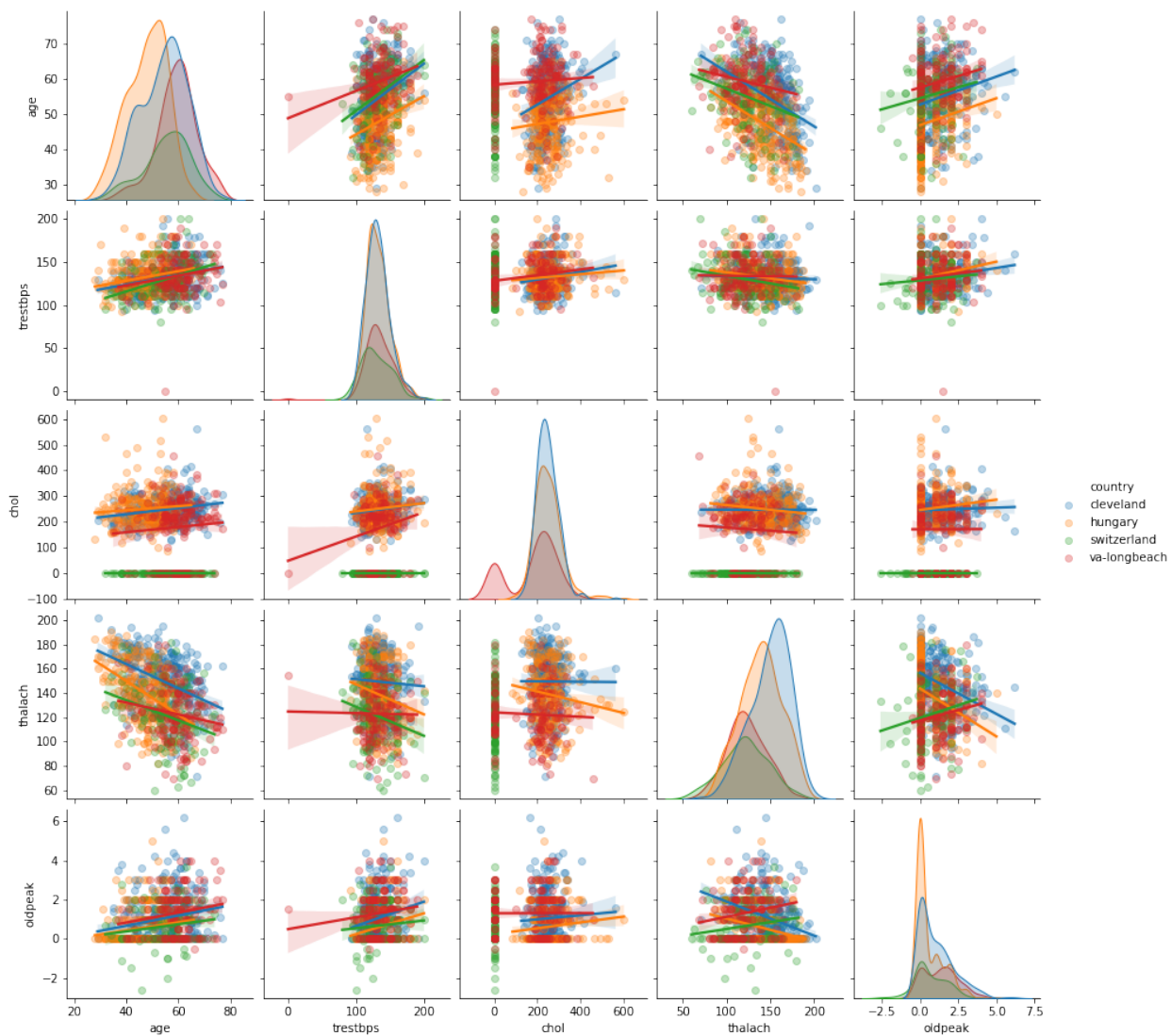
The features of interest are:

- **age** - age in years
- **sex** - 1 = male; 0 = female
- **cp** - chest pain type
- **trestbps** - resting blood pressure (in mm Hg on admission to the hospital)
- **chol** - serum cholesterol in mg/dl
- **fbs** - fasting blood sugar > 120 mg/dl (1 = true; 0 = false)
- **restecg** - resting electrocardiograph results
- **thalach** - maximum heart rate achieved
- **exang** - exercise induced angina (1 = yes; 0 = no)
- **oldpeak** - ST depression induced by exercise relative to rest
- **slope** - the slope of the peak exercise ST segment
- **ca** - number of major vessels (0-3) colored by fluoroscopy
- **thal** - 3 = normal; 6 = fixed defect; 7 = reversible defect
- **target** (the predicted attribute) - 0/1

**Table 1***Summary Table of Descriptive Statistics*

	age	trestbps	chol	thalach	oldpeak
count	920.000000	861.000000	890.000000	865.000000	858.000000
mean	53.510870	132.132404	199.130337	137.545665	0.878788
std	9.424685	19.066070	110.780810	25.926276	1.091226
min	28.000000	0.000000	0.000000	60.000000	-2.600000
25%	47.000000	120.000000	175.000000	120.000000	0.000000
50%	54.000000	130.000000	223.000000	140.000000	0.500000
75%	60.000000	140.000000	268.000000	157.000000	1.500000
max	77.000000	200.000000	603.000000	202.000000	6.200000

Below we have two pair plots of the continuous features. The first is by country to assess whether there is justification for implementing a hierarchical model with country clusters. Indeed, there appears to be some variation in the relationship between the features when controlling for country.

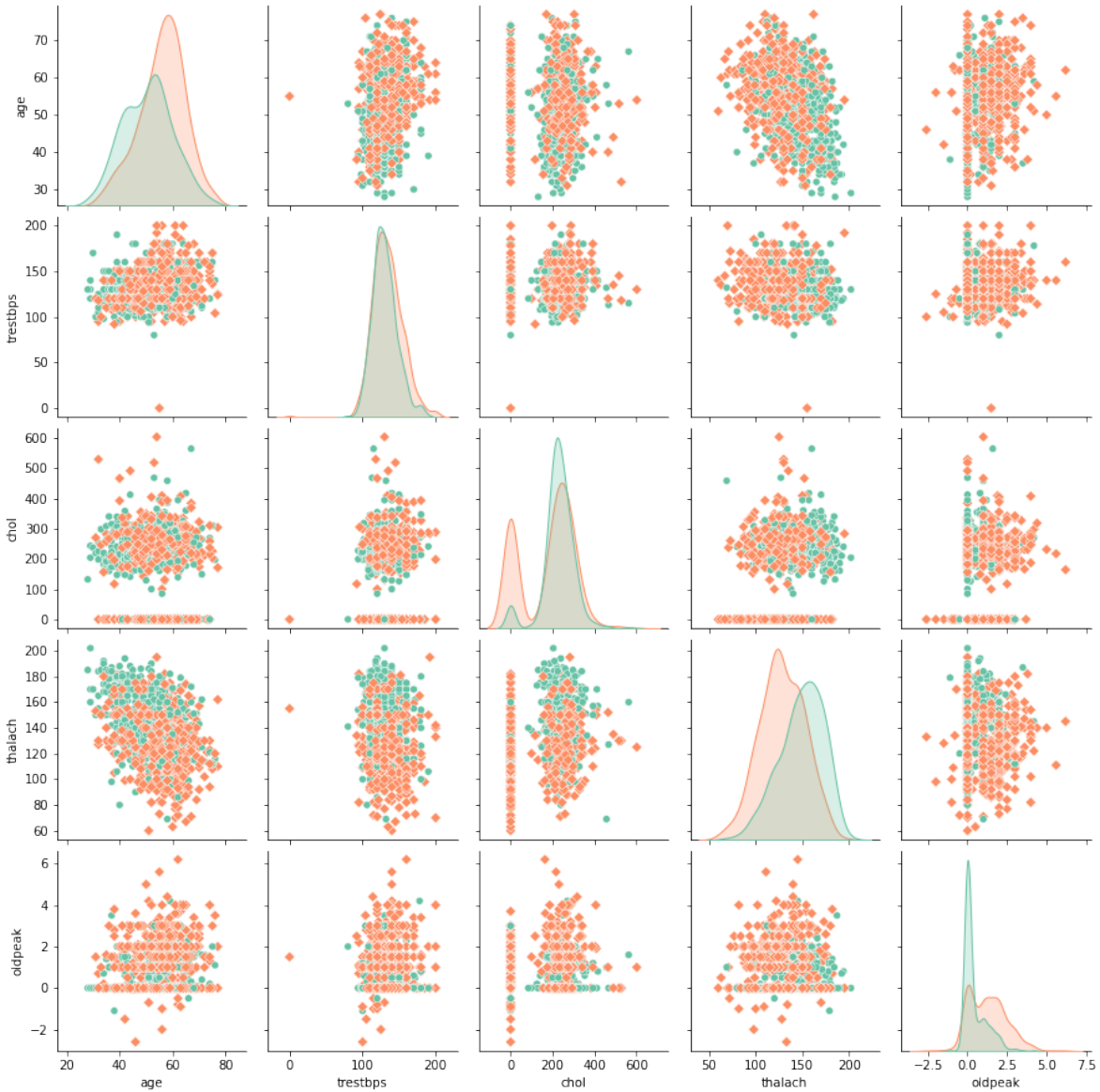


**Figure 1**

*Pair plot of continuous features by country*

The same is true when looking at the pair plot that controls for the target variable (presence or absence of heart disease).





**Figure 2**

*Pair plot of continuous features by "target" (response)*

Next we computed the pairwise correlation of for the continuous variables to assess the possibility of highly correlated variables. Notably, this does not appear to be the case.



**Figure 3**

*Correlation Matrix of continuous features*

## Imputing

During our exploratory phase, we discovered a large number of missing values. Of course, treatment of missing values is fraught with problems. We discussed the various ways missing values may occur. For instance, was it completely unsystematic (Missing Completely at Random (MCAR)), because the probability of the missing data on a variable is related to some other measured variable but unrelated to the variable with missing values itself (Missing at Random (MAR)), or because the missing values on a

variable are related to the variable with the missing values itself (Missing Not at Random (MNAR)).[\(Reference\)](#)

Ultimately, we were not able to determine the cause of the missing values. We chose to impute the missing values primarily because dropping missing values drops the entire observation. In this case, we would lose a significant amount of valuable information – particularly from the Switzerland and Long beach data. So, we implemented a K Nearest Neighbors method of imputation, which is a distance based method that replaces the missing values with the mean value from the parameter 'n' nearest neighbors. [\(Reference\)](#) and [\(Reference\)](#)

## Scaling: Z-Score vs. Mean Absolute Scaling

### *Z-Score*

$$x_{i,f} = \frac{x_{i,f} - \mu_f}{\sigma_f}$$

The data was recorded on wildly differing scales. Therefore we explored a number of methods to scale the data. For instance, we looked at z-score scaling. However, z-score scaling assumes a normal distribution with a mean of zero, and transforms all variables so that they are flipped along the axis. We were concerned we would this would distort the value of variables such as age, which is likely skewed and all positive. So, we then explored Mean Absolute Scaling.

### *Maximum Absolute Scaling*

$$x_{i,f} = \frac{x_{i,f}}{\max(|x_f|)}$$

**Table 2***Summary Table of Scaled Descriptive Statistics*

	age	trestbps	chol	thalach	oldpeak
count	920.000000	920.000000	920.000000	920.000000	920.000000
mean	0.694946	0.660533	0.333164	0.677513	0.146564
std	0.122399	0.094100	0.183888	0.127336	0.175852
min	0.363636	0.000000	0.000000	0.297030	-0.419355
25%	0.610390	0.600000	0.293532	0.589109	0.000000
50%	0.701299	0.650000	0.371476	0.688119	0.096774
75%	0.779221	0.700000	0.447761	0.772277	0.241935
max	1.000000	1.000000	1.000000	1.000000	1.000000

We chose to implement Maximum Absolute Scaling. This method re-scales each feature to be  $-1 \leq x_i \leq 1$ . Importantly, if your values are all negative, they will be re-scaled to  $-1$  to  $0$ . If all values are positive, they will be between  $0$  to  $1$ . If its a mix of both, its  $-1$  to  $1$ . This will preserve important characteristics of the features.

### Approach

Given our full model, which is a binary classification model:

$$target_i = \alpha + \beta_1 * age_i + \beta_2 * sex_i + \beta_3 * cp_i + \beta_4 * trestbps_i + \beta_5 * chol_i + \beta_6 * fbs_i + \beta_7 * restecg_i + \beta_8 * thalach_i + \beta_9 * exang_i + \beta_{10} * oldpeak_i + \beta_{11} * slope_i + \beta_{12} * ca_i + \beta_{13} * thal_i$$

Using the form:

$$p(M_i|X, y) = \frac{p(X, y|M_i)p(M_i)}{p(X, y)} \text{ for model } M_i$$

We implemented three Bayesian methods.

- Bayesian Logistic Model

- Bayesian Model Averaging
- Bayesian Hierarchical Model

### Bayesian Logistic Model

The model assumes a Bernoulli distribution with parameters  $p$ , which is the probability of a patient having heart disease. Bayesian Logistic Regression is relative fast (computationally light) and generally effective. So, we implement it here to serve as a baseline model to which we can compare other models.

### Bayesian Model Averaging

Bayesian Model Averaging (BMA) allows a parameter estimate to be achieved by averaging the estimates of different models with each estimate weighted by its probability. In addition, the BMA allowed us to see which variables have higher probability in determining the likelihood of the presence or absence of heart disease.

### Bayesian Hierarchical Model

Our problem statement suggests that model parameters are allowed to vary by group (or country in this case), which implies a hierarchical model is likely appropriate. A hierarchical model is a particular multilevel model where parameters are nested within one another. A strength of multilevel models is the ability to handle predictors on multiple levels simultaneously. So, in this case, rather than fitting separate regression models for each country, multilevel modeling shares strength among countries, allowing for more reasonable inference in countries. ([Reference](#)) A general expression of the model, where "i" is the particular country, is:

$$y_i = \alpha_j[i] + \beta_{x_i} + \epsilon_i \text{ for model } M_i$$

## Prior Rationale

Since we don't have a strong assumption as to the baseline distribution of the parameters, I will leverage the given weakly-informative priors.

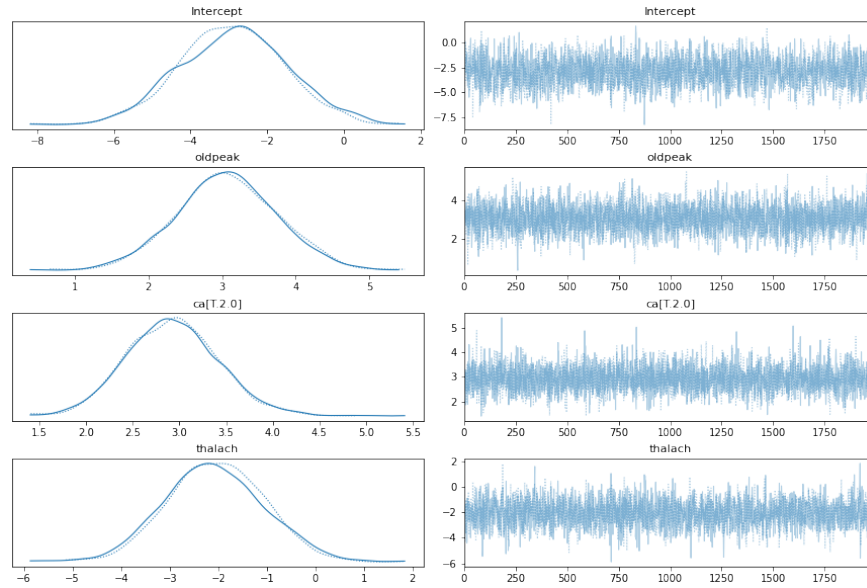
For our priors, we used normal priors from a normal distribution ( $\mu \sim N(\mu_0, \sigma_0^2)$ ) for the regression coefficients, and retained the Half Cauchys for the variance distributions. We can use the Half-Cauchy distribution when the data has a parameter  $\mu$ , where  $\mu = 0$ , positive scale parameter  $\sigma \geq 1$  and values at the peak and to the right have nonzero probability density. Our choice of link function was a Bernoulli distribution since estimating log odds we want a value that ranges 1 with probability  $p$  and  $q$  where  $q = 1 - p$ .

## Findings

**Note:** The Supplemental Images section at the end of this paper contains a comprehensive list of diagnostic plots. We summarize the finding here.

### Bayesian Logistic Model

For our Baseline model which was a Bayesian Logistic Model, we used `pm.glm.families.Binomial()` since our target was a 0/1 binary response. From this model, we can find an initial estimate of which variables play a role in predicting the posterior probability of having heart disease. Below, we display plots for just three of the variables that had the largest magnitude of predictive value. 'oldpeak' and 'ca' show a strong positive association and 'thalach' shows a negative association with the presence of heart disease.



**Figure 4**

*Bayesian Logistic Model Trace*

From the table below, the HDI as the summary credible interval for the posterior distribution. All the values inside the HDI have higher probability density (i.e., credibility) than any value outside the HDI. The HDI therefore includes the most credible values.

[\(Reference\)](#)

**Table 3***Summary Table Bayesian Logistic Model*

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
oldpeak	3.069	0.680	1.856	4.411	0.012	0.008	3424.0	2780.0	1.0
ca[T.2.0]	2.928	0.495	1.974	3.811	0.008	0.006	4051.0	2912.0	1.0
Intercept	-2.899	1.374	-5.681	-0.465	0.030	0.022	2069.0	2320.0	1.0
thalach	-2.066	1.001	-3.926	-0.254	0.020	0.014	2600.0	2770.0	1.0
ca[T.1.0]	2.017	0.340	1.388	2.653	0.006	0.004	3606.0	3103.0	1.0
ca[T.3.0]	1.877	0.780	0.416	3.284	0.013	0.011	3644.0	2468.0	1.0
trestbps	1.285	1.073	-0.637	3.347	0.017	0.013	4176.0	3002.0	1.0
cp[T.4.0]	1.253	0.399	0.512	2.020	0.008	0.006	2204.0	2635.0	1.0
sex[T.1.0]	1.183	0.275	0.678	1.688	0.005	0.003	3748.0	2826.0	1.0
exang[T.1.0]	1.107	0.231	0.690	1.553	0.004	0.003	4280.0	3175.0	1.0
cp[T.2.0]	-0.889	0.454	-1.772	-0.054	0.009	0.006	2574.0	2960.0	1.0
chol	-0.717	0.676	-2.027	0.530	0.011	0.009	3627.0	2918.0	1.0
fbs[T.1.0]	0.711	0.288	0.203	1.254	0.005	0.003	4021.0	2990.0	1.0
slope[T.2.0]	0.686	0.248	0.244	1.157	0.004	0.003	3113.0	2786.0	1.0
thal[T.7.0]	0.655	0.252	0.188	1.132	0.005	0.003	3104.0	2843.0	1.0
thal[T.6.0]	0.568	0.314	-0.042	1.142	0.005	0.004	3337.0	2832.0	1.0
age	0.281	0.940	-1.421	1.980	0.016	0.013	3326.0	2912.0	1.0
cp[T.3.0]	-0.232	0.411	-1.045	0.494	0.009	0.006	2292.0	2590.0	1.0
slope[T.3.0]	0.186	0.413	-0.603	0.935	0.007	0.006	3287.0	2791.0	1.0
restecg[T.1.0]	-0.010	0.263	-0.484	0.518	0.004	0.004	3683.0	2696.0	1.0
restecg[T.2.0]	0.008	0.268	-0.472	0.520	0.004	0.004	3997.0	3144.0	1.0

We have a large number of predictors and full output can be found in the Supplemental Images at the end of the paper. We show just a sample of the posterior probabilities below. By way of example, we can interpret the mean of intercept as the expected mean posterior probability when all predictor are zero is negative and large in magnitude. Specifically, the expected trace mean of the odds ratio when all predictors are



zero is -2.90 with a credible interval between -5.68 and -0.47.

### Bayesian Model Averaging

From the table below, we can see that the same variables rise to the top in terms of the magnitude of predictive value, and they have the same sign (positive vs. negative) on the expected value .

	coef	std err	z	P> z	[0.025	0.975]
const	-4.7944	1.380	-3.473	0.001	-7.500	-2.089
oldpeak	3.0286	0.649	4.665	0.000	1.756	4.301
thalach	-2.2712	0.926	-2.452	0.014	-4.087	-0.456
ca	1.3421	0.189	7.104	0.000	0.972	1.712
exang	1.1168	0.215	5.194	0.000	0.695	1.538
sex	1.1096	0.256	4.337	0.000	0.608	1.611
trestbps	1.0633	1.041	1.022	0.307	-0.976	3.103
cp	0.7241	0.105	6.896	0.000	0.518	0.930
chol	-0.6771	0.603	-1.122	0.262	-1.859	0.505
fbs	0.5516	0.266	2.071	0.038	0.030	1.074
slope	0.2868	0.180	1.596	0.110	-0.065	0.639
age	0.1689	0.897	0.188	0.851	-1.589	1.926
thal	0.1561	0.058	2.672	0.008	0.042	0.271
restecg	-0.0003	0.124	-0.002	0.998	-0.243	0.242

**Table 4**

*Summary Table statsmodels*

	Variable Name	Probability	Avg. Coefficient
3	cp	1.0	0.709164
12	ca	1.0	1.363645
10	oldpeak	1.0	3.35622
2	sex	0.999998	1.217282
9	exang	0.999995	1.130296
8	thalach	0.924357	-3.031
0	const	0.918035	-3.157973
13	thal	0.683891	0.120862
6	fbs	0.373102	0.232027
11	slope	0.169034	0.062245
5	chol	0.036271	-0.029707
4	trestbps	0.025566	0.030885
1	age	0.020963	0.018108
7	restecg	0.0	0.0

**Table 5**

*Summary Table Bayesian Model Averaging*

## Bayesian Hierarchical Model

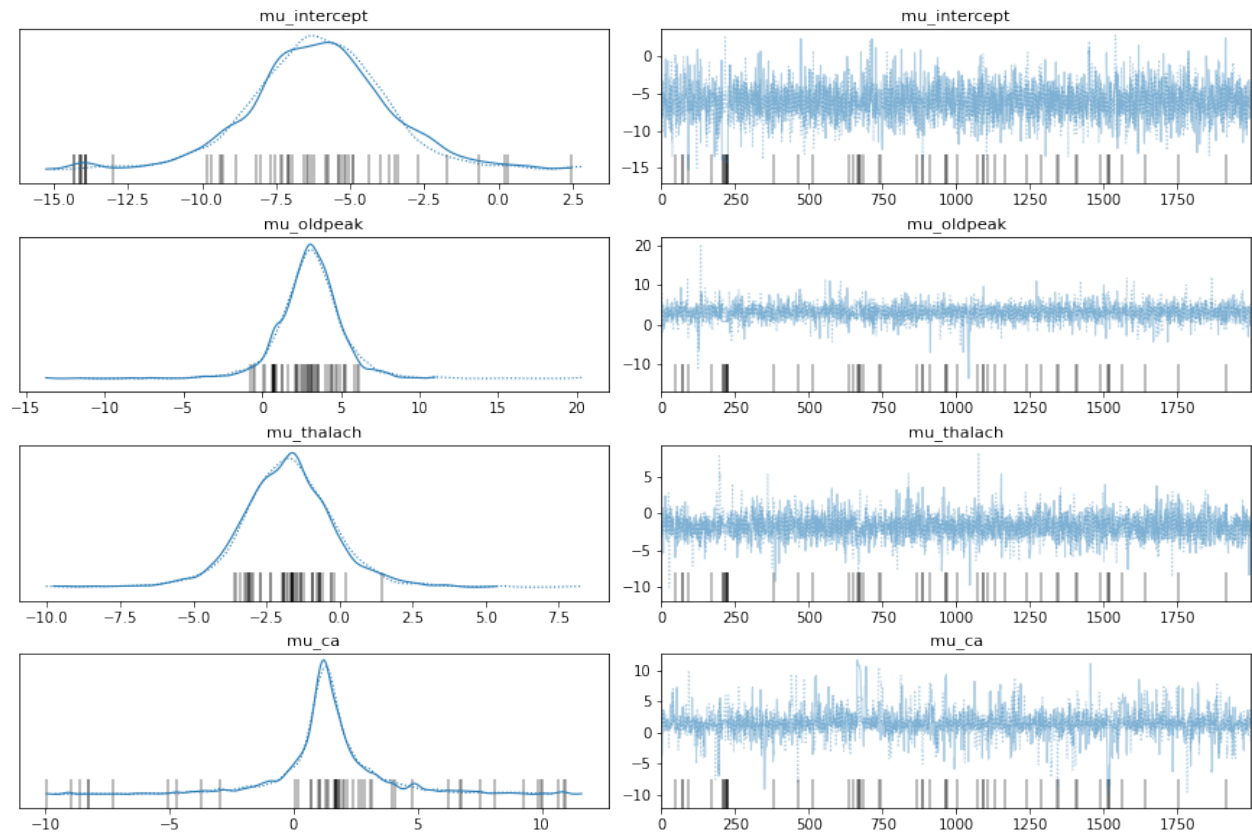
As mentioned above, since we don't have a strong assumption as to the baseline distribution of the parameters, I will leverage the given weakly-informative priors. For each of the parameters, we set  $\mu_n$  parameters to be normally distributed so that the regression coefficients could range from negative infinity to positive infinity. Since we have hierarchical parameters, we prepared the model as the usual normal mu and normal offset with  $(\mu = 0, \sigma = 1)$ .

For the variance distribution, we account for the variation and uncertainty at the

country-level via sigma. Here we use a half-Cauchy distribution, which is a heavy tailed or weakly informative distribution with a positive scale parameter (sigma greater than or equal to 1). The last building block is the pymc Deterministic class to bring it all together. Ultimately, our Hierarchical Model achieve good stability with minimal divergences. (see Supplemental Images)

As for this model, the variables ‘oldpeak‘ and ‘ca‘ show a strong positive association and ‘thalach‘ shows a negative association with the presence of heart disease.

Last, we can also see by looking at the marginal plots that there is some differences in the probability of an individual having heart disease between countries (each ‘rainbow’ color corresponding to a single country). The different widths are related to how much confidence we have in each parameter estimate – the more measurements per country, the higher our confidence will be. ([Reference](#)) See Figures 24 (offsets) and 25 (marginals) and Table 8 in the Supplemental Images section at the end of the paper.

**Figure 5**

*Bayesian Hierarchical Model Trace*

**Table 6***Summary Table Bayesian Hierarchical Model Mu*

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
mu_intercept	-6.067	2.344	-10.550	-1.659	0.057	0.053	2468.0	1085.0	1.0
mu_oldpeak	3.050	2.034	-0.595	6.906	0.045	0.033	2345.0	2151.0	1.0
mu_thalach	-1.771	1.602	-4.614	1.339	0.031	0.024	3014.0	2439.0	1.0
mu_ca	1.429	2.104	-3.151	5.631	0.074	0.059	1182.0	956.0	1.0
mu_trestbps	1.344	1.872	-2.087	4.952	0.038	0.032	3022.0	2130.0	1.0
mu_chol	1.132	2.275	-2.567	5.860	0.052	0.043	2485.0	2123.0	1.0
mu_exang	1.081	0.487	0.269	1.845	0.018	0.015	1508.0	998.0	1.0
mu_sex	1.009	0.841	-0.618	2.452	0.023	0.017	1500.0	1015.0	1.0
mu_age	0.840	1.697	-2.253	4.171	0.032	0.026	3039.0	2578.0	1.0
mu_fbs	0.839	1.605	-1.698	3.531	0.065	0.052	1148.0	886.0	1.0
mu_cp	0.739	0.208	0.361	1.139	0.005	0.004	2394.0	1693.0	1.0
mu_slope	0.370	0.569	-0.642	1.453	0.013	0.010	2551.0	948.0	1.0
mu_thal	0.130	0.239	-0.380	0.542	0.008	0.007	1351.0	958.0	1.0
mu_restecg	-0.103	0.478	-1.017	0.656	0.013	0.013	2053.0	1552.0	1.0

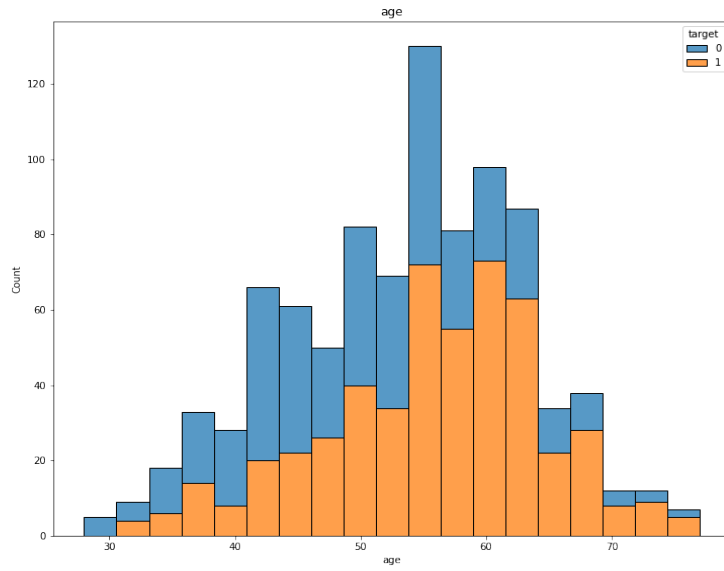
## Summary

In summary, from our models we can see that the variables old peak (ST depression induced by exercise relative to rest), thalach (maximum heart rate achieved), and ca (number of major coronary artery calcification vessels (0-3) colored by fluoroscopy) play a large role in determining the likelihood that a patient has heart disease. Conversely, variables such as restecg (resting electrocardiographic results) play a much smaller role. We can also find some differences in likelihood of having heart disease between countries, which is an interesting observation as well as justifies the use of Hierarchical Bayesian Analysis for problems of this nature.

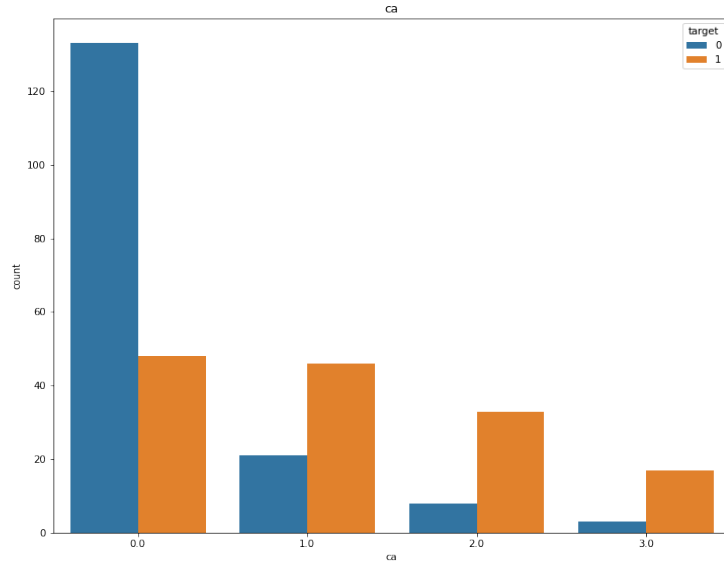
## Additional Resources

- <https://docs.pymc.io/notebooks/GLM-hierarchical.html>
- [https://docs.pymc.io/notebooks/posterior\\_predictive.html](https://docs.pymc.io/notebooks/posterior_predictive.html)
- [https://docs.pymc.io/notebooks/multilevel\\_modeling.html](https://docs.pymc.io/notebooks/multilevel_modeling.html)
- [https://docs.pymc.io/pymc-examples/examples/case\\_studies/rugby\\_analytics.html](https://docs.pymc.io/pymc-examples/examples/case_studies/rugby_analytics.html)
- [https://pymc3.readthedocs.io/en/latest/notebooks/multilevel\\_modeling.html](https://pymc3.readthedocs.io/en/latest/notebooks/multilevel_modeling.html)
- [https://twiecki.io/blog/2018/08/13/hierarchical\\_bayesian\\_neural\\_network/](https://twiecki.io/blog/2018/08/13/hierarchical_bayesian_neural_network/)
- <https://twiecki.io/blog/2013/08/12/bayesian-glms-1/>
- <https://www.kaggle.com/billbasener/bayesian-model-averaging-logistic-regression>
- <https://towardsdatascience.com/data-normalization-with-pandas-and-scikit-learn-7c1cc6ed6475>
- <https://towardsdatascience.com/all-about-feature-scaling-bcc0ad75cb35>
- <https://stats.stackexchange.com/questions/237847/what-are-the-properties-of-a-half-cauchy-distribution>
- <https://winderresearch.com/covid-19-hierarchical-bayesian-logistic-model-with-pymc3/>

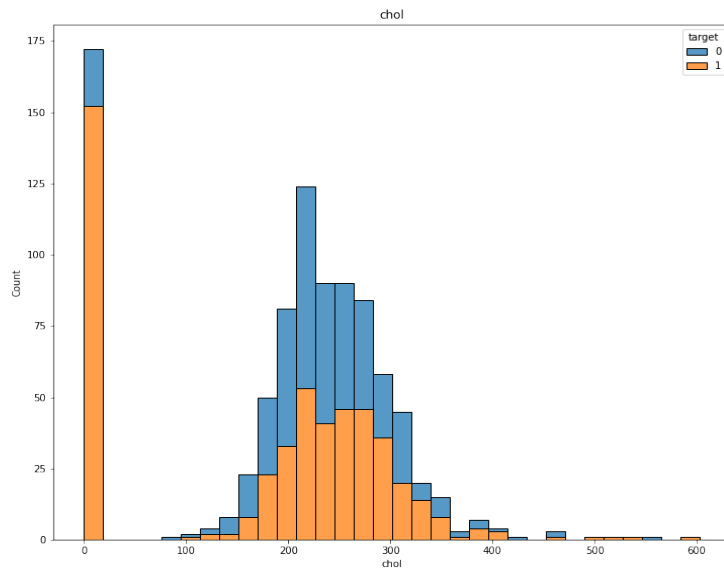
## Supplemental Images

**Figure 6***Age Histogram*

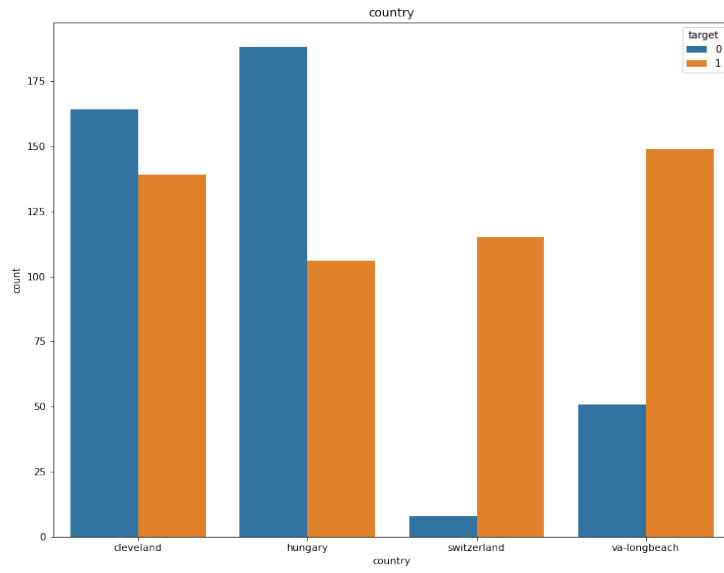


**Figure 7**

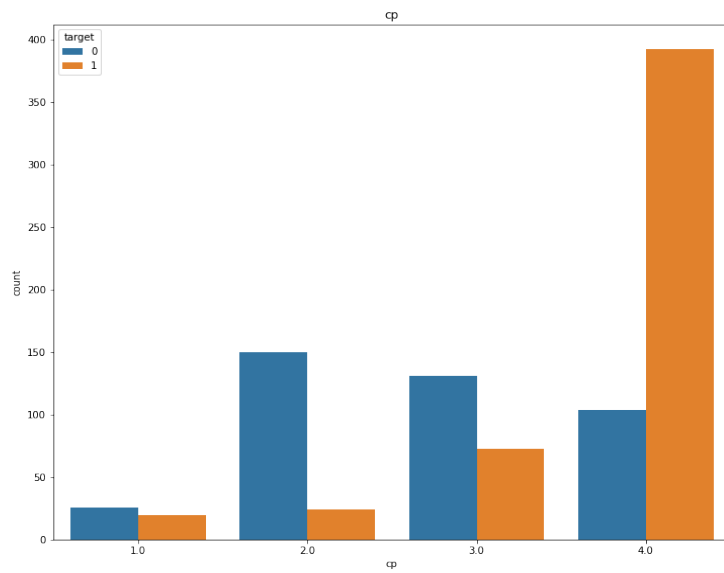
*Ca Histogram*

**Figure 8**

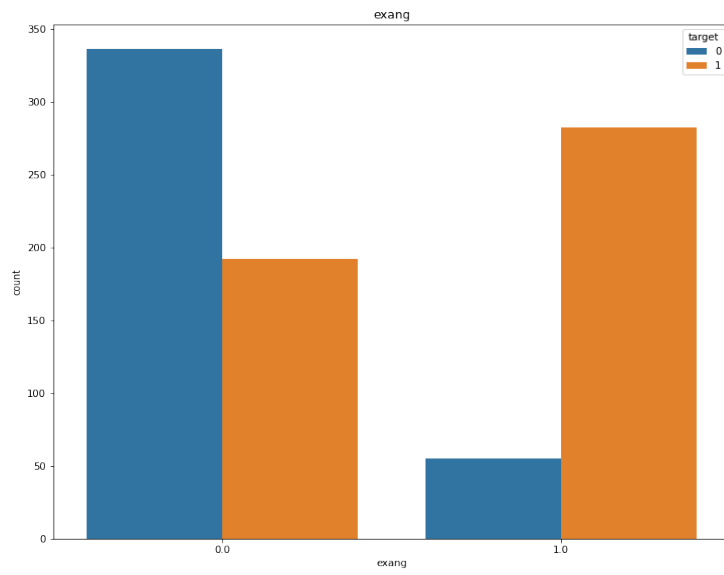
*Chol Histogram*

**Figure 9**

*Country Histogram*

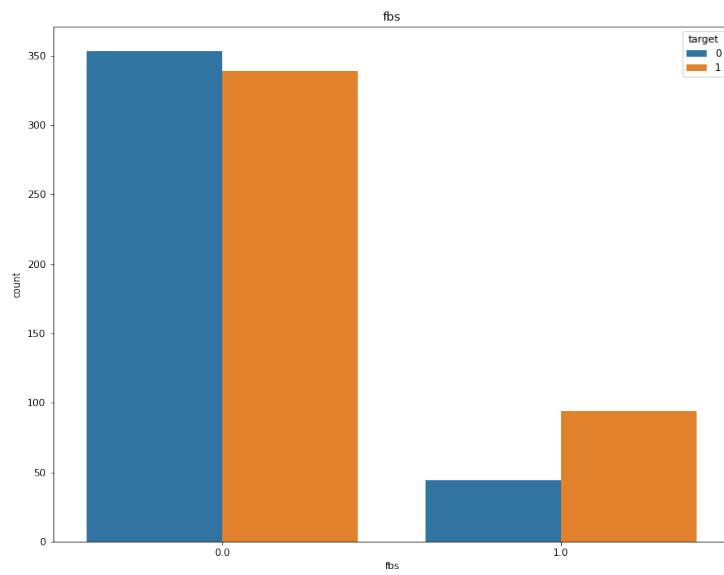
**Figure 10**

*Cp Histogram*



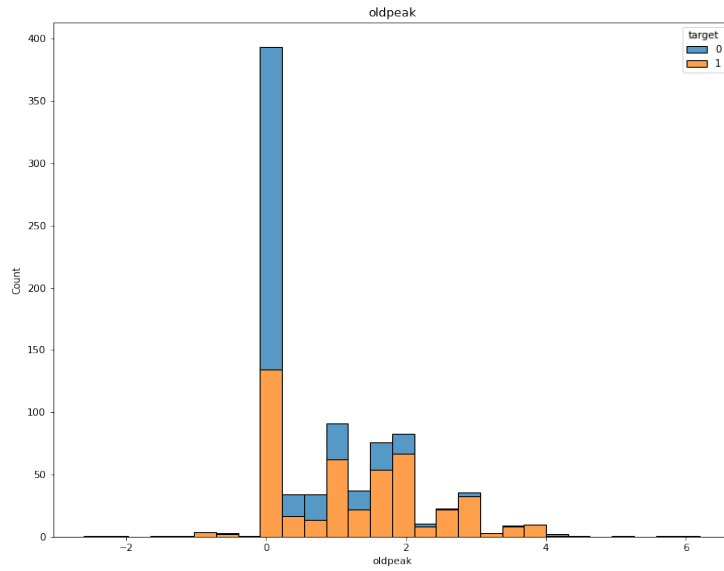
**Figure 11**

*Exang Histogram*



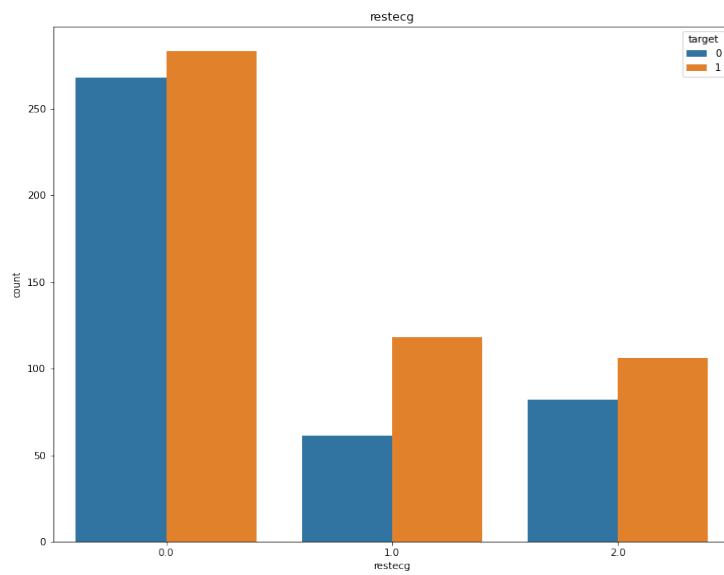
**Figure 12**

*Fbs Histogram*



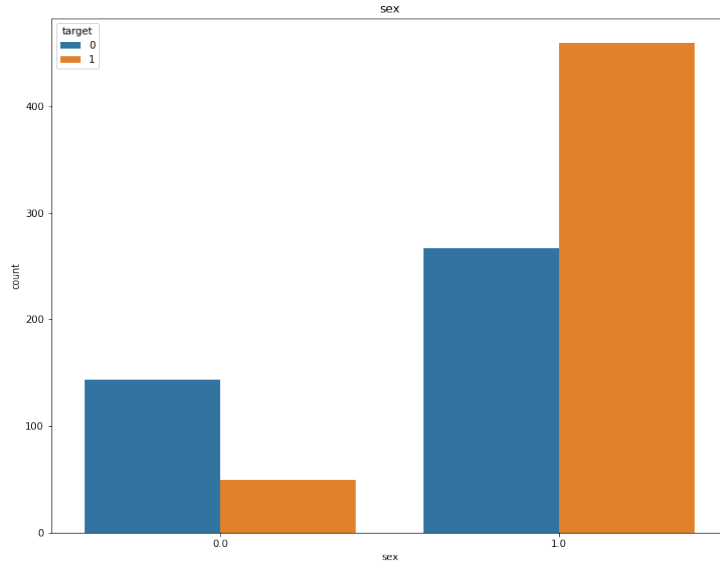
**Figure 13**

*Oldpeak Histogram*

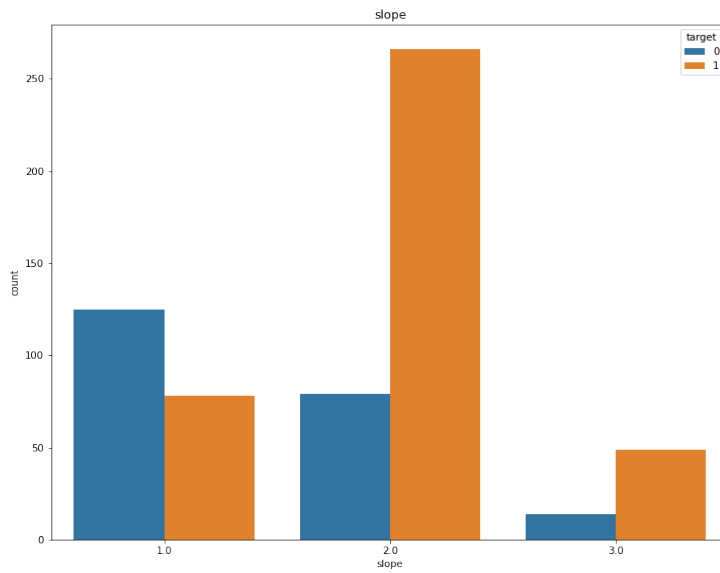


**Figure 14**

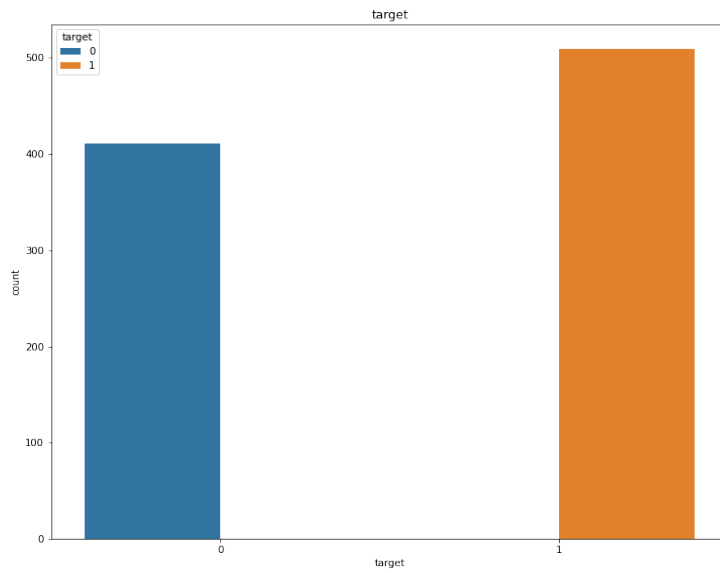
*Restecg Histogram*

**Figure 15**

*Sex Histogram*

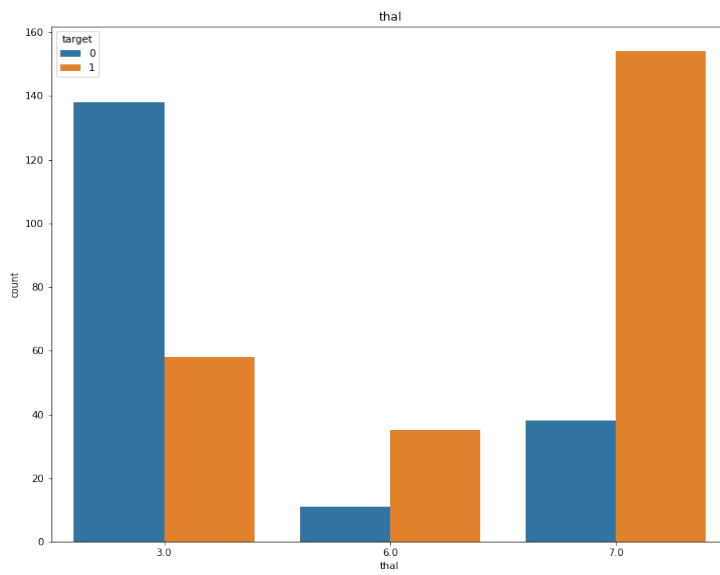
**Figure 16**

*Slope Histogram*



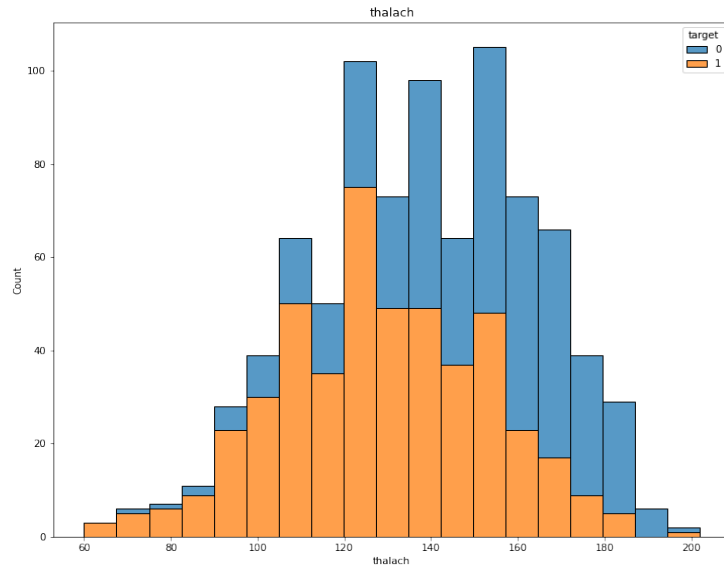
**Figure 17**

*Target Histogram*



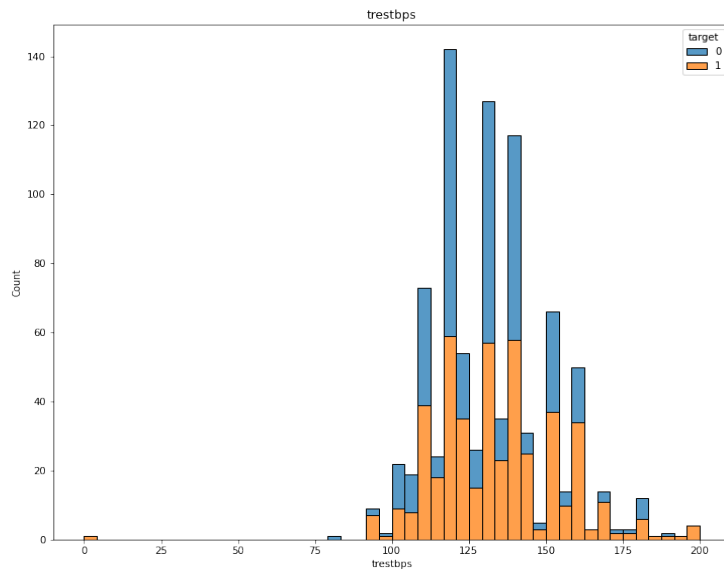
**Figure 18**

*Thal Histogram*



**Figure 19**

*Thalach Histogram*



**Figure 20**

*Trestbps Histogram*

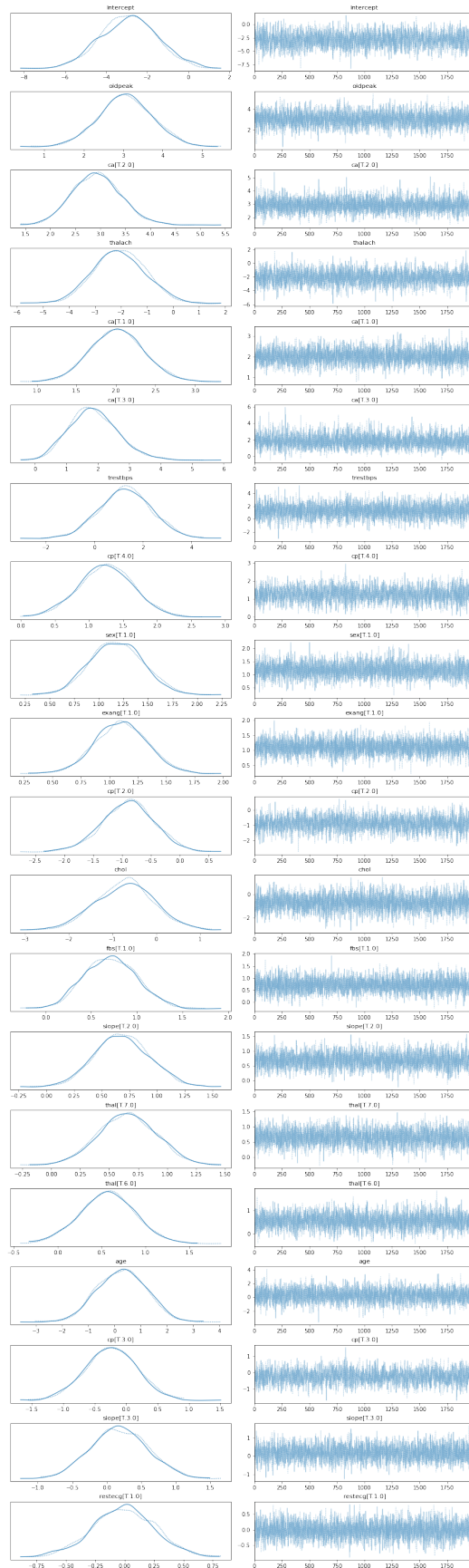
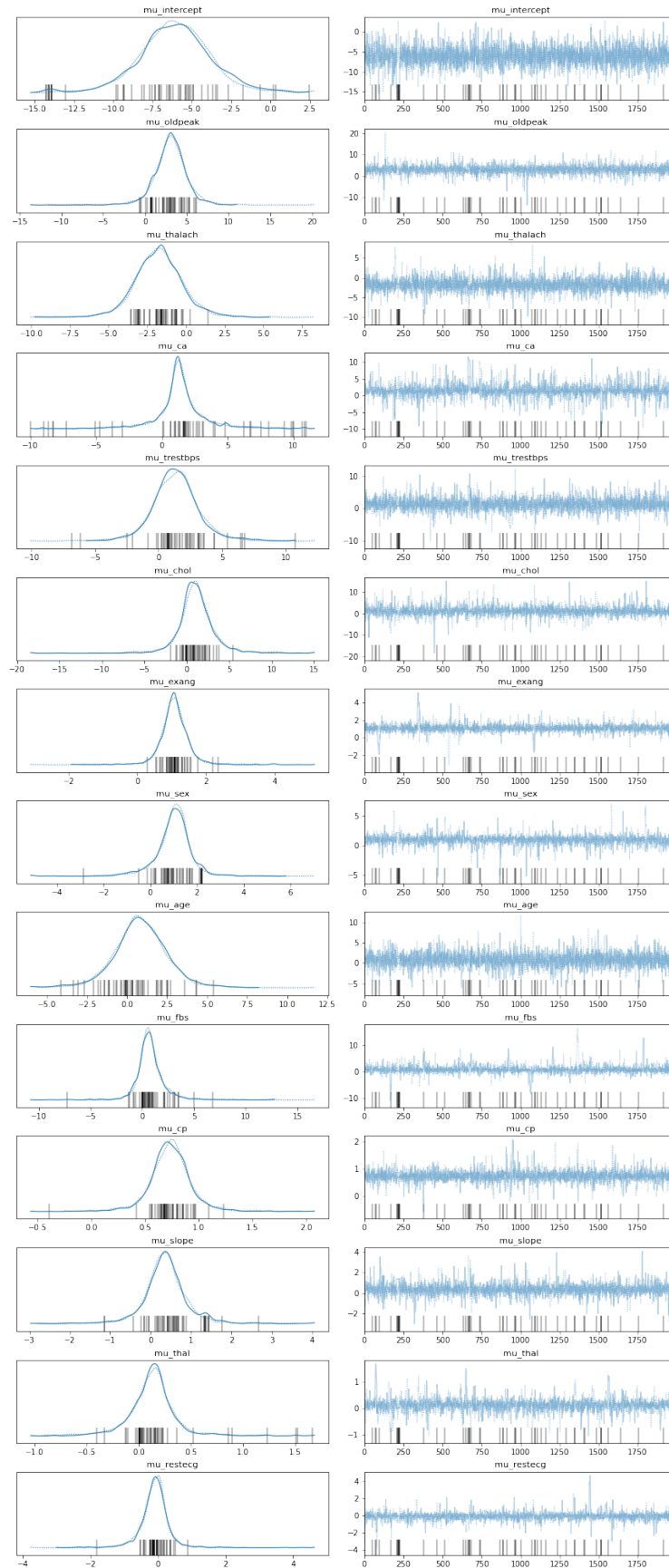


Figure 21

Trace Plot Of Logistic Model



**Figure 22**

*Trace Plot Of Mus Hierarchical Model*

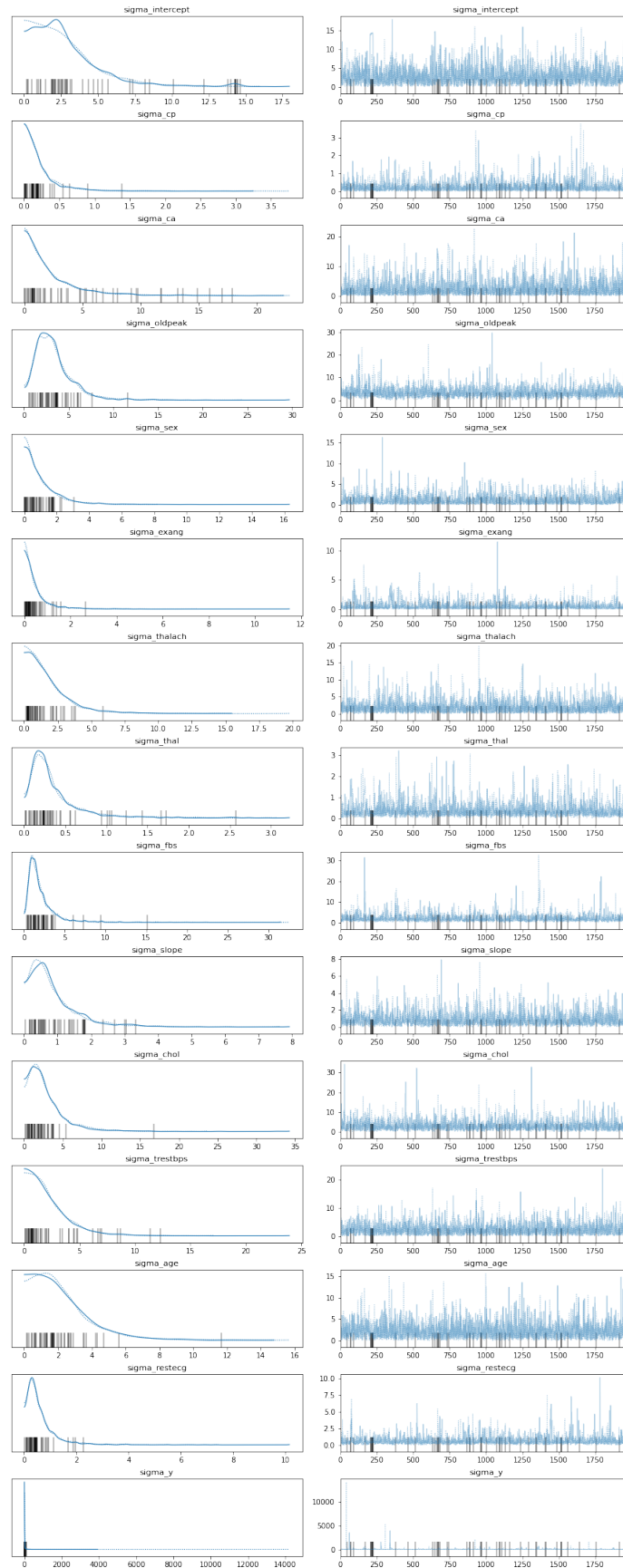
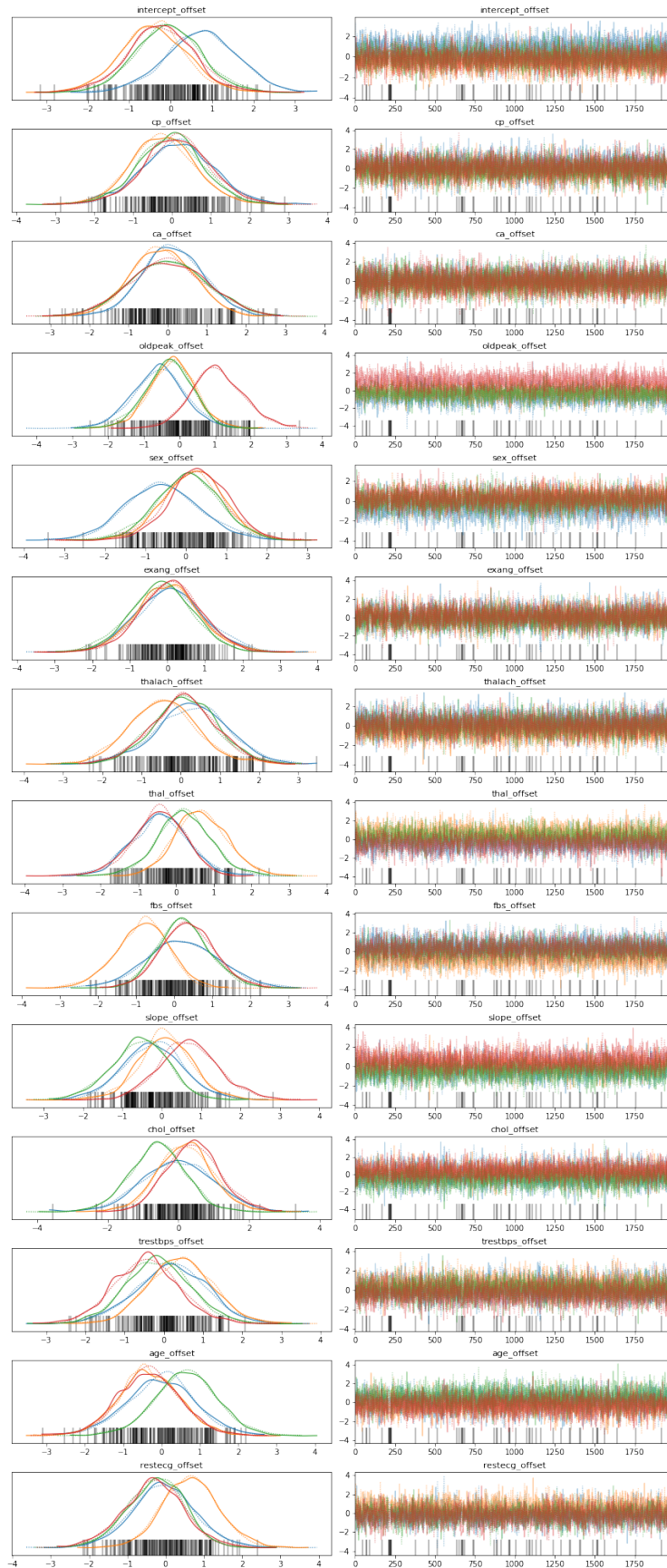


Figure 23

Trace Plot Of Sigmas Hierarchical Model

**Figure 24**

*Trace Plot Of Offsets Hierarchical Model*

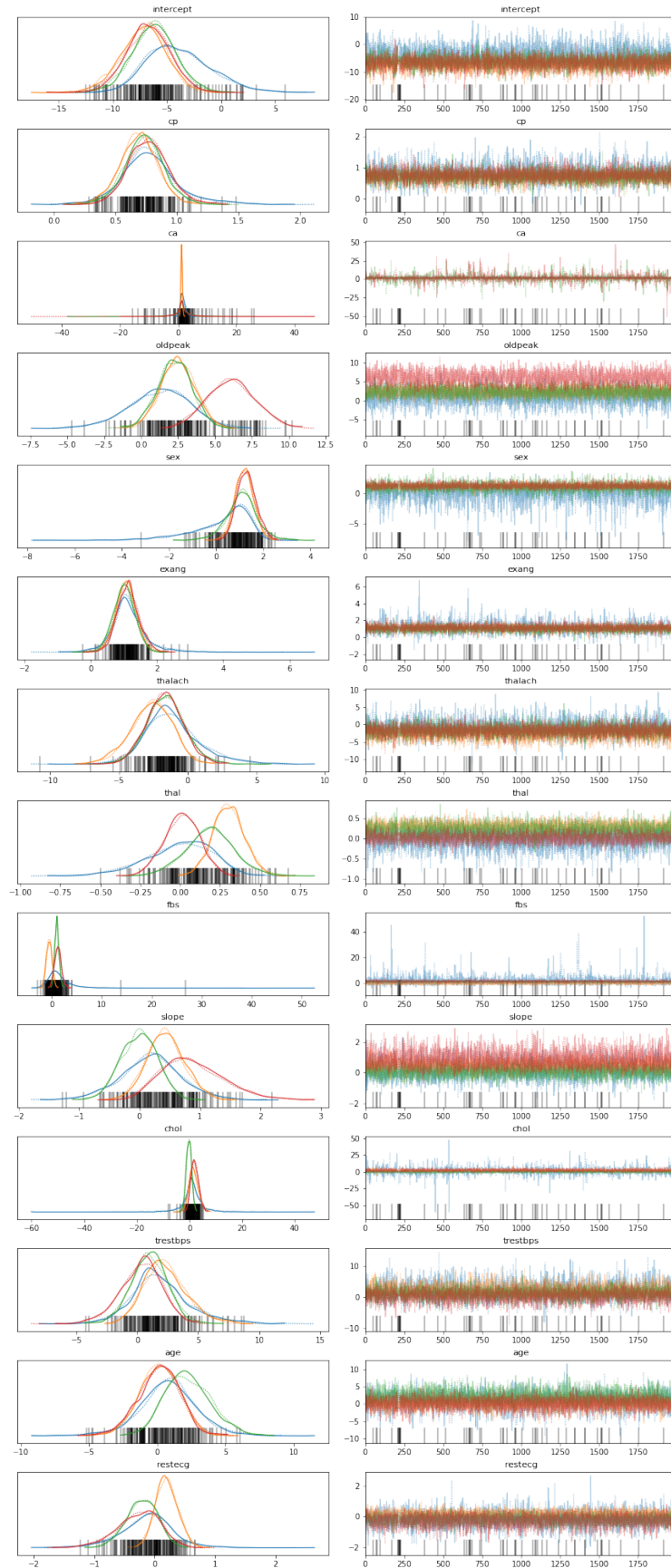


Figure 25

*Trace Plot Of Marginals Hierarchical Model*

## Supplemental Tables

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
intercept_offset[0]	0.731	0.884	-0.941	2.393	0.013	0.011	4418.0	3102.0	1.0
intercept_offset[1]	-0.429	0.803	-1.982	1.062	0.014	0.012	3053.0	2939.0	1.0
intercept_offset[2]	-0.060	0.799	-1.539	1.467	0.013	0.013	3933.0	2815.0	1.0
intercept_offset[3]	-0.242	0.815	-1.874	1.243	0.014	0.013	3582.0	2465.0	1.0
cp_offset[0]	0.151	0.936	-1.767	1.820	0.015	0.017	3825.0	916.0	1.0
cp_offset[1]	-0.227	0.839	-1.792	1.398	0.013	0.013	3915.0	3091.0	1.0
cp_offset[2]	-0.042	0.847	-1.582	1.649	0.012	0.015	4579.0	2764.0	1.0
cp_offset[3]	0.123	0.869	-1.478	1.856	0.014	0.015	4056.0	2730.0	1.0
ca_offset[0]	0.143	0.792	-1.343	1.604	0.015	0.012	2597.0	2743.0	1.0
ca_offset[1]	-0.150	0.804	-1.652	1.342	0.016	0.012	2595.0	2978.0	1.0
ca_offset[2]	0.020	0.977	-1.906	1.782	0.014	0.019	4734.0	2538.0	1.0
ca_offset[3]	0.017	0.995	-1.920	1.780	0.012	0.018	6845.0	2843.0	1.0
oldpeak_offset[0]	-0.572	0.765	-1.944	0.951	0.013	0.012	3474.0	2681.0	1.0
oldpeak_offset[1]	-0.200	0.664	-1.589	0.987	0.012	0.010	3042.0	2667.0	1.0
oldpeak_offset[2]	-0.304	0.663	-1.596	0.874	0.013	0.010	2494.0	2173.0	1.0
oldpeak_offset[3]	1.059	0.744	-0.297	2.533	0.014	0.010	2945.0	2561.0	1.0
sex_offset[0]	-0.559	0.928	-2.302	1.222	0.015	0.012	4047.0	3192.0	1.0
sex_offset[1]	0.166	0.766	-1.258	1.641	0.015	0.012	2609.0	2802.0	1.0
sex_offset[2]	0.050	0.805	-1.458	1.588	0.013	0.014	3693.0	2614.0	1.0
sex_offset[3]	0.272	0.765	-1.107	1.773	0.015	0.012	2444.0	2479.0	1.0
exang_offset[0]	0.067	0.933	-1.591	1.919	0.015	0.020	4097.0	2269.0	1.0
exang_offset[1]	0.022	0.875	-1.758	1.612	0.016	0.016	2809.0	2169.0	1.0
exang_offset[2]	-0.156	0.846	-1.896	1.320	0.015	0.014	3230.0	2599.0	1.0
exang_offset[3]	0.089	0.840	-1.499	1.692	0.014	0.013	3800.0	2709.0	1.0
thalach_offset[0]	0.281	0.948	-1.647	1.954	0.016	0.016	3569.0	2381.0	1.0
thalach_offset[1]	-0.435	0.895	-2.238	1.115	0.017	0.017	3115.0	1243.0	1.0
thalach_offset[2]	0.086	0.849	-1.485	1.701	0.014	0.014	3780.0	2359.0	1.0
thalach_offset[3]	0.070	0.845	-1.540	1.685	0.016	0.014	2836.0	2354.0	1.0
thal_offset[0]	-0.421	0.779	-1.860	1.028	0.014	0.012	2979.0	2763.0	1.0
thal_offset[1]	0.679	0.708	-0.699	1.972	0.014	0.010	2319.0	1914.0	1.0
thal_offset[2]	0.209	0.729	-1.160	1.639	0.014	0.012	2692.0	2696.0	1.0
thal_offset[3]	-0.475	0.721	-1.809	0.874	0.016	0.011	2064.0	2798.0	1.0
fbs_offset[0]	0.231	0.957	-1.647	1.971	0.013	0.016	5522.0	2975.0	1.0
fbs_offset[1]	-0.839	0.706	-2.263	0.375	0.014	0.010	2501.0	1420.0	1.0
fbs_offset[2]	0.255	0.688	-1.000	1.560	0.015	0.010	2154.0	2274.0	1.0

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
sigma_intercept	2.770	2.326	0.001	6.739	0.073	0.068	1521.0	1292.0	1.0
sigma_cp	0.237	0.299	0.000	0.700	0.008	0.006	1566.0	1738.0	1.0
sigma_ca	2.297	2.668	0.001	6.872	0.080	0.060	1381.0	1346.0	1.0
sigma_oldpeak	3.317	2.216	0.054	6.768	0.048	0.034	1792.0	1824.0	1.0
sigma_sex	1.063	1.140	0.000	3.008	0.028	0.020	1465.0	2396.0	1.0
sigma_exang	0.499	0.655	0.000	1.546	0.021	0.015	1365.0	1084.0	1.0
sigma_thalach	1.775	1.744	0.001	4.725	0.037	0.026	1924.0	2139.0	1.0
sigma_thal	0.361	0.353	0.001	0.920	0.010	0.007	1281.0	1282.0	1.0
sigma_fbs	2.032	2.019	0.000	4.820	0.070	0.050	1171.0	1069.0	1.0
sigma_slope	0.855	0.764	0.001	2.161	0.019	0.013	1313.0	1541.0	1.0
sigma_chol	2.794	2.587	0.001	7.042	0.061	0.043	1434.0	1788.0	1.0
sigma_trestbps	2.192	1.988	0.002	5.572	0.041	0.029	1856.0	1958.0	1.0
sigma_age	2.194	1.836	0.002	5.325	0.035	0.025	2309.0	2372.0	1.0
sigma_restecg	0.623	0.683	0.000	1.608	0.019	0.013	1297.0	1810.0	1.0
sigma_y	26.628	266.007	0.000	58.015	4.443	3.142	3938.0	1551.0	1.0

**Table 8**

*Summary Table Hierarchical Sigmas*

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
intercept[0]	-3.887	2.984	-8.826	1.962	0.062	0.044	2465.0	2114.0	1.0
intercept[1]	-7.202	2.092	-11.338	-3.537	0.031	0.022	4494.0	3207.0	1.0
intercept[2]	-6.219	2.039	-10.158	-2.397	0.032	0.023	4211.0	3154.0	1.0
intercept[3]	-6.726	2.045	-10.663	-3.036	0.030	0.022	4521.0	3175.0	1.0
cp[0]	0.782	0.240	0.334	1.270	0.004	0.003	3480.0	1338.0	1.0
cp[1]	0.685	0.154	0.381	0.969	0.002	0.002	4620.0	3937.0	1.0
cp[2]	0.730	0.163	0.399	1.027	0.002	0.002	4696.0	3216.0	1.0
cp[3]	0.767	0.165	0.462	1.083	0.003	0.002	3989.0	3398.0	1.0
ca[0]	1.647	0.899	0.006	3.428	0.015	0.011	3771.0	3708.0	1.0
ca[1]	1.182	0.250	0.752	1.692	0.004	0.003	3973.0	3472.0	1.0
ca[2]	1.363	3.989	-6.483	8.703	0.098	0.069	2370.0	1788.0	1.0
ca[3]	1.566	4.242	-6.005	9.526	0.120	0.086	2088.0	1426.0	1.0
oldpeak[0]	1.275	2.079	-2.637	5.226	0.038	0.032	2988.0	2474.0	1.0
oldpeak[1]	2.505	1.165	0.290	4.587	0.019	0.014	3892.0	3786.0	1.0
oldpeak[2]	2.278	1.168	-0.004	4.359	0.019	0.013	3867.0	3193.0	1.0
oldpeak[3]	6.119	1.631	3.100	9.245	0.032	0.023	2579.0	2820.0	1.0
sex[0]	0.259	1.314	-2.402	2.030	0.028	0.020	2212.0	2703.0	1.0
sex[1]	1.197	0.395	0.461	1.935	0.006	0.004	4123.0	3848.0	1.0
sex[2]	1.095	0.671	-0.157	2.448	0.010	0.008	4501.0	3457.0	1.0
sex[3]	1.288	0.402	0.534	2.037	0.007	0.005	3270.0	3536.0	1.0
exang[0]	1.147	0.571	0.029	2.254	0.011	0.010	4032.0	2232.0	1.0
exang[1]	1.073	0.316	0.449	1.652	0.005	0.003	4454.0	3462.0	1.0
exang[2]	1.010	0.323	0.417	1.629	0.005	0.003	4560.0	3849.0	1.0
exang[3]	1.129	0.339	0.559	1.844	0.005	0.004	3994.0	2536.0	1.0
thalach[0]	-1.127	2.063	-4.716	3.154	0.036	0.029	3353.0	2574.0	1.0
thalach[1]	-2.631	1.516	-5.760	-0.033	0.023	0.017	4207.0	3727.0	1.0
thalach[2]	-1.605	1.436	-4.330	1.112	0.021	0.016	4941.0	3601.0	1.0
thalach[3]	-1.681	1.381	-4.212	1.034	0.019	0.015	5140.0	3907.0	1.0
thal[0]	-0.011	0.215	-0.414	0.367	0.004	0.003	3216.0	3198.0	1.0
thal[1]	0.303	0.100	0.111	0.490	0.002	0.001	3289.0	3383.0	1.0
thal[2]	0.188	0.156	0.106	0.405	0.002	0.002	4098.0	3500.0	1.0