First three observations of the dataset sorted by FDI					
country	year	FDI	inflation	GDP1	GDP
CYP	2020	-296.01	-1.21	-0.03	-3.22
CHE	2020	-32.55	-0.70	-0.02	-2.14
NLD	2020	-23.71	2.36	-0.04	-3.87

Table 1, First three observations of the dataset sorted by FDI (ascending order).

The original variable GDP1 recorded GDP growth in proportions (e.g. 0.03 instead of 3).

A new variable, GDP, was created by multiplying GDP1 by 100 to express GDP growth in percentage terms.

This ensures that GDP growth is now measured consistently with other economic indicators in the dataset.

(All corresponding R code is provided in the Appendix.)

The dataset was sorted in ascending order with respect to inward Foreign Direct Investment (FDI).

Numerical descriptive statistics for FDI, GDP and Inflation							
Variable	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	SD
FDI	-296.01	0.10	2.04	1.12	4.64	191.86	58.03
GDP	-15.31	-6.06	-3.41	-4.14	-2.23	7.16	3.80
Inflation	-2.52	0.96	2.07	2.74	4.09	14.79	3.36

Table 2, Descriptive statistics (minimum, quartiles, median, mean, maximum, and standard deviation) for FDI, GDP growth, and inflation using data from the World Bank World Development Indicators, 2020.

FDI shows the greatest variability (SD ≈ 58) and extreme outliers. GDP growth is mostly negative (mean ≈ –4 %), reflecting weak performance. Inflation averages 2.7 % with moderate spread (SD ≈ 3.4).

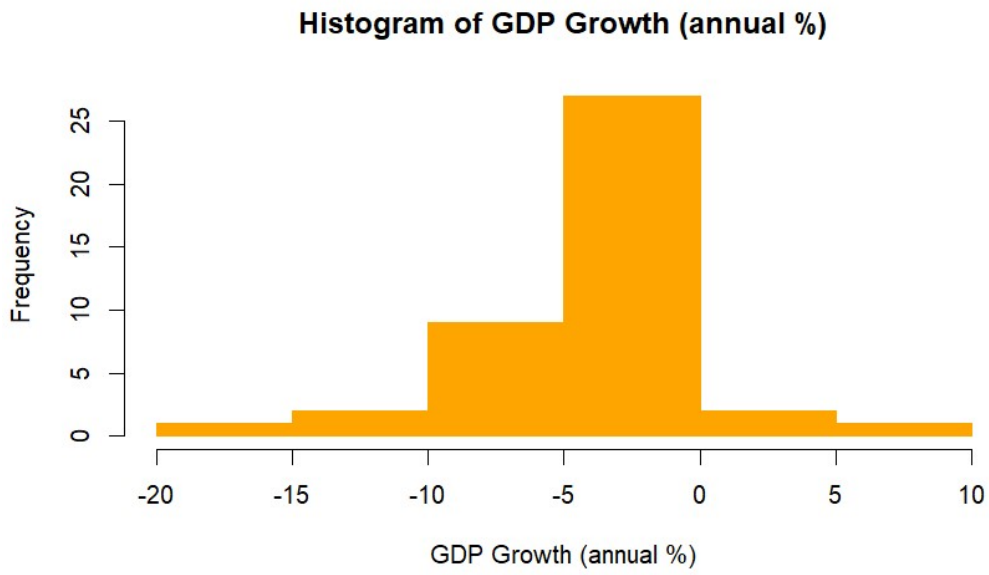

Figure 1, Histogram of GDP growth (annual %).

Clustered below zero.

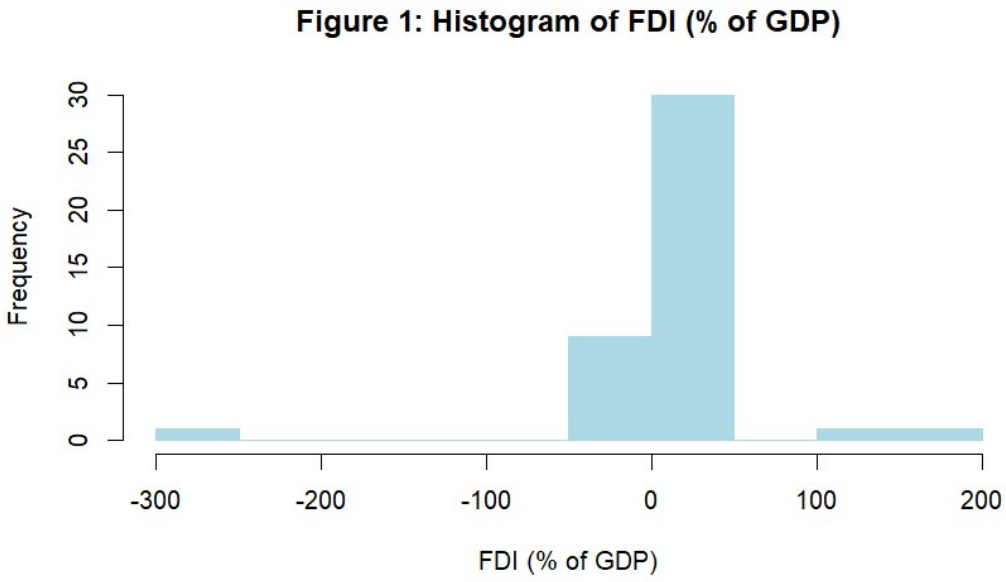

Figure 2, Histogram of FDI (% of GDP).

very wide spread, strong right skew.

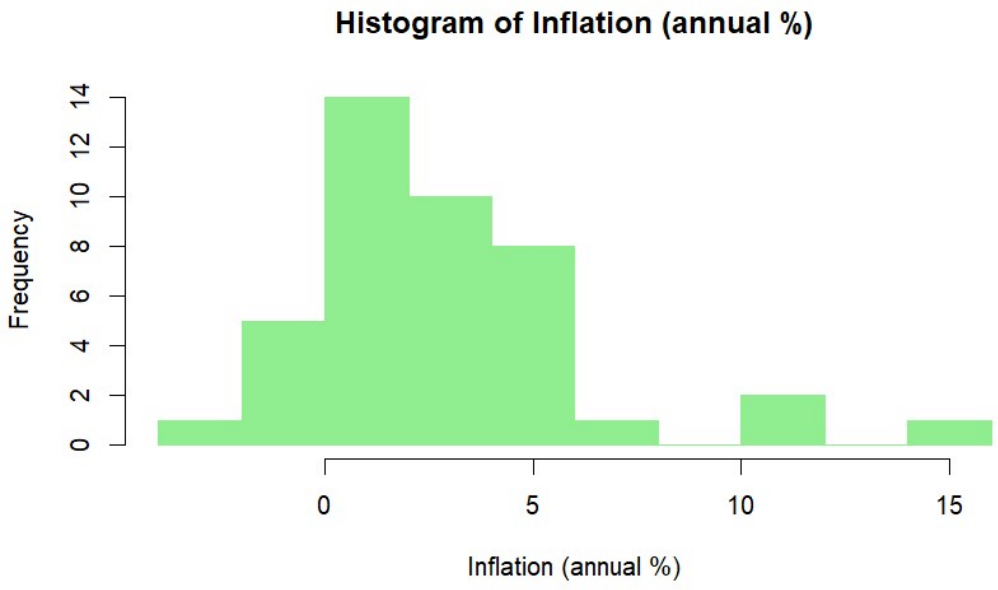

Figure 3, Histogram of inflation (annual %).

Cantered around 2–4 %, mild right tail.

Relationship between covariance and the correlation coefficient for GDP	
Statistic	Value
Covariance	-3.74
Correlation Coefficient	-0.02

Table 3, Covariance and correlation coefficient between GDP growth and FDI (% of GDP), sample 2020.

The covariance between GDP growth and FDI is negative (–3.74), indicating that the two variables tend to move in opposite directions. However, because covariance is unbounded and depends on the units of measurement, its value alone does not convey the strength of the relationship. The correlation coefficient, which standardises covariance to a range between –1 and +1, provides a clearer interpretation. Here, the correlation of –0.02 shows an extremely weak negative association between GDP and FDI. This suggests little linear relationship between the variables. It is also

important to note that correlation measures association, not causation; a relationship does not imply that FDI directly influences GDP growth.

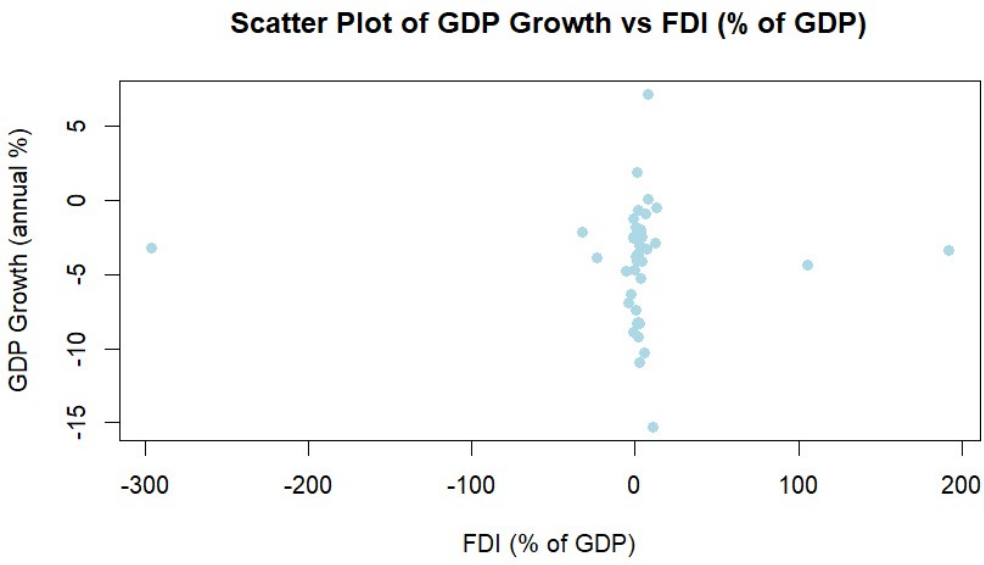

Figure 4, Scatter plot of GDP growth and FDI (% of GDP), sample 2020.

Figure 4 shows GDP growth (annual %) plotted against FDI inflows (% of GDP) for 2020. Most countries cluster near the centre, where FDI is close to zero and GDP growth lies between –10 % and 3 %. A few extreme values indicate the wide FDI dispersion noted earlier. The points form a loose cloud with no clear trend, suggesting no systematic relationship between the variables. This visual pattern mirrors **Table 3**, where the covariance (–3.74) and correlation (–0.02) both indicated an extremely weak negative association. Although GDP and FDI may move in opposite directions for some countries, the relationship is negligible. The scatter plot confirms that variations in FDI inflows are not linearly related to GDP growth, and correlation should not be interpreted as evidence of causation.

Regression Output Table				
Variable	Coefficient	Std. Error	t Statistics	P value
(Intercept)	-4.1349	0.5944	-6.9570	0.0000
FDI	-0.0011	0.0104	-0.1070	0.9153
R^2	0.0003			

Table 4, Estimated regression results for GDP growth (dependent variable) and FDI (% of GDP) as explanatory variable, 2020 sample.

Estimated regression equation: $\widehat{GDP_i} = -4.1349 - 0.0011 FDI_i$

The estimated model indicates a negative intercept (–4.13), meaning that when FDI equals zero, predicted GDP growth is around –4.1%. The slope coefficient on FDI (–0.0011) is negative, extremely small, and statistically insignificant (p = 0.915), implying no meaningful linear effect of FDI inflows on GDP growth. The R^2 value of 0.0003 shows that the model explains almost none of the variation in GDP growth across countries. These findings align with the near-zero correlation and scattered data pattern, confirming that FDI inflows alone do not significantly explain economic performance in the 2020 sample.

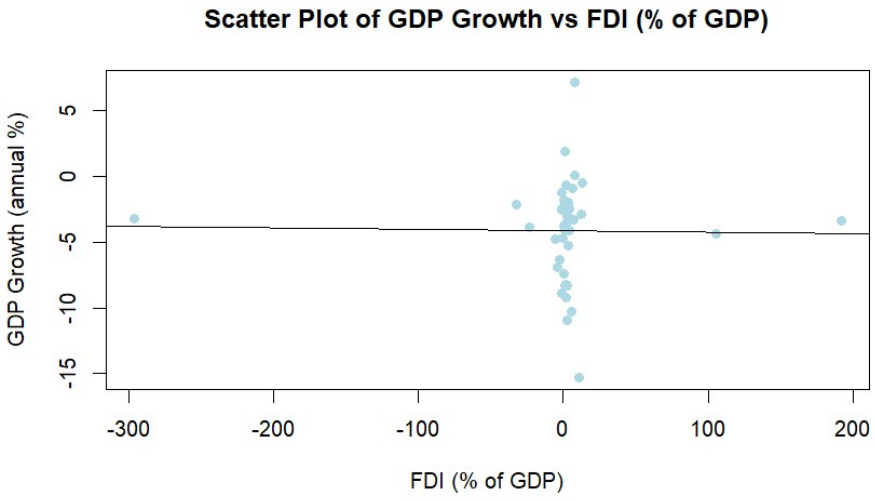

Figure 5, Scatter plot of GDP growth (annual %) and FDI (% of GDP) with fitted regression line, 2020 sample.

Figure 5 reproduces the scatter plot of GDP growth and FDI (% of GDP) with a fitted regression line added. The line slopes very slightly downward, reflecting the small negative coefficient from the regression, but the overall fit remains weak.

The differences between predicted and observed GDP values arise because the model includes only FDI as an explanatory variable, while GDP growth is influenced by many other factors. Omitted variables such as domestic demand, fiscal policy, investment efficiency, or external shocks likely explain most of the variation, leaving the simple FDI–GDP model with limited predictive accuracy.

Appendixes

A1. Data Management:

```
library(readxl)

sample2020$GDP <- sample2020$GDP1*100

round(sample2020$GDP, 2)

sample2020 <- apply_labels(sample2020,
                           country = " three letter country code for European
countries",
                           year = " year for which data was recorded",
                           FDI = "Foreign direct investment, net inflows (%
of GDP)",
                           GDP1 = "GDP growth in proportions",
                           inflation = "Inflation, GDP deflator (annual %)",
                           GDP = "GDP growth (annual %), values converted
from proportions to percentages")


sample2020.FDI <- sample2020[order(sample2020$FDI),]


save(sample2020, file = "sample2020.2.Rdata")

save(sample2020.FDI, file = "sample2020.FDI.2.Rdata")
```

A2. Descriptive Statistics:

```
summary(sample2020$FDI)
sd(sample2020$FDI)
summary(sample2020$GDP)
sd(sample2020$GDP)
summary(sample2020$inflation)
sd(sample2020$inflation)

FDI_summary <- summary(sample2020$FDI)
GDP_summary <- summary(sample2020$GDP)
inflation_summary <- summary(sample2020$inflation)

desc_table <- rbind(FDI_summary, GDP_summary, inflation_summary)
rownames(desc_table) <- c("FDI", "GDP", "Inflation")

desc_table <- rbind(FDI_summary, GDP_summary, inflation_summary)
rownames(desc_table) <- c("FDI", "GDP", "Inflation")

desc_table <- cbind(
  desc_table,
  SD = c(sd(sample2020$FDI), sd(sample2020$GDP), sd(sample2020$inflation))
)

desc_table
write.csv(desc_table, file = "desc_table.csv")
```

A3. Visualisations:

```
hist(sample2020$GDP,
     main = "Histogram of GDP Growth (annual %)",
     xlab = "GDP Growth (annual %)",
     ylab = "Frequency",
     col = "orange", border = "orange")


hist(sample2020$FDI,
     main = "Histogram of FDI (% of GDP)",
     xlab = "FDI (% of GDP)",
     ylab = "Frequency",
     col = "lightblue", border = "lightblue")


hist(sample2020$inflation,
     main = "Histogram of Inflation (annual %)",
     xlab = "Inflation (annual %)",
     ylab = "Frequency",
     col = "lightgreen", border = "lightgreen")
```

A4. Relationship Analysis:

```
# Covariance
cov(sample2020$GDP, sample2020$FDI)
# Correlation coefficient
cor(sample2020$GDP, sample2020$FDI)

cov_gdp_fdi <- cov(sample2020$GDP, sample2020$FDI)
cor_gdp_fdi <- cor(sample2020$GDP, sample2020$FDI)

cov_cor_table <- data.frame(
  Statistic = c("Covariance", "Correlation Coefficient"),
  Value = c(round(cov_gdp_fdi, 3), round(cor_gdp_fdi, 3))
)
cov_cor_table
write.csv(cov_cor_table, file = "cov_cor_table.csv", row.names = FALSE)

plot(sample2020$GDP~sample2020$FDI,
     main = "Scatter Plot of GDP Growth vs FDI (% of GDP)",
     xlab = "FDI (% of GDP)",
     ylab = "GDP Growth (annual %)",
     pch = 19, col = "lightblue"
     )

abline(regression_model)
```

A5. Regression Analysis:

```
regression_model <- lm(GDP ~ FDI, data = sample2020)
regression_model

summary(regression_model)

regression_model_sum <- summary(regression_model)
regression_model_sum$coefficients

regression_output <- regression_model_sum$coefficients

regression_output <- as.data.frame(regression_output)

write.csv(regression_output, "Regression_Output.csv")
```