Project 1

Archit Singh, Sreeram Kondapalli

2023-10-12

Introduction

- In our project we wish to analyze how the rise of AI tools like ChatGPT has impacted the education industry and the tech industry as a whole, seeing how companies that embraced it fared versus companies that did not. We intend to delve into various aspects of this transformation, examining the extent to which companies that have adopted AI technologies have fared compared to those that have not embraced these innovations.
- Our primary motivating question for this project is: "What is the impact of AI tools, such as ChatGPT, on the education industry and the tech sector, and how has this adoption influenced the competitive landscape and performance of companies?"

Methods/ Datasets

There are many datasets needed to analyze this topic as a whole. We have datasets accross the range from datasets of companies' stocks to datasets for ChatGPT user growth, as well as a student survey on how they feel about AI.

```
# Read data from the CSV file
chatgpt_users <- read.csv("Number of ChatGPT users recorded over past months.csv")
head(chatgpt_users)
```

ChatGPT User Growth

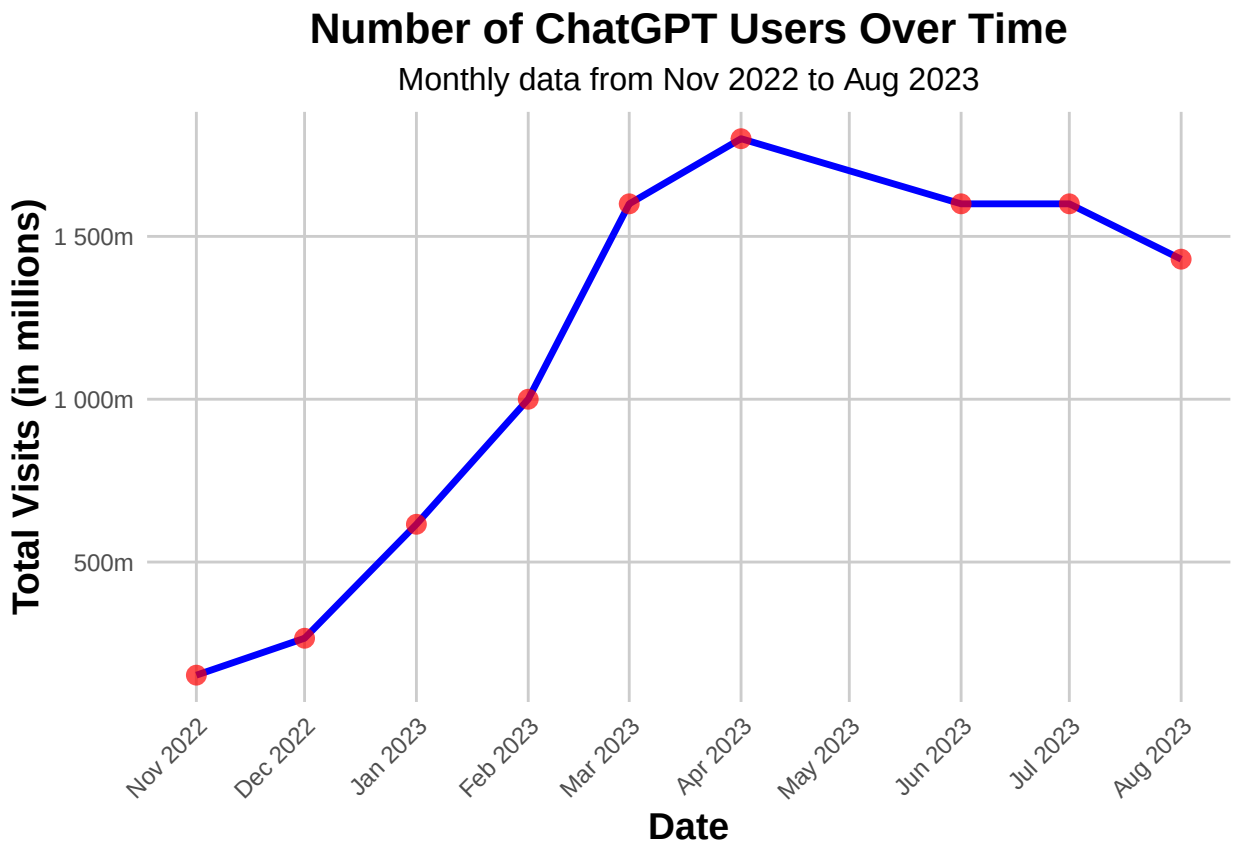
```
##   Sr.No       Date Number.of.ChatGPT.users.recorded.over.past.months
## 1     1 2023-08-01                                         1430000000
## 2     2 2023-07-01                                         1600000000
## 3     3 2023-06-01                                         1600000000
## 4     4 2023-04-01                                         1800000000
## 5     5 2023-03-01                                         1600000000
## 6     6 2023-02-01                                         1000000000
```

```
#Renaming for simplicity
chatgpt_users <- chatgpt_users %>%
  rename(TotalVisits = 'Number.of.ChatGPT.users.recorded.over.past.months')
chatgpt_users$Date <- as.Date(chatgpt_users$Date, format="%Y-%m-%d")
```

```r
label_format_millions <- function(x) {
  scales::number_format(scale = 1e-6, suffix = "m")(x)
}

ggplot(chatgpt_users, aes(x=Date, y=TotalVisits)) +
  geom_line(color = "blue", size = 1.2) +
  geom_point(color = "red", size = 3, shape = 19, alpha = 0.7) +
  labs(title="Number of ChatGPT Users Over Time",
       subtitle="Monthly data from Nov 2022 to Aug 2023",
       x="Date", y="Total Visits (in millions)") +
  scale_x_date(date_breaks = "1 month", date_labels = "%b %Y") +
  scale_y_continuous(labels = label_format_millions) +
  theme_minimal() +
  theme(
    plot.title = element_text(size = 16, face = "bold", hjust = 0.5),
    plot.subtitle = element_text(size = 12, hjust = 0.5),
    axis.text.x = element_text(angle = 45, hjust = 1),
    axis.title = element_text(size = 14, face = "bold"),
    panel.grid.major = element_line(color = "gray80"),
    panel.grid.minor = element_blank()
  )
```

```
## Warning: Using ‘size‘ aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use ‘linewidth‘ instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```



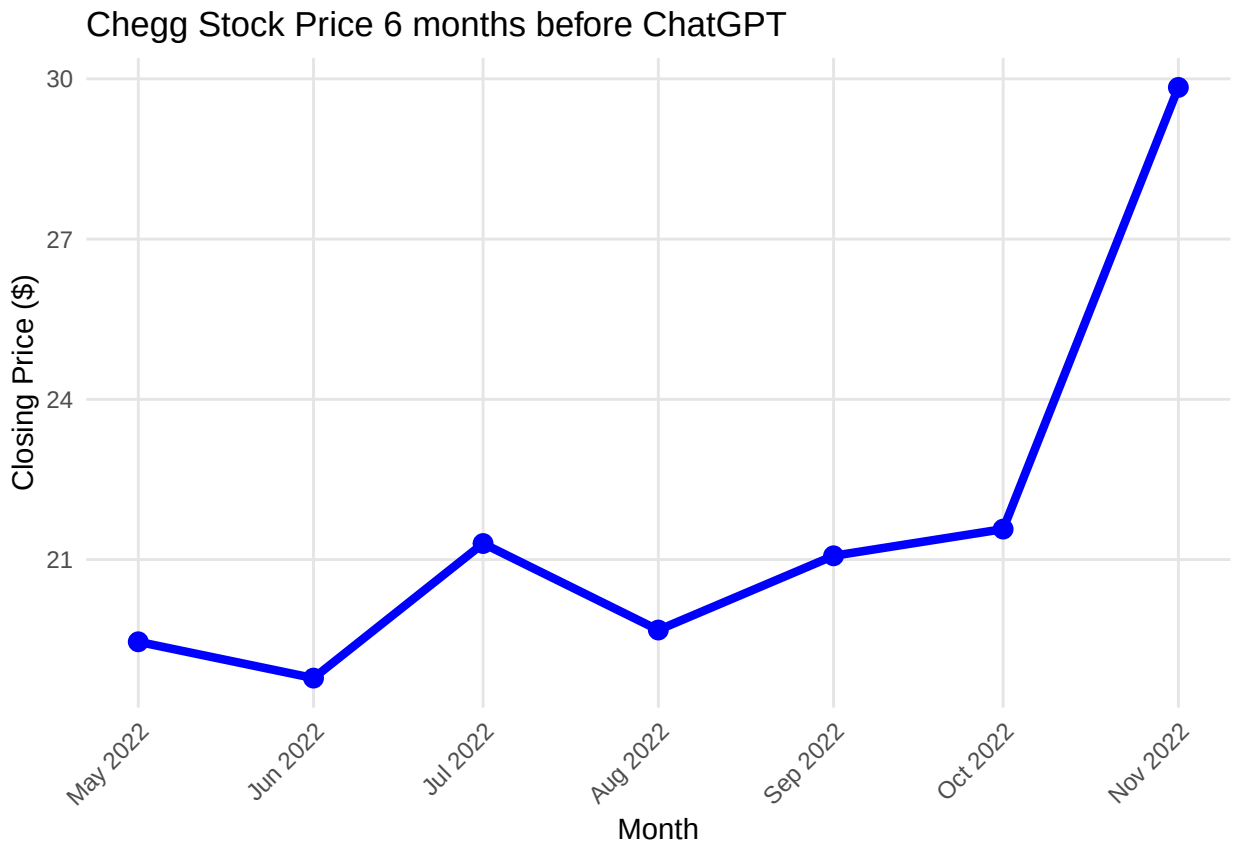

```r
# Assuming you've already loaded the data
chegg_stock_price <- read.csv("chegg-stock-dataset-Monthly.csv")

# Ensure the Date column is a Date type
chegg_stock_price$Date <- as.Date(chegg_stock_price$Date, format="%Y-%m-%d")

# Filter the data
filtered_chegg_stock_price <- chegg_stock_price %>%
  filter(Date >= as.Date("2022-05-01") & Date <= as.Date("2022-11-01"))

# Enhanced graph
ggplot(data = filtered_chegg_stock_price, aes(x = Date, y = Close)) +
  geom_line(aes(color = "Stock Price"), size = 1.5) +
  geom_point(aes(color = "Stock Price"), size = 3) +
  scale_color_manual(name = "", values = c("Stock Price" = "blue")) +
  labs(title = "Chegg Stock Price 6 months before ChatGPT",
       x = "Month",
       y = "Closing Price ($)") +
  theme_minimal() +
  theme(
    panel.grid.major = element_line(color = "gray90", size = 0.5),
    panel.grid.minor = element_blank(),
    axis.text.x = element_text(angle = 45, hjust = 1),
    legend.position = "none"
  ) +
  scale_x_date(date_breaks = "1 month", date_labels = "%b %Y")
```

```
## Warning: The ‘size‘ argument of ‘element_line()‘ is deprecated as of ggplot2 3.4.0.
## i Please use the ‘linewidth‘ argument instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```



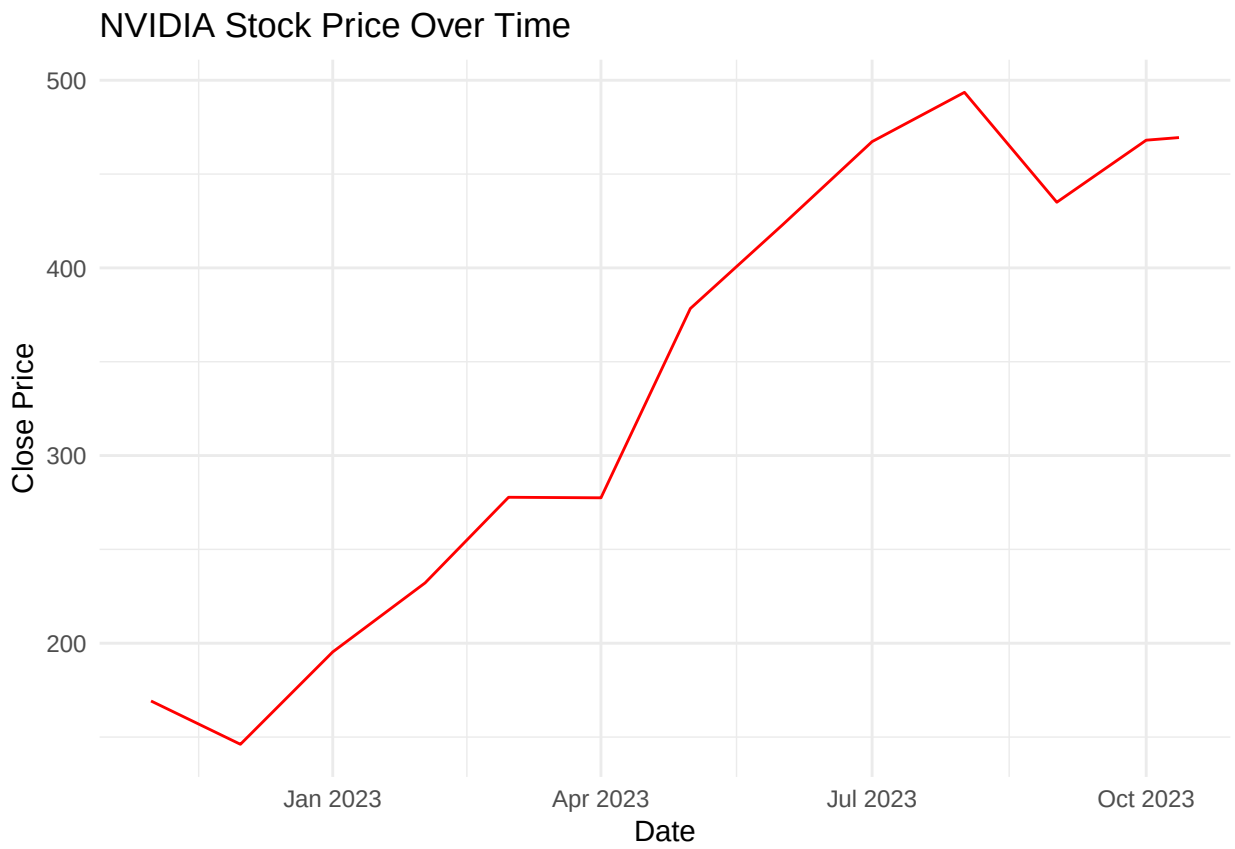
ChatGPT User growth VS Chegg stock

```
# 1_mil_users
```

Time taken by services to reach 1 million users

Students' Perspective on AI

Stock of Companies That Released Public AI The following are the companies that announced generative AI


NVIDIA Stock Price Over Time
500
400
300
200
Close Price
Jan 2023
Apr 2023
Jul 2023
Oct 2023
Date


Microsoft Stock Price Over Time
340
320
300
280
260
240
Close Price
Jan 2023
Apr 2023
Jul 2023
Oct 2023
Date

Salaries for AI and ML Specialists

ML Model (Companies that integrate AI)

Results

Conclusions

Appendix