Gini Index

RID	Age	income	student	Credit rating	buy_computer
1	Youth	high	no	Fair	no
2	Youth	high	no	excellent	no
3	Middle	high	no	Fair	yes
4	senior	Medium	no	Fair	yes
5	senior	low	yes	Fair	yes
6	senior	low	yes	excellent	no
7	Middle	low	yes	excellent	yes
8	youth	Medium	no	Fair	no
9	youth	~~Medium~~ low	yes	Fair	yes
10	senior	Medium	yes	Fair	yes
11	youth	Medium	yes	excellent	yes
12	Middle	Medium	no	excellent	yes
13	Middle	high	yes	Fair	yes
14	senior	Medium	no	excellent	no

	buy_Computer
yes	9
NO	5
	14

$$Gini(D) = 1 - \left(\frac{9}{14}\right)^2 - \left(\frac{5}{14}\right)^2$$

$$= 0.4592$$

Consider each of possible splitting subsets for Age attribute

1 {youth, Middle}, {senior}
2 {youth, senior}, {middle}
3 {Middle, senior}, {youth}

Distinct value	YES	NO	Total
youth	2	3	5
Middle	4	0	4
senior	3	2	5
			14

consider subset {youth, Middle}

Gini Age ∈ {youth, Middle} (D) =

$$\frac{9}{14}\left(1-\left(\frac{6}{9}\right)^2-\left(\frac{3}{9}\right)^2\right)$$

$$+$$

$$\frac{5}{14}\left(1-\left(\frac{3}{5}\right)^2-\left(\frac{2}{5}\right)^2\right)$$

$$= 0.4571$$

Gini Age ∈ {youth, senior} (D) =

$$\frac{10}{14}\left(1-\left(\frac{5}{10}\right)^2-\left(\frac{5}{10}\right)^2\right)$$

$$+$$

$$\frac{4}{14}\left(1-\left(\frac{4}{4}\right)^2-\left(\frac{0}{4}\right)^2\right)$$

$$= 0.3571$$

Gini Age ∈ {Middle, senior} (D) =

$$\frac{9}{14}\left(1-\left(\frac{7}{9}\right)^2-\left(\frac{2}{9}\right)^2\right) +$$

$$\frac{5}{14}\left(1-\left(\frac{2}{5}\right)^2-\left(\frac{3}{5}\right)^2\right) = 0.3937$$

DATE / / PAGE NO

Similarly For income

{high, Low} {medium}
{high, Medium} {Low}
{Low, Medium} {high}

	yes	No	total
High	2	2	4
LOW	3	1	4
Medium	4	2	6
			14

Gini Income ∈ {high, LOW} (D)

$$= \frac{8}{14}\left(1-\left(\frac{5}{8}\right)^2-\left(\frac{3}{8}\right)^2\right) + \frac{6}{14}\left(1-\left(\frac{4}{6}\right)^2-\left(\frac{2}{6}\right)^2\right)$$

$$= 0.4583$$

Gini Income ∈ {high, Medium} (D)

$$= \frac{10}{14}\left(1-\left(\frac{6}{10}\right)^2-\left(\frac{4}{10}\right)^2\right) + \frac{4}{14}\left(1-\left(\frac{3}{4}\right)^2-\left(\frac{1}{4}\right)^2\right)$$

$$= 0.4500$$

Gini Income ∈ {Low, Medium} (D)

$$= \frac{10}{14}\left(1-\left(\frac{7}{10}\right)^2-\left(\frac{3}{10}\right)^2\right) + \frac{4}{14}\left(1-\left(\frac{2}{4}\right)^2-\left(\frac{2}{4}\right)^2\right)$$

$$= 0.4429$$

DATE / / PAGE NO

For Student Attribute

	YES	NO	Total
yes	6	1	7
NO	3	4	7
			14

$$\text{Gini student (D)} = \frac{7}{14}\left(1-\left(\frac{6}{7}\right)^2-\left(\frac{1}{7}\right)^2\right) + \frac{7}{14}\left(1-\left(\frac{3}{7}\right)^2-\left(\frac{4}{7}\right)^2\right)$$

$$= 0.3673$$

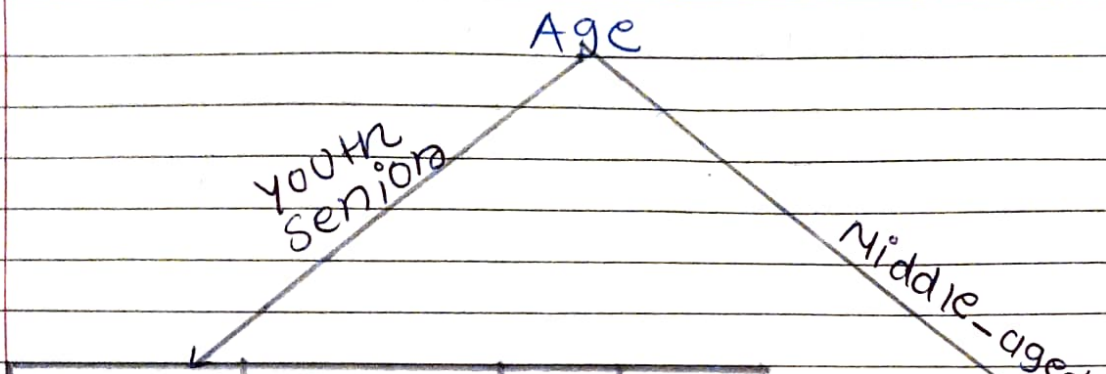
For credit rating

	YES	NO	Total
Fair	6	2	8
excellent	3	3	6
			14

$$\text{Gini Credit\_Rating} = \frac{8}{14}\left(1-\left(\frac{6}{8}\right)^2-\left(\frac{2}{8}\right)^2\right) + \frac{6}{14}\left(1-\left(\frac{3}{6}\right)^2-\left(\frac{3}{6}\right)^2\right)$$

$$= 0.4286$$

Attribute	Gini Index	Δ gini
Age	0.3571	0.4592 − 0.3571 = 0.1021
Income	0.4429	0.4592 − 0.4429 = 0.0163
student	0.3673	0.4592 − 0.3673 = 0.0919
credit_rating	0.4286	0.4592 − 0.4286 = 0.306



Income	student	CR	class
high	no	Fair	no
high	no	exe	no
Medium	no	Fair	no
low	yes	Fair	yes
Medium	yes	exe	yes
Medium	no	Fair	yes
low	yes	Fair	yes
low	yes	exe	no
Medium	yes	Fair	yes
Medium	no	exe	no

Income	Student	CR	class
high	no	fair	yes
Low	yes	excellent	yes
Medium	no	excellent	yes
high	yes	Fair	yes