STAT Midterm Project

Bhargav Dasari

Introduction:  
  
Insurance companies are crucial for protecting against unforeseen circumstances. They provide a safety net in trying times by offering protection for people's health, their belongings, and enterprises. Insurers support economic stability by sharing risks and assisting in recovery. They also promote safety procedures and risk management, which helps to build a prepared society. Both individuals and businesses must understand the importance of insurance firms in order to feel secure and confident when confronting life's difficulties.

Bad Graph:

A "bad" graph can be compared to an ambiguous road sign that points you in the incorrect direction. It either fails to effectively convey its point or even deceives you into thinking something that isn't true. What goes wrong with a graph is as follows: It occasionally distorts the data, making things appear worse or better than they are. Sometimes it's just a mess—too disorganized, too intricate, or with tiny, difficult-to-read labeling. And things can go wrong if the data is inaccurate. It's like attempting to find your way around on a map that has half the streets missing. Then there are the design decisions that simply leave you perplexed—vibrant hues that detract rather than clarify, or ornate effects that do nothing. The worst part is that some graphs merely leave you wondering what the numbers mean. A poor graph ultimately leaves you with more confusion than when you began since it is unable to captivate, educate, or inspire confidence. This is like a bad storyteller.

The article "Top 25 Largest U.S. Insurance Brokerage Firms" on the website HowMuch.net served as the source for the graph showing the top 25 largest U.S. insurance brokerage firms. Based on the brokerage income these companies received from clients in the United States in 2018, this thorough research rates these firms. The information offers insightful information about the state of the insurance market and the major participants in it.

A diagram of a company's financial report

Description automatically generated with medium confidence

FIG-1 : GRAPH CHOSEN TO REDESIGN

Strength of Bad Graph:

The exploding pie chart easily captures the reader's attention due to its appealing visuals. To improve clarity, revenue ranges are distinguished using colour coding. In addition, it provides a concise overview of the top 25 companies and their revenue positions, making comparisons easier. The segment size representation on the graph, wherein larger segments correspond to higher earnings, appropriately represents the revenue of each organization. The data visualization is guaranteed to be understandable and straightforward thanks to this natural approach.

Weakness of Bad Graph:

Certain firm names and their accompanying revenues seem small, making the graph difficult to read, especially in print or on smaller displays. This could affect the graph's legibility. Furthermore, the absence of comprehensive context regarding the data, such as the methodology used to calculate revenues and the duration of the data, may cause viewers to misinterpret or draw insufficient conclusions.

Redesigned Graph:

A well-made graph is a powerful tool for communicating information in an understandable, accurate, and visually appealing manner. It accomplishes this by using appropriate scales and axes to precisely portray the data and by appropriately representing the data without distorting patterns or correlations. Clarity and simplicity are crucial because they maximize visual appeal and comprehension while reducing needless complication. The graph is more credible and reliable when it uses reliable and up-to-date data. While this is going on, thoughtful design decisions about fonts, colour schemes, and other visual components encourage interaction without overpowering the information. Providing relevant context helps viewers understand the data, whether by explaining words, explaining trends, or drawing comparisons with reference points. Additionally, it is imperative to assure usage through clear labelling for a variety of audiences, including individuals with low understanding or visual impairments as well as other formats. In the end, a well-designed graph achieves its goals by offering insightful analysis and making complex data easier to understand.

By following all the principles of designing a good graph I have redesigned the exploding pie chart(fig.1) into horizontal bar charts(fig 2&3) using software’s such as R and Tableau

R Studio:

A graph of a business insurance firm

Description automatically generated

FIG-2 : GRAPH REDESIGNED USING R STUIDO

I have written the R code and designed the above graph (Fig 2). This version looks much better and is more understandable compared to the first figure. I utilized Colour Brewer to select colours that are friendly to colorblind individuals and applied a high-contrast background from the R file for improved visibility. Despite numerous adjustments, this graph still falls short as it remains difficult to discern values that are very close to each other.

R code:

library(ggplot2)

library(readxl)

hw <- theme\_gray()+ theme(

plot.title=element\_text(hjust=0.5),

plot.subtitle=element\_text(hjust=0.5),

plot.caption=element\_text(hjust=-.5),

# strip.text.y = element\_blank(),

strip.background=element\_rect(fill=rgb(.9,.95,1),

colour=gray(.5), linewidth=.2), #changed size to linewidth for ggplot2 3.4.0

panel.border=element\_rect(fill=FALSE,colour=gray(.70)),

panel.grid.minor.y = element\_blank(),

panel.grid.minor.x = element\_blank(),

panel.spacing.x = unit(0.20,"cm"),

panel.spacing.y = unit(0.20,"cm"),

# axis.ticks.y= element\_blank()

axis.ticks=element\_blank(),

axis.text=element\_text(colour="black"),

axis.text.y=element\_text(margin=margin(0,3,0,3)),

axis.text.x=element\_text(margin=margin(-1,0,3,0))

)

# Read the data

insurance\_data <- read\_excel("C:/Users/Bhargav Dasari/Documents/SPRING-2024/STAT-515/MIDTERM/MID TERM DATA.xlsx")

# Define revenue ranges and reverse their order

# Define revenue ranges

insurance\_data$Revenue\_Group <- cut(insurance\_data$`REVENUE IN BILLION DOLLARS`,

breaks = c(-Inf, 1, 2, 4, Inf),

labels = c("<1", "1-2", "2-4", ">4"),

right = FALSE,

ordered\_result = TRUE)

# Reorder company levels by revenue in ascending order

insurance\_data$COMPANY <- factor(insurance\_data$COMPANY, levels = insurance\_data$COMPANY[order(insurance\_data$`REVENUE IN BILLION DOLLARS`)])

# Define your custom fill colors

custom\_colors <- c("<1" = "#a6cee3", "1-2" = "#1f78b4", "2-4" = "#b2df8a", ">4" = "#33a02c")

ggplot(insurance\_data, aes(x = `REVENUE IN BILLION DOLLARS`, y = COMPANY, fill = Revenue\_Group)) +

geom\_bar(stat = "identity", position = position\_dodge(width = 0.8)) +

labs(x = "Revenue (Billion Dollars)", y = "Company",

title = "Top 25 Business Insurance Firms in the U.S. by Revenue Group",

fill = "Brokerage Revenue") + # Change legend title to "Brokerage Revenue"

scale\_fill\_manual(values = custom\_colors, guide = guide\_legend(reverse = TRUE)) + # Use custom colors

theme\_minimal() + hw +

theme(axis.title.x = element\_text(margin = margin(t = 20)),

legend.position = "right",

plot.title = element\_text(hjust = 0.5)) # Adjust the title to be centered

Tableau:

A graph of a company's insurance

Description automatically generated

FIG-3 : GRAPH REDESIGNED USING TABLEAU

So far, I have achieved the best outcomes using Tableau. I can display the difference between very close values by showing the values directly on the bars instead of relying solely on the x-axis. Additionally, I have utilized Tableau's colourblind-friendly colour palette in the graph.

I am exceedingly pleased with Mr. Daniel Berhane Araya's review and have integrated numerous of his suggestions into my redesigned graphs. His critique has fundamentally altered my perspective on graph design, enabling me to pinpoint precise issues. For example, I have decided to eliminate the x-axis, instead displaying numbers directly on the bars. This modification has enhanced clarity and minimized clutter. Furthermore, I have adopted a visually captivating color scheme by employing the Color Brewer palette. These adjustments have markedly enhanced both the efficacy and aesthetics of my graphs.

(ChatGPT, 2023) (Irena, 2020)The challenges I encountered while redesigning the exploding pie chart using R Studio were significant. Despite the ease of creating plots with the various libraries available, I struggled to achieve the desired customization. Despite my efforts, I found it challenging to create a perfect graph that met my requirements. To overcome this hurdle, I turned to Tableau, where I was able to create a flawless graph. Tableau's flexibility allowed me to differentiate similar values effectively by representing them directly on the bars themselves.

I also felt that if I had the opportunity to utilize Color Brewer in Tableau, the graph would have looked even better. Tableau offers a very limited selection of colors to choose from, which can be restrictive in achieving the desired visual appeal.

Conclusion:

In conclusion, the graph created using Tableau proves to be highly effective in the realm of data analytics. When analysing extensive datasets, Tableau offers convenience and efficiency, requiring significantly less time compared to other methods. This facilitates the analysis of large datasets within a remarkably short timeframe.

# References

ChatGPT. (2023). San Franciso: OpenAI.

Irena. (2020). *The Top 25 Business Insurance Firms in the U.S*. Retrieved from Howmuch.net: <iframe width="800" height="600" src="https://cdn.howmuch.net/articles/americas-top-25-brokers-revenue-361e.jpg"></iframe>