OhLord

Jennifer

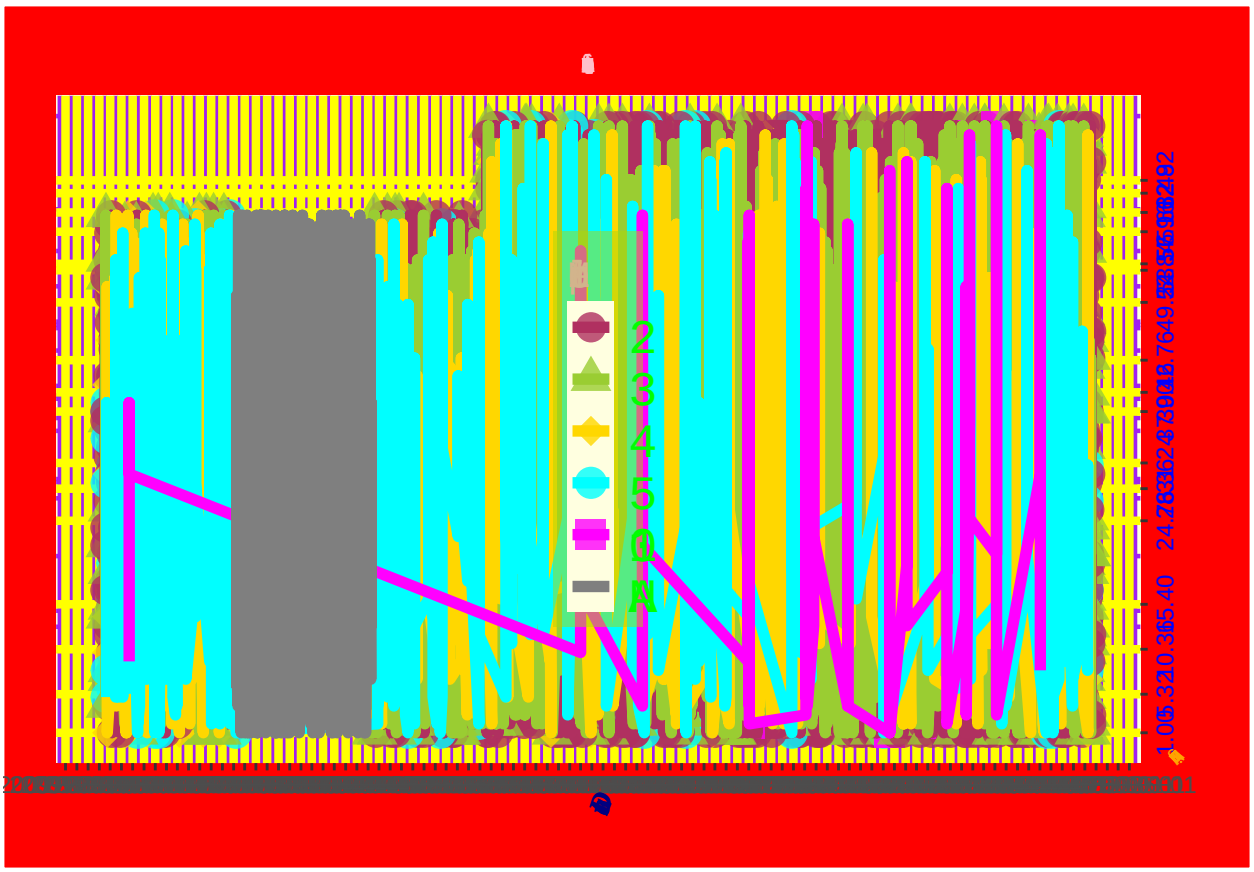
2024-10-06

This data is a set of the winning lotto numbers for Powerball that was last updated around October 4, 2024.

Yeah.

```
# Libraries
library(ggplot2)
library(dplyr)
library(lubridate)
library(extrafont)
library(tidyr)



# Read and clean the dataset
lottery_data <- read.csv("/Users/jenni/Downloads/Lottery_Powerball_Winning_Numbers__Beginning_2010.csv")




# Conversions / cleaning
lottery_data <- lottery_data %>%
  mutate(Draw.Date = mdy(Draw.Date)) %>%  # Date format
  mutate(Winning.Numbers = strsplit(as.character(Winning.Numbers), " ")) %>%
  unnest_wider(Winning.Numbers, names_sep = "_") %>%
  pivot_longer(cols = starts_with("Winning.Numbers"), names_to = "Number_Index", values_to = "Winning_Nu
  mutate(Winning_Number = as.numeric(Winning_Number))  # Numeric conversion



set.seed(1)  # Seed for reproducibility
random_breaks <- sample(seq(1, 70, by = .72), size = 17, replace = FALSE)


ggplot(lottery_data, aes(x = Draw.Date, y = Winning_Number, color = as.factor(Multiplier), shape = as.fa
  geom_point(size = 5, alpha = 0.8) +  # Winning numbers' points
  geom_line(aes(color = as.factor(Multiplier)), size = 2) +
  scale_color_manual(values = c("maroon", "yellowgreen", "gold", "cyan", "magenta")) +  # Custom colors
```

```r
  scale_shape_manual(values = c(16, 17, 18, 19, 15)) +  # Custom shapes for multipliers
  labs(
    title = "Powerball Winning Numbers Over Time (2010 - October 2024)",
    x = "Draw Date",
    y = "Winning Numbers",
    color = "Multiplier",
    shape = "Multiplier"
  ) +
  theme(
    plot.title = element_text(family = "Comic Sans MS", size = 8, face = "bold", hjust = 0.5, color = "
    axis.title.x = element_text(family = "Times New Roman", size = 11, face = "italic", angle = 72, hju
    axis.title.y.right = element_text(family = "Courier New", size = 7, angle = -134, hjust = 0.5, colo
    axis.text.y.right = element_text(angle = 90, hjust = 0.5, color = "blue"),  # Y-axis ticks rotated
    legend.title = element_text(family = "Arial Black", size = 14, color = "tan"),
    legend.text = element_text(family = "Georgia", size = 18, color = "green"),
    panel.background = element_rect(fill = "lightyellow"),
    plot.background = element_rect(fill = "red"),  # Background
    panel.grid.major = element_line(color = "yellow", size = 1.5),  # Gridlines
    panel.grid.minor = element_line(color = "purple", size = 0.8)  # Minor gridlines
  ) +
  scale_x_date(date_labels = "%Y-%m-%d", date_breaks = "2 months") +
  scale_y_continuous(breaks = random_breaks, position = "right") +  # Shift Y-axis to the right
  theme(
    plot.title = element_text(hjust = 0.5, vjust = 5, angle = 182),  # Title rotate
    axis.title.x = element_text(angle = 72, vjust = 0.3, hjust = 0.5),  # X-axis rotate
    legend.position = c(0.5, 0.5),  # Legend in the middle
    legend.background = element_rect(fill = rgb(173/255, 216/255, 10/255, alpha = 0.5)),  # Semi-transp
    plot.margin = margin(t = 20, r = 20, b = 20, l = 20, unit = "pt")  # Plot size
  )
```



This plot stands as the accumulation of my education and knowledge from all the courses I've taken at UVA. I tried to violate every single aesthetic and usability rule while keeping in the essentials for a time series.

I find that wherever your eyes go, there is something more to admire. The label text being rotated against convention is a choice you rarely see but will keep in mind as this distinct aesthetic choice lingers in the mind for years to come. This idea comes from the recent ICA where my group forgot to describe it as the convention is so standard, we do not think to consider it. The unique font choices especially bring them out, accessibility and readability be damned.

The axes labels are inspired by the many times I looked at a chart and found it entirely useless because the scaling was impractical. The y-axes ticks are randomized and inconsistent to bring a sense of excitement in addition to defying conventions, while the x-axis ticks are left alone so that every single date between 2010 and October 4th can be represented fully. Scaling or minimizing data as appropriate is a skill that should be applied by statisticians with their best judgement for each situation.

Legend placement is far more important than expected when reading unfamiliar data, or one where there may be more information than can be easily glimpsed. Your eyes have to go back and forth often to figure it out so I figure the most ideal place to put it is dead center so you can look at both at once.Transparency is lowered for the ease of viewers.

The grid lines utilizes two colors I love with no regard for readability. With the same sense and limitations as a toddler with the limited Cra-Z-Art crayon set I had instead of the nice 64 Crayola pack my peers did, I manually set the colors to the various winning numbers based on their multipliers. It is important to consider accessibility and color theory when making graphs. A red used in tandem with a blue may work but perhaps not green. Of note, the bright red background is a shining example of how values affect how difficult or easy it may be to even look at a plot.

Nothing done here is from beyond the scope of the course and are typically features of most visualizations. Happy learnings!