Final Project

Melissa Wu

7/12/2021

Rationale

I want to study the relationship between the health of the technology sector and the growth of the US economy. Technology plays a large role in today's society, so I want to learn how it can impact the economic wellbeing of an entire nation. I can measure associations and possibly correlations between technological indicators such as the San Francisco Tech Pulse and economic indicators such as real GDP growth rate, CPI, interest rates, and wage rates. This can be done by creating linear regressions, if appropriate, or overlaying time trends on line graphs to visualize any consistent patterns. Note that the economic indicators were transformed into percentage change trends in order to visualize a more clear and direct relationship between indexed measures.

Setup

```r
# Read csv files
tech_pulse <- read_csv("data/sf_tech_pulse.csv")
real_gdp <- read_csv("data/real_gdp.csv")
cpi <- read_csv("data/cpi.csv")
interest_rates <- read_csv("data/interest_rates.csv")
wage_rates <- read_csv("data/wage_rates.csv")

# Merge into data object
data <- tech_pulse %>%
  merge(real_gdp) %>%
  merge(cpi) %>%
  merge(interest_rates) %>%
  merge(wage_rates) %>%
  rename(date = DATE,
         sftp = SFTPAGRM158SFRBSF,
         gdp = GDPC1,
         cpi = CPIAUCSL,
         interest = INTDSRUSM193N,
         wage = LES1252881600Q) %>%
  mutate(year = substr(date, 1, 4),
         month = substr(date, 6, 7)) %>%
  relocate(year, .after = date) %>%
  relocate(month, .after = year)

# Reformat into date and numerics
```

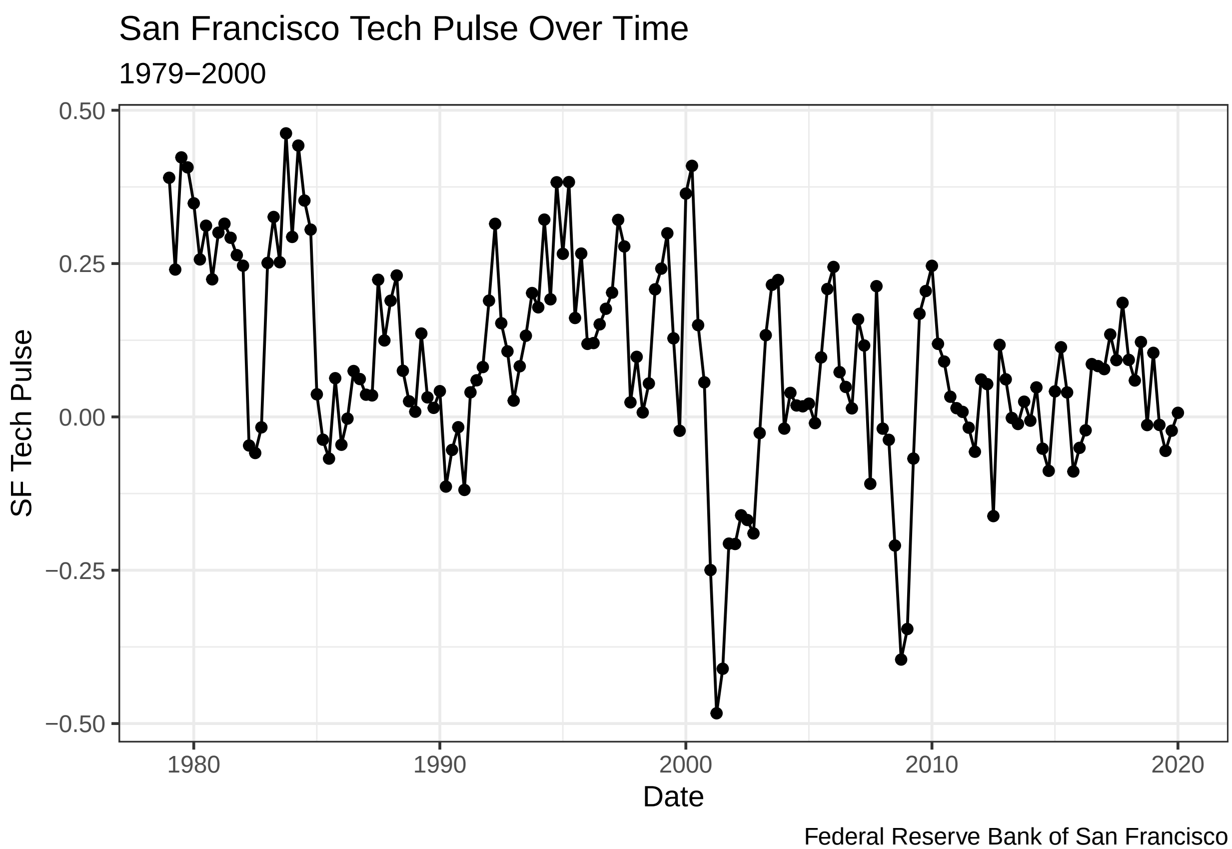
```r
data$date <- as.Date(data$date)
data$year <- as.numeric(data$year)
data$month <- as.numeric(data$month)

# Calculate percentage change in indicators
data <- data %>%
  mutate(gdp_change = (gdp - lag(gdp))/lag(gdp)) %>%
  mutate(cpi_change = (cpi - lag(cpi))/lag(cpi)) %>%
  mutate(interest_change = (interest - lag(interest))/lag(interest)) %>%
  mutate(wage_change = (wage - lag(wage))/lag(wage))
```

Change Over Time Plots

San Francisco Tech Pulse

The Tech Pulse Index measures the health of activity in the US information technology sector. The index is determined by indicators including investment in IT goods, consumption of personal computers, employment in the IT sector, industrial production, and more. The SF Tech Pulse was measured monthly by the Federal Reserve Bank of San Francisco, but was discontinued in March of 2020. The units used are Percent Change at Annual Rate, Seasonally Adjusted.

```r
# SFTP vs. Time
sftp_time <- data %>%
  ggplot(aes(x = date, y = sftp)) +
  geom_line() +
  geom_point() +
  labs(title = "San Francisco Tech Pulse Over Time",
       x = "Date", y = "SF Tech Pulse",
       subtitle = "1979-2000",
       caption = "Federal Reserve Bank of San Francisco") +
  theme_bw()
sftp_time
```



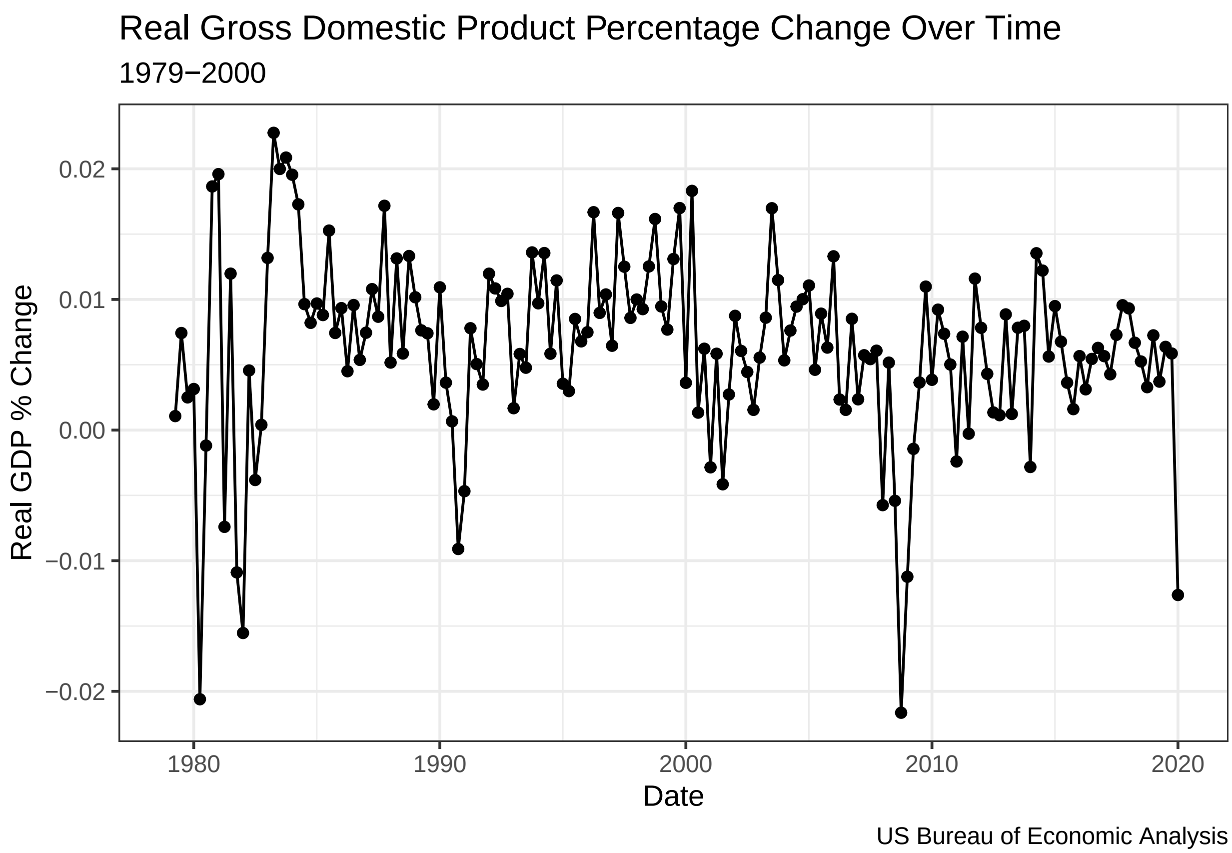

```
# Animated GIF for Shiny App
# p <- data %>%
#   ggplot(aes(x = date, y = sftp)) +
#   geom_line() +
#   geom_point() +
#   labs(title = "San Francisco Tech Pulse Over Time",
#        x = "Date", y = "SF Tech Pulse",
#        subtitle = "1979-2000",
#        caption = "Federal Reserve Bank of San Francisco") +
#   theme_bw() +
#   transition_reveal(date)
# anim_save("www/sftp.gif", animate(p))
```

Real Gross Domestic Product

Real GDP measures the total value of the goods and services produced in the US, adjusted for inflation. GDP includes consumption, investment, government expenditures, and net exports. The measure is computed quarterly by the US Bureau of Economic Analysis, and is measured in units of Billions of Chained 2012 Dollars, Seasonally Adjusted Annual Rate.

```
# GDP % change vs. Time
gdp_time <- data %>%
  ggplot(aes(x = date, y = gdp_change)) +
  geom_line() +
  geom_point() +
```

```
  labs(title = "Real Gross Domestic Product Percentage Change Over Time",
       x = "Date", y = "Real GDP % Change",
       subtitle = "1979-2000",
       caption = "US Bureau of Economic Analysis") +
  theme_bw()
gdp_time
```



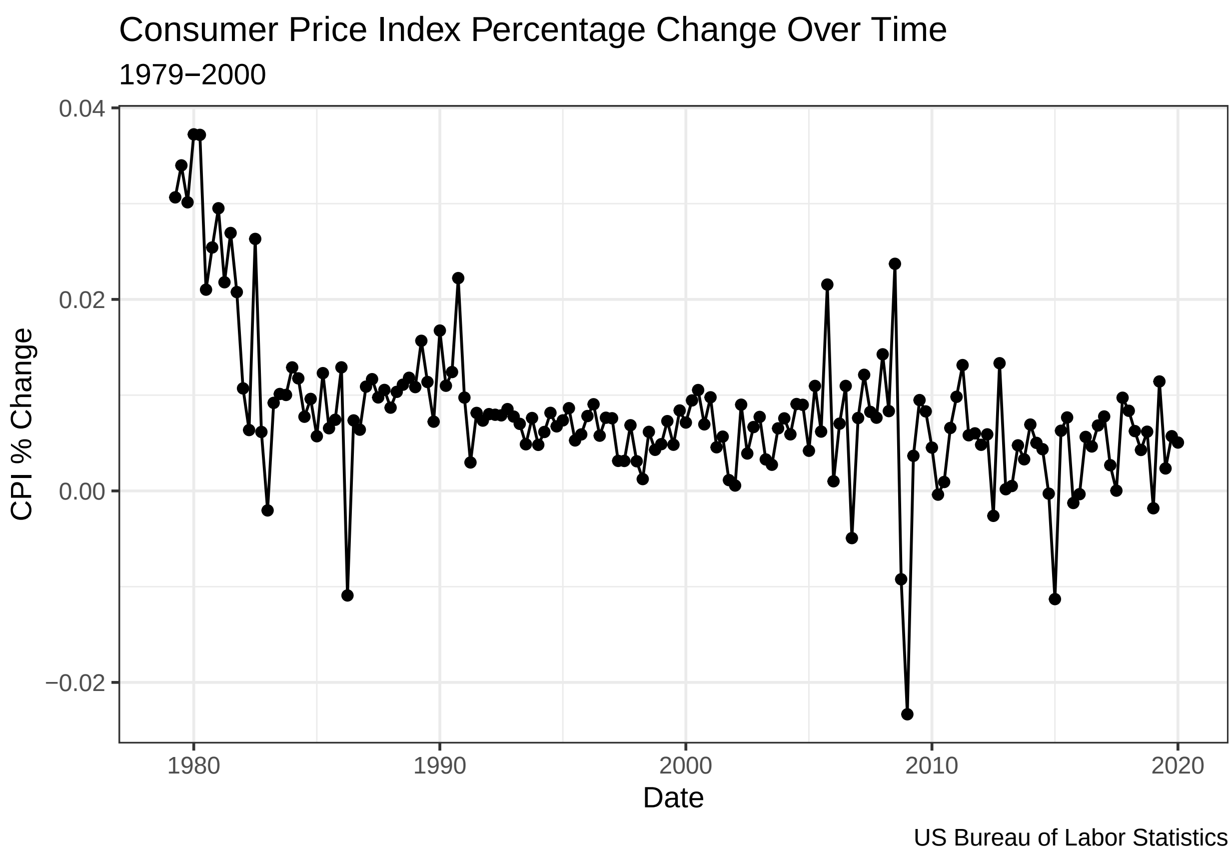

```
# Animated GIF for Shiny App
# p <- data %>%
#   ggplot(aes(x = date, y = gdp_change)) +
#   geom_line() +
#   geom_point() +
#   labs(title = "Real Gross Domestic Product Percentage Change Over Time",
#        x = "Date", y = "Real GDP % Change",
#        subtitle = "1979-2000",
#        caption = "US Bureau of Economic Analysis") +
#   theme_bw() +
#   transition_reveal(date)
# anim_save("www/gdp.gif", animate(p))

# GDP % change vs. SF Tech Pulse
gdp_sftp <- data %>%
  ggplot(aes(x = sftp, y = gdp_change)) +
  geom_point() +
  geom_smooth(method = lm) +
```

```r
  labs(title = "Real GDP % Change vs. SF Tech Pulse",
       x = "SF Tech Pulse", y = "Real GDP % Change") +
  theme_bw()
```

Consumer Price Index

The CPI measures the average change in prices of goods and services, including food, clothing, shelter, fuel, service fees, and sales taxes. Prices are weighted based on their spending importance. The CPI can be used to determine changes in the overall price level, with increases suggesting inflation and decreases suggesting deflation. This index accounts for about 88% of the US population, including wage-earners, self-employed, unemployed, retirees, and those not in the labor force. The CPI is measured monthly by the US Bureau of Labor Statistics.

```r
# CPI % change vs. Time
cpi_time <- data %>%
  ggplot(aes(x = date, y = cpi_change)) +
  geom_line() +
  geom_point() +
  labs(title = "Consumer Price Index Percentage Change Over Time",
       x = "Date", y = "CPI % Change",
       subtitle = "1979-2000",
       caption = "US Bureau of Labor Statistics") +
  theme_bw()
cpi_time
```

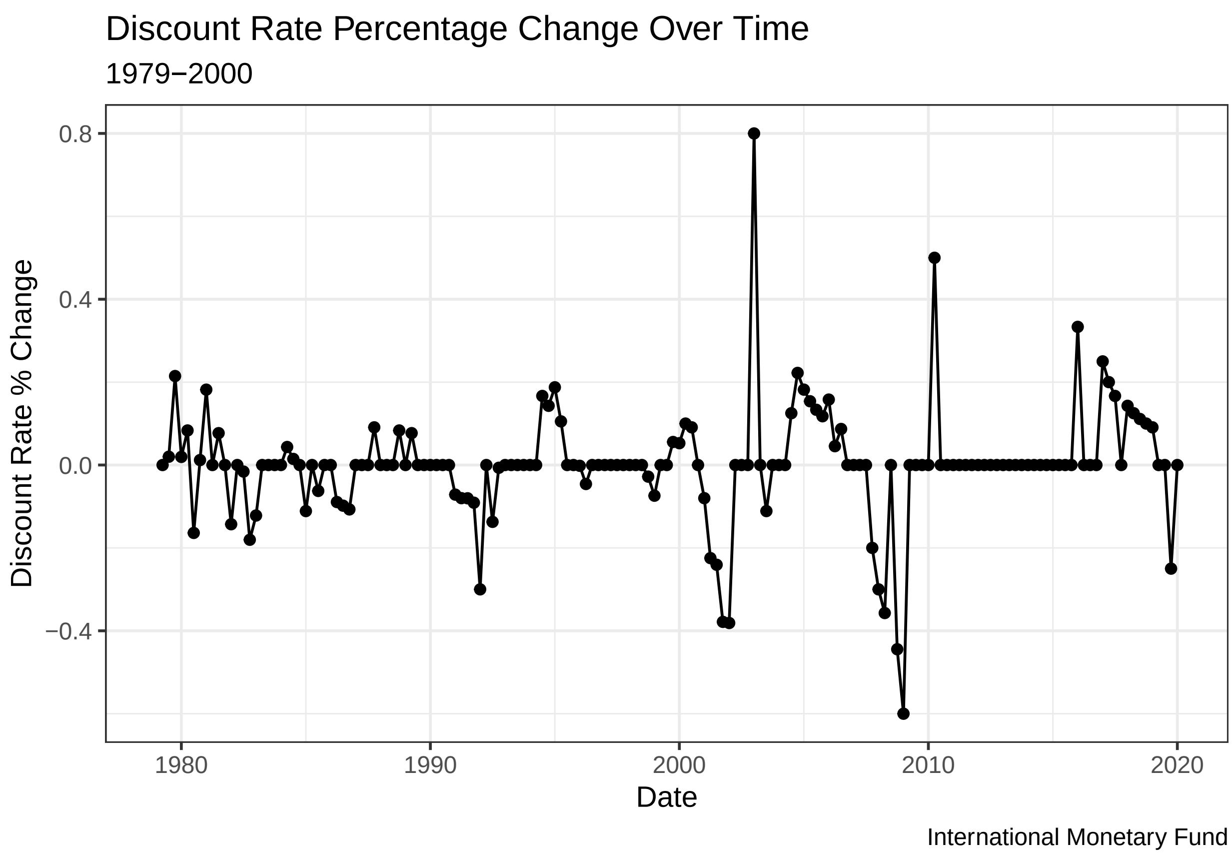


```r
# Animated GIF for Shiny App
# p <- data %>%
#   ggplot(aes(x = date, y = cpi_change)) +
#   geom_line() +
#   geom_point() +
#   labs(title = "Consumer Price Index Percentage Change Over Time",
#        x = "Date", y = "CPI % Change",
#        subtitle = "1979-2000",
#        caption = "US Bureau of Labor Statistics") +
#   theme_bw() +
#   transition_reveal(date)
# anim_save("www/cpi.gif", animate(p))

# CPI % change vs. SF Tech Pulse
cpi_sftp <- data %>%
  ggplot(aes(x = sftp, y = cpi_change)) +
  geom_point() +
  geom_smooth(method = lm) +
  labs(title = "CPI % Change vs. SF Tech Pulse",
       x = "SF Tech Pulse", y = "CPI % Change") +
  theme_bw()
```

Interest Rates

Interest rates affect incentives for consumers to spend and save, with higher interest rates promoting higher saving, and lower interest rates promoting more spending. The discount rate is a type of interest rate at which the Federal Reserve makes loans to banks. The interest rate is measured by the International Monetary Fund each month, and is measured in Percent per Annum, Not Seasonally Adjusted.

```r
# Interest % change vs. Time
interest_time <- data %>%
  ggplot(aes(x = date, y = interest_change)) +
  geom_line() +
  geom_point() +
  labs(title = "Discount Rate Percentage Change Over Time",
       x = "Date", y = "Discount Rate % Change",
       subtitle = "1979-2000",
       caption = "International Monetary Fund") +
  theme_bw()
interest_time
```

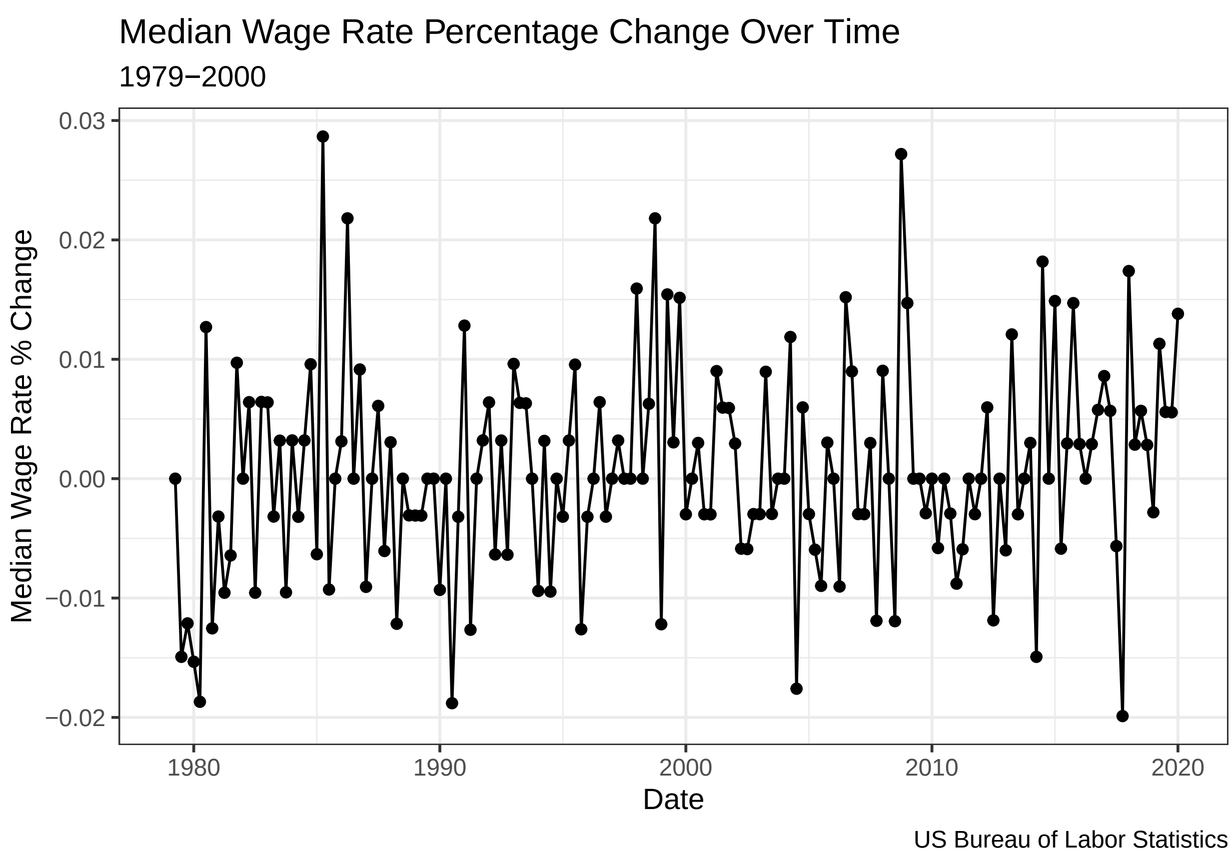



```r
# Animated GIF for Shiny App
# p <- data %>%
#   ggplot(aes(x = date, y = interest_change)) +
#   geom_line() +
#   geom_point() +
#   labs(title = "Discount Rate Percentage Change Over Time",
#        x = "Date", y = "Discount Rate % Change",
#        subtitle = "1979-2000",
#        caption = "International Monetary Fund") +
#   theme_bw() +
#   transition_reveal(date)
# anim_save("www/interest.gif", animate(p))

# Interest % change vs. SF Tech Pulse
interest_sftp <- data %>%
  ggplot(aes(x = sftp, y = interest_change)) +
  geom_point() +
  geom_smooth(method = lm) +
  labs(title = "Discount Rate % Change vs. SF Tech Pulse",
       x = "SF Tech Pulse", y = "Discount Rate % Change") +
  theme_bw()
```

Wage Rates

The median wage rate measures the usual weekly real earnings of workers aged 16 and above. This data covers workers in the public and private sectors, but excludes those who are self-employed. Usual weekly earnings are measured prior to taxes and deductions and include overtime pay and tips. This data is measured quarterly by the US Bureau of Labor Statistics, and is measured in 1982-84 CPI Adjusted Dollars, Seasonally Adjusted.

```
# Wage % change vs. Time
wage_time <- data %>%
  ggplot(aes(x = date, y = wage_change)) +
  geom_line() +
  geom_point() +
  labs(title = "Median Wage Rate Percentage Change Over Time",
       x = "Date", y = "Median Wage Rate % Change",
       subtitle = "1979-2000",
       caption = "US Bureau of Labor Statistics") +
  theme_bw()
wage_time
```




```
# Animated GIF for Shiny App
# p <- data %>%
#   ggplot(aes(x = date, y = wage_change)) +
#   geom_line() +
#   geom_point() +
```

	Min	Median	Max	Mean	SD
SF Tech Pulse	-0.483	0.075	0.462	0.090	0.167
GDP % Change	-0.022	0.007	0.023	0.006	0.007
CPI % Change	-0.023	0.007	0.037	0.008	0.008
Interest Rate % Change	-0.600	0.000	0.800	0.002	0.140
Wage Rate % Change	-0.020	0.000	0.029	0.001	0.009

```
#   labs(title = "Median Wage Rate Percentage Change Over Time",
#        x = "Date", y = "Median Wage Rate % Change",
#        subtitle = "1979-2000",
#        caption = "US Bureau of Labor Statistics") +
#   theme_bw() +
#   transition_reveal(date)
# anim_save("www/wage.gif", animate(p))

# Wage % change vs. SF Tech Pulse
wage_sftp <- data %>%
  ggplot(aes(x = sftp, y = wage_change)) +
  geom_point() +
  geom_smooth(method = lm) +
  labs(title = "Wage Rate % Change vs. SF Tech Pulse",
       x = "SF Tech Pulse", y = "Wage Rate % Change") +
  theme_bw()
```

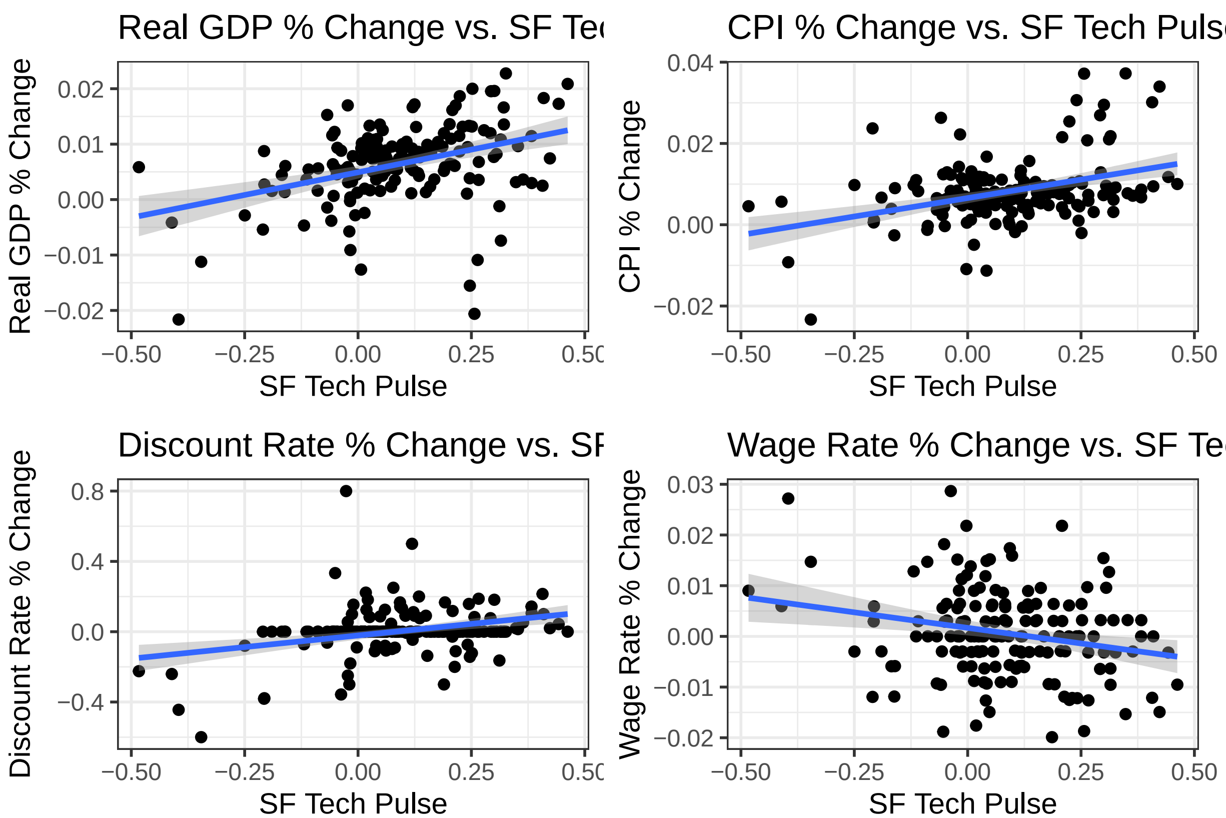
Diagnostics

The following table shows the summary statistics of each technological and economic indicator, measured quarterly from 1979 to 2020. Out of all the variables we're looking at, they each seem to be centered at a little above 0. The discount rate percentage change has the largest range of 1.4, and the GDP percentage change has the smallest range of 0.045. With regard to spread, the SF Tech Pulse has the largest standard deviation of 0.167, and the GDP percentage change has the smallest standard deviation of 0.007.

```
# Summary Statistics
response <- data %>% select(sftp, gdp_change, cpi_change, interest_change, wage_change)
sum_min <- sapply(response, min, na.rm=TRUE)
sum_median <- sapply(response, median, na.rm=TRUE)
sum_max <- sapply(response, max, na.rm=TRUE)
sum_mean <- sapply(response, mean, na.rm=TRUE)
sum_sd <- sapply(response, sd, na.rm=TRUE)

stats <- cbind(sum_min, sum_median, sum_max, sum_mean, sum_sd)
rownames(stats) <- c("SF Tech Pulse", "GDP % Change", "CPI % Change",
                     "Interest Rate % Change", "Wage Rate % Change")
colnames(stats) <- c("Min", "Median", "Max", "Mean", "SD")

stat_table <- stats %>%
  round(digits = 3) %>%
  kbl() %>%
  kable_styling(bootstrap_options = c("striped", "hover", "responsive"))
stat_table
```

The next graphs show each of the economic indicators plotted against the SF Tech Pulse, as well as the regression line that best describes the pattern between the variables. Real GDP, CPI, and Discount Rate all show positive relationships with the SF Tech Pulse, while the Wage Rate shows a downward trend.

```r
# 4 Indicator vs. SFTP Graphs
(gdp_sftp + cpi_sftp) / (interest_sftp + wage_sftp)
```



This table summarizes the regression statistics from running linear regressions of each economic indicator against the SF Tech Pulse. As expected, all indicators but Wage Rate have a positive slope estimate. The absolute values of t statistics of each model are all relatively high, and thus all p-values are near 0. The correlations (absolute value) range from 0.2 to 0.4, suggesting moderate linear relationships. The coefficients of determination are relatively low - the GDP vs. SFTP model captures 14.8% of the variation in GDP, while the wage rate vs. SFTP model captures 5.4% of the variation in wage rate.

```r
# Linear regressions
gdp_lm <- lm(gdp_change ~ sftp, data = data)
gdp_coef <- coef(summary(gdp_lm))
gdp_cor <- cor(data$sftp, data$gdp_change, use = "complete.obs")
gdp_reg <- c(gdp_coef["(Intercept)",][1],
             gdp_coef["sftp",],
             gdp_cor,
             gdp_cor^2)

cpi_lm <- lm(cpi_change ~ sftp, data = data)
cpi_coef <- coef(summary(cpi_lm))
cpi_cor <- cor(data$sftp, data$cpi_change, use = "complete.obs")
```

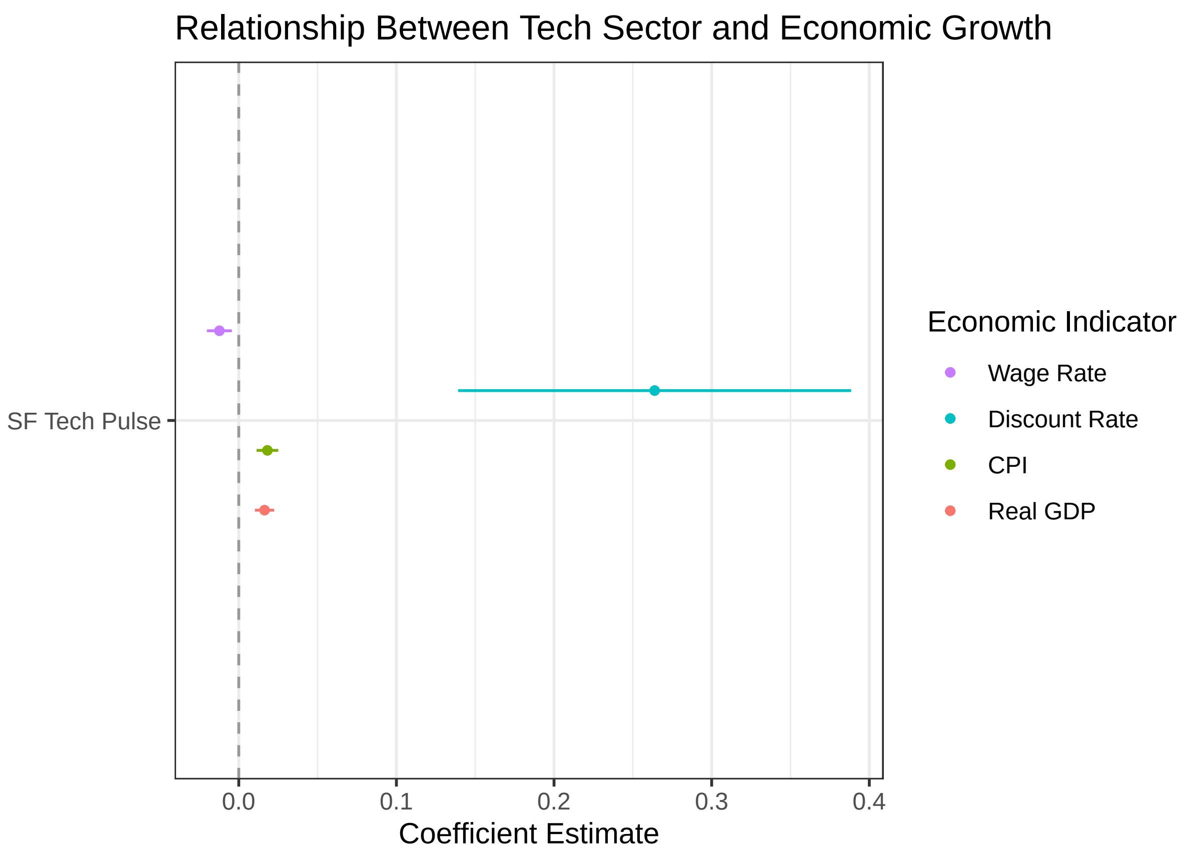
	Y-Intercept	Slope Estimate	Std. Error	t value	Pr(>|t|)	Correlation	R^2
GDP % Change	0.005	0.016	0.003	5.295	0.000	0.384	0.148
CPI % Change	0.007	0.018	0.003	5.215	0.000	0.379	0.144
Interest Rate % Change	-0.021	0.264	0.063	4.178	0.000	0.312	0.097
Wage Rate % Change	0.002	-0.012	0.004	-3.049	0.003	-0.233	0.054

```
cpi_reg <- c(cpi_coef["(Intercept)",][1],
             cpi_coef["sftp",],
             cpi_cor,
             cpi_cor^2)

interest_lm <- lm(interest_change ~ sftp, data = data)
interest_coef <- coef(summary(interest_lm))
interest_cor <- cor(data$sftp, data$interest_change, use = "complete.obs")
interest_reg <- c(interest_coef["(Intercept)",][1],
                  interest_coef["sftp",],
                  interest_cor,
                  interest_cor^2)

wage_lm <- lm(wage_change ~ sftp, data = data)
wage_coef <- coef(summary(wage_lm))
wage_cor <- cor(data$sftp, data$wage_change, use = "complete.obs")
wage_reg <- c(wage_coef["(Intercept)",][1],
              wage_coef["sftp",],
              wage_cor,
              wage_cor^2)

reg <- rbind(gdp_reg, cpi_reg, interest_reg, wage_reg)
rownames(reg) <- c("GDP % Change", "CPI % Change",
                   "Interest Rate % Change", "Wage Rate % Change")
colnames(reg) <- c("Y-Intercept", "Slope Estimate", "Std. Error",
                   "t value", "Pr(>|t|)", "Correlation", expression("$R^2$"))

reg_table <- reg %>%
  round(digits = 3) %>%
  kbl() %>%
  kable_styling(bootstrap_options = c("striped", "hover", "responsive"))
reg_table
```

The following graph is a coefficient plot that visualizes the confidence intervals for each linear model. None of the intervals capture the value of 0, and all p-values are near 0, so each model is statistically significant.

```
# Coefficient plot
coef_plot <- dwplot(list(gdp_lm, cpi_lm, interest_lm, wage_lm),
       model_order = c("Model 1", "Model 2", "Model 3", "Model 4")) %>%
  relabel_predictors(c(sftp = "SF Tech Pulse")) +
  theme_bw() +
  xlab("Coefficient Estimate") + ylab("") +
  geom_vline(xintercept = 0,
             colour = "grey60",
             linetype = 2) +
  ggtitle("Relationship Between Tech Sector and Economic Growth") +
```

```
  scale_color_discrete(
    name = "Economic Indicator",
    labels = c("Real GDP", "CPI", "Discount Rate", "Wage Rate"))
coef_plot
```



Conclusion

From the above analysis, we can conclude that the San Francisco Tech Pulse measure is positively correlated with real GDP, CPI, and the discount rate, and is negatively correlated with the wage rate. Increases in real GDP and CPI are consistent with economic growth, since they indicate the amount of economic activity and consumer confidence. However, increases in the discount rate and decreases in the median wage rate suggest the opposite. When the discount rate increases, banks are less incentivized to take out loans from the Federal Reserve, limiting the growth of the money supply. Lower wage rates imply that workers across the US are worse off in terms of their livelihood and self sufficiency. Overall, the SF Tech Pulse has a strong association with these economic indicators, but not all of the trends suggest that higher technology activity is related to greater economic wellbeing.