**DECISION TREE ANALYSIS**

The model summary table provides some very broad information about the specifications used to build the model and the resulting model.

* The Specifications section provides information on the settings used to generate the tree Model, including the variables used in the analysis.
* The Results section displays information on the number of total and terminal nodes, depth of the tree (number of levels below the root node), and independent variables included in the final model.

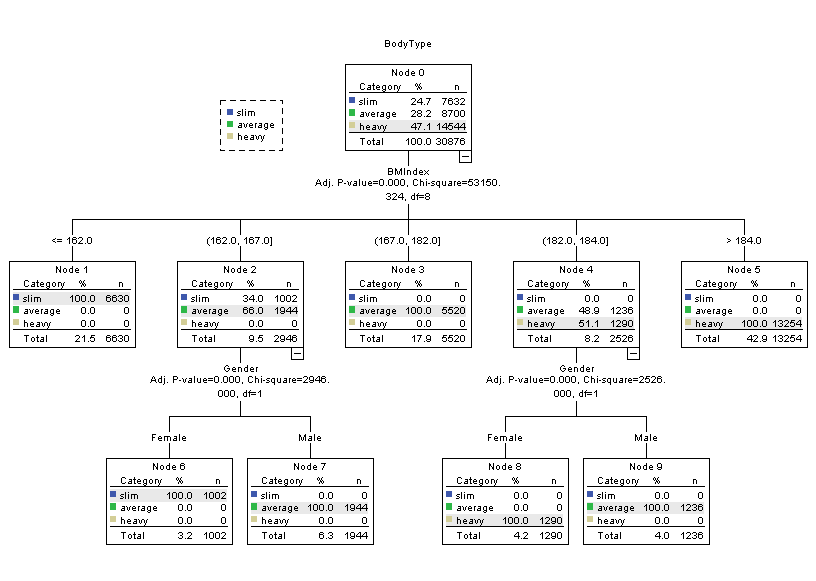
Four independent variables were specified, but only two were included in the final model. The Variables for Ht\_cent and Height did not make a significant contribution to the model, so they were automatically dropped from the final model.

|  |  |  |
| --- | --- | --- |
| **Model Summary** | | |
| Specifications | Growing Method | CHAID |
| Dependent Variable | BodyType |
| Independent Variables | Gender, Height, Ht\_Cent., BMIndex(Body mass index) |
| Validation | None |
| Maximum Tree Depth | 3 |
| Minimum Cases in Parent Node | 100 |
| Minimum Cases in Child Node | 50 |
| Results | Independent Variables Included | BMIndex, Gender |
| Number of Nodes | 10 |
| Number of Terminal Nodes | 7 |
| Depth | 2 |

The tree diagram is a graphic representation of the tree model. This tree diagram shows that:

Using the CHAID method, Body mass index (BMindex) is the best predictor of Body type.

* For the <=162.0 BMindex category, BMindex is the only significant predictor of Body type. Of the Population in this category, 100% have **Slim** population under <= 162.00 BMindex. Since there are no child nodes below it, this is considered a terminal node (Node 1).
* For the 162-167Bmindex categories(Node 2), the next best predictor is Gender you can see that gender predict that is 100% female population is **Slim** BMindex under the 162-167 category that is good for health and 100% male population is **Average** BMindex under the 162-167category(Node 6 and 7).
* For the 167-182 BMindex category, BMindex is the only significant predictor of Body type. The Population in this category, 100% **Average** population under 162-182 BMindex (Node 3).
* For the 182-184BMindex category(Node 4), the next best predictor have Gender you can see gender predict that is 100% female population is **Heavy** BMindex under the 182-184 category and 100% male population is **Average** BMindex under the 182-184 category(Node 8 and 9).
* For the >184 BMindex category, BMindex is the only significant predictor of Body type. The Population in this category, 100% Heavy population under >184 category (Node 5).



**RCODE**

#DECISION TREE

height <- read.csv("C:/Users/WELCOME/Desktop/height.csv")

View(height)

library("party")

str(height)

height <- ctree(BodyType ~ Gender + Weight + Height + BMIndex + BodyType, data=height)

print(height)

plot(height)

plot(height, type="simple")

**OUTPUT**

|  |
| --- |
| > height <- read.csv("C:/Users/WELCOME/Desktop/height.csv")  > View(height)  > library("party")  > str(height)  'data.frame': 168 obs. of 6 variables:  $ Gender : Factor w/ 2 levels "Female","Male": 1 1 2 2 1 1 1 2 1 1 ...  $ Weight : int 179 160 191 132 167 128 150 150 215 89 ...  $ Height : num 4.1 5.4 5.8 5.1 5.1 5.2 5.2 5.2 5.2 5.3 ...  $ Ht\_Cent.: int 147 163 173 155 180 157 157 157 157 160 ...  $ BMIndex : int 162 161 182 143 174 142 154 154 184 119 ...  $ BodyType: Factor w/ 3 levels "average","heavy",..: 3 3 1 3 1 3 3 3 2 3 ...  > height <- ctree(BodyType ~ Gender + Weight + Height + BMIndex + BodyType, data=height)  > print(height)  Conditional inference tree with 3 terminal nodes  Response: BodyType  Inputs: Gender, Weight, Height, BMIndex, BodyType  Number of observations: 168  1) BodyType == {average}; criterion = 1, statistic = 334  2)\* weights = 48  1) BodyType == {heavy, slim}  3) BodyType == {slim}; criterion = 1, statistic = 119  4)\* weights = 54  3) BodyType == {heavy}  5)\* weights = 66  > plot(height)  > plot(height, type="simple") |
|  |
| |  | | --- | | > | |

