Muhammad Nadeem

201980050

Data Visualization

CP-02

- **Build dataset on your own & plot the following distribution shapes:**

1st way:

Load required library

library(ggplot2)

Set seed for reproducibility

set.seed(123)

Create datasets for each distribution shape

Bimodal distribution

bimodal_data <- c(rnorm(1000, mean = -2, sd = 1), rnorm(1000, mean = 2, sd = 1))

Comb distribution

comb_data <- c(runif(1000, min = -3, max = -2), runif(1000, min = 1, max = 2))

Edge peak distribution

edge_peak_data <- c(rexp(1000, rate = 1), rexp(1000, rate = 2))

Normal distribution

normal_data <- rnorm(2000)

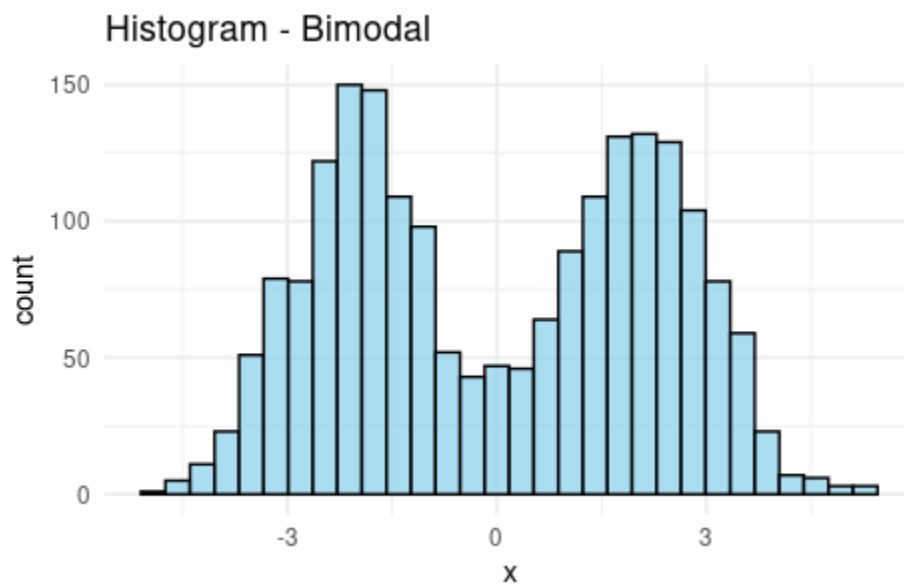# Skewed distribution

skewed_data <- c(rnorm(1000), rnorm(1000, mean = 4, sd = 1))

```r
# Uniform distribution
uniform_data <- runif(2000)
# Plot distributions
# Bimodal distribution
ggplot(data.frame(x = bimodal_data), aes(x)) +
  geom_histogram(bins = 30, fill = "skyblue", color = "black", alpha = 0.7) +
  labs(title = "Histogram - Bimodal") +
  theme_minimal()
```
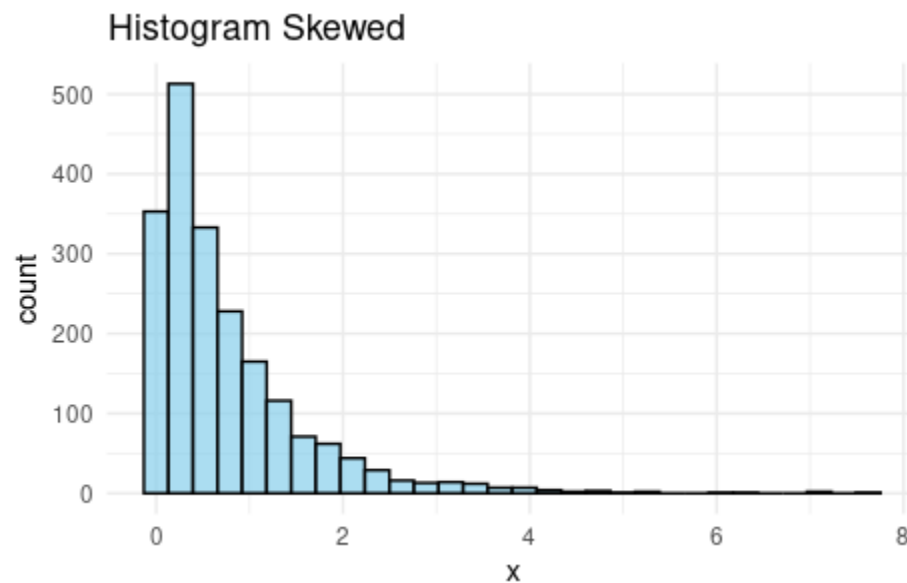

Histogram - Bimodal

```r
ggplot(data.frame(x = bimodal_data), aes(x)) +
  geom_density(fill = "skyblue", color = "black") +
  labs(title = "Density Plot - Bimodal") +
  theme_minimal()
```
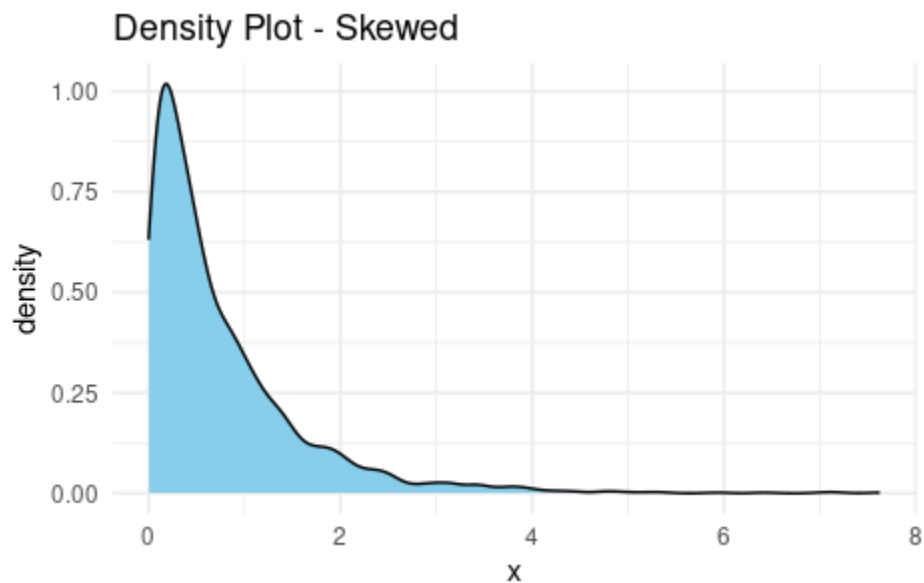
Density Plot - Bimodal
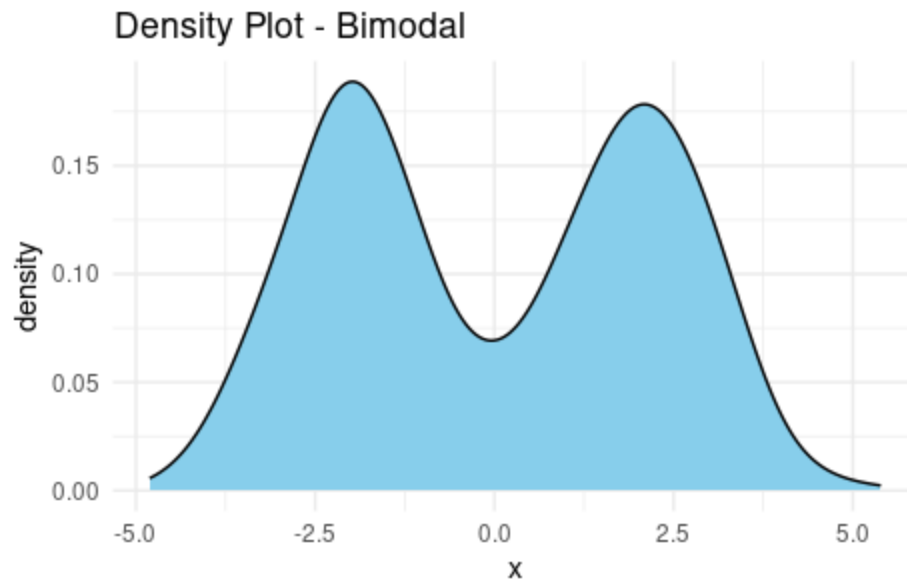
Edge peak distribution

```
ggplot(data.frame(x = edge_peak_data), aes(x)) +
  geom_histogram(bins = 30, fill = "skyblue", color = "black", alpha = 0.7) +
  labs(title = "Histogram Skewed") +
  theme_minimal()
```
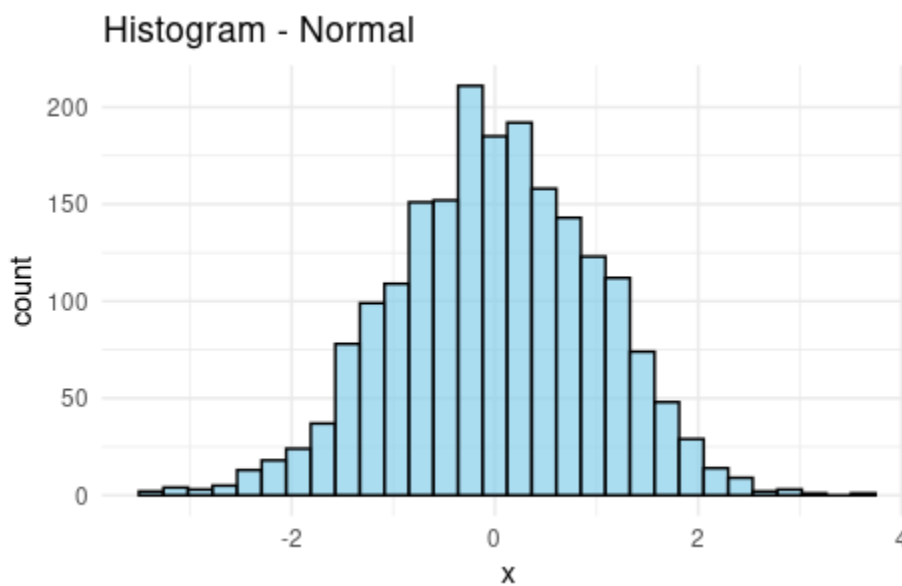

Histogram Skewed

```
ggplot(data.frame(x = edge_peak_data), aes(x)) +
  geom_density(fill = "skyblue", color = "black") +
  labs(title = "Density Plot - Skewed") +
```
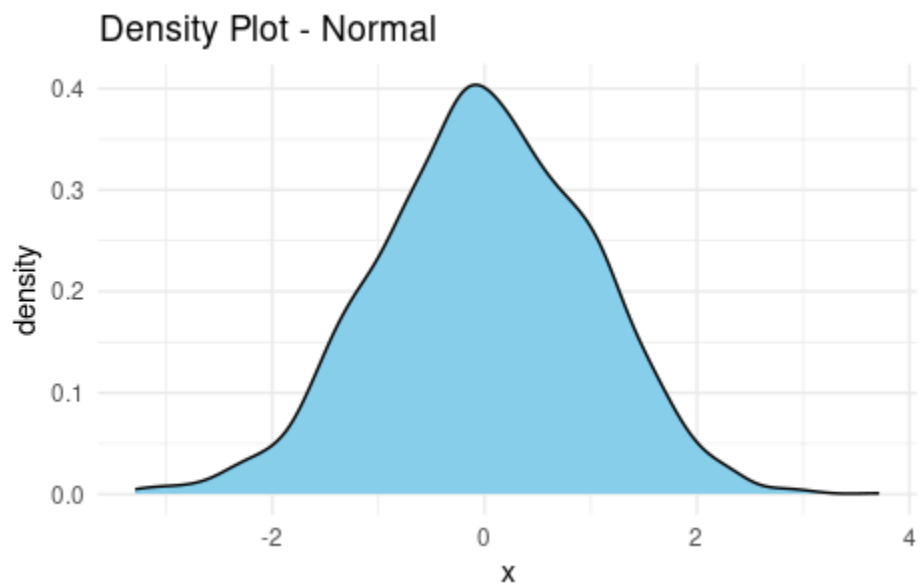
theme_minimal()

Density Plot - Skewed

Normal distribution

ggplot(data.frame(x = normal_data), aes(x)) +

  geom_histogram(bins = 30, fill = "skyblue", color = "black", alpha = 0.7) +

  labs(title = "Histogram - Normal") +

  theme_minimal()

Histogram - Normal

ggplot(data.frame(x = normal_data), aes(x)) +

  geom_density(fill = "skyblue", color = "black") +

```
labs(title = "Density Plot - Normal") +

theme_minimal()
```

Density Plot - Normal


```
# Uniform distribution
ggplot(data.frame(x = uniform_data), aes(x)) +
  geom_histogram(bins = 30, fill = "skyblue", color = "black", alpha = 0.7) +
  labs(title = "Histogram - Uniform") +
  theme_minimal()
```
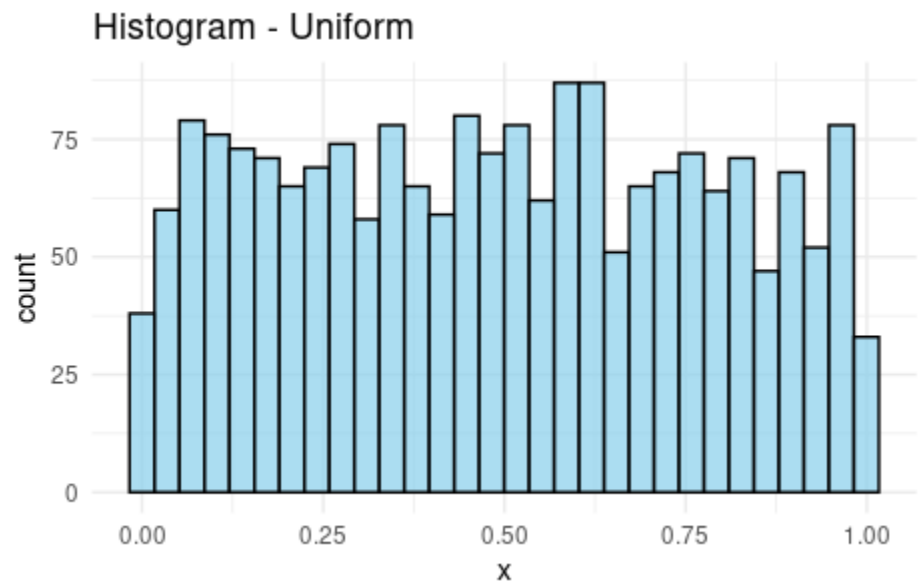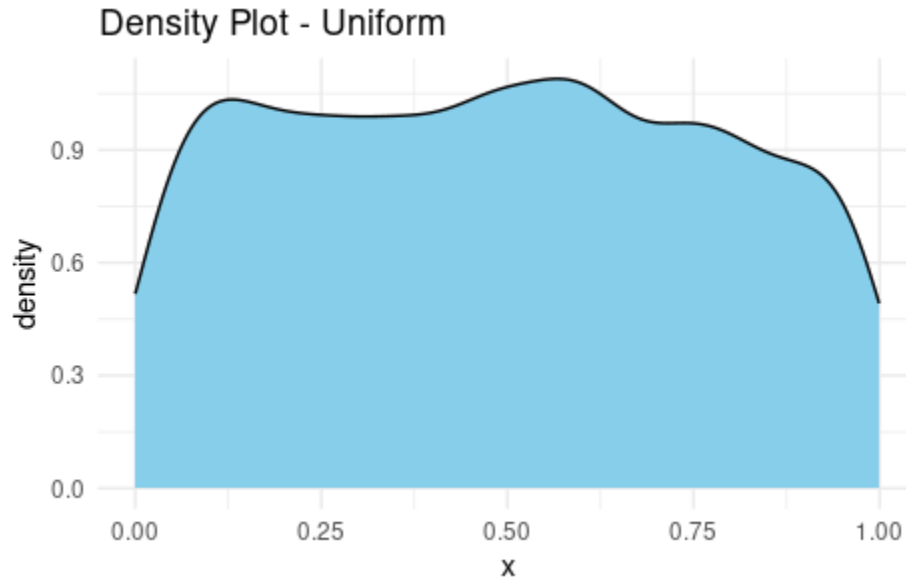
Histogram - Uniform


```
ggplot(data.frame(x = uniform_data), aes(x)) +
```

```
geom_density(fill = "skyblue", color = "black") +

labs(title = "Density Plot - Uniform") +

theme_minimal()
```


Density Plot - Uniform

2ⁿᵈ Way:

```
# Build dataset with different distributions
data <- data.frame(

  type = c( rep("edge peak", 1000), rep("comb", 1000), rep("normal", 1000), rep("uniform",
1000), rep("bimodal", 1000), rep("skewed", 1000) ),

  value = c( rnorm(900), rep(3, 100), rnorm(360, sd=0.5), rep(c(-1,-0.75,-0.5,-
0.25,0,0.25,0.5,0.75), 80), rnorm(1000), runif(1000), rnorm(500, mean=-2), rnorm(500,
mean=2), abs(log(rnorm(1000))) )

)


# Represent it
data %>%

  ggplot( aes(x=value)) +

  geom_density(fill="#69b3a2", color="#e9ecef", alpha=0.9, adjust = 0.5) +

  facet_wrap(~type, scale="free") +
```
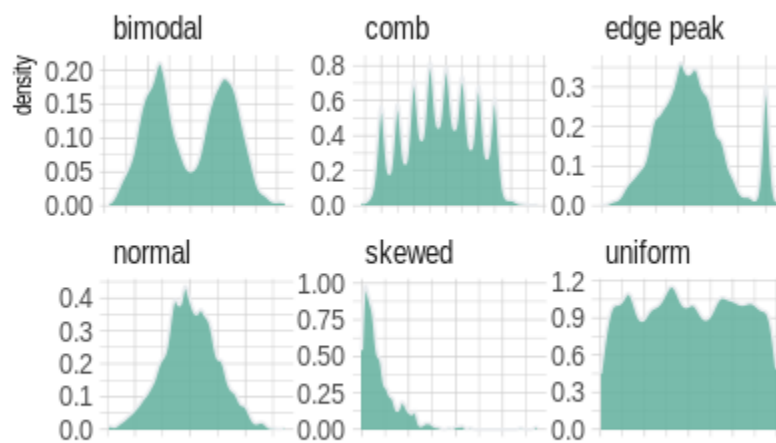
```
  theme_ipsum() +

  theme(

    panel.spacing = unit(0.1, "lines"),

    axis.title.x=element_blank(),

    axis.text.x=element_blank(),

    axis.ticks.x=element_blank()

  )
```


- **Plot histogram or density plot with your own choice of data and play with its bin-width to See the difference.**
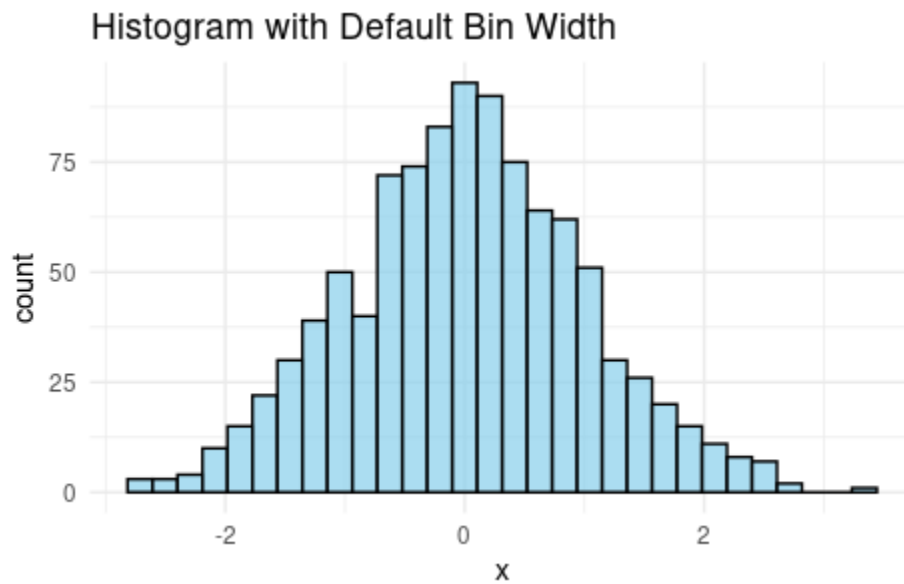
Task 2

Generate random data

set.seed(123)

data <- rnorm(1000)

# Load required library

library(ggplot2)

Plot histogram with default bin width

```
ggplot(data.frame(x = data), aes(x)) +
  geom_histogram(bins = 30, fill = "skyblue", color = "black", alpha = 0.7) +
  labs(title = "Histogram with Default Bin Width") +
  theme_minimal()
```

Histogram with Default Bin Width

Plot histogram with smaller bin width

```
ggplot(data.frame(x = data), aes(x)) +
  geom_histogram(bins = 50, fill = "lightgreen", color = "black", alpha = 0.7) +
  labs(title = "Histogram with Smaller Bin Width") +
  theme_minimal()
```

Histogram with Smaller Bin Width
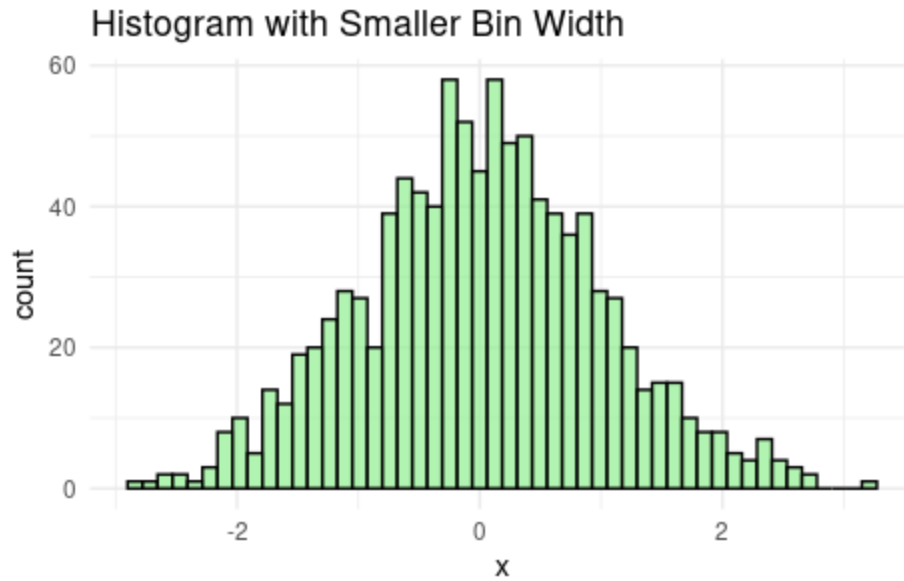
Plot histogram with larger bin width

ggplot(data.frame(x = data), aes(x)) +

  geom_histogram(bins = 20, fill = "salmon", color = "black", alpha = 0.7) +

  labs(title = "Histogram with Larger Bin Width") +

  theme_minimal()

Histogram with Larger Bin Width