

Loan Risk data with ggplot in R programing

Jintawee.s

Install & Load Package

```
library(ggplot2)
library(dplyr)

##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

Review Data

```
data <- read.csv("Loan_Risk.csv")
head(data)
```

	Id	Income	Age	Experience	Marital_Status	House_Ownership	Car_Ownership
## 1	1	1303834	23	3	single	rented	no
## 2	2	7574516	40	10	single	rented	no
## 3	3	3991815	66	4	married	rented	no
## 4	4	6256451	41	2	single	rented	yes
## 5	5	5768871	47	11	single	rented	no
## 6	6	6915937	64	0	single	rented	no

	Profession	CITY	CURRENT_JOB_YRS	CURRENT_HOUSE_YRS
## 1	Mechanical_engineer	Rewa	3	13
## 2	Software_Developer	Parbhani	9	13
## 3	Technical_writer	Alappuzha	4	10
## 4	Software_Developer	Bhubaneswar	2	12
## 5	Civil_servant	Tiruchirappalli[10]	3	14
## 6	Civil_servant	Jalgaon	0	12

	Risk_Flag	Rev_State	State_GDP	Literacy_Rate	Population	Unemployment
## 1	0	Madhya Pradesh	5.64514e+12	75.37	72627000	47
## 2	0	Maharashtra	1.88931e+13	94.00	112374000	65
## 3	0	Kerala	5.14400e+12	66.41	33406000	116
## 4	1	Odisha	3.81470e+12	74.43	41974000	78
## 5	1	Tamil Nadu	1.24560e+13	75.84	72147000	58
## 6	0	Maharashtra	1.88931e+13	94.00	112374000	65

	Poverty_Rate	Region	Per_Capita_Income	Income_Category	Age_Group
## 1	31.65	Central	77727.89	Low	Young

## 2	17.35	Western	168126.67	High	Young
## 3	7.05	Southern	153984.31	Medium	Senior
## 4	32.59	Eastern	90882.53	Medium	Middle-Age
## 5	11.28	Southern	172646.84	Medium	Middle-Age
## 6	17.35	Western	168126.67	High	Senior

Chart 1

This is a histogram chart to show Age distributin.

```
ggplot(data, aes(Age)) +
  geom_histogram(aes(y=..count..), color="black", fill="#FFE082")+
  geom_density(alpha=.2, fill="#FF6666") +
  geom_vline(aes(xintercept=mean(Age)), color="red",
             linetype="dashed") +
  theme_minimal() +
  labs(x = "Age")
```

```
## Warning: The dot-dot notation (`..count..`) was deprecated in ggplot2 3.4.0.
## i Please use `after_stat(count)` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

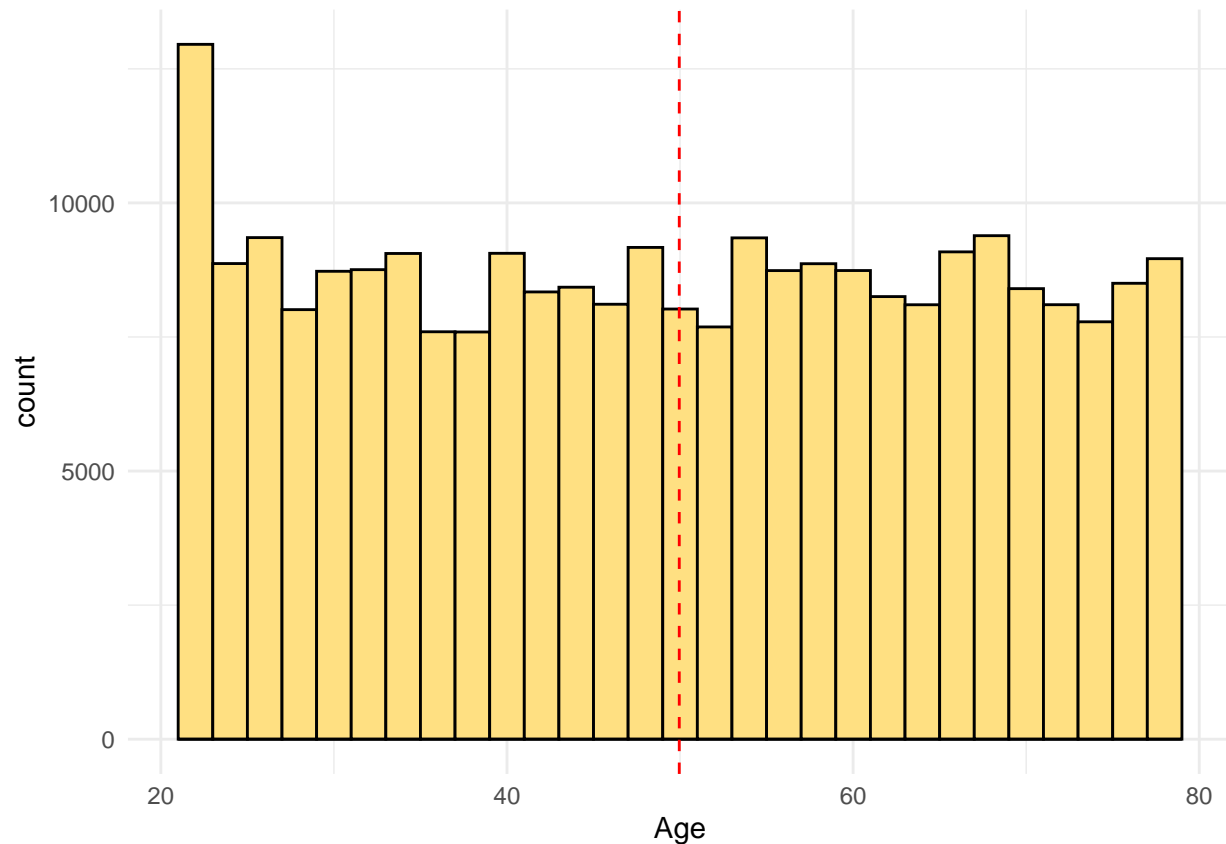


Chart 2 This is a bar chart to show numbers by Profession.

```
ggplot(data, aes(x=Profession))+
  geom_bar(width=0.7, fill="#558B2F")+
  theme_minimal() +
  coord_flip()
```

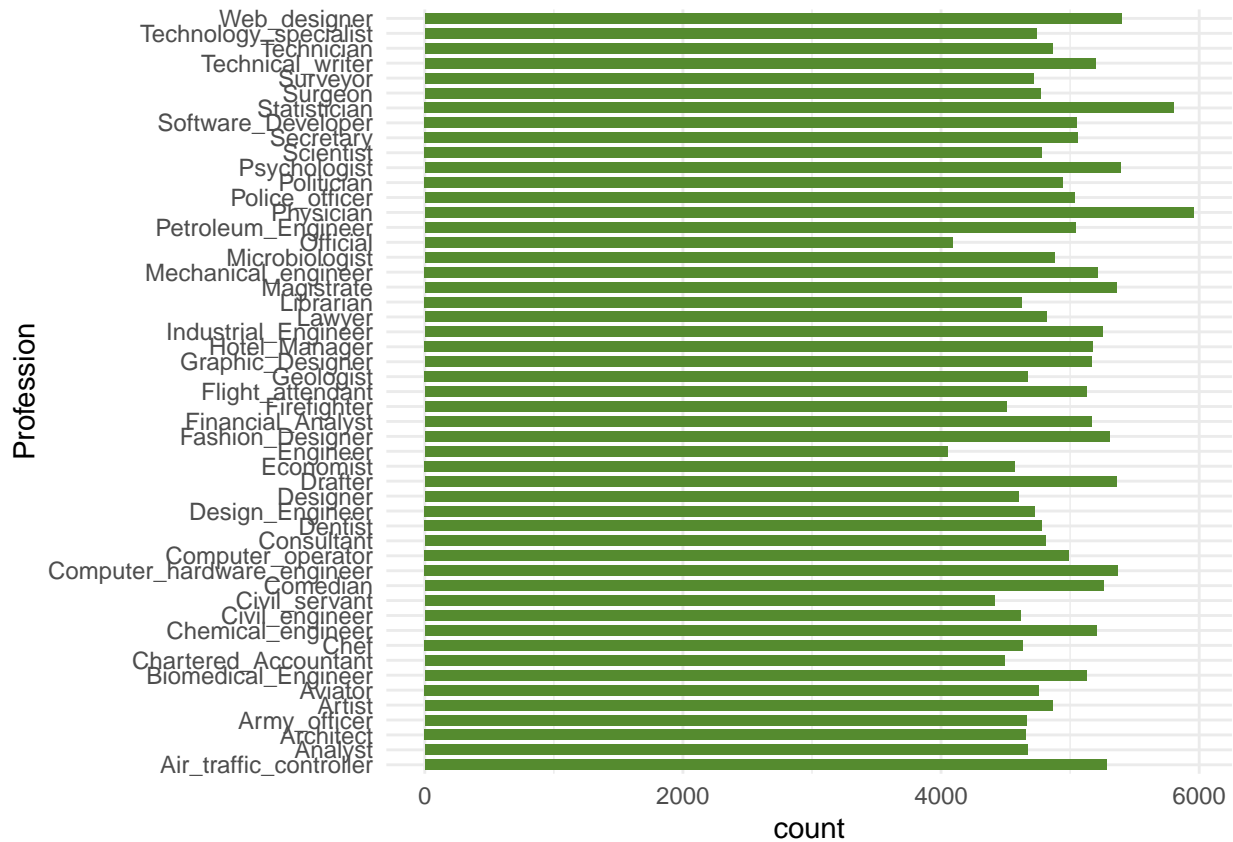


Chart 3

This is a bar chart to show numbers by House Owner ship.

```
ggplot(data, aes(x=factor(House_Ownership)))+
  geom_bar(width=0.5, fill="#80CBC4")+
  theme_minimal()
```

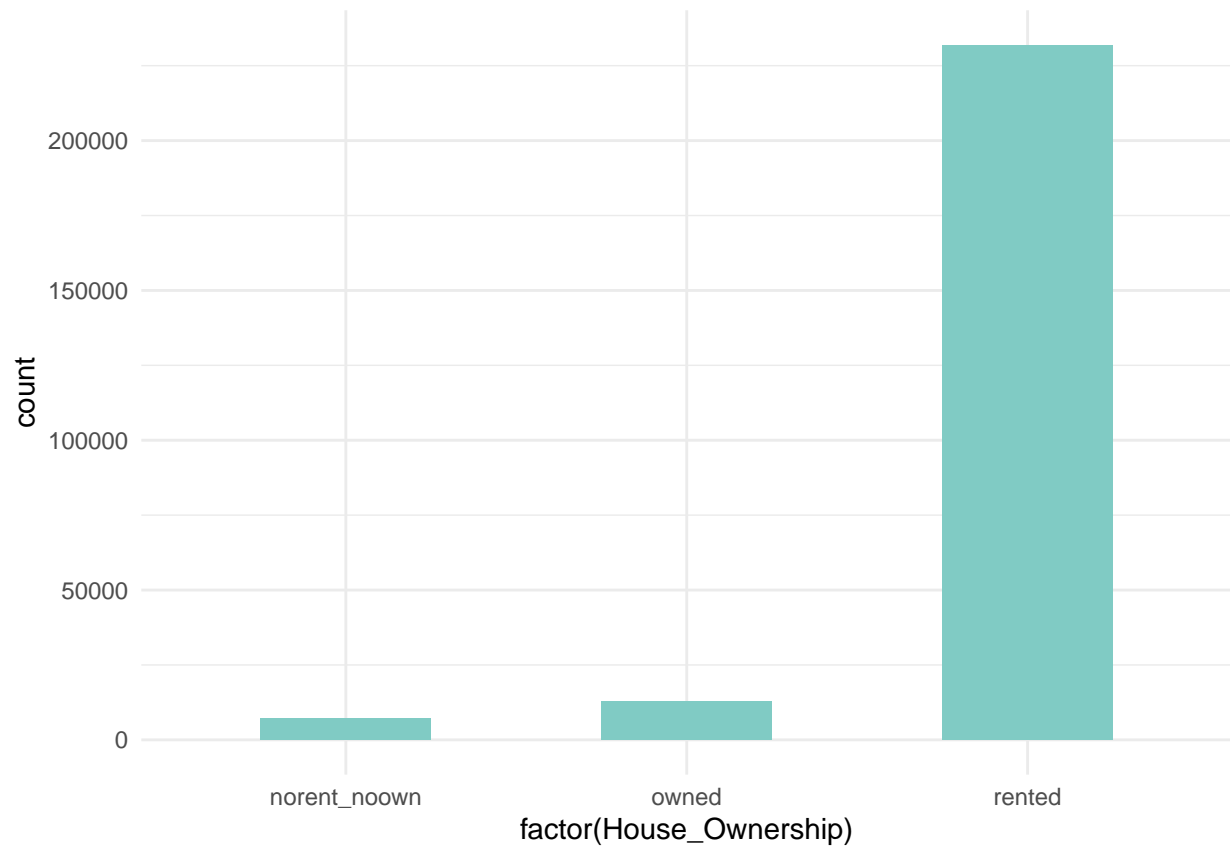


Chart 4

This is a boxplot to show Income distributin according Age group.

```
set.seed(99)
ggplot(sample_n(data, 1000),
  aes(Age_Group, Income, fill=Age_Group)) +
  geom_boxplot(width=0.3) +
  scale_fill_brewer(palette="BuPu") +
  labs(title = "User Income according Age Group",
    x = "Age_Group",
    y = "Income (USD)")
```

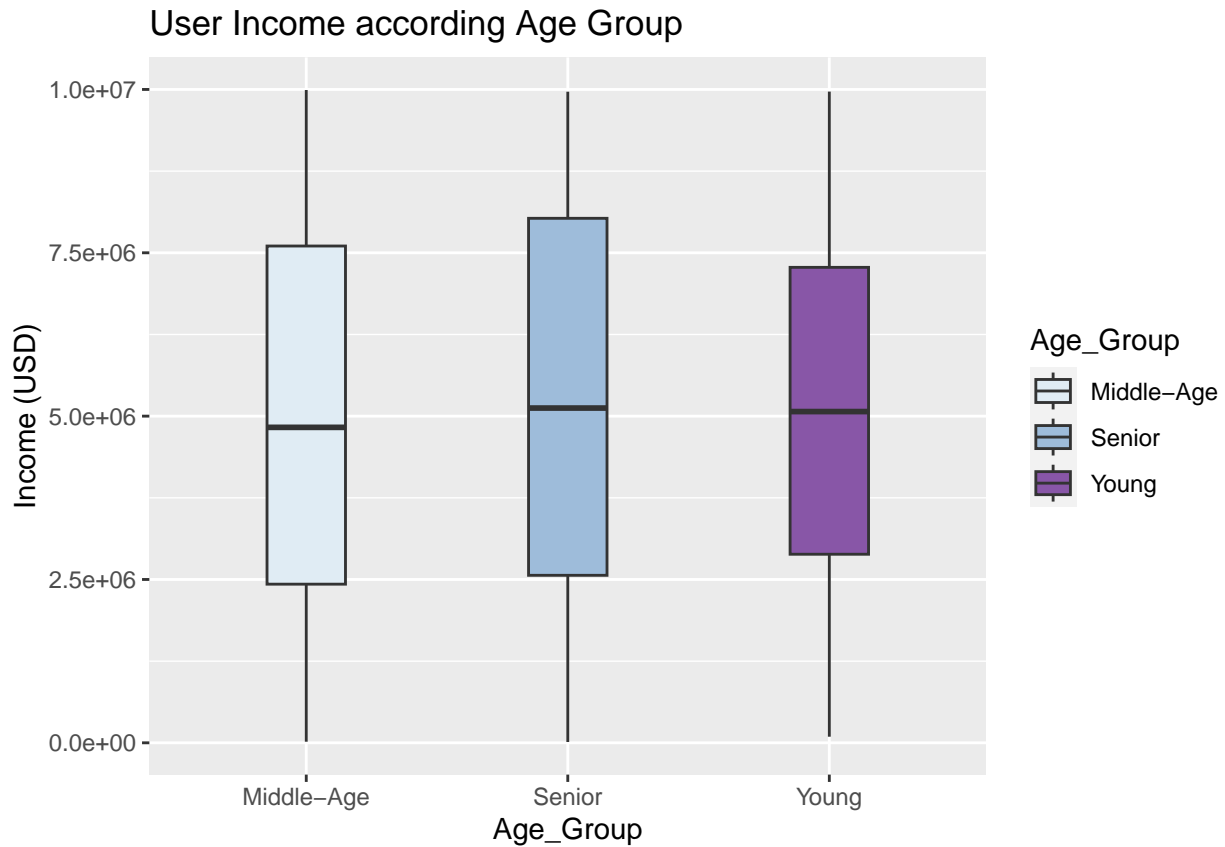


Chart 5 This is a violin plot to show the relationship between distributions of data and Income according Region. #

```
set.seed(55)
ggplot(sample_n(data, 1000),
  aes(Region, Income)) +
  geom_violin(color="#6495ED", fill="#CCCCFF", alpha=0.3) +
  geom_jitter(width=0.05,color="#0000FF", alpha=0.3 ) +
  labs(x = "Region",
    y = "Income (USD)")
```

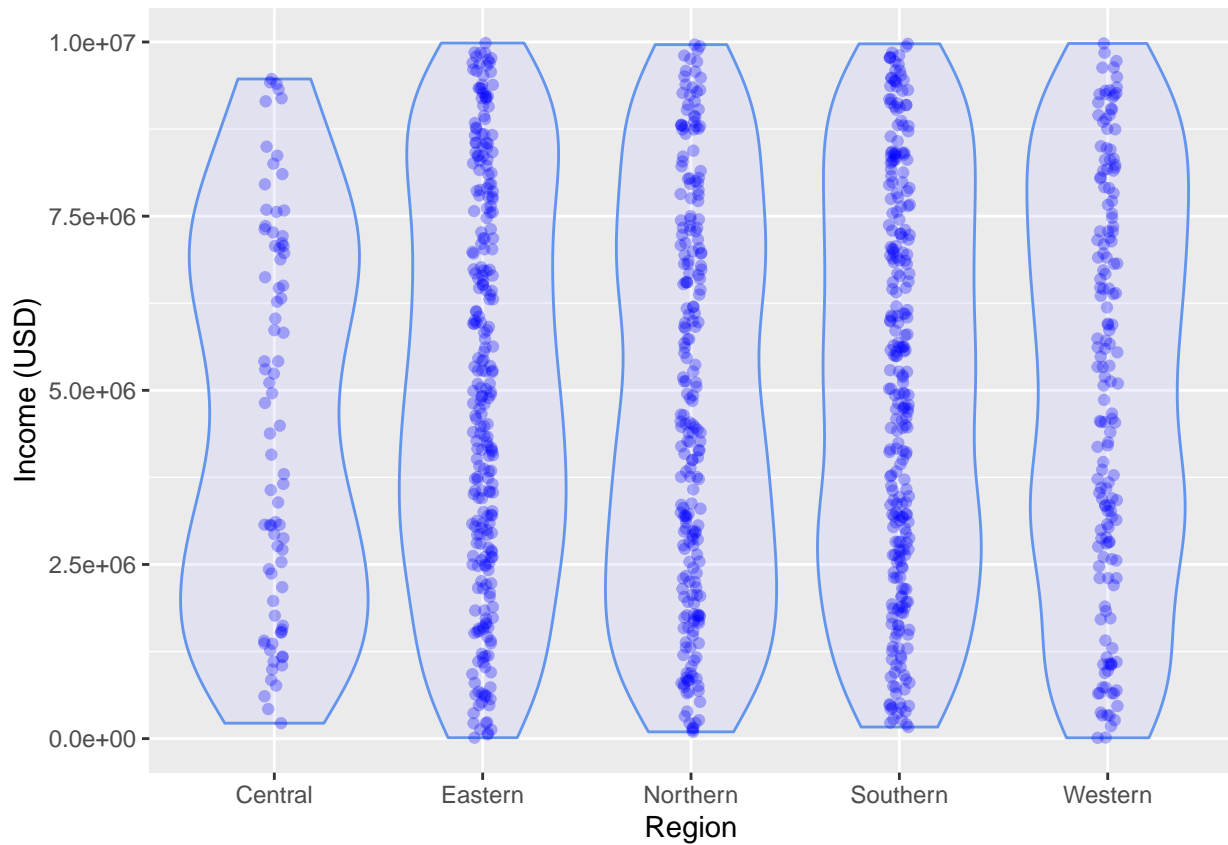


Chart 6 This is a stacked bar chart to show clarity distributin.

```
ggplot(sample_n(data, 5000), aes(Region, fill=Income_Category)) +
  geom_bar(position = "fill") +
  scale_fill_brewer(direction = -1) +
  theme_minimal() +
  labs(title = " ",
       x = "Region",
       y = "Income_Category")
```

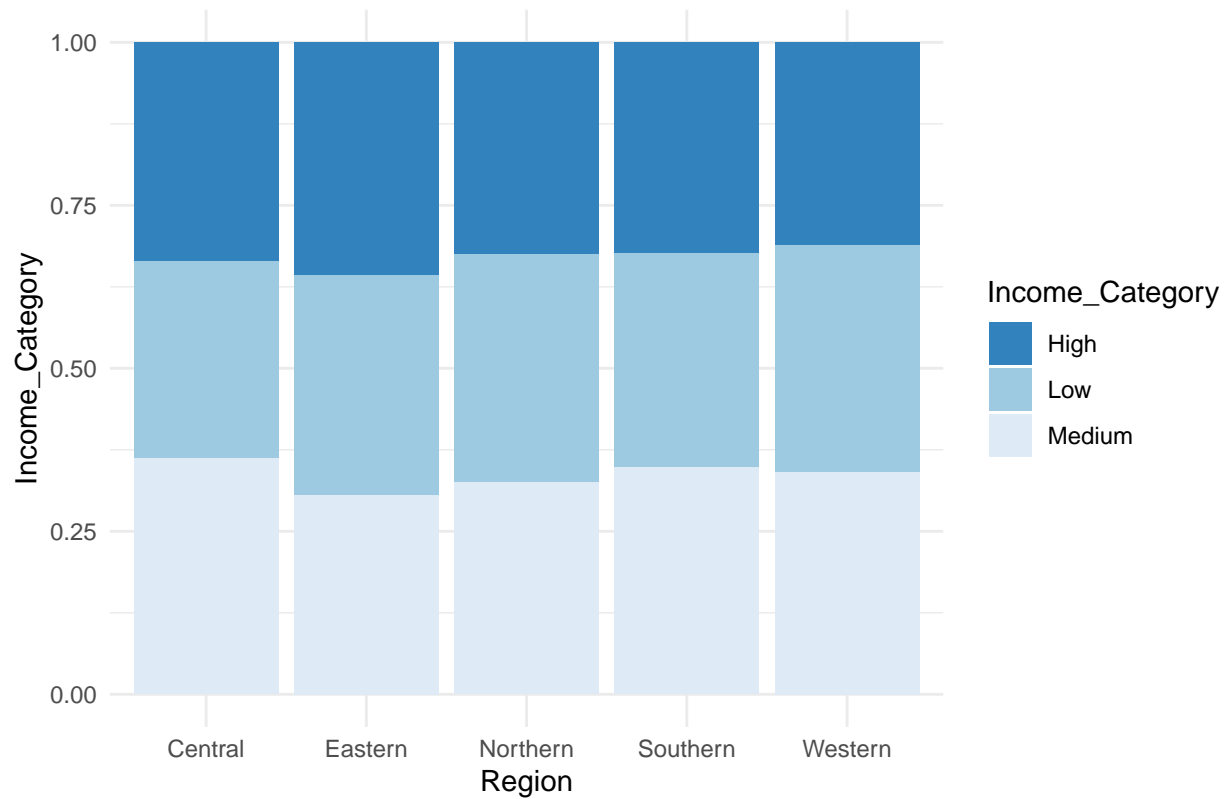


Chart 7 This is a scatter plot to show the relationship between Age and Income according Risk Flag. #

```
set.seed(19)
ggplot(sample_n(data, 1500), aes(Age, Income)) +
  geom_count(aes(color = ..n.., size = ..n..)) +
  scale_colour_gradient(low = "#E8DAEF", high = "#633974") +
  facet_wrap(~ Risk_Flag, ncol=3)
```

