

SA2 No2

Zyann Lynn Mayo

2025-05-10

Test using Shapiro-Wilk normality test the Ethereum returns for trading data every five minutes, from August 7, 2015 to April 15, 2025.

NOTE: Due to source limitation, we were only able to retrieve the **DAILY** returns from **March 10, 2016** to **April 15, 2025**.

First, we retrieve the datas from the CSV file

```
suppressPackageStartupMessages(library(dplyr))
library(dplyr)
library(stats)

eth_data <- read.csv("Ethereum_Historical_Data.csv", stringsAsFactors = FALSE)

head(eth_data)
```

##	Date	Price	Open	High	Low	Vol.	Change..
## 1	4/15/2025	1,588.80	1,625.32	1,660.34	1,584.81	543.64K	-2.15%
## 2	4/14/2025	1,623.73	1,596.46	1,690.44	1,596.14	747.30K	1.62%
## 3	4/13/2025	1,597.77	1,643.09	1,648.38	1,566.77	665.93K	-2.82%
## 4	4/12/2025	1,644.12	1,566.90	1,666.39	1,546.84	570.51K	4.94%
## 5	4/11/2025	1,566.77	1,519.63	1,587.87	1,507.56	653.52K	2.93%
## 6	4/10/2025	1,522.13	1,669.39	1,669.39	1,475.06	1.03M	-8.84%

```
eth_data$Price <- as.numeric(gsub(",", "", eth_data$Price))

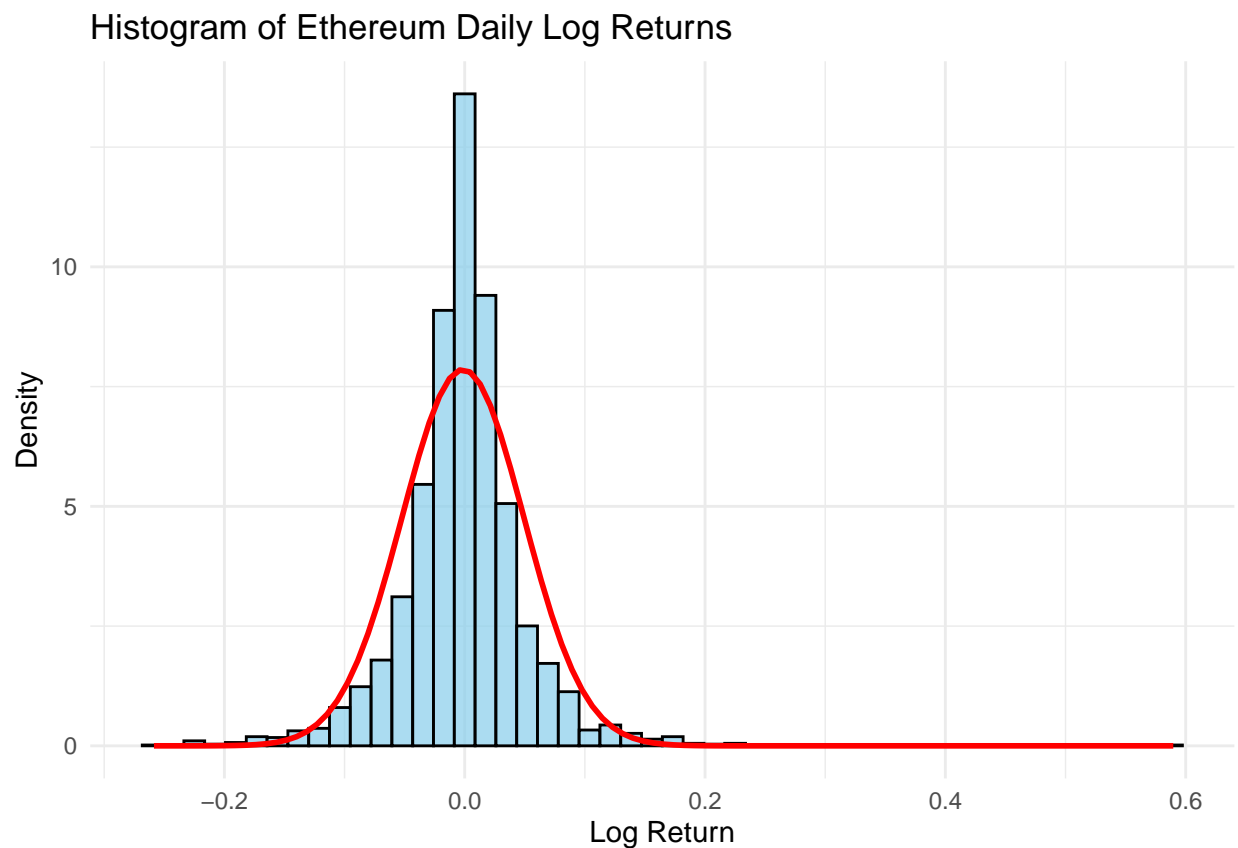
# Sort data by Date
eth_data <- eth_data %>%
  mutate(Date = as.Date(Date, format = "%b %d, %Y")) %>%
  arrange(Date)

# Calculate log returns
eth_data <- eth_data %>%
  mutate(log_return = log(Price / lag(Price))) %>%
  filter(!is.na(log_return))

library(ggplot2)

ggplot(eth_data, aes(x = log_return)) +
```

```
geom_histogram(aes(y = after_stat(density)),
               bins = 50,
               fill = "skyblue",
               color = "black",
               alpha = 0.7) +
stat_function(fun = dnorm,
              args = list(mean = mean(eth_data$log_return, na.rm = TRUE),
                           sd = sd(eth_data$log_return, na.rm = TRUE)),
              color = "red",
              linewidth = 1) +
labs(title = "Histogram of Ethereum Daily Log Returns",
     x = "Log Return",
     y = "Density") +
theme_minimal()
```



Using the built in function, we run the Ethereum data through the Shapiro-Wilk test

```
shapiro_test <- shapiro.test(eth_data$log_return)
```

After testing the normality of Ethereum daily log returns using the Shapiro-Wilk test. The result yielded a W statistic of **0.918** and a p-value of **< 2.2e-16**.

Since the p-value is less than 0.05, the daily returns of Ethereum do **not** follow a normal distribution.

This means the distribution of returns is significantly different from a normal distribution, possibly due to skewness, outliers, or fat tails.