EXPT NO:4	Visual encoding of data
DATE: 20.01.2026	

PRE-LAB QUESTIONS

1) How does the human visual system process visual variables?

Humans **detect visual features quickly (pre-attentively)** like **color, size, position, and shape** before thinking.
So we understand patterns fastest when data is shown using **clear visual cues**.

2) Why is improper encoding misleading in analytics?

Wrong encoding can **hide truth or exaggerate results**.
Example: using similar colors for different severities may look "same", causing **wrong decisions**.

3) What are perceptual limitations in visualization?

Common limitations are:

- Hard to compare **angles/areas** accurately (pie charts, bubbles)
- Too many colors confuse users
- Small differences in shades may not be noticed
- Overloaded charts reduce clarity **4) How does color choice affect**

interpretation?

Color strongly changes meaning:

- **Red = danger / high severity**
- **Green = safe / low severity**

Bad color selection can cause confusion, **especially for color-blind users** or if contrast is low.

5) Why should AI dashboards be perceptually optimized?

Because AI dashboards are used for **fast decision-making**.
If visuals are not clear:

- users may **misread predictions**
- high-risk areas may be ignored
- wrong actions may be taken

So dashboards must be **simple, readable, and accurate visually**.

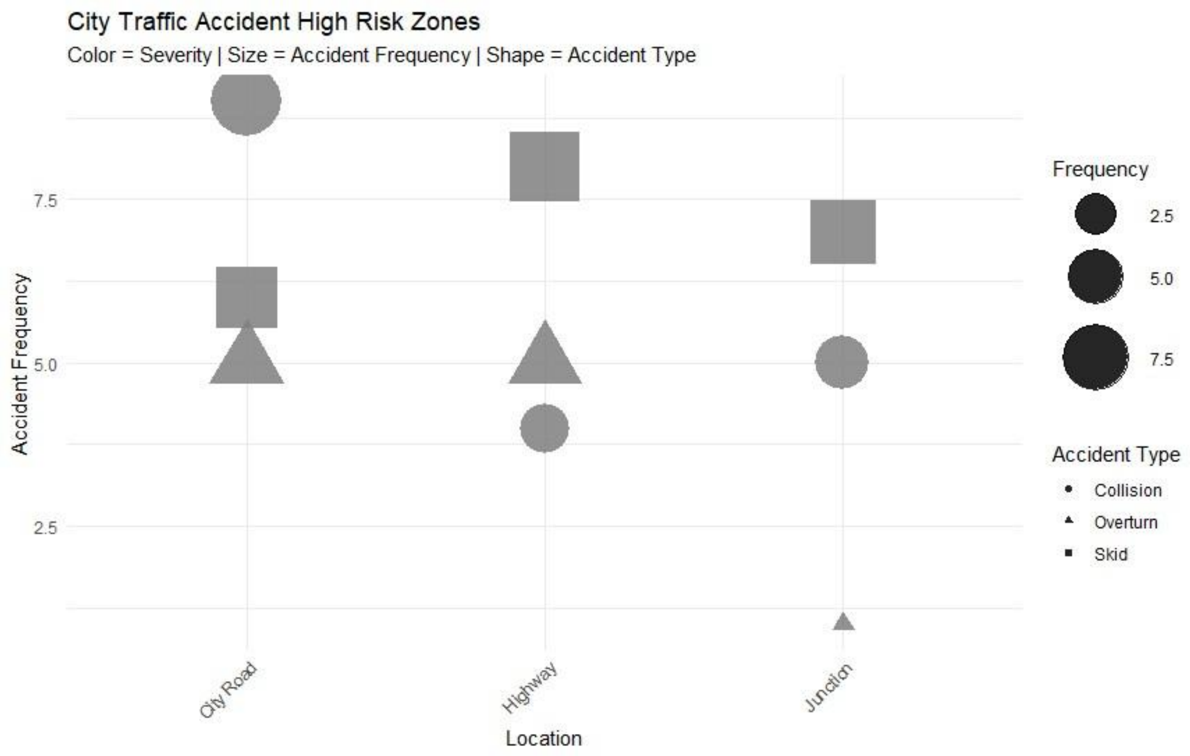
OBJECTIVE: To apply effective visual encoding principles for meaningful data communication.
SCENARIO: A city traffic department analyzes accident data to identify high-risk zones and accident severity patterns.
IN-LAB TASKS (Using R Language) • Encode severity using color gradients • Represent accident frequency using size • Use shape to indicate accident type

CODE:

```
# ==============================================
# Roll no: 23BAD101
# ==============================================
# TRAFFIC ACCIDENT VISUAL ENCODING (R)
# Dataset Name: Traffic_accidents
# Severity  -> Color Gradient
# Frequency -> Size
# Accident Type -> Shape
# ==============================================
library(ggplot2)
library(dplyr)
colnames(Traffic_accidents)
Traffic_accidents$Severity <- as.numeric(Traffic_accidents$Severity)
acc_summary <- Traffic_accidents %>%
  group_by(Location, Accident_Type) %>%
  summarise(
    Accident_Frequency = n(),                # Size Encoding
    Avg_Severity = mean(Severity, na.rm = TRUE), # Color Encoding
    .groups = "drop"
  )

ggplot(
  acc_summary,
  aes(
    x = Location,
    y = Accident_Frequency
  )
) +
  geom_point(
    aes(
      color = Avg_Severity,
      size  = Accident_Frequency,
      shape = Accident_Type
    ),
    alpha = 0.85
  ) +
  scale_color_gradient(
    low  = "lightgreen",
    high = "red"
  ) +
  scale_size(
    range = c(4, 18)
  ) +
  labs(
    title    = "City Traffic Accident High Risk Zones",
    subtitle = "Color = Severity | Size = Accident Frequency | Shape = Accident Type",
    x        = "Location",
    y        = "Accident Frequency",
    color    = "Average Severity",
    size     = "Frequency",
    shape    = "Accident Type"
  ) +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1)
  )
```

OUTPUT:



POST-LAB QUESTIONS

1) Which visual variable best conveys severity and why?

Color gradient is best for severity because it **instantly shows intensity levels** (low → high) and is easy to compare across zones.

2) How can poor encoding mislead policy decisions?

Poor encoding can show **wrong high-risk areas**, so officials may:

- place traffic signals in wrong zones
- ignore real accident hotspots
- waste budget and manpower

  This leads to **incorrect safety policies**.

3) What precautions should be taken for color-blind users?

- Use **color-blind friendly palettes** (avoid only red–green)
- Add **labels/legends**
- Use **shape + size along with color** (not color alone)
- Ensure **high contrast** between levels

4) How does visual encoding affect AI explainability?

Good encoding makes AI results **easy to understand and trust**.
Example: If severity is clearly shown using colors, users can **quickly relate AI predictions to real-world risk**, improving explainability.

5) Suggest improvements for large-scale public dashboards.

- Add **filters** (location, date, severity, accident type)
- Use **interactive maps** with zoom
- Show **top hotspots list**

LEARNING OUTCOME: Students learn perceptually effective visual communication for AI systems.

ASSESSMENT

Description	Max Marks	Marks Awarded
Pre Lab Exercise	**5**	
In Lab Exercise	**10**	
Post Lab Exercise	**5**	
Viva	**10**	
Total	**30**	
Faculty Signature		