EXPT NO:4	Visual encoding of data
DATE: 20.01.2026	

PRE-LAB QUESTIONS

1. How does the human visual system process visual variables?
2. Why is improper encoding misleading in analytics?
3. What are perceptual limitations in visualization?
4. How does color choice affect interpretation?
5. Why should AI dashboards be perceptually optimized?

ANSWERS:

1. The human visual system processes visual variables such as color, size, shape, and position by quickly detecting patterns, contrasts, and differences, allowing rapid interpretation of visual information.
2. Improper encoding is misleading in analytics because it can distort data perception, exaggerate or hide patterns, and lead to incorrect conclusions or decisions.
3. Perceptual limitations in visualization include limited ability to distinguish many colors or shapes, difficulty comparing areas or volumes, and reduced accuracy when data is cluttered or complex.
4. Color choice affects interpretation by influencing attention, readability, and emotional response; poor color selection can cause confusion, misinterpretation, or accessibility issues.
5. AI dashboards should be perceptually optimized to ensure accurate, efficient, and unbiased interpretation of data, enabling users to make reliable and informed decisions.

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.

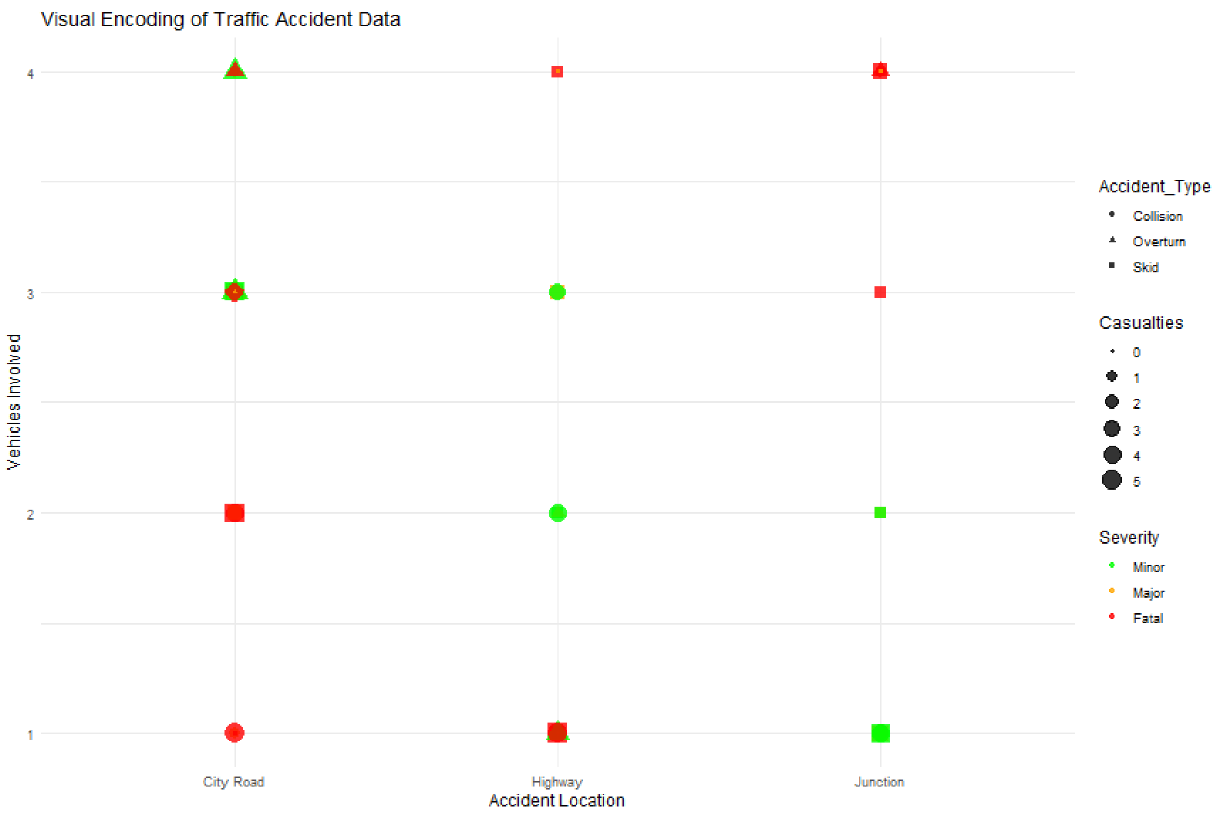
IN-LAB TASKS (Using R Language) • Encode severity using color gradients • Represent accident frequency using size • Use shape to indicate accident type

CODE:

```r
"Student Name; Sai Vaishnavi R"
"Roll no: 23BAd094"# Load libraries
library(ggplot2)
library(dplyr)

# Load traffic accident dataset
traffic <- read.csv("C:/Users/student/Downloads/4.traffic_accidents.csv")

# View dataset structure
str(traffic)
head(traffic)

# Convert Severity to ordered factor
traffic$Severity <- factor(
  traffic$Severity,
  levels = c("Minor", "Major", "Fatal"),
  ordered = TRUE
)

# Visual Encoding Plot
ggplot(
  traffic,
  aes(
    x = Location,
    y = Vehicles_Involved,
    color = Severity,
    size = Casualties,
    shape = Accident_Type
  )
) +
  geom_point(alpha = 0.8) +
  scale_color_manual(values = c("green", "orange", "red")) +
  labs(
    title = "Visual Encoding of Traffic Accident Data",
    x = "Accident Location",
    y = "Vehicles Involved"
  ) +
  theme_minimal()

```

OUTPUT:

```
[1] "Student Name; Sai Vaishnavi R"
> "Roll no: 23BAd094"# Load libraries
[1] "Roll no: 23BAd094"
> library(ggplot2)
> library(dplyr)
>
> # Load traffic accident dataset
> traffic <- read.csv("C:/Users/student/Downloads/4.traffic_accidents.csv")
>
> # View dataset structure
> str(traffic)
'data.frame':	50 obs. of  8 variables:
 $ Accident_ID      : int  7001 7002 7003 7004 7005 7006 7007 7008 7009 7010 ...
 $ Location         : chr  "Highway" "Highway" "Highway" "Junction" ...
 $ Date             : chr  "2024-06-01" "2024-06-02" "2024-06-03" "2024-06-04" ...
 $ Time             : chr  "Evening" "Evening" "Night" "Afternoon" ...
 $ Accident_Type    : chr  "Overturn" "Skid" "Overturn" "Collision" ...
 $ Severity         : chr  "Minor" "Minor" "Fatal" "Major" ...
 $ Vehicles_Involved: int  1 1 4 1 1 1 1 3 4 2 ...
 $ Casualties       : int  0 1 0 3 2 1 0 4 1 1 ...
> head(traffic)
  Accident_ID Location       Date      Time Accident_Type Severity Vehicles_Involved Casualties
1        7001  Highway 2024-06-01   Evening      Overturn    Minor                 1          0
2        7002  Highway 2024-06-02   Evening          Skid    Minor                 1          1
3        7003  Highway 2024-06-03     Night      Overturn    Fatal                 4          0
4        7004 Junction 2024-06-04 Afternoon     Collision    Major                 1          3
5        7005  Highway 2024-06-05   Evening          Skid    Minor                 1          2
6        7006  Highway 2024-06-06 Afternoon      Overturn    Minor                 1          1
>
> # Convert Severity to ordered factor
> traffic$Severity <- factor(
+   traffic$Severity,
+   levels = c("Minor", "Major", "Fatal"),
+   ordered = TRUE
+ )
>
> # Visual Encoding Plot
> ggplot(
+   traffic,
+   aes(
+     x = Location,
+     y = Vehicles_Involved,
+     color = Severity,
+     size = Casualties,
+     shape = Accident_Type
+   )
+ ) +
+   geom_point(alpha = 0.8) +
+   scale_color_manual(values = c("green", "orange", "red")) +
+   labs(
+     title = "Visual Encoding of Traffic Accident Data",
+     x = "Accident Location",
+     y = "Vehicles Involved"
+   ) +
+   theme_minimal()
>
```


Visual Encoding of Traffic Accident Data
Vehicles Involved
4
3
2
1
City Road
Highway
Junction
Accident Location
Accident_Type
Collision
Overturn
Skid
Casualties
0
1
2
3
4
5
Severity
Minor
Major
Fatal

POST-LAB QUESTIONS

1. Which visual variable best conveys severity and why?
2. How can poor encoding mislead policy decisions?
3. What precautions should be taken for color-blind users?
4. How does visual encoding affect AI explainability?
5. Suggest improvements for large-scale public dashboards.

ANSWERS:

1. Color is the most effective visual variable for conveying severity because distinct colors such as green, orange, and red allow users to immediately differentiate between minor, major, and fatal accidents without additional effort.
2. Poor visual encoding can mislead policy decisions by hiding critical accident patterns, underrepresenting fatal incidents, and causing incorrect identification of high-risk locations, which may result in improper resource allocation.
3. For color-blind users, visuals should not rely only on color; instead, color-blind-safe palettes, additional encodings like shapes or labels, and clear legends should be used to ensure accessibility.
4. Visual encoding improves AI explainability by clearly showing how variables like severity, casualties, and location influence outcomes, making AI-driven insights easier to interpret and trust.
5. Large-scale public dashboards can be improved by adding interactive filters, using aggregated visualizations such as heatmaps, applying accessible color schemes, incorporating time-based analysis, and clearly highlighting high-risk areas.

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	