```
library(tidyverse)
library(haven)
library(dplyr)
library(stargazer)
library(carData)
library(broom)
library(ggplot2)
library(lubridate)
library(quantmod)

start_date <- as.Date("2009-01-01")
end_date <- as.Date("2014-12-31")
getSymbols("FEDFUNDS", src = "FRED", from = start_date, to = end_date)
FEDFUNDS <- as.data.frame(FEDFUNDS) # Ensure it's a data frame
FEDFUNDS$date <- as.Date(row.names(FEDFUNDS), format="%Y-%m-%d") # Convert
row names to Date
row.names(FEDFUNDS)
str(FEDFUNDS)

firms_final <- read_csv("/Users/ananya/Downloads/Firms final returns.csv")
daily_prices <- read_csv("/Users/ananya/Downloads/firm daily returns.csv")
factors <- read_csv("/Users/ananya/Downloads/Famafrench factors.csv")

firms_final <- firms_final %>%
  mutate(
    start_date = evtdate - 60,
    end_date = evtdate - 6
  )

daily_prices <- daily_prices %>%
  group_by(tic) %>%
  arrange(tic, date) %>%
  mutate(daily_return = (prccd / lag(prccd) - 1)) %>%
  ungroup() %>%
  na.omit()

filtered_returns <- firms_final %>%
  left_join(daily_prices, by = "tic") %>%
  filter(date >= start_date & date <= end_date)

average_daily_returns <- filtered_returns %>%
  group_by(tic, start_date, end_date) %>%
  summarise(
    avg_daily_return = mean(daily_return, na.rm = TRUE),  # Calculating
average, ignoring NAs
    .groups = 'drop'  # Drop grouping after summarisation
  )

ticker_rf_data <- firms_final %>%
  rowwise() %>%  # Ensure we iterate over each row individually
  mutate(
    rf_dates = list(factors$date[factors$date >= start_date & factors$date
<= end_date]),
    rf_values = list(factors$rf[factors$date >= start_date & factors$date
<= end_date])
  ) %>%
  ungroup() %>%
  select(tic, start_date, end_date, rf_dates, rf_values) %>%
  tidyr::unnest(c(rf_dates, rf_values))

mean_rf_per_ticker <- ticker_rf_data %>%
  group_by(tic) %>%
  summarise(
    mean_rf = mean(rf_values, na.rm = TRUE),  # Calculate mean, ignoring
NA values
    .groups = 'drop'  # Drop grouping after summarisation
  )

excess_returns <- average_daily_returns %>%
  left_join(mean_rf_per_ticker, by = "tic") %>%
  mutate(
    excess_return = avg_daily_return - mean_rf  # Calculate excess return
  )
excess_returns <- excess_returns %>%
  select(tic, excess_return)

factors_mean <- firms_final %>%
  mutate(
    factor_means = map2(start_date, end_date, ~ factors %>%
                          filter(date >= .x & date <= .y) %>%
                          summarise(
                            mean_mktrf = mean(mktrf, na.rm = TRUE),
                            mean_SMB = mean(smb, na.rm = TRUE),
                            mean_HML = mean(hml, na.rm = TRUE)
                          ))
  ) %>%
  unnest(factor_means) %>%
  select(tic, mean_mktrf, mean_SMB, mean_HML)
print(factors_mean)

combined_data <- average_daily_returns %>%
  inner_join(mean_rf_per_ticker, by = "tic") %>%
  inner_join(factors_mean, by = "tic")
combined_data <- combined_data %>%
  distinct(start_date, .keep_all = TRUE)

print(combined_data$avg_daily_return[1:10])

combined_data <- combined_data %>%
  mutate(
    excess_return = avg_daily_return - mean_rf,
    market_premium = mean_mktrf - mean_rf,
    adjusted_SMB = mean_SMB,
    adjusted_HML = mean_HML
  )

regression_results <- combined_data %>%
  group_by(tic) %>%
  do({
    model <- lm(avg_daily_return - mean_rf ~ mean_mktrf - mean_rf, data =
.)
    tidy(model)  # Tidies the output to be more readable
  }) %>%
  ungroup()
write_csv(regression_results, "/Users/ananya/Downloads/Regression Results
2.csv")
```

```
write_csv(combined_data, "/Users/ananya/Downloads/firms regression
data.csv")
```