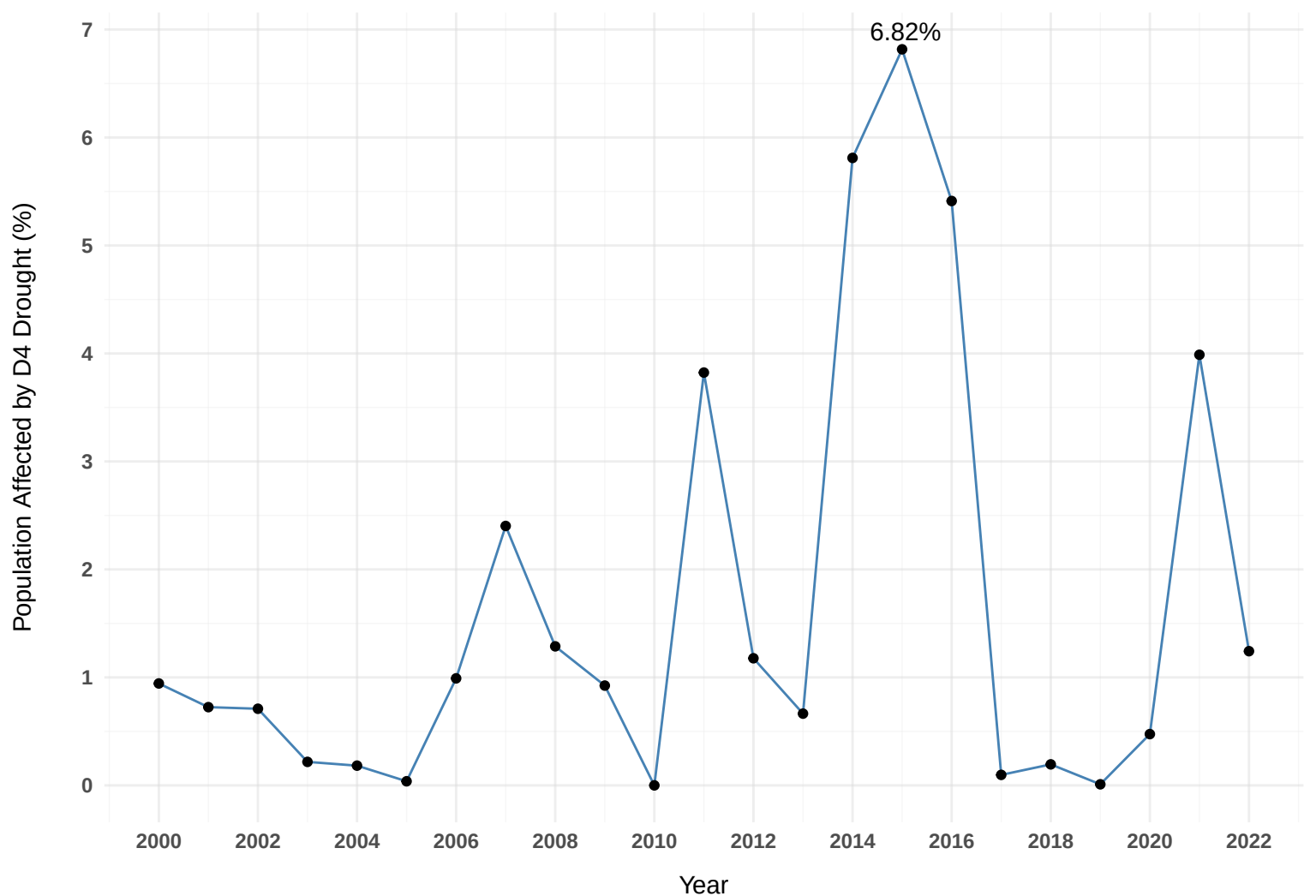
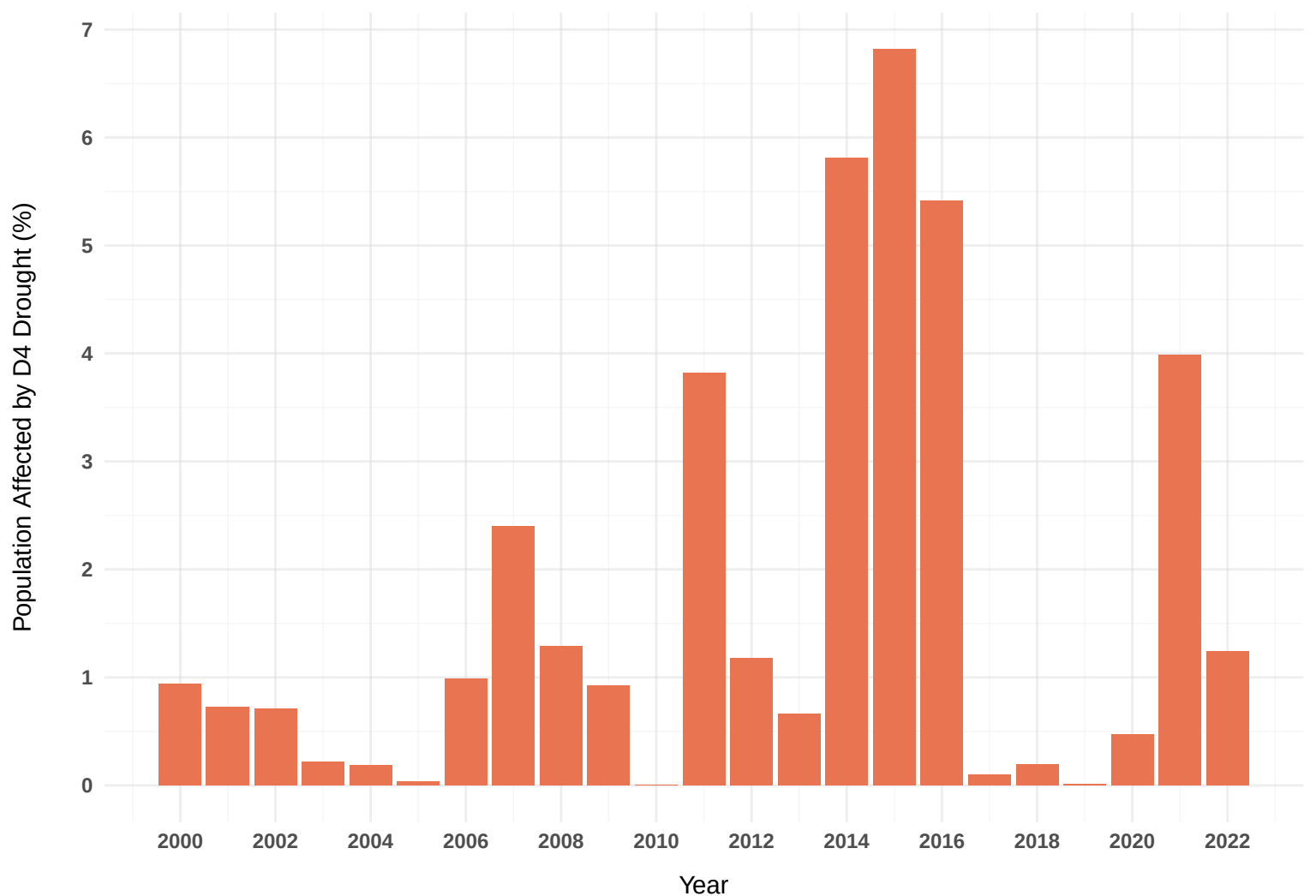
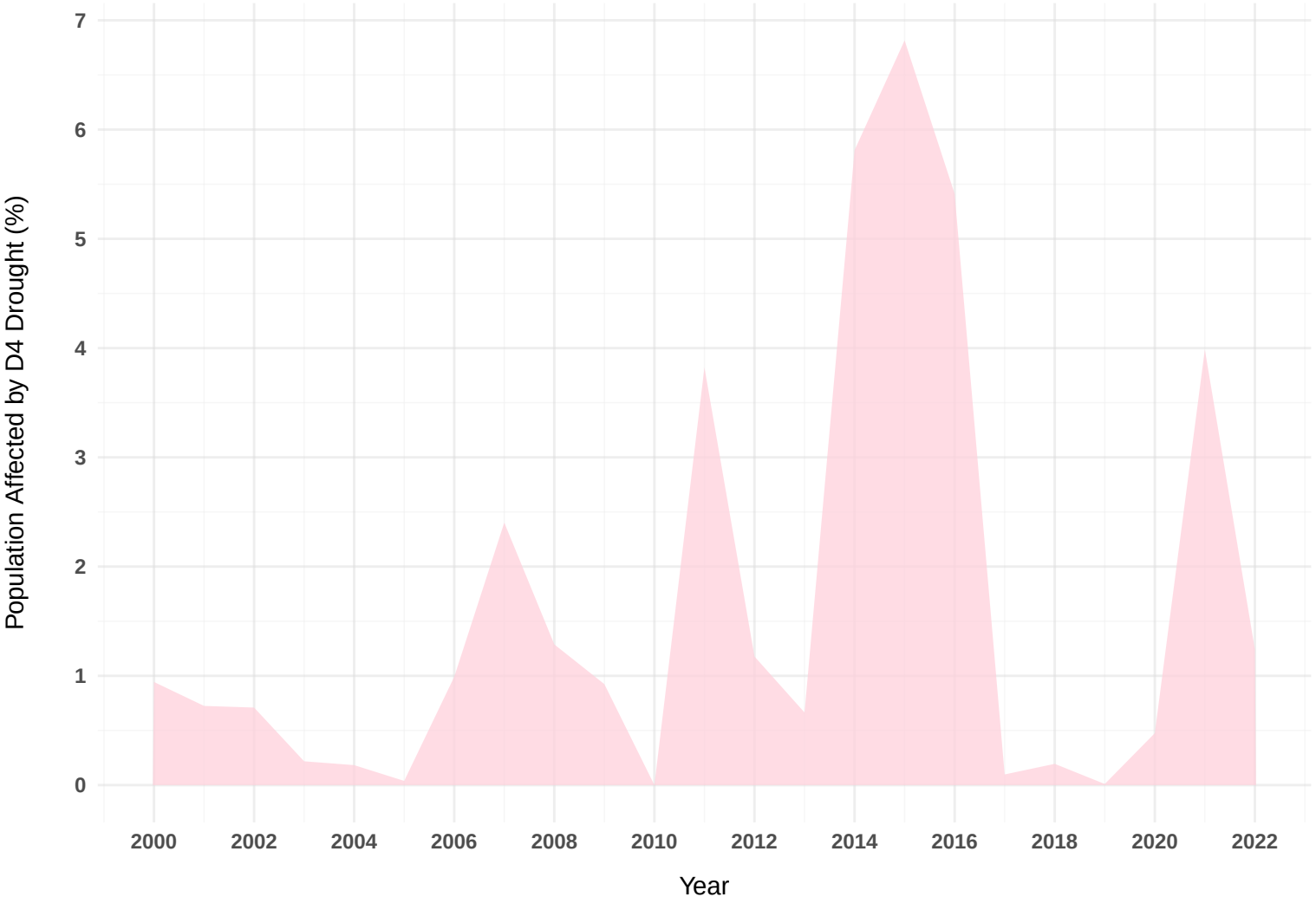


Figure−1: This line plot shows a fluctuating trend in the % of the U.S. population affected by exceptional drought conditions. Notably, there was a significant peak in 2012−14 where the impact reached a high of 6.82%, indicating the most severe drought period, whereas the years 2000 and 2006 saw comparatively lower impact, over the past two decades.



Figure−2: This bar chart displays the biennial avg % of the U.S. population affected by exceptional drought from 2000 to 2023. Peaks in the chart, especially during 2014−16, highlight the years with the most severe impact on the population. The data reveals significant fluctuations over time, underscoring the variable nature of drought conditions.



Figure−3: This area plot gives clear visual representation of the biennial change in % of the U.S. population affected by exceptional drought from 2000 to 2023,highlighting the variability over the past two decades. The filled area indicates the prevalence of drought conditions, with peaks such as the one in 2014−16 highlighting times when the drought was most widespread.

```
# Load necessary libraries
library(ggplot2)
library(dplyr)
library(lubridate)
library(scales)
library(pdftools)

# Read the data
data <- read.csv("E:/MSDS/SEMESTER-1/INFO-526-DATA VIZ & ANALYSIS-MEGHEAN/Wk-4- Advanced Plots/ASSIGNMENT-2-Multiple Plots/Drought Monitor Edited.csv")

# Inspect the data
head(data) # to view the first few rows.
str(data) # function is used to give the structure
summary(data) # function provides summary statistics
unique(data$Type) # using unique function wherever there are categorical items to know the types

# Clean the data
# Count the number of duplicated rows
sum(duplicated(data))
# Check for missing values
any(is.na(data))
# Count the number of NA rows
sum(is.na(data))

# Rename the columns for better understanding
data <- data %>%
  rename(
    `Abnormal_Dry` = D0,
    `Moderate_Drought` = D1,
    `Severe_Drought` = D2,
    `Extreme_Drought` = D3,
    `Exceptional_Drought` = D4
  )

# Convert MapDate to a Date type and extract the Year
data <- data %>%
  mutate(
    MapDate = as.Date(as.character(MapDate), format = "%Y%m%d"),
    Year = as.numeric(format(MapDate, "%Y"))
  )

# Data Summarization
drought_summary <- data %>%
  filter(Type == "Percent Population") %>%
  mutate(Year_Group = as.numeric(format(MapDate, "%Y"))) %>%
  group_by(Year_Group) %>%
  summarize(Exceptional_Drought = mean(Exceptional_Drought, na.rm = TRUE)) %>%
  ungroup()
```

```
# Find the row with the maximum Exceptional_Drought value
max_value <- drought_summary[which.max(drought_summary$Exceptional_Drought), ]

# Create a line plot for Exceptional Drought Percent Population
line_plot <- ggplot(drought_summary, aes(x = Year_Group, y = Exceptional_Drought)) +
  geom_line(color = "steelblue", group = 1) +
  geom_point() +
  geom_text(data = max_value, aes(label = sprintf(" %.2f%%", Exceptional_Drought)),
            vjust = -0.5, hjust = 0.5) +
  theme_minimal() +
  scale_x_continuous(breaks = seq(2000, 2023, by = 2)) + # Set x-axis breaks to every 2 years
  scale_y_continuous(breaks = seq(0, 7, by = 1)) + # Set y-axis breaks from 0 to 7
  labs(x = "Year", y = "Population Affected by D4 Drought (%)",
       caption = "Figure-1: This line plot shows a fluctuating trend in the % of the U.S. population affected by exceptional drought conditions. \n Notably, there was a significant peak in 2012-14 where the impact reached a high of 6.82%, indicating the most \n severe drought period, whereas the years 2000 and 2006 saw comparatively lower impact, over the past two decades.") +
  theme(
    # Make major grid lines lighter and semi-transparent
    panel.grid.major = element_line(color = "#DDDDDD7F", size = 0.5, linetype = "solid"),
    # Make minor grid lines even lighter and more transparent
    panel.grid.minor = element_line(color = "#EEEEEE7F", size = 0.25, linetype = "solid"),
    axis.text.x = element_text(angle = 0, hjust = 0.5, face = "bold"),
    axis.text.y = element_text(angle = 0, hjust = 0.5, face = "bold"),
    axis.title.y = element_text(margin = margin(r = 20, unit = "pt")),  #space between the margin and y-label
    axis.title.x = element_text(margin = margin(t = 10, unit = "pt")),  #space between the margin and x-label
    plot.caption = element_text(hjust = 0)  # Left-align the plot caption
  )



# Create an bar plot for Exceptional Drought Percent Population
bar_plot <- ggplot(drought_summary, aes(x = Year_Group, y = Exceptional_Drought, fill = Exceptional_Drought)) +
  geom_bar(stat = "identity", fill = "#E97451") + # Define the colors
  theme_minimal() +
  scale_x_continuous(breaks = seq(2000, 2023, by = 2)) + # Set x-axis breaks to every 2 years
  scale_y_continuous(breaks = seq(0, 7, by = 1)) + # Set y-axis breaks from 0 to 7
  labs(x = "Year", y = "Population Affected by D4 Drought (%)",
       caption = "Figure-2: This bar chart displays the biennial avg % of the U.S. population affected by exceptional drought from 2000 to 2023. \n Peaks in the chart, especially during 2014-16, highlight the years with the most severe impact on the population.\n The data reveals significant fluctuations over time, underscoring the variable nature of drought conditions.")+
  theme(
    # Make major grid lines lighter and semi-transparent
    panel.grid.major = element_line(color = "#DDDDDD7F", size = 0.5, linetype = "solid"),
    # Make minor grid lines even lighter and more transparent
    panel.grid.minor = element_line(color = "#EEEEEE7F", size = 0.25, linetype = "solid"),
    axis.text.x = element_text(angle = 0, hjust = 0.5, face = "bold"),
```

```
    axis.text.y = element_text(angle = 0, hjust = 0.5, face = "bold"),
    axis.title.y = element_text(margin = margin(r = 20, unit = "pt")),#space between the margin and y-label
    axis.title.x = element_text(margin = margin(t = 10, unit = "pt")), #space between the margin and x-label
    plot.caption = element_text(hjust = 0)  # Left-align the plot caption
  )


# Create an area plot for Exceptional Drought Percent Population
area_plot <- ggplot(drought_summary, aes(x = Year_Group, y = Exceptional_Drought, group = 1)) +
  geom_area(fill = "#FFD1DC", alpha = 0.75) +
  theme_minimal() +
  scale_x_continuous(breaks = seq(2000, 2023, by = 2)) + # Set x-axis breaks to every 2 years
  scale_y_continuous(breaks = seq(0, 7, by = 1)) + # Set y-axis breaks from 0 to 7
  labs(x = "Year", y = "Population Affected by D4 Drought (%)",
       caption = "Figure-3: This area plot gives clear visual representation of the biennial change in % of the U.S. population affected by \n exceptional drought from 2000 to 2023,highlighting the variability over the past two decades. The filled area indicates the prevalence \n of drought conditions, with peaks such as the one in 2014-16 highlighting times when the drought was most widespread.") +
  theme(
    # Make major grid lines lighter and semi-transparent
    panel.grid.major = element_line(color = "#DDDDDD7F", size = 0.5, linetype = "solid"),
    # Make minor grid lines even lighter and more transparent
    panel.grid.minor = element_line(color = "#EEEEEE7F", size = 0.25, linetype = "solid"),
    axis.text.x = element_text(angle = 0, hjust = 0.5, face = "bold"),
    axis.text.y = element_text(angle = 0, hjust = 0.5, face = "bold"),
    axis.title.y = element_text(margin = margin(r = 20, unit = "pt")), #space between the margin and y-label
    axis.title.x = element_text(margin = margin(t = 10, unit = "pt")), #space between the margin and x-label
    plot.caption = element_text(hjust = 0)  # Left-align the plot caption
    )


# Printing the plots
print(line_plot)
print(bar_plot)
print(area_plot)

# Save the plots to a PDF file with adjusted dimensions to fit the captions
ggsave("line_plot.pdf", line_plot, width = 8, height = 6, units = "in")
ggsave("area_plot.pdf", area_plot, width = 8, height = 6, units = "in")
ggsave("bar_plot.pdf", bar_plot, width = 8, height = 6, units = "in")

# Combine the individual PDFs into one
pdf_combine(c("line_plot.pdf", "bar_plot.pdf", "area_plot.pdf"), output = "combined_plots.pdf")
```